Master's Thesis
Computer Security

Radboud University (RU)

Stichting Internet Domeinregistratie Nederland (SIDN)

Resilience of the Domain Name System: A case study for .nl

Author:
Lars Bade
s4051513

Supervisor (SIDN):
Jelte Jansen
jelte.jansen@sidn.nl

Supervisor (RU):
Harald Vranken
harald.vranken@ou.nl

Second assessor (RU):
Erik Poll
erik.poll@cs.ru.nl

August 2016

Abstract

This master's thesis shows a method for analyzing the resilience of the network infrastructure serving the Domain Name System (DNS) data of a top-level domain (TLD) at the autonomous system (AS) level. The provided concepts are illustrated at the example of the country-code top-level domain (ccTLD) of the Netherlands, .nl. The described approach analyzes the distribution of DNS data and its users over autonomous systems and maps the obtained results onto the global AS topology as inferred by another research project. Furthermore, it investigates the existence of bottlenecks in the availability of second level domain names from the viewpoint of the most important resolvers of the TLD zone in the case of extraordinary circumstances such as complete failure of certain autonomous systems or other parts of the DNS infrastructure. To show some concrete results, 60 different such failure scenarios are simulated and their impact on the availability of the .nl-zone investigated. This is done by analyzing the existence and length of shortest paths in the AS topology between DNS resolvers and authoritative name servers. In most of the cases the impact on the DNS is negligible or negative impact can easily be avoided by placing multiple name servers at different locations in the network topology. However, some bottlenecks in this topology are identified.

Some of the observed phenomenons can be used directly by SIDN, the registry of .nl, to improve the performance of the .nl-zone and others can be exploited by registrars to improve the availability of their domains and the ones of their customers. Also owners of domain names could keep the results of this analysis in mind when choosing a registrar which is most of the times also the hosting provider of registered domain names. Of course, the shown methodology can also be extended and applied on other TLDs.

This kind of work was, to the best of our knowledge, never done before and can be improved in many aspects, which were for various different reasons, including constraints on time, available data and computational power, outside the possibilities for this thesis. Therefore this thesis is mainly meant as a guideline on DNS resilience analysis on second-level domain names. As such it introduces the reader to the necessary concepts of the DNS and inter-autonomous system routing by means of the Border Gateway Protocol (BGP), shows important aspects to keep in mind during this kind of work and also suggests how to improve the described approach to get more accurate results in future research. Next to providing the necessary background considerations also software tools are developed and made publicly available in order to enable other researchers to easily reproduce and improve upon the performed study.

Keywords: DNS, BGP, AS topology, network resilience, failure scenario, security incident, simulation, graph analysis, path finding

Word count: 27,681

Contents

Figures

Tables

Listings

Chapter 1

Introduction

The Domain Name System (DNS) is used to translate domain names into IP addresses and as such it is a crucial part of the infrastructure of the Internet [1]. Without the DNS, many services, while theoretically reachable, will not be able to be mapped into their IP addresses and therefore unavailable for anyone who does not know its IP address. Therefore the specifications of the DNS require the DNS data to be stored redundantly, preferably on servers located in different networks, to guarantee the availability of the data in case of failure of some parts of the network [50]. As it cannot be guaranteed that all parts of the global network are constantly connected to the Internet, risks associated with malfunction of parts of this global network should be analyzed from the viewpoint of the DNS. As it is a distributed system, there are many physical locations where something can break resulting in DNS data becoming unavailable.

The DNS data is split up over a large number of name servers which are located in various autonomous systems (ASs). Routers at the border of these autonomous systems are responsible for the exchange of traffic between different autonomous systems. These routers utilize the Border Gateway Protocol (BGP) to exchange their routing information. At every point in time one of the autonomous systems serving the DNS data or connections between two autonomous systems may be interrupted for various reasons. One possible scenario for an autonomous system to suddenly be disconnected from the Internet is that the company owning that AS becomes insolvent. A possible cause for a connection to fail may be physical damage on the cable connecting two routers or failure of an upstream provider, so that that link is not connected to the Internet anymore. But not only the name servers may fail when a certain autonomous systems suddenly becomes unavailable, it might also be the case that public resolvers are not reachable anymore during such an incident, which would cause the whole DNS to be unavailable, rather than only certain domain names.

Although failure of huge parts of the DNS sounds to be a worst-case scenario, multiple incidents in the past show that this kind of things happen. Different incidents in the past suffered from different kinds of failures which range from power outage and DDoS attacks [78, 79, 93, 95] to administrative errors as happened for example in February of 2008 when Pakistan Telecom successfully brought down YouTube for an estimated two-thirds of the Internet for about two hours. In this case routers of Pakistan Telecom were manually configured to announce a prefix to its providers which actually belongs to the IP address range of YouTube instead of to their own range. This fraudulent route was then propagated to the rest of the Internet and thereby the real AS of YouTube became unavailable for large portions of the Internet [67, 68]. Similar issues could in the future also target the infrastructure of the DNS.

1.1 Research method

This study analyzes the impact of such failures on the availability of the .nl-zone and underlying second level domains. To accomplish this, the full DNS hierarchy of the .nl-zone is extracted from public DNS data and the responsible name servers are mapped to the autonomous system they belong to. As a next step, the network topology of the autonomous systems involved in the DNS tree of the .nl-zone is inferred as accurate as possible given publicly available data and used to build up a graph containing the autonomous systems and the interconnections between these. As a last step, failures of ASs and connections are simulated within the graph representation and the impact on the availability of DNS data is analyzed. For this analysis also the resolvers responsible for the majority of DNS queries within the .nl-zone are investigated. These are used as origins of DNS queries in the performed simulations.

This analysis identifies bottlenecks and single points of failure which might then be mitigated in order to improve the overall stability of the DNS in the Netherlands. The fundamental question is the impact of failure of certain parts of the infrastructure currently in place on the availability of the DNS data of the .nl-zone.

This research project was performed in collaboration with Stichting Internet Domeinregistratie Nederland (SIDN), the registry of .nl, who provided working space, necessary data and shared their knowledge and thereby facilitated this project.

Although this study focuses only on the .nl domain and domain names within the .nl-zone, the used methodology may of course be adapted in order to analyze the potential risks for other top level domains or even the entire global DNS.

1.2 Research question

The main research question answered by this thesis can be formulated in the following way:

What is the impact on the availability of the .nl-zone and underlying second level domains for the users of the Internet when parts of the DNS infrastructure become unavailable?

This research question is tackled by answering the following sub-questions:

- How are domain names within the .nl-zone distributed over autonomous systems?
- How are autonomous systems serving parts of DNS data within the .nl-zone interconnected?
- Where in the network topology do the majority of DNS requests originate from?
- How does the reachability of autonomous systems change when connections or even whole ASs become unavailable? And what is the impact hereof for the reachability of authoritative name servers?

1.3 Related work

A lot of previous research focuses on obtaining the network topology of the Internet at the AS level. This includes research by e.g. Gao [15] or Magoni & Pansiot [23]. Previous research utilized different data sources for inferring the AS-level topology of the Internet and it is not known yet which of these data sources produces the best image of the actual topology [24]. Although different data sources have been utilized, they are all dependent on data collected from the Border Gateway Protocol which was never intended to reveal the network topology and should therefore be used with care [33]. Probably one of the most complete views on the connections of ASs is given by the methodology used by Zhang et al. [38]. Unfortunately, these researchers discontinued their topology analysis in the beginning of 2015. Outdated data sets dating from September 1999 up to February 2015 are however still available at [98].

To the best of our knowledge, no data exists which perfectly shows the connections of autonomous systems within the Internet and up to now there is no way known to perfectly infer this topology [36]. This is due to the way the Internet is constructed. There is no authority which manages the Internet at the topmost level. Some researchers even state that it is impossible to obtain this topology [36]. This may be the reason that most of the research done in this field was performed in the beginning of this millennium. More recently, far less research directly aiming at inferring the network topology was performed. A good overview of research related to the inference of the network topology of the Internet at the AS level is given in a survey by Haddadi et al. [17]. Another, more recent survey by Motamedi et al. [27] inventories topology inference techniques at various levels.

Research on the stability of a network, e.g. the one performed by Rexford et al. [31] focuses on the availability of the most popular domain names, based on the traffic received at the prefixes corresponding to those domains, rather than on the availability of the domain names within the DNS. This is however also a very important factor, as without the DNS entries also the servers belonging to that DNS entry are not reachable, except for those users who by chance know the IP addresses belonging to that domain.

Research on the stability of the DNS mainly focuses on the performance of name servers serving the root zone [7] and not on the impact of certain failure scenarios on a top level domain or even second level domains. Other recent research on the availability of DNS does only investigate DDoS attacks on the DNS [20, 37]. Also at SIDN some research takes place regarding the resilience of the DNS network. This research also focuses mainly on the enhancement of the availability of the authoritative name servers of the .nl-zone by using anycasting techniques to prevent DDoS attacks [28].

1.4 The role of SIDN

SIDN is the registry for the national top level domain (TLD) of the Netherlands, .nl. This means that SIDN is responsible for maintaining the authoritative DNS name servers of the .nl domain. They also provide registry services for the TLDs .aw and .amsterdam. SIDN aims to develop a useful and secure internet for everybody. Next to the registry services SIDN therefore acts as a trust framework manager and sponsors worthwhile initiatives to strengthen the role of the Internet. [92]

The research and development department of SIDN, SIDN Labs, develops and evaluates new technologies to further stabilize and secure the .nl domain, the DNS and the Internet as a whole. One of their projects (ENTRADA) aims at collecting and analyzing DNS data to get detailed insight into the usage of the domain name system. One part of this research will depend on the DNS query data collected by ENTRADA. Furthermore they have a lot of expertise in the field of DNS, which they are willing to share with the research community, making them a great partner to collaborate for the purpose of this thesis. [84]

1.5 Outline

In Chapter 2 necessary background information on the underlying techniques such as the Border Gateway Protocol and the Domain Name System is given. Chapter 3 covers the data and methodology used for the analysis performed, followed by Chapter 4 stating the results obtained during the analysis. In Chapter 5 conclusions are drawn from the results of the analysis. Eventually, Chapter 6 provides an overview of future work associated with the obtained results.

Chapter 2

Background knowledge

This chapter provides background information on the techniques which have to be known in order to understand and perform the research. This includes information on the infrastructure of the internet and also the working of the Domain Name System.

2.1 Autonomous Systems (AS)

The Internet is not a single network, but rather a network of networks [1, 9]. Each of these networks forms an autonomous system (AS) and together these ASs, of which there are currently approximately 54,700, form the Internet [75, 27]. Autonomous systems are identified by a 32-bit (formerly 16-bit [48]) number (ASN) which is associated to a specific AS. These numbers are managed and distributed by Regional Internet Registries (RIRs) who get these numbers assigned from the Internet Assigned Numbers Authority (IANA) [60]. Each AS roughly translates to one Internet Service Provider (ISP), however, there are exceptions as large ISPs may have a larger number of ASs [8] and also other organizations might own an AS.

Each autonomous system owns a certain range of IP addresses which are denoted by prefixes according to the concept of classless inter-domain routing (CIDR) [47]. A prefix consists of an IP address and a mask length [31] indicating the number of fixed bits in the address. For example the prefix `192.0.2.0/24` is used for the range of IP addresses from `192.0.2.0` up to `192.0.2.255`, i.e. the first 24 bits of the given IP address are fixed, whereas the last 8 bits are variable. Although the IP addresses are assigned to autonomous systems in the same way as its ASN, there is no direct relation between AS numbers and IP addresses. Both are independently used to identify certain parts of the Internet. An AS number identifies an autonomous system, thus a network as a whole, whereas an IP address is used to identify a single machine which is located within an AS.

Currently two versions of the Internet Protocol (IP) and thus the corresponding IP addresses are being used. As version 4 is still the most commonly used version of this protocol [61] these IP addresses will be used for examples throughout this thesis. However, this study does not take the network layer at IP level into account and therefore these IP versions might be exchanged in future research without loss of generality. Also, the example IP addresses in the block `192.0.2.0/24` are only meant for demonstration purposes and are not assigned to any ISP as required by RFC 5737 [44].

2.2 Border Gateway Protocol (BGP)

The Border Gateway Protocol (BGP) is used to share routing information between the routers at the borders of different ASs. That is also why this protocol is called Border Gateway Protocol as it is the protocol used by the gateways (routers) at the border of an AS. The fourth version of the protocol, which is currently in use, is standardized and described in detail in RFC 4271 [54].

The basic working of BGP is easily explained: each BGP-speaking router at the border of an AS announces the prefixes of that AS to its direct neighboring routers at neighboring ASs in the network topology which store this announcement together with the incoming port in their own routing table. Furthermore the announcement is propagated to the neighboring BGP-speaking routers of the receiver, that in turn might propagate the announcement further to their neighbors, according to their configuration and policies, more on this later. This way a global routing table is created where every BGP router in theory knows at least one path to every other AS in the network topology. The announcements of routers are based on trust which does not seem to be a big issue in practice, as fraud, such as announcing a prefix which is not owned by the AS the router is located in, is easily detect and traceable [19, 39, 40].

Whenever a packet has to be transmitted via a router, the router looks up the best route to the destination in its routing table and forwards the packet to the next router on this path. Hereby, the best route is selected first on the basis of the prefix length and secondly on the length of the path to reach the destination. If this procedure does not yield a single path, several tie breaking rules are being applied such that at the end only one path remains. The exact specification of the tie breaking rules can be as basic as selecting the route advertised from the lowest IP address and is outside the scope of this master's thesis. Further information can be found in Section 9.1 of RFC 4271 [54].

Ideally all BGP routers would have the same image of the Internet, although every router has its own distinct routing table [1], dependent on the location of the router within the Internet. Therefore BGP data can be used to retrieve an overview of the connectivity of the Internet, the network topology [1]. In practice it is not the case that all routers have the same image of the Internet at the same point in time, as the network needs some time to converge after internal changes which might occur for several reasons. Possible reasons for the network to change are broken connections or failure of some routers within the network.

Each AS may have its own routing policies and preferences, e.g. preferring one route over another. These policies mainly rely on commercial contracts between the owners of different ASs [15]. To comply with a routing policy, routers may add their own AS number to a route multiple times prior to propagating a route to their neighbors and thereby artificially increase the length of a path, making it less probable that this path is being selected as best path by other routers. Another way of enforcing routing policies with BGP is to append local preferences to certain paths. If this technique is used, the path with the highest local preference is likely to be chosen by the router. Routing policies also enable an autonomous system to stop the propagation of incoming routes, such that neighboring ASs are not aware of their connection to other neighbors and thereby prohibit transit traffic for some neighbors.

The contractual agreements between the owners of two connected ASs are also used to classify the connections between the ASs. Essentially, this classification boils down to who is paying whom for providing the connection. The commonly used classification scheme is based on customer-provider, peer, and sibling relationships. In customer-provider relationships, as the name suggests, the customer is paying the provider for providing upstream connectivity. The traffic transmitted via a provider AS is called transit traffic and has to be paid for. For peering relationships normally nobody pays as both parties benefit from exchanging traffic with each other. Therefore peering relationships are mostly found between ASs of roughly the same size as otherwise one AS would send considerably more traffic to the other one. However, there are some exceptions where one party pays another party to peer with them [11, 14, 21, 35]. A last commonly used relationship is the sibling relationship, which applies to ASs with common administrative boundaries [57], e.g. ASs belonging to the same organization. Figure 2.1 illustrates an example network of ASs with their relationships.

These relationships between autonomous systems do not only reflect the contractual agreements between the owners of those ASs, but also constrain the ways traffic may flow from one autonomous system to another. Paths between two ASs are only valid if every transit-provider in the path is paid by one of its direct neighbors on the path. This means that every valid path matches the following pattern:

1. zero or more customer-to-provider links,
2. followed by zero or one peer-to-peer link,
3. followed by zero or more provider-to-customer links,

where sibling-to-sibling links may appear anywhere and in any number within the path [57]. All paths between ASs that do not match this pattern are invalid. Figure 2.2 illustrates some valid and invalid paths between AS D and AS G from the example network of autonomous systems shown in Figure 2.1.

Not only BGP relationships can be classified, also ASs can be classified based on their size and BGP relationships to other ASs. ASs are commonly

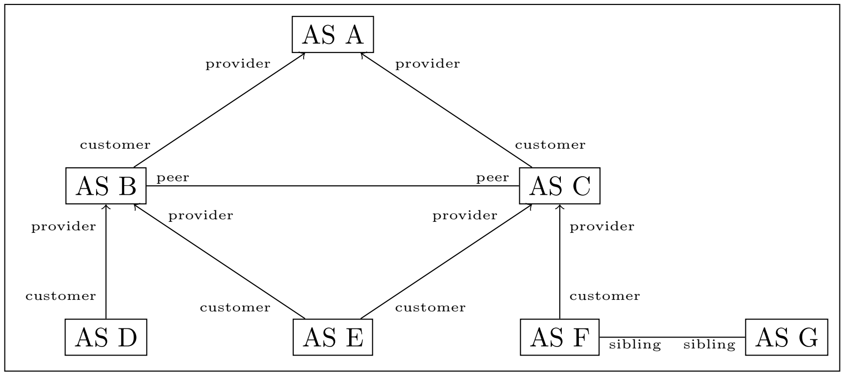
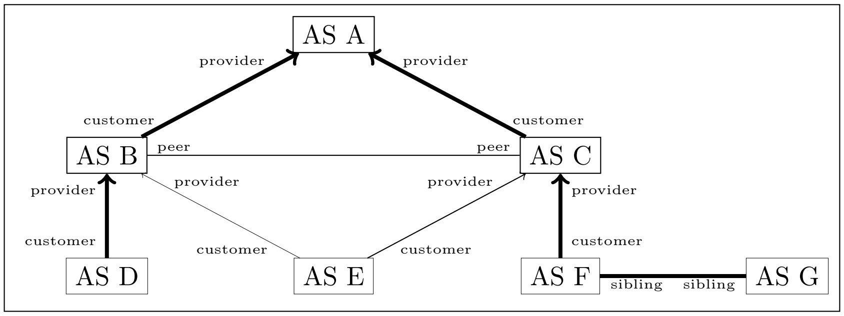


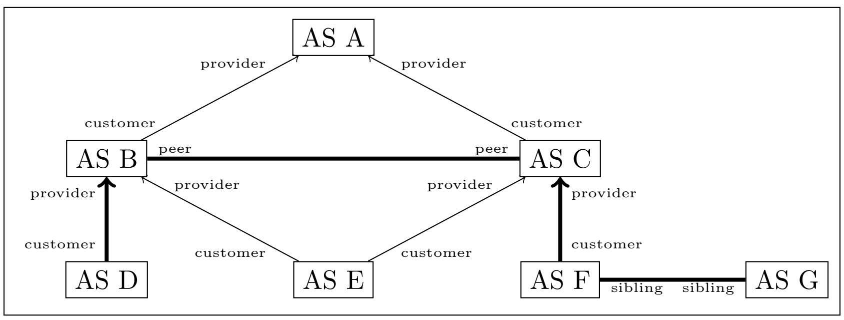
Figure 2.1: Some example ASs and their relationships with each other. In this example, AS A is the provider for AS B and AS C, AS D and AS E are customers of AS B and AS F is a customer of AS C. Furthermore, AS B and AS C peer with each other and AS F and AS G are sibling of each other. The arrows indicate money-flow from the source AS to the destination AS. This example was adapted from the one shown at the website of the Center for Applied Internet Data Analysis (CAIDA) [57].

classified into Tier-1 to Tier-5, where Tier-1 ASs are the largest ASs without any upstream providers, belonging to the largest global ISPs, which all peer with each other to ensure reachability with every other network [3, 36]. Tier-2 ASs are customers of Tier-1 ASs, which are generally national ISPs. The next step are Tier-3 and Tier-4 ASs which belong to regional ISPs and Tier-5 ASs are those not connected to any other network. Generally it holds that Tier-n ASs have customer relationships with one or more Tier-$(n-k)$ ASs [3]. However, these classifications are not official and different definitions of the different levels exist [36].

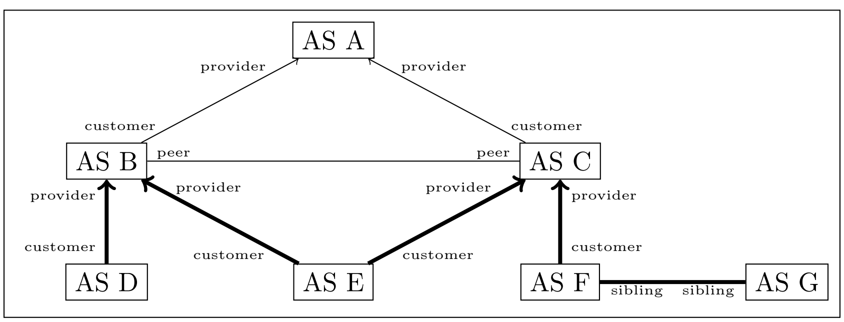
ASs that have connections to multiple transit providers are called multihomed, whereas ASs with just one transit provider are called singlehomed. Using multiple transit providers makes sense as this redundancy ensures that the AS remains connected in case of failure of one of the transit ASs. Routing policies may be implemented in order to use transit providers only as a backup and not when the preferred transit provider is functioning normally. The notion of singlehomed and multihomed AS is however a bit misleading, as every AS has to have at least two upstream providers in order to get assigned an ASN by IANA [33]. In that sense, singlehomed ASs do not really exist, although there was a proposal by RIPE to remove the requirement of being multihomed in order to be assigned an ASN in 2014 [64]. However, this proposal was withdrawn one year later [64]. Singlehomed therefore more refers to the number of upstream providers actually used by the AS, rather than the number of connected upstream providers.



(a) Valid path between AS D and AS G



(b) Valid path between AS D and AS G



(c) Invalid path between AS D and AS G

Figure 2.2: Three different paths between AS D and AS G. However, only the paths shown in (a) and (b) are valid paths in terms of traffic flow. The path shown in (c) is not a valid path as AS E would have to offer transit for which its operator is not paid by one of the AS' neighbors on the path. In practice, the path shown in (b) is most probably the selected one as AS B and AS C do not have to pay AS A for transit when going for this option.

2.2.1 Route loops

Due to the way the network is constructed and the routes being advertised it happens very often that route loops are being created. To manage this problem the AS path which has to be traversed in order to reach a certain destination prefix is being stored. If a router receives a route with its own AS number in the path, it can simply ignore this route.

2.2.2 Internet Exchanges

The European Internet Exchange Association defines an Internet Exchange Point (IXP) as *[...] a network facility that enables the interconnection and exchange of Internet traffic between more than two independent Autonomous Systems* [59]. Internet Exchanges are thus central places to facilitate ISPs to easily peer with each other. The worldwide biggest internet exchange points, such as the AMS-IX in Amsterdam, DE-CIX in Frankfurt and LINX in London, can in terms of handled traffic even be compared to the largest Tier-1 providers in the world [4, 9].

These Internet Exchanges provide several ways of exchanging routing information between the connected parties: direct connections and route servers.

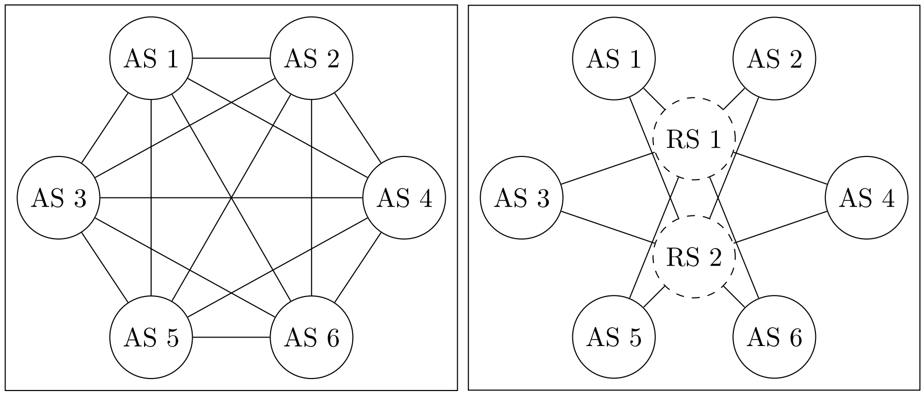
Direct connections

The easiest way to illustrate an internet exchange is to imagine it as a giant switch where every participant connects its own router [62]. That way, the routers of the participants are directly connected to each other and can directly exchange routing information via BGP sessions. This approach is referred to as bilateral peering. Bilateral peering does however have a major drawback: for every peering connection, a separate BGP session is necessary. This disadvantage can be circumvented by using route servers.

Route servers

Route servers (RS) are used to lower the amount of necessary BGP sessions while maximizing the number of parties to peer with. With this approach, which is also called multilateral peering, participants establish a single BGP session with a route server, which then broadcasts and sometimes filters according to the operators policies all incoming routes to the connected routers. Figure 2.3 demonstrates that the number of necessary BGP sessions to establish peerings between n autonomous systems reduces from $n \cdot (n-1)/2$ when using bilateral peering to $c \cdot n$, where c is the number of redundant route servers, when using multilateral peering.

On retrieving new routes from a route server, the router can still decide whether or not to accept the incoming routes and thereby follow its own routing policy. The route servers are however not meant to forward any traffic [9, 71] and are therefore sometimes also called route reflectors. These route servers, or rather the AS the route server belongs to, does therefore also not appear in the



(a) bilateral peering

(b) multilateral peering

Figure 2.3: Illustration of the different amounts of necessary BGP sessions when using either bilateral (a) or multilateral peering (b) in a network with six ASs who are all peering with each other. This example was adapted from [16].

AS path of certain routes which makes it complicated to retrieve information about the way in which two routers are connected to each other.

2.2.3 Anycast

Anycast can be considered a hack on the Border Gateway Protocol. Anycasting means that multiple servers at different locations within the global network topology announce the same IP address prefixes to their neighboring routers. As a result routers within different parts of the network will route the traffic belonging to a certain IP address to different servers. This shortens the mean path length to reach a certain IP address [6]. This principle is depicted in Figure 2.4.

Although it is possible by using anycast to announce the same IP address range from different autonomous systems, in practice, the routers announcing the same IP addresses most of the times belong to the same AS but are located at different sides.

A disadvantage of this technique is, that the servers whose routers announce the same IP address have to stay synchronized, because otherwise the responses of requests to the same IP address may be different according to the network location of a client. Furthermore it should only be used with stateless protocols as changes in the network topology might cause the best reachable server to change during a session. That means, that the first packets of an established session reach a different side than the last packets of the same session, causing the session to be destroyed. However, the advantages of this technique are worth the effort. Next to faster response times due to shorter network paths, anycasting can also be used for efficient load balancing [25] and the introduced redundancy ensures availability of the served data, even when one of the anycast nodes stops

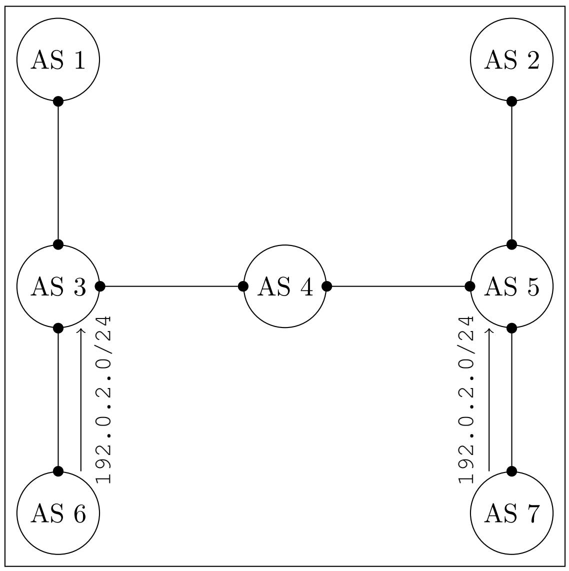


Figure 2.4: Illustration of a network of ASs where anycast is used to shorten the routes to reach the prefix `192.0.2.0/24`. This prefix is announced by the routers at the borders of both, AS 6 and AS 7. Traffic to this prefix will therefore reach one of these two systems, dependent on the originating AS. Traffic to one IP address in the range of the prefix `192.0.2.0/24` originating from for AS 1 will for example most probable reach AS 6 as the shortest path from AS 1 to AS 6 has length 2, whereas the shortest path from AS 1 to AS 7 has length 4. However, if the same packet is sent from within AS 2, it will most probably reach AS 7 since the shortest path from AS 2 to AS 7 has length 2 whereas the shortest path from AS 2 to AS 6 has length 4.

working. Anycasting is therefore also used in the DNS infrastructure [10, 22, 34].

2.2.4 Looking glasses

Looking glasses are applications which enable a user to view the routing table of routers within a specific network. These are typically provided by ISPs or IXPs and can be used to get detailed insight into the structure of a network. However, not every network (AS) has a publicly accessible looking glass, making it difficult to retrieve a full overview of the whole Internet. Dependent on the implementation and the policy of the owner of a looking glass several different functions may be provided, e.g. `ping`, `traceroute`, `bgp route` or `bgp summary`. All these commands equal their counterparts on a local shell but will be issued from within the network the looking glass is located in.

Additionally, the usage of different looking glasses might vary. Some of them are only available through public (web-)interfaces, whereas others are also available via `telnet` connections.

One example looking glass can be found at [70]. This specific looking glass provides routing information from several routers within the network of Hurricane Electric (AS6939), a globally operating ISP [69].

2.3 Domain Name System (DNS)

To send a packet to a remote host over the internet, knowledge of the IP address of that specific host is necessary, similar to the working of the telephone network where a phone number is necessary to reach a specific person. However, remembering the IP addresses of all hosts somebody might want to communicate with is rather difficult. Furthermore one probably want to reach a certain remote application, e.g. a website, regardless of the physical machine this application is running on. Moving applications across several machines leads to changing IP addresses making it even more difficult to remember them.

Therefore domain names were introduced which are more easy to remember and can automatically be translated into the corresponding IP addresses. This translation process is facilitated by the Domain Name System. To better understand the translation of domain names into IP addresses, the DNS can be compared with a telephone book which maps persons names on telephone numbers. The Domain Name System is described in many different RFCs which partly obsolete each other. The most important ones are RFC 1034 [50] and RFC 1035 [51]. The term DNS is often used for both the Domain Name System itself and the DNS protocol which is one part of the DNS.

2.3.1 History

The initial idea for the system as used today dates back to 1983 [26]. At that time, a centralized host file, called `hosts.txt` was being used to translate domain names into the corresponding IP addresses [26]. This host file contained a complete list of all hosts connected to the internet with their corresponding IP addresses. Users could access the most recent version of this file using FTP connections with the server hosting the file. This approach however had a lot of downsides and became impractical as the number of hosts grew and the file simply became too large to be distributed to the users efficiently [26]. Therefore a new, distributed system for translating domain names into IP addresses had to be developed to replace the host file. The `hosts.txt` file is however still locally present on any machine and can be used to locally overrule the DNS.

2.3.2 Architecture

A complete description of the Domain Name System is rather complex and outside the scope of this thesis. The following section therefore provides an overview of those aspects of the DNS which are relevant for this study. For a complete description of the DNS please refer to the RFCs describing the system [50, 51].

The DNS is a distributed system which consists of three components: Authoritative servers, recursive resolvers, and stub resolvers. The authoritative servers form a distributed database that stores information about existing domain names. Recursive resolvers are systems capable of querying the database in order to resolve a domain name into an IP address. Stub resolvers are pieces of software built into operating systems and end-user applications that ask recursive resolvers to perform a look-up. For this research, however, only recursive resolvers are relevant and stub resolvers are not taken into account. The DNS database is distributed in a hierarchical structure amongst a set of name servers, where every name server is serving its own zone within the domain name space. If necessary, a name server can also be used as a resolver. [50]

The domain namespace

The domain namespace is built up in a tree structure. The basis of the namespace is formed by the root zone, denoted by a single dot (.). A top level domain (TLD) is a direct child of the root zone. In the beginnings of DNS, two types of TLDs were introduced, ccTLDs which were based on country codes from ISO 3166 [94], e.g. .nl for the Netherlands and .de for Germany, and gTLDs which were based on more generic terms, e.g. .com for commercial applications and .edu for educational institutes [53]. Later, many more gTLDs were added to the system when the Internet Corporation for Assigned Names and Numbers (ICANN) started their new gTLD program which enabled any entity to apply as registry for a newly introduced TLD [5].

A domain name is formed by traversing the DNS tree and concatenating the labels on the way to the root with dots in between two labels. A label consists of up to 63 alphanumeric characters and the minus symbol which might be located at any point between two other characters. There is just one single zone with an empty label, the root zone. As the root zone is the root of the domain name

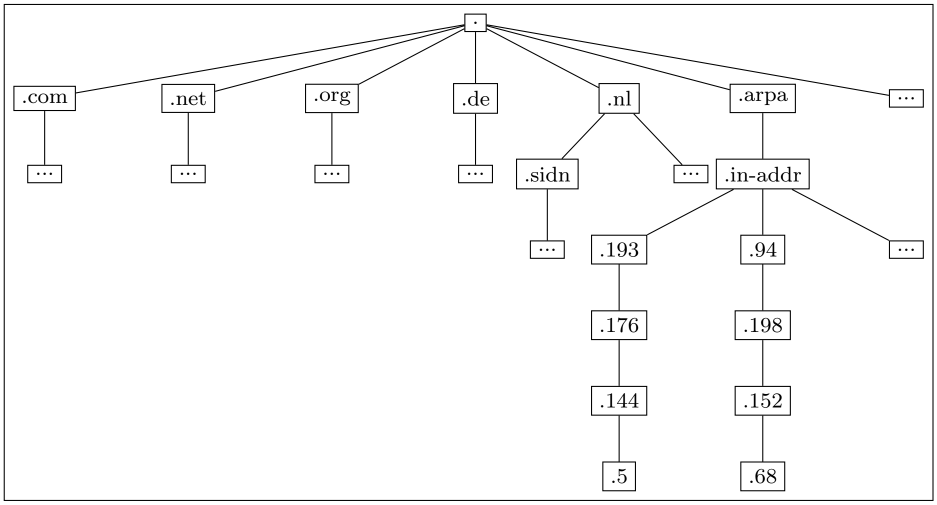


Figure 2.5: A (very) small part of the global DNS tree.

tree and does have an empty label, a full domain name always ends with a dot. This final dot is however often omitted in practice. Direct child domains of a top level domain are called second level domains and any child domain may be called subdomain of its parent, e.g. `example.nl` is a subdomain of the .nl domain, even though the term subdomain is commonly only used for domains in the third or lower level of the DNS tree. [50]

Figure 2.5 provides a very small part of the global DNS hierarchy.

Name servers

The name servers are responsible for storing the distributed database of domain names and corresponding IP addresses.

Delegation Every subdomain is delegated to a set of authoritative name servers serving the DNS data of that domain. To mitigate availability issues with the DNS, every zone should contain at least two distinct authoritative name servers, preferably located in different networks [53]. This policy is also enforced by SIDN as it can be found in their technical requirements for registering .nl-domain names [55]. At the point of writing, there are 13 active name servers for the root zone.

Resource Records The domain data is stored on the name servers in the form of Resource Records (RR). Every RR contains the following fields [51]:

`owner` The **`owner`** field specifies the domain the RR belongs to.

`type` The **`type`** field is used to distinguish different types of Resource Records. The most commonly used types are:

- **A** Contains the IPv4 address belonging to a domain name
- **AAAA** Contains the IPv6 address belonging to a domain name
- **NS** Contains the authoritative name server for a domain
- **MX** Contains the server associated with the mail exchange of a domain
- **CNAME** Is used to indicate that the domain of the record is actually an alias of some other domain name. In this case the **RDATA** field of the record contains the domain of which the domain in the **owner** field is an alias of.
- **SOA** The **s**tart **o**f **a**uthority is used to indicate that a name server is the primary authoritative name server for a domain. It holds version and refresh information about the current zone file. A detailed explanation of this record type is given in the following paragraph about synchronization of multiple name servers.
- **PTR** Points to another domain name within the DNS. This is used for reverse DNS lookups which map IP addresses onto domain names hosted on those IP addresses. Reverse DNS is explained in detail later on in this thesis.
- **TXT** Can hold arbitrary text without specific formatting

class The **class** field was introduced to enable the DNS to serve different RRs for different applications. However, the by far most used class is IN, identifying the internet. Other classes are CH for the Chaosnet or HS for Hesoid.

TTL The **T**ime **T**o **L**ive field informs the resolver on how long (in seconds) the obtained information may be cached. It is the responsibility of the resolver to remove the result from its cache after the specified time.

RDATA The **RDATA** field holds the actual data of the RR. The kind and format of data held in this field is dependent on the type of the record.

Depending on the application, multiple resource records of the same type might be stored by a name server. For example multiple A or AAAA records might be stored to facilitate load distribution over multiple servers as the resolver would typically select one of them. Also multiple NS records are very commonly used as every zone should be served by multiple authoritative name servers. However, there should always be just a single SOA record as every zone just has a single primary authoritative name server. [50]

Wildcard records The Domain Name System supports wildcard matching. That is, the DNS can be configured such that a record is matched for any domain name. For example a resource record belonging to the domain `*.example.com` does match all subdomains of `example.com`. [50]

Synchronization of multiple name servers As previously mentioned, every zone within DNS is served by multiple redundant name servers. These multiple name servers have to stay synchronized in order not to send different answers for the same query. In every zone there is one master name server, called primary name server, holding a special `SOA` RR. All other name servers within the same zone are called secondary name servers or slaves. The `SOA` RR on the primary name server specifies `SERIAL`, `REFRESH`, `RETRY` and `EXPIRE` values in the `RDATA` field. Every time records are being updated on the master name server, the `SERIAL` value of the `SOA` record is incremented somehow. This `SERIAL` value can thus be seen as a version number of the current zone file. This number might simply be increased by one at every change or hold the timestamp of the last modification. The slave servers use the `SERIAL` value in order to detect whether their local copy of the zone data is up to date. The `REFRESH` value hereby indicates the time in seconds between two subsequent checks. The `RETRY` value specifies the interval in which the secondary server retries to obtain the current `SERIAL` value in case of failure during a regular refresh. If the secondary server is not able to check for updates against the master server for `EXPIRE` seconds, it has to assume that its data is outdated. [50]

Resolvers

Resolvers are the applications capable to query the DNS. Resolvers can give answers to queries in two different ways: *recursive* or *non-recursive*. If a resolver operates recursively it is guaranteed to respond with the final answer to the query. In the non-recursive case it returns the answer to the first step of the resolution process as described later. This might be for example a `CNAME` RR in the case that the queried domain name is an alias of some other domain name, or a referral to another name server specified in an `NS` record in the case that the queried name server does not know the answer on the query. If the resolver operates recursively and the first resolving step returns a `CNAME` or `NS` record, the resolver will automatically perform the following step in finding the answer to the query. [50]

When a user wants to resolve a particular domain name, e.g. `example.nl`, the first step is to send a query to one of the root name servers, asking for the authoritative name servers of the TLD. The IP addresses of these very important name servers are hard-coded in the resolving software. The root name server will then respond with a referral to the name servers which serve the .nl-zone. The next step would then be to query one of those name servers in order to retrieve the IP address which belongs to the queried domain name and thereby traversing the DNS hierarchy tree up to the domain name in question. [50]

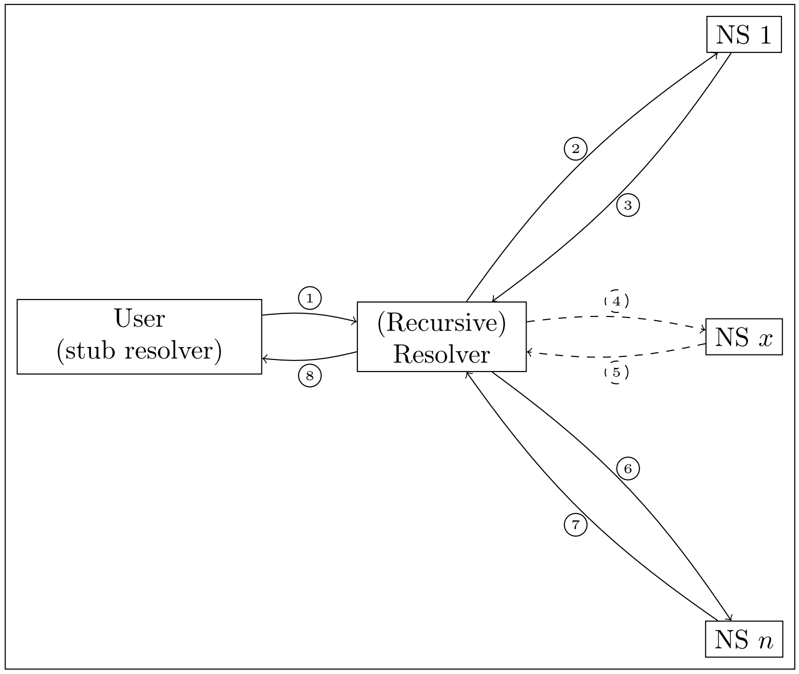


Figure 2.6: The resolution process of a DNS query and the involved steps. The arrows indicate the direction of a message. Dashed arrows indicate optional steps. An explanation of the numbered steps in this figure is given in the main text.

Figure 2.6 illustrates the necessary steps taken to resolve a DNS query for a user in more detail. These steps taken in the resolution process as depicted in Figure 2.6 are:

1. The query is sent to a resolver.
2. The resolver redirects the query to one of the name servers of the root zone. The IP addresses of these servers are hard-coded into the software of the resolver.
3. The name server responds with a referral to an authoritative name server of the TLD in question. A resolver acting non-recursively would stop after this step and directly respond to the user with the obtained referral.

4–5. A recursive resolver goes on with finding the final answer to the query of the user. The resolver therefore sends a query to the authoritative name server of the TLD as obtained in the previous step, possibly followed by an uncertain number of more queries to obtained referrals until it obtains a referral to the authoritative name server of the domain in question.

6. The resolver sends a final query to the authoritative name server of the domain in question.
7. The name server responds with the answer to the query.
8. The resolver redirects the response of the name server to the user.

Caching The resolution process can be optimized by using caching techniques at multiple levels. The given example is simpler than resolving a query in the real system as it assumes that all caches are empty. A resolver caches all responses obtained from name servers, such that these answers can be reused to answer subsequent queries. In the case that one user queries a domain in the .nl-zone, for example, the resolver will query one of the name servers of the root zone for an authoritative name server for the .nl-zone. The answer is then stored in the cache of the resolver. When another user queries another domain in the .nl-zone at a later point in time, the resolver still has the authoritative name server of the .nl-zone in its cache and does not have to query the authoritative name server of the root zone again in order to be able to answer the query. Although it would theoretically be possible to implement a recursive resolver in every browser, this is not done as it would bypass most of the advantages of caching as it would imply that much more queries are sent to the authoritative name servers of frequently used zones, such as the root zone. Therefore more servers would be necessary and the system would become more complex. By sharing resolvers amongst multiple people or systems, the caches can also be used by all users of those resolvers. [50]

Queries

The DNS protocol supports several query types (qtypes), most of which match the previously described types of resource records, i.e. every resource record type is also a qtype. However, there are some special query types which do not match a resource record type [50]:

ANY The **ANY** request is used to retrieve all known resource records associated to a given domain. This query type is sometimes also denoted as an asterisk (*).

AXFR The **AXFR** request is being used for full zone transfers. If an AXFR query is issued to a authoritative name server, the name server will initiate a zone transfer and send all the data contained in its zonefile to the origin of the query. This request has to be used with care and should be secured such that it can only be called by authorized slaves as it otherwise may be used in DoS attacks by DNS amplification. This is due to the fact, that the query is typically much smaller than the response. An attacker might thus send `AXFR` requests to the authoritative name server with a spoofed source IP and the server will respond and send large packets to the spoofed address in the source field of the packet. That way a DoS attack might be executed.

IXFR The **IXFR** request is also used for zone transfers. However, this request is used for incremental zone transfers, rather than full zone transfers. It therefore also contains a version number as parameter and is not answered with a copy of the whole zonefile, but instead with only the changes applied since the specified version of the file.

A DNS query contains the domain name to query, the query type, a random ID to link a response to the query and several flags, for instance to indicate whether recursive resolving is desired by the user. [50]

2.3.3 Responses

A DNS response contains several parts. In the header of a response some of the query arguments, such as the ID of the query and whether recursion was desired are repeated. Furthermore some flags are contained indicating whether the answer had to be truncated in order for the packet to fit into the packet size of the transport protocol used or whether recursion is available on the queried server. It also contains a response code (rcode) which indicates whether an error occurred during the resolution process and the number of resource records contained in the question, answer, authority and additional sections, respectively. [50]

The most commonly used response codes are the following [51]:

NOERROR The response code **NOERROR** indicates that no error occurred during the resolution process.

FORMERR The **FORMERR** response code indicates that there was an error in the formatting of the query.

SERVFAIL The **SERVFAIL** response code is send after an server error.

NXDOMAIN The **NXDOMAIN** response code indicates that the queried domain name does not exist.

A reply on a DNS query can either be *authoritative* or *non-authoritative* which is also indicated by a flag in the header of the response. A response is only authoritative if all subsequent answers used to retrieve the final reply are authoritative, i.e. directly extracted from an authoritative name server. A reply is however non-authoritative if at any step in the resolution process a cached reply is being used. [50]

The question section of the response contains a copy of the question as received by the name server, i.e. the domain name, type and class of the query. The actual answer to the query can be found in the answer section of the response. The optional authority section might contain the authoritative name servers of the queried domain. [50]

Sometimes it also happens that the name servers for a particular zone are located in the very same zone, e.g. the authoritative name servers of `example.com` might be `ns1.example.com` and `ns2.example.com`. In these cases it is impossible for the resolver to perform the next step in the resolution process as it is not able to resolve the IP addresses of the name servers. To circumvent this issue, so called glue records are being used. These are contained in the optional additional section in the DNS response and specify the IP addresses of the name servers in question. [50]

An example DNS response can be found in Listing 2.1. In this example, the usage of glue records can be shown, as the name servers serving `sidn.nl` are all located within the `sidn.nl` zone, as can be seen in the answer section. This means, that resolving `sidn.nl` requires to resolve one of the name servers `ns1.sidn.nl`, `ns2.sidn.nl` or `ns3.sidn.nl`, respectively. This introduces an endless loop and the domain name would thus not be resolvable. Therefore the resolved IP addresses of the name servers are also placed in the additional section of the response such that these IP addresses can be used to retrieve other records associated with `sidn.nl`.

Listing 2.1: Contents of an example DNS response for the name servers of `sidn.nl.`. This response was retrieved by issuing the command `dig NS sidn.nl`.

```
; <<>> DiG 9.9.5-11ubuntu1.3-Ubuntu <<>> NS sidn.nl
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 43609
;; flags: qr rd ra ad; QUERY: 1, ANSWER: 3, AUTHORITY: 0, ADDITIONAL: 7

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 4096
;; QUESTION SECTION:
;sidn.nl.   IN NS

;; ANSWER SECTION:
sidn.nl.  86400 IN NS ns2.sidn.nl.
sidn.nl.  86400 IN NS ns3.sidn.nl.
sidn.nl.  86400 IN NS ns1.sidn.nl.

;; ADDITIONAL SECTION:
ns1.sidn.nl.  61788 IN A 94.198.152.68
ns1.sidn.nl.  61788 IN AAAA 2a00:d78:0:147:94:198:152:68
ns2.sidn.nl.  61788 IN A 194.171.17.5
ns2.sidn.nl.  61788 IN AAAA 2001:610:0:800d::5
ns3.sidn.nl.  61788 IN A 194.0.28.3
ns3.sidn.nl.  61788 IN AAAA 2001:678:2c:0:194:0:28:3

;; Query time: 6 msec
;; SERVER: 127.0.1.1#53(127.0.1.1)
;; WHEN: Tue Mar 15 17:01:45 CET 2016
;; MSG SIZE rcvd: 222
```

2.3.4 Reverse DNS

Next to mapping domain names to IP addresses, DNS is also used to map IP addresses to domain names. This process is called reverse DNS and is in particular necessary for mail programs to detect spam mails. When a server receives a new mail, it may check whether the server the mail is coming from may actually send emails for a certain domain. For the reverse DNS look-up the special domain `in-addr.arpa` is used. To facilitate reverse look-ups, the domain is added as a `PTR` record to the reversed IP address used as subdomain of `in-addr.arpa`. This domain thus only contains RRs of the type `PTR` which maps an IP address to a domain name. For example querying the `PTR` RR for the domain `5.144.176.193.in-addr.arpa.` yields the domain name `ns1.dns.nl.`. And indeed the domain `ns1.dns.nl.` is located on the server with IP address `193.176.144.5` as can be seen by querying the `A` record of that domain. [50]

However, the reverse DNS look-up does not automatically match the forward look-up. It is the responsibility of an ISP to administer the reverse look-up of its IP addresses, possibly on behalf of its customers.

2.3.5 DNSSEC

The DNS in its original form is subject to several security vulnerabilities, including cache-poisoning and man-in-the-middle-attacks, which might be exploited by malicious people in order to feed falsified answers on DNS queries to users of the system and thereby redirecting the user to a different and unintended location [45].

These attacks can be mitigated by a rather new extension to the DNS called DNS Security Extensions, or DNSSEC in short. A detailed explanation of DNSSEC can be found in the RFCs 4033, 4034 and 4035 [41, 42, 43].

DNSSEC adds origin authentication and integrity protection to the data within the DNS by means of adding digital signatures to the stored resource records.

Additional resource records

To be able to do so, some additional resource records are defined for DNSSEC. These additional RRs are the following [41]:

`RRSIG` Contains a digital signature over the set of resource records.

`DNSKEY` Holds the public key corresponding to the private key used to sign the resource record as stored in `RRSIG`.

`DS` The **D**elegation **S**igner record contains the public key corresponding to the private key used to sign a child zone.

`NSEC` The **N**ext **Sec**ure resource record is used for authenticating negative responses such as NxDomain. It contains a signature over the next secure domain name in the same zone where next domain name is defined by the ordered canonical representations of existing domains.

Chain of trust

DNSSEC adds digital signatures to resource records such that a security-aware resolver is able to verify the signatures in order to verify the correctness of the received information. For this to work, the resolver however has to know the public key corresponding to the private key used to sign the zone. [41]

This public key, which is also stored in the DNS has itself also be signed by another key in order to make sure nobody modified the public key. This is done by forming an authentication chain of trusted data from a zone up to a previously authenticated public key or a trust anchor. As the root zone does not have a parent zone which may hold a signature on its public key, every security-aware resolver has to have at least one trust anchor configured which may be used to authenticate a public key in various ways. [41]

The typical chain of trust along the way of resolving a resource record then becomes `DNSKEY`→[`DS`→`DNSKEY`]*→`RRset`, where arrows indicate a key signing a record and * denotes zero or more occurrences of a `DNSKEY` record signed by a `DS` record. [41]

This authentication chain, which is typically built up from the root zone down the DNS tree to the resource record in question [41] is also the reason why a domain name can only be signed with DNSSEC if all its parent zones are properly signed with DNSSEC.

Authenticated denial of existence The previously described approach for signing and validating DNS data can only be applied on actually existing resource records. However, also NxDomain responses should be authenticated by the authoritative name servers of a zone. For this purpose, the `NSEC` resource record is used. This record contains a signature over the next secured existing domain name in the same zone, where next domain name is defined by the ordering of the canonical representation of existing names. Chains of these records therefore define the gaps of a zone, i.e. domain names which do not exist in the zone. These signatures can be verified the same way as the signatures contained in the `RRSIG` field and be used to validate whether a domain name exists or not.

Other extensions

Next to defining some more resource records to be used within the DNS, DNSSEC also defines three additional query header flags: Checking Disabled (CD), Authenticated Data (AD) and DNSSEC OK (DO). The CD-flag can be used to manually disable the DNSSEC checks, the AD-flag signals that a resolver already validated the retrieved data and the DO-flag signals that a resolver is capable of retrieving DNSSEC data [41, 88]. The DO-flag is introduced to keep backwards compatibility and not send DNSSEC record to old resolvers which might fail when getting those. However, although the DO bit is set, this does not necessarily mean the resolver validates, it merely signals it can receive DNSSEC records without failing [99].

Adoption

Although every domain should be secured using this technique in order to be able to defend against the in RFC 3833 [45] listed attacks, adoption is very slow and suffers from bootstrapping problems [29].

SIDN is therefore actively stimulating the adoption of DNSSEC and .nl thereby became the market leader in terms of the total amount of signed domain names [90]. Nevertheless, only roughly 44.5% of the domains contained in the .nl-zone are currently signed with DNSSEC [83], although the .nl-zone itself is already signed since 15 May 2012 [89].

2.4 DNS meets BGP

As previously explained, the global DNS is a distributed system. Both compartments of this system, the name servers and the resolvers are located in different ASs all over the world. To successfully resolve a DNS query, all intermediate name servers involved in this resolution and thus the ASs they are located in have to be reachable from the AS the resolver is located in. This reachability is ensured by means of the Border Gateway Protocol as this protocol is responsible for exchanging traffic between different autonomous systems as extensively described Section 2.2.

Chapter 3

Research method

In this chapter the performed analysis, i.e. data sources, data representations and the used methodology, will be described in detail. The following gives an overview of the general steps performed for the purpose of this assessment:

1. Obtaining a list of domain names within the .nl-zone and the corresponding name servers by resolving the `NS` records of domains and `A` records of obtained name servers.

2. Mapping of all obtained name servers onto the autonomous systems they are located in.

3. Identifying the ASs the most commonly used resolvers are located in.

4. Obtaining the network topology of the previously identified ASs, including relationships.

5. Creating a directed graph with labeled edges representing the obtained topology.

6. Simulating failure of ASs and connections by removing those from the graph.

7. Analyzing the impact of these failures on the reachability of DNS data from the viewpoint of the most commonly used resolvers.

All these steps were automated by developing Python scripts. These scripts are made available such that future research can utilize and refine them.

After describing the approach taken to perform each of these steps in Section 3.1, some limitations of this methodology will be discussed in Section 3.2.

3.1 Data and methodology

As the data within the DNS and also the BGP routing tables are constantly changing, the simulations cannot be performed in real time. Therefore a copy of the available data was made at a certain point in time and afterwards the data within that copy was used to perform the simulations. The methodology described here was translated into Python scripts to automatically perform the analysis. The complete source code is given in Appendix C. This code can either be used for verification of the obtained results or to repeat and/or extend the analysis with another scope, e.g. another TLD.

3.1.1 Obtaining name servers

All DNS data is in principle publicly accessible by querying the global DNS. To simplify the process of obtaining all existent domain names for analysis, SIDN provided a copy of the zonefile of the .nl-zone. This zonefile contains all registered domain names along with their name servers as known by SIDN and possibly glue records. However, the name servers that are known to SIDN are not guaranteed to match the name servers actually in place. Therefore this zonefile was only used to extract all currently registered domain names in the .nl-zone.

The name servers of a domain name and corresponding IP addresses of these name servers could then be retrieved by actively resolving the domain names of the name servers using the commandline tool `dig`. In particular the commands `dig +short NS <domainname>` and `dig +short A <domainname>` could be used to retrieve the name servers associated with a domain name and the IPv4 address(es) of these name servers, respectively.

To carry out this process more efficiently, a tool called spark [91], was used. This tool was developed by SIDN and is capable of resolving multiple DNS queries in parallel. Spark was used in two subsequent steps, firstly to retrieve all name servers of a domain and secondly to retrieve the IPv4 address(es) of these name servers. The name servers associated with all domain names were retrieved by calling spark with the command shown in Listing 3.1.

Listing 3.1: The command issued on the command line to retrieve the name servers associated with a domain name. The variables `<input_file>` and `<ip>` have to be replaced by proper values as described in the text. `<output_file>` should be replaced by the path to a file where the output should be stored.

```
spark -goroutines 500 -insecure -names <input_file> -print_rrs -
    resolver <ip> -rrtype NS > <output_file>
```

The arguments provided to `spark` are the following:

`-goroutines` The amount of queries resolved in parallel.

`-insecure` Turns off the DNSSEC check, which was the original purpose of the tool.

`-names` Specifies the path to a file containing the domain names which should be resolved. This file should contain one domain name per line.

-print_rrs Specifies that the obtained resource records should be printed.

-resolver Specifies the IP address of the resolver that should be used by the tool.

-rrtype Specifies the RR type that should be resolved.

Afterwards the IP addresses of these name servers were resolved by calling the command shown in Listing 3.2.

Listing 3.2: The command issued on the command line to retrieve the IPv4 addresses of the previously identified name servers. Again the variables <input_file>, <ip> and <output_file> have to be replaced by proper values.

```
spark -goroutines 500 -insecure -names <input_file> -print_rrs -
    resolver <ip> -rrtype A > <output_file>
```

By using the parallel resolving provided by spark all required (roughly 6 million) DNS queries where resolved within 1.5 hours. This process could have taken several days without parallelizing.

If a server failure occurs at one of the authoritative name servers of a domain name, it might happen that the NS set cannot be retrieved correctly. To overcome this problem, the resolution process was repeatedly executed on those domains which could not be resolved properly in previous executions.

3.1.2 Mapping IPs onto ASNs

As this study is performed on the autonomous system level, the IP addresses of the name servers which were obtained in the previous step have to be mapped onto the autonomous systems these IP addresses belong to. This was done by using the GeoLite database provided by MaxMind [76]. The main advantage of this method, in contrast to other services providing the same information as the RIPE database [65] and the service provided by Team Cymru [77], is that the dataset can be downloaded and therefore the look-up of AS numbers which belong to a given IP address can be done in memory, rather than issuing queries to network services. This results in a major performance speedup as a huge amount of IP addresses have to mapped onto autonomous systems.

As a backup method, the 'Network Info' functionality of the RIPE DB was used. This tool can also be accessed by means of an API, called RIPEstat. To obtain the ASN of any given IP address the following URL may be called: https://stat.ripe.net/data/network-info/data.json?resource=<ip>

3.1.3 Mapping ASNs onto organizations

To better understand and be able to interpret the results obtained during this research it is necessary to map the ASNs investigated throughout the analysis onto the companies these ASNs belong to. Otherwise one would not be able to draw conclusions from the results as they would only show some ASNs and not the companies managing the belonging autonomous systems.

For this purpose, the AS to Organization Mapping dataset provided by CAIDA [58] was being used which directly provides the necessary information. This dataset is based on `whois` data associated with an ASN.

In some situations this dataset is not very clear as some whois entries are misinterpreted by the data crawler. As these are very easily spotted these data points were manually corrected for the purpose of this research paper. A list of all mentioned ASs and the organizations they belong to is shown in Appendix A.1.

3.1.4 DNS Resolver location

This section describes how the locations, in terms of the AS topology, of the most important resolvers for the .nl-zone were obtained. Knowledge of these resolvers is important for the purpose of this study as we are analyzing the reachability of authoritative name servers in the case of an incident. Due to the previously described architecture of the DNS, the name servers must be reachable for resolvers, rather than users. In turn, of course, the resolvers must be reachable for the users of those resolvers. However, we do not have any information about the users of a resolver and hence exclude this aspect from the analysis. As we cannot analyze reachability of the authoritative name servers for all resolvers due to the huge amount of distinct resolvers being used, we just focus on the most important ones for the users of the .nl-zone, i.e. the most used ones. A remaining major challenge is to define the importance of a resolver appropriately, which might be tackled in a master's thesis of its own. The remainder of this section explains some caveats related to this part of the analysis and how the list of ASs, the most important resolvers reside in, was eventually obtained.

For this part of the study, the DNS look-ups received at two out of the seven authoritative name servers of the .nl-zone were investigated using ENTRADA [81], an open-source platform for storing and analyzing large amounts of DNS traffic developed by SIDN Labs. Within this platform SIDN collects the DNS queries received at two of the authoritative name servers of the .nl-zone, `ns1.dns.nl` and `ns2.sidn.nl`, for about 2 years now, resulting in a database containing more than 100 billion queries, which grows by another 400 million queries every day[1] [80]. ENTRADA provides a simple SQL interface to retrieve the stored data. For a more accurate analysis, only a subset of the available data was being used, representing the queries received in the first week of June 2016. This is roughly the same timeframe as the AS topology was modeled for. However this data covers a whole week rather than a few days, as in the case of the AS topology, as usage of the DNS varies a lot between weekdays

[1] In the meantime, also `ns3.dns.nl` and `ns4.dns.nl` were integrated into ENTRADA. The database therefore now roughly grows by 800 million queries every day and there are plans to also collect queries received by `ns5.dns.nl` in the near future.

and the weekend. To get a representative view on the overall usage of a resolver, therefore the DNS traffic of a whole week was investigated.

It was desired to perform the analysis from the viewpoint of those resolvers which are most important for actual human users of the DNS. At first glance one might think that these coincide with those resolvers issuing the most requests to the authoritative name servers of .nl. However, this is not the case as there are some use-cases of the DNS where a lot of requests are automatically generated by special applications whereby the resolvers serving those applications would be counted as disproportionately important as all those queries are generated by a single application for a single user. Next to these resolvers, there are also some resolvers sending more requests than others to get the same information due to DNSSEC validation or misconfiguration.

As can be seen, there are several caveats which have to be taken into account when using purely the amount of received queries as a measure for the importance of a resolver for human users as a lot of the received requests are either generated automatically or unnecessarily. Therefore just counting the amount of requests received does indeed provide a good metric for the usage of the DNS, but not for the usage of the DNS by actual human users and correctly configured resolvers.

In general, the following caveats have to be taken into account when trying to classify a resolver according to its usage based on the number of DNS requests originating from that resolver:

- Some resolvers are constantly querying the whole zone in a linear fashion and thereby producing disproportionately many DNS requests. There are two main reasons for these kinds of scans: generating profit (e.g. domainers) and research (e.g. the OpenINTEL project, a collaboration of the University of Twente, SURFnet and SIDN [32, 96]).

- Some resolvers send the same query multiple times in very short intervals. One possible explanation for this would be misconfiguration or absence of caching at the resolver. These resolvers therefore tend to send more queries than other resolvers, so the unnecessary ones have to be negated. Another reason for this phenomenon would be that resolvers send the same query to multiple authoritative name servers and just use the fastest reply for optimizing the performance of the resolver at cost of the performance of the authoritative name servers.

- Resolvers which validate DNSSEC entries send additional queries to the name servers in order to obtain the necessary records for validation of the resource records in question. We should therefore just look at the basic query types `A`, `AAAA`, `CNAME`, `MX`, `NS`, `PTR`, `SOA` and `TXT` such that we exclude additional queries.

Although all these caveats were tried to be taken into account for as far as possible for filtering those requests which are actually generated by users, by introducing some measures for classification of resolvers according to their usage, the resulting classification of resolvers is not very accurate and leaves a lot of possibilities for future research to increase its accuracy. Therefore the eventually used approach was a combination of both, investigating purely the amount of requests received and those requests which are most probably generated by human users. This was done by generating two distinct rankings of the top 30 most important resolver locations in the following ways:

1. By investigating only the amount of queries originating from resolvers located in an AS.
2. By investigating the amount of queries originating from the human users of resolvers located in an AS. This was done in several subsequent steps:
   (a) Retrieve all requests of resolvers, only counting basic query types, i.e. not counting DNSSEC related requests and other unusual query types.
   (b) Remove duplicates, i.e. identical queries received in short time intervals from the same resolver.
   (c) Excluding resolvers issuing only queries of type `NS` as these most probably belong to domainers.
   (d) Excluding resolvers querying mainly distinct domain names as these most probably belong to automated scans.
   (e) Excluding resolvers with mainly NxDomain responses as these are most probably used by botnets.
   (f) Excluding resolvers scanning the zone in a linear fashion.
   (g) Combine all approaches to exclude automated resolvers to a single SQL query in order to retrieve the ranking of the top 30 most important resolver locations.

   Hereby, step (f) was done by manually inspecting the requests originating from a resolver, whereas the other steps were performed automatically.

Although these two rankings of most important ASs are generated with different measures for the importance of a resolver in mind, they show a lot of overlapping ASs occurring in both lists. Both lists were then combined in order to retrieve the final list of ASs eventually taken into account during the analysis.

The remainder of this section gives and explains the SQL queries issued to the ENTRADA database in order to produce the two rankings of the top 30 most important ASs for both approaches.

ENTRADA Database

The ENTRADA database consists of a single table called `dns.queries` in which all incoming DNS queries are being stored. Table 3.1 explains those columns of this table that are necessary for the purpose of this research.

Table 3.1: The columns of the `dns.queries` table that were utilized for the purpose of this study. The complete documentation of the data model used by ENTRADA, including the other columns of the table not used during this research, can be found at [82].

column	protocol	source	description
`rcode`	DNS	response	rcode (-1 is no matching server response is found)
`unixtime`	-	META	packet timestamp, seconds since January 1, 1970, 00:00:00 GMT
`qname`	DNS	query	qname from question
`qtype`	DNS	query	qtype from question
`src`	IP	request	source IP address
`ipv`	IP	request	IP version, 4 or 6
`asn`	-	META	autonomous system number of the source IP address
`year`	-	META	year part of timestamp
`month`	-	META	month part of timestamp
`day`	-	META	day part of timestamp

1. Approach: All queries per AS

The first approach taken is to take the amount of queries received from a certain resolver as a measure for its importance. This approach has however several downsides as it only takes the resolvers into account and not the users which are served by these resolvers.

The query issued to the ENTRADA database to retrieve the number of queries received from resolvers per autonomous system is shown in Listing 3.3.

Listing 3.3: The SQL query used within ENTRADA to retrieve the origins of DNS queries.

```sql
SELECT asn, COUNT(1) as tot
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND asn != "UNKN"
GROUP BY asn
ORDER BY tot DESC
LIMIT 30
```

The result of this query forms the first list of most important autonomous systems and is shown in Table 4.3

2. Approach: Queries per AS originating from human usage

(a) Queries per ASN counting only basic query types The SQL query used to retrieve the amount of DNS requests per autonomous system logged by ENTRADA, only counting the basic query types is shown in Listing 3.4.

Listing 3.4: The SQL query used within ENTRADA to retrieve the origins of DNS queries.

```
SELECT asn, COUNT(1) as tot
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND asn != "UNKN"
AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
AND rcode = 0
AND ipv = 4
GROUP BY asn
ORDER BY tot DESC
```

Next to the constraint of only counting the basic query types, some additional properties must be met by the logged requests to be counted:

1. The AS, the source IP address originates from, must be known. As the DNS query packet only contains an IP address and not the ASN this IP address belongs to, this information is inferred by ENTRADA utilizing the IP2ASN database from MaxMind [76]. However, this database is not complete an therefore it happens that the source AS is not known.

2. The queried domain name must exists. Some resolvers query the DNS in a brute-force way which does not reflect how the DNS is used by legitimate users. Assuming that legitimate users browse almost only existent domain names, the `rcode` of a query should be 0, specifying NoError.

3. This study focuses on the reachability of name servers using IPv4. Therefore only queries using this protocol are considered here.

(b) Remove duplicate queries As previously said, some resolvers issue the same queries multiple times in very short intervals. This should not be the case when the resolver is well configured. As these resolvers tend to send more queries than those which are well configured, we have to exclude the additional queries from our search. The query shown in Listing 3.5 therefore lists the queries originating from one AS, where duplicates of a single resolver in intervals of 5 minutes (300 seconds) are removed.

Listing 3.5: The SQL query used to list the DNS queries where duplicate requests within timeframes of 5 minutes are removed from the dataset.

```
SELECT asn, src, qname, qtype, (unixtime DIV 300) * 300 AS slice
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND asn != "UNKN"
AND rcode = 0
AND ipv = 4
AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
GROUP BY asn, src, qname, qtype, slice
```

This query groups identical DNS queries received from resolvers within timeslices of 5 minutes. The interval of 5 minutes is arbitrarily chosen as it is considerably shorter than the recommended TTL value for DNS entries, but still longer than the interval in which misconfigured resolvers tend to resend the same query.

The downside of this approach is, that it distributes the timeslices according to the clock, i.e. the transitions of one slice to another is dependent on the actual time, rather than dependent on the timestamp of the first query received. It might therefore happen that two identical queries are received directly after each other, where one of the queries falls in one slice and the other one falls within another slice. However, this is not very probable in slices of 5 minutes (i.e. it happens with a probability of 1/300) and looking at the great amount of queries received in total, negligible. Fixing this issue would make the SQL query disproportionately complex.

The constraints which must be met by DNS requests considered in this SQL query are identical to those explained for the previous query in step (a) and will also be part of all following SQL queries in this section.

(c) Exclude resolvers with just `NS` queries It was discovered, that some resolvers only send DNS queries of the `NS` type. Although QNAME minimization, a technique for increasing the privacy in DNS [46], might be a reason for this, the most probable reason is that these resolvers belong to domainers, searching for recently removed domain names for reselling purposes. As these domainers produce disproportionate amounts of DNS queries, such as the resolvers constantly scanning the whole zone, it was decided to exclude these resolvers from this study. For identifying these domainers, the query shown in Listing 3.6 was used.

Listing 3.6: The SQL query to obtain those resolvers which are only issuing queries of the type NS.

```
SELECT src, COUNT(1) AS tot
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND asn != "UNKN"
AND rcode = 0
AND ipv = 4
GROUP BY src
HAVING COUNT(DISTINCT qtype) = 1
AND MIN(qtype) = 2
```

This query groups the DNS queries originating from resolvers and retrieves those resolvers who are only issuing queries of the type `NS`. Please note, that the usage of the group function `MIN(qtype)` is necessary, as we want to address the qtype in the query, but cannot group by this property as otherwise the `HAVING`-clause would not work. However, as we know that the resolvers identified by the query only issue queries of a single qtype (see the `HAVING`-clause), we can apply any group function on the qtype-column without interfering with the result.

(d) Exclude resolvers requesting mainly distinct domains It can be assumed that a resolver shared by multiple people queries well known domain names more often than less popular domain names. To be more precise, more popular domain names are queried multiple times a day after the TTL of cached results has expired, whereas less popular domain names are queried just once or just a few times a day. A scanning resolver queries every domain name presumably just one single time during its scanning period. Also a resolver querying just a small number of domain names very often is suspicious. These resolvers are most probably monitors measuring the activity of a certain system. Therefore also the percentage of distinct domain names queried is an indicator for the kind of usage of a resolver.

