# Border Gateway Protocol
# Complexity and Stability

Jeffrey de Looff
Faculty of Exact Sciences, department of Computer Science
Vrije Universiteit Amsterdam

December 17, 2013

**Master's Thesis**

Supervisors

**Benno Overeinder**
NLnetLabs, Amsterdam

**Spyros Voulgaris**
VU Amsterdam

**Abstract**

In recent years the Internet has grown to a very large and complex system. Covering all continents and an ever increasing number of users, the Internet consists of many important subjects to study. One such subject is the Border Gateway Protocol which assists in routing traffic from one source in the network to another destination in the network. However, previous studies have not focused on how small changes in the network could affect the global behaviour of the network. With a simulator running on a super-computer, real-world scenarios could be investigated. Scenarios are, for example, disconnecting links between Tier 1 Autonomous Systems (AS) and disconnecting a Tier 1 AS completely. Results include the average number of received update messages over all the ASes in a specific Tier, Path Diversity of ASes and Area of Impact. The number of received updates increases if more links are disconnected. Tier 1 ASes receive most of the withdrawals and convert those into announcements. These announcements are in turn received by both Tier 2 and 3 ASes. Path Diversity and Area of Impact might be indicators for the network's stability. ASes with a high Path Diversity are able to replace withdrawn routes with alternative paths. A small Area of Impact means the failure is local and the network remains stable.

In this study sources of instability and complexity are given. However, since many potential parameters were discovered, future study would be necessary to investigate those parameters into more detail.

**Keywords:** Border Gateway Protocol, Complexity, Stability, Simulation, Network Topology, Failure types

# Contents

# 1 Introduction

The Internet consists of a large number of networks inter-connected with each other. These inter-connected networks have evolved over the last few decades with constant changes such as new networks and connections added, different relations between the networks and more users every day. With more users comes an increasing volume of traffic, maintenance and strain on the Internet as a whole. Knowledge of the existing technologies used in the Internet is highly relevant for its future. In recent years there have a number of studies concerning its stability [1, 2], scalability [3, 4, 5], complexity [6], security [7, 8, 9, 10], topology [11, 12] and robustness [13].

However, previous studies have not yet focused on the impact of small changes in the network on the global behaviour. Today's Internet is facilitated by a robust yet efficient mechanism that interconnects individual networks into a global infrastructure. The Border Gateway Protocol (BGP) achieves interconnection and assists in routing of traffic, keeps track of reachability and makes the Internet scalable. This scalability comes from the dynamic nature of BGP by means of adjusting its configuration and routes when there are changes in the network. Nevertheless, some small changes could potentially bring the Internet to its knees if no proper measurements are taken. Therefore, it is crucial to study the overall behaviour and complexity.

With a generated topology a BGP simulator [14] is set-up and used to investigate the impact of small changes in the network on the overall stability. This thesis reviews the research conducted with the BGP simulator and attempts to discover possible sources for causing instability. Such as: the growth of the Internet in all its different parameters; the ever growing complexity of all the individual and overall pieces; or the overall structural topology of the Internet as it is designed today. These different sources could be researched with the simulator and may even lead to further observations of BGP. Observations that are achieved by simulating different (multiple) failure types such as disconnecting random or targeted links. Different failure types could lead to answers to unanswered questions about BGP's stability and complexity:

- How does the network react to failures of different types?
- How is BGP's stability or complexity defined in the face of failures?

- Which impact does the number of ASes and type of topology have on the overall stability and complexity?

Chapter 2 describes further details regarding BGP. In Chapter 3 the BGP simulator to perform the experiments with, the experimental set-up and details are explained. In Chapter 4 the results are discussed and explains the observations discovered in the experiments. A conclusion, possible improvements and suggestions for future research are given in Chapter 5.

pagebreak
for new chapters.

# 2 Background

Below is a description on the intricate details of BGP.

## 2.1 Border Gateway Protocol

Today's Internet relies on BGP as the standard routing information protocol. It has gone through various development phases and improvements since the original version, BGP-1, as proposed in 1989 [15]. Version 4 of BGP has been deployed in 1993 and is the first of its kind which handles aggregation (classless interdomain routing [CIDR]) and supernetting.

BGP is developed in such a manner that it imposes no restriction on the underlying network topology. Routing is done within an AS via an intra-autonomous system routing protocol (Interior Gateway Protocol [IGP]). For the purposes of this research IGP is not considered and abstracted. Instead, BGP is able to construct a graph of Autonomous Systems (ASes) based on information received from neighbouring BGP routers. Such a directed graph environment is referred to as a tree. The whole Internet is a graph of interconnected ASes.

### 2.1.1 Autonomous Systems

The Internet is divided into different regions of administrative control known as an Autonomous System (AS) [16]. An AS is identified by its officially assigned AS number (ASN) [17]. This ASN is globally unique and used during exchange of routing information. As of writing there are 44703 unique ASes [18]. Each AS contains a number of BGP speakers to connect with neighbouring ASes. Connections between ASes form paths and the collection of unique path information form a route to a specific destination in the network. With this path information associated with a unique destination it ensures loop-free routing. Figure 1 displays the interconnections between ASes. Each AS could possibly contain multiple BGP speakers and do not necessarily have to connect with neighbouring ASes.

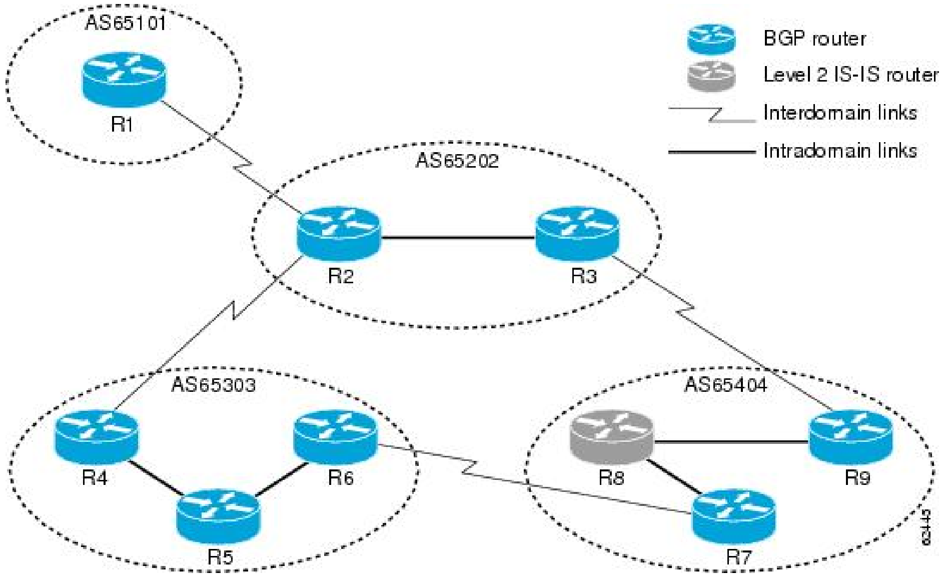

Figure 1: Autonomous Systems and their interconnections

### 2.1.2 AS relations

ASes are connected with one another and, for administrative and scalability reasons, specific relations exists between networks. Such a relation could be customer-provider, peering or merely a stub. In a customer-provider relationship the customer pays the provider for transit services. In a peering relationship ASes have a (private) contract to exchange traffic between their networks without having to pay for the bandwidth. ASes can have one or multiple providers, customers or peers. These types of relationships are depicted in figure 2.

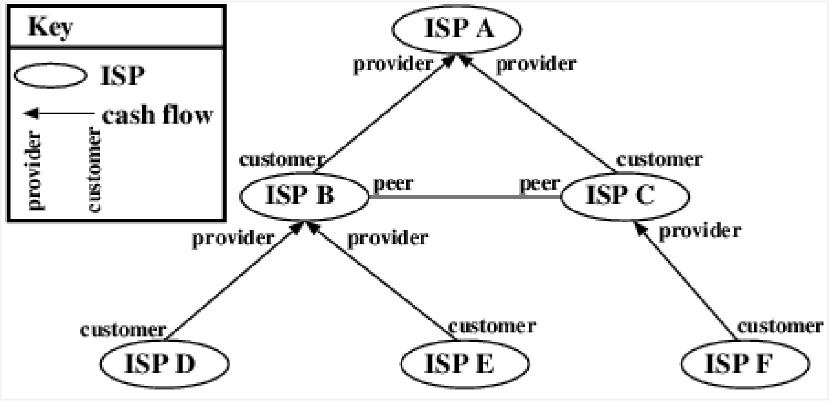

Figure 2: AS relationships [19]: *ASes D, E and F are customers of both B and C and are in turn providers of the ASes below. ASes B and C are also peers of one another and customers of A. This AS A is a top Tier AS because it is only provider and is a customer of no other AS.*

### 2.1.3 Routing information

Each AS has (unique) prefixes assigned to it [20]. Such prefixes are used to identify and locate destination networks and are IP addresses from IPv4 or IPv6 divided into two parts, a network section and a host section. The network section identifies a particular network while the host section identifies a specific node. With this information a data packet with a certain destination received at a router can be identified as lying in the same or some other network. How to reach a certain network is accomplished by means of exchanging BGP routing information.