It can be expected that three different types of resolvers can be identified by this measure:

Crawlers These resolvers query a high percentage of distinct domain names, as typically every domain name is just queried once to get information about its existence or DNS data.

Users These resolvers are actually used by human users, e.g. for surfing the web, and show a medium percentage of distinct domain names. This is due to the fact that some domain names are only queried once in a while, whereas frequently used domain names are queried more often.

Monitors These resolvers monitor certain systems and therefore send a lot of queries regarding just a small amount of domain names. Therefore the percentage of distinct domain names queried is low for these resolvers.

Listing 3.7 shows the query used to obtain the percentage of distinct domain names queried by every resolver.

Listing 3.7: The SQL query to obtain those resolvers which are mainly querying distinct domain names.

```
SELECT src, COUNT(1) AS tot, (COUNT(DISTINCT domainname) * 100) / COUNT
    (1) AS prcnt_distinct
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND rcode = 3
AND ipv = 4
AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
GROUP BY src
```

(e) Exclude resolvers with mainly NxDomain responses Another way of identifying possible abuse of the DNS system is to analyze the percentage of successful resolved queries. A lot of resolvers mainly send queries resulting in NxDomain responses, which should not be the case when the resolver is used by legitimate users as it can be assumed that users mainly look up existent domain names and just make a typo once in a while. This behavior would result in a low percentage of NxDomain responses with regard to the total number of queries.

Resolvers with a high percentage of NxDomain responses are thus most probably not used by actual human users and can be identified by issuing the query shown in Listing 3.8 to ENTRADA.

Listing 3.8: The SQL query to obtain those resolvers which are mainly issuing queries resulting in NxDomain responses.

```
SELECT suc.src AS src, suc.amount AS success, nx.amount AS nxdomain, (
    suc.amount + nx.amount) AS tot, (nx.amount * 100) / (nx.amount +
    suc.amount) AS prcnt_nx
FROM (
  SELECT src, COUNT(1) AS amount
  FROM dns.queries
  WHERE year=2016 AND month=6 AND day<=7
  AND rcode = 0
  AND ipv = 4
  AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
  GROUP BY src
) suc
INNER JOIN (
  SELECT src, COUNT(1) AS amount
  FROM dns.queries
  WHERE year=2016 AND month=6 AND day<=7
  AND rcode = 3
  AND ipv = 4
  AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
  GROUP BY src
) nx
ON (suc.src = nx.src)
ORDER BY prcnt_nx DESC
```

(f) Exclude resolvers scanning the zone in a linear fashion Although the previously defined measures can be used to filter a lot of the queries which are not originating from actual users, still some resolvers performing linear scans of the whole .nl-zone are missed. Therefore these resolvers were manually identified in three subsequent steps:

First the busiest ASs were identified using the query previously seen in Listing 3.4.

Next, the busiest resolvers within those ASs were identified using the query displayed in Listing 3.9.

Listing 3.9: The SQL query to list the busiest resolvers within a specified autonomous system. The variable `<asn>` has to be replaced with the ASN of the resolver to be investigated.

```
SELECT src, COUNT(1) AS tot
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND asn="<asn>"
AND rcode = 0
AND ipv = 4
AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
GROUP BY src
```

```
ORDER BY tot DESC
```

In the next and last step, the actual DNS queries originating from these busiest resolvers were manually inspected in order to identify linear scans of the zone. This was done using the query shown in Listing 3.10 which displays the DNS queries issued by a single resolver ordered by the timestamp these queries were received at the two name servers.

Listing 3.10: The SQL query to retrieve the DNS queries sent by a single resolver ordered by the timestamp of receiving the query. The variable <ip> has to be replaced with the IP address of the resolver to be investigated.

```
SELECT qname, qtype, unixtime
FROM dns.queries
WHERE year=2016 AND month=6 AND day<=7
AND src = "<ip>"
AND rcode = 0
AND ipv = 4
AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
ORDER BY unixtime
```

Of course there are too many distinct resolvers which send too many queries to the DNS in order to investigate all resolvers and all queries. However, only investigating those resolvers that send the most queries in an AS turns out to be enough, as the other resolvers account for too less queries to influence the total amount of queries received from that AS significantly.

A bigger problem is the fact that not all queries of all resolvers could be investigated by hand. Therefore the chances are high, that scanning resolvers were missed (false negatives) or resolvers which mainly send legitimate queries and just a few queries for scanning purposes were identified as scanning resolvers (false positives).

All these difficulties, and those explained previously eventually resulted in the choice of the twofold approach of investigating both, those resolvers issuing the most queries and those resolvers issuing the most legitimate queries originating from human users.

(g) Combine everything to a single query The final list of resolver locations, where those resolvers sending disproportionately many queries are excluded was retrieved by combining and optimizing the previously mentioned queries from steps (a) to (f), resulting in a single SQL query shown in Listing 3.11. The list of manually excluded resolvers which has to be fed into the query was obtained by repeatedly performing the in step (f) mentioned steps to identify suspicious resolvers until no more malicious resolvers could be identified in the topmost ASs.

The used thresholds of 90% as an upper bound and 1% as lower bound for distinct domain names queried for legitimate resolvers are motivated by investigating the 'normal' behavior of a resolver, shown in Figure 4.5a.

For the percentage of NxDomain responses, an upper bound threshold of 10% was being used. This threshold is again motivated by looking into the turning point in Figure 4.6.

The list of manually excluded resolvers and the reasons they were excluded is described in Section 4.2.2.

Listing 3.11: The final SQL query to obtain a list of the most important ASs in terms of DNS queries originating from resolvers within those ASs. The variable <src_list> has to be substituted by a list of IP addresses belonging to those resolvers which should be manually excluded from the result.

```
SELECT asn, COUNT(1) AS tot
FROM (
  SELECT asn
  FROM dns.queries
  WHERE year=2016 AND month=6 AND day<=7
  AND asn != "UNKN"
  AND rcode = 0
  AND ipv = 4
  AND qtype IN (1, 2, 5, 6, 12, 15, 16, 28)
  AND src NOT IN (
    SELECT src
    FROM dns.queries
    WHERE year=2016 AND month=6 AND day<=7
    GROUP BY src
    HAVING COUNT(DISTINCT qtype) = 1
    AND MIN(qtype) = 2
  )
  AND src NOT IN (
    SELECT src
    FROM dns.queries
    WHERE year=2016 AND month=6 AND day<=7
    GROUP BY src
    HAVING COUNT(1) >= 10000
    AND (
      (COUNT(DISTINCT domainname) * 100) / COUNT(1) >= 90
      OR (COUNT(DISTINCT domainname) * 100) / COUNT(1) <= 1
    )
  )
  AND src NOT IN (
    SELECT src
    FROM (
      SELECT suc.src AS src, suc.amount AS success, nx.amount AS nxdomain
      FROM (
        SELECT src, COUNT(1) AS amount
        FROM dns.queries
        WHERE year=2016 AND month=6 AND day<=7
        AND rcode = 0
        GROUP BY src
      ) suc
      INNER JOIN (
        SELECT src, COUNT(1) AS amount
        FROM dns.queries
        WHERE year=2016 AND month=6 AND day<=7
        AND rcode = 3
        GROUP BY src
      ) nx
      ON (suc.src = nx.src)
    ) tmp
```

```
    WHERE (success + nxdomain) >= 10000
    AND (nxdomain * 100) / (nxdomain + success) >= 10
  )
  AND src NOT IN (<src_list>)
  GROUP BY asn, src, qname, qtype, (unixtime div 300) * 300
) AS queries
GROUP BY asn
ORDER BY tot DESC
LIMIT 30
```

As can be seen, this query gets rather complicated and due to the beforehand mentioned problems still misses some crawlers and other abnormal use of the DNS. Future research may be performed to be able to classify DNS queries according to their origin with higher accuracy.

The obtained list of ASs where the most legitimate queries originate from is given in Section 4.2.

The complete list

The final list of resolving ASs investigated was obtained by combining the results of both described approaches. This list in given in Table 4.6.

3.1.5 Obtaining network topology and graph creation

This is the most difficult part of the research, as there simply is no data perfectly representing the network topology. This is mainly due to the earlier mentioned fact that the Internet is a global self-managing network without an authority which does know everything about the connected parties and their relationships to each other.

To obtain an estimate of the network topology mainly the AS relationship dataset provided by CAIDA [57] was being used. This dataset provides connections inferred from several sources such as the RouteViews project [85], RIPE RIS [66] and CAIDAs ark monitors [56]. As such it provides, to the best of my knowledge, the most complete view on the network topology of the Internet currently available. For detailed information and further background reading about the methodology used to infer this dataset, please refer to [57, 12, 13, 16]. However, this dataset also has its limitations which have to be known when working with it. Therefore Section 3.2.2 lists the problems associated with the network topology inferred with this dataset.

A big advantage of the CAIDA dataset in contrast to other available datasets (such as the 'Neighbors'-tool of RIPEstat [63]) is that this dataset also provides the relationships between two interconnected ASs, enabling us to limit our analysis to the paths on which traffic might be exchanged, rather than analyzing all existing connections.

However, this dataset does not specify the locations, e.g. certain IXPs, at which peering between two neighboring ASs takes place. This information is however necessary to be able to analyze the impact on the reachability of ASs when a certain IXP stops working properly. To estimate the peering relationships at an IXP, the members of that IXP where assumed to all peer with each other as there is no reason why not to. Of course, there are some special situations where this simple rule does not apply, however, this assumption had to be made to be able to say something about the impact of IXP failures. Although there might be some backup routes available at other internet exchanges, the original customer-provider infrastructure should be capable of handling such issues (as happened on 13th of May 2015 [95]).

This data was being used to create a network topology graph as shown in Figure 2.1. The vertices of this graph are given by the autonomous systems, whereas edges represent connections between two ASs. The edges are enriched with certain properties, e.g. whether they are directed or undirected, which are determined by the kind of relationship these edges represent. There are three different types of edges, one per type of AS relationship:

customer-to-provider These edges are the only directed ones. The direction of these edges is given by the money flow involved, i.e. the source of the edge is at the customer's side whereas its destination lays at the provider's side.

peer-to-peer These edges are undirected and labeled with `peer`. If there is evidence that this peering takes place at a certain IXP, a second label will be added to the edge representing this IXP.

sibling-to-sibling In a real world scenario some edges of this type exist. However, unfortunately the used dataset misses these relationship. Nevertheless the developed method is capable of handling relations of this type to facilitate future research when better data gets available. Edges representing sibling-to-sibling relations are also undirected but labeled with `sibling`.

This mixture of directed and undirected edges with certain labels is necessary to be able to distinguish valid and invalid paths between two indirectly interconnected autonomous systems. This however also complicates the analysis of the graph as the standard graph libraries which could be used for graph analysis can handle either directed or undirected graphs. Therefore a special library had to implemented to perform the required analysis. For more details on this implementation please refer to appendix C.3.

3.1.6 Simulation & Selection of failure scenarios

In the resolution process of a DNS request three different sources of possible network failures can be identified:

1. Failure of the resolver
2. Failure of any part of the network necessary for transmitting the request to the authoritative name server, i.e. transit providers or connections

3. Failure of the authoritative name server

As the first failure can not be analyzed using the described approach and available data as every other autonomous system and domain name would have to be counted as unreachable from a non-existing resolver, this case is not considered in the following analysis. However, it might happen that failure of an AS is simulated which serves both sides of the DNS request, the resolver and the authoritative name server. In these cases no conclusions can be drawn due to the reasons mentioned before and this particular AS is therefore not considered as a resolver location in these simulations.

The second and third failure scenarios, however, can be modeled by the described approach. In an analysis on the autonomous system level, failure of parts of the network infrastructure can be modeled by removing ASs or connections between ASs from the network topology used as a model. Incidents resulting in failure of all authoritative name servers located in an AS can be modeled by removing that AS from the topology.

After removing the interesting parts from the network the availability and length of the paths between resolver locations and name server locations are investigated as described in Section 3.1.7.

Selection of scenarios

Ideally one would want to know the impact of failure of any part of the infrastructure. However, due to the huge amount of possible points of failure and the limited computational means during the analysis, it was necessary to focus on some scenarios. It was therefore chosen to focus on those scenarios which have presumably the highest impact on the DNS.

The ASs whose failure is analyzed are selected on basis of the number of domains hosted in those ASs. This choice is based on the assumption, that failure of those autonomous systems hosting the most domain names has the greatest impact on the availability of the DNS data. Another approach taken is to select those ASs which are most commonly traversed during the computation of the baseline. These ASs seem to be important transit providers and failure of these might thus have a big impact on the availability of paths between resolvers and name servers.

To analyze the failure of an IXP, especially the AMS-IX was chosen. This is the largest internet exchange point of the Netherlands and the majority of the Dutch ASs are members of the AMS-IX. As it is not exactly known who peers with whom at the AMS-IX, all peering connections between participants are removed from the topology to simulate this scenario. In theory there is no reason for the members of the AMS-IX not to peer with each other and so it can be assumed, that all ASs are peering with each other.

A last categorie of analyzed failure scenarios is failure of individual connections. The simulated failures are based on the number of occurences of connections on shortest paths discovered during the baseline computations. These connections are assumed to be importtant for the DNS traffic. Future research might use other selection mechanisms, one of which would for example be to cluster the logical links between autonomous systems, such that they map onto physical cables. One could then simulate the impact of accidental cutting of a cable.

Co-location of name servers might also be an interesting factor for future analysis. However, as no locations of data centers are taken into account in this analysis this is left for future research. By this one would also be able to analyze failure of a data center by investigating failure of the name servers located that data center rather than all name servers located in an autonomous system. Furthermore, as a data center might host servers located in various autonomous systems, this would add another aspect to the analysis.

3.1.7 Analysis of simulations

The simulated failures were analyzed by investigating whether all ASs involved in the resolution process of a DNS query are still reachable in the partial network graph from the viewpoint of the ASs the most commonly used resolvers are located in. Reachability is defined as: there is a path between the resolver AS and the name server AS which fulfills the constraints given on an AS path as described earlier, i.e. it consists of zero or more upstream links, followed by zero or more peering links, followed by zero or more downstream links. Sibling links may occur at any position.

For investigating whether a valid path exists between two vertices of the AS graph, breadth-first search is being used as search algorithm. This choice was made as this algorithm is guaranteed to find the shortest path between two nodes in a graph if it exists [2]. Other, more sophisticated search algorithms as for example A* would require a heuristic to estimate the distance between a node and the goal [18]. However, such a heuristic does not exists in this case due to the structure of the network. Due to the huge size of the AS topology graph and the great amount of paths which have to be investigated it is necessary to implement some optimalizations of the general breadth-first search:

- **Limiting path length:** Not all pairs of nodes in the graph are connected by a valid path. Due to the large number of possible paths, a maximum length had to be defined to stop the search at a given point. Given the amount of paths which had to be investigated and the available hardware it was decided not to investigate any paths containing more than 20 vertices. Ignoring paths of a certain length is justifiable as also the IP packets containing the DNS queries and responses on the network layer contain a Time To Live and do not traverse an infinite amount of machines before being discarded. However, due to the high connectivity of the graph, most of the autonomous systems are connected by multiple links of size less than 6.

- **Marking of already seen vertices:** Once a node is reached, this node is being marked such that a vertex is not considered as possible next hop if it was already found on a shorter path. However, it is also necessary to

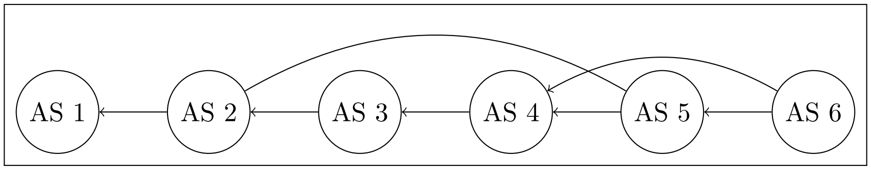


Figure 3.1: Illustration of a situation where the shortest path from AS6 to AS1 is invalid whereas a longer path is valid. The arrows indicate money-flow, i.e. AS1 is a provider of AS2, AS2 is a provider of AS3, etc. Furthermore, AS2 and AS5 have a peering contract. In this situation the path AS6→AS4→AS3→AS2→AS1 would be valid. However, the shorter variant AS6→AS5→AS2→AS1 would be invalid as AS2 would have to provide transit traffic without being paid for it. Using simply breadth-first search with marking of already seen nodes one would first investigate the path AS6→AS5→AS2 and hereby mark AS2 as already seen. The valid path starting with AS6→AS4→AS3→AS2 investigated later would thus not be considered because already a shorter path to AS2 is known. By also storing the relation used to reach a certain node in the marker, this problem can be circumvented as AS2 would be marked with 'peer' in the first case and 'provider' in the latter case. Although this adds some complexity compared to simple marking with boolean values, some paths can be pruned by this, e.g. after finding the path AS6→AS4, paths starting with AS6→AS5→AS4 will not be considered anymore, as a shorter path with same validity features is already known.

store the relation by which one came to the node as some shorter paths may be invalid whereas a longer path using another relation to reach a node may be valid. This problem is illustrated in Figure 3.1. Storing the last relation on the path which was not a sibling-to-sibling relation is enough as the allowed next hops can completely be determined by this, as follows directly from the structure of valid paths for traffic flow between two distinct ASs.

- **Caching of baseline paths:** When simulating failure scenarios one might skip a lot of the calculations of the shortest paths from one autonomous system to another by investigating the shortest paths discovered during the baseline computations, i.e. by investigating the shortest path discovered in the full topology. If this path is not affected by the simulated failure scenario, i.e. all autonomous systems and connections on these paths are still contained in the reduced topology, this path will also always be the shortest path in the reduced topology. This optimization might drastically reduce the number of shortest paths that have to be calculated and thereby also drastically decrease the computation time necessary for a simulation.

Next to analyzing whether there is a path between two ASs, also the length of the shortest path was analyzed as a quality measure of the network. This may reveal some ASs used for storing DNS data which are poorly connected to those ASs requesting this data. As another quality measure also the amount of different paths was considered to be investigated but this investigation was discontinued due to performance issues on the large scale of the whole Internet.

3.2 Limitations

While performing this research we struggled with some limitations of the methods and available data which might be circumvented in future work. As these might influence the accuracy of the results it is important to fully understand these. Therefore, this section describes the encountered problems in detail.

3.2.1 Network location of resolvers

As previously described, the initial idea to identify the most important resolving ASs was to simply count the number of DNS queries originating from a single AS. However, this did not seem to work in practice as there are some resolvers which are used for scanning the whole zone and therefore produce disproportionately much traffic on the name servers of the .nl-zone. Therefore it was decided to investigate these resolvers and exclude them from this study. The downside of this approach is, that also the legit queries originating at those resolvers are not taken into account.

This does however not seem to be a great issue, as it may be assumed that the big autonomous systems use several different resolvers. This means that excluding a small amount of the resolvers should not have such a huge impact on the overall traffic originating at a certain AS. Nevertheless, a better approach to tackle the problem of resolver scanning the whole zone would be to exclude just the queries belonging to the scan, instead of all queries of that resolver. This is however too complicated to be done within the scope of this study and may be done in further research.

Also, the conclusions over the availability of the DNS data in the network topology are drawn from the perspective of those ASs where the majority of DNS queries originate from. However, there are also a lot of ASs where only a minority of the total amount of DNS queries on the .nl-zone originate. The users of these resolvers are not taken into account for this analysis. This abstraction was necessary to give the study a clear scope. In future work, the availability of DNS data can be analyzed from arbitrary resolver locations.

3.2.2 Network topology

The greatest issue is contained in the process of obtaining the network topology. As previously described, not all data about the network topology is publicly available and therefore we must rely on those sources which provide publicly available data and draw conclusion from that to infer the full topology. However, this comes with some limitations. For example, although the utilized RIPEstat database uses probes at several locations all over the world, it does not give a full overview of the full network. Roughan et al. [33] give a detailed summary of the problems associated with inferring the AS topology by means of BGP traffic. The main problem is that the BGP protocol was never intended to reveal the network topology [33]. Furthermore, due to the information-hiding nature of the protocol, it is impossible to get a full view on the connections within the Internet using this measure. However, as there is no authority for the internet, it is in general impossible to get a complete map of these connections [36].

Another problem is that the data available at different sources which we have to combine in order to draw our conclusions about the topology are collected at different points in time. This problem can partly be circumvented by trying to synchronize the data as well as possible, but as the network is constantly changing, some errors will be introduced.

Furthermore we only analyze logical connections between autonomous systems, rather than physical links. The latter would be more interesting as it is possible that several logical connections can originate from the same physical link. Also it is possible that several physical cables are bundled in the same underground pipe, which might get damaged during construction works. However, this information is kept secret by the owners of physical infrastructure.

One last problem in obtaining the network topology is the usage of anycast. Since some ASs might use anycasting there might appear logical connections in the obtained topology which are physically distributed over two different networks.

Future research might thus use more sophisticated methods to retrieve the network topology or at best use insider information of all involved ISPs in order to retrieve a full view on the physical infrastructure.

3.2.3 Graph analysis

Due to the computational constraints given by the available hardware and the time available for this thesis, just a few failure scenarios are investigated. Ideally one would like to know the impact of all possible incidents in order to configure the network in such a way that the impact of these incidents can be reduced to a minimum. Future research might thus either use better hardware or apply more optimizations to the code in order to simulate more failure scenarios.

Chapter 4

Results

This chapter summarizes the results obtained during the analysis. Its structure basically follows the steps of the research mentioned earlier, but does not exactly match these as it also gives some other insights which were obtained during the analysis and are closely related to answering the research question.

4.1 Domains and name servers

4.1.1 Analyzed domains

In theory all second level domains within the .nl-zone were analyzed. In the morning of the 2nd of June 2016, when the zonefile was retrieved, the number of registered domain names within the .nl-zone was 5,626,381. However, some domain names were excluded from the analysis due to several errors within their configuration or while retrieving the necessary information. As previously explained the resolution process was repeated 5 times in case of occurrence of any error during the process to obtain a higher accuracy for the classification of errors. Domain names were excluded from the analysis in either of the following situations:

1. NoError: It happens quite often that no answer on a DNS query can be retrieved, although no particular error occurred. This is the case when a name server who is marked as authoritative for a certain domain name in the zonefile does not have any resource records regarding that domain name, although it knows that a domain name exist as otherwise a Servfail would be sent back.

2. Servfail: This error occurs when no answer at all was retrieved from the queried name server. This might be due to lame delegation. This is the case when a name server who is marked as authoritative for a certain domain name in the zonefile does not have any data regarding that domain name [30]. Other reasons for this error include network failures or when the

server is too busy to respond within the timeout period. Errors occurring due to the last two reasons can sometimes be resolved by repeating the resolution process.

3. NxDomain: This can happen in extraordinary cases when a domain was either removed from the zone or put in quarantine in the time between retrieving the zonefile and querying the name servers of a domain.

4. No `A` record for any name server: In some cases, there is no `A` record configured for the domain name belonging to a name server of a certain domain name. In this case the IP address of the name server cannot be resolved and therefore it is not possible to tell in which autonomous system this name server resides. In this case the name server was excluded from the analysis and if there are no name servers for a domain name, that domain name has to be excluded as well.

5. ASN look-up failed: It can happen that an IP address is not included in the MaxMind dataset providing the mapping of IP addresses to AS numbers. As previously said, the RIPEstat API was queried in these cases. However, it turned out that also this database does not contain the IP addresses missing in the dataset provided by MaxMind. Therefore these IP addresses had to be excluded from the analysis. Therefore the name servers belonging to that IP address had to be excluded and again, if there are no name servers left for a domain name, that domain name has to be excluded as well.

Table 4.1 shows the total amount of analyzed domain names and the numbers of excluded domains due to the beforehand mentioned reasons. The numbers are not surprising as there are some possible explanations for domain names not being resolvable:

1. When resolving the `A` records of a domain name belonging to a name server, NoError can occur when name servers are configured for IPv6 only. In this case no `A` record exists, although the name server is aware of the existence of the domain name. The same would hold for domain names that are only used for email transport. These do have an `MX` record, but no `A` record.

2. A lot of domain names are only registered but not used by their owners. Therefore they are also not well configured and as SIDN does not perform any checks on the configuration of a domain name, Servfail responses are quite common.

3. Some registrars remove the DNS entries of domain names after these have been resigned. However, the domain names stay registered at SIDN until the end of the contract period.

4. The zonefile is constantly changing. Therefore some domains might have been removed between the time of retrieval of the zonefile and the time

Table 4.1: Amount of name servers (a) and domains (b) analyzed during this research. Also the number of name servers and domain names which were excluded from the analysis due to the described reasons are given. Please note that a name server is here identified by the domain name belonging to that name server, rather than by its IP address as the tables lists errors during resolution of those domain names.

(a) amount of analyzed name servers

Aspect	**Amount**
Total number of name servers	69,996
Error during resolution	1,936
– NoError	48
– Servfail	249
– NxDomain	1,639
ASN unknown	31
Total number of name servers investigated	68,029

(b) amount of analyzed domains

Aspect	**Amount**
Total number of domains in zonefile	5,626,381
Error during resolution	259,364
– NoError	9,219
– Servfail	245,419
– NxDomain	4,726
No name servers included in analysis	2,229
Total number of domains investigated	5,364,788

of resolvong individual domain names. However, this should only be the case for a small subset of the NxDomain errors.

4.1.2 Distribution of DNS data

This section is meant to give some insight into the distribution of DNS data over servers, autonomous systems and geographic locations.

Distribution of domains over name servers

Surprisingly, a rather small amount of name servers is utilized to serve the roughly 5.6 million domain names within the .nl-zone. Looking at the name servers specified for all domain names (excluding those which do not resolve to IP addresses) one sees only 68,060 distinct name servers. After further investigation it turned out, that some of the domain names of name servers point to the very same IP address, such that in fact only 45,031 distinct name servers are responsible for serving the whole .nl-zone. In general, there are only a few name servers that host most of the domain names and most of the name servers host just a small amount of domain names. On average every name server serves 310.8 domain names, with a median of 5 domain names per name server. This low median also indicates that there are a lot of name servers serving just very few domains, whereas the high mean indicates that there are also some name servers which host a very large amount of domain names. The name server hosting the most domains serves 437,228 domain names, which is roughly 8.15% of the whole zone. Not surprisingly, those name servers hosting the most

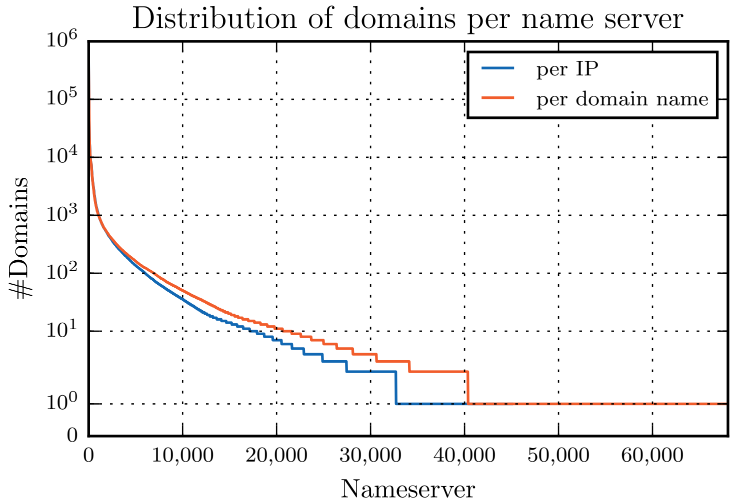


Figure 4.1: Distribution of the amount of domain names hosted on distinct name servers. Every position on the x-axis represents one of the analyzed name servers, whereas the y-axis shows the number of domain names hosted on that server. Distinct name servers are here specified as either having distinct IP addresses (blue) or distinct domain names (orange). Please also note the logarithmic scale of the y-axis.

domain names belong to big Dutch hosting- and infrastructure providers, such as Schuberg Philis, TransIP, Hostnet and LeaseWeb.

Figure 4.1 shows the number of domains hosted on distinct name servers, ranked by the number of domains hosted on a single server. A name server is in this figure either specified by its IP address or by its associated domain name. As the area under both given curves, which represents the number of hostings, is identical in both cases the figure also illustrates that some domains belonging to name servers map to the same IP address. This figure again shows that almost 75% of the name servers host less than 10 domain names, whereas only roughly 10% host more than 100 domain names.

Distribution of name servers over ASs

Figure 4.2 shows the distribution of name servers per autonomous system. This figure is also ranked by the total number of name servers located in an autonomous system and shows, that also on the autonomous systems level, there are some very important ASs in terms of the amount of name servers located within that AS, but also a lot of ASs providing only a very small portion of the name servers of the .nl-zone. In total 3,806 autonomous systems are incorporated in hosting the domain names within the analyzed part of the .nl-zone.

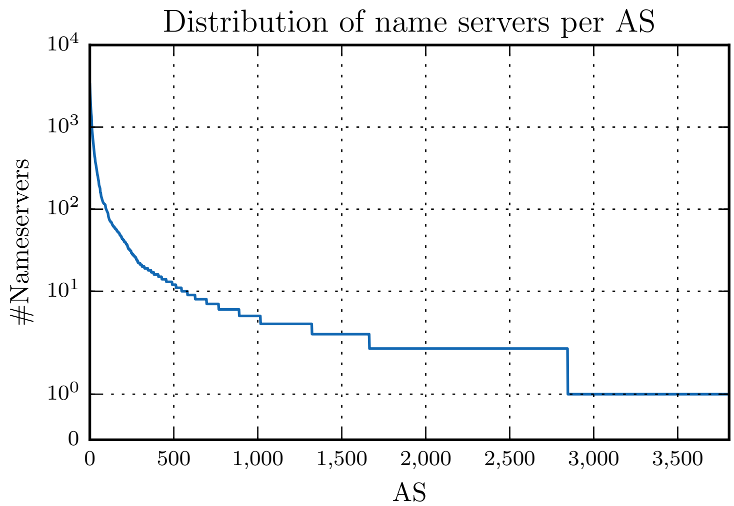


Figure 4.2: Distribution of the amount of name servers located in distinct autonomous systems. Every position on the x-axis represents one autonomous system, whereas the y-axis shows the number of name servers within that AS. Please also note the logarithmic scale of the y-axis.

Distribution of domains over ASs

Another interesting aspect is the distribution of domains over autonomous systems. This distribution is depicted in Figure 4.3 and roughly follows the previously shown distributions. This figure shows the number of domains by the autonomous systems their name servers are located in. This figure is ranked by the total number of domains served by name servers located in an autonomous system.

Although SIDN recommends to place the name servers of a domain in different autonomous systems [55], a lot of domains have all their name servers in a single AS. In total, this is the case for 1,715,627 domain names which is roughly 31.98% of the analyzed domain names.

For a subset containing 16,095 of these 1,715,627 domain names, the situation is even worse because these are, against the technical requirements of SIDN, served by a single name server. This is possible by mapping the domain names of the two name servers, which have to be reported to SIDN, to the very same IP address. This is of course not a desirable situation as this also means that there is no backup name server in case of failure of the first one.

These are however not the only cases when the DNS records do not match the recommendations or requirements of SIDN. Section 4.1.3 therefore provides some more insight into other types of configuration errors within the .nl-zone and also quantifies these errors in terms of the amounts of affected domain names.

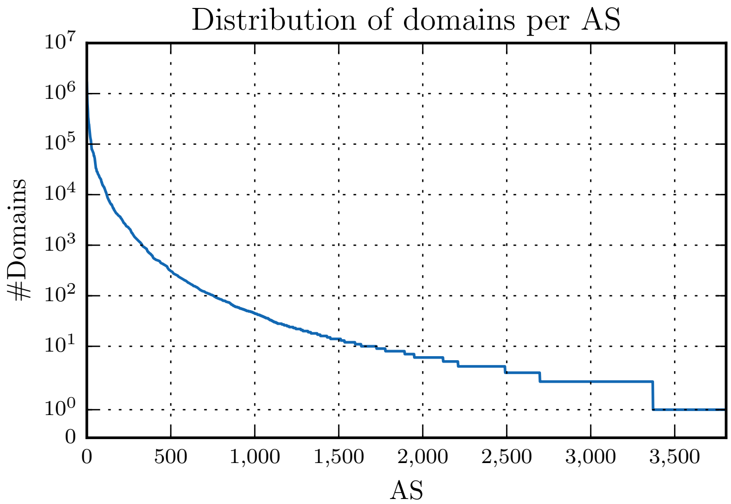


Figure 4.3: Distribution of the amount of domain names hosted on name servers located in distinct autonomous systems. Every position on the x-axis represents one autonomous system, whereas the y-axis shows the number of domain names hosted on servers within that AS. Please also note the logarithmic scale of the y-axis.

Distribution of DNS data over countries

Although the distribution of DNS data over geographical locations is not important for this study, it is interesting how many of the .nl domains and associated name servers are actually located within the Netherlands. Figure 4.4 shows the geographical locations of the name servers hosting domains within the .nl-zone and the locations were the most domains are hosted. This mapping is done using the GeoLite Country database by MaxMind [76]. Unsurprisingly, most of the domains are hosted on name servers that are located within the Netherlands, as they belong to Dutch Internet Service Providers (ISPs). However, a large portion of the domains is also hosted in Germany, most of them hosted by Strato (~34.65%) and Hetzner Online (~26.58%), and the United States of America where the most domain names in the .nl-zone are hosted by Akamai (~10.78%). Other Countries only account for a small amount of the name servers serving the .nl-zone.

An interesting observation is, that although more than 75% of the domain names within the .nl-zone are hosted in the Netherlands, less than 50% of the name servers are geographically located there. This indicates that those name servers hosting a lot of the domain names are located within the Netherlands which makes sense, as these are operated by the big Dutch hosting providers. Furthermore there are a lot of name servers located in unspecified countries, however, these account for only a small portion of the domain names.

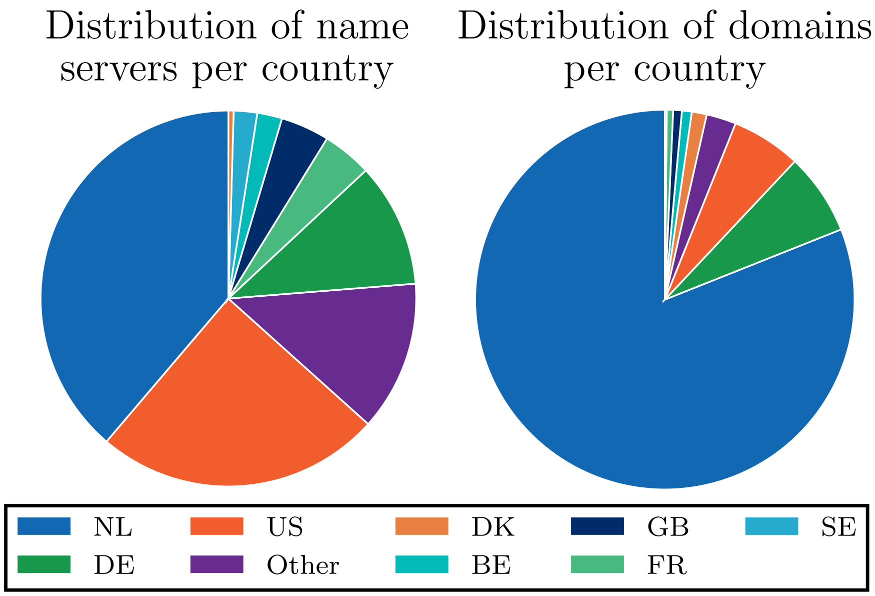


Figure 4.4: Distribution of name servers and hosted domain names per country. Please note that a single domain name might be hosted in multiple countries when the corresponding name servers are located in different countries.

4.1.3 Configuration errors

The data provided in the zonefile of .nl and the data retrieved by querying the zone a domain name resides in should theoretically be identical. However, during this research a lot of mismatches were discovered which might slow down the resolution process of a domain. One example for this type of mismatches is that the list of name servers specified for a domain name in the .nl zonefile does not exactly match the set of name servers specified by the domain itself, i.e. the set of name servers which can be retrieved using `dig NS <domainname>`. Such a mismatch might however not only affect the efficiency of the DNS lookups, but might also form a security risk in some special cases. If, for example, the zonefile specifies a non-existent name server which resides in an unregistered zone, a person with malicious intents might register the domain name belonging to that zone, set up his own name server for the misconfigured domain name and forward incoming requests to his own server, e.g. for phishing purposes.

These mismatches should be investigated by the owners of the corresponding domain names. A complete list of domain names for which configuration errors were found is too long to be published in this thesis, but was handed out to SIDN for further analysis. However, we can provide some insight into the frequency these mismatches occur.

As some of these mismatches might be the result of updates which are taken

Table 4.2: Discovered configuration errors when comparing the contents of the zonefile of .nl with the data provided by individual domains.

Error	**#occurrences**
All name servers in single AS:	1,715,627
– single IP	16,095
Name server mismatch	16,394
– more name servers in zonefile	16,380
– less name servers in zonefile	943

out in between the time of retrieving the zonefile and the `NS` records associated with the contained domain names, this analysis step should be performed twice in future research. The intersection of the results of both runs then provides the set of domain names which are not properly configured.

Table 4.2 shows the amount of errors grouped by the previously described types.

4.2 DNS Resolver locations

This section shows the origins of DNS queries received at the authoritative name servers of the .nl-zone. These are the locations of resolvers used throughout the analysis of reachability of DNS data in case of failure scenarios.

4.2.1 Top locations based on all queries

Table 4.3 shows the number of queries originating from the top ASs. As can be seen in this table, the list contains the big ISPs as well as hosting parties and content provider networks.

4.2.2 Top locations based on human-usage generated queries

Resolver usage

One of the investigated measures to filter out resolvers automatically scanning the zone was the ratio of distinct domain names queried. This ratio was investigated as percentage of distinct domain names queried with respect to the total amount of queries received. This measure turned out to be very good for classifying resolvers according to the way they are being used. In general, the previously mentioned three different types of resolvers, crawlers, users and monitors can be distinguished using this measure.

Figure 4.5 shows the percentage of distinct domain names queried per resolver. Figure 4.5a shows this percentage for all resolvers, regardless of the AS they are located in, whereas Figure 4.5b shows the same for resolvers located in the autonomous system of Google (AS15169). The three steps in the latter

Table 4.3: The top 30 most commonly used autonomous systems as resolvers which were investigated in this analysis. This list was generated by analyzing only the amount of queries received without filtering any resolvers.

Rank	**ASN**	**Owner**	**#Queries**	**%Queries**
1	AS15169	Google Inc.	312,002,843	10.271
2	AS49544	i3d B.V.	292,870,607	9.641
3	AS20857	Transip B.V.	185,141,551	6.095
4	AS35470	XL Internet Services B.V.	156,150,180	5.140
5	AS60781	LeaseWeb Netherlands B.V.	152,175,858	5.010
6	AS8075	Microsoft Corporation	90,334,724	2.974
7	AS24793	NL Hosting Internet	75,276,923	2.478
8	AS8737	KPN B.V.	69,947,928	2.303
9	AS32934	Facebook, Inc.	64,762,352	2.132
10	AS13414	Twitter Inc.	49,701,823	1.636
11	AS17204	Nominum, Inc	45,911,528	1.511
12	AS2637	Georgia Institute of Technology	43,479,222	1.431
13	AS16276	OVH SAS	41,776,230	1.375
14	AS36692	OpenDNS, LLC	39,085,634	1.287
15	AS24940	Hetzner Online GmbH	36,573,903	1.204
16	AS16509	Amazon.com, Inc.	33,020,255	1.087
17	AS1103	SURFnet, The Netherlands	32,803,930	1.080
18	AS9143	Ziggo B.V.	30,235,064	0.995
19	AS13238	YANDEX LLC	27,448,351	0.904
20	AS14618	Amazon.com, Inc.	27,059,825	0.891
21	AS55967	Beijing Baidu Netcom Science and Technology Co., Ltd.	22,501,572	0.741
22	AS6830	Liberty Global Operations B.V.	21,506,235	0.708
23	AS4134	China Telecom Backbone	20,805,449	0.685
24	AS8972	PlusServer AG	16,614,202	0.547
25	AS3356	Level 3 Communications, Inc.	16,117,270	0.531
26	AS34173	SafeBrands S.A.S.	15,495,227	0.510
27	AS13335	CloudFlare, Inc.	14,582,505	0.480
28	AS31615	T-mobile Netherlands bv.	14,357,175	0.473
29	AS36351	SoftLayer Technologies Inc.	13,881,090	0.457
30	AS393406	Digital Ocean, Inc.	13,635,435	0.449
		Sum	1,975,254,891	65.024

graph very well show the resolvers used for different purposes as it clearly shows that the resolvers with prefix `66.249.64.0/19` query way more distinct domain names than resolvers with IP prefix `74.125.0.0/16`. By looking at the reverse DNS entries of IP addresses in both address ranges and looking up documentations of Google's tools, it can be verified that the former address range is used by Googlebot [74], whereas the latter is used for Google's Public DNS Service.

To draw some meaningful conclusions from the percentage of distinct domain names queried, it is necessary that only resolvers with a certain amount of queries in total are being considered. A personal resolver which just sent a single query would otherwise be counted as a crawler since a single query always results in 100% distinct queries. For the analysis and for generating Figure 4.5 therefore only resolvers with more than 10,000 queries in total where considered.

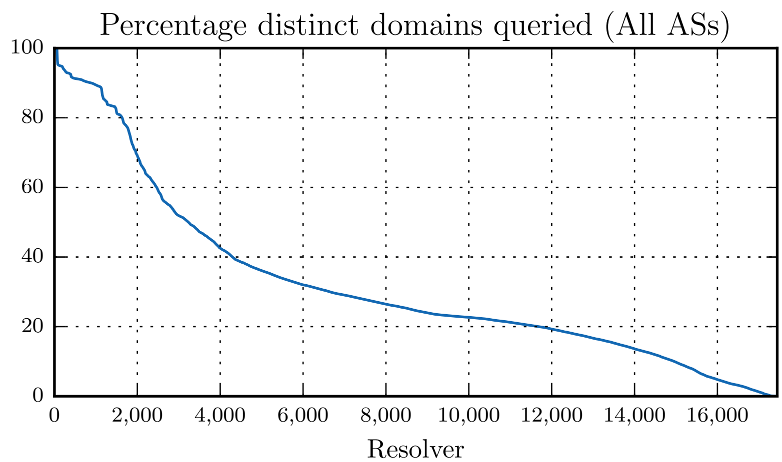
Although this distinction can be made easily, it is hard to identify strong boundaries between the different categories of resolvers. In general, one may assume that resolvers with more than 90% distinct domain names queried, are most probably crawlers. However, this percentage is influenced by a lot of factors. If, for example, the same scan is performed multiple times within the same measurement period, this percentage already drops to only 50%. In the case of Google as shown in Figure 4.5b this is however not the case due to the load balancing applied for the Googlebot. In this application, the scan is distributed over a lot of different machines and IP addresses, whereby the probability that the same machine queries the same domain name again gets very small.

Another reason why it is hard to define strong boundaries in terms of percentages for the different types of resolvers is that different clusters of resolvers use different caching approaches. This is a possible explanation of why the resolvers used for Google's Public DNS service seem to query 40% to 60% (see Figure 4.5b) distinct domain names, whereas the most resolvers located in other ASs seem to query 20% to 40% distinct domain names (see Figure 4.5a). The resolvers provided by Google all share the same cache as described in [73]. Therefore different resolvers within the same cluster tend to send less identical queries to the authoritative name servers.

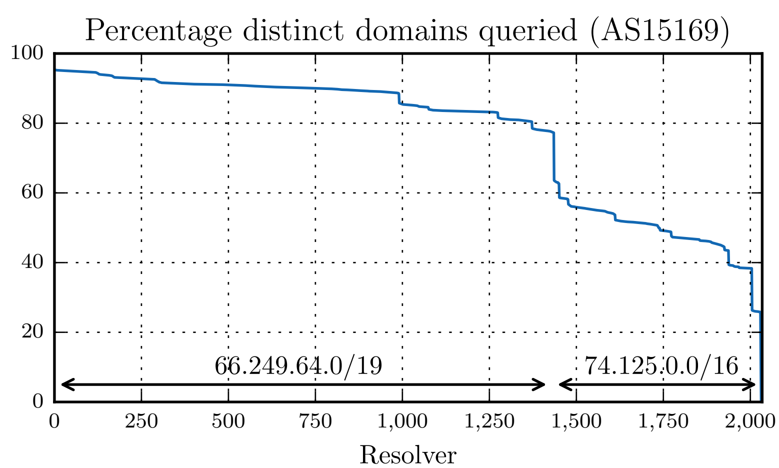
Last but not least it might of course also be the case that the same resolver is used for multiple purposes.

To actually calculate a strong boundary one would thus have to know how often a scan is being performed and over how many machines this scan is distributed as well as how multiple machines share their caches and of course for which purposes the machines are actually used. All these factors influence the measurable results.

In contrast to separating crawlers from actual users, monitors are relatively easy to spot as these typically query only a small amount of distinct domain names very often, which results in a very small percentage of distinct domain names queried, which is typically neither the case for crawlers nor users.



(a)



(b)

Figure 4.5: The percentage of distinct domain names queried in the first week of June 2016 for all resolvers (a) and for resolvers located in the autonomous system operated by Google (b). Note that the shown IP addresses in (b) are purely indicative to illustrate the IP address ranges the resolvers with different percentages reside in and some exceptions exist.

NxDomain responses

Another investigated measure for analysing the origin of DNS queries is the ratio of NxDomain responses to the queries. As previously explained, it can be assumed that users mainly look up existent domain names and just make a typo once in a while. This results in a low percentage of NxDomain responses with regard to the total number of queries.

Some resolvers, however, query a lot more non-existent domain names. This phenomenon is also depicted in Figure 4.6. One possible explanation for this could be spambots sending emails to randomly generated addresses, possibly due to a spamtrap.

Another possible explanation for this behavior is that the resolver is used by a botnet which makes use of a domain generation algorithm to contact the command and control (C&C) server. These botnets randomly generate a long list of domain names on a daily basis where just a few are being registered and linked to the C&C server. This behavior can for example be found in the Conficker worm [52].

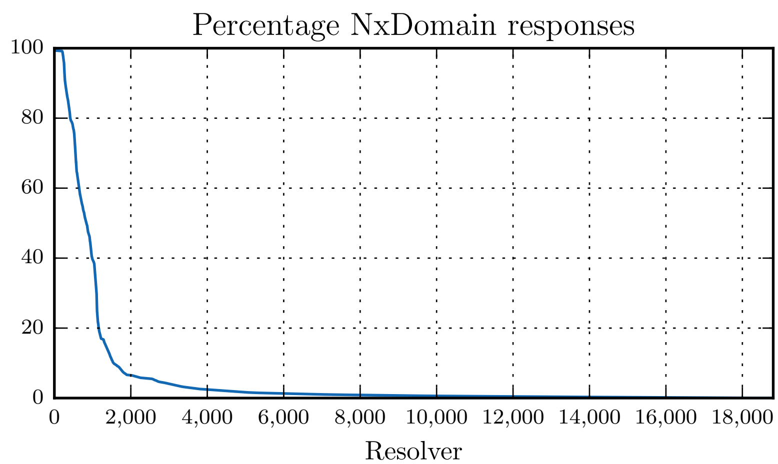


Figure 4.6: The percentage of NxDomain responses on queries per resolver in the first week of June 2016.

By analyzing Figure 4.6 one can see that the normal behavior of a resolver would be to query less than 10% non-existent domain names.

Manually excluded resolvers

Some resolvers are producing disproportionately many DNS queries as they are not used by human users. These resolvers were manually identified and excluded from the analysis. These are resolvers which are being used to generate maps of the whole zone, i.e. they are scanning the whole zone by querying (large portions) of all existent domain names. Table 4.4 shows which resolvers were not investigated and the reason for this.

Locations of legitimate resolvers

Table 4.5 shows the locations of the most important legitimate resolvers and the amount of queries counted.

4.2.3 Investigated resolver locations

As can be seen by comparing the two tables 4.3 and 4.5 the identified resolvers show a lot of overlap. Combining the autonomous systems contained in both rankings, the list of eventually investigated resolver locations is shown in Table 4.6.

This list contains autonomous systems belonging to all kinds of different companies working with the DNS, such as internet service providers, hosting providers, owners of internet infrastructure, search engines and content providers. As such, this list seems to be representative for all use-cases of the DNS.

4.3 AS Topology

The topology of autonomous systems was inferred from the dataset provided by CAIDA. This topology contains 54,466 autonomous systems with 488,140 connections. Figure 4.7 shows a small part of this topology as inferred for the autonomous system managed by SIDN (AS1140).

4.4 Reachability of DNS data

The reachability of DNS data was being analyzed from the viewpoint of the in section 4.2.3 identified most important resolver locations in terms of the amount

Table 4.4: These resolvers were manually excluded from the analysis as they are known to automatically resolve (large parts) of the .nl-zone.

ASN	IP Address(es)	Reason
AS1103	145.0.9.[182-184]	Resolvers used for the OpenINTEL project [100]
AS1140	94.198.159.3	Resolver used during this research
AS2637	130.207.54.[130-160]	Manually identified (distributed) linear scan
AS16276	51.255.67.42	Manually identified linear scan
AS20857	149.210.129.73	Manually identified linear scan
AS24940	148.251.78.115	Manually identified linear scan
AS35470	178.18.83.160	Manually identified linear scan
AS49544	213.163.65.[43-79]	Manually identified (distributed) linear scan

Table 4.5: The top 30 most commonly used autonomous systems as resolvers which were investigated in this analysis. This list was generated by analyzing the amount of queries received after filtering resolvers which automatically generate a lot of queries.

Rank	**ASN**	**Owner**	**#Queries**	**%Queries**
1	AS15169	Google Inc.	154,986,892	13.182
2	AS8075	Microsoft Corporation	54,020,566	4.595
3	AS32934	Facebook, Inc.	36,924,255	3.141
4	AS8737	KPN B.V.	34,713,139	2.952
5	AS14618	Amazon.com, Inc.	25,376,610	2.158
6	AS6830	Liberty Global Operations B.V.	19,575,407	1.665
7	AS16509	Amazon.com, Inc.	15,394,007	1.309
8	AS16276	OVH SAS	15,242,918	1.296
9	AS9143	Ziggo B.V.	14,138,184	1.202
10	AS24940	Hetzner Online GmbH	12,657,640	1.077
11	AS8972	PlusServer AG	11,855,099	1.008
12	AS36692	OpenDNS, LLC	11,737,892	0.998
13	AS13127	Tele 2 Nederland B.V.	11,566,830	0.984
14	AS3356	Level 3 Communications, Inc.	11,243,341	0.956
15	AS13238	YANDEX LLC	10,841,475	0.922
16	AS4134	China Telecom Backbone	10,602,541	0.902
17	AS393406	Digital Ocean, Inc.	9,483,821	0.807
18	AS31615	T-mobile Netherlands bv.	9,481,710	0.806
19	AS9121	Turk Telekomunikasyon Anonim Sirketi	9,406,272	0.800
20	AS36351	SoftLayer Technologies Inc.	9,286,600	0.790
21	AS60781	LeaseWeb Netherlands B.V.	8,704,417	0.740
22	AS23033	Wowrack.com	8,093,633	0.688
23	AS1103	SURFnet, The Netherlands	7,930,617	0.675
24	AS5432	Proximus NV	7,733,331	0.658
25	AS20940	Akamai International B.V.	7,127,354	0.606
26	AS3215	Orange S.A.	6,742,155	0.573
27	AS3320	Deutsche Telekom AG	6,506,647	0.553
28	AS36647	Yahoo	6,300,666	0.536
29	AS13335	CloudFlare, Inc.	6,180,157	0.526
30	AS197902	Hostnet B.V.	6,028,291	0.513
		Sum	559,882,467	47.620

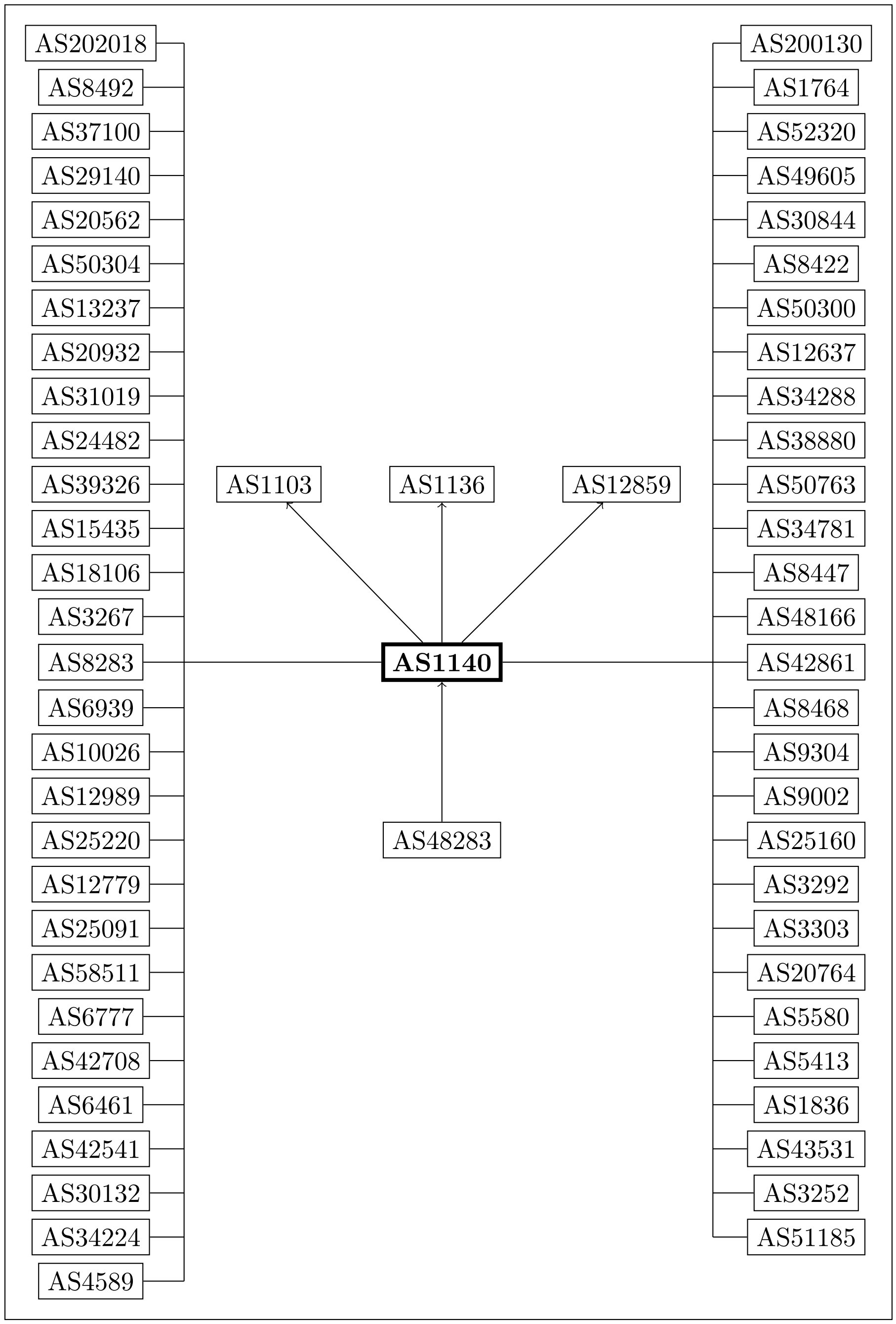


Figure 4.7: Connections of AS1140 (SIDN) in the AS topology provided by the dataset of CAIDA. Again arrows indicate customer-provider relationships, whereas simple lines indicate peering connections.

of unreachable hosting ASs and domains and the mean length of the shortest valid path between the autonomous systems providing the resolver service and the AS hosting the DNS data.

4.4.1 Baseline

Table 4.6 shows the baseline measurements used for the impact analysis. This baseline indicates the reachability of ASs and domains from the viewpoint of the analyzed resolver locations in the graph as acquired from the data set provided by CAIDA. Due to the issues related to the analysis with regard to completeness of the data and computational capabilities of the available hardware described beforehand, it cannot be assumed that all ASs connected to the global internet are reachable from every other AS in the used model. Therefore the conclusions about the impact of shutting down parts of the infrastructure can only be investigated with this baseline in mind.

This baseline analysis suggests that there are some ASs which are not reachable from all resolving ASs and one AS (AS26850) is even not reachable from any resolver AS. By manual inspection of the connectivity of this specific AS one sees that this AS is very poorly connected as it just has one peering connection with one other AS, at least according to the available data. However, this does not have a huge influence on the rest of the analysis, as this specific AS just contains a single name server serving a single domain name in the .nl-zone.

As the mean path length to the closest DNS entry, shown in Table 4.6 suggests, some resolvers have significantly shorter paths to reach a domain name than others. This does however not indicate that autonomous systems with longer paths are generally less well connected to the internet. The mean path length is highly influenced by the length of the paths to those ASs hosting the most domain names and less influenced by the paths to those ASs hosting a low amount of domain names. Nevertheless, it was chosen to analyze the connectivity weighted by the number of domain names served by name servers located in an autonomous system, as just the mean path length to every other autonomous system would be an indicator for the general connectivity and not applied on the case of the DNS. In such a calculation all paths would be weighted equally, however, many autonomous systems host just a single domain name and its reachability is therefore negligible when analyzing the whole zone. Therefore the mean path length is weighted by the number of domain names hosted on name servers located in an AS.

4.4.2 Simulation results

As previously described, four different approaches of selecting failure scenarios are taken:

- Failure of the top 20 ASs in which the most domain names are hosted. These autonomous systems and references to extensive results of these simulations are depicted in Table 4.7.

Table 4.6: The selected resolver locations with reachability baseline in the full topology, without removing any vertices or edges.

Resolver location (ASN)	Owner	Baseline		
		#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	SURFnet, The Netherlands	2	0	1.182
AS2637	Georgia Institute of Technology	3	0	2.189
AS3215	Orange S.A.	6	1	2.836
AS3320	Deutsche Telekom AG	5	1	1.819
AS3356	Level 3 Communications, Inc.	3	0	1.741
AS4134	China Telecom Backbone	1	0	2.099
AS5432	Proximus NV	1	0	2.277
AS6830	Liberty Global Operations B.V.	4	1	1.909
AS8075	Microsoft Corporation	1	0	1.645
AS8737	KPN B.V.	1	0	2.886
AS8972	PlusServer AG	1	0	2.223
AS9121	Turk Telekomunikasyon Anonim Sirketi	1	0	2.097
AS9143	Ziggo B.V.	1	0	1.940
AS13127	Tele 2 Nederland B.V.	5	1	2.037
AS13238	YANDEX LLC	2	0	1.663
AS13414	Twitter Inc.	1	0	1.791
AS13335	CloudFlare, Inc.	1	0	1.645
AS14618	Amazon.com, Inc.	1	0	2.652
AS15169	Google Inc.	4	1	1.651
AS16276	OVH SAS	3	0	1.667
AS16509	Amazon.com, Inc.	1	0	1.825
AS17204	Nominum, Inc	1	0	2.359
AS20857	Transip B.V.	1	0	1.553
AS20940	Akamai International B.V.	1	0	1.621
AS23033	Wowrack.com	1	0	2.500
AS24793	NL Hosting Internet	1	0	2.223
AS24940	Hetzner Online GmbH	1	0	1.611
AS31615	T-mobile Netherlands bv.	1	0	2.001
AS32934	Facebook, Inc.	1	0	1.759
AS34173	SafeBrands S.A.S.	1	0	2.219
AS35470	XL Internet Services B.V.	1	0	2.639
AS36351	SoftLayer Technologies Inc.	1	0	1.495
AS36647	Yahoo	1	0	2.650
AS36692	OpenDNS, LLC	1	0	1.908
AS49544	i3d B.V.	1	0	1.711
AS55967	Beijing Baidu Netcom Science and Technology Co., Ltd.	1	0	2.254
AS60781	LeaseWeb Netherlands B.V.	2	0	1.825
AS197902	Hostnet B.V.	1	0	1.773
AS393406	Digital Ocean, Inc.	2	0	2.140
			Mean	2.000

- Failure of the top 20 ASs providing the most transit connections on paths in the baseline computations. These autonomous systems and references to extensive results of these simulations are depicted in Table 4.8.
- Failure of the top 20 most traversed connections in the baseline computations. These connections and references to extensive results of these simulations are depicted in Table 4.9.
- Failure of all peerings at the AMS-IX. Extensive results can be found in Table B.40 and Figure B.40.

As there is some overlap between between those ASs in which the most domain names are hosted and those ASs providing transit traffic in the most shortest paths on the baseline, in total 60 different scenarios are simulated. As it is not possible to clearly describe the impact on the reachability of DNS data in each simulation, this section summarizes some general observations. Extensive results for every single simulation can be found in Appendix B.1.

Please note that the numbers of autonomous systems and domain names becoming unreachable in certain simulations, shown in the tables 4.7, 4.8, 4.9 and 4.10, are those becoming unavailable unexpectedly. When removing an autonomous system from the topology, it is clear that that AS will become unavailable for all other autonomous systems. The same holds for domain names which are hosted on name servers only located in a single AS. Therefore the given numbers are exclusive the ASs and domains which are expected to become unreachable in a certain simulation or already are unreachable in the baseline computations. If you are interested in the total numbers of unavailable domain names in certain scenarios, please refer to the extensive results shown in Appendix B.1

Failure of ASs

Failures of two distinct types of autonomous systems were being analyzed. Those autonomous systems responsible for hosting DNS data and those which provide transit traffic on the shortest paths between the ASs of the most used resolvers and the ASs hosting the DNS data.

In most of the cases one can say that failure of the first type only results in that AS being unreachable and the amount of unreachable domain names increases with regard to the baseline by exactly that amount of domain names which is only hosted on name servers located in that single AS. Furthermore, the mean length of the shortest path to reach the AS hosting a domain name slightly increases for a subset of the resolvers. In some cases, however, this mean length of the shortest path decreases a little. This is not due to the fact that removing one autonomous system from the topology leads to a shorter path between the resolver location and the name server location, but is rather a side effect of taking the mean path length over fewer domain names as some domain names are not reachable anymore. Sometimes it is also the case, e.g. by removing AS21155, that a small amount of other ASs becomes unreachable.

Table 4.7: The top 20 autonomous systems in terms of the number of domain names hosted on name servers within those ASs and the results of a simulated failure of those ASs. A resolver location is only counted as affected by the simulated failure if the number of unreachable ASs increases by more than the one which was removed in the simulation. The average amounts of unreachable ASs and unreachable domains are only calculated over the affected resolver locations, whereas the mean length of shortest paths is calculated over all resolver locations..

Simulated failure (ASN)	**# Domains hosted**	**# Resolver locations affected**	**Mean unreachable ASs**	**Mean unreachable domains**	**Mean shortest path length**	**Extensive results shown in**
AS20857	1,102,720	0	0	0	2.086	Table B.1 (p. 99) Fig. B.1 (p. 100)
AS60781	910,200	0	0	0	2.049	Table B.2 (p. 101) Fig. B.2 (p. 102)
AS21155	554,539	39	1	1	1.986	Table B.3 (p. 103) Fig. B.3 (p. 104)
AS12859	554,182	0	0	0	2.065	Table B.4 (p. 105) Fig. B.4 (p. 106)
AS8455	491,008	39	5.41	11,474.67	2.046	Table B.5 (p. 107) Fig. B.5 (p. 108)
AS197902	458,068	0	0	0	2.015	Table B.6 (p. 109) Fig. B.6 (p. 110)
AS25151	291,815	0	0	0	2.003	Table B.7 (p. 111) Fig. B.7 (p. 112)
AS24940	248,358	0	0	0	2.022	Table B.8 (p. 113) Fig. B.8 (p. 114)
AS48635	235,129	39	1	34,770	2.003	Table B.9 (p. 115) Fig. B.9 (p. 116)
AS3265	234,512	0	0	0	2.004	Table B.10 (p. 117) Fig. B.10 (p. 118)
AS35470	206,455	0	0	0	1.974	Table B.11 (p. 119) Fig. B.11 (p. 120)
AS15879	199,021	0	0	0	1.998	Table B.12 (p. 121) Fig. B.12 (p. 122)
AS25459	187,562	0	0	0	2.000	Table B.13 (p. 123) Fig. B.13 (p. 124)
AS6724	172,155	39	1.36	0.36	2.004	Table B.14 (p. 125) Fig. B.14 (p. 126)
AS49544	149,090	38*	4	1,425	2.020	Table B.15 (p. 127) Fig. B.15 (p. 128)
AS34233	148,421	0	0	0	1.997	Table B.16 (p. 129) Fig. B.16 (p. 130)
AS61387	130,468	0	0	0	1.989	Table B.17 (p. 131) Fig. B.17 (p. 132)
AS8315	113,760	0	0	0	2.002	Table B.18 (p. 133) Fig. B.18 (p. 134)
AS50673	112,783	39	3.05	514.92	2.016	Table B.19 (p. 135) Fig. B.19 (p. 136)
AS25525	99,106	39	1	15	2.001	Table B.20 (p. 137) Fig. B.20 (p. 138)

*In these situations the removed AS coincides with an investigated resolver location. Therefore actually one more resolver location is affected but from this no conclusions can be drawn.

Table 4.8: The top 20 autonomous systems in terms of the number of paths on the baseline on which they provide transit traffic and the results of a simulated failure of those ASs. A resolver location is only counted as affected by the simulated failure if the number of unreachable ASs increases by more than the one which was removed in the simulation. The average amounts of unreachable ASs and unreachable domains are only calculated over the affected resolver locations, whereas the mean length of shortest paths is calculated over all resolver locations.

Simulated failure (ASN)	**# Paths**	**# Resolver locations affected**	**Mean unreachable ASs**	**Mean unreachable domains**	**Mean shortest path length**	**Extensive results shown in**
AS174	22,165	39	11.08	94.72	2.040	Table B.21 (p. 139) Fig. B.21 (p. 140)
AS2914	17,849	39	3.80	2.67	2.037	Table B.22 (p. 141) Fig. B.22 (p. 142)
AS1299	13,140	39	6.64	113.49	2.023	Table B.23 (p. 143) Fig. B.23 (p. 144)
AS3356	12,932	15*	7.73	21.80	2.010	Table B.24 (p. 145) Fig. B.24 (p. 146)
AS6453	6,693	0	0	0	2.001	Table B.25 (p. 147) Fig. B.25 (p. 148)
AS3320	6,345	38*	2	1	2.007	Table B.26 (p. 149) Fig. B.26 (p. 150)
AS20562	5,011	1	1	0	2.049	Table B.27 (p. 151) Fig. B.27 (p. 152)
AS6939	4,861	39	4.59	9.69	2.011	Table B.28 (p. 153) Fig. B.28 (p. 154)
AS43531	4,374	39	1.15	0.03	2.007	Table B.29 (p. 155) Fig. B.29 (p. 156)
AS10310	4,026	39	102.41	137,559.59	1.932	Table B.30 (p. 157) Fig. B.30 (p. 158)
AS4436	3,990	2	1	0	2.000	Table B.31 (p. 159) Fig. B.31 (p. 160)
AS49685	3,946	39	100.46	178,109.51	1.923	Table B.32 (p. 161) Fig. B.32 (p. 162)
AS8455	491,008	39	5.41	11,474.67	2.046	Table B.5 (p. 107) Fig. B.5 (p. 108)
AS5511	3,649	14	109.57	83,677.43	2.107	Table B.33 (p. 163) Fig. B.33 (p. 164)
AS16509	3,592	14	107.71	83,359.43	2.059	Table B.34 (p. 165) Fig. B.34 (p. 166)
AS9002	3,473	39	1.39	0	2.001	Table B.35 (p. 167) Fig. B.35 (p. 168)
AS8220	3,192	39	10	29	2.001	Table B.36 (p. 169) Fig. B.36 (p. 170)
AS1136	2,997	39	40.56	51,100.72	2.106	Table B.37 (p. 171) Fig. B.37 (p. 172)
AS3257	2,982	7	1	0	2.002	Table B.38 (p. 173) Fig. B.38 (p. 174)
AS701	2,856	39	6	11	2.000	Table B.39 (p. 175) Fig. B.39 (p. 176)

*In these situations the removed AS coincides with an investigated resolver location. Therefore actually one more resolver location is affected but from this no conclusions can be drawn.

However, the impact of this on the reachability of domain names is rather small because these ASs mostly host just a small number of domain names which are not hosted in another AS which is still reachable. In some simulations, however, removal of one AS leads to a lot of ASs becoming unreachable. This is for example the case when removing AS8455 from the AS topology. In this case, also a higher amount of domain names becomes unreachable.

When removing ASs responsible for providing transit traffic, the situation changes. In this case most of the time some additional autonomous systems become unavailable from the viewpoint of either all or some of the investigated resolver locations. As a general remark it holds, that the reachability of domain names is highly dependent on the location of the resolver trying to access a certain domain name. In some simulations it is the case that there is almost no impact on the availability for some resolvers, whereas there is a huge impact on the reachability for other resolvers. This means that the impact on most resolver locations is rather low as just one or two domain names become unreachable, but the impact on a special AS is quite high as no domain names outside of that AS can be reached anymore. This is for example the case when removing AS10310 from the AS topology. In this scenario, resolvers in AS36647 are not able to reach any of the domain names not hosted in that AS. The same holds for AS35470 when removing AS49685 from the network topology. Also the cases when AS5511, AS16509 and AS1136 are removed are interesting as in these cases the resolvers located in AS3215, AS14618 and AS8737, respectively, loose their connection to roughly half of the ASs where name servers are located, resulting in roughly 1.1 million unreachable domain names.

Failure of connections

Removal of a single connection in general does not break anything as in most of the cases no autonomous systems and thus no domains other than those which are already unreachable in the baseline computations become unreachable. The only difference here is the mean length of the shortest path from the AS a resolver is located in to the AS the name servers are located in, which slightly increases. However, most of the times only one resolver is affected by this.