This routing information is shared among peers and has two forms: announcements and withdrawals. An announcement includes a route which a router has learned from its peers or could have made a policy decision preferring another route to a network destination. A withdrawal is sent if a router makes a local decision that a network is unreachable. A distinction is made between explicit and implicit withdrawals. Explicit would mean a withdrawal message from one of its peers has been received. An implicit withdrawal occurs if an existing route is replaced by the announcement of a new route to this destination prefix [16]. Upon receipt of an announcement or withdrawal each BGP-speaker decides what to do with the message. The message could affect the routing table or has to be propagated to its neighbours. Such decisions are made following the BGP-speaker's AS policies.

As a result, BGP performs peer-to-peer information exchange by distributing routing information and keeping routing tables up-to-date. Distributing this information is done in an incremental manner, meaning messages are only sent if a new route is either announced, updated or withdrawn. These messages should be exchanged with all other ASes according to AS policies. Each route to a certain AS is stored as a list of all the ASes that have to be traversed in order to reach the destination, as well as the next hop. If a new announcement with route has been received by a BGP-speaker it checks whether the new route is preferred than the current one by evaluating its policies. If the policy depicts the new route more preferred then the old route is replaced and, if allowed by policies, announced to its neighbours.

In turn, BGP requires all BGP-speakers involved in the process to store the information received from every neighbour and all information sent to other BGP speakers. One route can be used per destination and it is prohibited to announce a different route from the one used, even though it limits the path diversity of an ISP [21]. To store only the next-hop AS would be sufficient. However, BGP requires a full path with all ASes to determine the preferred paths (length-wise) and avoid routing loops. These are avoided by checking if the BGP speaker's own AS is not already included in the path.

In the core of the Internet there are mainly providers which peer with one another to provide connectivity to all their customers. These customers might peer with one another as well but are mostly depending on the provider for reachability. This reachability is achieved by BGP's vector-path based algorithm. In this algorithm information is not broadcasted periodically and BGP speakers have no complete view of the network topology. Each speaker knows how to route traffic to neighbours and which neighbours are able to reach certain networks. However, some networks may be *more preferred* and policies indicate that traffic should be directed to specific providers and/or customers.

In contrast, link-state algorithms, such as the Open Shortest Path First (OSPF) and Intermediate system to Intermediate system (IS-IS), have global knowledge, which are commonly deployed inside ASes. Since these algorithms require global knowledge they are not scalable.

## 2.2 BGP (In)stability and Complexity

### 2.2.1 Instability issues

Even though BGP's principles seem to be simple, the overall impact of small changes could affect the whole network. BGP only knows about reachability and faces a few problems which cause route instabilities, such as: route flapping; convergence time; faulty hardware; software problems; insufficient CPU power; insufficient memory; network upgrades and routine maintenance; human error [15]; no verification to determine if the routing information distributed is valid [22]; no mechanism to convey capacity information [23]; mechanisms to direct traffic away from congestion are limited [10, 24, 25]; can be slow to converge to a stable state [26, 27]; no assurance to cope with extreme conditions.

**Route flapping** A BGP-speaker advertises a destination network repeatedly in quick sequence. Its state could go *up* and *down* because of an unreliable connection, misconfiguration, software bug or power-problem. A possible solution is route dampening which will be discussed further on.

**Convergence time** Convergence relates to the state of the network in which the set of routers maintain the same topological information. If the routers are said to be converged, they must have gathered all available topology information from one another through the BGP process. The information reflects the real state of the network. After enabling a router it will participate in the protocol and communicate information about the topology of the network. Any change in the network will result in a non-convergent state. This state could be caused by a flapping router, certain (mis-)configuration or hardware condition. Convergence time is the measurement of how long it will take for all routers to reach the state of convergence. The state is reached slower in a larger network and all routes running BGP should quickly and reliably converge. If a BGP session dies: routers are forced to withdraw all routes learned via that session, remove the routes from their forwarding tables, recalculate best routes to the affected prefixes and send out updated advertisements.

**Faulty hardware** Either faulty interfaces, systems or faulty lines could be the cause of route instability. An interface which is not functioning properly might cause certain route information to change state repeatedly. Failures in hardware are beyond the control of service users. In order to reduce connectivity loss due to failures, system and link redundancy are used. However, if a physical failure happens, routing is interrupted completely, which in turn could cause a kind of cascading effect throughout the network.

**Software problems** Software is deployed on many devices running the Internet and could cause software problems (bugs). These bugs could produce system failures and network instability. Since the Internet is a live network it is almost impossible to prevent every situation which could occur. New software, new features or new standards should be experimented in test environments in order to gain a stable version.