In some situations however, the connection is very important for the reachability of certain autonomous systems. Some autonomous systems as for example AS26647 and AS35470 do only have a single connection to the rest of the network. Removing these connections thus results in those autonomous systems being completely isolated from the rest of the Internet.

One simulated event, removal of the connection AS174↔AS2914 does not have any impact on the reachability of the DNS data as for all paths alternatives of the same length exist. This would also be the desired situation for all other events, as this suggests that the rest of the network is stable enough to catch this scenario.

Table 4.9: The top 20 connections in terms of the number of paths on the baseline on which they occur and the results of a simulated failure of those ASs. A resolver location is only counted as affected by the simulated failure if the number of unreachable ASs increases by more than the one which was removed in the simulation. The average amounts of unreachable ASs and unreachable domains are only calculated over the affected resolver locations, whereas the mean length of shortest paths is calculated over all resolver locations.

Simulated failure (connection)	**# Paths**	**# Resolver locations affected**	**Mean unreachable ASs**	**Mean unreachable domains**	**Mean shortest path length**	**Extensive results shown in**
AS35470↔AS49685	3,842	39	98.51	173,998.69	1.923	Table B.41 (p. 179) Fig. B.41 (p. 180)
AS10310↔AS36647	3,842	39	98.51	137,558.62	1.932	Table B.42 (p. 181) Fig. B.42 (p. 182)
AS1299↔AS23033	3,594	0	0	0	2.015	Table B.43 (p. 183) Fig. B.43 (p. 184)
AS14618↔AS16509	3,594	14	107.71	71,123.64	2.048	Table B.44 (p. 185) Fig. B.44 (p. 186)
AS3215↔AS5511	3,479	0	0	0	2.021	Table B.45 (p. 187) Fig. B.45 (p. 188)
AS17204↔AS2914	3,128	0	0	0	2.016	Table B.46 (p. 189) Fig. B.46 (p. 190)
AS2914↔AS5432	3,041	0	0	0	2.012	Table B.47 (p. 191) Fig. B.47 (p. 192)
AS1136↔AS8737	2,826	0	0	0	2.028	Table B.48 (p. 193) Fig. B.48 (p. 194)
AS31615↔AS3320	2,681	0	0	0	2.002	Table B.49 (p. 195) Fig. B.49 (p. 196)
AS24793↔AS41887	2,657	14	107.43	59,304.86	2.004	Table B.50 (p. 197) Fig. B.50 (p. 198)
AS1299↔AS2637	2,634	0	0	0	2.003	Table B.51 (p. 199) Fig. B.51 (p. 200)
AS55967↔AS6453	2,396	0	0	0	2.001	Table B.52 (p. 201) Fig. B.52 (p. 202)
AS30781↔AS34173	2,250	0	0	0	2.002	Table B.53 (p. 203) Fig. B.53 (p. 204)
AS43531↔AS9143	2,225	1	3	1	2.004	Table B.54 (p. 205) Fig. B.54 (p. 206)
AS2914↔AS393406	2,087	1	1	0	2.003	Table B.55 (p. 207) Fig. B.55 (p. 208)
AS174↔AS2914	2,084	0	0	0	2.000	Table B.56 (p. 209) Fig. B.56 (p. 210)
AS197902↔AS8455	2,047	0	0	0	2.003	Table B.57 (p. 211) Fig. B.57 (p. 212)
AS3320↔AS8972	1,892	0	0	0	2.001	Table B.58 (p. 213) Fig. B.58 (p. 214)
AS1136↔AS286	1,869	0	0	0	2.004	Table B.59 (p. 215) Fig. B.59 (p. 216)
AS1299↔AS13127	1,800	0	0	0	2.005	Table B.60 (p. 217) Fig. B.60 (p. 218)

Failure of the AMS-IX

Surprisingly, removing all peerings between participants of the AMS-IX does not lead to many changes in the amounts of unreachable ASs or domains. In fact, only resolvers located in AS3320 and AS3356 can not reach 116 and 57 ASs anymore, respectively. Other resolvers are not affected at all by this scenario. Furthermore, this results in only 6,096 and 1,406 domain names becoming unreachable respectively for the previously mentioned resolver locations. The amount of domains which can not be reached by removing these connections is quite small for those resolvers (at most 6,006) compared to the great amount of domain names which is still reachable. This was not expected as a lot of connections were removed and the data set only contains relationships between ASs and thus lacks possible backup peerings at other IXPs, which most probably exist in practice. However, the mean length of the shortest path between the AS of the resolver and the AS of the name servers does increase, which follows the expectations. The impact on the availability of data in case of failure of the AMS-IX is thus quite small and only the mean path length increases significantly. This can be explained as in these cases the tgsus132raditional customer-provider links have to be utilized. In practice however, there still might be availability issues in case of failure of the AMS-IX as the available bandwidth of the remaining customer-provider links is not taken into account and these might become overloaded. Extensive results regarding this simulation are depicted in Table B.40 and Figure B.40.

Identified network bottlenecks

This section is meant to show the identified bottlenecks in the network infrastructure which might be overcome by the operators of the affected autonomous systems by adding additional connections to other ASs. Due to the incompleteness of the used data set it might however be the case that these connections already exist in practice. A bottleneck is here defined as an AS or connection resulting in at least one AS becoming unreachable from at least one of the resolver locations. In total this is the case for 29 of the 60 simulated cases. However, a great portion of these bottlenecks only affect a small number of ASs and domain names. All bottlenecks are depicted in Table 4.10.

Table 4.10: The bottlenecks in the network infrastructure discovered in the results of the analysis. Every AS or connection whose removal leads to at least one AS becoming unreachable from at least one of the investigated resolver locations is considered as bottleneck. However, some bottlenecks are worse than others as the number of affected autonomous systems varies. Also the impact of failure of one of these bottlenecks on the availability of DNS data for domain names within the .nl-zone varies a lot. When no comment is given, the AS and domains becoming unreachable are the same for all investigated resolver locations.

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS21155	AS59980	1	-
AS8455	AS8608, AS9989, AS17819, AS21073, AS34215, AS44187, AS49820, AS50554, AS200831, AS202016	11,076 or 12,272	Not all ASs become unreachable from all resolver locations. Although the amount of unreachable ASs varies between four and ten for the different resolver locations, the amount of unreachable domain names is either 11,076 or 12,272. This indicates that storing DNS data on name servers located in multiple ASs actually solves a lot of the issues as the domain names stored on name servers located in some of the unreachable ASs have to be stored additionally on a name server located in another AS to reach this situation. The ASs becoming unreachable from all resolver locations are AS44187, AS202016, AS200831, and AS50554.
AS48635	AS199835	34,770	-

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS6724	AS6786, AS50575	0 or 1	AS6786 becomes unreachable from all resolver locations whereas AS50575 just becomes unreachable for only those resolvers located in AS1103, AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS15169, AS16276, AS24940, AS60781, and AS393406. For these resolvers one domain name becomes unreachable.
AS49544	AS5419, AS59743, AS62370, AS198485	1,425	-
AS50673	AS21100, AS57043, AS61044, AS200429, AS203318	418 or 688	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between two and five for the different resolver locations and the amount of unreachable domain names is either 418 or 688. The ASs becoming unreachable from all resolver locations are AS50673 and AS200429.
AS25525	AS47344	1	-

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS174	AS12212, AS12409, AS18986, AS27172, AS32650, AS41420, AS44204, AS44312, AS47223, AS57977, AS63363, AS196878, AS197415, AS198156	10, 12 or 14	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between ten and 14 for the different resolver locations and the amount of unreachable domain names is either 94 or 96. The ASs becoming unreachable from all resolver locations are AS41420, AS63363, AS44204, AS27172, AS12409, AS44312, AS32650, AS197415, AS57977, and AS196878.
AS2914	AS3949, AS4713, AS16266, AS17819, AS21371, AS47886, AS203849	1, 5 or 6	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between three and seven. The amount of unreachable domain names becoming unavailable is either 1, 5 or 6. The ASs becoming unreachable from all resolver locations are AS16266, AS3949, and AS203849.
AS1299	AS224, AS1653, AS1759, AS3249, AS3301, AS3308, AS13065, AS29217, AS34225, AS41483, AS42708, AS50066	4, 7, 41 or 455	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between three and twelve. The amount of unreachable domain names becoming unavailable is either 4, 7, 41 or 455. The ASs becoming unreachable from all resolver locations are AS3249, AS1759, and AS41483.

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS3356	AS11790, AS12401, AS16294, AS19891, AS21473, AS32421, AS61130, AS200081, AS201117	21 to 23	This only affects resolvers located in AS1103, AS2637, AS3215, AS3320, AS4134, AS6830, AS13127, AS13238, AS13414, AS15169, AS16276, AS24940, AS55967, AS60781, and AS393406. Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between six and nine. The amount of unreachable domain names becoming unavailable varies between 21 and 23. The ASs becoming unreachable from all affected resolver locations are AS32421, AS201117, AS61130, AS19891, AS16294, and AS11790.
AS3320	AS8292, AS21207	1	-
AS20562	AS9989	0	This only affects the resolvers located in AS20857. This bottleneck does however not have any impact on the availability of DNS data as all domains hosted on name servers located in AS9989 are additionally stored on name servers located in other ASs that are still reachable.

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS6939	AS4224, AS31064, AS32097, AS33387, AS46816, AS47066, AS53841	0 or 42	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between three and seven. The amount of unreachable domain names becoming unavailable is either 0 or 42. The ASs becoming unreachable from all resolver locations are AS46816, AS31064, and AS47066.
AS42531	AS17819, AS21371, AS43531, AS47886	0 or 1	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs varies between one and four. The only resolver location for which a domain name becomes unavailable is AS9143. The only AS becoming unreachable from all resolver locations is AS43531.
AS10310	AS7280, AS10880, AS26101, AS36646, AS36647	1	The situation is worse for resolvers located in AS36647 as these loose their connection to the rest of the network (3,804 ASs) and thereby to 5,364,786 domain names.
AS4436	AS9989	0	This only affects resolvers located in AS4134 and AS6830. The availability of DNS data is not influenced by this scenario.
AS49685	AS28878, AS35470, AS61331	47,051	The situation is worse for resolvers located in AS35470 as these loose their connection to the rest of the network (3,804 ASs) and thereby to 5,158,333 domain names.

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS5511	AS3215, AS16017, AS25186	29	This only affects resolvers located in AS1103, AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS15169, AS16276, AS24940, AS60781, and AS393406. The situation is worse for resolvers located in AS3215 as these loose their connection to 1,495 other ASs and thereby to 1,171,107 domain names.
AS16509	AS14618	13,237	This only affects resolvers located in AS1103, AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS13414, AS14618, AS15169, AS55967, and AS393406. The situation is worse for resolvers located in AS14618 as these loose their connection to 1,495 other ASs and thereby to 994,927 domain names.
AS9002	AS12594	0	This only affects resolvers located in AS1103, AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS15169, AS16276, AS24940, AS60781, and AS393406. For resolvers located in AS13238 additionally AS17819 becomes unreachable. The availability of DNS data is not influenced by this.

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS8220	AS15404, AS21234, AS42479, AS24771, AS31743, AS34601, AS41167, AS47393, AS49322, AS57533	29	-
AS1136	AS8737, AS12871, AS21286, AS59599	1,311 or 61,754	Not all ASs become unreachable from all resolver locations. The amount of unreachable ASs is either two or four. The amount of unreachable domain names becoming unavailable is either 1,311 or 61,754, dependent on the ASs becoming unreachable. The ASs becoming unreachable from all resolver locations are AS59599 and AS21286. The situation is worse for resolvers located in AS8737 as these loose their connection to 1,481 other ASs and thereby to 1,157,386 domain names.
AS3257	AS9989	0	This only affects resolvers located in AS1103, AS4134, AS6830, AS13414, AS17204, AS24940, and AS55967. The availability of DNS data is not influenced by this.
AS701	AS3378, AS3707, AS11486, AS33214, AS33478, AS36157	11	-

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS35470↔AS49685	AS35470	42,832	This is the only connection of AS35470 to the rest of the network. Therefore removing this connection results in AS35470 and 42,832 domains becoming unreachable from all other ASs. Additionally, AS35470 can not reach any other AS anymore resulting in 5,158,333 unreachable domain names.
AS10310↔AS36647	AS36647	0	This is the only connection of AS36647 to the rest of the network. Therefore removing this connection results in AS36647 becoming unreachable from all other ASs. This does however not have any impact on the DNS as there are no domain names hosted only on name servers located in AS36647. Additionally, AS36647 can not reach any other AS anymore resulting in 5,364,786 unreachable domain names.
AS14618↔AS16509	AS14618	301	This only affects the resolvers located in AS1103, AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS13414, AS14618, AS15169, AS55967, and AS393406. The situation is worse for resolvers located in AS14618 as these loose their connection to 1,495 other ASs and 991,818 domain names.

Continued on next page

Table 4.10 – *Continued from previous page*

Failing ASN/connection	**ASs becoming unreachable**	**# Domains becoming unreachable**	**Comments**
AS24793↔AS41887	AS24793	5,432	This only affects the resolvers located in AS2637, AS3215, AS3320, AS3356, AS4134, AS6830, AS13127, AS13238, AS15169, AS16276, AS24793, AS24940, AS60781, and AS393406. The situation is worse for resolvers located in AS24793 as these loose their connection to 1,491 other ASs and 759,652 domain names.
AS43531↔AS9143	AS17819, AS21371, AS47886	1	Only resolvers located in AS9143 are affected.
AS2914↔AS393406	AS17819	0	Only resolvers located in AS393406 are affected. As there are no domain names hosted only on name servers located in AS17819, this does not influence the availability of DNS data.

Chapter 5

Conclusion

This chapter is meant to draw conclusions from the results and to provide some recommendations to SIDN, associated registrars of .nl and the owners of second-level domain names within the .nl-zone.

5.1 Methodology

In this thesis a new method for investigation of the resilience of the DNS infrastructure of a TLD is shown. Although there are still some issues to be solved in order to improve the quality of the results, especially regarding the availability of accurate information about the Internet's infrastructure, the approach could already be applied for research on other TLDs than .nl. An interesting follow up research might be to compare the results obtained during this research with those obtained for other TLDs using the same method and AS topology. This would enable a comparison between the setups of multiple TLDs and would also indicate whether incidents with negligible impacts for the availability of the .nl-zone have greater impacts on the availability of other zones.

5.2 Connectivity of ASs

Most of the autonomous systems investigated have multiple connections to the rest of the Internet. Therefore failure of single connections is not that a great issue for most of the ASs. However, this analysis also reveals some autonomous systems which are poorly connected to the rest of the network, i.e. which have just a single upstream provider and no peers at all. Thereby all domain names which are hosted solely on name servers within that AS are at risk of becoming unavailable to the rest of the network in case of failure of a single connection. The operators of these autonomous systems should therefore consider to add additional connections to lower the impact of failure of parts of the Internet's infrastructure. These observations might however also be due to the incompleteness of the underlying data set regarding the AS topology. Furthermore, only the connectivity in terms of BGP sessions is investigated. The physical layer is not considered here which means that two autonomous systems might be connected via multiple physical links which show up as a single connection in the AS topology graph. In that case failure of one of those links would only lower the bandwidth of the connection, but does not influence the AS topology. As the bandwidth of connections was not considered in this research the reachability of ASs would thus not change. Nevertheless, the available bandwidth does play an important role in the traffic flow of the Internet and future research might take this aspect into account, assuming that this kind of data becomes available one day.

5.3 Resilience of the infrastructure serving the .nl-zone

In most of the analyzed scenarios the impact of failure of certain parts of the Internet's infrastructure is negligible. However, operators of the analyzed autonomous systems can use the provided knowledge about the structure of the network to further strengthen the connectivity of their network to other autonomous systems. Especially the in Section 4.4.2 identified bottlenecks in the connectivity of certain autonomous systems should be overcome by adding additional connections to the AS topology if they do not already exist and just do not show up in the inferred topology provided by CAIDA.

This research also showed, that there are a lot of domain names with corresponding name servers located in just a single autonomous system. This is an unwanted situation as failure or unreachability of that AS due to failures in other parts of the system can not be handled. Most of the issues regarding unreachable domain names can be overcome simply by using at least one additional name server located in another autonomous system. Nevertheless, in some situations this is not a big issue in terms of the DNS as there are a lot of applications hosted in a single autonomous system and even though the DNS data would be available when hosted in multiple ASs, the application itself would be unreachable.

Although the DNS is a distributed systems which generally increases its resilience in case of network failure scenarios, most of the data is hosted on a very small set of authoritative name servers. Although it is not known how much load-balancing is applied and how many physical servers are reachable via the same IP address, it should be considered to limit the amount of domain names hosted on a single IP address. A DoS attack might also be carried out on the routing level by fraudulently announcing the IP address of a huge name server and thereby redirecting a lot of queries. A similar incident happened in 2008 when a Pakistani ISP hijacked the IP address range of Youtube [67].

5.4 Optimizing the zonefile

As shown in Section 4.1.3 there are a lot of inconsistencies between the name servers known to SIDN and the name servers actually serving a domain. These inconsistencies should be resolved in order to improve the performance of the DNS look-ups. SIDN could use the gathered data to inform the registrars over inconsistencies and ask them to update their data.

This could also be used as another indicator for SIDN's Registrar ScoreCard, a program for registrars which is meant to reward those registrars who stimulate an active and secure usage of the .nl-zone [86, 87].

5.5 Lessons learned

The purpose of this research was to investigate the impact of failure of parts of the Internet's infrastructure on the Domain Name System and especially on the availability of domain names in the .nl-zone. From the results obtained for the simulated events it can be concluded that the availability of most of the domain names in the .nl-zone is not at risk even though some ASs or interconnections between those drop out.

Most of the problems regarding the availability of DNS data in extraordinary circumstances can be circumvented by hosting this data on name servers located in different autonomous systems as there are just very few situations in which a large number of autonomous systems becomes unreachable. Thus, by hosting the data on name servers located in different autonomous systems the chances are high that one of the two ASs stays reachable during an incident.

However, only 60 different failure scenarios were considered in order to draw this conclusion. Although these scenarios were selected such that their impact is expected to be the worst, it is not entirely clear whether these scenarios are representative for all possible events.

5.6 Final remark

This thesis gave necessary background knowledge and showed a framework for analyzing the impact of failures of parts of the Internet's infrastructure on the DNS. This framework was applied on the domain names contained in the .nl-zone, 60 failure scenarios were simulated and the results were used to identify bottlenecks in the connectivity of autonomous systems. However, there are still some issues which have to be solved for a more accurate analysis. This mainly affects the availability of more accurate data. Chapter 6 will therefore summarize some possible enhancements on the shown methodology and also summarize some other interesting follow up research questions that came up during this project.

Chapter 6

Future work

Future research might extend and improve the shown methodology by using a more accurate network topology or by analyzing more resolver locations and failure scenarios.

The public nature of the used data assures that this research might be repeated to perform risk assessments for other top level domain names. This may be done by the registries of those TLDs, but also by the rest of the research community. The researchers do however need access to the list of domain names they want to include in their analysis. Some registries do have programs to grant access to their zonefiles for academic purposes, such as Verisign, the registry for .com, .net, .tv, .cc and .name [97]. However, not all zonefiles might be available even when maintained by the same company. In the case of Verisign, only the zonefiles of .com, .net and .name might be accessed. If access of the zonefile is not granted, to some extent a brute-force approach to generate the list of registered domain names may be applied. However, there is a big chance, that it is infeasible to generate a full list using only brute-force as the domain name space is simply too big and therefore another source of uncertainty would be introduced as you can never be sure to have found all domain names.

As another solution to this problem a security vulnerability of DNSSEC might be exploited once DNSSEC is fully adopted to obtain a list of all existing domain names by sequentially scanning the `NSEC` records provided by DNSSEC [41]. However, since the first publication of DNSSEC yet another RFC (RFC 5155) was published proposing the `NSEC3` resource record to circumvent this problem [49]. However, for zones not implementing this RFC, enumeration of all domains contained in the zonefile is possible.

To get a more accurate view on the relationships between autonomous systems, more insider information might be gathered from the owner of those systems. However, as mentioned earlier not every ISP wants to share this information as it might reveal security vulnerabilities. A more accurate sight on the AS relationship graph would also be very interesting for other types of research regarding the infrastructure of the Internet as a whole.

One could of course also analyze the impact of failure of other ASs or other connections between those. It would also be possible to analyze the impact of simultaneous failure of multiple ASs, although this incident is less probable to happen in practice, unless those multiple ASs fall within the same administrative boundary. Last but not least one could simulate simultaneous failure of multiple connections. For more realistic predictions one could take co-location into account when analyzing simultaneous failure of different parts of the network infrastructure.

While performing the mapping of name servers onto the belonging IP addresses it was discovered that there are some mismatches between the DNS records in the .nl zonefile maintained by SIDN and the zonefiles of second-level domains. Future research might investigate these mismatches to quantify the impact of this on the performance of the DNS in general. These might be lame delegations which can result in domain names being only sometimes resolvable, namely in those cases the correct name server is queried. Although this is not a big issue for the whole zone in practice, as only some domain names are affected, the zonefile of .nl might be optimized by removing the additional name servers.

As a last idea, one could also investigate the resilience of the AS topology in general, i.e. apply the proposed path finding technique using all ASs as source and goal rather than using only those responsible for resolving and storing DNS data.

Acknowledgments

By this I want to thank SIDN and all its employees for giving me the opportunity to perform my master's thesis in such a wonderful working environment. Especially I would like to thank my supervisor Jelte Jansen, for standing open for any arising question during the project and keeping me motivated in periods of stagnation. Furthermore, I want to thank the rest of the SIDN Labs department, Cristian Hesselman, Marco Davids, Maarten Wullink, Giovane Moura, and Moritz Müller for helping out whenever necessary and reviewing the final thesis. I would also like to thank Arjen Zonneveld and Erwin Heringa for their explanations of the Border Gateway Protocol and derivation of the AS topology.

From the Radboud University Nijmegen I would like to thank Dr. ir. Harald Vranken for supervising the project from the academical perspective and for keeping to ask critical questions which steer me into the right direction. Furthermore I want to thank Dr. ir. Erik Poll for his support in finding a suitable research topic and bringing me in contact with the people from SIDN and thereby enabling the collaboration on this thesis.

Last but not least, I want to thank Marijke Kröhnke for emotionally supporting me during the formation phase of this thesis and during the rest of my preceding study period.

Thank you all!
Lars Bade

Bibliography

Books

[1] Iljitsch van Beijnum. *BGP - Building Reliable Networks with the Border Gateway Protocol*. 1st Edition. O'Reilly Media, Inc., Sept. 2002.

[2] Steven S Skiena. *The algorithm design manual: Text*. Vol. 1. Springer Science & Business Media, 1998.

Scientific publications

[3] Sharad Agarwal et al. "The impact of BGP dynamics on intra-domain traffic". In: *ACM SIGMETRICS Performance Evaluation Review*. Vol. 32. 1. ACM. 2004, pp. 319–330.

[4] Bernhard Ager et al. "Anatomy of a large European IXP". In: *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*. ACM. 2012, pp. 163–174.

[5] M. Tariq Banday. "Recent Developments in the Domain Name System". In: *International Journal of Computer Applications* 31.2 (Oct. 2011), pp. 18–25.

[6] Erol Basturk et al. "Using Network Layer Anycast for Load Distribution in the Internet". In: *TECH. REP., IBM T.J. WATSON RESEARCH CENTER*. 1997.

[7] Nevil Brownlee, KC Claffy, and Evi Nemeth. "DNS measurements at a root server". In: *Global Telecommunications Conference, 2001. GLOBECOM'01. IEEE*. Vol. 3. IEEE. 2001, pp. 1672–1676.

[8] Matthew Caesar and Jennifer Rexford. "BGP routing policies in ISP networks". In: *Network, IEEE* 19.6 (2005), pp. 5–11.

[9] Nikolaos Chatzis et al. "There is more to IXPs than meets the eye". In: *ACM SIGCOMM Computer Communication Review* 43.5 (2013), pp. 19–28.

[10] Lorenzo Colitti et al. "Evaluating the effects of anycast on DNS root name servers". In: *RIPE document RIPE-393* 6 (2006).

[11] Amogh Dhamdhere, Constantine Dovrolis, and Pierre Francois. "A value-based framework for internet peering agreements". In: *Teletraffic Congress (ITC), 2010 22nd International*. IEEE. 2010, pp. 1–8.

[12] Xenofontas Dimitropoulos et al. "AS relationships: Inference and validation". In: *ACM SIGCOMM Computer Communication Review* 37.1 (2007), pp. 29–40.

[13] Xenofontas Dimitropoulos et al. "Inferring as relationships: Dead end or lively beginning?" In: *Experimental and Efficient Algorithms*. Springer, 2005, pp. 113–125.

[14] Peyman Faratin et al. "The growing complexity of Internet interconnection". In: *Communications & strategies* 72 (2008), p. 51.

[15] Lixin Gao. "On inferring autonomous system relationships in the Internet". In: *IEEE/ACM Transactions on Networking (ToN)* 9.6 (2001), pp. 733–745.

[16] Vasileios Giotsas, Shi Zhou, Matthew Luckie, et al. "Inferring multilateral peering". In: *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*. ACM. 2013, pp. 247–258.

[17] Hamed Haddadi et al. "Network topologies: inference, modeling, and generation". In: *Communications Surveys & Tutorials, IEEE* 10.2 (2008), pp. 48–69.

[18] Peter E Hart, Nils J Nilsson, and Bertram Raphael. "A formal basis for the heuristic determination of minimum cost paths". In: *IEEE transactions on Systems Science and Cybernetics* 4.2 (1968), pp. 100–107.

[19] Christopher Kruegel et al. "Topology-based detection of anomalous BGP messages". In: *Recent Advances in Intrusion Detection*. Springer. 2003, pp. 17–35.

[20] Wei-min Li et al. "Improving DNS cache to alleviate the impact of DNS DDoS attack". In: *Journal of Networks* 6.2 (2011), pp. 279–286.

[21] Steffen Lippert and Giancarlo Spagnolo. "Internet peering as a network of relations". In: *Telecommunications policy* 32.1 (2008), pp. 33–49.

[22] Ziqian Liu et al. "Two days in the life of the DNS anycast root servers". In: *Passive and Active Network Measurement*. Springer, 2007, pp. 125–134.

[23] Damien Magoni and Jean Jacques Pansiot. "Analysis of the autonomous system network topology". In: *ACM SIGCOMM Computer Communication Review* 31.3 (2001), pp. 26–37.

[24] Priya Mahadevan et al. "The Internet AS-level topology: three data sources and one definitive metric". In: *ACM SIGCOMM Computer Communication Review* 36.1 (2006), pp. 17–26.

[25] Hirokazu Miura et al. "Server load balancing with network support: Active anycast". In: *Active Networks*. Springer, 2000, pp. 371–384.

[26] P. Mockapetris and K. J. Dunlap. "Development of the Domain Name System". In: *Symposium Proceedings on Communications Architectures and Protocols*. SIGCOMM '88. Stanford, California, USA: ACM, 1988, pp. 123–133. ISBN: 0-89791-279-9. DOI: 10.1145/52324.52338. URL: http://doi.acm.org/10.1145/52324.52338.

[27] Reza Motamedi, Reza Rejaie, and Walter Willinger. "A Survey of Techniques for Internet Topology Discovery". In: *Communications Surveys & Tutorials, IEEE* 17.2 (2015), pp. 1044–1065.

[28] Giovane CM Moura et al. "Anycast vs. DDoS: Evaluating the November 2015 Root DNS Event (extended)". not yet published. May 2016.

[29] Andy Ozment and Stuart E Schechter. "Bootstrapping the Adoption of Internet Security Protocols." In: *WEIS*. 2006.

[30] Vasileios Pappas et al. "Distributed DNS troubleshooting". In: *Proceedings of the ACM SIGCOMM workshop on Network troubleshooting: research, theory and operations practice meet malfunctioning reality*. ACM. 2004, pp. 265–270.

[31] Jennifer Rexford et al. "BGP routing stability of popular destinations". In: *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*. ACM. 2002, pp. 197–202.

[32] Roland van Rijswijk-Deij et al. "A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements". In: (2016).

[33] Matthew Roughan et al. "10 lessons from 10 years of measuring and modeling the internet's autonomous systems". In: *Selected Areas in Communications, IEEE Journal on* 29.9 (2011), pp. 1810–1821.

[34] Sandeep Sarat, Vasileios Pappas, and Andreas Terzis. "On the use of anycast in DNS". In: *Computer Communications and Networks, 2006. ICCN 2006. Proceedings. 15th International Conference on*. IEEE. 2006, pp. 71–78.

[35] Gireesh Shrimali and Sunil Kumar. "Paid peering among internet service providers". In: *Proceeding from the 2006 workshop on Game theory for communications and networks*. ACM. 2006, p. 11.

[36] Lakshminarayanan Subramanian et al. "Characterizing the Internet hierarchy from multiple vantage points". In: *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*. Vol. 2. IEEE. 2002, pp. 618–627.

[37] Li Wei-Min, Chen Lu-Ying, and Lei Zhen-Ming. "Alleviating the impact of DNS DDoS attacks". In: *Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010 Second International Conference on*. Vol. 1. IEEE. 2010, pp. 240–243.

[38] Beichuan Zhang et al. "Collecting the Internet AS-level topology". In: *ACM SIGCOMM Computer Communication Review* 35.1 (2005), pp. 53–61.

[39] Ke Zhang et al. "On detection of anomalous routing dynamics in BGP". In: *Networking 2004*. Springer. 2004, pp. 259–270.

[40] Changxi Zheng et al. "A light-weight distributed scheme for detecting IP prefix hijacks in real-time". In: *ACM SIGCOMM Computer Communication Review*. Vol. 37. 4. ACM. 2007, pp. 277–288.

Technical reports

[41] R. Arends et al. *DNS Security Introduction and Requirements*. RFC 4033. Internet Engineering Task Force, Mar. 2005.

[42] R. Arends et al. *Protocol Modifications for the DNS Security Extensions*. RFC 4035. Internet Engineering Task Force, Mar. 2005.

[43] R. Arends et al. *Resource Records for the DNS Security Extensions*. RFC 4034. Internet Engineering Task Force, Mar. 2005.

[44] J. Arkko et al. *IPv4 Address Blocks Reserved for Documentation*. RFC 5737. Internet Engineering Task Force, Jan. 2010.

[45] D. Atkins. *Threat Analysis of the Domain Name System (DNS)*. RFC 3833. Internet Engineering Task Force, Aug. 2004.

[46] S. Bortzmeyer. *DNS Query Name Minimisation to Improve Privacy*. RFC 7816. Internet Engineering Task Force, Mar. 2016.

[47] V. Fuller and T. Li. *Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan*. RFC 4632. Internet Engineering Task Force, Aug. 2006.

[48] J. Hawkinson and T. Bates. *Guidelines for creation, selection, and registration of an Autonomous System (AS)*. RFC 1930. Internet Engineering Task Force, Mar. 1996.

[49] B. Laurie et al. *DNS Security (DNSSEC) Hashed Authenticated Denial of Existence*. RFC 5155. Internet Engineering Task Force, Mar. 2008.

[50] P. Mockapetris. *Domain Names - Concepts and Facilities*. RFC 1034. Internet Engineering Task Force, Nov. 1987.

[51] P. Mockapetris. *Domain Names - Implementation and Specification*. RFC 1035. Internet Engineering Task Force, Nov. 1987.

[52] Dave Piscitello. *Conficker Summary and Review*. Tech. rep. Internet Corporation for Assigned Names and Numbers, May 2010.

[53] J. Postel. *Domain Name System Structure and Delegation*. RFC 1591. Internet Engineering Task Force, Mar. 1994.

[54] Y. Rekhter, T. Li, and S. Hares. *A Border gateway Protocol 4 (BGP-4)*. RFC 4271. Internet Engineering Task Force, Jan. 2006.

[55] *Technical requirements for the registration of .nl domain names.* https://www.sidn.nl/downloads/terms-and-conditions/Technical+requirements+for+the+registration+of+nl+domain+names.pdf. SIDN. May 2010.

Online resources

[56] Center for Applied Internet Data Analysis. *Archipelago (Ark) Measurement Infrastructure.* URL: http://www.caida.org/projects/ark/ (visited on 2016-05-19).

[57] Center for Applied Internet Data Analysis. *AS Relationships.* URL: http://www.caida.org/data/as-relationships/ (visited on 2016-03-10).

[58] Center for Applied Internet Data Analysis. *The CAIDA AS Organizations Dataset, <date range used>.* URL: http://www.caida.org/data/as-organizations (visited on 2016-05-03).

[59] European Internet Exchange Association. *What is an IXP?* URL: https://euro-ix.net/ixps/what-is-ixp/ (visited on 2016-03-09).

[60] Internet Assigned Number Authority. *About the Internet Assigned Number Authority.* URL: https://www.iana.org/about (visited on 2016-03-02).

[61] Internet Assigned Number Authority. *Number Resources.* URL: https://www.iana.org/numbers (visited on 2016-03-22).

[62] Iljitsch van Beijnum. *Transit vs peering: what makes sense when?* URL: http://packetpushers.net/transit-vs-peering-makes-sense/ (visited on 2016-03-10).

[63] RIPE Network Coordination Centre. *ASN Neighbours.* URL: https://stat.ripe.net/widget/asn-neighbours (visited on 2016-05-19).

[64] RIPE Network Coordination Centre. *Remove Multihoming Requirement for AS Number Assignments.* URL: https://www.ripe.net/participate/policies/proposals/2014-03 (visited on 2016-08-04).

[65] RIPE Network Coordination Centre. *RIPE Database.* URL: https://apps.db.ripe.net/search/query.html (visited on 2016-02-22).

[66] RIPE Network Coordination Centre. *Routing Information Service (RIS).* URL: https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris (visited on 2016-05-19).

[67] cnet. *How Pakistan knocked YouTube offline (and how to make sure it never happens again).* URL: http://www.cnet.com/news/how-pakistan-knocked-youtube-offline-and-how-to-make-sure-it-never-happens-again/ (visited on 2016-08-04).

[68] Dyn. *Pakistan hijacks YouTube.* URL: http://research.dyn.com/2008/02/pakistan-hijacks-youtube-1/ (visited on 2016-08-08).

[69] Hurricane Electric. *About Hurricane Electric*. URL: http://he.net/about_us.html (visited on 2016-03-30).

[70] Hurricane Electric. *Looking Glass*. URL: https://lg.he.net/ (visited on 2016-03-30).

[71] Amsterdam Internet Exchange. *Ams-IX Route Servers*. URL: https://ams-ix.net/technical/specifications-descriptions/ams-ix-route-servers (visited on 2016-03-09).

[72] Amsterdam Internet Exchange. *Connected Parties*. URL: https://ams-ix.net/connected_parties (visited on 2016-03-16).

[73] Google. *Google Public DNS - Performance Benefits*. URL: https://developers.google.com/speed/public-dns/docs/performance (visited on 2016-07-06).

[74] Google. *Verifying Googlebot*. URL: https://support.google.com/webmasters/answer/80553 (visited on 2016-07-06).

[75] Geoff Huston. *CIDR REPORT for 09 Aug 16*. URL: http://www.cidr-report.org/as2.0/ (visited on 2016-08-09).

[76] MaxMind Inc. *GeoLite Legacy Downloadable Databases*. URL: http://dev.maxmind.com/geoip/legacy/geolite/ (visited on 2016-05-03).

[77] Team Cymru Inc. *IP to ASN mapping*. URL: http://www.team-cymru.org/IP-ASN-mapping.html (visited on 2016-04-13).

[78] ISPam. *Nameserver storing TransIP legt tienduizenden domeinnamen korte tijd plat*. URL: http://www.ispam.nl/archives/61431/nameserver-storing-transip-legt-tienduizenden-domeinnamen-korte-tijd-plat/ (visited on 2016-06-21).

[79] ISPam. *TransIP getroffen door ernstige stroomstoring – update*. URL: http://www.ispam.nl/archives/48039/transip-getroffen-door-ernstige-stroomstoring/ (visited on 2016-06-21).

[80] SIDN Labs. *About SIDN Labs and ENTRADA*. URL: http://entrada.sidnlabs.nl/docs/introduction/about/ (visited on 2016-05-03).

[81] SIDN Labs. *ENTRADA - An open source platform for network data analytics*. URL: http://entrada.sidnlabs.nl (visited on 2016-03-15).

[82] SIDN Labs. *ENTRADA Data Model*. URL: http://entrada.sidnlabs.nl/docs/concepts/data_model/ (visited on 2016-05-19).

[83] SIDN Labs. *.nl stats and data - Insight into the use of .nl*. URL: http://stats.sidnlabs.nl/ (visited on 2016-04-07).

[84] SIDN Labs. *SIDN Labs, the research- and developmentteam of SIDN*. URL: https://www.sidnlabs.nl/over-sidnlabs?language_id=2 (visited on 2016-04-12).

[85] University of Oregon. *Route Views Project*. URL: http://www.routeviews.org/ (visited on 2016-05-19).

[86] Vereniging van Registrars. *Resultaten Registrar Score Card*. URL: https://www.verenigingvanregistrars.nl/nieuws/resultaten-registrar-score-card/ (visited on 2016-07-08).

[87] Vereniging van Registrars. *Uitbreiding Registrar Score Card*. URL: https://www.verenigingvanregistrars.nl/nieuws/uitbreiding-registrar-score-card/ (visited on 2016-07-08).

[88] SIDN. *DNSSEC FAQ*. URL: https://www.dnssec.nl/wat-is-dnssec/faq.html (visited on 2016-08-10).

[89] SIDN. *DNSSEC overzicht*. URL: https://www.dnssec.nl/wat-is-dnssec/overzicht.html (visited on 2016-07-20).

[90] SIDN. *SIDN Annual Report 2015*. URL: http://jaarverslag.sidn.nl/jaarverslag2015/en/sidn-labs (visited on 2016-07-20).

[91] SIDN. *spark: Scan the DNS efficiently*. URL: https://github.com/SIDN/spark (visited on 2016-05-30).

[92] SIDN. *What we do*. URL: https://www.sidn.nl/a/about-sidn/what-we-do (visited on 2016-04-12).

[93] Skynews. *Google Outage: Internet Traffic Plunges 40%*. URL: http://news.sky.com/story/1129847/google-outage-internet-traffic-plunges-40-percent (visited on 2016-06-21).

[94] International Organization for Standardization. *Country Codes - ISO 3166*. URL: http://www.iso.org/iso/country_codes (visited on 2016-04-12).

[95] Tweakers. *Internetknooppunt AMS-IX kampt met uitval - update 2*. URL: https://tweakers.net/nieuws/103067/internetknooppunt-ams-ix-kampt-met-uitval.html (visited on 2016-06-21).

[96] University of Twente, SURFnet, and SIDN. *OpenINTEL Open Access*. URL: http://openintel.nl/ (visited on 2016-06-16).

[97] Verisign. *Zone Files For Top-Level Domains (TLDs)*. URL: https://www.verisign.com/en_US/channel-resources/domain-registry-products/zone-file/index.xhtml (visited on 2016-03-31).

[98] Yu Zhang. *Internet AS-level Topology Archive*. URL: http://irl.cs.ucla.edu/topology/ (visited on 2016-03-31).

Other

[99] Jelte Jansen. personal communication. Aug. 5, 2016.

[100] Roland van Rijswijk-Deij. personal communication. July 12, 2016.

Appendix A

Additional information

A.1 Autonomous system operators

Table A.1 shows the operators of all autonomous systems mentioned in the results. This list is shown here for reference purposes.

Table A.1: Complete list of the operators of autonomous systems mentioned throughout the thesis.

ASN	Operator
AS174	Cogent Communications
AS224	UNINETT AS
AS701	Verizon Business/UUNet
AS1103	SURFnet, The Netherlands
AS1136	KPN B.V.
AS1140	Stichting Internet Domeinregistratie Nederland
AS1299	TeliaSonera AB
AS1653	SUNET Swedish University Network
AS1759	TeliaSonera Finland Oyj
AS2637	Georgia Institute of Technology
AS2914	NTT America, Inc.
AS3215	Orange S.A.
AS3249	AS Eesti Telekom
AS3257	Tinet Spa
AS3265	Xs4all Internet BV
AS3301	TeliaSonera AB
AS3308	TeliaSonera AB
AS3320	Deutsche Telekom AG
AS3356	Level 3 Communications, Inc.
AS3378	MCI

Continued on next page

Table A.1 – *Continued from previous page*

ASN	Operator
AS3707	MCI Communications Services, Inc. d/b/a Verizon Business
AS3949	NTT America, Inc.
AS4134	China Telecom Backbone
AS4224	The Calyx Institute
AS4436	nLayer Communications, Inc.
AS4713	NTT Communications Corporation
AS5419	Cubic Circle B.V.
AS5432	Proximus NV
AS5511	Orange S.A.
AS6453	TATA COMMUNICATIONS (AMERICA) INC
AS6724	Strato AG
AS6786	CRONON AG
AS6830	Liberty Global Operations B.V.
AS6939	Hurricane Electric, Inc.
AS7280	Yahoo! Inc.
AS8075	Microsoft Corporation
AS8220	COLT Technology Services Group Limited
AS8292	CamData GmbH
AS8315	Amsio B.V.
AS8455	Schuberg Philis B.V.
AS8608	EspritXB B.V.
AS8737	KPN B.V.
AS8972	PlusServer AG
AS9002	RETN Limited
AS9121	Turk Telekomunikasyon Anonim Sirketi
AS9143	Ziggo B.V.
AS9989	Equinix Singapore Pte Ltd
AS10310	Yahoo!
AS10880	Yahoo
AS11486	Verizon Online LLC
AS11790	Random House, Inc.
AS12212	Ravand Cybertech Inc.
AS12401	internic Datenkommunikations GmbH
AS12409	H.R.Net SARL
AS12594	Externet Nyrt
AS12859	BIT BV
AS12871	KPN B.V.
AS13065	Uniwersytet Zielonogorski
AS13127	Tele 2 Nederland B.V.
AS13238	YANDEX LLC
AS13335	CloudFlare, Inc.

Continued on next page

Table A.1 – *Continued from previous page*

ASN	Operator
AS13414	Twitter Inc.
AS14618	Amazon.com, Inc.
AS15169	Google Inc.
AS15404	COLT Technology Services Group Limited
AS15879	IS Group B.V.
AS16017	Hager Electro GmbH & Co. KG
AS16266	Canon Europa N.V.
AS16276	OVH SAS
AS16294	Descartes Systems (Belgium) NV
AS16509	Amazon.com, Inc.
AS17204	Nominum, Inc
AS17819	Equinix Asia Pacific
AS18986	Pacific Online Inc.
AS19891	F-Secure Corporation
AS20562	Open Peering B.V.
AS20857	Transip B.V.
AS20940	Akamai International B.V.
AS21073	Zoranet Internetdiensten
AS21100	ITL Company
AS21155	ProServe B.V.
AS21207	RWE IT GmbH
AS21234	GRENKE Service AG
AS21286	KPN Corporate Market B.V.
AS21371	Equinix Limited
AS21473	MAnet GmbH
AS23033	Wowrack.com
AS24771	Fiat Information Technology, Excellence and Methods S.p.A.
AS24793	NL Hosting Internet
AS24940	Hetzner Online GmbH
AS25151	Cyso Hosting B.V.
AS25186	Orange S.A.
AS25459	NedZone Internet BV
AS25525	Reasonnet IP Networks B.V.
AS26101	Yahoo!
AS27172	AIT Worldwide Logistics, Inc.
AS28878	Signet B.V.
AS29217	CGI Sverige AB
AS31064	SABRI-BERISHA
AS31615	T-mobile Netherlands bv.
AS31743	Interpublic GIS (UK) Ltd
AS32097	WholeSale Internet, Inc.

Continued on next page

Table A.1 – *Continued from previous page*

ASN	Operator
AS32421	Black Lotus Communications
AS32650	SANDHILLS PUBLISHING
AS32934	Facebook, Inc.
AS33214	Tenneco Automotive
AS33387	DataShack, LC
AS33478	Live Nation Entertainment, Inc.
AS34173	SafeBrands S.A.S.
AS34215	Inet Technology Consultancy B.V.
AS34225	SpeedPartner GmbH
AS34233	Superior B.V.
AS34601	Intercom Factory S.L.
AS35470	XL Internet Services B.V.
AS36157	Gateway, Inc
AS36351	SoftLayer Technologies Inc.
AS36646	Yahoo
AS36647	Yahoo
AS36692	OpenDNS, LLC
AS41167	REWE Touristik Gesellschaft mbH
AS41420	Xentech B.V.
AS41483	Tele2 Sverige AB
AS42479	Kroll Ltd
AS42708	Portlane AB
AS43531	IX Reach Ltd
AS44187	OneXS International B.V.
AS44204	FX MEDIA SRL
AS44312	Moving Art Studio ASBL
AS46816	DirectSpace Networks, LLC.
AS47066	prgmr.com, Inc.
AS47223	Lukman Multimedia Sp. z o.o
AS47344	TMG Online Media B.V.
AS47393	WPP Group Technology Services
AS47886	Equinix (Netherlands) B.V.
AS48635	PCextreme B.V.
AS49322	ViM Internetdienstleistungen GmbH
AS49544	i3d B.V.
AS49685	Signet B.V.
AS49820	Pictura Imaginis B.V.
AS50066	Serious Tubes Networks
AS50554	Networkconcepts BV
AS50575	Hatchery Group GmbH & Co.KG
AS50673	Serverius Holding B.V.

Continued on next page

Table A.1 – *Continued from previous page*

ASN	Operator
AS53841	RAM Host
AS55967	Beijing Baidu Netcom Science and Technology Co., Ltd.
AS57043	HOSTKEY B.V.
AS57533	DELTICOM AG
AS57977	ISVTEC SARL
AS59599	NCCW B.V.
AS59743	Angelo Kreikamp trading as Elitehosting
AS59980	mijndomein.nl BV
AS60781	LeaseWeb Netherlands B.V.
AS61044	SIT Internetdiensten B.V.
AS61130	Source Managed Services Limited
AS61331	Unigarant NV
AS61387	Denkers-ICT B.V.
AS62370	Snel.com B.V.
AS63363	Uniregistry Corp.
AS196878	Marcel Edler trading as Optimate-Server
AS197415	Financial Art S.A.
AS197902	Hostnet B.V.
AS198156	Dediserv Dedicated Servers Sp. z o.o.
AS198485	DirectVPS B.V.
AS199835	Vevida bv
AS200081	Netversor GmbH
AS200429	HostSlim BV
AS200831	Mihos
AS201117	PAN NET SRL
AS202016	Domino ICT B.V.
AS203849	Umicore NV
AS393406	Digital Ocean, Inc.

A.2 List of ASs peering at the AMS-IX

Listing A.1 shows all autonomous systems which were connected to the Amsterdam Internet Exchange on the 1st of June, 2016 [72].

Listing A.1: List of all autonomous systems peering at the AMS-IX.

```
AS10026, AS10310, AS109, AS1101, AS1103, AS11179, AS112, AS1126,
AS1136, AS1140, AS11816, AS12041, AS12189, AS12301, AS12306, AS12310,
AS12315, AS12322, AS12329, AS12350, AS12355, AS12389, AS1239, AS12399,
AS1241, AS12414, AS1248, AS12496, AS12552, AS1257, AS12578, AS12615,
AS12637, AS12654, AS1267, AS12695, AS12713, AS12714, AS12715, AS1273,
AS12731, AS12732, AS12759, AS12779, AS12849, AS12850, AS12859,
AS12871, AS12874, AS12876, AS12883, AS12902, AS12956, AS12989,
AS12993, AS13002, AS13004, AS13030, AS13101, AS13122, AS13127,
```

```
AS13157, AS13193, AS13213, AS13237, AS13238, AS13285, AS132876,
AS133229, AS13335, AS13414, AS134390, AS13445, AS13720, AS13768,
AS15133, AS15169, AS15224, AS15399, AS15412, AS15426, AS15435,
AS15447, AS1547, AS15480, AS15516, AS15542, AS15547, AS15557, AS15589,
AS15600, AS15703, AS15704, AS15720, AS15763, AS15772, AS15802,
AS15830, AS15879, AS15958, AS15966, AS16080, AS16097, AS16147,
AS16211, AS16243, AS16265, AS16276, AS16298, AS16347, AS16509,
AS16637, AS1668, AS1680, AS16839, AS17451, AS17511, AS17557, AS1764,
AS18001, AS18106, AS1836, AS18403, AS18705, AS19551, AS196640,
AS196689, AS196752, AS19679, AS196844, AS197000, AS197038, AS197156,
AS197219, AS197426, AS197440, AS197541, AS197595, AS197612, AS197692,
AS197902, AS197981, AS197985, AS197999, AS198028, AS198126, AS198435,
AS198721, AS198792, AS199139, AS199275, AS199332, AS199524, AS199547,
AS199636, AS199660, AS199670, AS199837, AS199939, AS199947, AS199981,
AS200020, AS200023, AS200043, AS200052, AS200130, AS20080, AS200904,
AS200981, AS201214, AS201791, AS201877, AS202018, AS202023, AS202112,
AS202189, AS202197, AS20278, AS203565, AS203923, AS203959, AS20473,
AS20485, AS20495, AS20504, AS20559, AS20562, AS20621, AS20634,
AS20640, AS20764, AS20766, AS20771, AS20804, AS20847, AS20857,
AS20886, AS20932, AS20940, AS20953, AS20969, AS21011, AS21073, AS2119,
AS21219, AS21221, AS21263, AS21267, AS21320, AS21339, AS21345,
AS21478, AS22363, AS22822, AS23148, AS23393, AS23576, AS23947,
AS24167, AS24218, AS24482, AS24586, AS24608, AS24642, AS24724,
AS24730, AS24778, AS2484, AS24875, AS24904, AS24940, AS24961, AS25074,
AS25091, AS251, AS25148, AS25151, AS25152, AS25160, AS25178, AS25180,
AS25182, AS25220, AS25229, AS25291, AS25358, AS25459, AS25542,
AS25596, AS25751, AS260, AS2603, AS2611, AS26120, AS2635, AS26496,
AS26667, AS2686, AS27281, AS27313, AS27381, AS2818, AS2828, AS28283,
AS2852, AS2854, AS286, AS28685, AS28788, AS28801, AS28836, AS28876,
AS28917, AS28929, AS29014, AS29017, AS2906, AS29073, AS29075, AS29119,
AS2914, AS29140, AS29141, AS29169, AS29263, AS29278, AS29314, AS29396,
AS29405, AS29438, AS29462, AS29467, AS29527, AS29535, AS29587,
AS29697, AS29791, AS29990, AS30071, AS30081, AS30094, AS30132,
AS30419, AS30740, AS30781, AS30818, AS30844, AS30889, AS30925,
AS30961, AS31019, AS31027, AS31042, AS31078, AS31122, AS31133,
AS31216, AS31220, AS31328, AS31334, AS31383, AS31424, AS31449,
AS31477, AS31499, AS31500, AS31529, AS31595, AS31638, AS31673, AS3209,
AS3216, AS3223, AS3225, AS32338, AS32421, AS3252, AS3257, AS32590,
AS3262, AS3265, AS3267, AS3292, AS32934, AS3302, AS3303, AS33031,
AS3320, AS3327, AS3333, AS3356, AS33796, AS33873, AS33891, AS33915,
AS33986, AS34019, AS34106, AS34141, AS34177, AS34215, AS34224,
AS34288, AS34305, AS34309, AS34442, AS34569, AS34624, AS34655,
AS34695, AS34756, AS34758, AS34762, AS34781, AS34803, AS34863,
AS34868, AS3491, AS34968, AS34984, AS35017, AS35156, AS35297, AS35320,
AS35366, AS35432, AS35574, AS35592, AS3561, AS35662, AS35699, AS35745,
AS35819, AS36089, AS36236, AS36351, AS36408, AS36692, AS36944,
AS37009, AS37090, AS37100, AS37187, AS37282, AS37662, AS38082, AS3856,
AS38880, AS38915, AS39063, AS39120, AS39122, AS39180, AS39216,
AS39244, AS39309, AS39318, AS39326, AS39386, AS39442, AS39537,
AS39591, AS39637, AS39647, AS39651, AS39704, AS39709, AS39737,
AS39792, AS39832, AS39912, AS39923, AS39995, AS40627, AS40999,
AS41059, AS41090, AS41095, AS41153, AS41313, AS41331, AS4134, AS41393,
AS41552, AS41690, AS41692, AS41887, AS41913, AS42, AS42093, AS42152,
AS42158, AS42184, AS42473, AS42525, AS42541, AS42567, AS42707,
AS42708, AS42794, AS42841, AS42861, AS42927, AS42949, AS43142,
AS43190, AS43293, AS43350, AS43366, AS43376, AS43404, AS43437,
AS43531, AS43821, AS43942, AS43996, AS44020, AS44034, AS44050,
```

```
AS44053, AS44066, AS44141, AS44160, AS44431, AS44444, AS44608,
AS44654, AS44735, AS44788, AS44814, AS44953, AS44981, AS45037,
AS45204, AS45352, AS45474, AS45629, AS4589, AS46235, AS4637, AS46391,
AS46489, AS4651, AS4657, AS46786, AS47065, AS47143, AS47172, AS47195,
AS47329, AS47463, AS47531, AS47541, AS4761, AS47622, AS4766, AS47748,
AS47764, AS47836, AS47869, AS47872, AS4788, AS47927, AS48056, AS48072,
AS48166, AS48173, AS48200, AS48283, AS48519, AS48522, AS48526,
AS48961, AS49206, AS49218, AS49285, AS49289, AS49360, AS49375,
AS49405, AS49453, AS49463, AS49535, AS49544, AS49605, AS49653,
AS49685, AS49820, AS49981, AS50010, AS50056, AS50072, AS50189,
AS50245, AS50266, AS50272, AS50295, AS50300, AS50304, AS50316,
AS50324, AS50352, AS50384, AS50473, AS50606, AS50618, AS50622,
AS50629, AS50673, AS50763, AS5089, AS50923, AS51028, AS51050, AS51088,
AS51092, AS51185, AS51313, AS51468, AS51555, AS51646, AS51752,
AS51852, AS51906, AS51942, AS52041, AS52233, AS52320, AS52438,
AS53264, AS53563, AS5390, AS5394, AS5400, AS5410, AS54104, AS54113,
AS5413, AS54183, AS54312, AS5466, AS55195, AS5524, AS5563, AS5568,
AS55799, AS5580, AS55819, AS5583, AS5588, AS559, AS5615, AS56396,
AS56550, AS56583, AS56630, AS56654, AS56665, AS56704, AS56953,
AS57023, AS57043, AS57119, AS5713, AS57463, AS57685, AS57724, AS57866,
AS57928, AS57976, AS58073, AS58138, AS58209, AS58291, AS58453,
AS58511, AS59318, AS59560, AS59605, AS59624, AS59701, AS59865,
AS59940, AS60239, AS60357, AS60447, AS60718, AS60764, AS60822,
AS60868, AS60893, AS61124, AS61180, AS61266, AS61319, AS61349,
AS61387, AS61955, AS62041, AS62052, AS62245, AS62267, AS62325,AS62713,
AS62715, AS63399, AS63541, AS6412, AS64518, AS6453, AS64597, AS6461,
AS64650, AS6507, AS6661, AS6663, AS6667, AS6677, AS6717, AS6724,
AS6730, AS6735, AS6762, AS6774, AS6805, AS6830, AS6834, AS6866,
AS6898, AS6908, AS6939, AS702, AS714, AS7160, AS7342, AS7385, AS7415,
AS7713, AS8075, AS8218, AS8220, AS8251, AS8262, AS8282, AS8283,
AS8304, AS8309, AS8315, AS8359, AS8365, AS8368, AS8399, AS8400,
AS8403, AS8422, AS8426, AS8447, AS8452, AS8455, AS8462, AS8468,
AS8473, AS8487, AS8529, AS8551, AS8560, AS8587, AS8608, AS8613,
AS8632, AS8648, AS8657, AS8674, AS8708, AS8717, AS8732, AS8763,
AS8764, AS8767, AS8781, AS8816, AS8839, AS8866, AS8881, AS8918,
AS8926, AS8928, AS8932, AS8966, AS9002, AS9009, AS9029, AS9031,
AS9036, AS9050, AS9066, AS9115, AS9121, AS9143, AS9145, AS9150,
AS9167, AS9269, AS9304, AS9318, AS9329, AS9498, AS9505, AS9583
```

Appendix B

Extensive results

This chapter features more extensive results on which the conclusions where based upon. They are given here for the sake of completeness and for future reference to be able to quantitatively compare the results of possible future research with those obtained during the analysis performed for this thesis.

B.1 Reachability of DNS data in failure scenarios

This section features extensive results for every simulated scenario and compares these to the baseline computations. For every scenario a table is given showing the exact results in numbers and three plots visualizing the obtained results.

B.1.1 Remarks about the plots

The numbers of unreachable ASs and unreachable domains plotted are split up into three categories, which form supersets of each other. The first category, shown in blue, shows those ASs and domain names which are always unreachable, even in the baseline calculations. The second category, shown in green, contains the number of autonomous systems and domain names which are expected to become unavailable in the simulation. By removing an autonomous system, for example, we expect that autonomous system to become unreachable. However, sometimes also autonomous systems will be removed from the topology which are not hosting any domain names and are thus not considered in the analysis. In this case we do not expect that any additional ASs becomes unreachable. Basically the same holds for the green areas in the plot showing the number of unreachable domains. This contains the number of domain names hosted solely on name servers located in that AS. Finally, the orange part shows the number of unreachable ASs and domain names which become unreachable in the simulation on top of those which were already unreachable in the baseline and those which were expected to become unavailable. In some situations when there is a resolver AS with a very large number of unreachable

ASs or domain names, the linear scale of the x-axis prevents the bars for other resolver locations to be shown as they are too small compared to the long one. However, it was chosen not to use a logarithmic scale here as in most of the cases a logarithmic scale would result in less clear plots as the orange part, which is the most important one, would then be too small to be seen. If not all bars show up in the plots one should therefore also consider to investigate the numbers shown in the table that corresponds to the graph.

The third plot shows the mean length of the shortest path from a resolver location to the name server of a domain. This facilitates a comparison between the performance of an resolver location to other resolver locations. Furthermore it features the mean path lengths of all investigated resolvers which allows for a comparison of the mean path lengths between multiple simulations.

Table B.1: Reachability of DNS data in a simulation where AS20857 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	37,926	1.25765826109
AS2637	3	0	2.18911427628	4	37,926	2.26855379396
AS3215	6	1	2.83567157466	7	37,927	2.90853206044
AS3320	5	1	1.81932553893	6	37,927	1.89131685621
AS3356	3	0	1.74099517073	4	37,926	1.74038730495
AS4134	1	0	2.09931352367	2	37,926	2.17416482725
AS5432	1	0	2.2768154119	2	37,926	2.36084659223
AS6830	4	1	1.90947226796	5	37,927	1.98006668468
AS8075	1	0	1.64468828964	2	37,926	1.71308079691
AS8737	1	0	2.88639271487	2	37,926	2.9604425645
AS8972	1	0	2.22328412605	2	37,926	2.30419672971
AS9121	1	0	2.09727243649	2	37,926	2.16953133008
AS9142	1	0	1.93983546041	2	37,926	2.01443814388
AS13127	5	1	2.03705757563	6	37,927	2.18391412879
AS13238	2	0	1.66266719207	3	37,926	1.73197465975
AS13414	1	0	1.79106909723	2	37,926	1.86457430284
AS13335	1	0	1.64517926897	2	37,926	1.71374272508
AS14618	1	0	2.65234637417	2	37,926	2.72089515366
AS15169	4	1	1.65053505386	5	37,927	1.7189451724
AS16276	3	0	1.66660210991	4	37,926	1.73615085204
AS16509	1	0	1.82501377501	2	37,926	1.89489647
AS17204	1	0	2.35903972347	2	37,926	2.45829458319
AS20857	1	0	1.55288727159	n/a	n/a	n/a
AS20940	1	0	1.62063850426	2	37,926	1.68876610657
AS23033	1	0	2.49950622466	2	37,926	2.56693696965
AS24793	1	0	2.22269584558	2	37,926	2.38936957631
AS24940	1	0	1.61063288987	2	37,926	1.6788677837
AS31615	1	0	2.00108969823	2	37,926	2.07592931824
AS32934	1	0	1.75897836038	2	37,926	1.82816412364
AS34173	1	0	2.21939263956	2	37,926	2.30019681381
AS35470	1	0	2.638705947	2	37,926	2.63763957091
AS36351	1	0	1.4947956937	2	37,926	1.56195204606
AS36647	1	0	2.64950842419	2	37,926	2.71792004373
AS36692	1	0	1.90835667691	2	37,926	1.98224113934
AS49544	1	0	1.71134721447	2	37,926	1.78027082361
AS55967	1	0	2.25354329006	2	37,926	2.3366163794
AS60781	2	0	1.82534314497	3	37,926	1.90086170808
AS197902	1	0	1.77303725702	2	37,926	1.8460898743
AS393406	2	0	2.14028867497	3	37,926	2.21922831866
		Mean	2.0003648877		**Mean**	2.0862540681

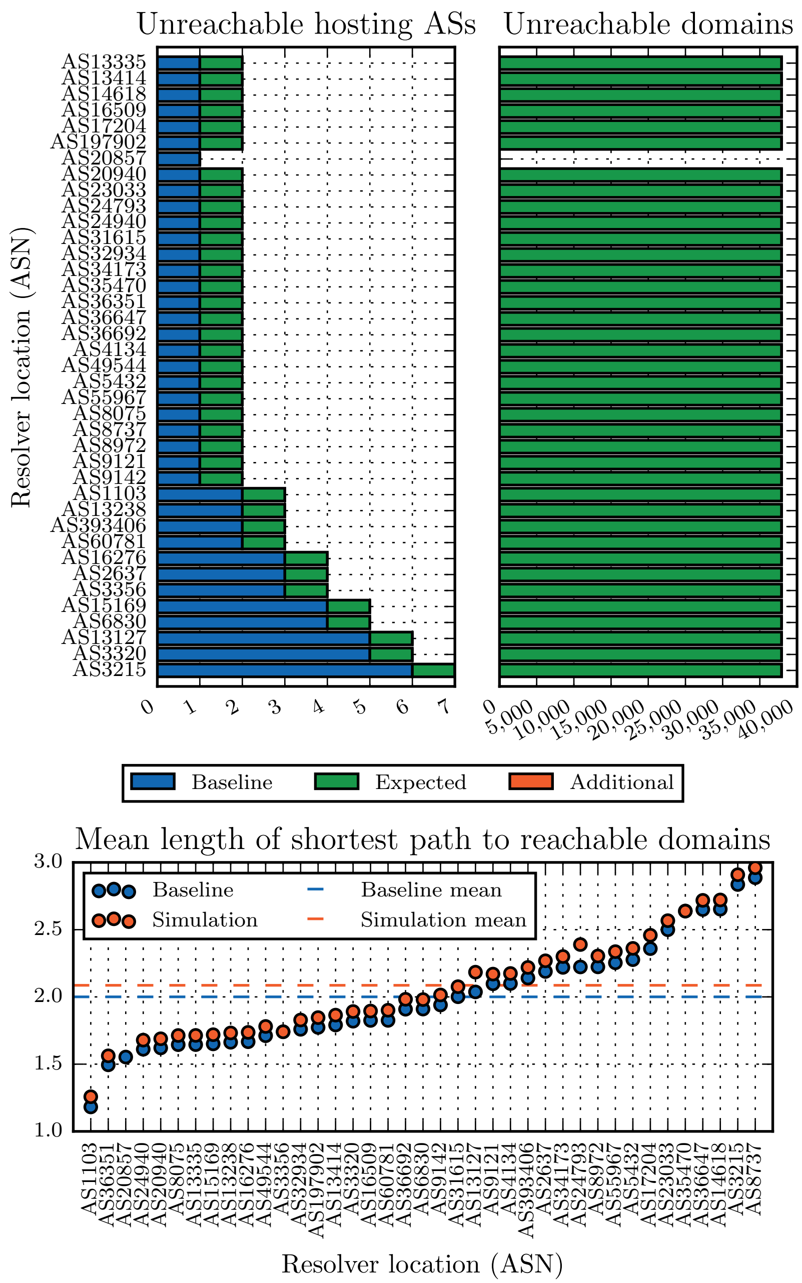


Figure B.1: Plots showing the reachability of autonomous systems and domains in a simulation where AS20857 was removed from the AS topology compared to the baseline.

Table B.2: Reachability of DNS data in a simulation where AS60781 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	77,122	1.17001300763
AS2637	3	0	2.18911427628	4	77,122	2.29940336625
AS3215	6	1	2.83567157466	7	77,123	3.0948819564
AS3320	5	1	1.81932553893	6	77,123	1.99419819523
AS3356	3	0	1.74099517073	4	77,122	1.82559866678
AS4134	1	0	2.09931352367	2	77,122	2.19438557579
AS5432	1	0	2.2768154119	2	77,122	2.41022844484
AS6830	4	1	1.90947226796	5	77,123	1.91479225707
AS8075	1	0	1.64468828964	2	77,122	1.64347199691
AS8737	1	0	2.88639271487	2	77,122	2.87017372126
AS8972	1	0	2.22328412605	2	77,122	2.32287515891
AS9121	1	0	2.09727243649	2	77,122	2.10097196003
AS9142	1	0	1.93983546041	2	77,122	1.92437343811
AS13127	5	1	2.03705757563	6	77,123	2.06635953677
AS13238	2	0	1.66266719207	3	77,122	1.66244293796
AS13414	1	0	1.79106909723	2	77,122	1.79504529976
AS13335	1	0	1.64517926897	2	77,122	1.64131093
AS14618	1	0	2.65234637417	2	77,122	2.80490692869
AS15169	4	1	1.65053505386	5	77,123	1.64943997019
AS16276	3	0	1.66660210991	4	77,122	1.66647590827
AS16509	1	0	1.82501377501	2	77,122	1.8239155045
AS17204	1	0	2.35903972347	2	77,122	2.47907923836
AS20857	1	0	1.55288727159	2	77,122	1.5489872091
AS20940	1	0	1.62063850426	2	77,122	1.61638291829
AS23033	1	0	2.49950622466	2	77,122	2.64339029734
AS24793	1	0	2.22269584558	2	77,122	2.21269516645
AS24940	1	0	1.61063288987	2	77,122	1.60897076328
AS31615	1	0	2.00108969823	2	77,122	2.03325380234
AS32934	1	0	1.75897836038	2	77,122	1.756836192
AS34173	1	0	2.21939263956	2	77,122	2.31849610017
AS35470	1	0	2.638705947	2	77,122	2.64011550654
AS36351	1	0	1.4947956937	2	77,122	1.48864470638
AS36647	1	0	2.64950842419	2	77,122	2.64858672995
AS36692	1	0	1.90835667691	2	77,122	1.92654452834
AS49544	1	0	1.71134721447	2	77,122	1.70851052241
AS55967	1	0	2.25354329006	2	77,122	2.35449118761
AS60781	2	0	1.82534314497	n/a	n/a	n/a
AS197902	1	0	1.77303725702	2	77,122	1.75514281727
AS393406	2	0	2.14028867497	3	77,122	2.23873614559
		Mean	2.0003648877		**Mean**	2.0487928577

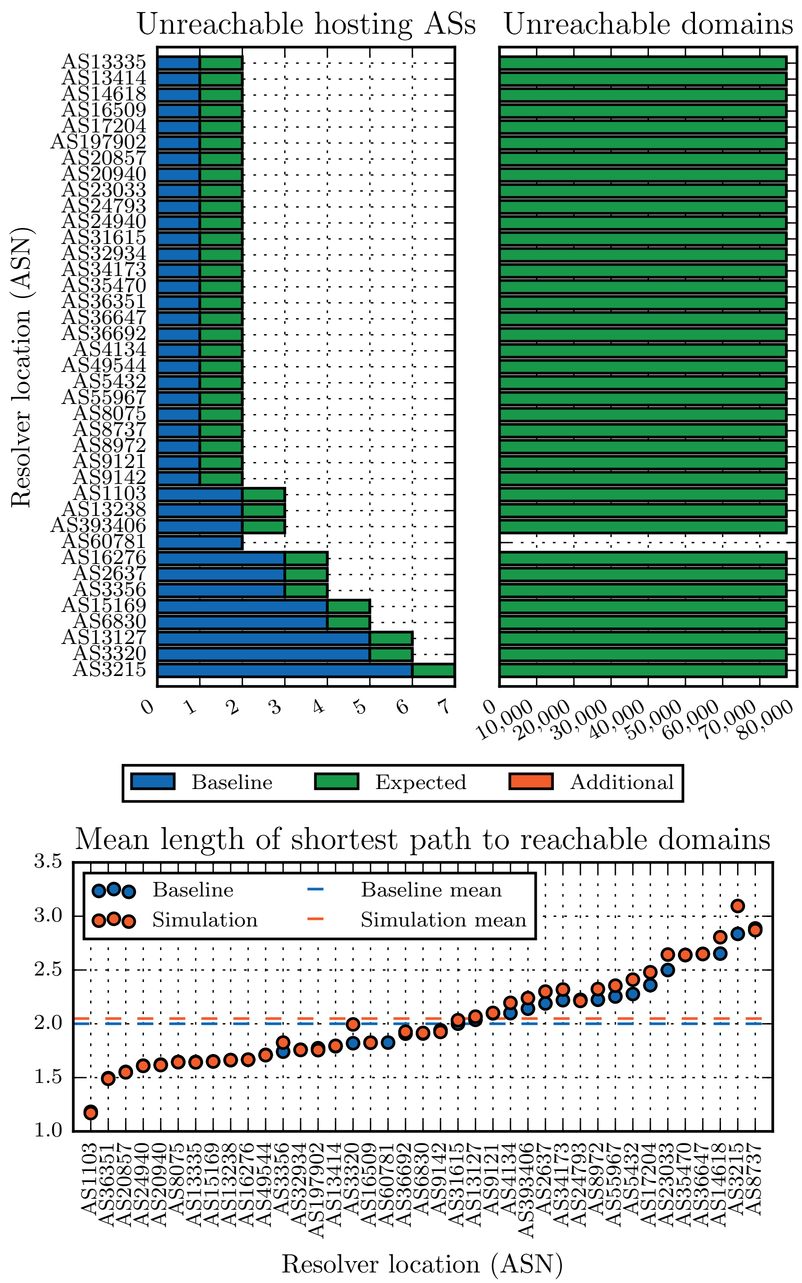


Figure B.2: Plots showing the reachability of autonomous systems and domains in a simulation where AS60781 was removed from the AS topology compared to the baseline.

Table B.3: Reachability of DNS data in a simulation where AS21155 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	80,739	1.16930861163
AS2637	3	0	2.18911427628	5	80,739	2.17679605166
AS3215	6	1	2.83567157466	8	80,740	2.81795188083
AS3320	5	1	1.81932553893	7	80,740	1.8013558923
AS3356	3	0	1.74099517073	5	80,739	1.73722896968
AS4134	1	0	2.09931352367	3	80,739	2.08562221887
AS5432	1	0	2.2768154119	3	80,739	2.26767049284
AS6830	4	1	1.90947226796	6	80,740	1.8928092629
AS8075	1	0	1.64468828964	3	80,739	1.62397945212
AS8737	1	0	2.88639271487	3	80,739	2.86942248265
AS8972	1	0	2.22328412605	3	80,739	2.21148706229
AS9121	1	0	2.09727243649	3	80,739	2.08354975512
AS9142	1	0	1.93983546041	3	80,739	1.92365513643
AS13127	5	1	2.03705757563	7	80,740	2.00706428102
AS13238	2	0	1.66266719207	4	80,739	1.64223306786
AS13414	1	0	1.79106909723	3	80,739	1.77259692331
AS13335	1	0	1.64517926897	3	80,739	1.62447793349
AS14618	1	0	2.65234637417	3	80,739	2.63175455035
AS15169	4	1	1.65053505386	6	80,740	1.62991554959
AS16276	3	0	1.66660210991	5	80,739	1.6462281103
AS16509	1	0	1.82501377501	3	80,739	1.80706026761
AS17204	1	0	2.35903972347	3	80,739	2.35120435106
AS20857	1	0	1.55288727159	3	80,739	1.53079466144
AS20940	1	0	1.62063850426	3	80,739	1.5995621918
AS23033	1	0	2.49950622466	3	80,739	2.47657904005
AS24793	1	0	2.22269584558	3	80,739	2.19555780047
AS24940	1	0	1.61063288987	3	80,739	1.58940369402
AS31615	1	0	2.00108969823	3	80,739	1.98582658866
AS32934	1	0	1.75897836038	3	80,739	1.7400158477
AS34173	1	0	2.21939263956	3	80,739	2.20755485046
AS35470	1	0	2.638705947	3	80,739	2.73609158431
AS36351	1	0	1.4947956937	3	80,739	1.4717965333
AS36647	1	0	2.64950842419	3	80,739	2.62887323717
AS36692	1	0	1.90835667691	3	80,739	1.89167662904
AS49544	1	0	1.71134721447	3	80,739	1.69165690931
AS55967	1	0	2.25354329006	3	80,739	2.2440437248
AS60781	2	0	1.82534314497	4	80,739	1.80746563857
AS197902	1	0	1.77303725702	3	80,739	1.75430829654
AS393406	2	0	2.14028867497	4	80,739	2.12722346065
		Mean	2.0003648877		**Mean**	1.9859436665

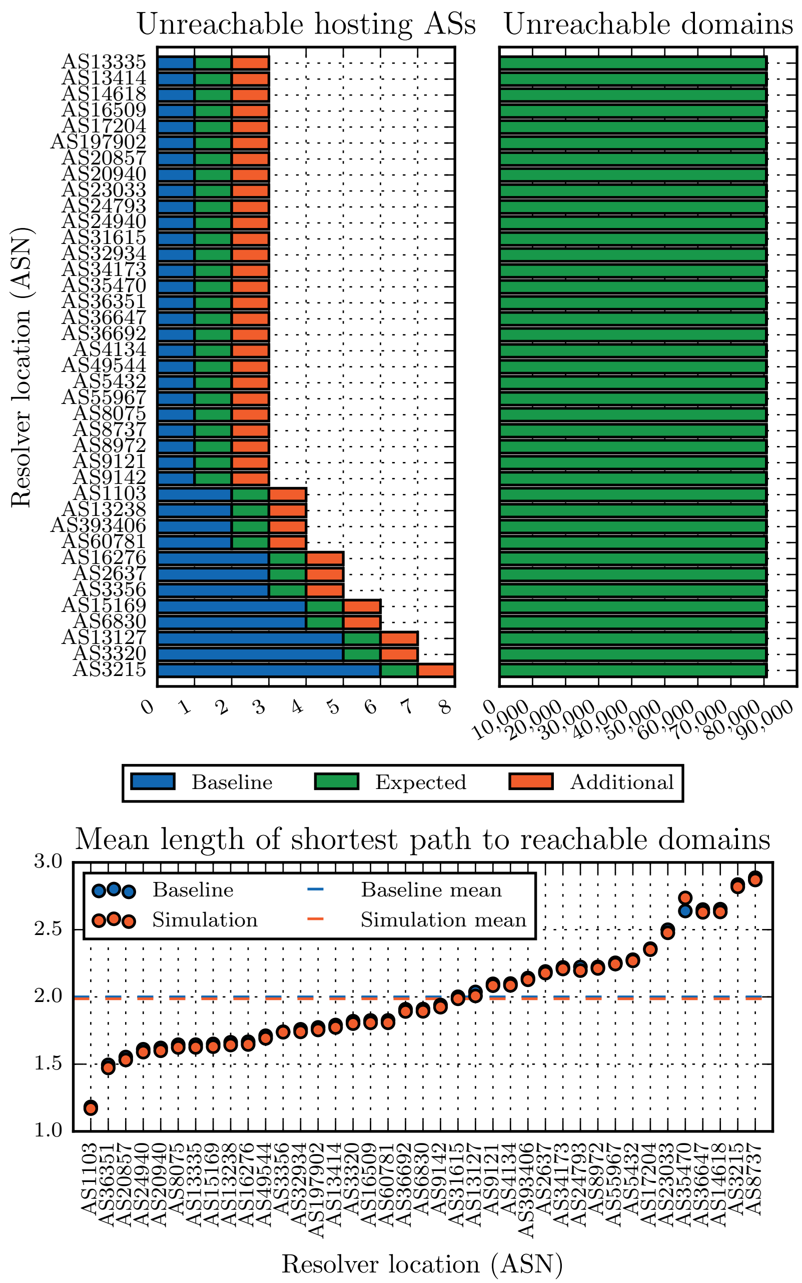


Figure B.3: Plots showing the reachability of autonomous systems and domains in a simulation where AS21155 was removed from the AS topology compared to the baseline.

Table B.4: Reachability of DNS data in a simulation where AS12859 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	15,863	1.20603261403
AS2637	3	0	2.18911427628	4	15,863	2.18687923274
AS3215	6	1	2.83567157466	7	15,864	2.83622594002
AS3320	5	1	1.81932553893	6	15,864	1.81978936324
AS3356	3	0	1.74099517073	4	15,863	1.74048187253
AS4134	1	0	2.09931352367	2	15,863	2.10067069551
AS5432	1	0	2.2768154119	2	15,863	2.33473697986
AS6830	4	1	1.90947226796	5	15,864	2.00994013001
AS8075	1	0	1.64468828964	2	15,863	1.71727608819
AS8737	1	0	2.88639271487	2	15,863	3.03589001528
AS8972	1	0	2.22328412605	2	15,863	2.22700037858
AS9121	1	0	2.09727243649	2	15,863	2.09856503877
AS9142	1	0	1.93983546041	2	15,863	2.05601574148
AS13127	5	1	2.03705757563	6	15,864	2.20488531899
AS13238	2	0	1.66266719207	3	15,863	1.73533429614
AS13414	1	0	1.79106909723	2	15,863	1.86739896334
AS13335	1	0	1.64517926897	2	15,863	1.71779301448
AS14618	1	0	2.65234637417	2	15,863	2.7438044093
AS15169	4	1	1.65053505386	5	15,864	1.72368685739
AS16276	3	0	1.66660210991	4	15,863	1.7392730315
AS16509	1	0	1.82501377501	2	15,863	1.92915380193
AS17204	1	0	2.35903972347	2	15,863	2.44098636642
AS20857	1	0	1.55288727159	2	15,863	1.65431895194
AS20940	1	0	1.62063850426	2	15,863	1.68946246208
AS23033	1	0	2.49950622466	2	15,863	2.49860055993
AS24793	1	0	2.22269584558	2	15,863	2.40424515206
AS24940	1	0	1.61063288987	2	15,863	1.68346163014
AS31615	1	0	2.00108969823	2	15,863	2.10852479704
AS32934	1	0	1.75897836038	2	15,863	1.8610571283
AS34173	1	0	2.21939263956	2	15,863	2.22325981389
AS35470	1	0	2.638705947	2	15,863	2.73348532649
AS36351	1	0	1.4947956937	2	15,863	1.5645530644
AS36647	1	0	2.64950842419	2	15,863	2.72234215286
AS36692	1	0	1.90835667691	2	15,863	2.02256995565
AS49544	1	0	1.71134721447	2	15,863	1.78448809808
AS55967	1	0	2.25354329006	2	15,863	2.26766312857
AS60781	2	0	1.82534314497	3	15,863	1.82652458204
AS197902	1	0	1.77303725702	2	15,863	1.88832241993
AS393406	2	0	2.14028867497	3	15,863	2.14655916843
		Mean	2.0003648877		**Mean**	2.0654168857

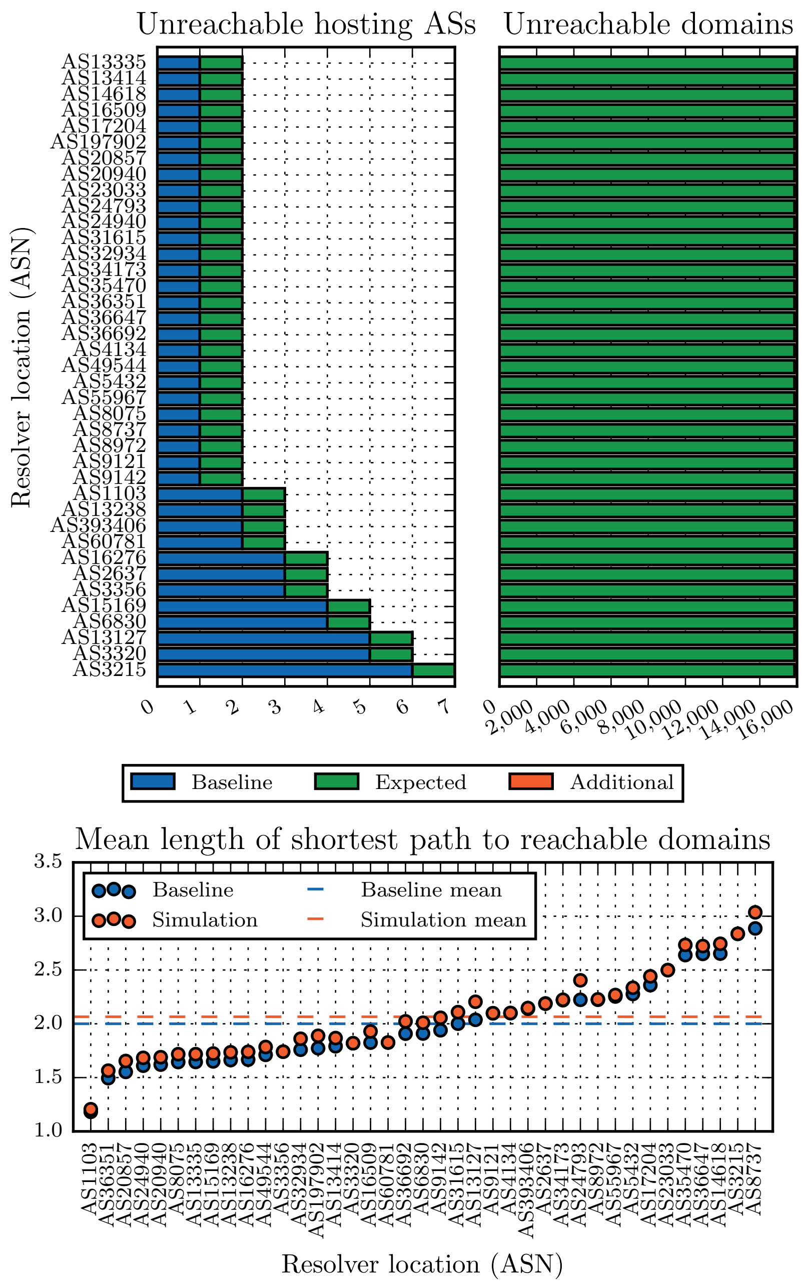


Figure B.4: Plots showing the reachability of autonomous systems and domains in a simulation where AS12859 was removed from the AS topology compared to the baseline.

Table B.5: Reachability of DNS data in a simulation where AS8455 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	8	11,369	1.18044057452
AS2637	3	0	2.18911427628	12	12,565	2.18778160028
AS3215	6	1	2.83567157466	15	12,566	2.833396111
AS3320	5	1	1.81932553893	14	12,566	1.81692725003
AS3356	3	0	1.74099517073	12	12,565	1.73821382256
AS4134	1	0	2.09931352367	10	12,565	2.09754246039
AS5432	1	0	2.2768154119	6	11,369	2.27628605196
AS6830	4	1	1.90947226796	13	12,566	1.90716528574
AS8075	1	0	1.64468828964	6	11,369	1.73530971516
AS8737	1	0	2.88639271487	6	11,369	3.00850017531
AS8972	1	0	2.22328412605	6	11,369	2.22268255109
AS9121	1	0	2.09727243649	6	11,369	2.09566297725
AS9142	1	0	1.93983546041	6	11,369	2.03330880695
AS13127	5	1	2.03705757563	14	12,566	2.03553514783
AS13238	2	0	1.66266719207	10	12,565	1.75367225917
AS13414	1	0	1.79106909723	6	11,369	1.78869765285
AS13335	1	0	1.64517926897	6	11,369	1.73580509951
AS14618	1	0	2.65234637417	6	11,369	2.74366624395
AS15169	4	1	1.65053505386	13	12,566	1.74342338565
AS16276	3	0	1.66660210991	12	12,565	1.77215448609
AS16509	1	0	1.82501377501	6	11,369	1.91709765292
AS17204	1	0	2.35903972347	6	11,369	2.35893566336
AS20857	1	0	1.55288727159	6	11,369	1.55020800726
AS20940	1	0	1.62063850426	6	11,369	1.71117896806
AS23033	1	0	2.49950622466	6	11,369	2.59100175047
AS24793	1	0	2.22269584558	6	11,369	2.22538997975
AS24940	1	0	1.61063288987	7	11,369	1.71560212268
AS31615	1	0	2.00108969823	6	11,369	1.99934004792
AS32934	1	0	1.75897836038	6	11,369	1.84997512804
AS34173	1	0	2.21939263956	6	11,369	2.21868884165
AS35470	1	0	2.638705947	12	12,565	2.7943486286
AS36351	1	0	1.4947956937	6	11,369	1.58479095322
AS36647	1	0	2.64950842419	6	11,369	2.74068385083
AS36692	1	0	1.90835667691	6	11,369	2.0209701501
AS49544	1	0	1.71134721447	6	11,369	1.70876350235
AS55967	1	0	2.25354329006	6	11,369	2.25292808204
AS60781	2	0	1.82534314497	11	12,565	1.82301073778
AS197902	1	0	1.77303725702	6	11,369	1.87297556944
AS393406	2	0	2.14028867497	11	12,565	2.13873917436
		Mean	2.0003648877		**Mean**	2.0456615505

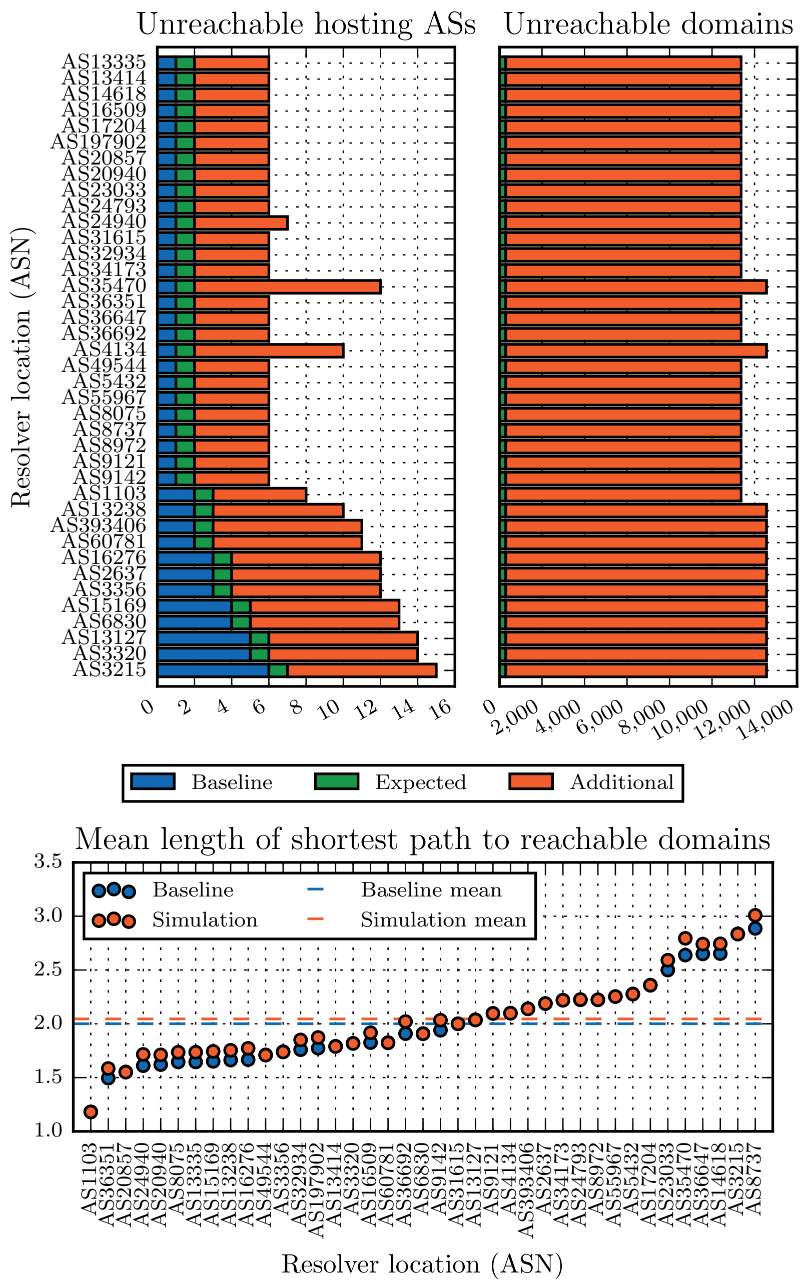


Figure B.5: Plots showing the reachability of autonomous systems and domains in a simulation where AS8455 was removed from the AS topology compared to the baseline.

Table B.6: Reachability of DNS data in a simulation where AS197902 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	7,155	1.26608914795
AS2637	3	0	2.18911427628	4	7,155	2.18803135638
AS3215	6	1	2.83567157466	7	7,156	2.83411663959
AS3320	5	1	1.81932553893	6	7,156	1.8190848121
AS3356	3	0	1.74099517073	4	7,155	1.73931379772
AS4134	1	0	2.09931352367	2	7,155	2.09811067686
AS5432	1	0	2.2768154119	2	7,155	2.27584961493
AS6830	4	1	1.90947226796	5	7,156	1.90935174346
AS8075	1	0	1.64468828964	2	7,155	1.64421415203
AS8737	1	0	2.88639271487	2	7,155	2.97029845083
AS8972	1	0	2.22328412605	2	7,155	2.22224683923
AS9121	1	0	2.09727243649	2	7,155	2.09740290162
AS9142	1	0	1.93983546041	2	7,155	2.02380920828
AS13127	5	1	2.03705757563	6	7,156	2.03710762516
AS13238	2	0	1.66266719207	3	7,155	1.66221706489
AS13414	1	0	1.79106909723	2	7,155	1.79079044795
AS13335	1	0	1.64517926897	2	7,155	1.64470578705
AS14618	1	0	2.65234637417	2	7,155	2.65188246377
AS15169	4	1	1.65053505386	5	7,156	1.65006872439
AS16276	3	0	1.66660210991	4	7,155	1.66615723772
AS16509	1	0	1.82501377501	2	7,155	1.82478045809
AS17204	1	0	2.35903972347	2	7,155	2.35818373524
AS20857	1	0	1.55288727159	2	7,155	1.55229146901
AS20940	1	0	1.62063850426	2	7,155	1.6201322487
AS23033	1	0	2.49950622466	2	7,155	2.49883819963
AS24793	1	0	2.22269584558	2	7,155	2.30705201345
AS24940	1	0	1.61063288987	2	7,155	1.61011327204
AS31615	1	0	2.00108969823	2	7,155	2.0010919001
AS32934	1	0	1.75897836038	2	7,155	1.75865685462
AS34173	1	0	2.21939263956	2	7,155	2.21835015575
AS35470	1	0	2.638705947	2	7,155	2.63822438006
AS36351	1	0	1.4947956937	2	7,155	1.49412100456
AS36647	1	0	2.64950842419	2	7,155	2.64904072377
AS36692	1	0	1.90835667691	2	7,155	1.90823466258
AS49544	1	0	1.71134721447	2	7,155	1.71096209837
AS55967	1	0	2.25354329006	2	7,155	2.25254641369
AS60781	2	0	1.82534314497	3	7,155	1.8237744168
AS197902	1	0	1.77303725702	n/a	n/a	n/a
AS393406	2	0	2.14028867497	3	7,155	2.13914054957
		Mean	2.0003648877		**Mean**	2.0146416644

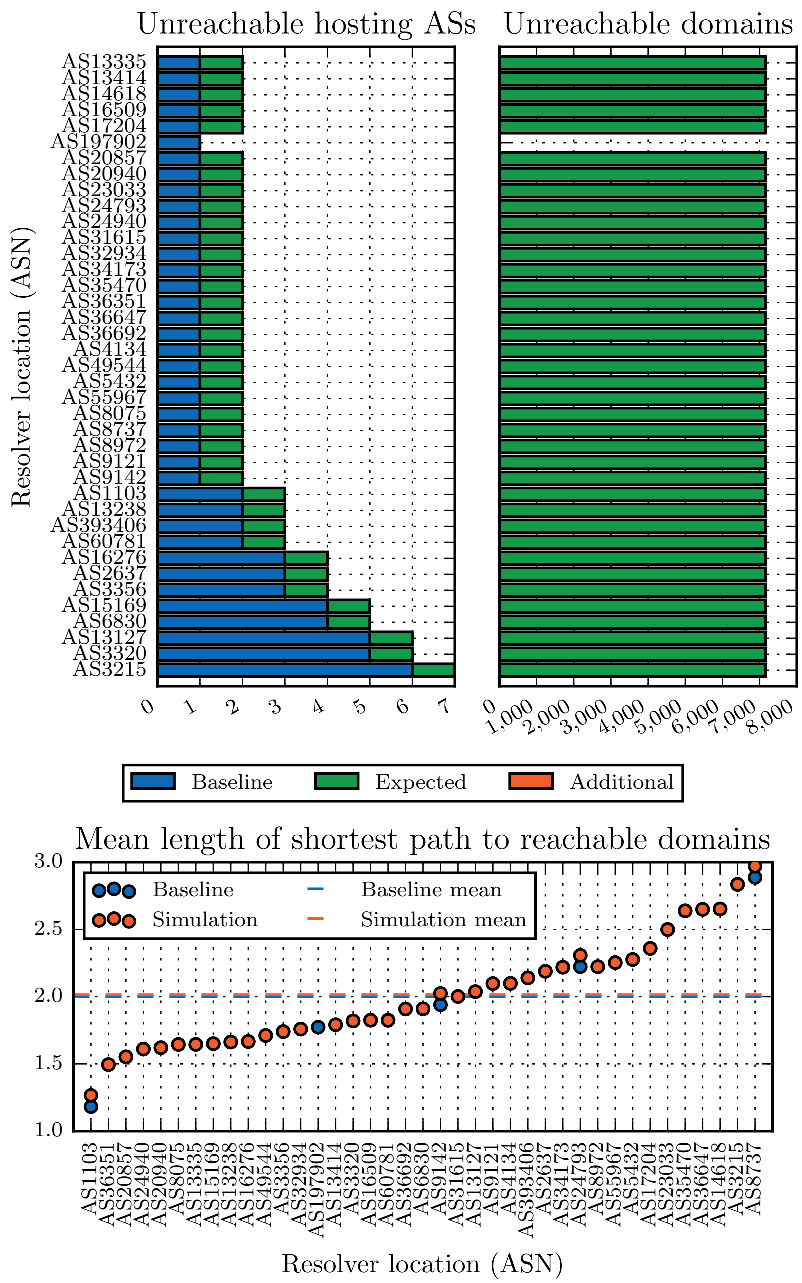


Figure B.6: Plots showing the reachability of autonomous systems and domains in a simulation where AS197902 was removed from the AS topology compared to the baseline.

Table B.7: Reachability of DNS data in a simulation where AS25151 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	1,131	1.18183079194
AS2637	3	0	2.18911427628	4	1,131	2.24187452703
AS3215	6	1	2.83567157466	7	1,132	2.83564233053
AS3320	5	1	1.81932553893	6	1,132	1.81929284801
AS3356	3	0	1.74099517073	4	1,131	1.79267372988
AS4134	1	0	2.09931352367	2	1,131	2.09934005847
AS5432	1	0	2.2768154119	2	1,131	2.27666795248
AS6830	4	1	1.90947226796	5	1,132	1.90945839927
AS8075	1	0	1.64468828964	2	1,131	1.64461336734
AS8737	1	0	2.88639271487	2	1,131	2.88615808207
AS8972	1	0	2.22328412605	2	1,131	2.22479550799
AS9121	1	0	2.09727243649	2	1,131	2.09729332058
AS9142	1	0	1.93983546041	2	1,131	1.94127644628
AS13127	5	1	2.03705757563	6	1,132	2.03685974641
AS13238	2	0	1.66266719207	3	1,131	1.66259606086
AS13414	1	0	1.79106909723	2	1,131	1.79103100739
AS13335	1	0	1.64517926897	2	1,131	1.64510445019
AS14618	1	0	2.65234637417	2	1,131	2.65227325312
AS15169	4	1	1.65053505386	5	1,132	1.65046639829
AS16276	3	0	1.66660210991	4	1,131	1.66654187619
AS16509	1	0	1.82501377501	2	1,131	1.82497706322
AS17204	1	0	2.35903972347	2	1,131	2.35890960216
AS20857	1	0	1.55288727159	2	1,131	1.55279317824
AS20940	1	0	1.62063850426	2	1,131	1.62055851073
AS23033	1	0	2.49950622466	2	1,131	2.49940590907
AS24793	1	0	2.22269584558	2	1,131	2.22253212687
AS24940	1	0	1.61063288987	2	1,131	1.6105558204
AS31615	1	0	2.00108969823	2	1,131	2.00109030089
AS32934	1	0	1.75897836038	2	1,131	1.75892753769
AS34173	1	0	2.21939263956	2	1,131	2.22090320093
AS35470	1	0	2.638705947	2	1,131	2.63863535644
AS36351	1	0	1.4947956937	2	1,131	1.4946891645
AS36647	1	0	2.64950842419	2	1,131	2.64943451828
AS36692	1	0	1.90835667691	2	1,131	1.90834275943
AS49544	1	0	1.71134721447	2	1,131	1.7112863481
AS55967	1	0	2.25354329006	2	1,131	2.25360793951
AS60781	2	0	1.82534314497	3	1,131	1.82531675683
AS197902	1	0	1.77303725702	2	1,131	1.77298977172
AS393406	2	0	2.14028867497	3	1,131	2.14032888382
		Mean	2.0003648877		**Mean**	2.0031044847

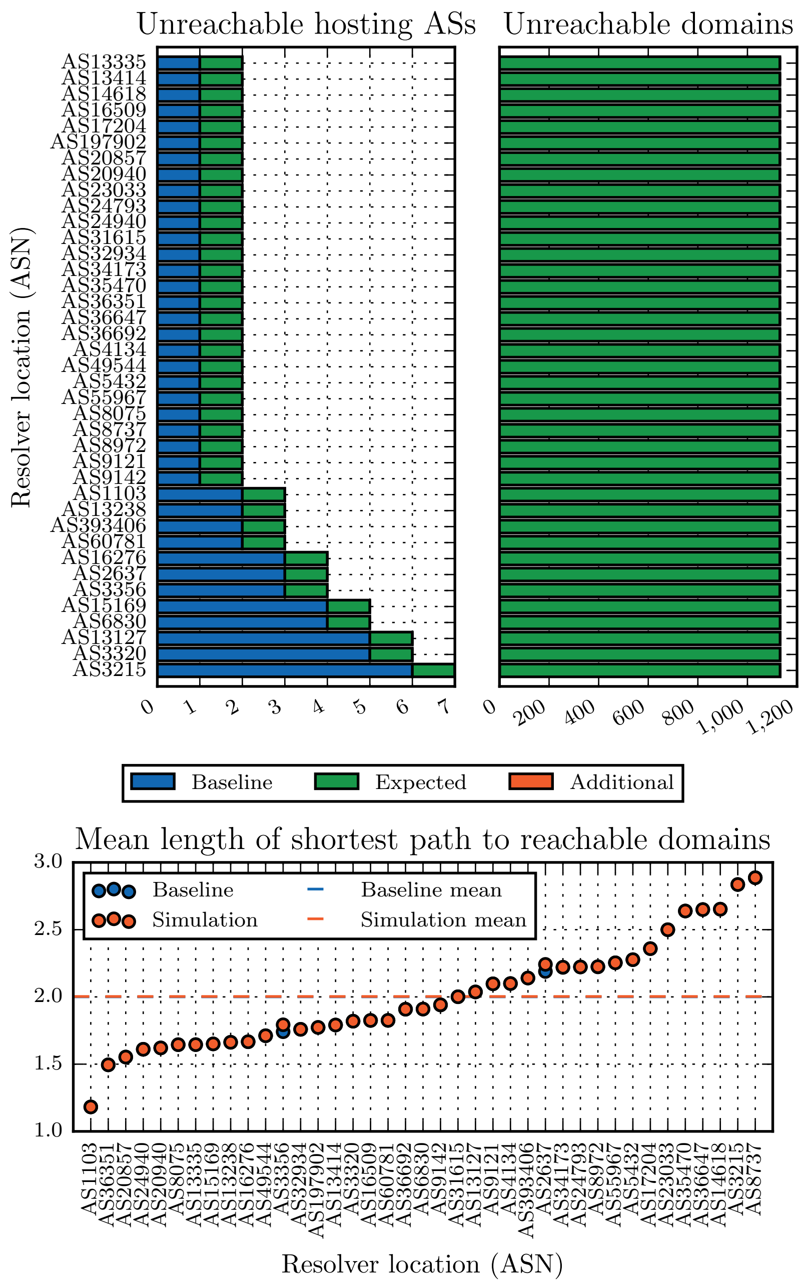


Figure B.7: Plots showing the reachability of autonomous systems and domains in a simulation where AS25151 was removed from the AS topology compared to the baseline.

Table B.8: Reachability of DNS data in a simulation where AS24940 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	3,629	1.18345641306
AS2637	3	0	2.18911427628	4	3,629	2.18969517599
AS3215	6	1	2.83567157466	7	3,630	2.83565136487
AS3320	5	1	1.81932553893	6	3,630	1.81929016082
AS3356	3	0	1.74099517073	4	3,629	1.77662143578
AS4134	1	0	2.09931352367	2	3,629	2.09957734885
AS5432	1	0	2.2768154119	2	3,629	2.27862743858
AS6830	4	1	1.90947226796	5	3,630	1.9094790715
AS8075	1	0	1.64468828964	2	3,629	1.68009212187
AS8737	1	0	2.88639271487	2	3,629	2.88754110818
AS8972	1	0	2.22328412605	2	3,629	2.24311310297
AS9121	1	0	2.09727243649	2	3,629	2.09742967892
AS9142	1	0	1.93983546041	2	3,629	1.94583018336
AS13127	5	1	2.03705757563	6	3,630	2.03835495988
AS13238	2	0	1.66266719207	3	3,629	1.69883993368
AS13414	1	0	1.79106909723	2	3,629	1.82767588128
AS13335	1	0	1.64517926897	2	3,629	1.6812073658
AS14618	1	0	2.65234637417	2	3,629	2.65218435044
AS15169	4	1	1.65053505386	5	3,630	1.68595851866
AS16276	3	0	1.66660210991	4	3,629	1.7023792803
AS16509	1	0	1.82501377501	2	3,629	1.8249690039
AS17204	1	0	2.35903972347	2	3,629	2.35866722102
AS20857	1	0	1.55288727159	2	3,629	1.55278177722
AS20940	1	0	1.62063850426	2	3,629	1.65597793313
AS23033	1	0	2.49950622466	2	3,629	2.49923067008
AS24793	1	0	2.22269584558	2	3,629	2.22225697839
AS24940	1	0	1.61063288987	n/a	n/a	n/a
AS31615	1	0	2.00108969823	2	3,629	2.00237859015
AS32934	1	0	1.75897836038	2	3,629	1.75888273413
AS34173	1	0	2.21939263956	2	3,629	2.23921916884
AS35470	1	0	2.638705947	2	3,629	2.63861041987
AS36351	1	0	1.4947956937	2	3,629	1.51602106932
AS36647	1	0	2.64950842419	2	3,629	2.68495133235
AS36692	1	0	1.90835667691	2	3,629	1.90944700577
AS49544	1	0	1.71134721447	2	3,629	1.74752007915
AS55967	1	0	2.25354329006	2	3,629	2.25526140896
AS60781	2	0	1.82534314497	3	3,629	1.8254882946
AS197902	1	0	1.77303725702	2	3,629	1.77425254502
AS393406	2	0	2.14028867497	3	3,629	2.14169790525
		Mean	2.0003648877		**Mean**	2.0221215535

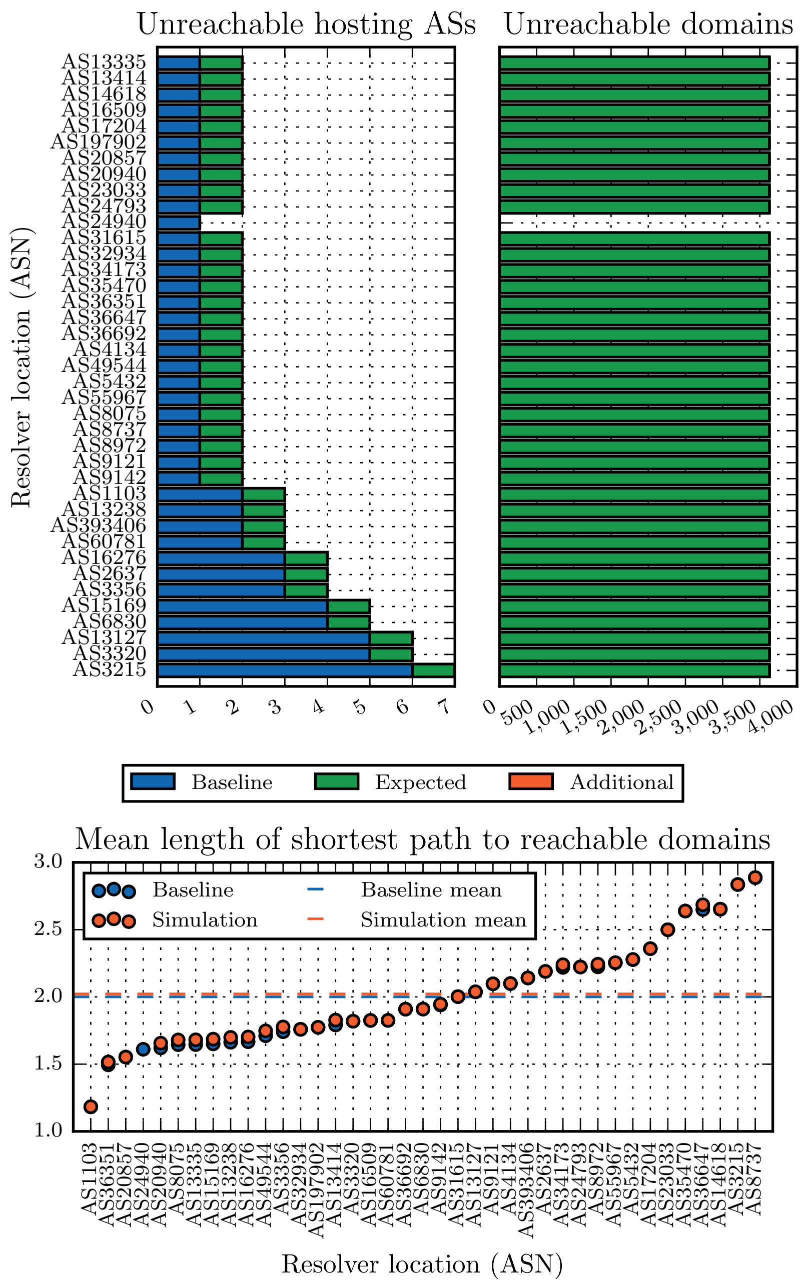


Figure B.8: Plots showing the reachability of autonomous systems and domains in a simulation where AS24940 was removed from the AS topology compared to the baseline.

Table B.9: Reachability of DNS data in a simulation where AS48635 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	39,761	1.18469990856
AS2637	3	0	2.18911427628	5	39,761	2.1942647427
AS3215	6	1	2.83567157466	8	39,762	2.83557545071
AS3320	5	1	1.81932553893	7	39,762	1.81927149276
AS3356	3	0	1.74099517073	5	39,761	1.73981352583
AS4134	1	0	2.09931352367	3	39,761	2.10142934487
AS5432	1	0	2.2768154119	3	39,761	2.28225997727
AS6830	4	1	1.90947226796	6	39,762	1.90970992442
AS8075	1	0	1.64468828964	3	39,761	1.64268087279
AS8737	1	0	2.88639271487	3	39,761	2.88724977357
AS8972	1	0	2.22328412605	3	39,761	2.21748471886
AS9121	1	0	2.09727243649	3	39,761	2.0989179961
AS9142	1	0	1.93983546041	3	39,761	1.940994102
AS13127	5	1	2.03705757563	7	39,762	2.04062797064
AS13238	2	0	1.66266719207	4	39,761	1.66087833921
AS13414	1	0	1.79106909723	3	39,761	1.79146753622
AS13335	1	0	1.64517926897	3	39,761	1.64317551817
AS14618	1	0	2.65234637417	3	39,761	2.65066468208
AS15169	4	1	1.65053505386	6	39,762	1.64857129336
AS16276	3	0	1.66660210991	5	39,761	1.6648691171
AS16509	1	0	1.82501377501	3	39,761	1.82462135873
AS17204	1	0	2.35903972347	3	39,761	2.38001741587
AS20857	1	0	1.55288727159	3	39,761	1.55054706765
AS20940	1	0	1.62063850426	3	39,761	1.61845151208
AS23033	1	0	2.49950622466	3	39,761	2.53330771844
AS24793	1	0	2.22269584558	3	39,761	2.2426342627
AS24940	1	0	1.61063288987	3	39,761	1.6083711876
AS31615	1	0	2.00108969823	3	39,761	2.00328542935
AS32934	1	0	1.75897836038	3	39,761	1.75807765857
AS34173	1	0	2.21939263956	3	39,761	2.21356473873
AS35470	1	0	2.638705947	3	39,761	2.63639376852
AS36351	1	0	1.4947956937	3	39,761	1.5078237162
AS36647	1	0	2.64950842419	3	39,761	2.6475369984
AS36692	1	0	1.90835667691	3	39,761	1.90911839508
AS49544	1	0	1.71134721447	3	39,761	1.70983752758
AS55967	1	0	2.25354329006	3	39,761	2.25847211667
AS60781	2	0	1.82534314497	4	39,761	1.82653740535
AS197902	1	0	1.77303725702	3	39,761	1.77290462565
AS393406	2	0	2.14028867497	4	39,761	2.14405992683
		Mean	2.0003648877		**Mean**	2.0025684389

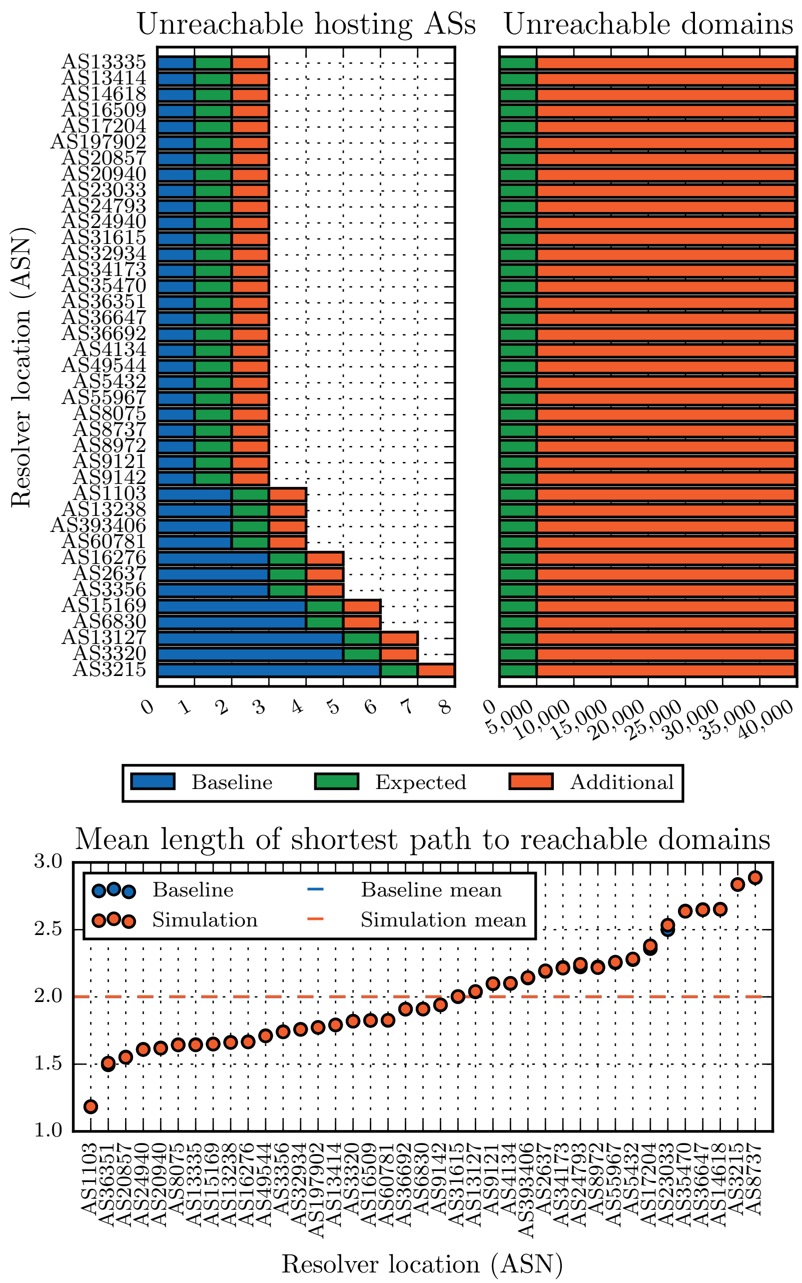


Figure B.9: Plots showing the reachability of autonomous systems and domains in a simulation where AS48635 was removed from the AS topology compared to the baseline.

Table B.10: Reachability of DNS data in a simulation where AS3265 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	46,424	1.18364576024
AS2637	3	0	2.18911427628	4	46,424	2.19751149038
AS3215	6	1	2.83567157466	7	46,425	2.87810967397
AS3320	5	1	1.81932553893	6	46,425	1.81792837382
AS3356	3	0	1.74099517073	4	46,424	1.738818742
AS4134	1	0	2.09931352367	2	46,424	2.10040098797
AS5432	1	0	2.2768154119	2	46,424	2.30658093354
AS6830	4	1	1.90947226796	5	46,425	1.90885860931
AS8075	1	0	1.64468828964	2	46,424	1.64171218818
AS8737	1	0	2.88639271487	2	46,424	2.93075991038
AS8972	1	0	2.22328412605	2	46,424	2.21654253075
AS9121	1	0	2.09727243649	2	46,424	2.09830711098
AS9142	1	0	1.93983546041	2	46,424	1.93061945365
AS13127	5	1	2.03705757563	6	46,425	2.04016649484
AS13238	2	0	1.66266719207	3	46,424	1.65985724181
AS13414	1	0	1.79106909723	2	46,424	1.7902702786
AS13335	1	0	1.64517926897	2	46,424	1.64222211943
AS14618	1	0	2.65234637417	2	46,424	2.64948262285
AS15169	4	1	1.65053505386	5	46,425	1.64761299671
AS16276	3	0	1.66660210991	4	46,424	1.66384136174
AS16509	1	0	1.82501377501	2	46,424	1.82365723745
AS17204	1	0	2.35903972347	2	46,424	2.37029112712
AS20857	1	0	1.55288727159	2	46,424	1.5491190524
AS20940	1	0	1.62063850426	2	46,424	1.61745247223
AS23033	1	0	2.49950622466	2	46,424	2.53781181581
AS24793	1	0	2.22269584558	2	46,424	2.23335465568
AS24940	1	0	1.61063288987	2	46,424	1.60736045897
AS31615	1	0	2.00108969823	2	46,424	2.00123665849
AS32934	1	0	1.75897836038	2	46,424	1.75704314334
AS34173	1	0	2.21939263956	2	46,424	2.21261707548
AS35470	1	0	2.638705947	2	46,424	2.63578668177
AS36351	1	0	1.4947956937	2	46,424	1.49051099173
AS36647	1	0	2.64950842419	2	46,424	2.646576466
AS36692	1	0	1.90835667691	2	46,424	1.90843932457
AS49544	1	0	1.71134721447	2	46,424	1.70895297877
AS55967	1	0	2.25354329006	2	46,424	2.27636807108
AS60781	2	0	1.82534314497	3	46,424	1.82938888726
AS197902	1	0	1.77303725702	2	46,424	1.77127364731
AS393406	2	0	2.14028867497	3	46,424	2.14204650152
		Mean	2.0003648877		**Mean**	2.004167593

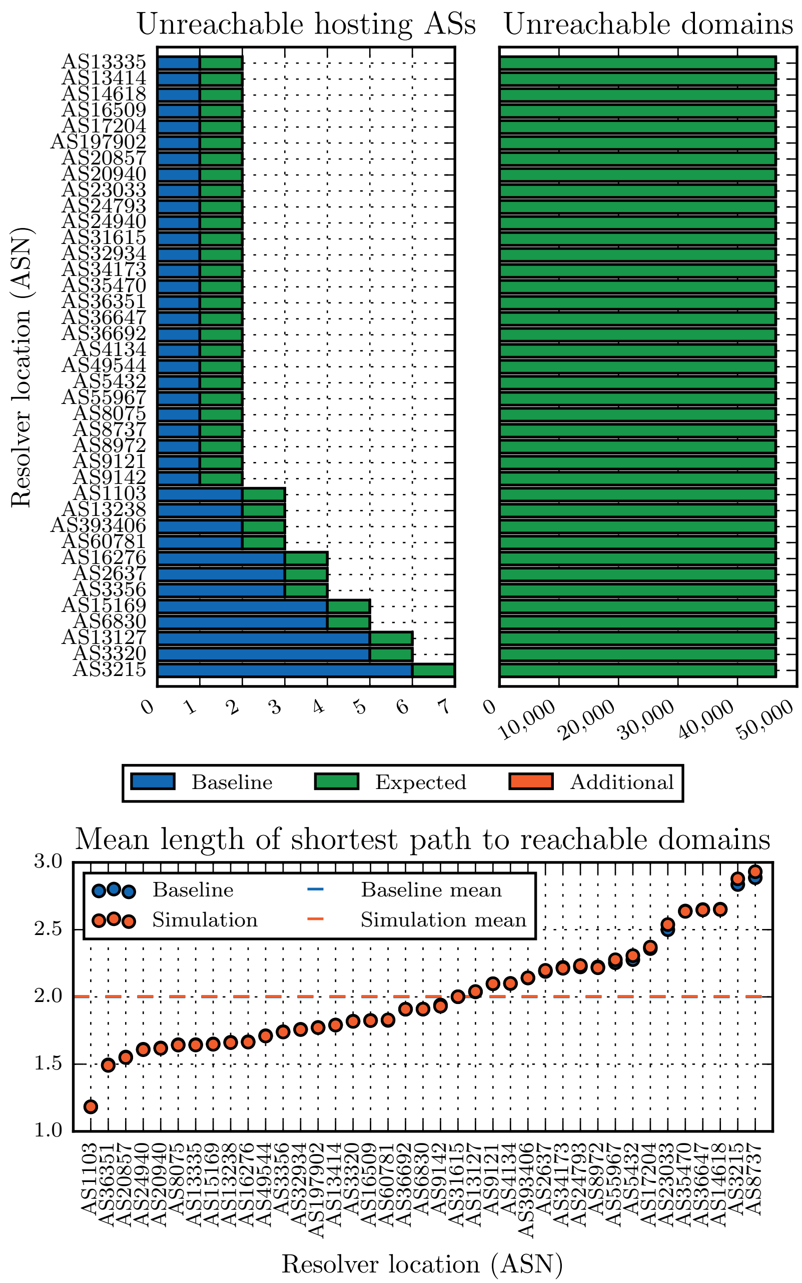


Figure B.10: Plots showing the reachability of autonomous systems and domains in a simulation where AS3265 was removed from the AS topology compared to the baseline.

Table B.11: Reachability of DNS data in a simulation where AS35470 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	42,832	1.17520701035
AS2637	3	0	2.18911427628	4	42,832	2.18258982224
AS3215	6	1	2.83567157466	7	42,833	2.82630086124
AS3320	5	1	1.81932553893	6	42,833	1.80982326983
AS3356	3	0	1.74099517073	4	42,832	1.73891347467
AS4134	1	0	2.09931352367	2	42,832	2.09206464691
AS5432	1	0	2.2768154119	2	42,832	2.26294693154
AS6830	4	1	1.90947226796	5	42,833	1.90069551509
AS8075	1	0	1.64468828964	2	42,832	1.63378051228
AS8737	1	0	2.88639271487	2	42,832	2.87743209452
AS8972	1	0	2.22328412605	2	42,832	2.21703467672
AS9121	1	0	2.09727243649	2	42,832	2.09000713272
AS9142	1	0	1.93983546041	2	42,832	1.93130326519
AS13127	5	1	2.03705757563	6	42,833	2.02125948077
AS13238	2	0	1.66266719207	3	42,832	1.65190411195
AS13414	1	0	1.79106909723	2	42,832	1.78133941731
AS13335	1	0	1.64517926897	2	42,832	1.63427544309
AS14618	1	0	2.65234637417	2	42,832	2.64150023037
AS15169	4	1	1.65053505386	5	42,833	1.6396743302
AS16276	3	0	1.66660210991	4	42,832	1.65587238978
AS16509	1	0	1.82501377501	2	42,832	1.81555728758
AS17204	1	0	2.35903972347	2	42,832	2.34583299824
AS20857	1	0	1.55288727159	2	42,832	1.54124066415
AS20940	1	0	1.62063850426	2	42,832	1.60953717017
AS23033	1	0	2.49950622466	2	42,832	2.48742999754
AS24793	1	0	2.22269584558	2	42,832	2.20839180181
AS24940	1	0	1.61063288987	2	42,832	1.59945102891
AS31615	1	0	2.00108969823	2	42,832	1.99305048745
AS32934	1	0	1.75897836038	2	42,832	1.74899040879
AS34173	1	0	2.21939263956	2	42,832	2.2131128104
AS35470	1	0	2.638705947	n/a	n/a	n/a
AS36351	1	0	1.4947956937	2	42,832	1.48268155543
AS36647	1	0	2.64950842419	2	42,832	2.63863944009
AS36692	1	0	1.90835667691	2	42,832	1.89957094722
AS49544	1	0	1.71134721447	2	42,832	1.70097591938
AS55967	1	0	2.25354329006	2	42,832	2.23948751173
AS60781	2	0	1.82534314497	3	42,832	1.81589099947
AS197902	1	0	1.77303725702	2	42,832	1.76316264171
AS393406	2	0	2.14028867497	3	42,832	2.13337126425
		Mean	2.0003648877		**Mean**	1.9736920934

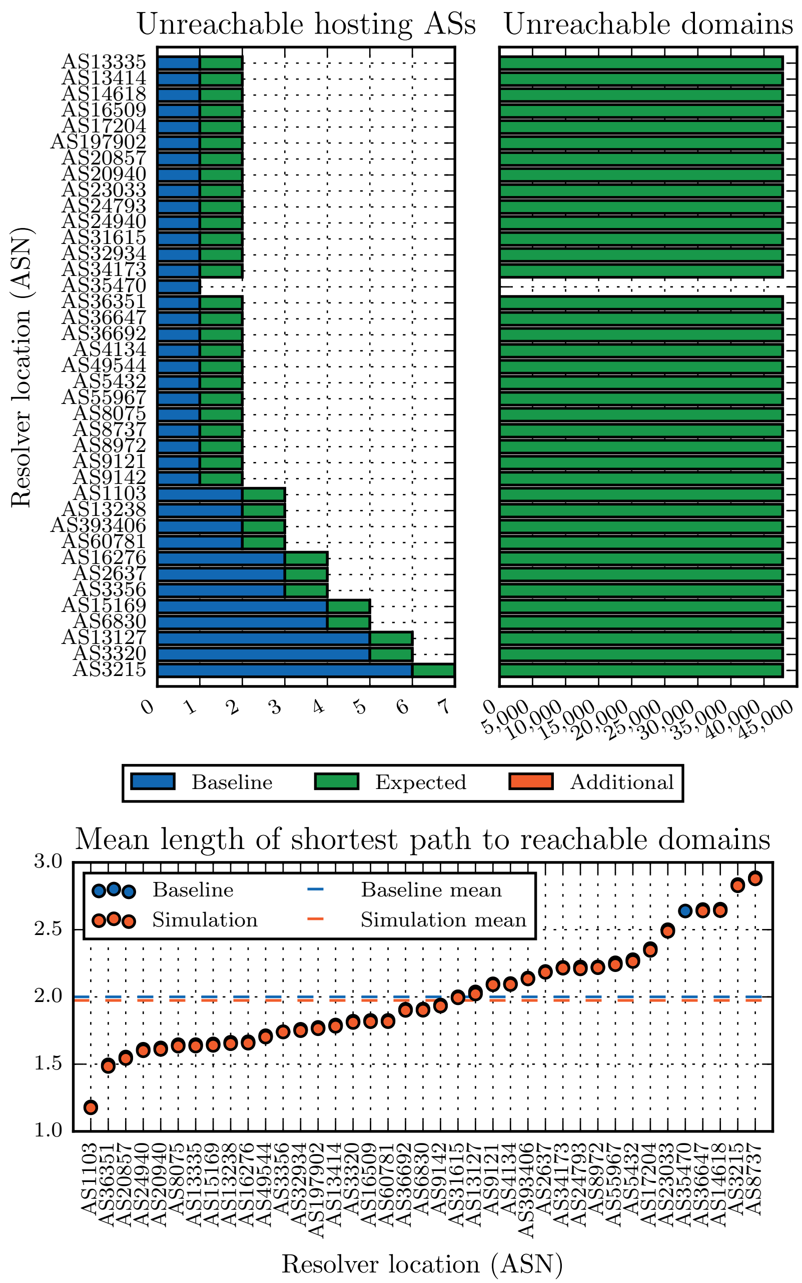


Figure B.11: Plots showing the reachability of autonomous systems and domains in a simulation where AS35470 was removed from the AS topology compared to the baseline.

Table B.12: Reachability of DNS data in a simulation where AS15879 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	54,242	1.18489567739
AS2637	3	0	2.18911427628	4	54,242	2.18083733763
AS3215	6	1	2.83567157466	7	54,243	2.83995465625
AS3320	5	1	1.81932553893	6	54,243	1.81870297681
AS3356	3	0	1.74099517073	4	54,242	1.74312057555
AS4134	1	0	2.09931352367	2	54,242	2.10989171358
AS5432	1	0	2.2768154119	2	54,242	2.26944216282
AS6830	4	1	1.90947226796	5	54,243	1.90918182597
AS8075	1	0	1.64468828964	2	54,242	1.64169258679
AS8737	1	0	2.88639271487	2	54,242	2.89006893077
AS8972	1	0	2.22328412605	2	54,242	2.21535619878
AS9121	1	0	2.09727243649	2	54,242	2.09949617233
AS9142	1	0	1.93983546041	2	54,242	1.94041949735
AS13127	5	1	2.03705757563	6	54,243	2.02782275642
AS13238	2	0	1.66266719207	3	54,242	1.65985512601
AS13414	1	0	1.79106909723	2	54,242	1.77872105806
AS13335	1	0	1.64517926897	2	54,242	1.64224224778
AS14618	1	0	2.65234637417	2	54,242	2.64941853437
AS15169	4	1	1.65053505386	5	54,243	1.64759944601
AS16276	3	0	1.66660210991	4	54,242	1.66389501193
AS16509	1	0	1.82501377501	2	54,242	1.82384956274
AS17204	1	0	2.35903972347	2	54,242	2.35251629493
AS20857	1	0	1.55288727159	2	54,242	1.54894355496
AS20940	1	0	1.62063850426	2	54,242	1.61738604656
AS23033	1	0	2.49950622466	2	54,242	2.4950677388
AS24793	1	0	2.22269584558	2	54,242	2.23061960107
AS24940	1	0	1.61063288987	2	54,242	1.60729009785
AS31615	1	0	2.00108969823	2	54,242	2.00192183629
AS32934	1	0	1.75897836038	2	54,242	1.75713928474
AS34173	1	0	2.21939263956	2	54,242	2.2114253412
AS35470	1	0	2.638705947	2	54,242	2.63865127992
AS36351	1	0	1.4947956937	2	54,242	1.49024205797
AS36647	1	0	2.64950842419	2	54,242	2.64656515545
AS36692	1	0	1.90835667691	2	54,242	1.91202976116
AS49544	1	0	1.71134721447	2	54,242	1.70902201017
AS55967	1	0	2.25354329006	2	54,242	2.24593233916
AS60781	2	0	1.82534314497	3	54,242	1.81335064229
AS197902	1	0	1.77303725702	2	54,242	1.77154590131
AS393406	2	0	2.14028867497	3	54,242	2.13151792678
		Mean	2.0003648877		**Mean**	1.9978879725

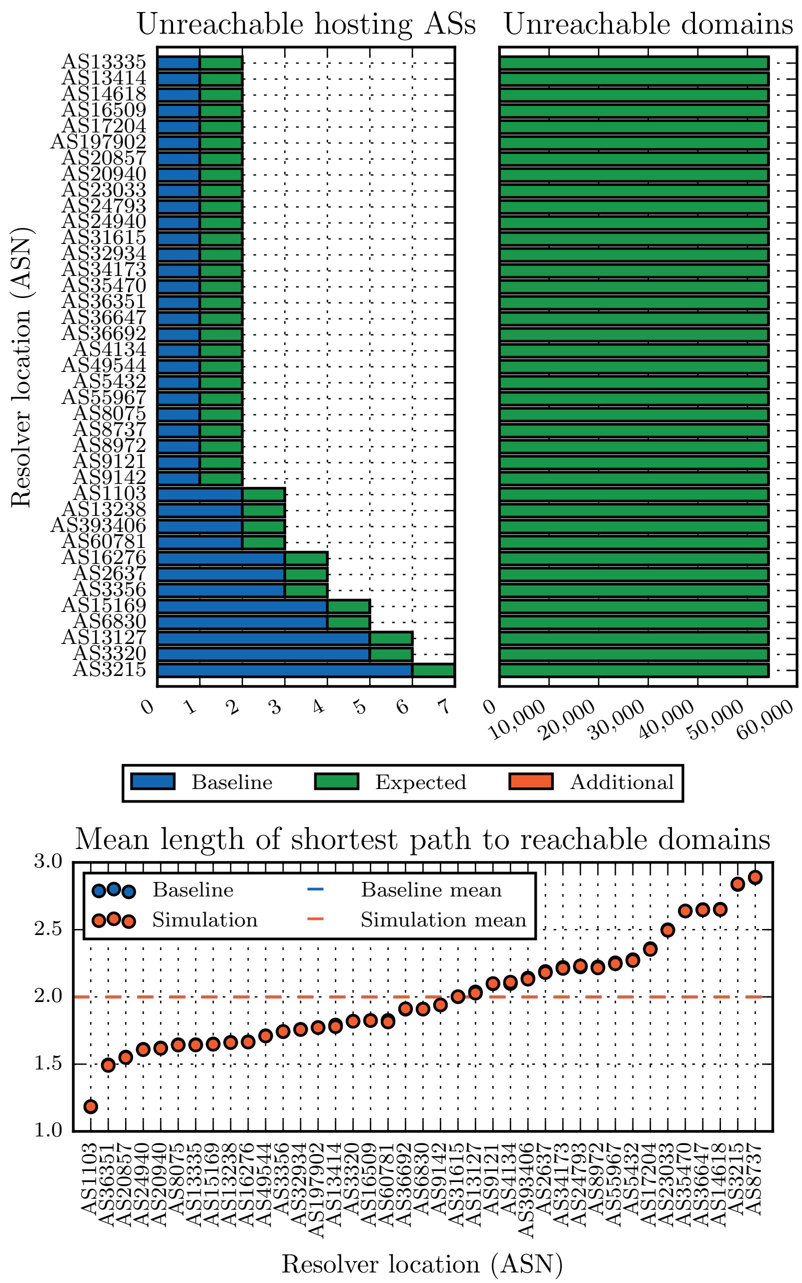


Figure B.12: Plots showing the reachability of autonomous systems and domains in a simulation where AS15879 was removed from the AS topology compared to the baseline.

Table B.13: Reachability of DNS data in a simulation where AS25459 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	48,207	1.19844445142
AS2637	3	0	2.18911427628	4	48,207	2.2110619588
AS3215	6	1	2.83567157466	7	48,208	2.84776736173
AS3320	5	1	1.81932553893	6	48,208	1.8312723593
AS3356	3	0	1.74099517073	4	48,207	1.73898338048
AS4134	1	0	2.09931352367	2	48,207	2.11382653626
AS5432	1	0	2.2768154119	2	48,207	2.27044655202
AS6830	4	1	1.90947226796	5	48,208	1.90890290375
AS8075	1	0	1.64468828964	2	48,207	1.6416241942
AS8737	1	0	2.88639271487	2	48,207	2.88765825255
AS8972	1	0	2.22328412605	2	48,207	2.2428071725
AS9121	1	0	2.09727243649	2	48,207	2.11148461765
AS9142	1	0	1.93983546041	2	48,207	1.93022263744
AS13127	5	1	2.03705757563	6	48,208	2.02856686065
AS13238	2	0	1.66266719207	3	48,207	1.6597663047
AS13414	1	0	1.79106909723	2	48,207	1.78010736599
AS13335	1	0	1.64517926897	2	48,207	1.64211134938
AS14618	1	0	2.65234637417	2	48,207	2.64934344083
AS15169	4	1	1.65053505386	5	48,208	1.64755651189
AS16276	3	0	1.66660210991	4	48,207	1.66377489593
AS16509	1	0	1.82501377501	2	48,207	1.8236370329
AS17204	1	0	2.35903972347	2	48,207	2.35341810837
AS20857	1	0	1.55288727159	2	48,207	1.54906226389
AS20940	1	0	1.62063850426	2	48,207	1.61734806636
AS23033	1	0	2.49950622466	2	48,207	2.49515826054
AS24793	1	0	2.22269584558	2	48,207	2.21585601724
AS24940	1	0	1.61063288987	2	48,207	1.60729254384
AS31615	1	0	2.00108969823	2	48,207	1.99203228541
AS32934	1	0	1.75897836038	2	48,207	1.75694229054
AS34173	1	0	2.21939263956	2	48,207	2.23894754919
AS35470	1	0	2.638705947	2	48,207	2.63560115044
AS36351	1	0	1.4947956937	2	48,207	1.49036401402
AS36647	1	0	2.64950842419	2	48,207	2.64691556472
AS36692	1	0	1.90835667691	2	48,207	1.89845842657
AS49544	1	0	1.71134721447	2	48,207	1.70887925906
AS55967	1	0	2.25354329006	2	48,207	2.24695513903
AS60781	2	0	1.82534314497	3	48,207	1.8407993784
AS197902	1	0	1.77303725702	2	48,207	1.76191202579
AS393406	2	0	2.14028867497	3	48,207	2.13253592864
		Mean	2.0003648877		**Mean**	2.000457549

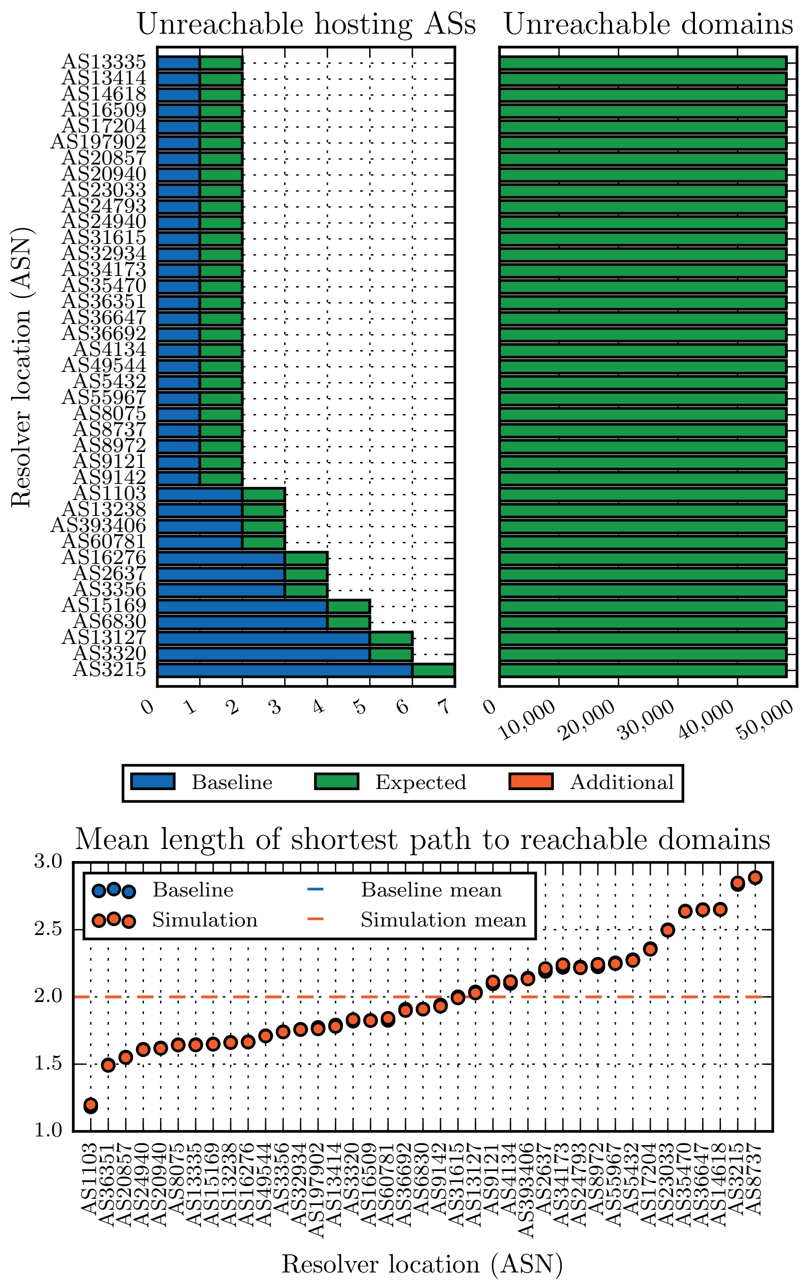


Figure B.13: Plots showing the reachability of autonomous systems and domains in a simulation where AS25459 was removed from the AS topology compared to the baseline.

Table B.14: Reachability of DNS data in a simulation where AS6724 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	5	165,009	1.18769451548
AS2637	3	0	2.18911427628	6	165,009	2.19543061349
AS3215	6	1	2.83567157466	9	165,010	2.83049834051
AS3320	5	1	1.81932553893	8	165,010	1.8465642187
AS3356	3	0	1.74099517073	6	165,009	1.7327763353
AS4134	1	0	2.09931352367	4	165,009	2.10250685654
AS5432	1	0	2.2768154119	3	165,008	2.28609614253
AS6830	4	1	1.90947226796	7	165,010	1.90660197416
AS8075	1	0	1.64468828964	3	165,008	1.66599952306
AS8737	1	0	2.88639271487	3	165,008	2.85105542927
AS8972	1	0	2.22328412605	3	165,008	2.23075553196
AS9121	1	0	2.09727243649	3	165,008	2.10040020924
AS9142	1	0	1.93983546041	3	165,008	1.93811795884
AS13127	5	1	2.03705757563	8	165,010	2.03840452419
AS13238	2	0	1.66266719207	5	165,009	1.6846019802
AS13414	1	0	1.79106909723	3	165,008	1.81711168549
AS13335	1	0	1.64517926897	3	165,008	1.66649531326
AS14618	1	0	2.65234637417	3	165,008	2.64131674802
AS15169	4	1	1.65053505386	7	165,010	1.67194618693
AS16276	3	0	1.66660210991	6	165,009	1.68863561317
AS16509	1	0	1.82501377501	3	165,008	1.81946351576
AS17204	1	0	2.35903972347	3	165,008	2.33870048348
AS20857	1	0	1.55288727159	3	165,008	1.53874336991
AS20940	1	0	1.62063850426	3	165,008	1.64109327702
AS23033	1	0	2.49950622466	3	165,008	2.5161762613
AS24793	1	0	2.22269584558	3	165,008	2.19803357065
AS24940	1	0	1.61063288987	4	165,009	1.63091700628
AS31615	1	0	2.00108969823	3	165,008	2.00116601087
AS32934	1	0	1.75897836038	3	165,008	1.75133255638
AS34173	1	0	2.21939263956	3	165,008	2.2267720942
AS35470	1	0	2.638705947	3	165,008	2.62728250041
AS36351	1	0	1.4947956937	3	165,008	1.51110470058
AS36647	1	0	2.64950842419	3	165,008	2.67088934532
AS36692	1	0	1.90835667691	3	165,008	1.87371484947
AS49544	1	0	1.71134721447	3	165,008	1.73477550973
AS55967	1	0	2.25354329006	3	165,008	2.26205147141
AS60781	2	0	1.82534314497	5	165,009	1.82007523781
AS197902	1	0	1.77303725702	3	165,008	1.765967214
AS393406	2	0	2.14028867497	5	165,009	2.14497135359
		Mean	2.0003648877		**Mean**	2.0040061546

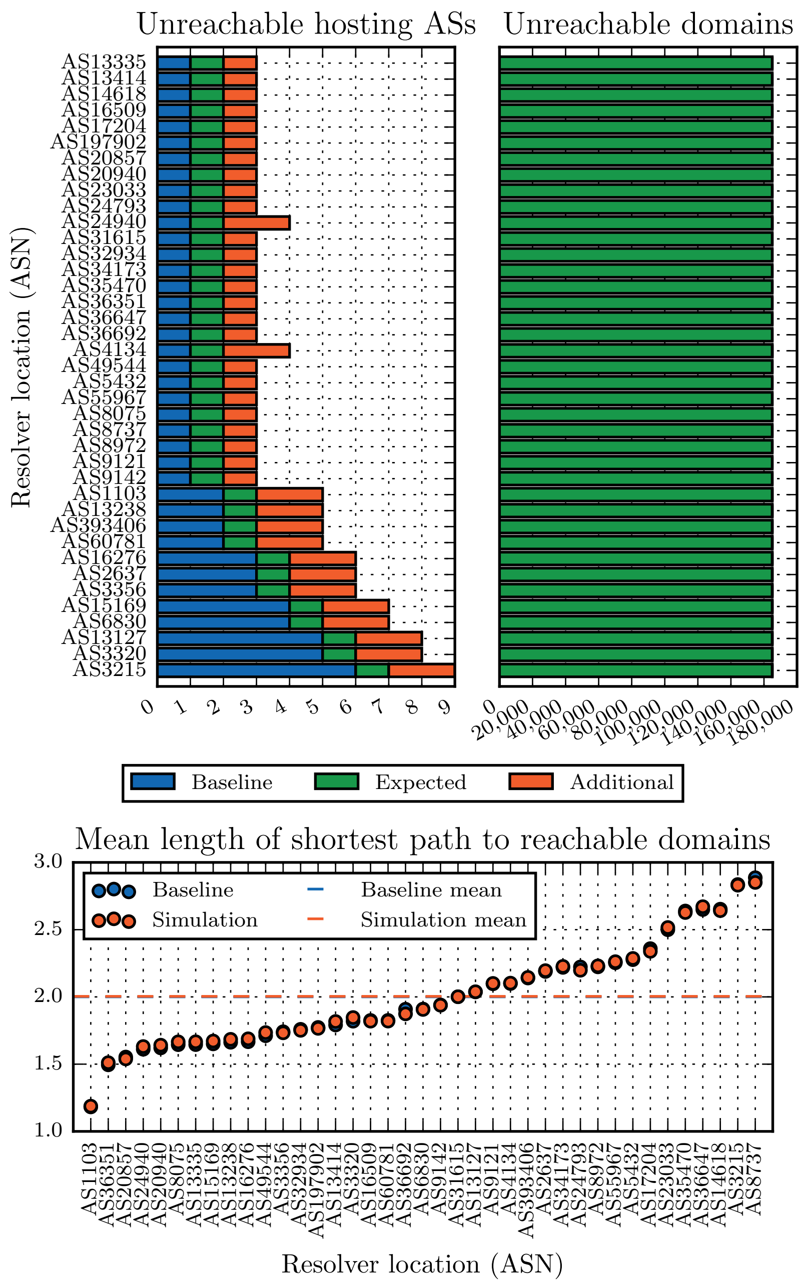


Figure B.14: Plots showing the reachability of autonomous systems and domains in a simulation where AS6724 was removed from the AS topology compared to the baseline.

Table B.15: Reachability of DNS data in a simulation where AS49544 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	7	17,892	1.1860810459
AS2637	3	0	2.18911427628	8	17,892	2.19533800545
AS3215	6	1	2.83567157466	11	17,893	2.8386358812
AS3320	5	1	1.81932553893	10	17,893	1.84979207559
AS3356	3	0	1.74099517073	8	17,892	1.7667699166
AS4134	1	0	2.09931352367	6	17,892	2.10320698214
AS5432	1	0	2.2768154119	6	17,892	2.28492100838
AS6830	4	1	1.90947226796	9	17,893	1.91251819982
AS8075	1	0	1.64468828964	6	17,892	1.66884244616
AS8737	1	0	2.88639271487	6	17,892	2.89307422475
AS8972	1	0	2.22328412605	6	17,892	2.23994930891
AS9121	1	0	2.09727243649	6	17,892	2.10135787193
AS9142	1	0	1.93983546041	6	17,892	1.94362486198
AS13127	5	1	2.03705757563	10	17,893	2.04669626017
AS13238	2	0	1.66266719207	7	17,892	1.68701354954
AS13414	1	0	1.79106909723	6	17,892	1.81909747263
AS13335	1	0	1.64517926897	6	17,892	1.66968798346
AS14618	1	0	2.65234637417	6	17,892	2.65462803092
AS15169	4	1	1.65053505386	9	17,893	1.67755136392
AS16276	3	0	1.66660210991	8	17,892	1.69123525126
AS16509	1	0	1.82501377501	6	17,892	1.82787415353
AS17204	1	0	2.35903972347	6	17,892	2.37122135908
AS20857	1	0	1.55288727159	6	17,892	1.55481965611
AS20940	1	0	1.62063850426	6	17,892	1.64437778479
AS23033	1	0	2.49950622466	6	17,892	2.50117376512
AS24793	1	0	2.22269584558	6	17,892	2.23444387173
AS24940	1	0	1.61063288987	6	17,892	1.63475107801
AS31615	1	0	2.00108969823	6	17,892	2.00472086983
AS32934	1	0	1.75897836038	6	17,892	1.78940080376
AS34173	1	0	2.21939263956	6	17,892	2.23599299481
AS35470	1	0	2.638705947	6	17,892	2.63866662078
AS36351	1	0	1.4947956937	6	17,892	1.50652284989
AS36647	1	0	2.64950842419	6	17,892	2.67368432077
AS36692	1	0	1.90835667691	6	17,892	1.91340545992
AS49544	1	0	1.71134721447	n/a	n/a	n/a
AS55967	1	0	2.25354329006	6	17,892	2.26176701398
AS60781	2	0	1.82534314497	7	17,892	1.82904417815
AS197902	1	0	1.77303725702	6	17,892	1.77580974083
AS393406	2	0	2.14028867497	7	17,892	2.14685305269
		Mean	2.0003648877		**Mean**	2.0203829293

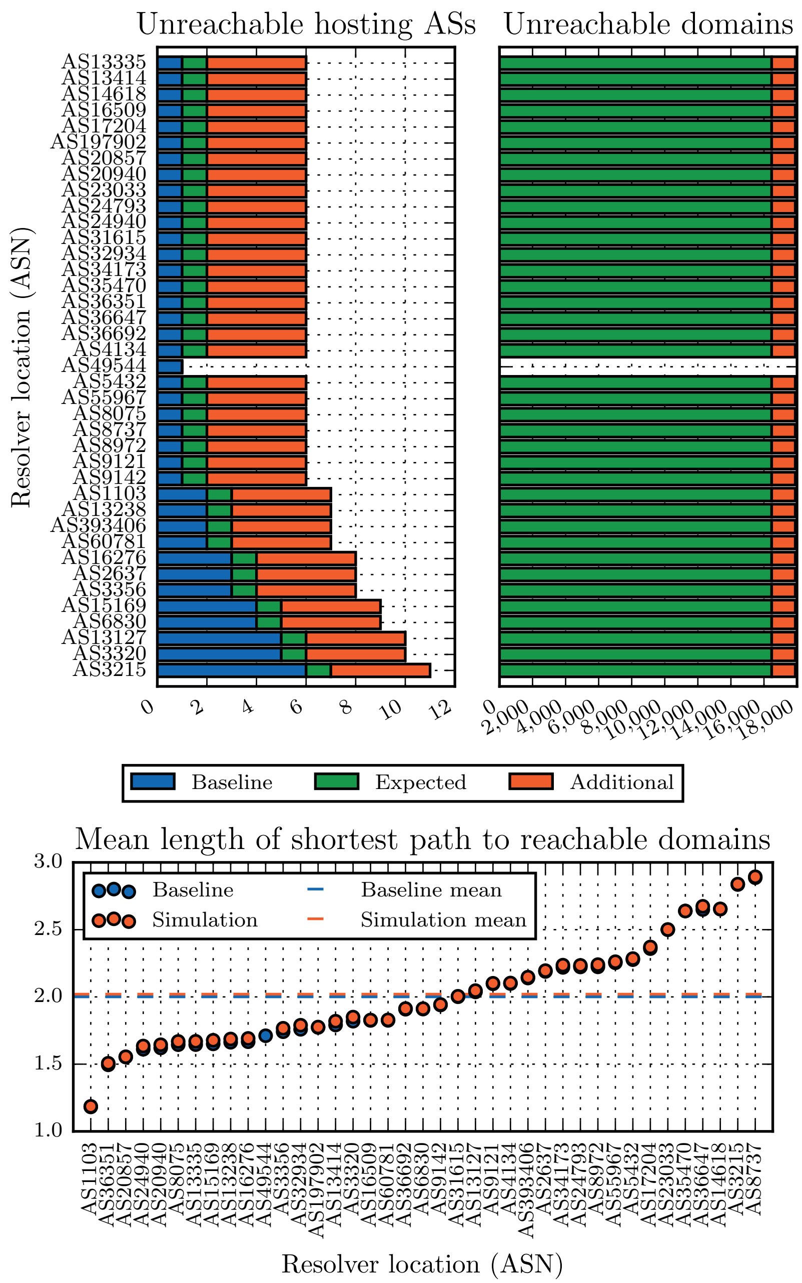


Figure B.15: Plots showing the reachability of autonomous systems and domains in a simulation where AS49544 was removed from the AS topology compared to the baseline.

Table B.16: Reachability of DNS data in a simulation where AS34233 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	72,131	1.1706409465
AS2637	3	0	2.18911427628	4	72,131	2.1780631165
AS3215	6	1	2.83567157466	7	72,132	2.81980351642
AS3320	5	1	1.81932553893	6	72,132	1.80323470862
AS3356	3	0	1.74099517073	4	72,131	1.73746551118
AS4134	1	0	2.09931352367	2	72,131	2.08703851393
AS5432	1	0	2.2768154119	2	72,131	2.26696043216
AS6830	4	1	1.90947226796	5	72,132	1.9082398327
AS8075	1	0	1.64468828964	2	72,131	1.63984724497
AS8737	1	0	2.88639271487	2	72,131	2.88484762946
AS8972	1	0	2.22328412605	2	72,131	2.21269865022
AS9121	1	0	2.09727243649	2	72,131	2.08496960978
AS9142	1	0	1.93983546041	2	72,131	1.95217751689
AS13127	5	1	2.03705757563	6	72,132	2.05080662715
AS13238	2	0	1.66266719207	3	72,131	1.65807136189
AS13414	1	0	1.79106909723	2	72,131	1.801382557
AS13335	1	0	1.64517926897	2	72,131	1.64034491561
AS14618	1	0	2.65234637417	2	72,131	2.64760969774
AS15169	4	1	1.65053505386	5	72,132	1.64577387988
AS16276	3	0	1.66660210991	4	72,131	1.66205971783
AS16509	1	0	1.82501377501	2	72,131	1.82263029703
AS17204	1	0	2.35903972347	2	72,131	2.35030533813
AS20857	1	0	1.55288727159	2	72,131	1.54679568315
AS20940	1	0	1.62063850426	2	72,131	1.61546969698
AS23033	1	0	2.49950622466	2	72,131	2.49268656556
AS24793	1	0	2.22269584558	2	72,131	2.24013874317
AS24940	1	0	1.61063288987	2	72,131	1.60532790997
AS31615	1	0	2.00108969823	2	72,131	2.01426618048
AS32934	1	0	1.75897836038	2	72,131	1.75569491845
AS34173	1	0	2.21939263956	2	72,131	2.2087541286
AS35470	1	0	2.638705947	2	72,131	2.63378242724
AS36351	1	0	1.4947956937	2	72,131	1.48791164816
AS36647	1	0	2.64950842419	2	72,131	2.64473325968
AS36692	1	0	1.90835667691	2	72,131	1.92143038931
AS49544	1	0	1.71134721447	2	72,131	1.70741482019
AS55967	1	0	2.25354329006	2	72,131	2.2433711461
AS60781	2	0	1.82534314497	3	72,131	1.80933432867
AS197902	1	0	1.77303725702	2	72,131	1.78310591448
AS393406	2	0	2.14028867497	3	72,131	2.12857209526
		Mean	2.0003648877		**Mean**	1.9965067045

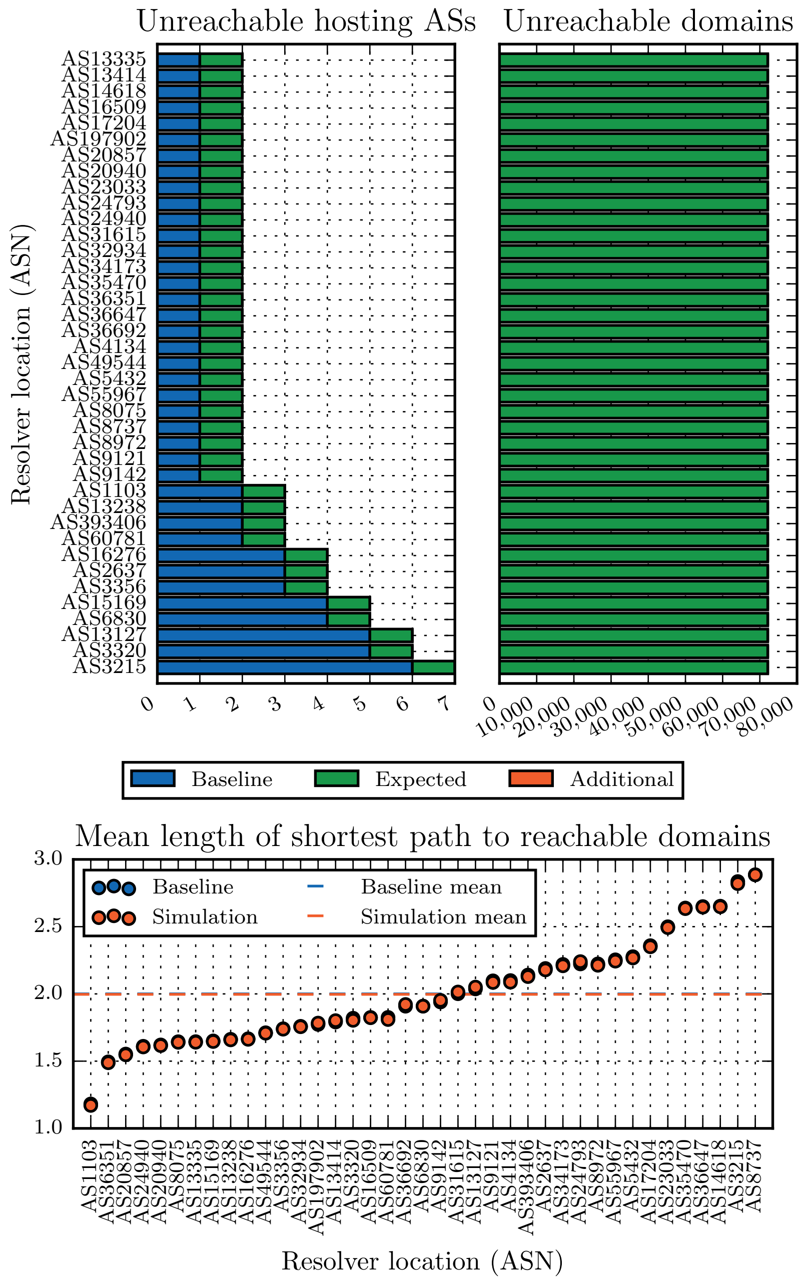


Figure B.16: Plots showing the reachability of autonomous systems and domains in a simulation where AS34233 was removed from the AS topology compared to the baseline.

Table B.17: Reachability of DNS data in a simulation where AS61387 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	57,693	1.17297994477
AS2637	3	0	2.18911427628	4	57,693	2.18029920324
AS3215	6	1	2.83567157466	7	57,694	2.82301425224
AS3320	5	1	1.81932553893	6	57,694	1.80649052005
AS3356	3	0	1.74099517073	4	57,693	1.73817992706
AS4134	1	0	2.09931352367	2	57,693	2.08952223392
AS5432	1	0	2.2768154119	2	57,693	2.26895561508
AS6830	4	1	1.90947226796	5	57,694	1.89761722705
AS8075	1	0	1.64468828964	2	57,693	1.62995480578
AS8737	1	0	2.88639271487	2	57,693	2.87428866451
AS8972	1	0	2.22328412605	2	57,693	2.21484051067
AS9121	1	0	2.09727243649	2	57,693	2.08745895824
AS9142	1	0	1.93983546041	2	57,693	1.92831049755
AS13127	5	1	2.03705757563	6	57,694	2.02659101195
AS13238	2	0	1.66266719207	3	57,693	1.6481291554
AS13414	1	0	1.79106909723	2	57,693	1.77792690728
AS13335	1	0	1.64517926897	2	57,693	1.63045112251
AS14618	1	0	2.65234637417	2	57,693	2.63769614073
AS15169	4	1	1.65053505386	5	57,694	1.63586512694
AS16276	3	0	1.66660210991	4	57,693	1.65210684942
AS16509	1	0	1.82501377501	2	57,693	1.8122405949
AS17204	1	0	2.35903972347	2	57,693	2.35207472261
AS20857	1	0	1.55288727159	2	57,693	1.5371558263
AS20940	1	0	1.62063850426	2	57,693	1.60564357714
AS23033	1	0	2.49950622466	2	57,693	2.48319447834
AS24793	1	0	2.22269584558	2	57,693	2.2142496036
AS24940	1	0	1.61063288987	2	57,693	1.59552919252
AS31615	1	0	2.00108969823	2	57,693	1.99023137894
AS32934	1	0	1.75897836038	2	57,693	1.74548731462
AS34173	1	0	2.21939263956	2	57,693	2.21090690858
AS35470	1	0	2.638705947	2	57,693	2.62390742958
AS36351	1	0	1.4947956937	2	57,693	1.48930591972
AS36647	1	0	2.64950842419	2	57,693	2.63482733963
AS36692	1	0	1.90835667691	2	57,693	1.89648951074
AS49544	1	0	1.71134721447	2	57,693	1.69733837438
AS55967	1	0	2.25354329006	2	57,693	2.24542975017
AS60781	2	0	1.82534314497	3	57,693	1.81257354541
AS197902	1	0	1.77303725702	2	57,693	1.7596997981
AS393406	2	0	2.14028867497	3	57,693	2.13094282277
		Mean	2.0003648877		**Mean**	1.988664276

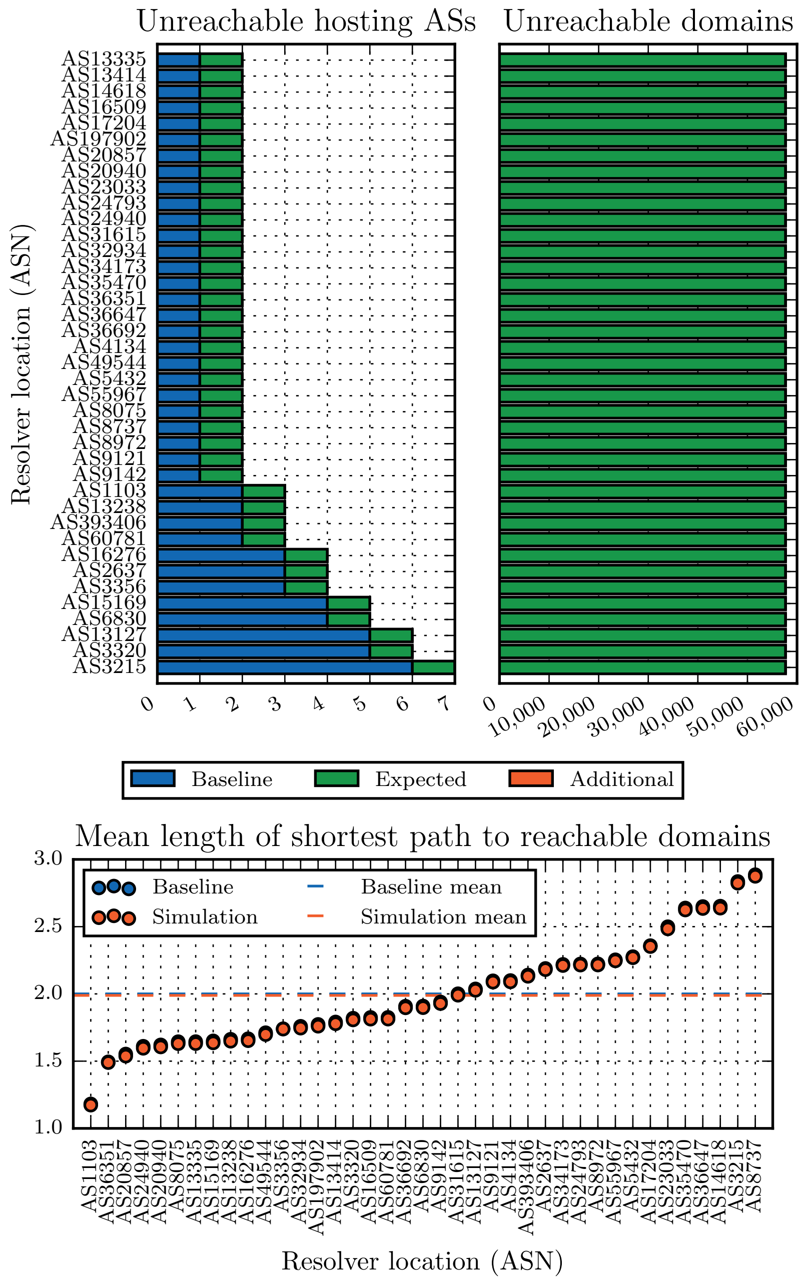


Figure B.17: Plots showing the reachability of autonomous systems and domains in a simulation where AS61387 was removed from the AS topology compared to the baseline.

Table B.18: Reachability of DNS data in a simulation where AS8315 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	2,976	1.18189764953
AS2637	3	0	2.18911427628	4	2,976	2.18937683753
AS3215	6	1	2.83567157466	7	2,977	2.83558390999
AS3320	5	1	1.81932553893	6	2,977	1.81922880161
AS3356	3	0	1.74099517073	4	2,976	1.7620660702
AS4134	1	0	2.09931352367	2	2,976	2.09937218985
AS5432	1	0	2.2768154119	2	2,976	2.27641718882
AS6830	4	1	1.90947226796	5	2,977	1.90942556535
AS8075	1	0	1.64468828964	2	2,976	1.64449462234
AS8737	1	0	2.88639271487	2	2,976	2.88633432131
AS8972	1	0	2.22328412605	2	2,976	2.24281642102
AS9121	1	0	2.09727243649	2	2,976	2.09734004102
AS9142	1	0	1.93983546041	2	2,976	1.95917648735
AS13127	5	1	2.03705757563	6	2,977	2.03710574655
AS13238	2	0	1.66266719207	3	2,976	1.66248369022
AS13414	1	0	1.79106909723	2	2,976	1.79097700553
AS13335	1	0	1.64517926897	2	2,976	1.6449961319
AS14618	1	0	2.65234637417	2	2,976	2.65215695739
AS15169	4	1	1.65053505386	5	2,977	1.6503448182
AS16276	3	0	1.66660210991	4	2,976	1.6664308633
AS16509	1	0	1.82501377501	2	2,976	1.82492019489
AS17204	1	0	2.35903972347	2	2,976	2.36033023911
AS20857	1	0	1.55288727159	2	2,976	1.55264283791
AS20940	1	0	1.62063850426	2	2,976	1.62043148846
AS23033	1	0	2.49950622466	2	2,976	2.49923160305
AS24793	1	0	2.22269584558	2	2,976	2.22293116581
AS24940	1	0	1.61063288987	2	2,976	1.6104205071
AS31615	1	0	2.00108969823	2	2,976	2.00109459265
AS32934	1	0	1.75897836038	2	2,976	1.75884812821
AS34173	1	0	2.21939263956	2	2,976	2.23896455154
AS35470	1	0	2.638705947	2	2,976	2.63850597522
AS36351	1	0	1.4947956937	2	2,976	1.49451883057
AS36647	1	0	2.64950842419	2	2,976	2.64931761875
AS36692	1	0	1.90835667691	2	2,976	1.90831942634
AS49544	1	0	1.71134721447	2	2,976	1.71119054529
AS55967	1	0	2.25354329006	2	2,976	2.2531321501
AS60781	2	0	1.82534314497	3	2,976	1.82531185353
AS197902	1	0	1.77303725702	2	2,976	1.77291594707
AS393406	2	0	2.14028867497	3	2,976	2.14141170933
		Mean	2.0003648877		**Mean**	2.0023708893

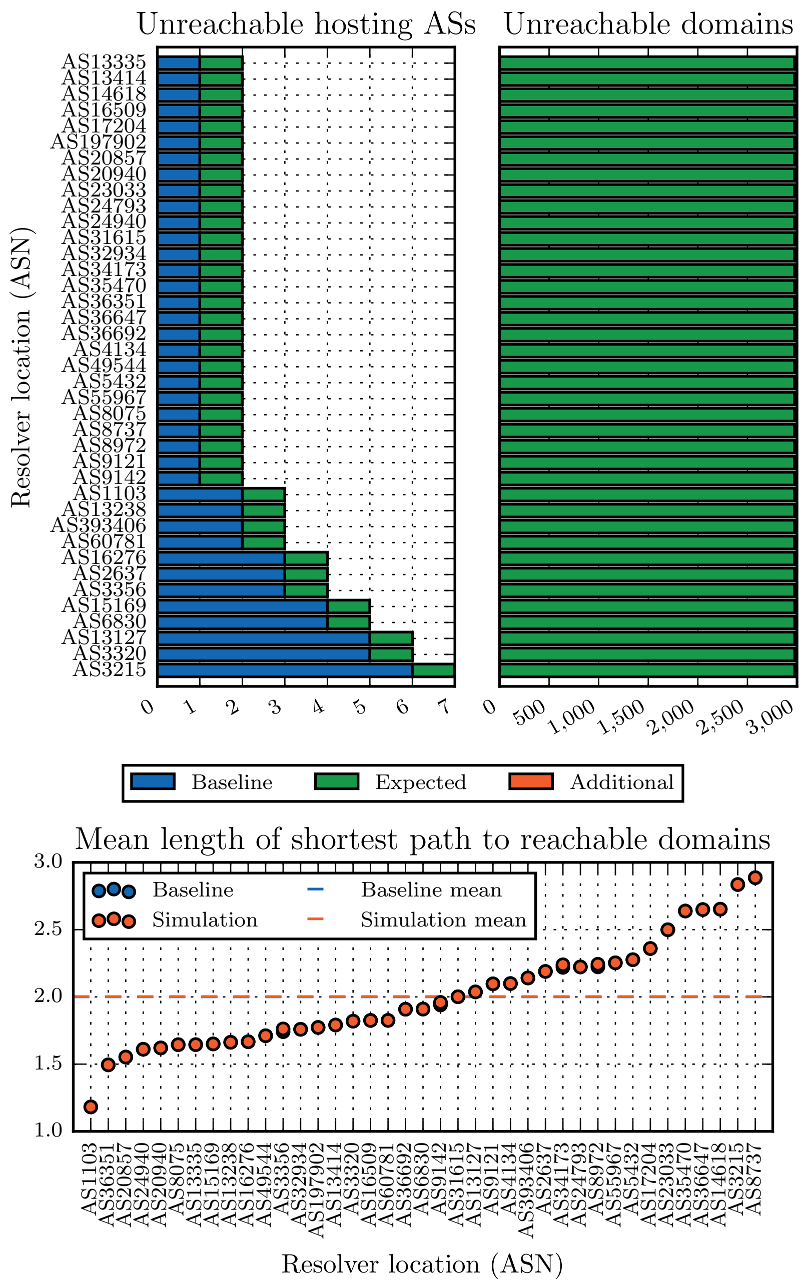


Figure B.18: Plots showing the reachability of autonomous systems and domains in a simulation where AS8315 was removed from the AS topology compared to the baseline.

Table B.19: Reachability of DNS data in a simulation where AS50673 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	7	23,267	1.19366936871
AS2637	3	0	2.18911427628	9	23,267	2.20106164518
AS3215	6	1	2.83567157466	12	23,268	2.86720427893
AS3320	5	1	1.81932553893	11	23,268	1.82904585212
AS3356	3	0	1.74099517073	9	23,267	1.73980201519
AS4134	1	0	2.09931352367	7	23,267	2.11051065792
AS5432	1	0	2.2768154119	4	22,997	2.29088446178
AS6830	4	1	1.90947226796	10	23,268	1.9198469724
AS8075	1	0	1.64468828964	4	22,997	1.67113202295
AS8737	1	0	2.88639271487	4	22,997	2.91888975065
AS8972	1	0	2.22328412605	4	22,997	2.23542160298
AS9121	1	0	2.09727243649	4	22,997	2.10851079722
AS9142	1	0	1.93983546041	4	22,997	1.95053849917
AS13127	5	1	2.03705757563	11	23,268	2.04948610133
AS13238	2	0	1.66266719207	8	23,267	1.68940625713
AS13414	1	0	1.79106909723	4	22,997	1.81842588001
AS13335	1	0	1.64517926897	4	22,997	1.67163428895
AS14618	1	0	2.65234637417	4	22,997	2.66160731485
AS15169	4	1	1.65053505386	10	23,268	1.67700560889
AS16276	3	0	1.66660210991	9	23,267	1.69324336645
AS16509	1	0	1.82501377501	4	22,997	1.83501806791
AS17204	1	0	2.35903972347	4	22,997	2.35768658864
AS20857	1	0	1.55288727159	4	22,997	1.56192314525
AS20940	1	0	1.62063850426	4	22,997	1.64646932087
AS23033	1	0	2.49950622466	4	22,997	2.52861727462
AS24793	1	0	2.22269584558	4	22,997	2.23718374605
AS24940	1	0	1.61063288987	7	23,267	1.63720764179
AS31615	1	0	2.00108969823	4	22,997	2.01210024877
AS32934	1	0	1.75897836038	4	22,997	1.76843197347
AS34173	1	0	2.21939263956	4	22,997	2.23161332969
AS35470	1	0	2.638705947	4	22,997	2.64670201436
AS36351	1	0	1.4947956937	4	22,997	1.51077531862
AS36647	1	0	2.64950842419	4	22,997	2.67619830128
AS36692	1	0	1.90835667691	4	22,997	1.91875964447
AS49544	1	0	1.71134721447	4	22,997	1.73787742725
AS55967	1	0	2.25354329006	4	22,997	2.26725792904
AS60781	2	0	1.82534314497	8	23,267	1.83565973063
AS197902	1	0	1.77303725702	4	22,997	1.78301434856
AS393406	2	0	2.14028867497	8	23,267	2.15208121432
		Mean	2.0003648877		**Mean**	2.0164590771

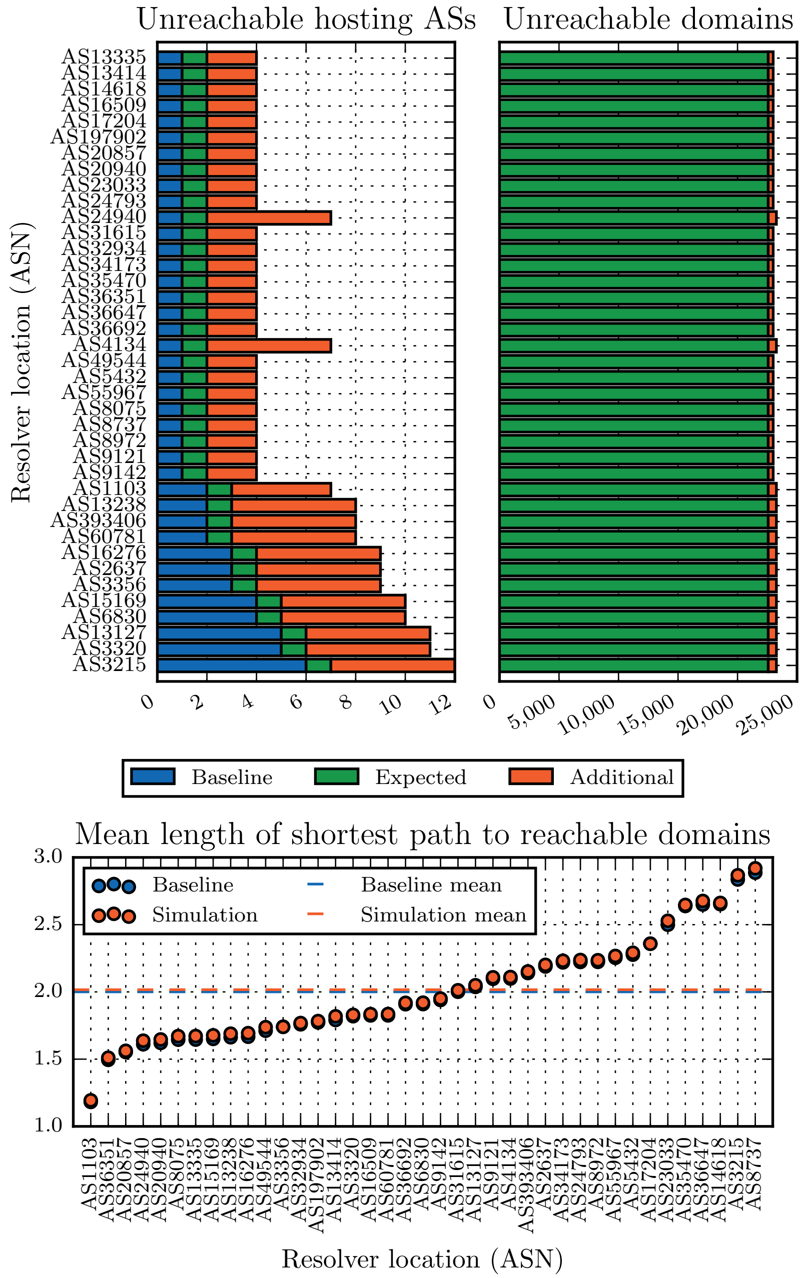


Figure B.19: Plots showing the reachability of autonomous systems and domains in a simulation where AS50673 was removed from the AS topology compared to the baseline.

Table B.20: Reachability of DNS data in a simulation where AS25525 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	8,399	1.18050798775
AS2637	3	0	2.18911427628	5	8,399	2.19193900966
AS3215	6	1	2.83567157466	8	8,400	2.83632010975
AS3320	5	1	1.81932553893	7	8,400	1.81994844287
AS3356	3	0	1.74099517073	5	8,399	1.76107467176
AS4134	1	0	2.09931352367	3	8,399	2.10037545817
AS5432	1	0	2.2768154119	3	8,399	2.27589874447
AS6830	4	1	1.90947226796	6	8,400	1.90954856146
AS8075	1	0	1.64468828964	3	8,399	1.64434920615
AS8737	1	0	2.88639271487	3	8,399	2.88464411379
AS8972	1	0	2.22328412605	3	8,399	2.22468383084
AS9121	1	0	2.09727243649	3	8,399	2.09834629262
AS9142	1	0	1.93983546041	3	8,399	1.93817065938
AS13127	5	1	2.03705757563	7	8,400	2.03572202014
AS13238	2	0	1.66266719207	4	8,399	1.66235779366
AS13414	1	0	1.79106909723	3	8,399	1.79112850094
AS13335	1	0	1.64517926897	3	8,399	1.64486690567
AS14618	1	0	2.65234637417	3	8,399	2.65201929882
AS15169	4	1	1.65053505386	6	8,400	1.65020513824
AS16276	3	0	1.66660210991	5	8,399	1.66632333835
AS16509	1	0	1.82501377501	3	8,399	1.82495744801
AS17204	1	0	2.35903972347	3	8,399	2.36112407071
AS20857	1	0	1.55288727159	3	8,399	1.55240592123
AS20940	1	0	1.62063850426	3	8,399	1.62026170989
AS23033	1	0	2.49950622466	3	8,399	2.49895031896
AS24793	1	0	2.22269584558	3	8,399	2.22165137745
AS24940	1	0	1.61063288987	3	8,399	1.61024189991
AS31615	1	0	2.00108969823	3	8,399	1.99952094592
AS32934	1	0	1.75897836038	3	8,399	1.7588184876
AS34173	1	0	2.21939263956	3	8,399	2.22079912419
AS35470	1	0	2.638705947	3	8,399	2.63819431337
AS36351	1	0	1.4947956937	3	8,399	1.49417639384
AS36647	1	0	2.64950842419	3	8,399	2.64917689884
AS36692	1	0	1.90835667691	3	8,399	1.90848573545
AS49544	1	0	1.71134721447	3	8,399	1.71111433468
AS55967	1	0	2.25354329006	3	8,399	2.25521970118
AS60781	2	0	1.82534314497	4	8,399	1.82611214383
AS197902	1	0	1.77303725702	3	8,399	1.77111091073
AS393406	2	0	2.14028867497	4	8,399	2.14181344932
		Mean	2.0003648877		**Mean**	2.0008350069

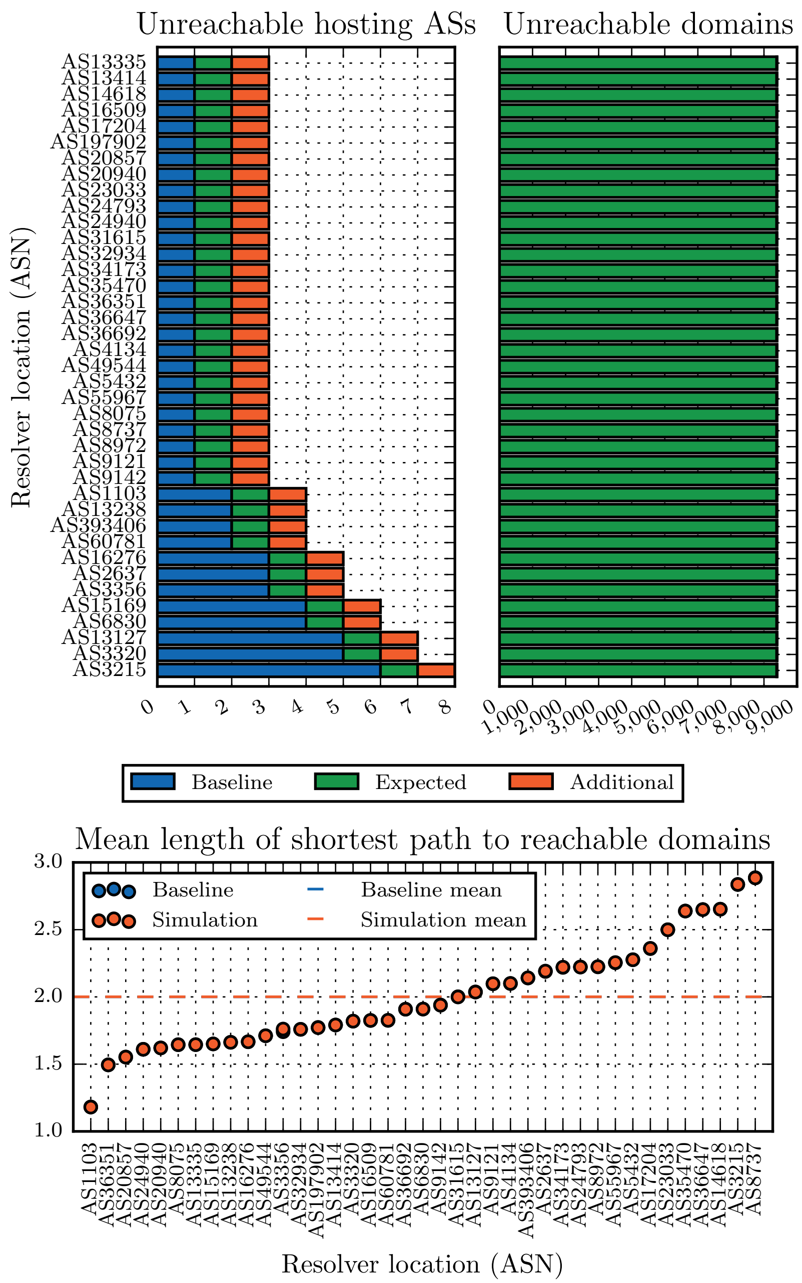


Figure B.20: Plots showing the reachability of autonomous systems and domains in a simulation where AS25525 was removed from the AS topology compared to the baseline.

Table B.21: Reachability of DNS data in a simulation where AS174 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	17	118	1.18512695096
AS2637	3	0	2.18911427628	18	118	2.50944997549
AS3215	6	1	2.83567157466	21	119	2.85193774304
AS3320	5	1	1.81932553893	20	119	1.83263030767
AS3356	3	0	1.74099517073	18	118	1.74526224353
AS4134	1	0	2.09931352367	16	118	2.11743555522
AS5432	1	0	2.2768154119	12	116	2.28124944079
AS6830	4	1	1.90947226796	17	119	1.91405452974
AS8075	1	0	1.64468828964	12	116	1.64503309802
AS8737	1	0	2.88639271487	12	116	2.88656175065
AS8972	1	0	2.22328412605	12	116	2.58340435352
AS9121	1	0	2.09727243649	12	116	2.12809413138
AS9142	1	0	1.93983546041	12	116	1.93997153973
AS13127	5	1	2.03705757563	18	119	2.03721851246
AS13238	2	0	1.66266719207	17	118	1.66502711257
AS13414	1	0	1.79106909723	12	116	1.79112273779
AS13335	1	0	1.64517926897	12	116	1.65531219802
AS14618	1	0	2.65234637417	12	116	2.65233811126
AS15169	4	1	1.65053505386	17	119	1.65052718071
AS16276	3	0	1.66660210991	16	118	1.68281254951
AS16509	1	0	1.82501377501	12	116	1.82500924567
AS17204	1	0	2.35903972347	12	116	2.36339108896
AS20857	1	0	1.55288727159	12	116	1.57005442271
AS20940	1	0	1.62063850426	12	116	1.62136100772
AS23033	1	0	2.49950622466	12	116	2.50231533261
AS24793	1	0	2.22269584558	12	116	2.24281093047
AS24940	1	0	1.61063288987	14	118	1.62791914507
AS31615	1	0	2.00108969823	12	116	2.00129588538
AS32934	1	0	1.75897836038	12	116	1.75943002666
AS34173	1	0	2.21939263956	12	116	2.71194231446
AS35470	1	0	2.638705947	12	116	2.70783041349
AS36351	1	0	1.4947956937	12	116	1.50426232955
AS36647	1	0	2.64950842419	12	116	2.64950252317
AS36692	1	0	1.90835667691	12	116	1.90836830285
AS49544	1	0	1.71134721447	12	116	1.71270992896
AS55967	1	0	2.25354329006	12	116	2.25364514364
AS60781	2	0	1.82534314497	15	118	1.91566154116
AS197902	1	0	1.77303725702	12	116	1.77522633257
AS393406	2	0	2.14028867497	15	118	2.1403612897
		Mean	2.0003648877		**Mean**	2.039683775

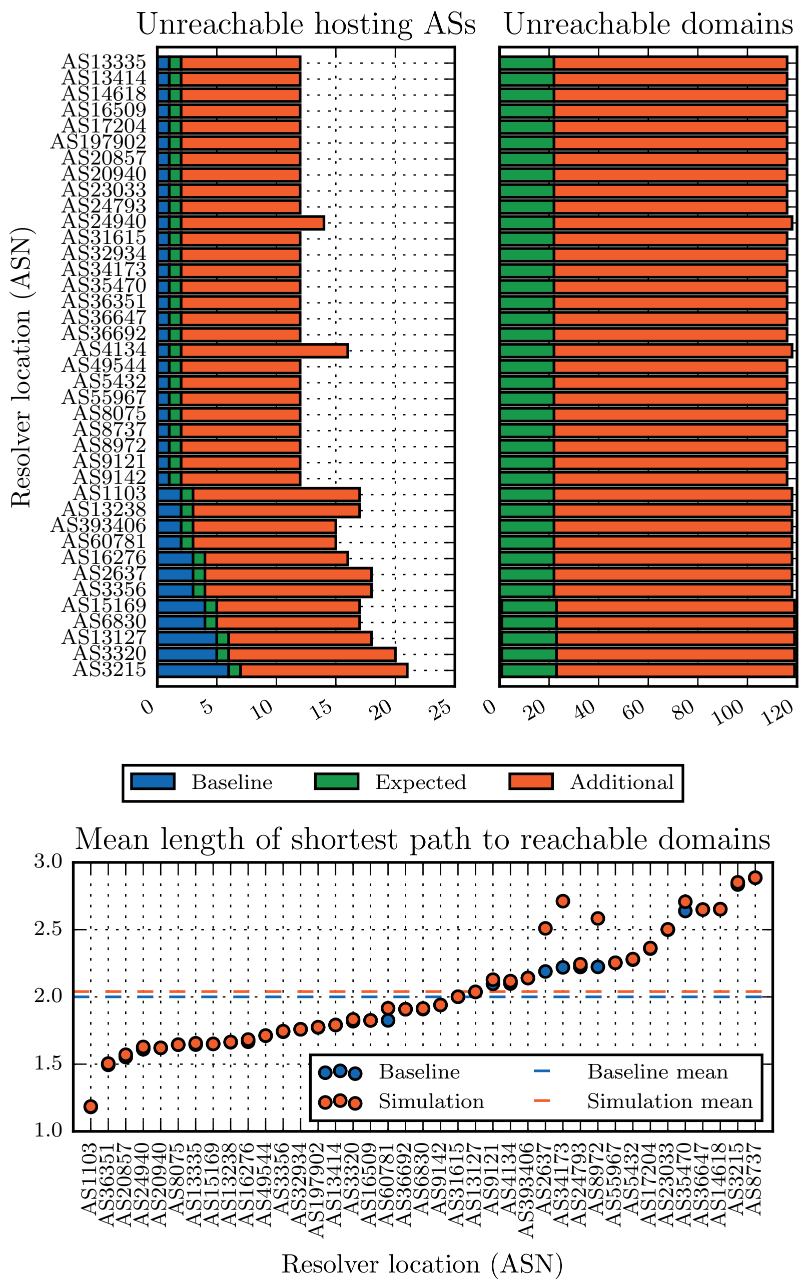


Figure B.21: Plots showing the reachability of autonomous systems and domains in a simulation where AS174 was removed from the AS topology compared to the baseline.

Table B.22: Reachability of DNS data in a simulation where AS2914 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	8	68	1.18178078259
AS2637	3	0	2.18911427628	9	68	2.18912524792
AS3215	6	1	2.83567157466	12	69	2.83573603762
AS3320	5	1	1.81932553893	11	69	1.81945484936
AS3356	3	0	1.74099517073	9	68	1.7410323372
AS4134	1	0	2.09931352367	8	68	2.09947192025
AS5432	1	0	2.2768154119	5	64	2.74091919733
AS6830	4	1	1.90947226796	10	69	1.9095149252
AS8075	1	0	1.64468828964	5	64	1.64472748272
AS8737	1	0	2.88639271487	5	64	2.8864004933
AS8972	1	0	2.22328412605	5	64	2.22330207481
AS9121	1	0	2.09727243649	5	64	2.10057907918
AS9142	1	0	1.93983546041	5	64	1.93984406281
AS13127	5	1	2.03705757563	10	69	2.03706587428
AS13238	2	0	1.66266719207	9	68	1.66266328904
AS13414	1	0	1.79106909723	6	68	1.79134978154
AS13335	1	0	1.64517926897	5	64	1.64517559524
AS14618	1	0	2.65234637417	5	64	2.6526701094
AS15169	4	1	1.65053505386	9	69	1.65053099706
AS16276	3	0	1.66660210991	8	68	1.66659974798
AS16509	1	0	1.82501377501	5	64	1.82533975653
AS17204	1	0	2.35903972347	5	64	2.95425691983
AS20857	1	0	1.55288727159	5	64	1.55654792306
AS20940	1	0	1.62063850426	5	64	1.62071991029
AS23033	1	0	2.49950622466	5	64	2.49956139403
AS24793	1	0	2.22269584558	5	64	2.22590258138
AS24940	1	0	1.61063288987	7	68	1.6106285137
AS31615	1	0	2.00108969823	5	64	2.00109940418
AS32934	1	0	1.75897836038	5	64	1.75910447583
AS34173	1	0	2.21939263956	5	64	2.21941091471
AS35470	1	0	2.638705947	5	64	2.63870816094
AS36351	1	0	1.4947956937	5	64	1.49481539032
AS36647	1	0	2.64950842419	5	64	2.64950442931
AS36692	1	0	1.90835667691	5	64	1.9448801094
AS49544	1	0	1.71134721447	5	64	1.71141013033
AS55967	1	0	2.25354329006	6	68	2.44381123339
AS60781	2	0	1.82534314497	10	69	1.85507926883
AS197902	1	0	1.77303725702	5	64	1.77309419832
AS393406	2	0	2.14028867497	8	68	2.24863366588
		Mean	2.0003648877		**Mean**	2.0371910837

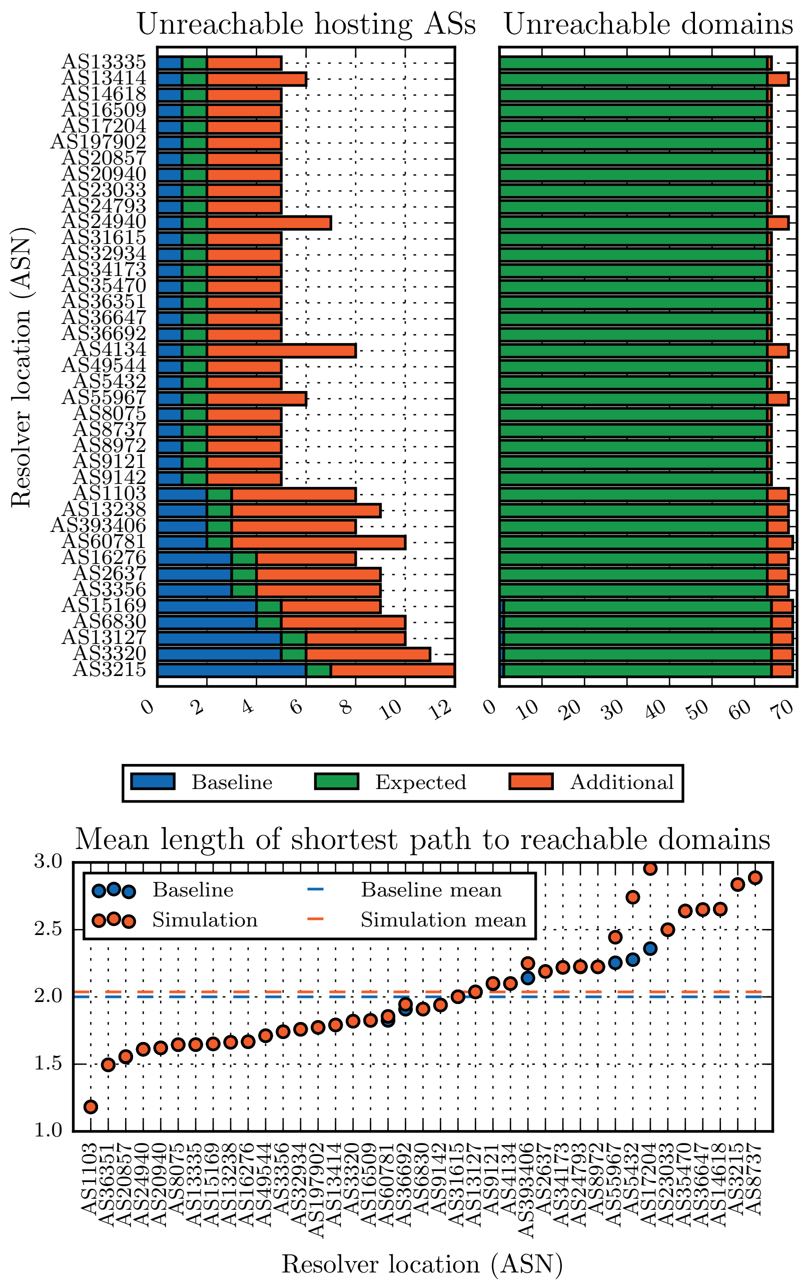


Figure B.22: Plots showing the reachability of autonomous systems and domains in a simulation where AS2914 was removed from the AS topology compared to the baseline.

Table B.23: Reachability of DNS data in a simulation where AS1299 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	6	32	1.18178645963
AS2637	3	0	2.18911427628	14	483	2.28279693269
AS3215	6	1	2.83567157466	19	484	2.8360911313
AS3320	5	1	1.81932553893	16	484	1.81943137451
AS3356	3	0	1.74099517073	16	483	1.74127589688
AS4134	1	0	2.09931352367	12	483	2.09975178518
AS5432	1	0	2.2768154119	8	69	2.27681524419
AS6830	4	1	1.90947226796	16	484	1.90983974808
AS8075	1	0	1.64468828964	8	69	1.64469415826
AS8737	1	0	2.88639271487	8	69	2.88680134039
AS8972	1	0	2.22328412605	8	69	2.22328010097
AS9121	1	0	2.09727243649	8	69	2.09847449605
AS9142	1	0	1.93983546041	8	69	1.93982778968
AS13127	5	1	2.03705757563	10	36	2.17973785181
AS13238	2	0	1.66266719207	8	69	1.66291114222
AS13414	1	0	1.79106909723	8	69	1.79792417832
AS13335	1	0	1.64517926897	7	69	1.64520583464
AS14618	1	0	2.65234637417	8	69	2.65284556377
AS15169	4	1	1.65053505386	11	70	1.65177200367
AS16276	3	0	1.66660210991	10	69	1.6667864617
AS16509	1	0	1.82501377501	8	69	1.82551891348
AS17204	1	0	2.35903972347	8	69	2.35918861733
AS20857	1	0	1.55288727159	8	69	1.55287518321
AS20940	1	0	1.62063850426	7	69	1.62064387715
AS23033	1	0	2.49950622466	8	69	3.08038463897
AS24793	1	0	2.22269584558	8	69	2.22267876472
AS24940	1	0	1.61063288987	8	69	1.61062042579
AS31615	1	0	2.00108969823	8	69	2.0010865434
AS32934	1	0	1.75897836038	8	69	1.75898420775
AS34173	1	0	2.21939263956	8	69	2.21938912364
AS35470	1	0	2.638705947	8	69	2.63869421679
AS36351	1	0	1.4947956937	8	69	1.49479982083
AS36647	1	0	2.64950842419	8	69	2.65086018485
AS36692	1	0	1.90835667691	8	69	1.9083487877
AS49544	1	0	1.71134721447	8	69	1.71133604575
AS55967	1	0	2.25354329006	8	69	2.25368635338
AS60781	2	0	1.82534314497	9	69	1.83421685274
AS197902	1	0	1.77303725702	8	69	1.77303769312
AS393406	2	0	2.14028867497	13	483	2.18983558914
		Mean	2.0003648877		**Mean**	2.0231855214

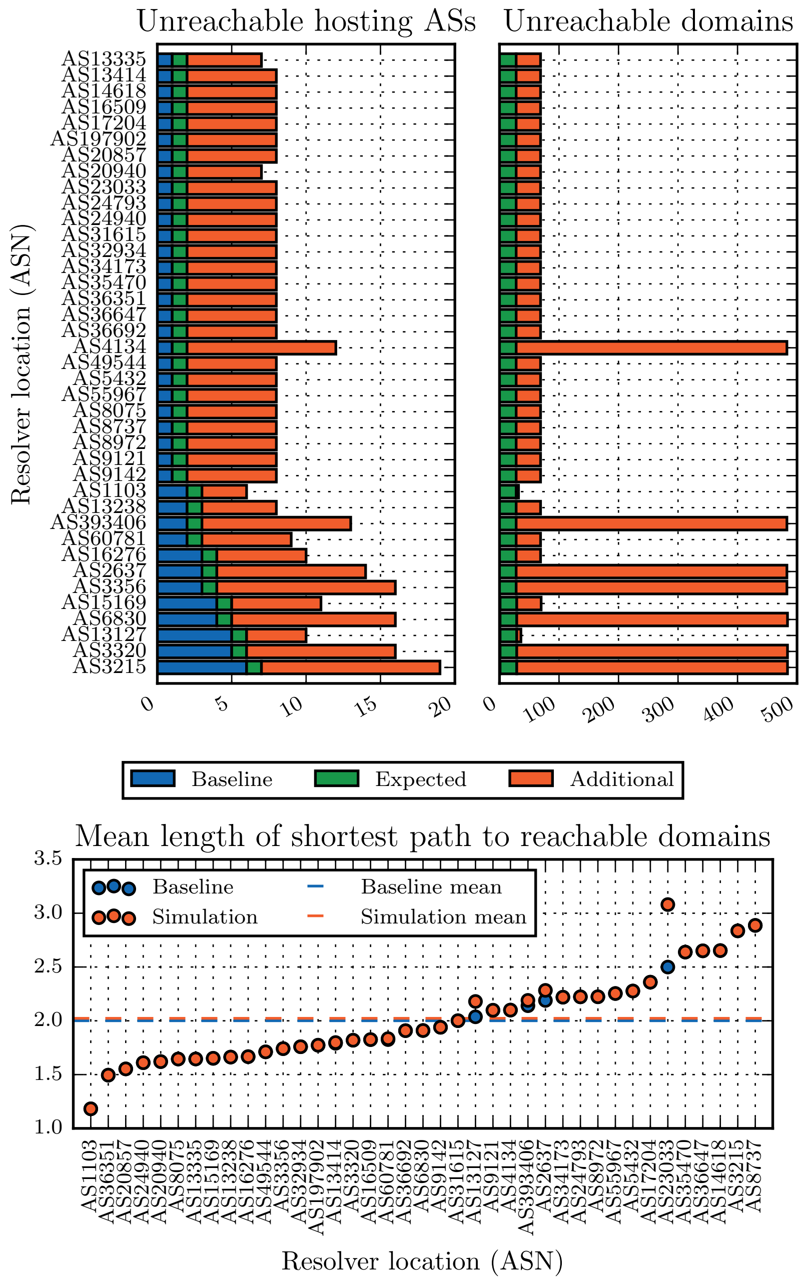


Figure B.23: Plots showing the reachability of autonomous systems and domains in a simulation where AS1299 was removed from the AS topology compared to the baseline.

Table B.24: Reachability of DNS data in a simulation where AS3356 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	11	260	1.18176398744
AS2637	3	0	2.18911427628	13	259	2.18980063301
AS3215	6	1	2.83567157466	16	261	2.8372617008
AS3320	5	1	1.81932553893	15	261	1.82114173347
AS3356	3	0	1.74099517073	n/a	n/a	n/a
AS4134	1	0	2.09931352367	11	260	2.101341255
AS5432	1	0	2.2768154119	2	237	2.27754289222
AS6830	4	1	1.90947226796	13	261	1.91039377749
AS8075	1	0	1.64468828964	2	237	1.64548253899
AS8737	1	0	2.88639271487	2	237	2.88649301684
AS8972	1	0	2.22328412605	2	237	2.22386533374
AS9121	1	0	2.09727243649	2	237	2.09773604538
AS9142	1	0	1.93983546041	2	237	1.94036984642
AS13127	5	1	2.03705757563	13	259	2.03777573017
AS13238	2	0	1.66266719207	9	258	1.66483997666
AS13414	1	0	1.79106909723	9	258	1.80709475015
AS13335	1	0	1.64517926897	2	237	1.6453632373
AS14618	1	0	2.65234637417	2	237	2.65360157821
AS15169	4	1	1.65053505386	13	259	1.65222053977
AS16276	3	0	1.66660210991	13	259	1.66917002406
AS16509	1	0	1.82501377501	2	237	1.8262786578
AS17204	1	0	2.35903972347	2	237	2.35972087878
AS20857	1	0	1.55288727159	2	237	1.55312699982
AS20940	1	0	1.62063850426	2	237	1.6206241678
AS23033	1	0	2.49950622466	2	237	2.5002558462
AS24793	1	0	2.22269584558	2	237	2.22269133055
AS24940	1	0	1.61063288987	8	258	1.6159607645
AS31615	1	0	2.00108969823	2	237	2.00136563153
AS32934	1	0	1.75897836038	2	237	1.75984383409
AS34173	1	0	2.21939263956	2	237	2.2201742513
AS35470	1	0	2.638705947	2	237	2.66771068073
AS36351	1	0	1.4947956937	2	237	1.49562358527
AS36647	1	0	2.64950842419	2	237	2.65068800725
AS36692	1	0	1.90835667691	2	237	1.91688847771
AS49544	1	0	1.71134721447	2	237	1.71299312841
AS55967	1	0	2.25354329006	9	258	2.25424091206
AS60781	2	0	1.82534314497	10	258	1.82601234405
AS197902	1	0	1.77303725702	2	237	1.77299386286
AS393406	2	0	2.14028867497	10	258	2.16296711921
		Mean	2.0003648877		**Mean**	2.0100899757

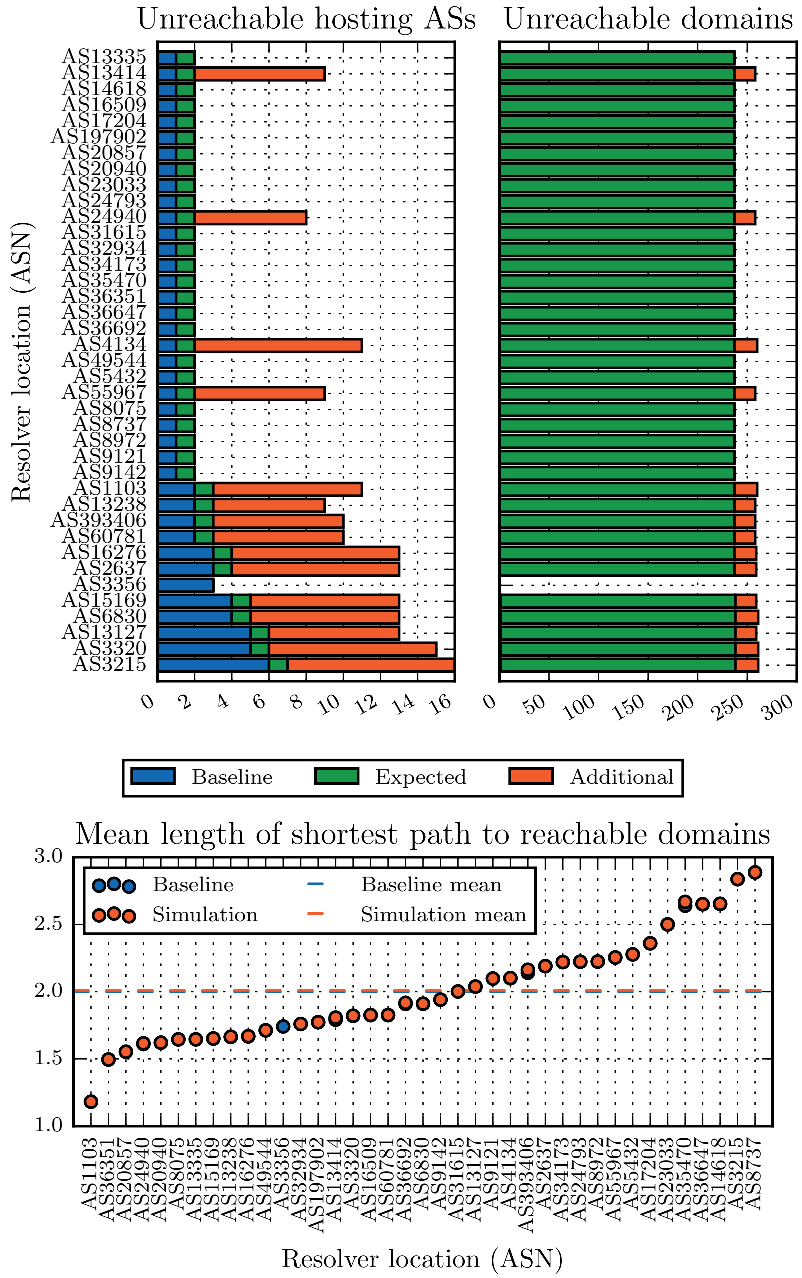


Figure B.24: Plots showing the reachability of autonomous systems and domains in a simulation where AS3356 was removed from the AS topology compared to the baseline.

Table B.25: Reachability of DNS data in a simulation where AS6453 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	109	1.18177583412
AS2637	3	0	2.18911427628	4	109	2.1891216604
AS3215	6	1	2.83567157466	7	110	2.83569395964
AS3320	5	1	1.81932553893	6	110	1.81934777819
AS3356	3	0	1.74099517073	4	109	1.74101786892
AS4134	1	0	2.09931352367	2	109	2.0993436886
AS5432	1	0	2.2768154119	2	109	2.27682364593
AS6830	4	1	1.90947226796	5	110	1.909495034
AS8075	1	0	1.64468828964	2	109	1.64470530296
AS8737	1	0	2.88639271487	2	109	2.88639133861
AS8972	1	0	2.22328412605	2	109	2.22329220444
AS9121	1	0	2.09727243649	2	109	2.09727478569
AS9142	1	0	1.93983546041	2	109	1.9398349836
AS13127	5	1	2.03705757563	6	110	2.0370637343
AS13238	2	0	1.66266719207	3	109	1.66266071092
AS13414	1	0	1.79106909723	2	109	1.79106522496
AS13335	1	0	1.64517926897	2	109	1.64520002036
AS14618	1	0	2.65234637417	2	109	2.65236950804
AS15169	4	1	1.65053505386	5	110	1.65057064003
AS16276	3	0	1.66660210991	4	109	1.66803941112
AS16509	1	0	1.82501377501	2	109	1.82504041714
AS17204	1	0	2.35903972347	2	109	2.35905428824
AS20857	1	0	1.55288727159	2	109	1.55287893274
AS20940	1	0	1.62063850426	2	109	1.62065465613
AS23033	1	0	2.49950622466	2	109	2.49951991536
AS24793	1	0	2.22269584558	2	109	2.22268079786
AS24940	1	0	1.61063288987	2	109	1.61164461098
AS31615	1	0	2.00108969823	2	109	2.00109251644
AS32934	1	0	1.75897836038	2	109	1.75899769585
AS34173	1	0	2.21939263956	2	109	2.21940063888
AS35470	1	0	2.638705947	2	109	2.63869935182
AS36351	1	0	1.4947956937	2	109	1.49478542891
AS36647	1	0	2.64950842419	2	109	2.64950167568
AS36692	1	0	1.90835667691	2	109	1.9083550013
AS49544	1	0	1.71134721447	2	109	1.7113417224
AS55967	1	0	2.25354329006	2	109	2.26510309377
AS60781	2	0	1.82534314497	3	109	1.82681125935
AS197902	1	0	1.77303725702	2	109	1.77301270029
AS393406	2	0	2.14028867497	3	109	2.14526386388
		Mean	2.0003648877		**Mean**	2.0008955359

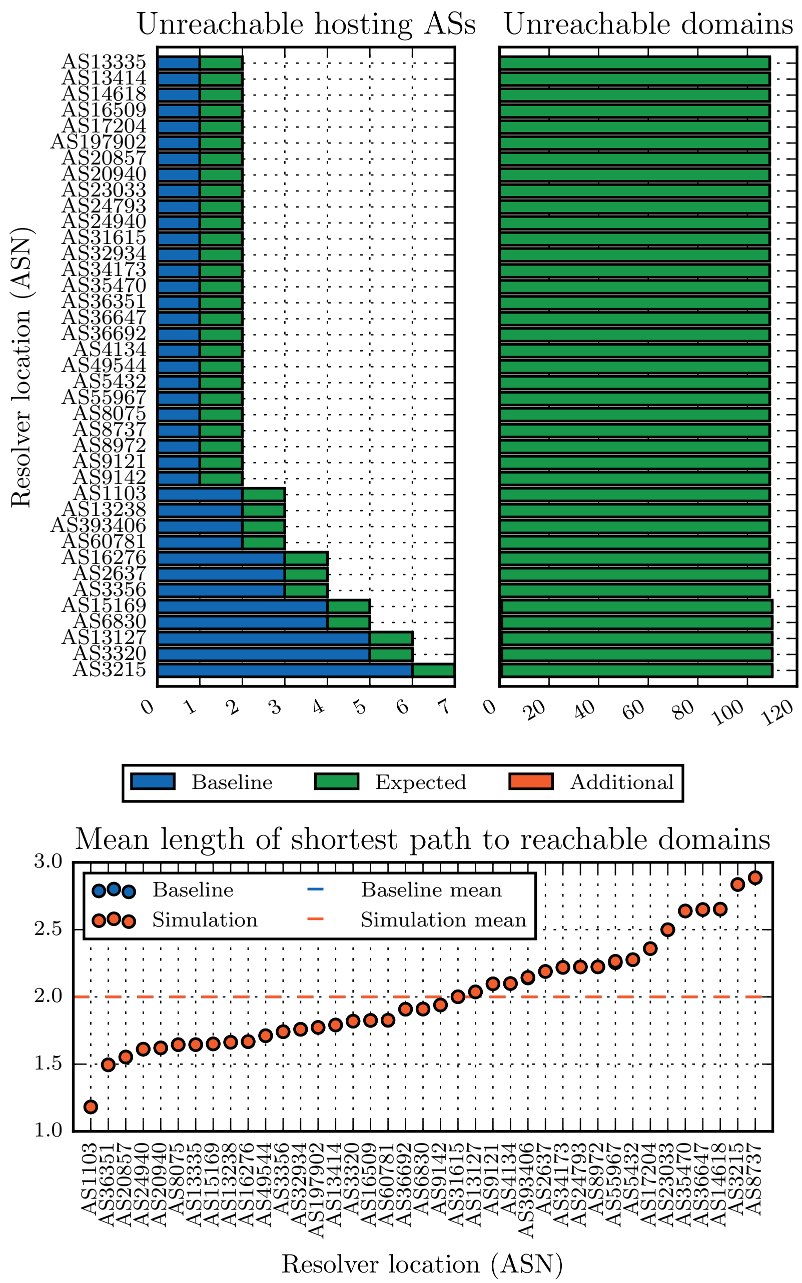


Figure B.25: Plots showing the reachability of autonomous systems and domains in a simulation where AS6453 was removed from the AS topology compared to the baseline.

Table B.26: Reachability of DNS data in a simulation where AS3320 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	5	46	1.18198526602
AS2637	3	0	2.18911427628	6	46	2.18915970237
AS3215	6	1	2.83567157466	9	47	2.83602172034
AS3320	5	1	1.81932553893	n/a	n/a	n/a
AS3356	3	0	1.74099517073	6	46	1.74126546999
AS4134	1	0	2.09931352367	4	46	2.09967972365
AS5432	1	0	2.2768154119	4	46	2.27684276336
AS6830	4	1	1.90947226796	7	47	1.90956804811
AS8075	1	0	1.64468828964	4	46	1.64472923395
AS8737	1	0	2.88639271487	4	46	2.88663499568
AS8972	1	0	2.22328412605	4	46	2.23091529844
AS9121	1	0	2.09727243649	4	46	2.09763134928
AS9142	1	0	1.93983546041	4	46	1.93984221422
AS13127	5	1	2.03705757563	8	47	2.03708007525
AS13238	2	0	1.66266719207	5	46	1.66266504522
AS13414	1	0	1.79106909723	4	46	1.79106805136
AS13335	1	0	1.64517926897	4	46	1.64517659936
AS14618	1	0	2.65234637417	4	46	2.65296933198
AS15169	4	1	1.65053505386	7	47	1.65053317579
AS16276	3	0	1.66660210991	6	46	1.66664864778
AS16509	1	0	1.82501377501	4	46	1.82564697426
AS17204	1	0	2.35903972347	4	46	2.35930656125
AS20857	1	0	1.55288727159	4	46	1.55289070751
AS20940	1	0	1.62063850426	4	46	1.62067868315
AS23033	1	0	2.49950622466	4	46	2.49961340173
AS24793	1	0	2.22269584558	4	46	2.22269421344
AS24940	1	0	1.61063288987	4	46	1.61062992405
AS31615	1	0	2.00108969823	4	46	2.05350546215
AS32934	1	0	1.75897836038	4	46	1.75927323252
AS34173	1	0	2.21939263956	4	46	2.21941204255
AS35470	1	0	2.638705947	4	46	2.63890938278
AS36351	1	0	1.4947956937	4	46	1.49479136182
AS36647	1	0	2.64950842419	4	46	2.6495057917
AS36692	1	0	1.90835667691	4	46	1.90856410243
AS49544	1	0	1.71134721447	4	46	1.71139600003
AS55967	1	0	2.25354329006	4	46	2.25379766632
AS60781	2	0	1.82534314497	5	46	1.82578640315
AS197902	1	0	1.77303725702	4	46	1.77306513529
AS393406	2	0	2.14028867497	5	46	2.14031355096
		Mean	2.0003648877		**Mean**	2.0068472976

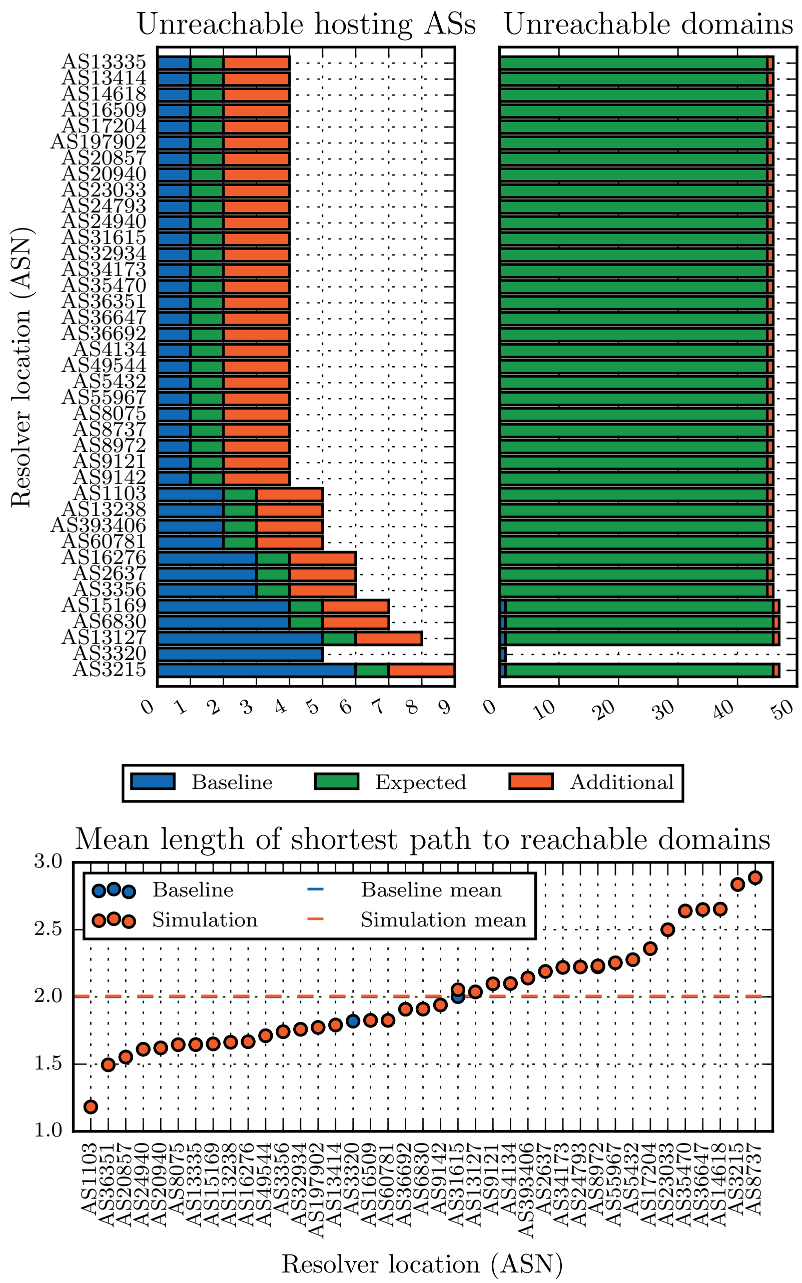


Figure B.26: Plots showing the reachability of autonomous systems and domains in a simulation where AS3320 was removed from the AS topology compared to the baseline.

Table B.27: Reachability of DNS data in a simulation where AS20562 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.29289787406
AS6830	4	1	1.90947226796	4	1	1.91176462365
AS8075	1	0	1.64468828964	1	0	1.66055117183
AS8737	1	0	2.88639271487	1	0	2.89421110396
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.97483292909
AS13127	5	1	2.03705757563	5	1	2.20599494444
AS13238	2	0	1.66266719207	2	0	1.67952545376
AS13414	1	0	1.79106909723	1	0	1.84876792895
AS13335	1	0	1.64517926897	1	0	1.65908438507
AS14618	1	0	2.65234637417	1	0	2.66861318658
AS15169	4	1	1.65053505386	4	1	1.6749308034
AS16276	3	0	1.66660210991	3	0	1.77321228723
AS16509	1	0	1.82501377501	1	0	1.84128058742
AS17204	1	0	2.35903972347	1	0	2.373691747
AS20857	1	0	1.55288727159	2	0	1.63317189794
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.68295597142
AS24940	1	0	1.61063288987	1	0	1.62948657058
AS31615	1	0	2.00108969823	1	0	2.38866363405
AS32934	1	0	1.75897836038	1	0	1.77275187761
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.66869949008
AS36351	1	0	1.4947956937	1	0	1.50924547251
AS36647	1	0	2.64950842419	1	0	2.67014782318
AS36692	1	0	1.90835667691	1	0	1.95963009163
AS49544	1	0	1.71134721447	1	0	1.72868974506
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	2.06632862286
AS393406	2	0	2.14028867497	2	0	2.14029240298
		Mean	2.0003648877		**Mean**	2.0490926917

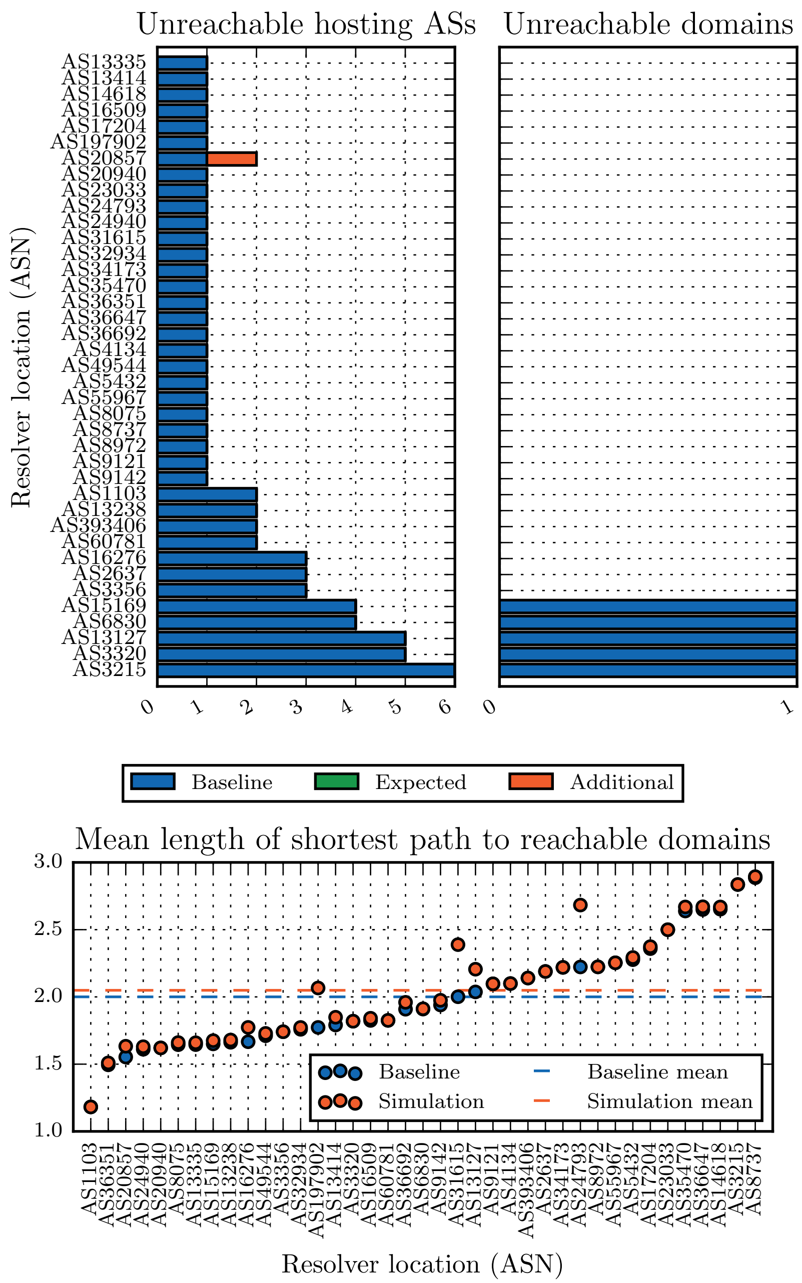


Figure B.27: Plots showing the reachability of autonomous systems and domains in a simulation where AS20562 was removed from the AS topology compared to the baseline.

Table B.28: Reachability of DNS data in a simulation where AS6939 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	10	200	1.18181433504
AS2637	3	0	2.18911427628	8	158	2.20720124221
AS3215	6	1	2.83567157466	14	201	2.92374585406
AS3320	5	1	1.81932553893	13	201	1.819384605
AS3356	3	0	1.74099517073	11	200	1.74104386022
AS4134	1	0	2.09931352367	9	200	2.0993951446
AS5432	1	0	2.2768154119	6	158	2.27684462861
AS6830	4	1	1.90947226796	12	201	1.90952947543
AS8075	1	0	1.64468828964	6	158	1.64473933897
AS8737	1	0	2.88639271487	6	158	2.9838814606
AS8972	1	0	2.22328412605	6	158	2.22331605348
AS9121	1	0	2.09727243649	6	158	2.10863582391
AS9142	1	0	1.93983546041	6	158	1.95120185362
AS13127	5	1	2.03705757563	10	159	2.04898381603
AS13238	2	0	1.66266719207	6	158	1.66273852996
AS13414	1	0	1.79106909723	6	158	1.80268723099
AS13335	1	0	1.64517926897	5	158	1.66530795227
AS14618	1	0	2.65234637417	6	158	2.65252310038
AS15169	4	1	1.65053505386	12	201	1.6506862504
AS16276	3	0	1.66660210991	11	200	1.67548747453
AS16509	1	0	1.82501377501	6	158	1.82519912836
AS17204	1	0	2.35903972347	6	158	2.3590907481
AS20857	1	0	1.55288727159	6	158	1.55625681547
AS20940	1	0	1.62063850426	5	158	1.62069070933
AS23033	1	0	2.49950622466	6	158	2.56416677385
AS24793	1	0	2.22269584558	6	158	2.26172876787
AS24940	1	0	1.61063288987	5	158	1.61949957406
AS31615	1	0	2.00108969823	6	158	2.00109476329
AS32934	1	0	1.75897836038	6	158	1.7590420961
AS34173	1	0	2.21939263956	6	158	2.21942426598
AS35470	1	0	2.638705947	6	158	2.65090472223
AS36351	1	0	1.4947956937	6	158	1.49485276711
AS36647	1	0	2.64950842419	6	158	2.64959615109
AS36692	1	0	1.90835667691	6	158	1.91989419587
AS49544	1	0	1.71134721447	6	158	1.71141812203
AS55967	1	0	2.25354329006	6	158	2.25721251978
AS60781	2	0	1.82534314497	7	158	1.82534098344
AS197902	1	0	1.77303725702	6	158	1.77305722855
AS393406	2	0	2.14028867497	10	200	2.14313568908
		Mean	2.0003648877		**Mean**	2.0113013859

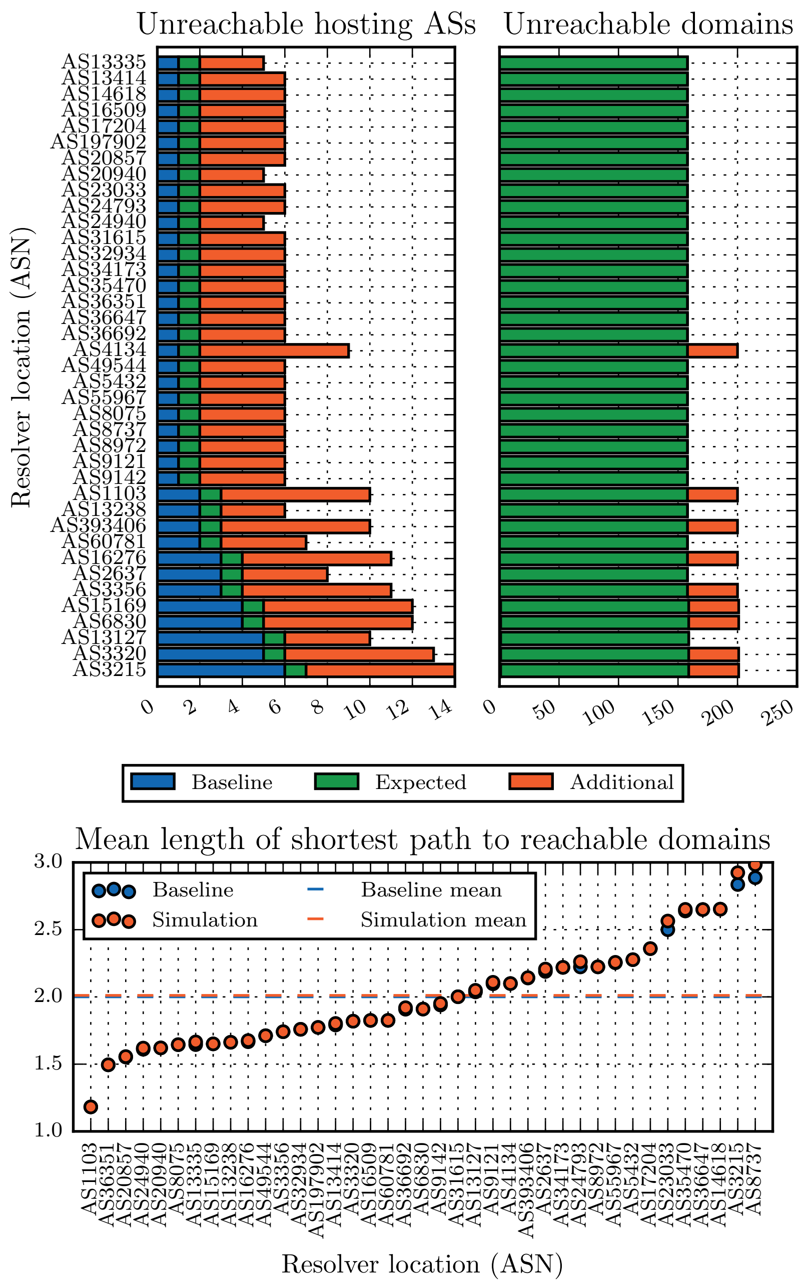


Figure B.28: Plots showing the reachability of autonomous systems and domains in a simulation where AS6939 was removed from the AS topology compared to the baseline.

Table B.29: Reachability of DNS data in a simulation where AS43531 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	0	1.18179301773
AS2637	3	0	2.18911427628	4	0	2.18911427628
AS3215	6	1	2.83567157466	7	1	2.83569002833
AS3320	5	1	1.81932553893	6	1	1.81932553893
AS3356	3	0	1.74099517073	4	0	1.7410136244
AS4134	1	0	2.09931352367	2	0	2.09931352367
AS5432	1	0	2.2768154119	2	0	2.27681653031
AS6830	4	1	1.90947226796	5	1	1.90947226796
AS8075	1	0	1.64468828964	2	0	1.64468847604
AS8737	1	0	2.88639271487	2	0	2.89230720767
AS8972	1	0	2.22328412605	4	0	2.22328412605
AS9121	1	0	2.09727243649	2	0	2.135124072
AS9142	1	0	1.93983546041	5	1	2.088895794
AS13127	5	1	2.03705757563	6	1	2.04636679891
AS13238	2	0	1.66266719207	3	0	1.66267036088
AS13414	1	0	1.79106909723	2	0	1.79109183811
AS13335	1	0	1.64517926897	2	0	1.64717133277
AS14618	1	0	2.65234637417	2	0	2.6596061205
AS15169	4	1	1.65053505386	5	1	1.65053878188
AS16276	3	0	1.66660210991	4	0	1.66660807473
AS16509	1	0	1.82501377501	2	0	1.83227389414
AS17204	1	0	2.35903972347	2	0	2.35904084187
AS20857	1	0	1.55288727159	2	0	1.5529167229
AS20940	1	0	1.62063850426	2	0	1.62063869066
AS23033	1	0	2.49950622466	2	0	2.49950622466
AS24793	1	0	2.22269584558	2	0	2.25579929719
AS24940	1	0	1.61063288987	2	0	1.61063717709
AS31615	1	0	2.00108969823	2	0	2.00145709392
AS32934	1	0	1.75897836038	2	0	1.76019574306
AS34173	1	0	2.21939263956	3	0	2.21939263956
AS35470	1	0	2.638705947	2	0	2.63872346866
AS36351	1	0	1.4947956937	2	0	1.49542516871
AS36647	1	0	2.64950842419	2	0	2.64950898339
AS36692	1	0	1.90835667691	2	0	1.90836581054
AS49544	1	0	1.71134721447	2	0	1.71350014204
AS55967	1	0	2.25354329006	2	0	2.25354329006
AS60781	2	0	1.82534314497	3	0	1.82534314497
AS197902	1	0	1.77303725702	2	0	1.77333158365
AS393406	2	0	2.14028867497	3	0	2.14220804252
		Mean	2.0003648877		**Mean**	2.0069923013

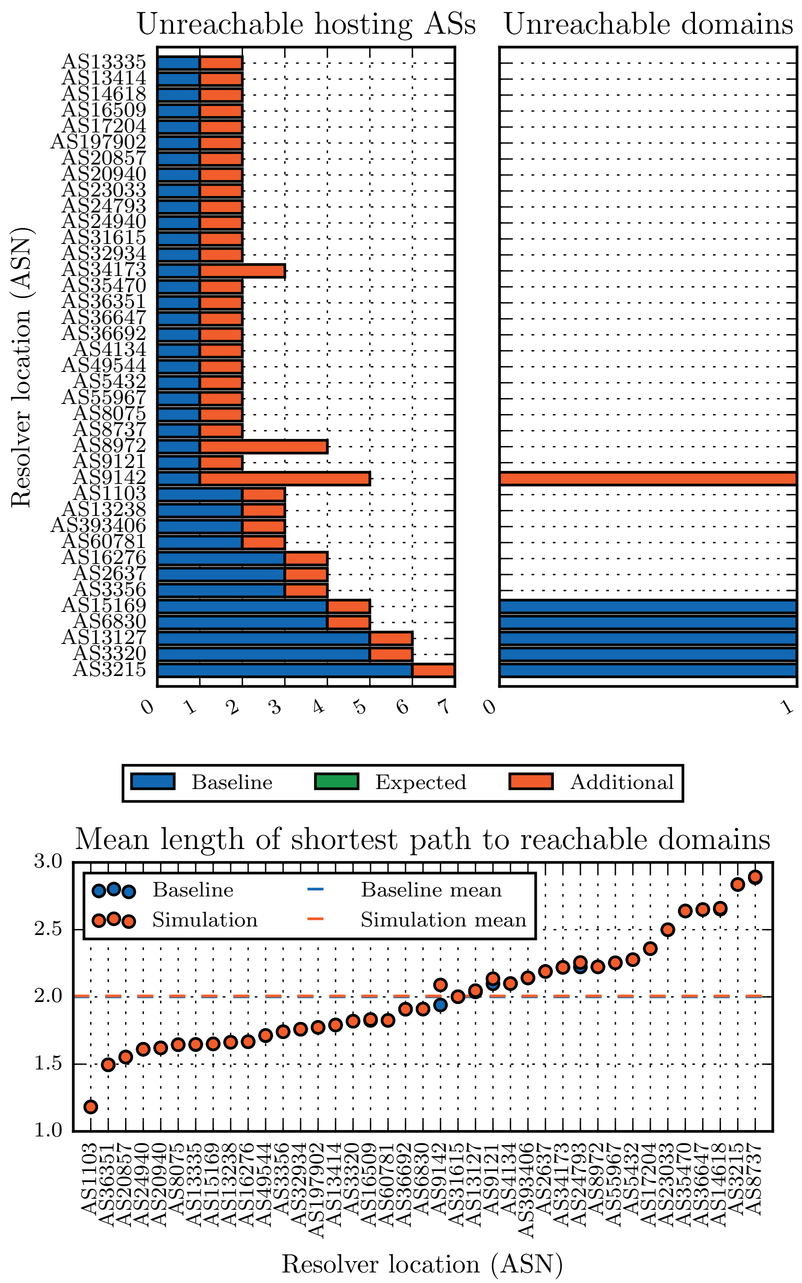


Figure B.29: Plots showing the reachability of autonomous systems and domains in a simulation where AS43531 was removed from the AS topology compared to the baseline.

Table B.30: Reachability of DNS data in a simulation where AS10310 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	7	1	1.18180498126
AS2637	3	0	2.18911427628	8	1	2.18911412513
AS3215	6	1	2.83567157466	11	2	2.83567135763
AS3320	5	1	1.81932553893	10	2	1.81932531885
AS3356	3	0	1.74099517073	8	1	1.7410068657
AS4134	1	0	2.09931352367	6	1	2.09931335578
AS5432	1	0	2.2768154119	6	1	2.2768152771
AS6830	4	1	1.90947226796	9	2	1.90948399433
AS8075	1	0	1.64468828964	6	1	1.64469996665
AS8737	1	0	2.88639271487	6	1	2.88640425053
AS8972	1	0	2.22328412605	6	1	2.22328398127
AS9121	1	0	2.09727243649	6	1	2.09727226822
AS9142	1	0	1.93983546041	6	1	1.93984700604
AS13127	5	1	2.03705757563	10	2	2.03706913938
AS13238	2	0	1.66266719207	7	1	1.66267887243
AS13414	1	0	1.79106909723	6	1	1.79108080153
AS13335	1	0	1.64517926897	6	1	1.64519094607
AS14618	1	0	2.65234637417	6	1	2.65234612297
AS15169	4	1	1.65053505386	9	2	1.65054673197
AS16276	3	0	1.66660210991	8	1	1.666613791
AS16509	1	0	1.82501377501	6	1	1.82501355599
AS17204	1	0	2.35903972347	6	1	2.359039604
AS20857	1	0	1.55288727159	6	1	1.55288700185
AS20940	1	0	1.62063850426	6	1	1.62063843355
AS23033	1	0	2.49950622466	6	1	2.49950613137
AS24793	1	0	2.22269584558	6	1	2.22269551429
AS24940	1	0	1.61063288987	6	1	1.61064456054
AS31615	1	0	2.00108969823	6	1	2.00110125528
AS32934	1	0	1.75897836038	6	1	1.75899005869
AS34173	1	0	2.21939263956	6	1	2.21939249405
AS35470	1	0	2.638705947	6	1	2.63870569325
AS36351	1	0	1.4947956937	6	1	1.49479559953
AS36647	1	0	2.64950842419	3,805	5,364,786	0.0
AS36692	1	0	1.90835667691	6	1	1.90835647343
AS49544	1	0	1.71134721447	6	1	1.7113589039
AS55967	1	0	2.25354329006	6	1	2.25354315092
AS60781	2	0	1.82534314497	7	1	1.82534292601
AS197902	1	0	1.77303725702	6	1	1.77304877155
AS393406	2	0	2.14028867497	7	1	2.14030044436
		Mean	2.0003648877		**Mean**	1.9324340955

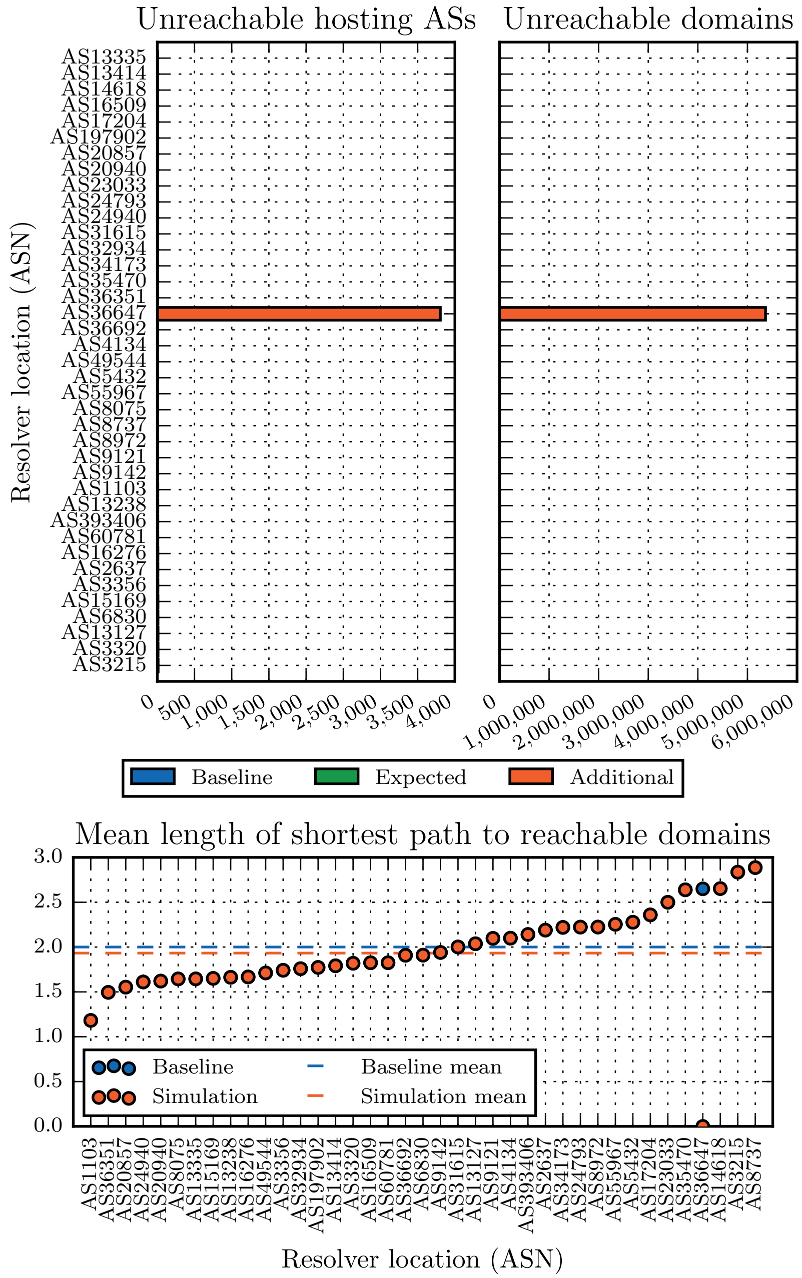


Figure B.30: Plots showing the reachability of autonomous systems and domains in a simulation where AS10310 was removed from the AS topology compared to the baseline.

Table B.31: Reachability of DNS data in a simulation where AS4436 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	2	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	5	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0003648877

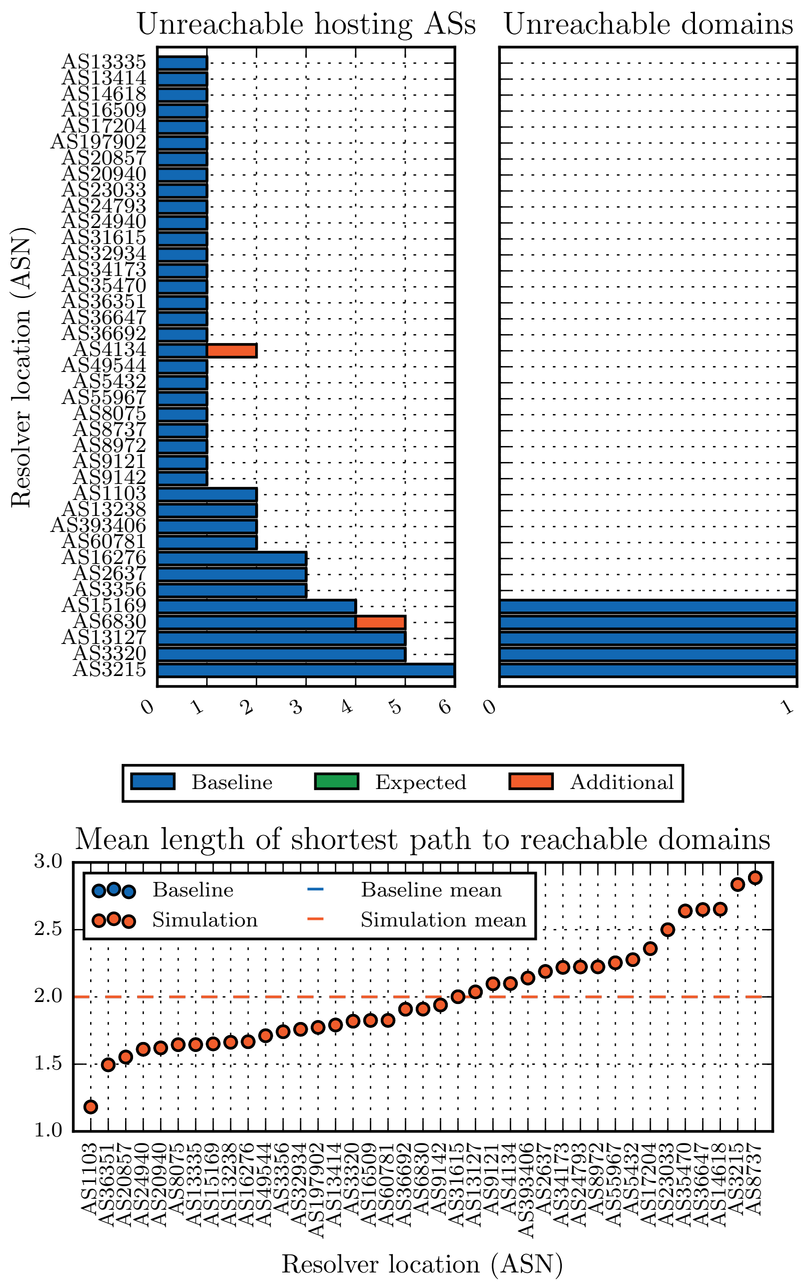


Figure B.31: Plots showing the reachability of autonomous systems and domains in a simulation where AS4436 was removed from the AS topology compared to the baseline.

Table B.32: Reachability of DNS data in a simulation where AS49685 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	5	47,051	1.17454624025
AS2637	3	0	2.18911427628	6	47,051	2.18193490953
AS3215	6	1	2.83567157466	9	47,052	2.8253632749
AS3320	5	1	1.81932553893	8	47,052	1.80887261045
AS3356	3	0	1.74099517073	6	47,051	1.7388511316
AS4134	1	0	2.09931352367	4	47,051	2.0913379131
AS5432	1	0	2.2768154119	4	47,051	2.26156239017
AS6830	4	1	1.90947226796	7	47,052	1.89981695218
AS8075	1	0	1.64468828964	4	47,051	1.63269018381
AS8737	1	0	2.88639271487	4	47,051	2.89362053821
AS8972	1	0	2.22328412605	4	47,051	2.21640709196
AS9121	1	0	2.09727243649	4	47,051	2.08927876651
AS9142	1	0	1.93983546041	4	47,051	1.9304489861
AS13127	5	1	2.03705757563	8	47,052	2.01968318848
AS13238	2	0	1.66266719207	5	47,051	1.65082816243
AS13414	1	0	1.79106909723	4	47,051	1.78036615951
AS13335	1	0	1.64517926897	4	47,051	1.63318550729
AS14618	1	0	2.65234637417	4	47,051	2.64041602659
AS15169	4	1	1.65053505386	7	47,052	1.63858867759
AS16276	3	0	1.66660210991	6	47,051	1.65479958862
AS16509	1	0	1.82501377501	4	47,051	1.81461117765
AS17204	1	0	2.35903972347	4	47,051	2.34451421723
AS20857	1	0	1.55288727159	4	47,051	1.54007691618
AS20940	1	0	1.62063850426	4	47,051	1.60842760746
AS23033	1	0	2.49950622466	4	47,051	2.48622355713
AS24793	1	0	2.22269584558	4	47,051	2.20696397735
AS24940	1	0	1.61063288987	4	47,051	1.59833346403
AS31615	1	0	2.00108969823	4	47,051	1.99224519753
AS32934	1	0	1.75897836038	4	47,051	1.74799148585
AS34173	1	0	2.21939263956	4	47,051	2.21248230215
AS35470	1	0	2.638705947	3,805	5,158,333	0.0
AS36351	1	0	1.4947956937	4	47,051	1.48147134768
AS36647	1	0	2.64950842419	4	47,051	2.63755296661
AS36692	1	0	1.90835667691	4	47,051	1.89869149226
AS49544	1	0	1.71134721447	4	47,051	1.69993890258
AS55967	1	0	2.25354329006	4	47,051	2.23808435806
AS60781	2	0	1.82534314497	5	47,051	1.8149451543
AS197902	1	0	1.77303725702	4	47,051	1.76217496277
AS393406	2	0	2.14028867497	5	47,051	2.13267730239
		Mean	2.0003648877		**Mean**	1.9225642228

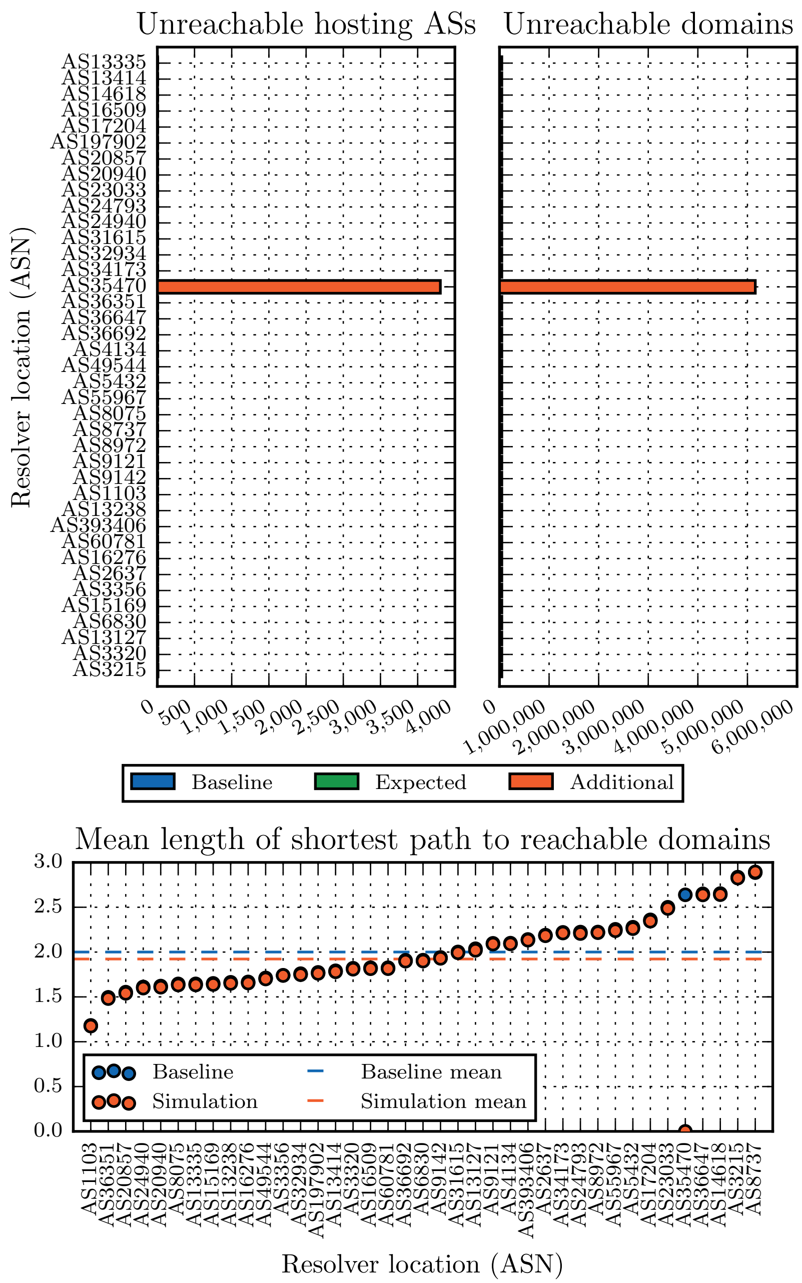


Figure B.32: Plots showing the reachability of autonomous systems and domains in a simulation where AS49685 was removed from the AS topology compared to the baseline.

Table B.33: Reachability of DNS data in a simulation where AS5511 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	5	29	1.18178207073
AS2637	3	0	2.18911427628	6	29	2.18910989291
AS3215	6	1	2.83567157466	1,501	1,171,108	6.98786531161
AS3320	5	1	1.81932553893	8	30	1.81932661268
AS3356	3	0	1.74099517073	6	29	1.74099507545
AS4134	1	0	2.09931352367	4	29	2.09931704295
AS5432	1	0	2.2768154119	1	0	2.27683181516
AS6830	4	1	1.90947226796	7	30	1.909471965
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88641992936
AS8972	1	0	2.22328412605	1	0	2.22329493728
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	8	30	2.0370525567
AS13238	2	0	1.66266719207	5	29	1.66265996292
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	7	30	1.65052794553
AS16276	3	0	1.66660210991	6	29	1.66660496772
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35905183951
AS20857	1	0	1.55288727159	1	0	1.55291970531
AS20940	1	0	1.62063850426	1	0	1.62064260508
AS23033	1	0	2.49950622466	1	0	2.49952262792
AS24793	1	0	2.22269584558	1	0	2.22272827929
AS24940	1	0	1.61063288987	4	29	1.61062537944
AS31615	1	0	2.00108969823	1	0	2.00111691273
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.63874378633
AS36351	1	0	1.4947956937	1	0	1.49480128572
AS36647	1	0	2.64950842419	1	0	2.64951401621
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354888208
AS60781	2	0	1.82534314497	5	29	1.82535077531
AS197902	1	0	1.77303725702	1	0	1.77306428511
AS393406	2	0	2.14028867497	5	29	2.14028421407
		Mean	2.0003648877		**Mean**	2.1068372891

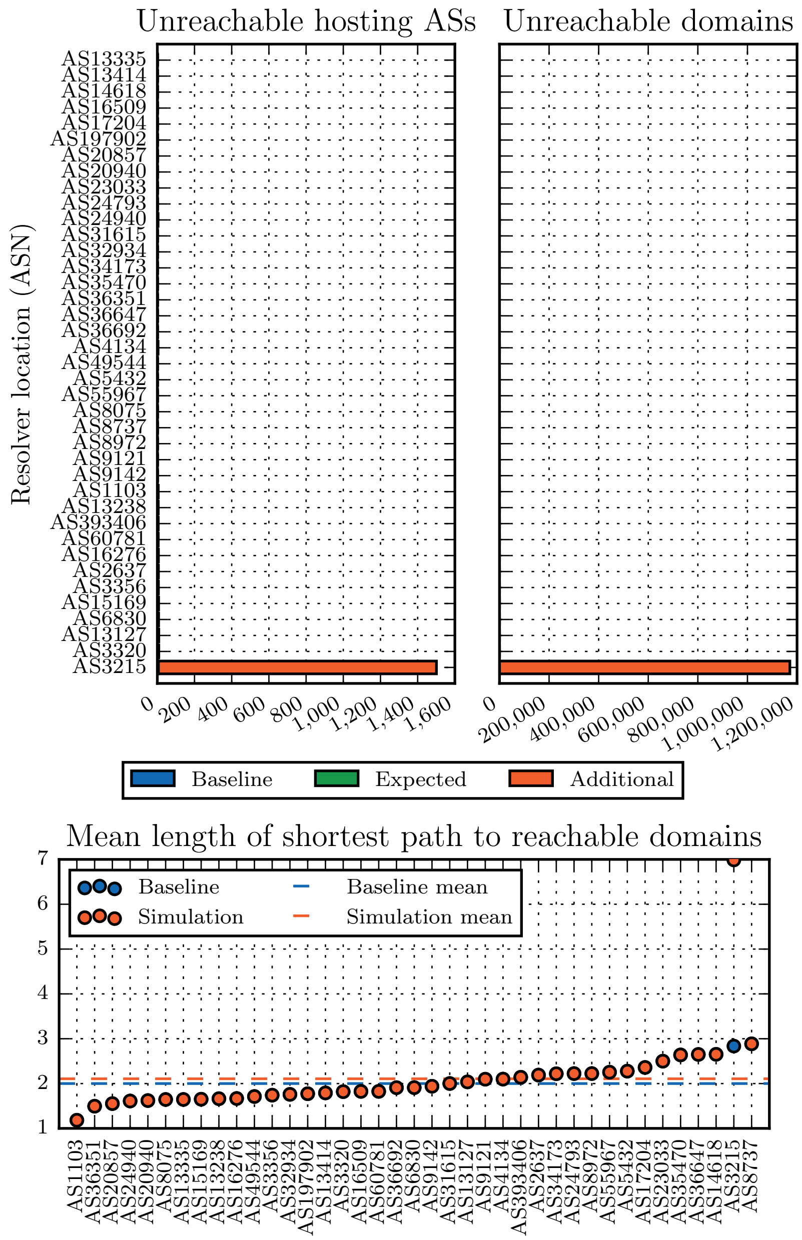


Figure B.33: Plots showing the reachability of autonomous systems and domains in a simulation where AS5511 was removed from the AS topology compared to the baseline.

Table B.34: Reachability of DNS data in a simulation where AS16509 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	28,293	1.18393627278
AS2637	3	0	2.18911427628	5	28,293	2.19038657396
AS3215	6	1	2.83567157466	8	28,294	2.83500178207
AS3320	5	1	1.81932553893	7	28,294	1.82569023782
AS3356	3	0	1.74099517073	5	28,293	1.74643656557
AS4134	1	0	2.09931352367	3	28,293	2.10004244359
AS5432	1	0	2.2768154119	2	15,057	2.28617644513
AS6830	4	1	1.90947226796	6	28,294	1.91653996051
AS8075	1	0	1.64468828964	2	15,057	1.64385405547
AS8737	1	0	2.88639271487	2	15,057	2.89436795981
AS8972	1	0	2.22328412605	2	15,057	2.2391021156
AS9121	1	0	2.09727243649	2	15,057	2.10756129607
AS9142	1	0	1.93983546041	2	15,057	1.94984140324
AS13127	5	1	2.03705757563	7	28,294	2.03773226392
AS13238	2	0	1.66266719207	4	28,293	1.66105599274
AS13414	1	0	1.79106909723	3	28,293	1.79017819749
AS13335	1	0	1.64517926897	2	15,057	1.6443510898
AS14618	1	0	2.65234637417	1,497	1,009,983	4.5183056417
AS15169	4	1	1.65053505386	6	28,294	1.64882149216
AS16276	3	0	1.66660210991	4	15,057	1.66591852936
AS16509	1	0	1.82501377501	n/a	n/a	n/a
AS17204	1	0	2.35903972347	2	15,057	2.3740343206
AS20857	1	0	1.55288727159	2	15,057	1.56170712135
AS20940	1	0	1.62063850426	2	15,057	1.61973788962
AS23033	1	0	2.49950622466	2	15,057	2.51219528608
AS24793	1	0	2.22269584558	2	15,057	2.22963864912
AS24940	1	0	1.61063288987	2	15,057	1.6121532092
AS31615	1	0	2.00108969823	2	15,057	2.00875184192
AS32934	1	0	1.75897836038	2	15,057	1.75846636027
AS34173	1	0	2.21939263956	2	15,057	2.22522534311
AS35470	1	0	2.638705947	2	15,057	2.65021568374
AS36351	1	0	1.4947956937	2	15,057	1.49929762824
AS36647	1	0	2.64950842419	2	15,057	2.65110918661
AS36692	1	0	1.90835667691	2	15,057	1.91582922581
AS49544	1	0	1.71134721447	2	15,057	1.71312295889
AS55967	1	0	2.25354329006	3	28,293	2.25666191011
AS60781	2	0	1.82534314497	3	15,057	1.82515625552
AS197902	1	0	1.77303725702	2	15,057	1.77521785675
AS393406	2	0	2.14028867497	4	28,293	2.14943422602
		Mean	2.0003648877		**Mean**	2.0585067177

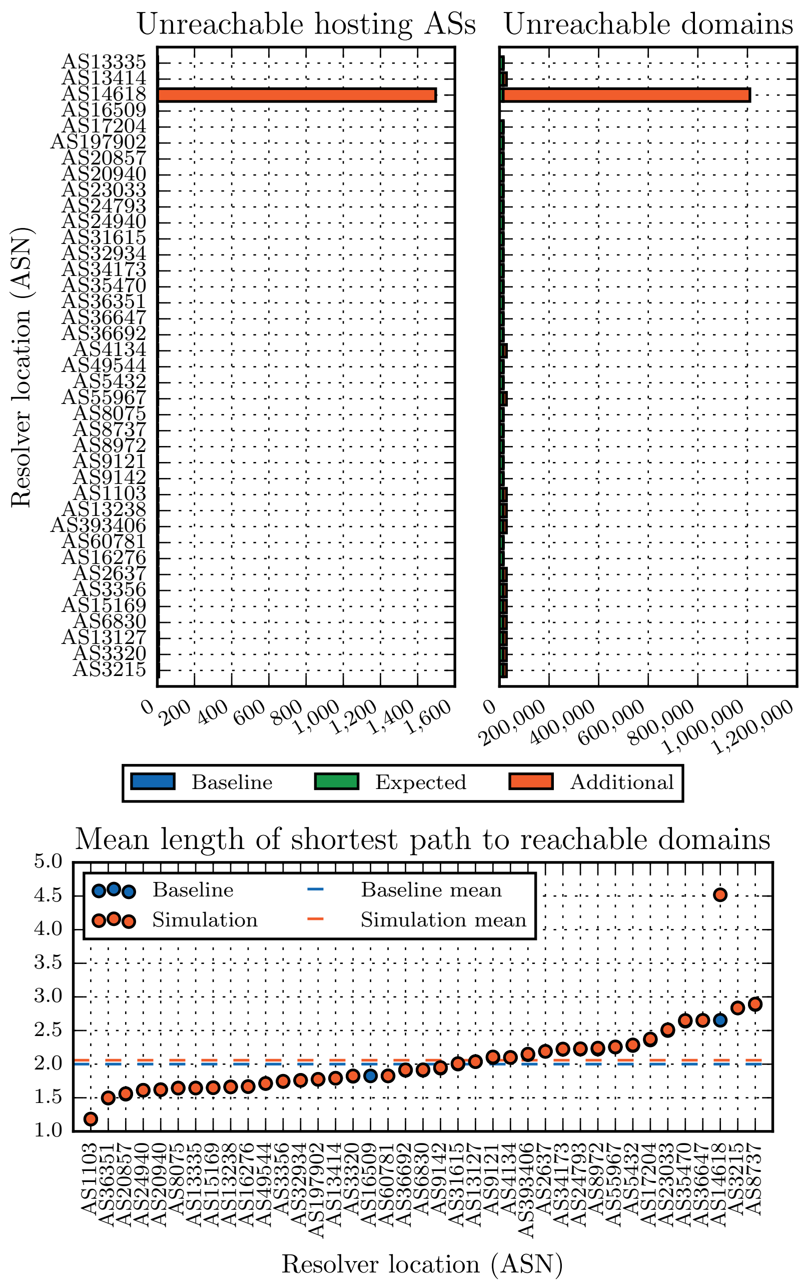


Figure B.34: Plots showing the reachability of autonomous systems and domains in a simulation where AS16509 was removed from the AS topology compared to the baseline.

Table B.35: Reachability of DNS data in a simulation where AS9002 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	0	1.18179283133
AS2637	3	0	2.18911427628	5	0	2.18911464908
AS3215	6	1	2.83567157466	8	1	2.83567194746
AS3320	5	1	1.81932553893	7	1	1.81932591173
AS3356	3	0	1.74099517073	5	0	1.74101138759
AS4134	1	0	2.09931352367	3	0	2.09936198784
AS5432	1	0	2.2768154119	2	0	2.27705120873
AS6830	4	1	1.90947226796	6	1	1.90952054574
AS8075	1	0	1.64468828964	2	0	1.64469089925
AS8737	1	0	2.88639271487	2	0	2.88647379915
AS8972	1	0	2.22328412605	2	0	2.22328449885
AS9121	1	0	2.09727243649	2	0	2.09732052786
AS9142	1	0	1.93983546041	2	0	1.93989193981
AS13127	5	1	2.03705757563	7	1	2.03712020626
AS13238	2	0	1.66266719207	5	0	1.67850453736
AS13414	1	0	1.79106909723	2	0	1.79107729886
AS13335	1	0	1.64517926897	2	0	1.64518225138
AS14618	1	0	2.65234637417	2	0	2.65237787588
AS15169	4	1	1.65053505386	6	1	1.65053822267
AS16276	3	0	1.66660210991	5	0	1.66661105714
AS16509	1	0	1.82501377501	2	0	1.82504639512
AS17204	1	0	2.35903972347	2	0	2.35905333072
AS20857	1	0	1.55288727159	2	0	1.55293126215
AS20940	1	0	1.62063850426	2	0	1.62064111387
AS23033	1	0	2.49950622466	2	0	2.49951312149
AS24793	1	0	2.22269584558	2	0	2.22270758882
AS24940	1	0	1.61063288987	3	0	1.61063568588
AS31615	1	0	2.00108969823	2	0	2.00109007103
AS32934	1	0	1.75897836038	2	0	1.7589867484
AS34173	1	0	2.21939263956	2	0	2.21939301236
AS35470	1	0	2.638705947	2	0	2.6387063198
AS36351	1	0	1.4947956937	2	0	1.49479830331
AS36647	1	0	2.64950842419	2	0	2.6495121522
AS36692	1	0	1.90835667691	2	0	1.90841632512
AS49544	1	0	1.71134721447	2	0	1.71134982408
AS55967	1	0	2.25354329006	2	0	2.25354384926
AS60781	2	0	1.82534314497	4	0	1.82534351777
AS197902	1	0	1.77303725702	2	0	1.77303744342
AS393406	2	0	2.14028867497	4	0	2.14529837899
		Mean	2.0003648877		**Mean**	2.0009212315

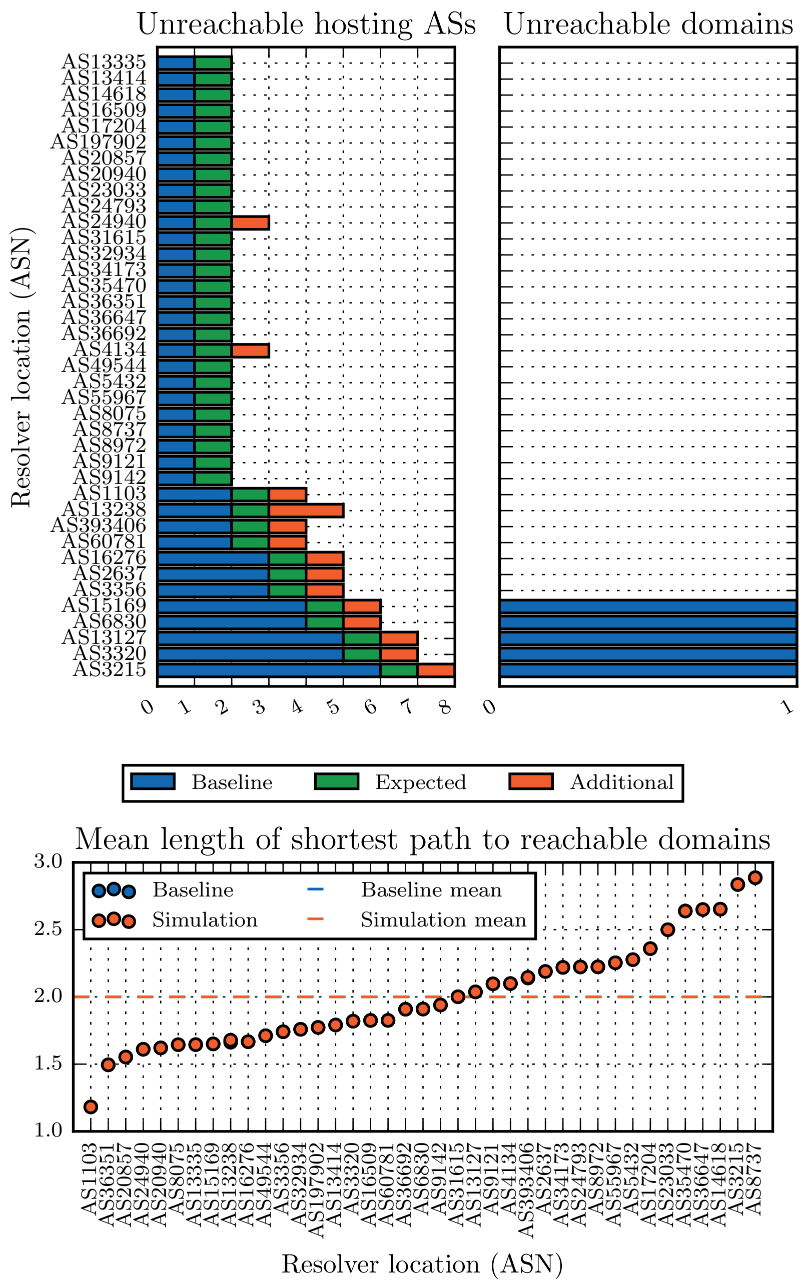


Figure B.35: Plots showing the reachability of autonomous systems and domains in a simulation where AS9002 was removed from the AS topology compared to the baseline.

Table B.36: Reachability of DNS data in a simulation where AS8220 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	13	565	1.1820259896
AS2637	3	0	2.18911427628	14	565	2.18915246439
AS3215	6	1	2.83567157466	17	566	2.83612870608
AS3320	5	1	1.81932553893	16	566	1.81976379799
AS3356	3	0	1.74099517073	14	565	1.74148203757
AS4134	1	0	2.09931352367	12	565	2.09998335267
AS5432	1	0	2.2768154119	12	565	2.27827646241
AS6830	4	1	1.90947226796	15	566	1.9099269195
AS8075	1	0	1.64468828964	12	565	1.64517768929
AS8737	1	0	2.88639271487	12	565	2.88988545033
AS8972	1	0	2.22328412605	12	565	2.22332554034
AS9121	1	0	2.09727243649	12	565	2.0979989087
AS9142	1	0	1.93983546041	12	565	1.94116426554
AS13127	5	1	2.03705757563	16	566	2.03707042699
AS13238	2	0	1.66266719207	13	565	1.66335758226
AS13414	1	0	1.79106909723	12	565	1.79175381038
AS13335	1	0	1.64517926897	12	565	1.64576379468
AS14618	1	0	2.65234637417	12	565	2.65296558327
AS15169	4	1	1.65053505386	15	566	1.65109460421
AS16276	3	0	1.66660210991	14	565	1.66715440428
AS16509	1	0	1.82501377501	12	565	1.82565378061
AS17204	1	0	2.35903972347	12	565	2.35942185849
AS20857	1	0	1.55288727159	12	565	1.55365353006
AS20940	1	0	1.62063850426	12	565	1.62114830051
AS23033	1	0	2.49950622466	12	565	2.49944996694
AS24793	1	0	2.22269584558	12	565	2.22308375323
AS24940	1	0	1.61063288987	12	565	1.61122496212
AS31615	1	0	2.00108969823	12	565	2.00111684395
AS32934	1	0	1.75897836038	12	565	1.75948576336
AS34173	1	0	2.21939263956	12	565	2.21943308472
AS35470	1	0	2.638705947	12	565	2.63943948639
AS36351	1	0	1.4947956937	12	565	1.49531199579
AS36647	1	0	2.64950842419	12	565	2.65016182213
AS36692	1	0	1.90835667691	12	565	1.90923513061
AS49544	1	0	1.71134721447	12	565	1.71187849573
AS55967	1	0	2.25354329006	12	565	2.25391207636
AS60781	2	0	1.82534314497	13	565	1.82534059453
AS197902	1	0	1.77303725702	12	565	1.77315633597
AS393406	2	0	2.14028867497	13	565	2.14064087194
		Mean	2.0003648877		**Mean**	2.0009282165

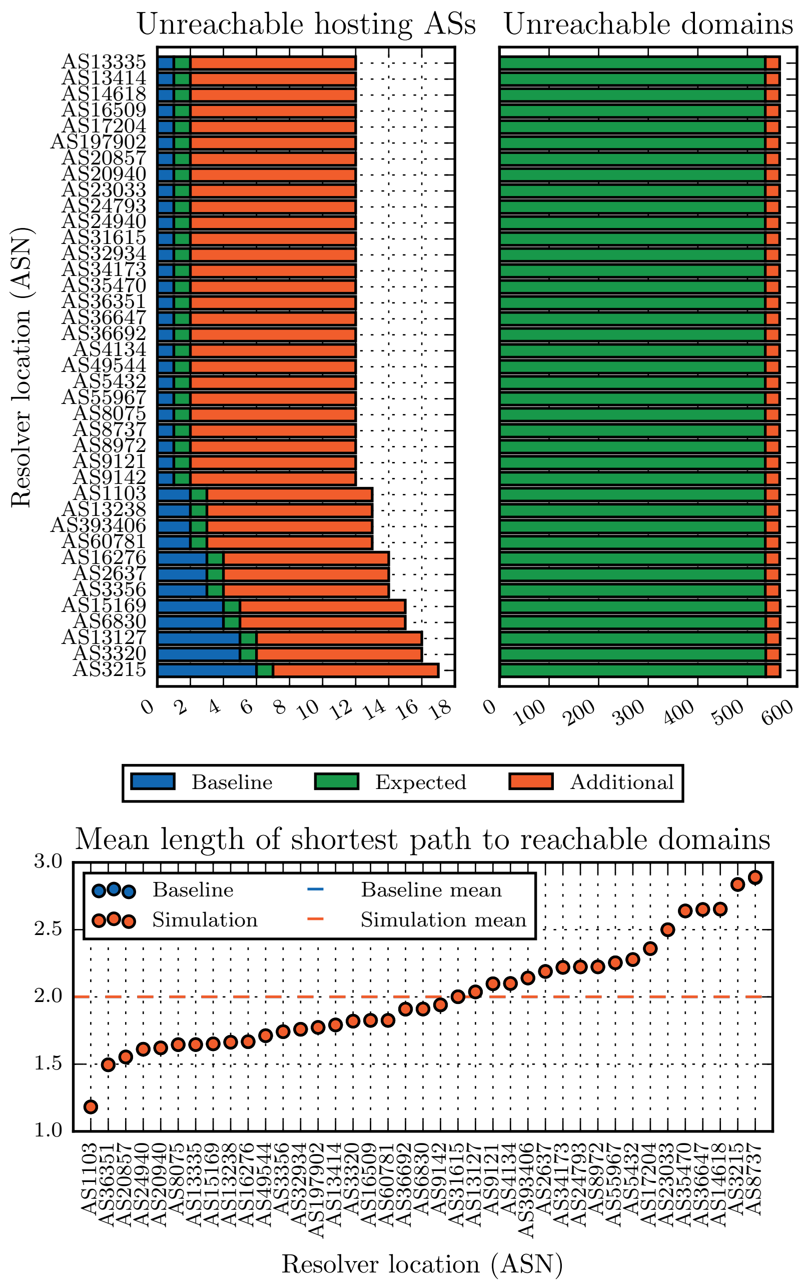


Figure B.36: Plots showing the reachability of autonomous systems and domains in a simulation where AS8220 was removed from the AS topology compared to the baseline.

Table B.37: Reachability of DNS data in a simulation where AS1136 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	5	15,040	1.18335760862
AS2637	3	0	2.18911427628	8	75,483	2.16613978585
AS3215	6	1	2.83567157466	11	75,484	2.82280768888
AS3320	5	1	1.81932553893	10	75,484	1.80599318927
AS3356	3	0	1.74099517073	8	75,483	1.72618463106
AS4134	1	0	2.09931352367	6	75,483	2.09021241921
AS5432	1	0	2.2768154119	4	15,040	2.29727119857
AS6830	4	1	1.90947226796	9	75,484	1.89729612819
AS8075	1	0	1.64468828964	4	15,040	1.64419090395
AS8737	1	0	2.88639271487	1,483	1,171,115	6.83081561199
AS8972	1	0	2.22328412605	4	15,040	2.2324169288
AS9121	1	0	2.09727243649	4	15,040	2.09458015593
AS9142	1	0	1.93983546041	4	15,040	1.93665140863
AS13127	5	1	2.03705757563	10	75,484	2.01217778369
AS13238	2	0	1.66266719207	6	75,448	1.6471833915
AS13414	1	0	1.79106909723	4	15,040	1.7874666246
AS13335	1	0	1.64517926897	4	15,040	1.64116664934
AS14618	1	0	2.65234637417	4	15,040	2.65190995912
AS15169	4	1	1.65053505386	9	75,484	1.63464342379
AS16276	3	0	1.66660210991	8	75,483	1.63593988246
AS16509	1	0	1.82501377501	4	15,040	1.82506297493
AS17204	1	0	2.35903972347	4	15,040	2.36849212337
AS20857	1	0	1.55288727159	4	15,040	1.62010958273
AS20940	1	0	1.62063850426	4	15,040	1.62007331934
AS23033	1	0	2.49950622466	4	15,040	2.52122436421
AS24793	1	0	2.22269584558	4	15,040	2.28857303185
AS24940	1	0	1.61063288987	6	75,483	1.59415934608
AS31615	1	0	2.00108969823	4	15,040	2.05832218639
AS32934	1	0	1.75897836038	4	15,040	1.75880209685
AS34173	1	0	2.21939263956	4	15,040	2.22851450199
AS35470	1	0	2.638705947	4	15,040	2.71703040966
AS36351	1	0	1.4947956937	4	15,040	1.51017487179
AS36647	1	0	2.64950842419	4	15,040	2.66036830146
AS36692	1	0	1.90835667691	4	15,040	1.90508412733
AS49544	1	0	1.71134721447	4	15,040	1.71103779094
AS55967	1	0	2.25354329006	4	15,040	2.25134997013
AS60781	2	0	1.82534314497	7	75,483	1.79699620271
AS197902	1	0	1.77303725702	4	15,040	1.84094746145
AS393406	2	0	2.14028867497	7	75,483	2.11643628038
		Mean	2.0003648877		**Mean**	2.1059272902

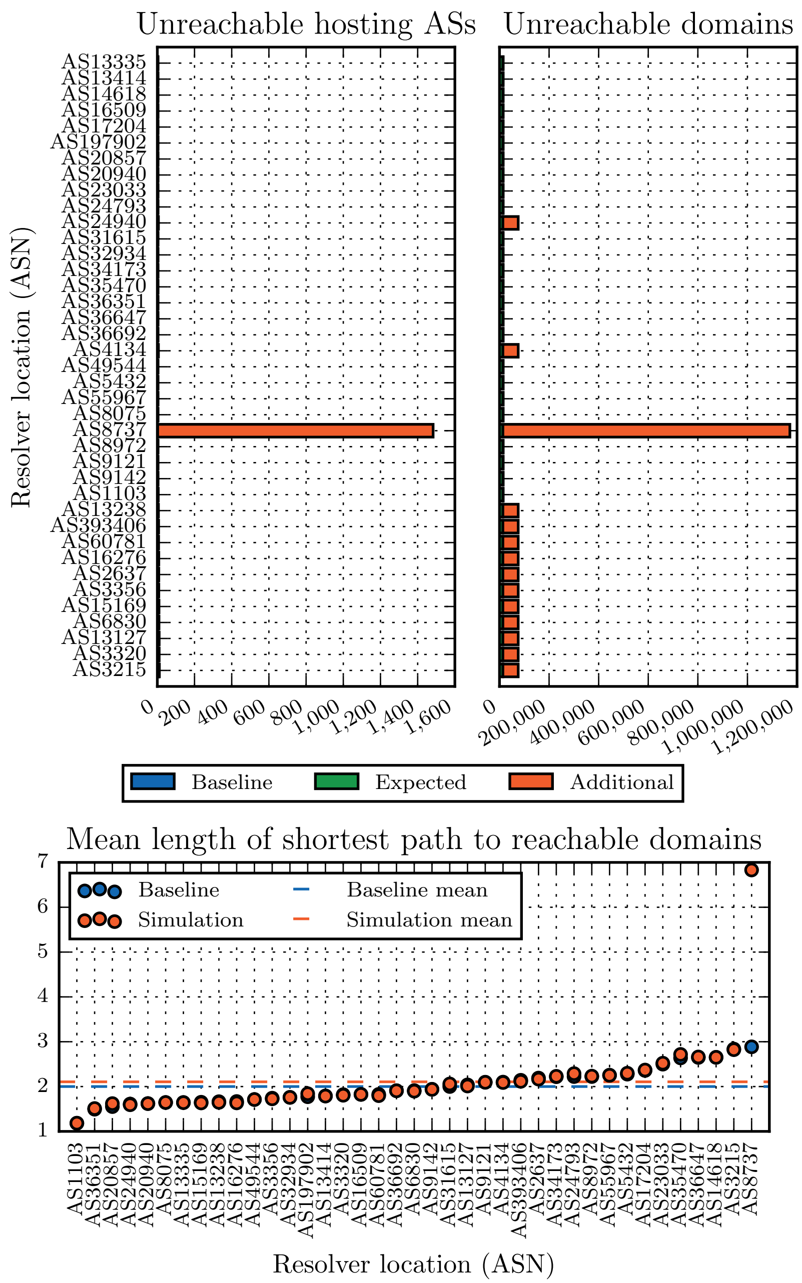


Figure B.37: Plots showing the reachability of autonomous systems and domains in a simulation where AS1136 was removed from the AS topology compared to the baseline.

Table B.38: Reachability of DNS data in a simulation where AS3257 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	4	8	1.18185834275
AS2637	3	0	2.18911427628	4	8	2.18911493109
AS3215	6	1	2.83567157466	7	9	2.83567598964
AS3320	5	1	1.81932553893	6	9	1.81933030233
AS3356	3	0	1.74099517073	4	8	1.74099627571
AS4134	1	0	2.09931352367	3	8	2.09931870459
AS5432	1	0	2.2768154119	2	8	2.27681899351
AS6830	4	1	1.90947226796	6	9	1.90947716579
AS8075	1	0	1.64468828964	2	8	1.64469241982
AS8737	1	0	2.88639271487	2	8	2.88639571427
AS8972	1	0	2.22328412605	2	8	2.22328483181
AS9121	1	0	2.09727243649	2	8	2.09735515715
AS9142	1	0	1.93983546041	2	8	1.93983853951
AS13127	5	1	2.03705757563	6	9	2.03706434133
AS13238	2	0	1.66266719207	3	8	1.66266668903
AS13414	1	0	1.79106909723	3	8	1.79372388057
AS13335	1	0	1.64517926897	2	8	1.64523838815
AS14618	1	0	2.65234637417	2	8	2.6523753071
AS15169	4	1	1.65053505386	5	9	1.65167940748
AS16276	3	0	1.66660210991	4	8	1.66660161274
AS16509	1	0	1.82501377501	2	8	1.82504296542
AS17204	1	0	2.35903972347	3	8	2.35904025887
AS20857	1	0	1.55288727159	2	8	1.55288660486
AS20940	1	0	1.62063850426	2	8	1.62069758685
AS23033	1	0	2.49950622466	2	8	2.49951051115
AS24793	1	0	2.22269584558	2	8	2.22269468645
AS24940	1	0	1.61063288987	3	8	1.61063734207
AS31615	1	0	2.00108969823	2	8	2.00109324147
AS32934	1	0	1.75897836038	2	8	1.75898266099
AS34173	1	0	2.21939263956	2	8	2.21939333952
AS35470	1	0	2.638705947	2	8	2.63870540824
AS36351	1	0	1.4947956937	2	8	1.49488087116
AS36647	1	0	2.64950842419	2	8	2.64950790154
AS36692	1	0	1.90835667691	2	8	1.91074414981
AS49544	1	0	1.71134721447	2	8	1.71134678402
AS55967	1	0	2.25354329006	3	8	2.29409109041
AS60781	2	0	1.82534314497	3	8	1.82534288452
AS197902	1	0	1.77303725702	2	8	1.77309563486
AS393406	2	0	2.14028867497	3	8	2.14028888417
		Mean	2.0003648877		**Mean**	2.0015766616

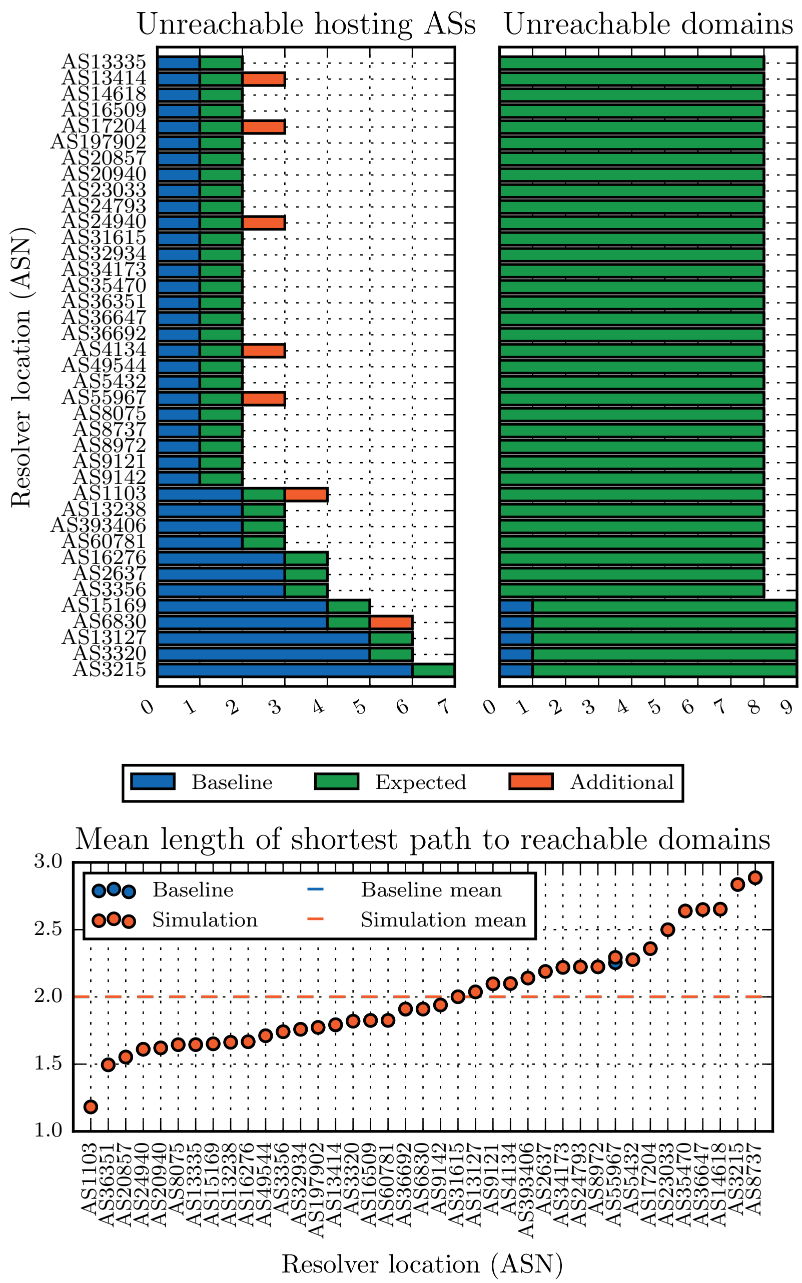


Figure B.38: Plots showing the reachability of autonomous systems and domains in a simulation where AS3257 was removed from the AS topology compared to the baseline.

Table B.39: Reachability of DNS data in a simulation where AS701 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	9	30	1.18178676466
AS2637	3	0	2.18911427628	10	30	2.18911719783
AS3215	6	1	2.83567157466	13	31	2.83577839593
AS3320	5	1	1.81932553893	12	31	1.81942947276
AS3356	3	0	1.74099517073	10	30	1.74109997133
AS4134	1	0	2.09931352367	8	30	2.09942759767
AS5432	1	0	2.2768154119	8	30	2.2768197559
AS6830	4	1	1.90947226796	11	31	1.90948704666
AS8075	1	0	1.64468828964	8	30	1.64470177406
AS8737	1	0	2.88639271487	8	30	2.88640028124
AS8972	1	0	2.22328412605	8	30	2.22328817069
AS9121	1	0	2.09727243649	8	30	2.09727372605
AS9142	1	0	1.93983546041	8	30	1.93984258004
AS13127	5	1	2.03705757563	12	31	2.03705014039
AS13238	2	0	1.66266719207	9	30	1.66266362807
AS13414	1	0	1.79106909723	8	30	1.79106643767
AS13335	1	0	1.64517926897	8	30	1.64517672559
AS14618	1	0	2.65234637417	8	30	2.65234293886
AS15169	4	1	1.65053505386	11	31	1.65053123562
AS16276	3	0	1.66660210991	10	30	1.66659856791
AS16509	1	0	1.82501377501	8	30	1.82501130526
AS17204	1	0	2.35903972347	8	30	2.35913847372
AS20857	1	0	1.55288727159	8	30	1.55288514412
AS20940	1	0	1.62063850426	8	30	1.62067068077
AS23033	1	0	2.49950622466	8	30	2.49960874284
AS24793	1	0	2.22269584558	8	30	2.22269205806
AS24940	1	0	1.61063288987	8	30	1.6106290349
AS31615	1	0	2.00108969823	8	30	2.00109324596
AS32934	1	0	1.75897836038	8	30	1.75899267031
AS34173	1	0	2.21939263956	8	30	2.21939666244
AS35470	1	0	2.638705947	8	30	2.63870262181
AS36351	1	0	1.4947956937	8	30	1.49479081815
AS36647	1	0	2.64950842419	8	30	2.64950478661
AS36692	1	0	1.90835667691	8	30	1.90835579163
AS49544	1	0	1.71134721447	8	30	1.71134392269
AS55967	1	0	2.25354329006	8	30	2.25354675831
AS60781	2	0	1.82534314497	9	30	1.82534291388
AS197902	1	0	1.77303725702	8	30	1.77302834536
AS393406	2	0	2.14028867497	9	30	2.14038303312
		Mean	2.0003648877		**Mean**	2.0003846005

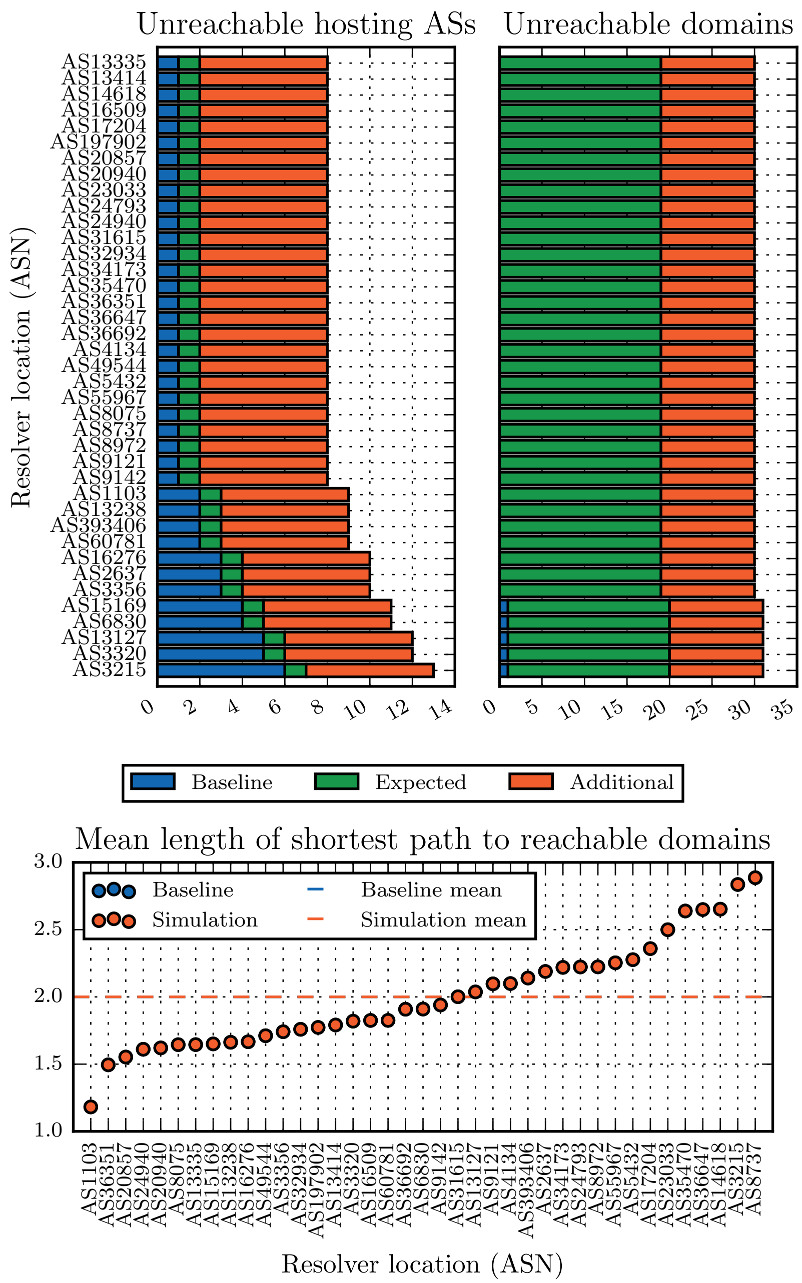


Figure B.39: Plots showing the reachability of autonomous systems and domains in a simulation where AS701 was removed from the AS topology compared to the baseline.

Table B.40: Reachability of DNS data in a simulation where all peerings at the AMS-IX were removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	2.02463676104
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	116	6,096	2.02275294046
AS3356	3	0	1.74099517073	57	1,406	1.78669056204
AS4134	1	0	2.09931352367	1	0	2.1553107038
AS5432	1	0	2.2768154119	1	0	2.3933115344
AS6830	4	1	1.90947226796	4	1	2.20994813028
AS8075	1	0	1.64468828964	1	0	2.11792283311
AS8737	1	0	2.88639271487	1	0	3.25843910328
AS8972	1	0	2.22328412605	1	0	2.22472686712
AS9121	1	0	2.09727243649	1	0	2.13586818342
AS9142	1	0	1.93983546041	1	0	2.86195521612
AS13127	5	1	2.03705757563	5	1	2.61053998975
AS13238	2	0	1.66266719207	2	0	2.13487783674
AS13414	1	0	1.79106909723	1	0	2.17799305397
AS13335	1	0	1.64517926897	1	0	2.08961360635
AS14618	1	0	2.65234637417	1	0	2.94191326852
AS15169	4	1	1.65053505386	4	1	2.44805320323
AS16276	3	0	1.66660210991	3	0	2.2039081507
AS16509	1	0	1.82501377501	1	0	2.14929424984
AS17204	1	0	2.35903972347	1	0	2.3766214061
AS20857	1	0	1.55288727159	1	0	1.7519044555
AS20940	1	0	1.62063850426	1	0	2.00698871978
AS23033	1	0	2.49950622466	1	0	2.49950641106
AS24793	1	0	2.22269584558	1	0	2.29068809429
AS24940	1	0	1.61063288987	1	0	2.07584381713
AS31615	1	0	2.00108969823	1	0	2.1386778005
AS32934	1	0	1.75897836038	1	0	2.09596539509
AS34173	1	0	2.21939263956	1	0	2.30310871557
AS35470	1	0	2.638705947	1	0	2.87775155328
AS36351	1	0	1.4947956937	1	0	1.98667608114
AS36647	1	0	2.64950842419	1	0	3.3126997749
AS36692	1	0	1.90835667691	1	0	2.23113010244
AS49544	1	0	1.71134721447	1	0	2.02118070649
AS55967	1	0	2.25354329006	1	0	2.28961330811
AS60781	2	0	1.82534314497	2	0	1.82792143883
AS197902	1	0	1.77303725702	1	0	2.16821261157
AS393406	2	0	2.14028867497	2	0	2.14029258938
		Mean	2.0003648877		**Mean**	2.2914698725

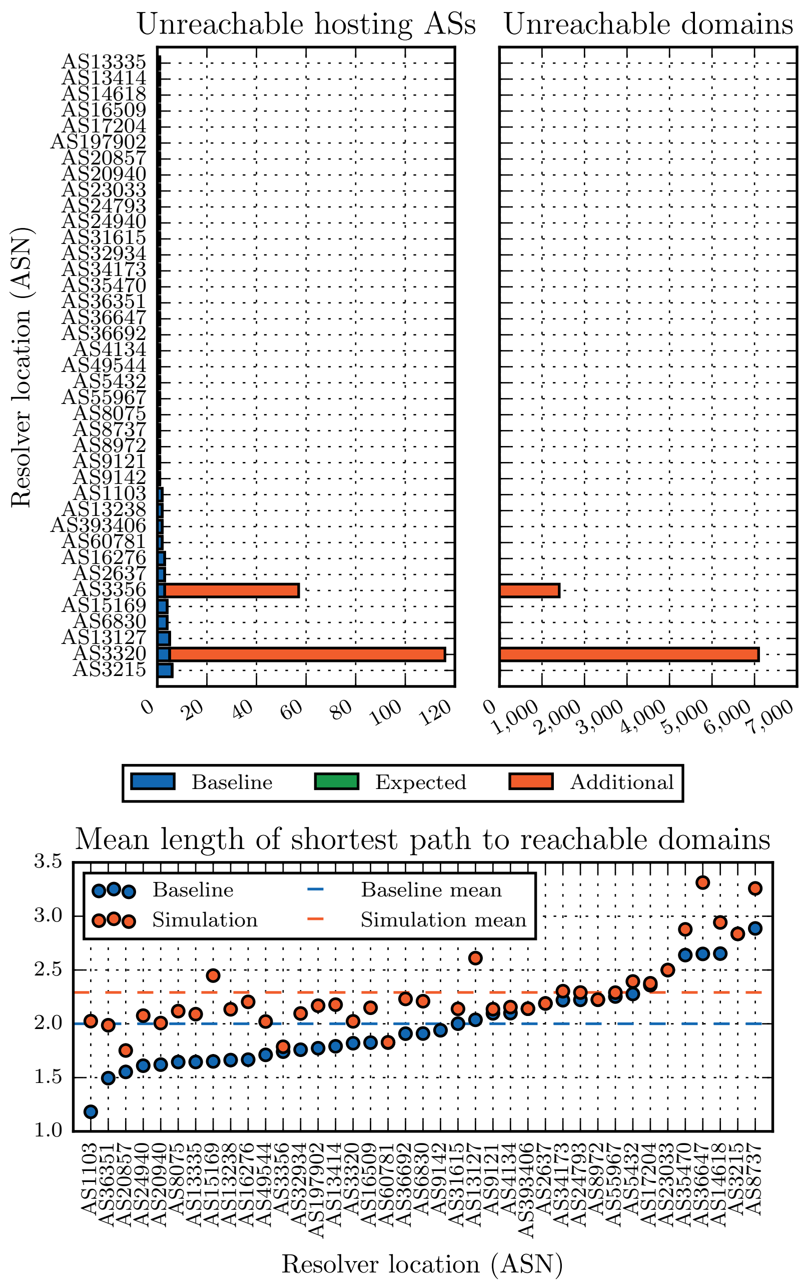


Figure B.40: Plots showing the reachability of autonomous systems and domains in a simulation where AMS-IX was removed from the AS topology compared to the baseline.

Table B.41: Reachability of DNS data in a simulation where the connection AS35470↔AS49685 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	3	42,832	1.17520701035
AS2637	3	0	2.18911427628	4	42,832	2.18258982224
AS3215	6	1	2.83567157466	7	42,833	2.82630086124
AS3320	5	1	1.81932553893	6	42,833	1.80982326983
AS3356	3	0	1.74099517073	4	42,832	1.73891347467
AS4134	1	0	2.09931352367	2	42,832	2.09206464691
AS5432	1	0	2.2768154119	2	42,832	2.26294693154
AS6830	4	1	1.90947226796	5	42,833	1.90069551509
AS8075	1	0	1.64468828964	2	42,832	1.63378051228
AS8737	1	0	2.88639271487	2	42,832	2.87743209452
AS8972	1	0	2.22328412605	2	42,832	2.21703467672
AS9121	1	0	2.09727243649	2	42,832	2.09000713272
AS9142	1	0	1.93983546041	2	42,832	1.93130326519
AS13127	5	1	2.03705757563	6	42,833	2.02125948077
AS13238	2	0	1.66266719207	3	42,832	1.65190411195
AS13414	1	0	1.79106909723	2	42,832	1.78133941731
AS13335	1	0	1.64517926897	2	42,832	1.63427544309
AS14618	1	0	2.65234637417	2	42,832	2.64150023037
AS15169	4	1	1.65053505386	5	42,833	1.6396743302
AS16276	3	0	1.66660210991	4	42,832	1.65587238978
AS16509	1	0	1.82501377501	2	42,832	1.81555728758
AS17204	1	0	2.35903972347	2	42,832	2.34583299824
AS20857	1	0	1.55288727159	2	42,832	1.54124066415
AS20940	1	0	1.62063850426	2	42,832	1.60953717017
AS23033	1	0	2.49950622466	2	42,832	2.48742999754
AS24793	1	0	2.22269584558	2	42,832	2.20839180181
AS24940	1	0	1.61063288987	2	42,832	1.59945102891
AS31615	1	0	2.00108969823	2	42,832	1.99305048745
AS32934	1	0	1.75897836038	2	42,832	1.74899040879
AS34173	1	0	2.21939263956	2	42,832	2.2131128104
AS35470	1	0	2.638705947	3,805	5,158,333	0.0
AS36351	1	0	1.4947956937	2	42,832	1.48268155543
AS36647	1	0	2.64950842419	2	42,832	2.63863944009
AS36692	1	0	1.90835667691	2	42,832	1.89957094722
AS49544	1	0	1.71134721447	2	42,832	1.70097591938
AS55967	1	0	2.25354329006	2	42,832	2.23948751173
AS60781	2	0	1.82534314497	3	42,832	1.81589099947
AS197902	1	0	1.77303725702	2	42,832	1.76316264171
AS393406	2	0	2.14028867497	3	42,832	2.13337126425
		Mean	2.0003648877		**Mean**	1.9230846039

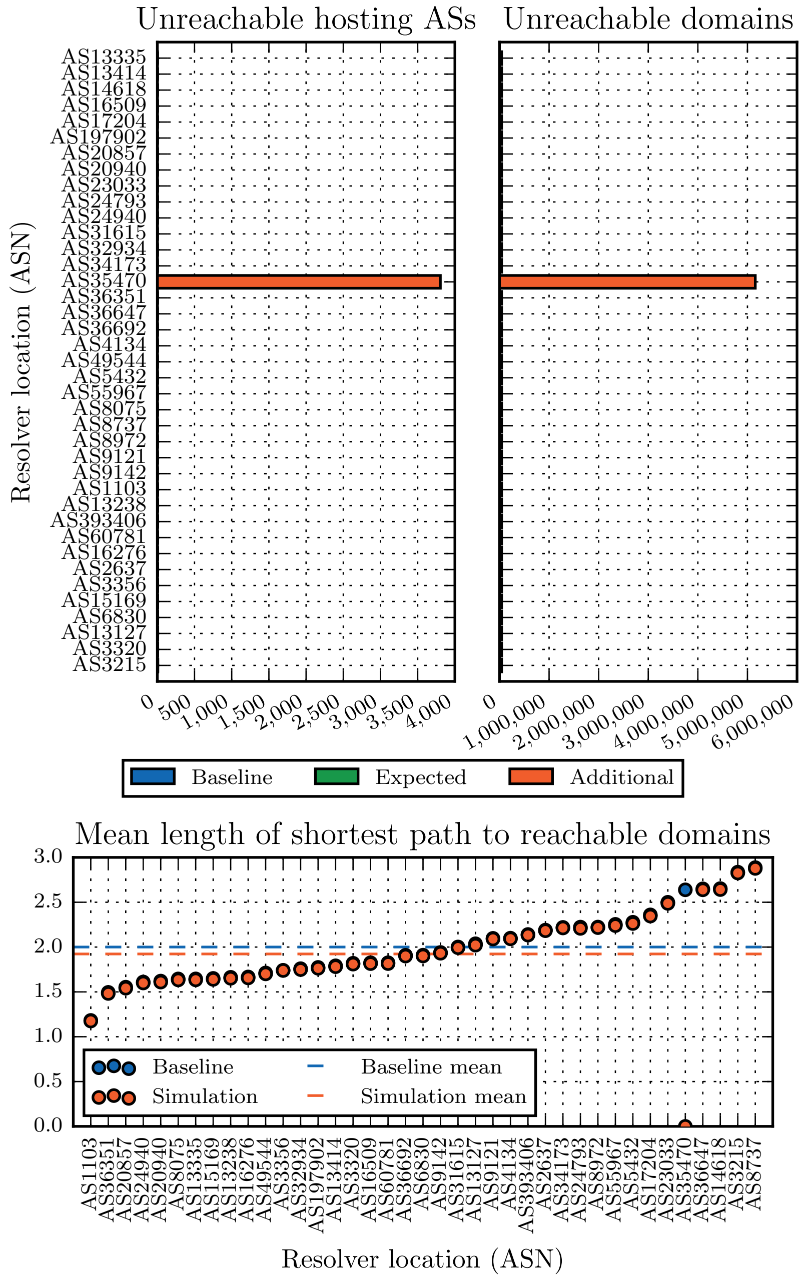
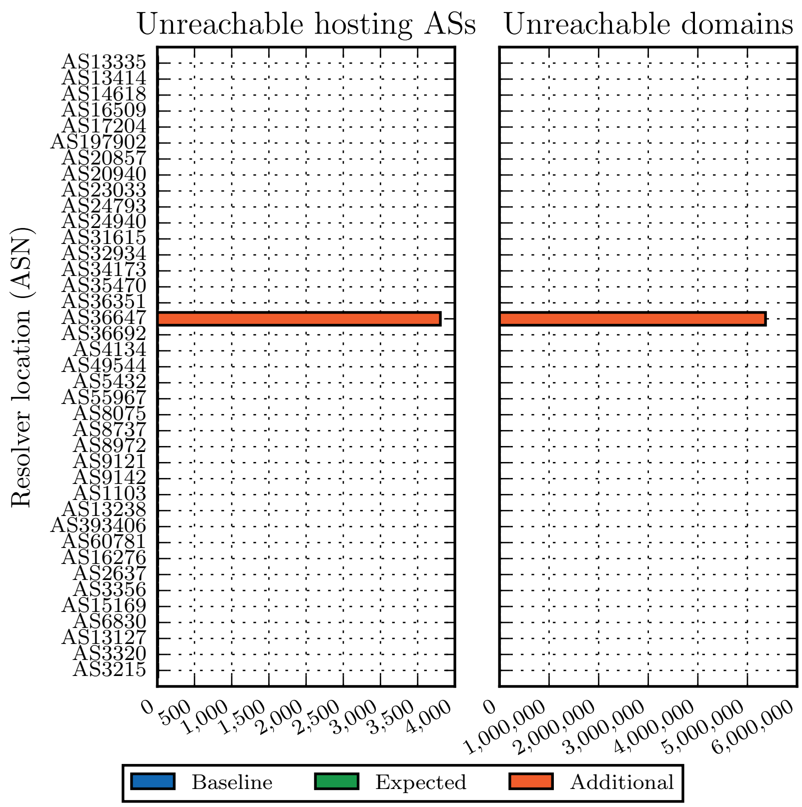


Figure B.41: Plots showing the reachability of autonomous systems and domains in a simulation where AS35470↔AS49685 was removed from the AS topology compared to the baseline.

Table B.42: Reachability of DNS data in a simulation where the connection AS10310↔AS36647 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	0	1.18179189933
AS2637	3	0	2.18911427628	4	0	2.18911427628
AS3215	6	1	2.83567157466	7	1	2.83567157466
AS3320	5	1	1.81932553893	6	1	1.81932553893
AS3356	3	0	1.74099517073	4	0	1.74099517073
AS4134	1	0	2.09931352367	2	0	2.09931352367
AS5432	1	0	2.2768154119	2	0	2.2768154119
AS6830	4	1	1.90947226796	5	1	1.90947226796
AS8075	1	0	1.64468828964	2	0	1.64468828964
AS8737	1	0	2.88639271487	2	0	2.88639271487
AS8972	1	0	2.22328412605	2	0	2.22328412605
AS9121	1	0	2.09727243649	2	0	2.09727243649
AS9142	1	0	1.93983546041	2	0	1.93983546041
AS13127	5	1	2.03705757563	6	1	2.03705757563
AS13238	2	0	1.66266719207	3	0	1.66266719207
AS13414	1	0	1.79106909723	2	0	1.79106909723
AS13335	1	0	1.64517926897	2	0	1.64517926897
AS14618	1	0	2.65234637417	2	0	2.65234637417
AS15169	4	1	1.65053505386	5	1	1.65053505386
AS16276	3	0	1.66660210991	4	0	1.66660210991
AS16509	1	0	1.82501377501	2	0	1.82501377501
AS17204	1	0	2.35903972347	2	0	2.35903972347
AS20857	1	0	1.55288727159	2	0	1.55288727159
AS20940	1	0	1.62063850426	2	0	1.62063850426
AS23033	1	0	2.49950622466	2	0	2.49950622466
AS24793	1	0	2.22269584558	2	0	2.22269584558
AS24940	1	0	1.61063288987	2	0	1.61063288987
AS31615	1	0	2.00108969823	2	0	2.00108969823
AS32934	1	0	1.75897836038	2	0	1.75897836038
AS34173	1	0	2.21939263956	2	0	2.21939263956
AS35470	1	0	2.638705947	2	0	2.638705947
AS36351	1	0	1.4947956937	2	0	1.4947956937
AS36647	1	0	2.64950842419	3,805	5,364,786	0.0
AS36692	1	0	1.90835667691	2	0	1.90835667691
AS49544	1	0	1.71134721447	2	0	1.71134721447
AS55967	1	0	2.25354329006	2	0	2.25354329006
AS60781	2	0	1.82534314497	3	0	1.82534314497
AS197902	1	0	1.77303725702	2	0	1.77303725702
AS393406	2	0	2.14028867497	3	0	2.14028867497
		Mean	2.0003648877		**Mean**	1.9324287742



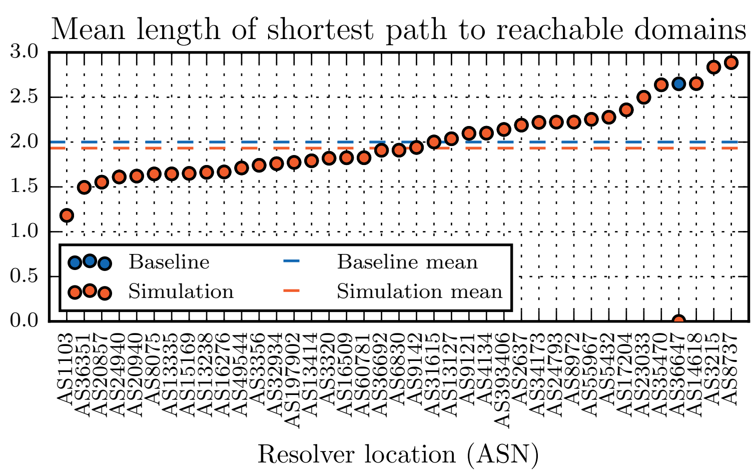


Figure B.42: Plots showing the reachability of autonomous systems and domains in a simulation where AS10310↔AS36647 was removed from the AS topology compared to the baseline.

Table B.43: Reachability of DNS data in a simulation where the connection AS1299↔AS23033 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911595388
AS3215	6	1	2.83567157466	6	1	2.83567325227
AS3320	5	1	1.81932553893	5	1	1.81932721653
AS3356	3	0	1.74099517073	3	0	1.74099684834
AS4134	1	0	2.09931352367	1	0	2.09931520127
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947394556
AS8075	1	0	1.64468828964	1	0	1.64468996725
AS8737	1	0	2.88639271487	1	0	2.88639439247
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727411409
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03714331995
AS13238	2	0	1.66266719207	2	0	1.66266886967
AS13414	1	0	1.79106909723	1	0	1.79107077484
AS13335	1	0	1.64517926897	1	0	1.64518094657
AS14618	1	0	2.65234637417	1	0	2.65234805178
AS15169	4	1	1.65053505386	4	1	1.65053673147
AS16276	3	0	1.66660210991	3	0	1.66660378751
AS16509	1	0	1.82501377501	1	0	1.82501545261
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62064018187
AS23033	1	0	2.49950622466	1	0	3.08039068832
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00109137584
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.6495101018
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534482257
AS197902	1	0	1.77303725702	1	0	1.77303893462
AS393406	2	0	2.14028867497	2	0	2.14037069126
		Mean	2.0003648877		**Mean**	2.015264566

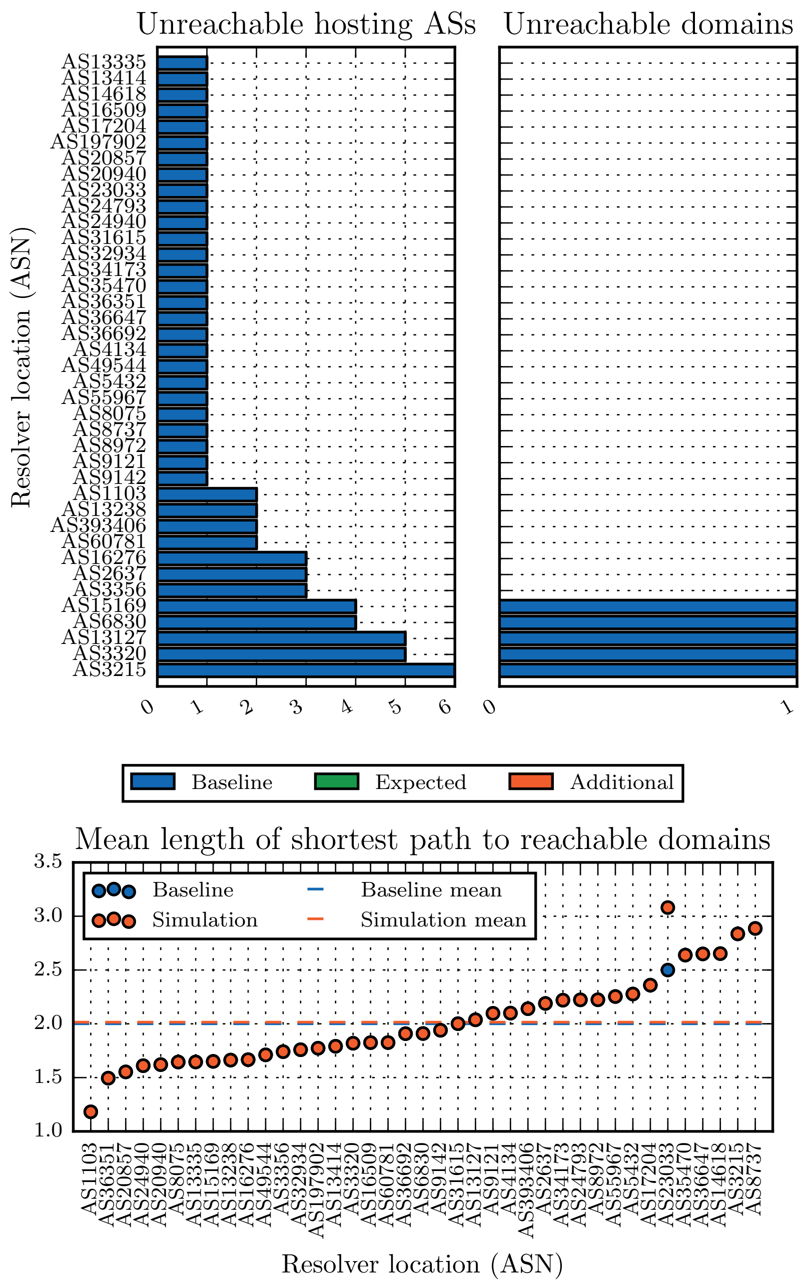


Figure B.43: Plots showing the reachability of autonomous systems and domains in a simulation where AS1299↔AS23033 was removed from the AS topology compared to the baseline.

Table B.44: Reachability of DNS data in a simulation where the connection AS14618↔AS16509 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	3	301	1.18174822681
AS2637	3	0	2.18911427628	4	301	2.18906877769
AS3215	6	1	2.83567157466	7	302	2.83560624448
AS3320	5	1	1.81932553893	6	302	1.8193167062
AS3356	3	0	1.74099517073	4	301	1.74098063804
AS4134	1	0	2.09931352367	2	301	2.09926298638
AS5432	1	0	2.2768154119	1	0	2.2769276251
AS6830	4	1	1.90947226796	5	302	1.90946830694
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88650604646
AS8972	1	0	2.22328412605	1	0	2.22356465903
AS9121	1	0	2.09727243649	1	0	2.09744075628
AS9142	1	0	1.93983546041	1	0	1.94000378021
AS13127	5	1	2.03705757563	6	302	2.03700373158
AS13238	2	0	1.66266719207	3	301	1.66259215466
AS13414	1	0	1.79106909723	2	301	1.79100126443
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1,496	991,818	4.50091402411
AS15169	4	1	1.65053505386	5	302	1.65045933571
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82511704097
AS17204	1	0	2.35903972347	1	0	2.35926414986
AS20857	1	0	1.55288727159	1	0	1.55305559139
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49967454446
AS24793	1	0	2.22269584558	1	0	2.22275642579
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00120191143
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21944986456
AS35470	1	0	2.638705947	1	0	2.63893037339
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.9084688901
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	2	301	2.25350140657
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	3	301	2.14030288451
		Mean	2.0003648877		**Mean**	2.047803577

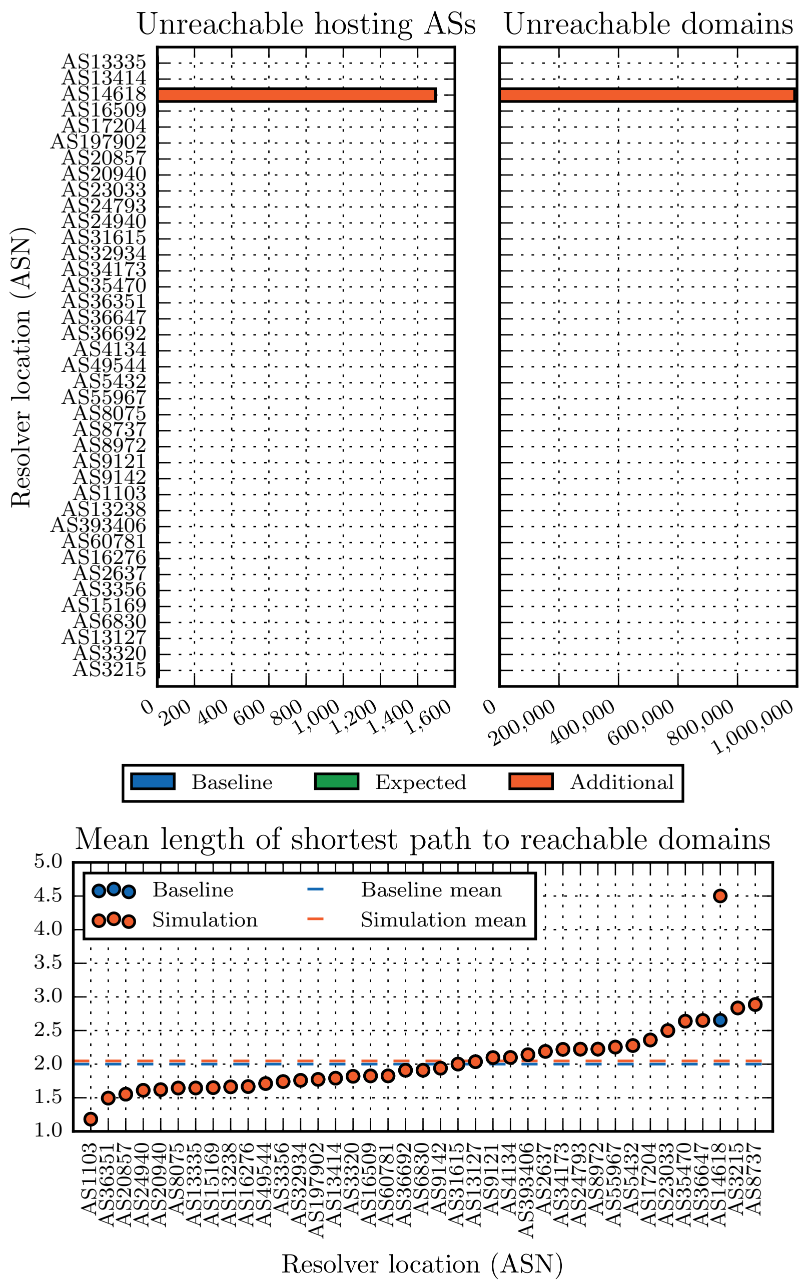


Figure B.44: Plots showing the reachability of autonomous systems and domains in a simulation where AS14618↔AS16509 was removed from the AS topology compared to the baseline.

Table B.45: Reachability of DNS data in a simulation where the connection AS3215↔AS5511 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179730494
AS2637	3	0	2.18911427628	3	0	2.1891196819
AS3215	6	1	2.83567157466	6	1	3.64959261197
AS3320	5	1	1.81932553893	5	1	1.81933299495
AS3356	3	0	1.74099517073	3	0	1.74100188116
AS4134	1	0	2.09931352367	1	0	2.0993219117
AS5432	1	0	2.2768154119	1	0	2.27682100392
AS6830	4	1	1.90947226796	4	1	1.90947785998
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639830689
AS8972	1	0	2.22328412605	1	0	2.22328953166
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03706316765
AS13238	2	0	1.66266719207	2	0	1.66267259769
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65054064588
AS16276	3	0	1.66660210991	3	0	1.66660882033
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35904643389
AS20857	1	0	1.55288727159	1	0	1.55289267721
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49951181668
AS24793	1	0	2.22269584558	1	0	2.2227012512
AS24940	1	0	1.61063288987	1	0	1.61063829549
AS31615	1	0	2.00108969823	1	0	2.00109529025
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.63871135262
AS36351	1	0	1.4947956937	1	0	1.49480128572
AS36647	1	0	2.64950842419	1	0	2.64951401621
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354888208
AS60781	2	0	1.82534314497	2	0	1.825351533
AS197902	1	0	1.77303725702	1	0	1.77304266264
AS393406	2	0	2.14028867497	2	0	2.14029426699
		Mean	2.0003648877		**Mean**	2.0212386201

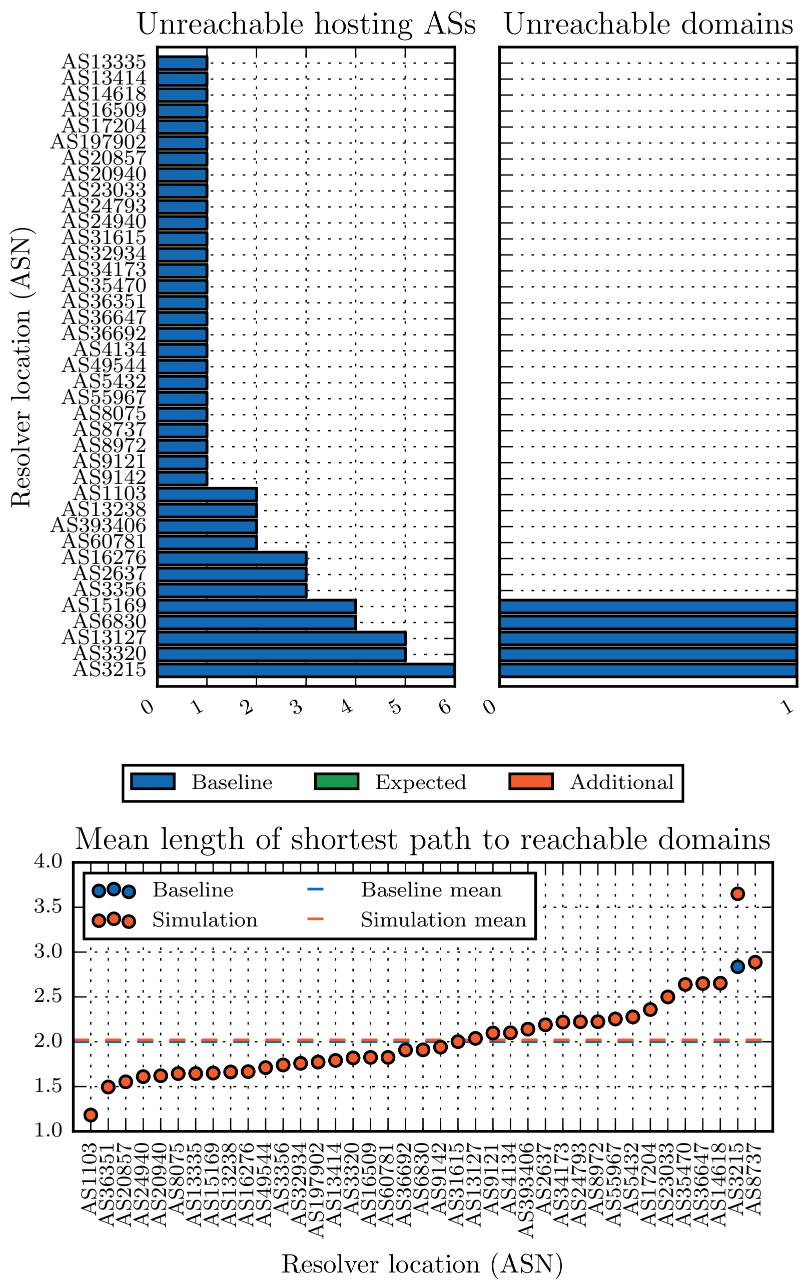


Figure B.45: Plots showing the reachability of autonomous systems and domains in a simulation where AS3215↔AS5511 was removed from the AS topology compared to the baseline.

Table B.46: Reachability of DNS data in a simulation where the connection AS17204↔AS2914 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.95412027465
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0156233633

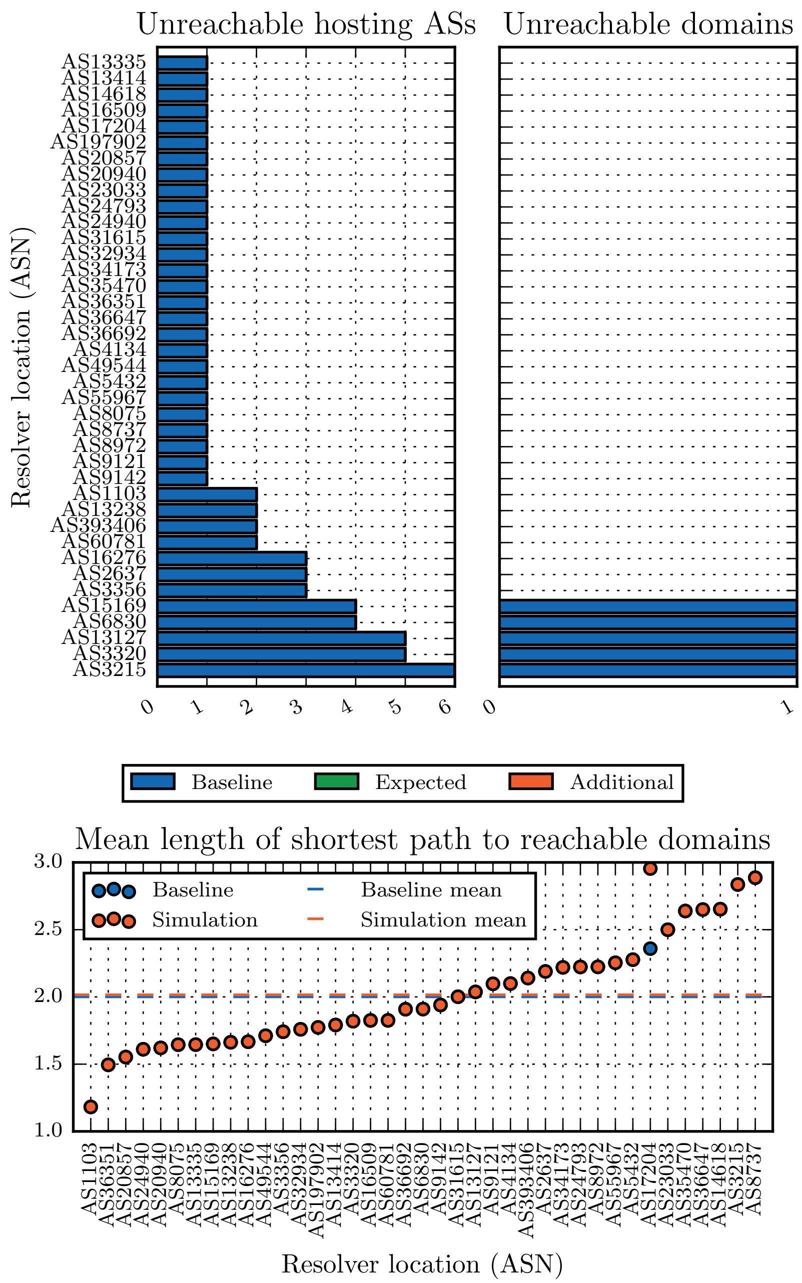


Figure B.46: Plots showing the reachability of autonomous systems and domains in a simulation where AS17204↔AS2914 was removed from the AS topology compared to the baseline.

Table B.47: Reachability of DNS data in a simulation where the connection AS2914↔AS5432 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81940960564
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09940448719
AS5432	1	0	2.2768154119	1	0	2.7408719599
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.0973604176
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65243044087
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.8250978417
AS17204	1	0	2.35903972347	1	0	2.35914392144
AS20857	1	0	1.55288727159	1	0	1.55298233593
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22278401309
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75906242707
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90844521722
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.2536391
AS60781	2	0	1.82534314497	2	0	1.82543951411
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.012291552

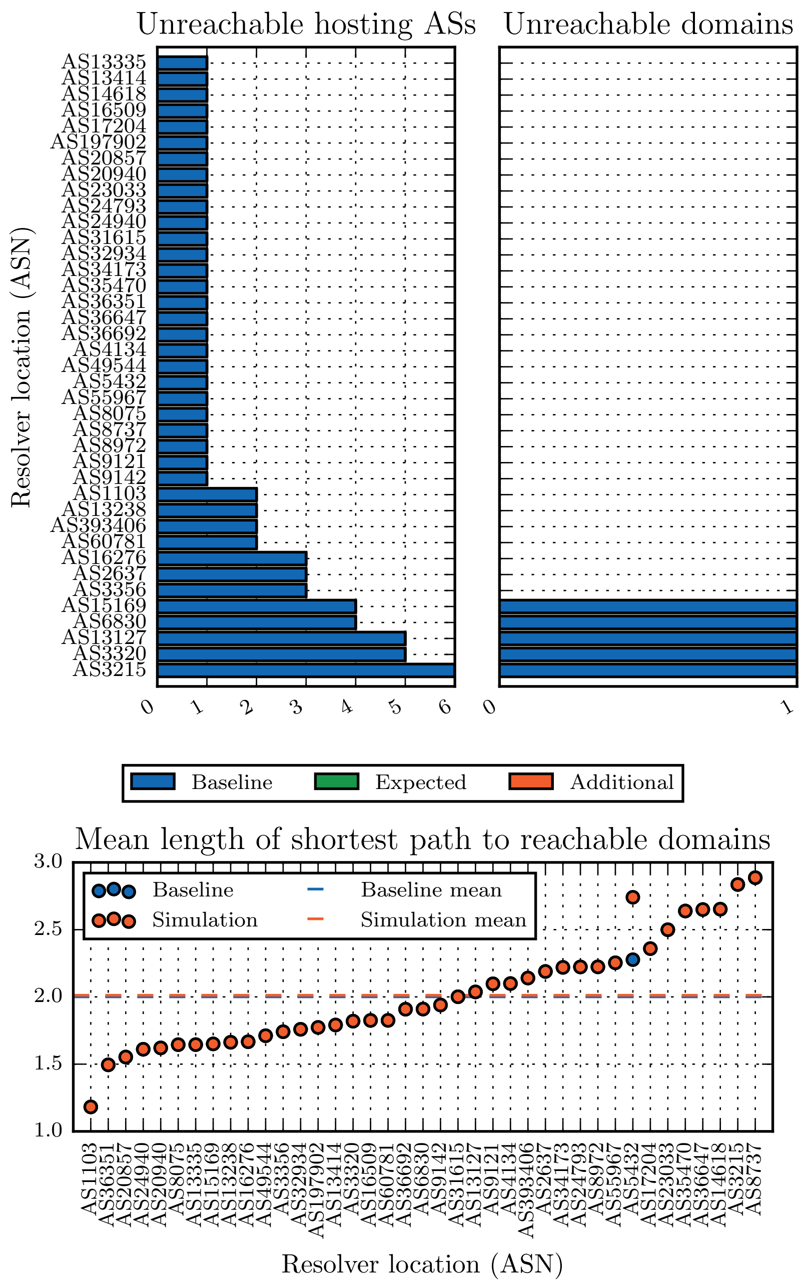


Figure B.47: Plots showing the reachability of autonomous systems and domains in a simulation where AS2914↔AS5432 was removed from the AS topology compared to the baseline.

Table B.48: Reachability of DNS data in a simulation where the connection AS1136↔AS8737 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.20037436708
AS3215	6	1	2.83567157466	6	1	2.84693166756
AS3320	5	1	1.81932553893	5	1	1.83058563182
AS3356	3	0	1.74099517073	3	0	1.75225526153
AS4134	1	0	2.09931352367	1	0	2.11057361447
AS5432	1	0	2.2768154119	1	0	2.2880755027
AS6830	4	1	1.90947226796	4	1	1.92073236086
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	3.67296135467
AS8972	1	0	2.22328412605	1	0	2.23454421685
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.04831766853
AS13238	2	0	1.66266719207	2	0	1.67392728287
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.66179514676
AS16276	3	0	1.66660210991	3	0	1.67786220071
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.37029981427
AS20857	1	0	1.55288727159	1	0	1.56414736239
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.51076631546
AS24793	1	0	2.22269584558	1	0	2.23395593638
AS24940	1	0	1.61063288987	1	0	1.62189298067
AS31615	1	0	2.00108969823	1	0	2.01234978903
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.23065273036
AS35470	1	0	2.638705947	1	0	2.6499660378
AS36351	1	0	1.4947956937	1	0	1.5060557845
AS36647	1	0	2.64950842419	1	0	2.66076851499
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.83660323577
AS197902	1	0	1.77303725702	1	0	1.78429734782
AS393406	2	0	2.14028867497	2	0	2.15154876577
		Mean	2.0003648877		**Mean**	2.0277513215

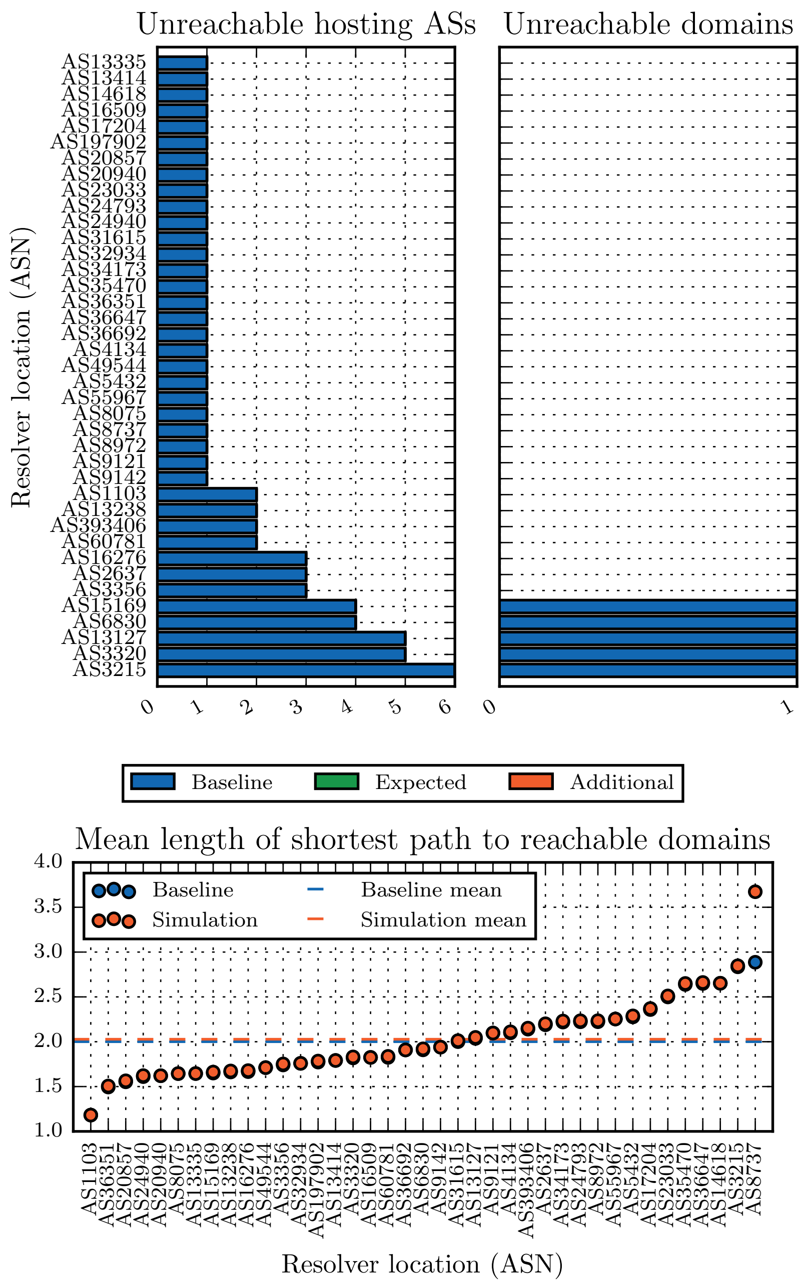


Figure B.48: Plots showing the reachability of autonomous systems and domains in a simulation where AS1136↔AS8737 was removed from the AS topology compared to the baseline.

Table B.49: Reachability of DNS data in a simulation where the connection AS31615↔AS3320 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.0535133914
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0017090849

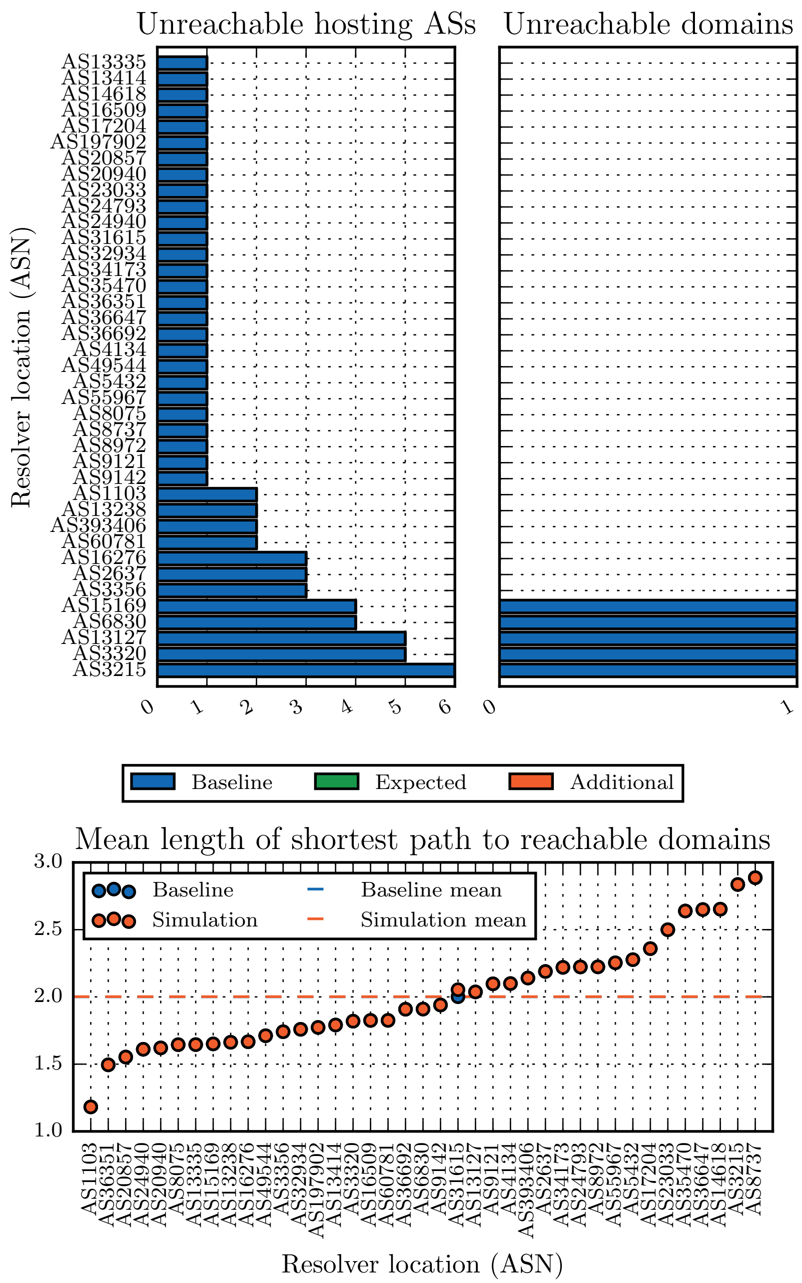


Figure B.49: Plots showing the reachability of autonomous systems and domains in a simulation where AS31615↔AS3320 was removed from the AS topology compared to the baseline.

Table B.50: Reachability of DNS data in a simulation where the connection AS24793↔AS41887 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	4	5,432	2.18829239931
AS3215	6	1	2.83567157466	7	5,433	2.83449146399
AS3320	5	1	1.81932553893	6	5,433	1.81812886066
AS3356	3	0	1.74099517073	4	5,432	1.73971910058
AS4134	1	0	2.09931352367	2	5,432	2.09840062873
AS5432	1	0	2.2768154119	1	0	2.27985299699
AS6830	4	1	1.90947226796	5	5,433	1.90836695834
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.89145535667
AS8972	1	0	2.22328412605	1	0	2.22429665441
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	6	5,433	2.03608158071
AS13238	2	0	1.66266719207	3	5,432	1.66131173223
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	5	5,433	1.64916729718
AS16276	3	0	1.66660210991	4	5,432	1.66525063832
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.36106478019
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.50153128139
AS24793	1	0	2.22269584558	1,492	759,652	2.36826686552
AS24940	1	0	1.61063288987	2	5,432	1.60922469043
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.22040516792
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.65052095255
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25455581842
AS60781	2	0	1.82534314497	3	5,432	1.82415256609
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	3	5,432	2.13941731059
		Mean	2.0003648877		**Mean**	2.0041279058

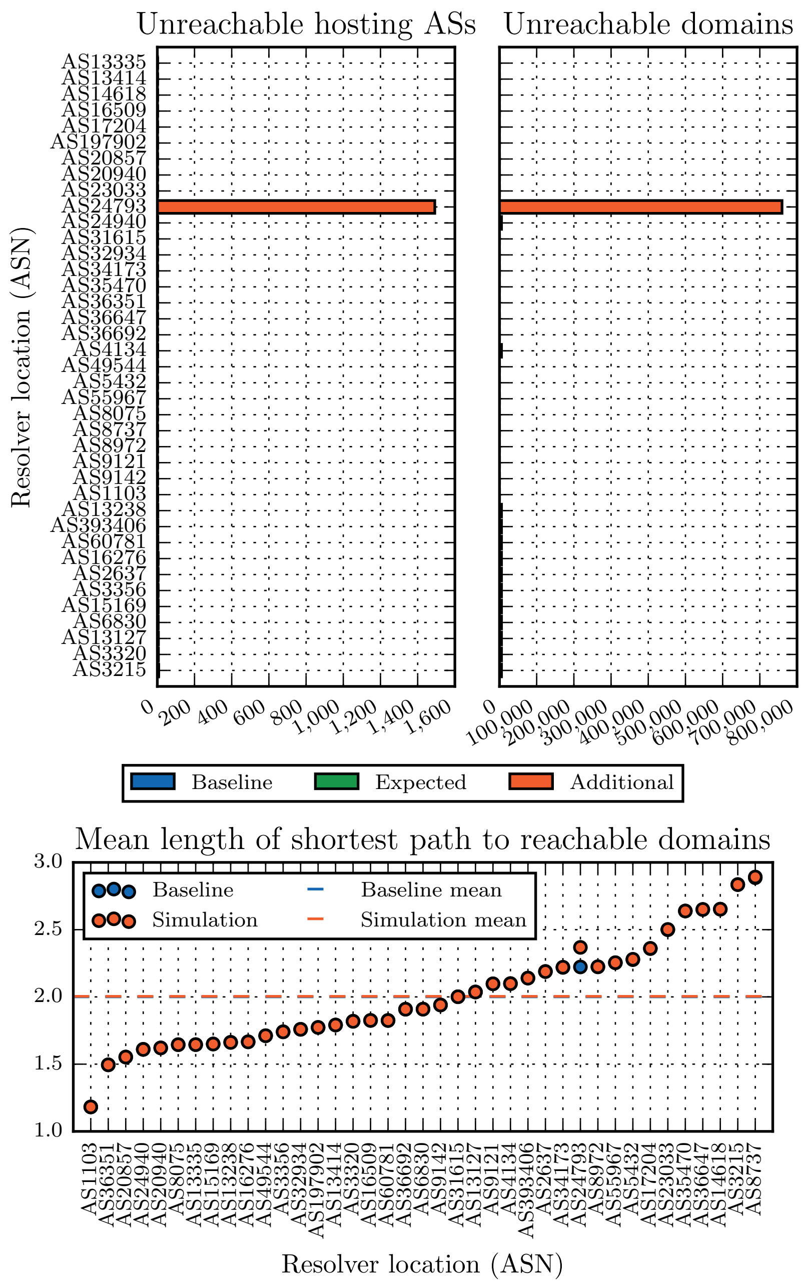


Figure B.50: Plots showing the reachability of autonomous systems and domains in a simulation where AS24793↔AS41887 was removed from the AS topology compared to the baseline.

Table B.51: Reachability of DNS data in a simulation where the connection AS1299↔AS2637 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.28277874168
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0027665406

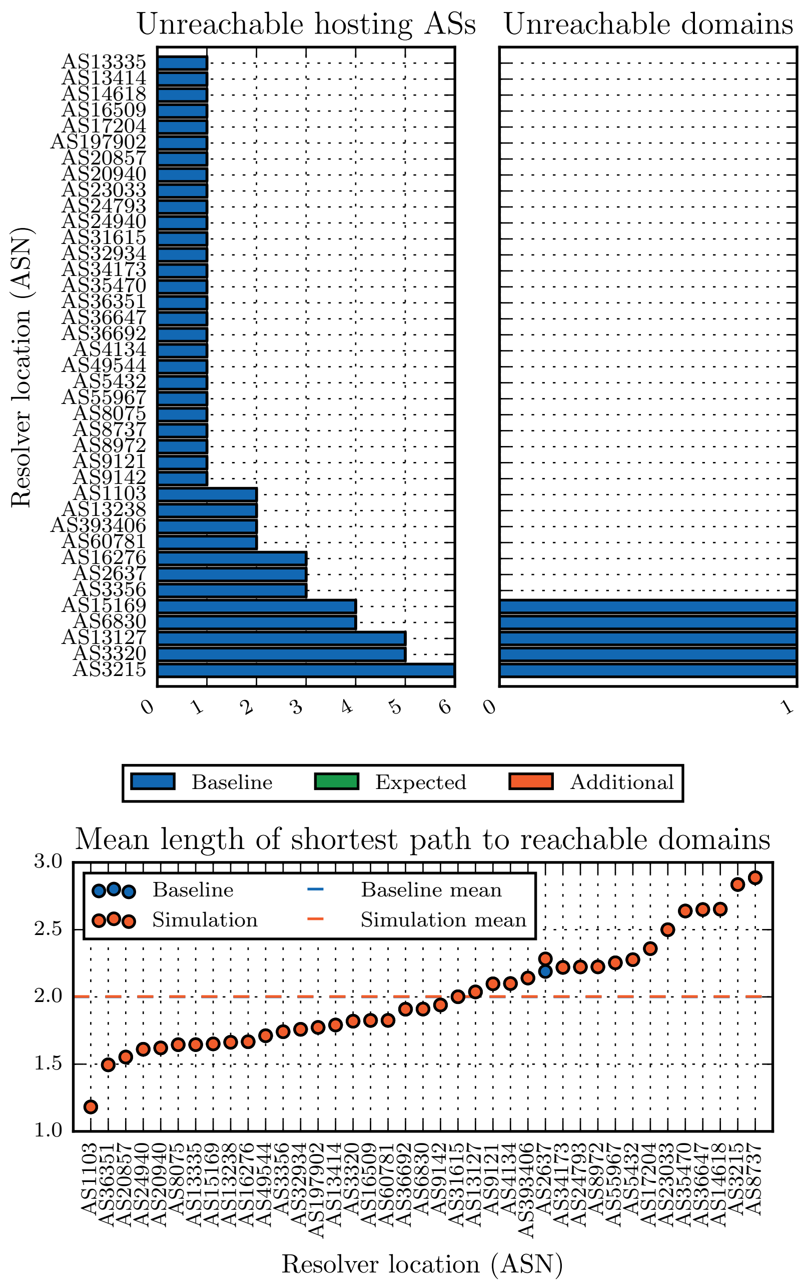


Figure B.51: Plots showing the reachability of autonomous systems and domains in a simulation where AS1299↔AS2637 was removed from the AS topology compared to the baseline.

Table B.52: Reachability of DNS data in a simulation where the connection AS55967↔AS6453 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	**Baseline**			**Simulation**		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.26509379308
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0006610544

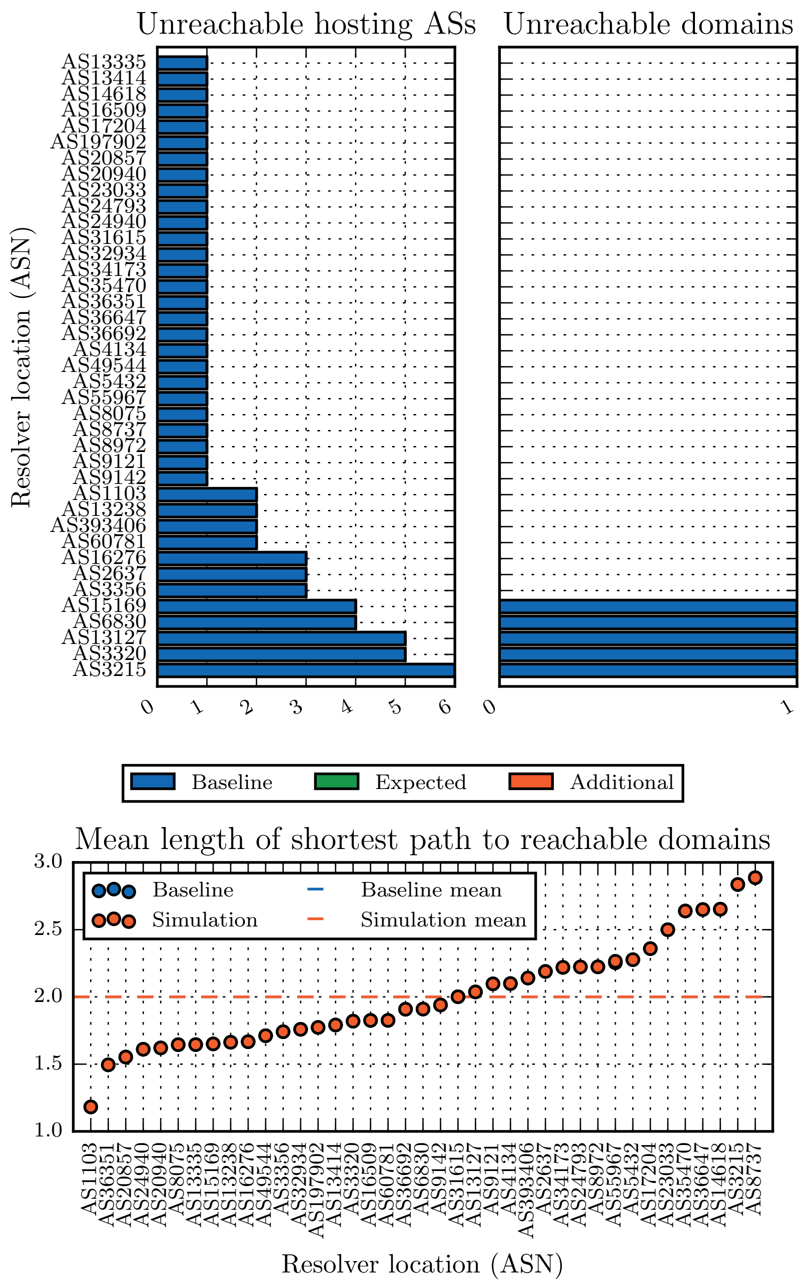


Figure B.52: Plots showing the reachability of autonomous systems and domains in a simulation where AS55967↔AS6453 was removed from the AS topology compared to the baseline.

Table B.53: Reachability of DNS data in a simulation where the connection AS30781↔AS34173 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.28645810422
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.002084515

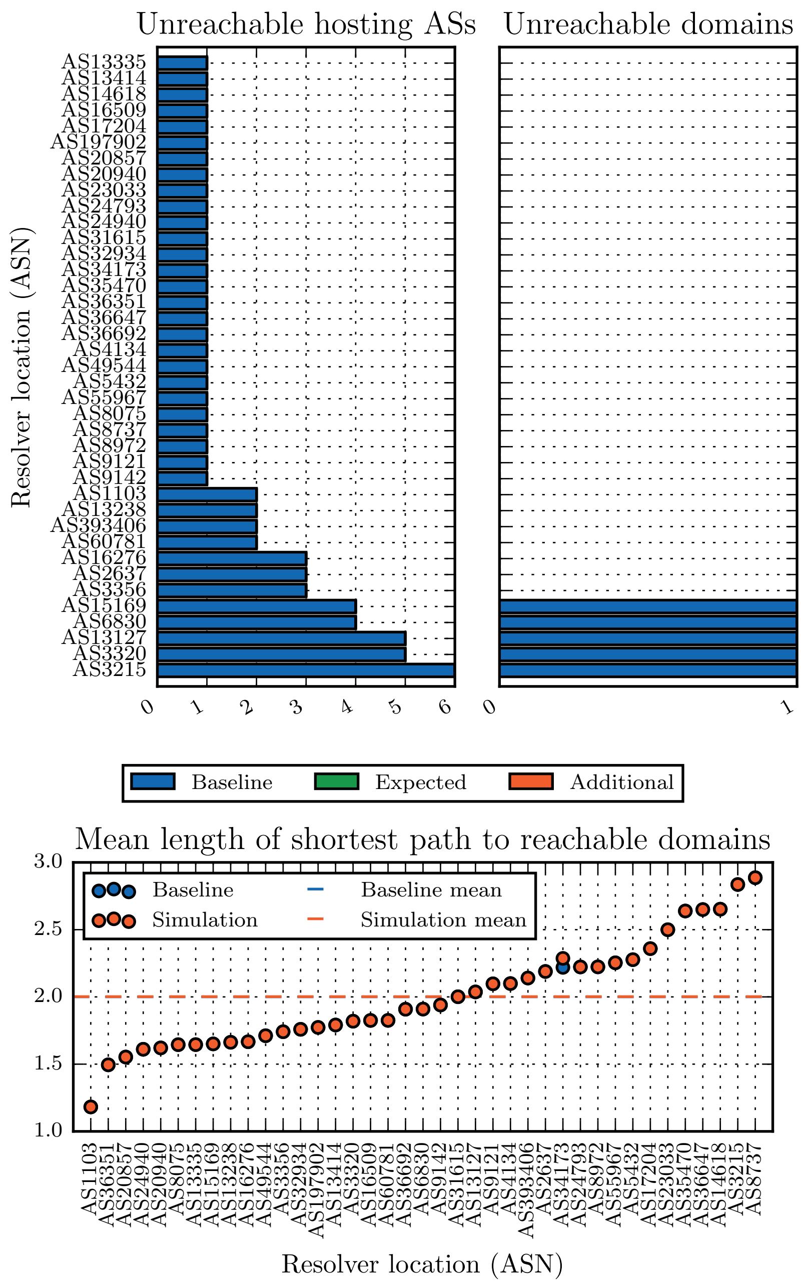


Figure B.53: Plots showing the reachability of autonomous systems and domains in a simulation where AS30781↔AS34173 was removed from the AS topology compared to the baseline.

Table B.54: Reachability of DNS data in a simulation where the connection AS43531↔AS9143 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.887454453
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09811105304
AS9142	1	0	1.93983546041	4	1	2.088895794
AS13127	5	1	2.03705757563	5	1	2.03791781482
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64594910367
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.2251205453
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.49542516871
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14219070726
		Mean	2.0003648877		**Mean**	2.0044045535

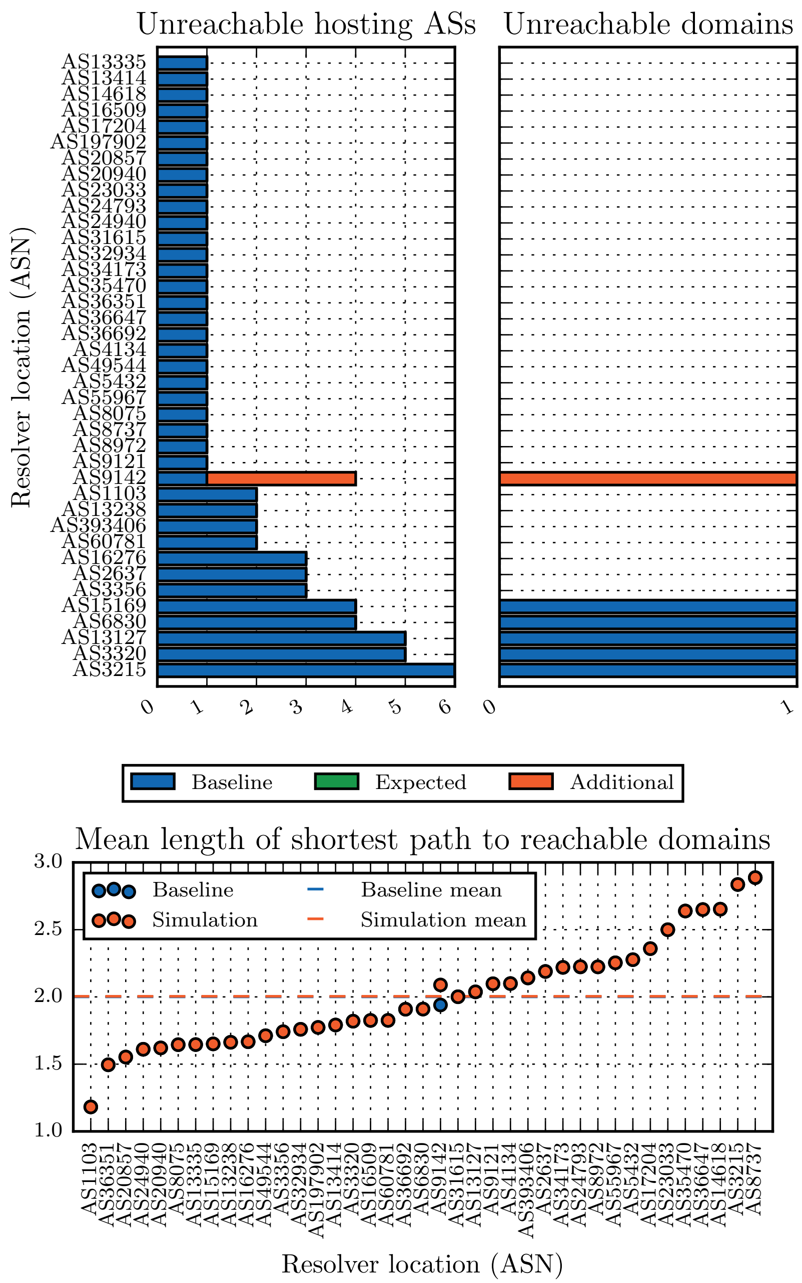


Figure B.54: Plots showing the reachability of autonomous systems and domains in a simulation where AS43531↔AS9143 was removed from the AS topology compared to the baseline.

Table B.55: Reachability of DNS data in a simulation where the connection AS2914↔AS393406 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.3591040317
AS20857	1	0	1.55288727159	1	0	1.55289565962
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22270423361
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	3	0	2.24861001031
		Mean	2.0003648877		**Mean**	2.0031444369

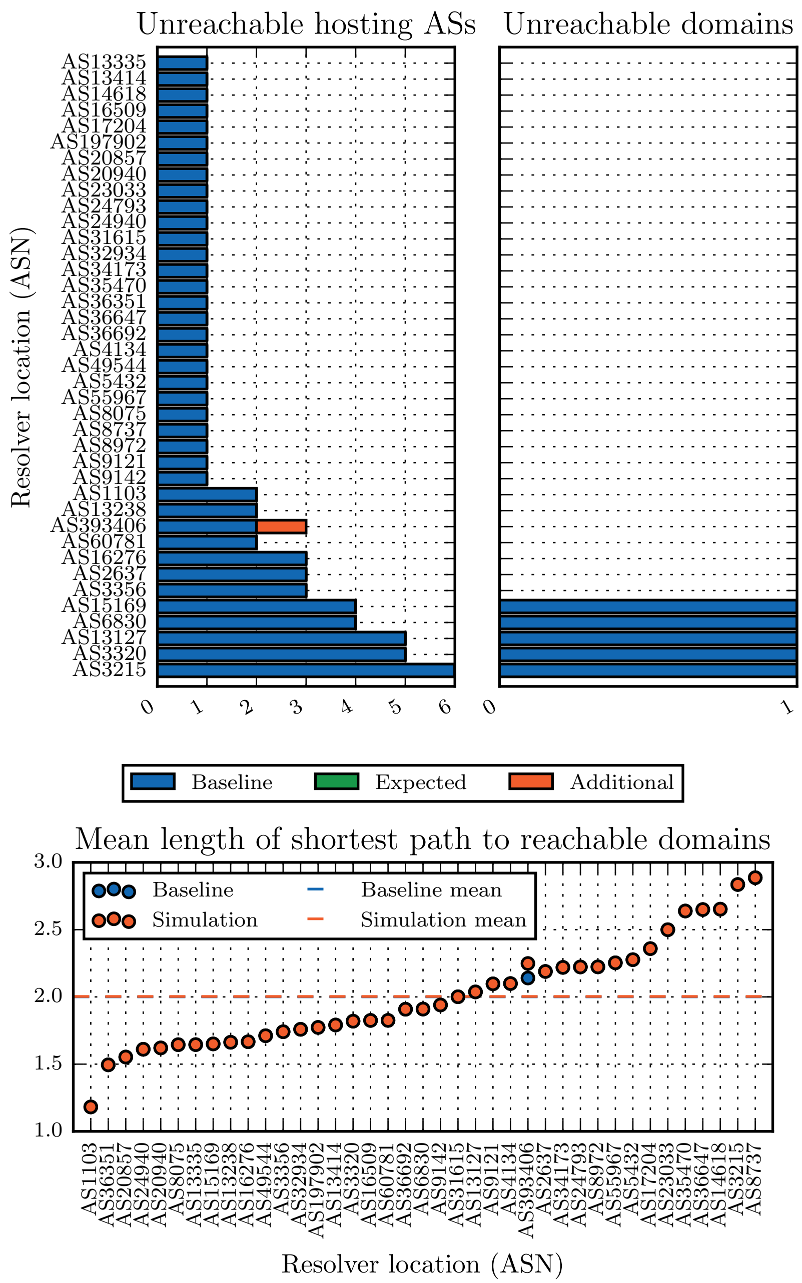


Figure B.55: Plots showing the reachability of autonomous systems and domains in a simulation where AS2914↔AS393406 was removed from the AS topology compared to the baseline.

Table B.56: Reachability of DNS data in a simulation where the connection AS174↔AS2914 was removed from the AS topology compared to the baseline.

	Baseline			Simulation		
Resolver location (ASN)	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0003648877

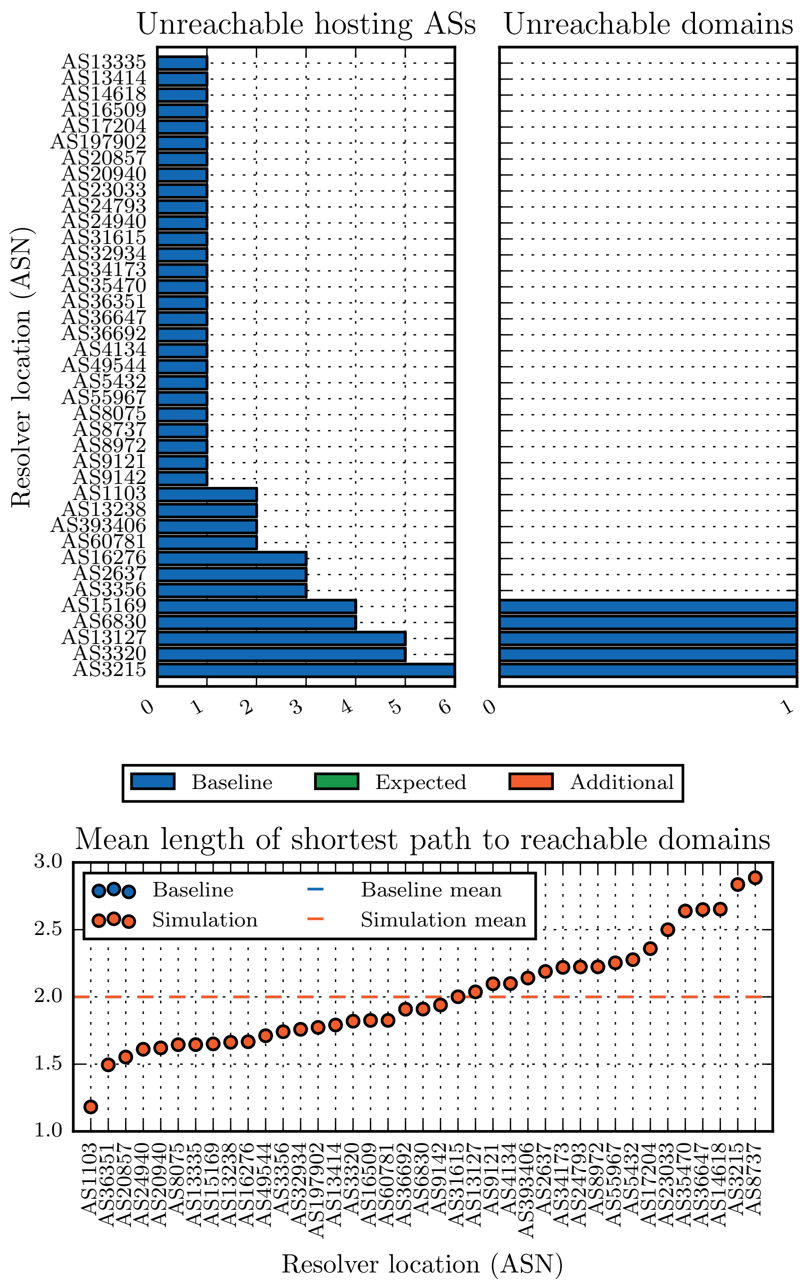


Figure B.56: Plots showing the reachability of autonomous systems and domains in a simulation where AS174↔AS2914 was removed from the AS topology compared to the baseline.

Table B.57: Reachability of DNS data in a simulation where the connection AS197902↔AS8455 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	**#Unreachable Domains**	**Mean length of shortest path to domain**	**#Unreachable ASs**	**#Unreachable Domains**	**Mean length of shortest path to domain**
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.87496542268
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0029784304

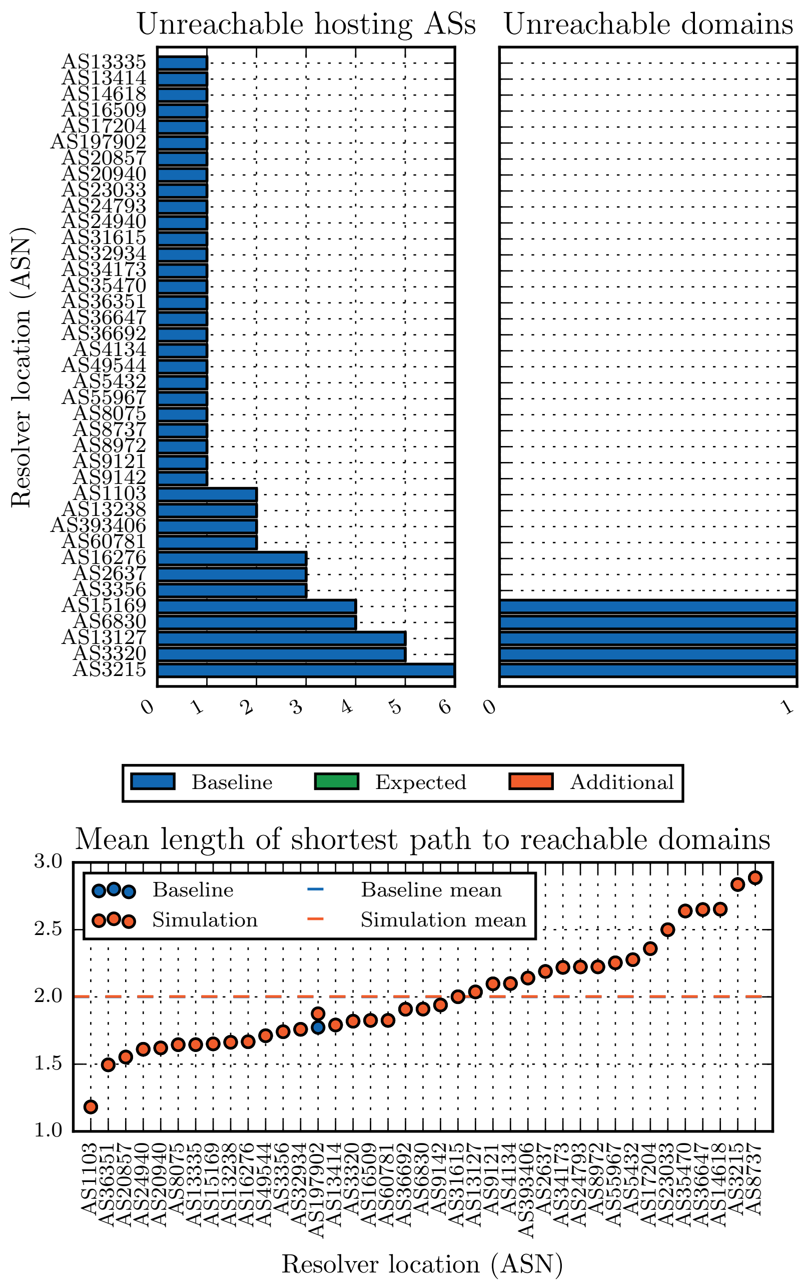


Figure B.57: Plots showing the reachability of autonomous systems and domains in a simulation where AS197902↔AS8455 was removed from the AS topology compared to the baseline.

Table B.58: Reachability of DNS data in a simulation where the connection AS3320↔AS8972 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.82012445228
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.2308915096
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65265747687
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82532562331
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00143211624
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64950842419
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0006051868

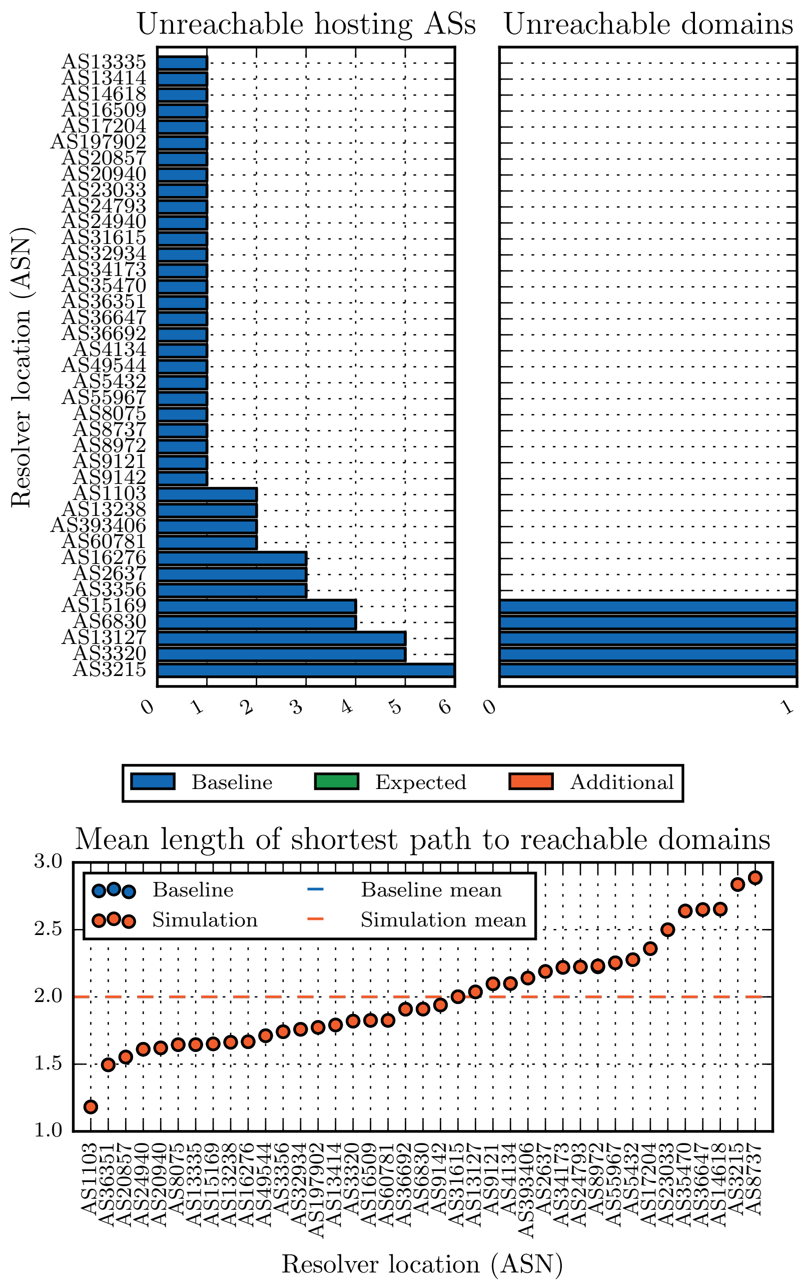


Figure B.58: Plots showing the reachability of autonomous systems and domains in a simulation where AS3320↔AS8972 was removed from the AS topology compared to the baseline.

Table B.59: Reachability of DNS data in a simulation where the connection AS1136↔AS286 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.18911427628
AS3215	6	1	2.83567157466	6	1	2.85079165305
AS3320	5	1	1.81932553893	5	1	1.83421354846
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.11443378564
AS5432	1	0	2.2768154119	1	0	2.29098782655
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64819504517
AS8737	1	0	2.88639271487	1	0	2.9208928666
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.03705757563
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.79106909723
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65589246024
AS15169	4	1	1.65053505386	4	1	1.66532389823
AS16276	3	0	1.66660210991	3	0	1.66660210991
AS16509	1	0	1.82501377501	1	0	1.82855986108
AS17204	1	0	2.35903972347	1	0	2.37322015334
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.49950622466
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.76248492951
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.66432708991
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82534314497
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14028867497
		Mean	2.0003648877		**Mean**	2.0042544865

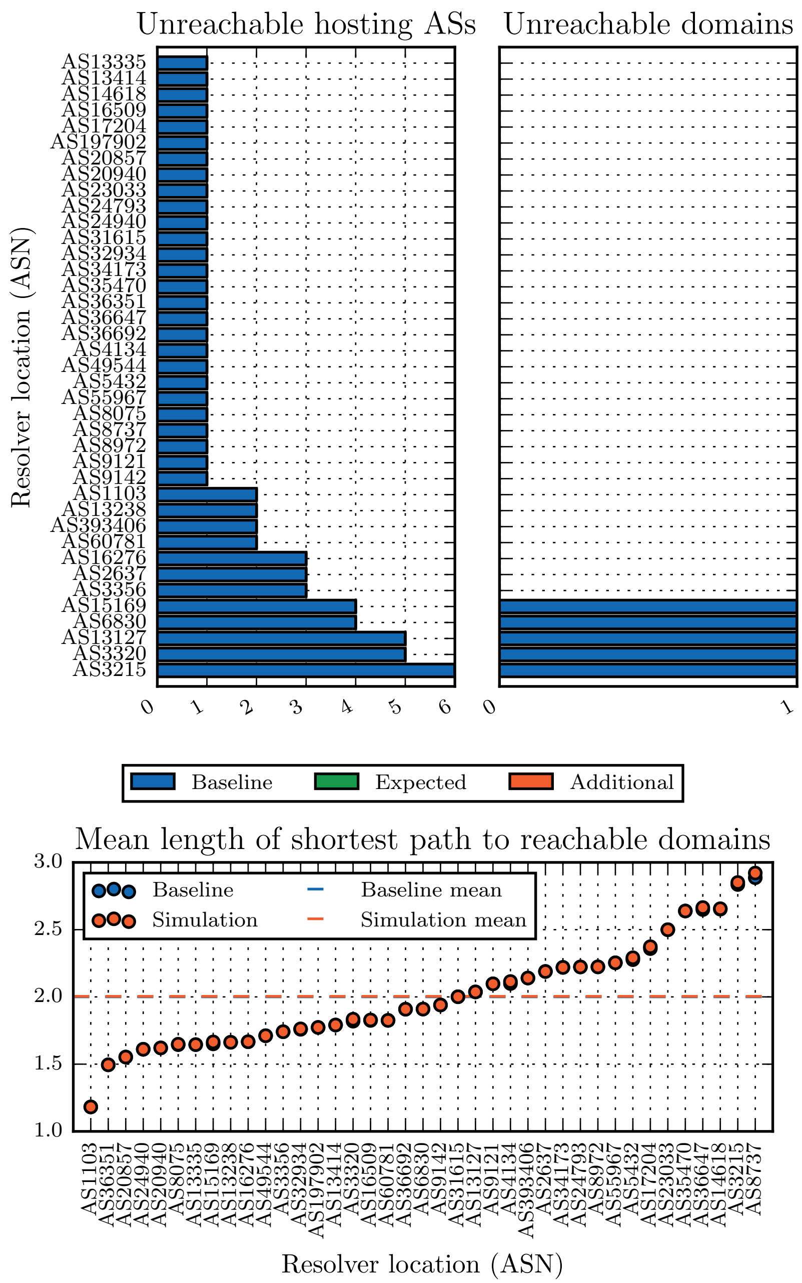


Figure B.59: Plots showing the reachability of autonomous systems and domains in a simulation where AS1136↔AS286 was removed from the AS topology compared to the baseline.

Table B.60: Reachability of DNS data in a simulation where the connection AS1299↔AS13127 was removed from the AS topology compared to the baseline.

Resolver location (ASN)	Baseline			Simulation		
	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain	#Unreachable ASs	#Unreachable Domains	Mean length of shortest path to domain
AS1103	2	1	1.18179189933	2	0	1.18179189933
AS2637	3	0	2.18911427628	3	0	2.1928171253
AS3215	6	1	2.83567157466	6	1	2.83567157466
AS3320	5	1	1.81932553893	5	1	1.81932553893
AS3356	3	0	1.74099517073	3	0	1.74099517073
AS4134	1	0	2.09931352367	1	0	2.09931352367
AS5432	1	0	2.2768154119	1	0	2.2768154119
AS6830	4	1	1.90947226796	4	1	1.90947226796
AS8075	1	0	1.64468828964	1	0	1.64468828964
AS8737	1	0	2.88639271487	1	0	2.88639271487
AS8972	1	0	2.22328412605	1	0	2.22328412605
AS9121	1	0	2.09727243649	1	0	2.09727243649
AS9142	1	0	1.93983546041	1	0	1.93983546041
AS13127	5	1	2.03705757563	5	1	2.17972773197
AS13238	2	0	1.66266719207	2	0	1.66266719207
AS13414	1	0	1.79106909723	1	0	1.7947561022
AS13335	1	0	1.64517926897	1	0	1.64517926897
AS14618	1	0	2.65234637417	1	0	2.65234637417
AS15169	4	1	1.65053505386	4	1	1.65053505386
AS16276	3	0	1.66660210991	3	0	1.66672345673
AS16509	1	0	1.82501377501	1	0	1.82501377501
AS17204	1	0	2.35903972347	1	0	2.35903972347
AS20857	1	0	1.55288727159	1	0	1.55288727159
AS20940	1	0	1.62063850426	1	0	1.62063850426
AS23033	1	0	2.49950622466	1	0	2.50590461357
AS24793	1	0	2.22269584558	1	0	2.22269584558
AS24940	1	0	1.61063288987	1	0	1.61063288987
AS31615	1	0	2.00108969823	1	0	2.00108969823
AS32934	1	0	1.75897836038	1	0	1.75897836038
AS34173	1	0	2.21939263956	1	0	2.21939263956
AS35470	1	0	2.638705947	1	0	2.638705947
AS36351	1	0	1.4947956937	1	0	1.4947956937
AS36647	1	0	2.64950842419	1	0	2.64962958462
AS36692	1	0	1.90835667691	1	0	1.90835667691
AS49544	1	0	1.71134721447	1	0	1.71134721447
AS55967	1	0	2.25354329006	1	0	2.25354329006
AS60781	2	0	1.82534314497	2	0	1.82903555555
AS197902	1	0	1.77303725702	1	0	1.77303725702
AS393406	2	0	2.14028867497	2	0	2.14654260336
		Mean	2.0003648877		**Mean**	2.004637894

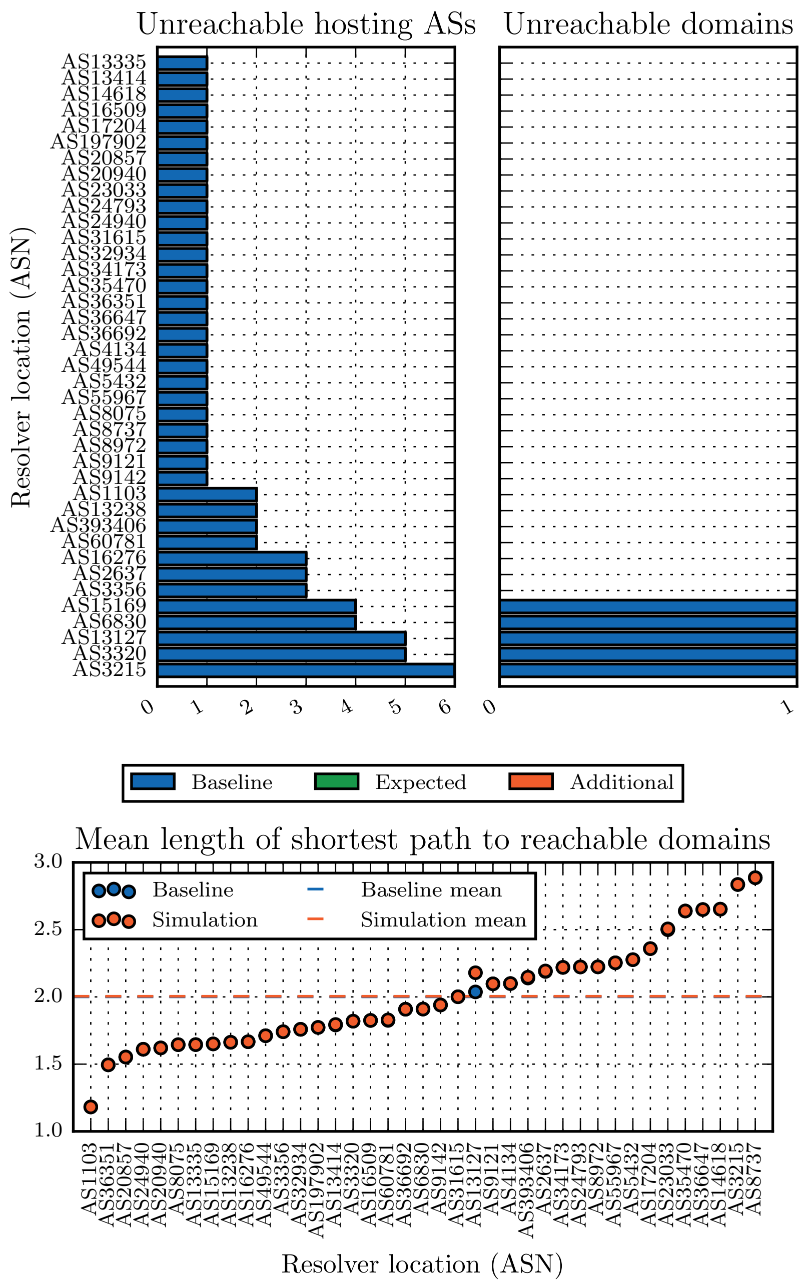


Figure B.60: Plots showing the reachability of autonomous systems and domains in a simulation where AS1299↔AS13127 was removed from the AS topology compared to the baseline.

Appendix C

Python Code

This chapter shows the scripts developed for the purpose of the performed analysis. The code is listed here for the sake of completeness and to enable other researchers to verify and/or adapt (parts of) the provided code. Please note that a lot of the code might require optimization and clarifications as well as documentation, however, the time needed for the analysis of the results did not allow for these and most of the code should be self explanatory. Nevertheless the code can be used as a guideline for future reference.

C.1 Environment and dependencies

Listing C.1 shows the environment settings and the imports which are necessary to run the code given in this chapter and to reproduce the figures shown in this thesis. Please note that not all imports are necessary for all parts of the code but it seemed to be useful to have them all collected in the same place.

Listing C.1: This listing shows the necessary imports and some environment settings necessary to run the code given in the other listings.

```
import operator
import copy
import matplotlib as mpl
mpl.use('pgf')
%matplotlib inline
mpl.rcParams.update({ # setup matplotlib to use latex for output
   "pgf.texsystem": "pdflatex", # change this if using xetex or lautex
   "text.usetex": True, # use LaTeX to write all text
   "font.family": "serif",
   "font.serif": [], # blank entries should cause plots to inherit
       fonts from the document
   "font.sans-serif": [],
   "font.monospace": [],
   "axes.labelsize": 10, # LaTeX default is 10pt font.
   "font.size": 10,
   "legend.fontsize": 8, # Make the legend/label fonts a little smaller
   "xtick.labelsize": 8,
```

```
    "ytick.labelsize": 8,
    "figure.figsize": [4.296388542963886, 2.6553141484273195], # default
         fig size of 0.9 textwidth
    "pgf.preamble": [
       r"\usepackage[utf8x]{inputenc}", # use utf8 fonts becasue your
           computer can handle it :)
       r"\usepackage[T1]{fontenc}", # plots will be generated using this
            preamble
       ]
    })
import matplotlib.pyplot as plt
plt.rc('text', usetex=True)
plt.rc('font', family='serif')

import numpy as np
import pandas as pd

import struct
import socket

import requests

import re

from ast import literal_eval

import datetime
```

C.2 The **DataWrapper** class

This class holds all data necessary for the analysis, some of which in multiple representations to enable faster access of necessary information. It is responsible for loading, filtering and storing of the data.

Listing C.2: The `DataWrapper` class contains all necessary data and preprocessing methods.

```
class DataWrapper(object):
    def __init__(self, timestamp):
       """
       Constructor.

       Parameters:
       timestamp - string - A timestamp used in the names of files
           containing data.
                       This can be used for distinguishing files
                           containging
                       data from different points in time.
       """
       # Initialize variables
       self.__datadir = "data/SIDN/DNS/"
       self.__timestamp = timestamp
       self.__tld = "nl"
```

```
    # The following variables contain the data used
    self.domains_to_ns = {} # Contains a one-to-many(unique) mapping
        of domains to nameservers
    self.domains_to_ns_zone = {} # Contains a one-to-many(unique)
        mapping of domains to nameservers as found in the zonefile
    self.ns_to_ips = {} # Contains a one-to-many(unique) mapping of
        nameservers to IPs
    self.ips_to_asn = {} # Contains a one-to-one mapping of IPs to
        ASNs
    self.ips_to_country = {} # Contains a one-to-one mapping of IPs
        to country

    # The following dictionaries contain mappings derived from the
        previous ones to allow some optimization of processing data
    self.domains_to_asn = {}
    self.ns_to_asn = {} # Contains a one-to-many(unique) mapping of
        nameservers to ASNs
    self.asn_to_domains = {} # Contains a one-to-many(unique) mapping
         of ASNs to domains
    self.asn_to_organisation = {} # Contains a one-to-one mapping of
        ASNs to organisations

    # Variables to hold the data that was filtered during
        preprocessing
    self.filtered_domains = {} # Contains a one-to-one mapping of
        domains to filter reason
    self.filtered_nameservers = {} # Contains a one-to-one mapping of
         nameservers to filter reason
    self.filtered_ips = {} # Contains a one-to-one mapping of ips to
        filter reason

    self.asgraph = ASGraph() # Contains the ASGraph object
        representing the AS topology
    self.asgraph_original = ASGraph()
    self.asgraph_shortest_paths = {} # Matrix of shortest paths
        between to given ASs (dynamic programming)

def preprocess(self):
    """
    Function to preprocess the given data.

    This function is used to create cache files for faster access and
         to filter
    out unusable data from the input. The preprocessed data is
        written back to disk.
    """
    # Read data for preprocessing
    self.readOriginalData()

    # Filter out unresolvable domains and nameservers
    self.filterUnresolvableData()

    # Generate mappings
    self.generateIPtoASNMapping()
    self.generateIPtoCountryMapping()

    # Filter out unusable data
```

```
    self.filterUnusableData()

    # Generate more mappings
    self.generateNStoASNMapping()
    self.generateDomainstoASNMapping()
    self.generateASNtoDomainMapping()
    self.generateASNtoOrganisationMapping()

    # Write preprocessed data back to files for later usage
    self.writePreprocessedData()

def init(self):
    """
    Initializes the object for analysis.

    This function should only be called if preprocessed data is
        available on disk, ie.
    after the preprocess() function was called at least once on an
        Analysis object belonging
    to the same timestamp.
    """
    # Read data from files
    self.readPreprocessedData()

def readOriginalData(self):
    """
    Reads the necessary data from disk.

    The follow files are read and expected to exist on disk:
    nl.zone_resolved.000000.txt: Containing all domains and
        nameservers
    nl.nameservers_resolved.000000.txt: Containing all nameservers
        contained in the zone_resolved
                           file and their corresponding IP
                               addresses
    if preprocess:
       nl.zone.000000.txt:
       nl.zone_error.000000.txt:
       nl.nameservers_error.000000.txt:
    else:
       nl.ips_to_asn.000000.txt:
       nl.ips_to_country.000000.txt:
       nl.domains_to_nameservers_zone.000000.txt:

    Parameters:
    preprocess - boolean - Specifies whether the data necessary for
        preprocessing should be
                 loaded. If set to True the data necessary for
                     preprocessing will
                 be loaded, if set to False the data necessary
                     for analysis will
                 be loaded.
                 Default: False
    """
    print "STATUS:␣Start␣reading␣data␣from␣files␣(" + str(datetime.
        datetime.now()) + ")"
    self.readDataFromFile("data/SIDN/DNS/nl.zone_resolved." + self.
```

```
        __timestamp + ".txt", self.domains_to_ns)
    self.readDataFromFile("data/SIDN/DNS/nl.nameservers_resolved." +
        self.__timestamp + ".txt", self.ns_to_ips)
    self.readZoneFile("data/SIDN/DNS/nl.zone." + self.__timestamp + "
        .txt", self.domains_to_ns_zone)

def readPreprocessedData(self):
    """
    Reads preprocessed data from disk. This function requires all
        files created by and specified in the documentation
    of writePreprocessedData to exist on disk. This function should
        therefore never be called before the
    data has been preprocessed.
    """
    self.readDataFromFile("data/SIDN/DNS/nl.ns_to_ips." + self.
        __timestamp + ".txt", self.ns_to_ips)
    self.readDataFromFile("data/SIDN/DNS/nl.domains_to_nameservers."
        + self.__timestamp + ".txt", self.domains_to_ns)
    self.readDataFromFile("data/SIDN/DNS/nl.
        domains_to_nameservers_zone." + self.__timestamp + ".txt",
        self.domains_to_ns_zone)
    self.readDataFromFile("data/SIDN/DNS/nl.ips_to_asn." + self.
        __timestamp + ".txt", self.ips_to_asn, False)
    self.readDataFromFile("data/SIDN/DNS/nl.ips_to_country." + self.
        __timestamp + ".txt", self.ips_to_country, False)
    self.readDataFromFile("data/SIDN/DNS/nl.ns_to_asn." + self.
        __timestamp + ".txt", self.ns_to_asn)
    self.readDataFromFile("data/SIDN/DNS/nl.asn_to_domains." + self.
        __timestamp + ".txt", self.asn_to_domains)
    self.readDataFromFile("data/SIDN/DNS/nl.domains_to_asns." + self.
        __timestamp + ".txt", self.domains_to_asn)
    self.readDataFromFile("data/SIDN/DNS/asn_to_organisation.txt",
        self.asn_to_organisation, False)
    self.readDataFromFile(self.__datadir + self.__tld + ".
        filtered_domains.txt", self.filtered_domains, False)
    self.readDataFromFile(self.__datadir + self.__tld + ".
        filtered_nameservers.txt", self.filtered_nameservers, False)
    self.readDataFromFile(self.__datadir + self.__tld + ".
        filtered_ips.txt", self.filtered_ips, False)
    self.asgraph_original.readFromFile("data/CAIDA/as-
        relationships20160601.txt")
    self.asgraph = self.asgraph_original.copy()

def writePreprocessedData(self):
    """
    Writes preprocessed data to disk.

    This function creates three distinct files:
       1. Mapping of nameserver IP addresses onto ASNs
       2. Mapping of nameserver IP addresses onto countries
       3. Mapping of domains onto nameservers as defined in the
           zonefile (for caching)
    """
    print "STATUS:␣Starting␣writing␣data␣(" + str(datetime.datetime.
        now()) + ")"
    self.writeDataToFile("data/SIDN/DNS/nl.ips_to_asn." + self.
        __timestamp + ".txt", self.ips_to_asn)
```

```
        self.writeDataToFile("data/SIDN/DNS/nl.ips_to_country." + self.
            __timestamp + ".txt", self.ips_to_country)
        self.writeDataToFile("data/SIDN/DNS/nl.domains_to_nameservers." +
             self.__timestamp + ".txt", self.domains_to_ns)
        self.writeDataToFile("data/SIDN/DNS/nl.
            domains_to_nameservers_zone." + self.__timestamp + ".txt",
            self.domains_to_ns_zone)
        self.writeDataToFile("data/SIDN/DNS/nl.ns_to_ips." + self.
            __timestamp + ".txt", self.ns_to_ips)
        self.writeDataToFile("data/SIDN/DNS/nl.ns_to_asn." + self.
            __timestamp + ".txt", self.ns_to_asn)
        self.writeDataToFile("data/SIDN/DNS/nl.asn_to_domains." + self.
            __timestamp + ".txt", self.asn_to_domains)
        self.writeDataToFile("data/SIDN/DNS/nl.domains_to_asns." + self.
            __timestamp + ".txt", self.domains_to_asn)
        self.writeDataToFile("data/SIDN/DNS/asn_to_organisation.txt",
            self.asn_to_organisation)
        self.writeDataToFile(self.__datadir + self.__tld + ".
            filtered_domains.txt", self.filtered_domains)
        self.writeDataToFile(self.__datadir + self.__tld + ".
            filtered_nameservers.txt", self.filtered_nameservers)
        self.writeDataToFile(self.__datadir + self.__tld + ".filtered_ips
            .txt", self.filtered_ips)

    def readDataFromFile(self, filepath, storage, sets=True):
        """
        Reads data from a file into memory.
        All files used during analysis follow the same format:
            1. One entry per line
            2. Every line contains a key-value pair, separated by a colon
                (:)

        """
        storage.clear()
        for line in open(filepath, 'r'):
            data = line.split(":")
            key = data[0].strip()
            value = data[1].strip()
            if sets:
                if key not in storage:
                    storage[key] = set()
                storage[key].add(value)
            else:
                storage[key] = value

    def writeDataToFile(self, filepath, storage):
        """
        Writes data from memory to file. For sepcifications on the format
             used see readDataFromFile.
        """
        f = open(filepath, 'w')
        for key, value in storage.items():
            if type(value) is set:
                for val in value:
                    f.write(key + "␣:␣" + val + "\n")
            else:
                f.write(key + "␣:␣" + value + "\n")
```

```
def resetGraph(self):
   self.asgraph = self.asgraph_original.copy()

def readZoneFile(self, filepath, storage):
   """
   Reads data from the zonefile. This file follows a special format
       and therefore requires an import function on its own.
   The zone file is needed during preprocessing.
   """
   for line in open(filepath, 'r'):
      if not (line.startswith(";") or line.startswith("nl.") or line
          .strip() == ""):
         data = re.split("\s+", line.strip())
         if(data[3] == "NS"):
            domain = data[0]
            nameserver = data[4]
            if domain not in storage:
               storage[domain] = set()
            storage[domain].add(nameserver)

def getASNforIP(self, ip, asndata):
   ip_integer = struct.unpack("!I", socket.inet_aton(ip))[0]
   line = asndata[(asndata[0] <= ip_integer) & (asndata[1] >=
       ip_integer)]
   if line.empty:
      return None
   else:
      return line.iat[0,2].split("␣")[0].strip()

def getCountryforIP(self, ip, countrydata):
   ip_integer = struct.unpack("!I", socket.inet_aton(ip))[0]
   line = countrydata[(countrydata[2] <= ip_integer) & (countrydata
       [3] >= ip_integer)]
   if line.empty:
      return None
   else:
      return line.iat[0,4]

def generateIPtoASNMapping(self):
   print "STATUS:␣Start␣generating␣IP2ASN␣mapping␣(" + str(datetime.
       datetime.now()) + ")"
   asndata = pd.read_csv("data/MaxMind/GeoIPASNum2.csv", header=None
       )
   for nameserver, ips in self.ns_to_ips.items():
      for ip in ips:
         asn = self.getASNforIP(ip, asndata)
         if asn is not None:
            self.ips_to_asn[ip] = asn

def generateIPtoCountryMapping(self):
   print "STATUS:␣Start␣generating␣IP2country␣mapping␣(" + str(
       datetime.datetime.now()) + ")"
   countrydata = pd.read_csv("data/MaxMind/GeoIPCountryWhois.csv",
       header=None)
   for nameserver, ips in self.ns_to_ips.items():
      for ip in ips:
```

```
            country = self.getCountryforIP(ip, countrydata)
            if country is not None:
               self.ips_to_country[ip] = country

   def generateASNtoOrganisationMapping(self):
      OrgData = {} # Holds a mapping from org_id to org_name
      readOrgData = False # These variables keep track of the
      readASData = False # position in the file currently parsed
      for line in open("data/CAIDA/as-organizations20160101.txt", 'r'):
            if "#␣format:org_id|changed|org_name|country|source" in
                 line:
               readOrgData = True
            elif "#␣format:aut|changed|aut_name|org_id|source" in line:
               readOrgData = False
               readASData = True
            elif readOrgData:
               temp = line.split("|")
               OrgData[temp[0]] = temp[2]
            elif readASData:
               temp = line.split("|")
               self.asn_to_organisation["AS" + temp[0]] = OrgData[temp
                   [3]]


   def generateNStoASNMapping(self):
      print "STATUS:␣Start␣generating␣NS2ASN␣mapping␣(" + str(datetime.
          datetime.now()) + ")"
      for nameserver, ips in self.ns_to_ips.items():
         asns = set()
         for ip in ips:
            asns.add(self.ips_to_asn[ip])
         self.ns_to_asn[nameserver] = asns

   def generateDomainstoASNMapping(self):
      for domain, nameservers in self.domains_to_ns.items():
         if domain not in self.domains_to_asn:
            self.domains_to_asn[domain] = set()
         for nameserver in nameservers:
            for asn in self.ns_to_asn[nameserver]:
               self.domains_to_asn[domain].add(asn)

   def generateASNtoDomainMapping(self):
      print "STATUS:␣Start␣generating␣ASN2Domain␣mapping␣(" + str(
          datetime.datetime.now()) + ")"
      for domain, nameservers in self.domains_to_ns.items():
         for nameserver in nameservers:
            for asn in self.ns_to_asn[nameserver]:
               if asn not in self.asn_to_domains:
                  self.asn_to_domains[asn] = set()
               self.asn_to_domains[asn].add(domain)

   def filterUnresolvableData(self):
      print "STATUS:␣Start␣filtering␣of␣data␣(" + str(datetime.datetime
          .now()) + ")"

      self.filterUnresolvableDomains()
      self.filterUnresolvableNameservers()
```

```
    print "STATUS:␣Start␣removing␣filtered␣data␣(" + str(datetime.
        datetime.now()) + ")"
    domains_prior_removal = len(self.domains_to_ns) # Number of
        domains before removal

    # Remove unresolvable domains
    print "Removing␣unresolvable␣domains:"
    self.removeDomains(set(self.filtered_domains.keys()))

    # Remove unresolvable nameservers
    print "Removing␣unresolvable␣nameservers:"
    self.removeNameservers(set(self.filtered_nameservers.keys()))

    domains_after_removal = len(self.domains_to_ns) # Number of
        domains after removal

    print "Domains␣removed:" + str(domains_prior_removal -
        domains_after_removal)


def removeDomains(self, domains):
    for domain in domains:
        del self.domains_to_ns[domain]

def removeNameservers(self, nameservers):
    print "STATUS:␣Start␣removing␣nameservers␣(" + str(datetime.
        datetime.now()) + ")"
    for nameserver in nameservers:
        del self.ns_to_ips[nameserver]
    print "STATUS:␣Still␣removing␣nameservers␣(" + str(datetime.
        datetime.now()) + ")"
    for domain, nsset in self.domains_to_ns.items():
        nsset -= nameservers
    print "STATUS:␣Finish␣removing␣nameservers␣(" + str(datetime.
        datetime.now()) + ")"

def removeIPs(self, ips):
    for nameserver, ipset in self.ns_to_ips.items():
        ipset -= ips
        if len(ips) == 0:
            for domain, nsset in self.domains_to_ns.items():
                nsset.discard(nameserver)

# Remove domains whose nameservers could not be found
def filterUnresolvableDomains(self):
    for domain, nameservers in self.domains_to_ns.items():
        nameserver = next(iter(nameservers))
        if nameserver.startswith("nodata"):
            self.filtered_domains[domain] = "unresolvable␣-␣" +
                nameserver

# Remove nameservers whose IP address could not be found
def filterUnresolvableNameservers(self):
    for nameserver, ips in self.ns_to_ips.items():
        ip = next(iter(ips))
        if ip.startswith("nodata"):
```

```
            self.filtered_nameservers[nameserver] = "unresolvable␣-␣" +
                 ip

   def filterUnusableData(self):
      self.filterUnknownASNs()

   # Remove nameservers whose nameservers AS is not known
   def filterUnknownASNs(self):
      print "STATUS:␣Start␣removing␣unknown␣ASs␣(" + str(datetime.
          datetime.now()) + ")"
      for domain, nameservers in self.domains_to_ns.items():
         for nameserver in nameservers:
            for ip in self.ns_to_ips[nameserver]:
               if ip not in self.ips_to_asn:
                  self.filtered_ips[ip] = "ASN␣unknown"

      # Remove IPs without ASN
      print "Removing␣IPs␣without␣ASN:"
      print len(self.filtered_ips)
      for nameserver, ips in self.ns_to_ips.items():
         ips -= set(self.filtered_ips.keys())
         if len(ips) == 0:
            self.filtered_nameservers[nameserver] = "ASN␣unkown"
            for domain, nsset in self.domains_to_ns.items():
               nsset.discard(nameserver)

      print self.domains_to_ns
      self.removeDomainsWithoutNameservers()

   def removeDomainsWithoutNameservers(self):
      # Remove domains who do not have any nameservers left
      print "Removing␣domains␣without␣nameservers:"
      domains_without_ns = set()
      for domain, nameservers in self.domains_to_ns.items():
         if len(nameservers) == 0:
            domains_without_ns.add(domain)
            self.filtered_domains[domain] = "No␣nameservers"
      print len(domains_without_ns)
      for domain in domains_without_ns:
         del self.domains_to_ns[domain]

      print len(self.domains_to_ns)
```

C.3 The **`ASGraph`** class

As the graph used is somewhat special in the sense that it contains directed and undirected edges at the same time, where edges are labeled according to some kind of relation between two nodes and the fact that path finding algorithms have to investigate the relations in order to only return valid paths, it was decided to create a custom representation class for the graphs. This was done as no standard graph library seemed to be suited for these special needs. Listing C.3 shows the implementation of this representation.

Listing C.3: The `ASGraph` class which is used to represent the structure of a network of autonomous systems and provides some functionality on these graphs, e.g. searching all paths or the shortest path between two given autonomous systems.

```
"""
A class representing a network of autonomous systems and their
    relationships.

This class is based on work found at:
   1. http://www.python-course.eu/graphs_python.php
   2. https://www.python.org/doc/essays/graphs/
"""
class ASGraph(object):
   MAX_PATH_LENGTH = 20

   def __init__(self, graph_dict={}):
      """ initializes a graph object """
      self.__graph_dict = graph_dict

      self.__marks = set()


   def copy(self):
      return ASGraph(copy.deepcopy(self.__graph_dict))

   def removeVertex(self, remove_vertex):
      del self.__graph_dict[remove_vertex]

      for vertex, neighbors in self.__graph_dict.items():
         if remove_vertex in neighbors.keys():
            del self.__graph_dict[vertex][remove_vertex]

   def removePeeringsBetweenASs(self, asns):
      for vertex, neighbors in self.__graph_dict.items():
         for neighbor, relation in neighbors.items():
            if vertex in asns and neighbor in asns and relation == "
                peer":
               del self.__graph_dict[vertex][neighbor]

   def removeConnection(self, connection):
      (start, goal) = tuple(connection)
      del self.__graph_dict[start][goal]
      del self.__graph_dict[goal][start]

   def readFromFile(self, filename):
      readData = False # This variables keeps track of the position in
          the file currently parsed
      for line in open(filename, 'r'):
         if "#" in line:
            readData = False
         else:
            readData = True

         if readData:
            temp = line.split("|")
            if temp[2] == "0":
```

```
            relation = "peer"
         elif temp[2] == "-1":
            relation = "provider"
         else:
            print "WARNING:␣Illegal␣relation␣found␣in␣the␣dataset:␣"
                 + line

         self.add_relation("AS" + temp[0], "AS" + temp[1], relation)

def contains_edge(self, edge):
   (start, goal, relation) = tuple(edge)
   if start in self.__graph_dict and goal in self.__graph_dict[start
       ]:
      return self.__graph_dict[start][goal] == relation
   else:
      return False

def contains_vertex(self, vertex):
   return vertex in self.vertices()

def markVertice(self, vertice, relation):
   self.__marks.add((vertice, relation))

def isMarked(self, vertice, relation):
   return (vertice, relation) in self.__marks

def clearMarks(self):
   self.__marks = set()

def vertices(self):
   """ returns the vertices of a graph """
   return list(self.__graph_dict.keys())

def edges(self):
   """ returns the edges of a graph """
   return self.__generate_edges()

def add_vertex(self, vertex):
   """ If the vertex "vertex" is not in
      self.__graph_dict, a key "vertex" with an empty
      list as a value is added to the dictionary.
      Otherwise nothing has to be done.
   """
   if vertex not in self.__graph_dict:
      self.__graph_dict[vertex] = {}

def generate_path(self, vertices):
   path = []
   for i in range(1, len(vertices)):
      start = vertices[i-1]
      goal = vertices[i]
      relation = self.__graph_dict[vertices[i-1]][vertices[i]]
      path.append((start, goal, relation))
   return path

def bfs_paths(self, start, goal):
   #print "START SEARCHING!!! FROM " + start + " TO " + goal
```

```
    self.clearMarks()

    if goal not in self.vertices():
       yield None
    elif start not in self.vertices():
       yield None
    elif start == goal:
       yield []

    queue = [(start, [start])]
    i = 0
    while i < len(queue):
       (vertex, path) = queue[i]
       i += 1
       #print "Current length: " + str(len(path))

       elements = len(path)
       if elements > 1:
          lastRelation = self.__graph_dict[path[elements-2]][path[
              elements-1]]
       else:
          lastRelation = "customer"

       nextRelation = set()
       nextRelation.add("sibling")
       if lastRelation != "peer":
          nextRelation.add(lastRelation)
       if lastRelation == "customer":
          nextRelation.add("peer")
          nextRelation.add("provider")
       elif lastRelation == "peer":
          nextRelation.add("provider")

       self.markVertice(vertex, lastRelation)

       if len(path) <= self.MAX_PATH_LENGTH:
          for nextVertex in set(self.__graph_dict[vertex].keys()) -
              set(path):
             if self.__graph_dict[vertex][nextVertex] in nextRelation
                 :
                if nextVertex == goal:
                   yield self.generate_path(path + [nextVertex])
                else:
                   if (not self.isMarked(nextVertex, self.
                       __graph_dict[vertex][nextVertex])):
                      queue.append((nextVertex, path + [nextVertex]))
    yield None

 def shortest_path(self, start, goal):
    pathgen = self.bfs_paths(start, goal)
    path = next(pathgen)
    return path

 def add_relation(self, vertex1, vertex2, relation):
    """ Adds the given relation between vertex1 and vertex2 to the
        graph.
       As the graph is directed, every relation corresponds to two
```

```
            distinct
        edges.
    """
    if relation == "customer":
        self.add_edge((vertex1, vertex2, "customer"))
        self.add_edge((vertex2, vertex1, "provider"))
    elif relation == "provider":
        self.add_edge((vertex1, vertex2, "provider"))
        self.add_edge((vertex2, vertex1, "customer"))
    elif relation == "peer":
        self.add_edge((vertex1, vertex2, "peer"))
        self.add_edge((vertex2, vertex1, "peer"))
    elif relation == "sibling":
        self.add_edge((vertex1, vertex2, "sibling"))
        self.add_edge((vertex2, vertex1, "sibling"))
    else:
        print "ERROR:␣Unknown␣relation␣specified!"

def add_edge(self, edge):
    """ assumes that edge is of type tuple of the form
        (source, destination, relation).
        There can be multiple edges between two vertices!
        If the source and/or destination vertex does not
        exist, these will be created in the graph.
    """
    (vertex1, vertex2, relation) = tuple(edge)
    self.add_vertex(vertex1)
    self.add_vertex(vertex2)
    self.__graph_dict[vertex1][vertex2] = relation

def get_neighbors(self, vertex):
    return self.__graph_dict[vertex]

def __generate_edges(self):
    """ A static method generating the edges of the
        graph "graph". Edges are represented as tuples
        of the form (source, destination, relation).
    """
    edges = []
    for vertex in self.__graph_dict:
        for neighbour, relation in self.__graph_dict[vertex].iteritems
            ():
            #if (vertex, neighbour, relation) not in edges:
                edges.append((vertex, neighbour, relation))
    return edges

def __generate_path(self, path):
    """ A static method that generates the path given an
        array of nodes on the path.
    """
    ret = str(path[0])
    lastnode = path[0]
    del path[0]

    for vertex in path:
        ret = ret + "␣-"+ self.__graph_dict[lastnode][vertex] + "-␣" +
             vertex
```

```
        lastnode = vertex

      return ret

   def __str__(self):
      res = "vertices:␣"
      for k in self.__graph_dict:
         res += str(k) + "␣"
      res += "\nedges:␣"
      for edge in self.__generate_edges():
         res += str(edge) + "␣"
      return res

   def get_degree_distribution(self):
      dist = {}
      degrees = {}
      for vertex, neighbors in self.__graph_dict.iteritems():
         degrees[vertex] = len(neighbors)

      for vertex, degree in degrees.iteritems():
         if degree not in dist:
            dist[degree] = 0

         dist[degree] = dist[degree] + 1

      return sorted(degrees)

   def plot_degree_distribution(self):
      dist = self.get_degree_distribution()
      plt.plot(dist.values(), dist.keys())
      plt.ylabel("Degree")
      plt.xlabel('Amount␣of␣nodes')
      plt.show()
```

C.4 The **`Analysis`** class

This class is responsible for performing the actual analysis. It is initialed with the data to be analyzed and afterwards the individual methods can be called.

Listing C.4: This class contains all methods which are actually used for the analysis part.

```
"""
This class contains all methods for analysis of the given data. It has
    to be instantiated with an DataWrapper object.
"""
class Analysis(object):
   def __init__(self, data):
      self.data = data

   def accumulateValues(self, values):
      """
      This function accumulates all given values and returns a
          dictionary which contains the number of
      occurences of every value in the given dictionary.
```

```
        """
        accumulated = {}
        for key, value in values.items():
            if type(value) is set:
                for v in value:
                    if v not in accumulated:
                        accumulated[v] = 0
                    accumulated[v] += 1
            else:
                if value not in accumulated:
                    accumulated[value] = 0
                accumulated[value] += 1
        return accumulated

    def countValues(self, values):
        """
        This function counts the number of occurences of values on a list
             and returns a dictionary with the
        number of occurences per value.
        """
        counts = {}
        for key, value in values.items():
            counts[key] = len(value)
        return counts

    # Analyses the mismatches between nameservers contained in the
         zonefile and nameservers found by resolution of domains
    def analyseZoneDNSMismatches(self):
        zone_mismatches = {"UnresolvableNameservers": set(), "
             NameserverMismatchZoneMore": set(), "
             NameserverMismatchZoneLess": set()}
        for domain, nameservers in self.data.domains_to_ns_zone.items():
            if domain not in self.data.domains_to_ns:
                zone_mismatches["UnresolvableNameservers"].add(domain)
            else:
                if len(nameservers.difference(self.data.domains_to_ns[
                     domain])) > 0:
                    zone_mismatches["NameserverMismatchZoneMore"].add(domain
                        )
                if len(self.data.domains_to_ns[domain].difference(
                     nameservers)) > 0:
                    zone_mismatches["NameserverMismatchZoneLess"].add(domain
                        )

        return zone_mismatches

    def analyseDomainsWithSingleAS(self):
        single_as = 0
        asn_to_domains = {}
        for domain, asns in self.data.domains_to_asn.items():
            if len(asns) == 1:
                asn = next(iter(asns))
                if asn not in asn_to_domains:
                    asn_to_domains[asn] = set()
                asn_to_domains[asn].add(domain)

                single_as += 1
```

```
    return single_as

def analyseDomainsWithSingleIP(self):
    domains_ips = {}
    for domain, nameservers in self.data.domains_to_ns.items():
        domains_ips[domain] = set()
        for nameserver in nameservers:
            for ip in self.data.ns_to_ips[nameserver]:
                domains_ips[domain].add(ip)

    single_ip = 0
    for domain, ips in domains_ips.items():
        if len(ips) == 1:
            single_ip += 1

    return single_ip

def analyseDomainsPerNameserver(self):
    domains_per_nameserver = {}
    for domain, nameservers in self.data.domains_to_ns.items():
        for nameserver in nameservers:
            if not nameserver in domains_per_nameserver:
                domains_per_nameserver[nameserver] = 0
            domains_per_nameserver[nameserver] += 1

    return domains_per_nameserver

def analyseDomainsPerNameserverIP(self):
    domains_per_ip = {}
    for domain, nameservers in self.data.domains_to_ns.items():
        for nameserver in nameservers:
            for ip in self.data.ns_to_ips[nameserver]:
                if not ip in domains_per_ip:
                    domains_per_ip[ip] = 0
                domains_per_ip[ip] += 1
    return domains_per_ip

def analyseNameserversPerAS(self):
    nameservers_per_as = {}
    for nameserver, ips in self.data.ns_to_ips.items():
        for ip in ips:
            asn = self.data.ips_to_asn[ip]
            if not asn in nameservers_per_as:
                nameservers_per_as[asn] = 0
            nameservers_per_as[asn] += 1
    return nameservers_per_as

def analyseDomainsPerAS(self):
    domains_per_as = {}
    for domain, nameservers in self.data.domains_to_ns.items():
        for nameserver in nameservers:
            for ip in self.data.ns_to_ips[nameserver]:
                asn = self.data.ips_to_asn[ip]
                if not asn in domains_per_as:
                    domains_per_as[asn] = 0
                domains_per_as[asn] += 1
```

```
    return domains_per_as

def analyseDomainsPerCountry(self):
    domains_per_country = {}
    for domain, nameservers in self.data.domains_to_ns.items():
        for nameserver in nameservers:
            for ip in self.data.ns_to_ips[nameserver]:
                # The country is not guaranteed to be known as this is
                    not a filtering criteria
                if ip in self.data.ips_to_country:
                    country = self.data.ips_to_country[ip]
                    if country not in domains_per_country:
                        domains_per_country[country] = 0
                    domains_per_country[country] += 1

    return domains_per_country

def analyseNameserversPerCountry(self):
    return self.accumulateValues(self.data.ips_to_country)

def loadBaselineShortestPaths(self):
    baseline_shortest_paths = {}
    for line in open("data/baseline_paths.txt", 'r'):
        # Load start : goal : path
        data = line.split(":")
        start = data[0].strip()
        goal = data[1].strip()
        path = literal_eval(data[2].strip())

        if start not in baseline_shortest_paths:
            baseline_shortest_paths[start] = {}
        baseline_shortest_paths[start][goal] = path
    return baseline_shortest_paths


def storeBaselineShortestPaths(self, baseline_shortest_paths):
    f = open("data/baseline_paths.txt", 'w')
    for start, paths in baseline_shortest_paths.items():
        for goal, path in paths.items():
            f.write(start + " : " + goal + " : " + str(path) + "\n")

def analyseMostImportantTransitASs(self):
    baseline_shortest_paths = self.loadBaselineShortestPaths()

    occurences = {}

    for start, goals in baseline_shortest_paths.items():
        for goal, path in goals.items():
            if path is not None:
                skipFirst = True
                for step in path:
                    if skipFirst:
                        skipFirst = False
                    else:
                        (vertex1, vertex2, relation) = tuple(step)
                        if vertex1 not in occurences:
```

```
                     occurences[vertex1] = 0
                  occurences[vertex1] += 1

      return occurences

   def analyseMostImportantConnections(self):
      baseline_shortest_paths = self.loadBaselineShortestPaths()

      occurences = {}

      for start, goals in baseline_shortest_paths.items():
         for goal, path in goals.items():
            if path is not None:
               for step in path:
                  connection = tuple(sorted(step[0:2]))
                  if connection not in occurences:
                     occurences[connection] = 0
                  occurences[connection] += 1

      return occurences


   def analyseReachability(self, asns, baseline=False):
      if baseline:
         baseline_shortest_paths = {}
      else:
         # Load shortest paths from baseline
         baseline_shortest_paths = self.loadBaselineShortestPaths()

      print "Resolver␣ASN␣&␣Unreachable␣ASs␣&␣Unreachable␣Domains␣&␣
          Mean␣path␣length␣\\\\"

      for start in asns:
         unreachableAS = set()
         status = 0
         paths = {}

         for asn, domains in self.data.asn_to_domains.items():
            shortest_path = None

            cached_path = False # cached_path indicated whether or not
                a cached path is available
            if not baseline:
               cached_path = True
               # Check whether baseline path is still applicable
               shortest_path = baseline_shortest_paths[start][asn]
               if shortest_path == []:
                  if not self.data.asgraph.contains_vertex(start):
                     shortest_path = None
               elif shortest_path != None:
                  for edge in shortest_path:
                     if not self.data.asgraph.contains_edge(edge):
                        cached_path = False
                        break

            if not cached_path:
               shortest_path = self.data.asgraph.shortest_path(start,
```

```
                    asn)

            if baseline:
               if start not in baseline_shortest_paths:
                  baseline_shortest_paths[start] = {}
               baseline_shortest_paths[start][asn] = shortest_path

            status +=1
            if shortest_path == None:
               unreachableAS.add(asn)
            else:
               paths[asn] = shortest_path

         unreachableDomains = set()
         tot_length = 0.0
         for domain, hosting_asns in self.data.domains_to_asn.items():
            if hosting_asns.issubset(unreachableAS):
               unreachableDomains.add(domain)
            else:
               # Calculate shortest path
               tot_length += min([len(paths[asn]) for asn in
                    hosting_asns - unreachableAS])

         countUnreachableDomains = len(unreachableDomains)
         countReachableDomains = len(self.data.domains_to_ns.keys()) -
             countUnreachableDomains
         meanpathlen = -1
         if countReachableDomains != 0:
            meanpathlen = tot_length / countReachableDomains

         print start + "␣&␣" + str(len(unreachableAS)) + "␣&␣" + str(
             countUnreachableDomains) + "␣&␣" + str(meanpathlen) + "␣
             \\\\"

         print start + ":␣" + str(unreachableAS)

      if baseline:
         self.storeBaselineShortestPaths(baseline_shortest_paths)


   def plotData(self, data, title="", xlabel="", xticks=[], xticklabels
       =[], ylabel="", filename="plot"):
      color_blue = "#1268b3"

      mpl.rcParams.update({ # setup matplotlib to use latex for output
         "figure.figsize": [4.296388542963886, 2.6553141484273195], #
             default fig size of 0.9 textwidth
      })

      plt.plot(data, color=color_blue)
      plt.gca().axes.set_yscale('symlog')
      plt.grid(True)
      plt.title(title, fontsize=12)
      plt.xlabel(xlabel)
      plt.gca().axes.get_xaxis().set_ticks(xticks)
      plt.gca().set_xticklabels(["{:,}".format(int(x)) for x in plt.gca
          ().get_xticks()])
```

```
    plt.ylabel(ylabel)
    plt.ylim(ymin=0)
    plt.xlim(xmin=0, xmax=len(data))
    plt.savefig('{}.pdf'.format(filename), bbox_inches='tight')
    plt.show()
```