**Insufficient CPU power** With an increasing network size there are more routing updates and peering sessions a router has to handle. Handling more updates and sessions requires more CPU power. In the initial start-up phase of BGP the routing tables have to be set-up. After sessions have been established between BGP speakers a system's processor might spend more than 90% of its time to process all incoming announcements. This could cause the links to become unstable and overloaded. An overloaded BGP speaker could trigger a race condition: its CPU would be too busy handling updates and some sessions might be dropped, producing more instability.

**Insufficient memory** A router has to run its own operating system, usually a Linux distribution. In addition, sufficient memory is required to store BGP's routing tables, cache tables, policies and related databases. If these operations cause the router to run out of memory it might stop functioning altogether. This could result in loss of certain or all routes or updates.

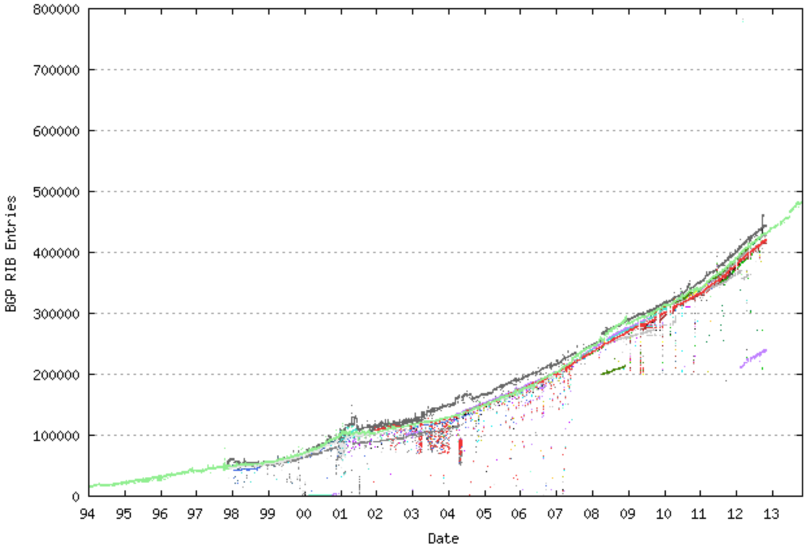

Figure 3: Size of BGP's routing table over time [28]

Each BGP speaker has a routing table consisting of route entries with additional necessary information corresponding to that entry. As of today, Internet routing tables consist of over 400000 entries [28], and, as can be seen in figure 3, increases every year. Router which store the complete table from the Internet from one or more providers are hardly able to keep up. If a router has insufficient memory to begin with, it could turn into a source of route flaps.

**Network upgrades and routine maintenance** The Internet's network nature is very dynamic. In order to improve performance, perform site consolidation or expand the network, changes are required. Such changes could be upgrading to newer versions of either software or hardware, add more links or higher bandwidth or reconfigure the network topology.

If administrators were to upgrade their systems they would prefer to bring it down during times there is minimal usage and users experience minimal downtime. Furthermore, downtime cannot exceed hours on end for some networks because of the differences in time zones. Even though these are obstacles for upgrading systems the most significant errors do not happen during the upgrade period. If hardware/software problems arise, network instability is caused the next day when users get back online. If this is the case, rolling back to the old set-up is not an option. Another option is to remedy the complete situation. However, if administrators start to add or change the configurations it could potentially result in an even worse situation.

In conclusion, to reduce the possibility of disruptions, such network changes are recommended to be simulated in a test environment, for example, a BGP simulator [14]. If multiple changes were to be applied they should be deployed one after another to analyze the effects of each change.

**Human error** Some of the network instabilities could be caused by human errors. Some private policies might allow traffic to flow inefficiently or, if an administrator implements a change, the effects are not known a priori. Consequently, even small changes could be catastrophic, one wrong setting and the complete AS could become disconnected. Even though it is another administrator's responsibility to send the correct routes, it is also the administrator's responsibility to protect the AS from benign or unwanted routes.

**No verification of routing information** The information exchanged between BGP-speakers is only concerning reachability. After a BGP-session between two speakers is established they will exchange information from their tables/configurations. However, their tables or configurations might not be valid. Traffic can be disrupted, modified, examined or all three.

**No mechanism to convey capacity information** BGP cannot help to reconfigure the interconnections to avoid congestion. If, for some reason, a route fails, BGP will find another route. However, this route could be suboptimal and may have insufficient capacity to cope with traffic it will receive now. Even if this new route is able to handle the extra traffic, some router might not be able to cope at all and degrade the overall performance.

### 2.2.2 Stability features of BGP

To achieve stability in BGP, effective routing policies have to be developed and configured correctly. BGP inhibits a buffer for possible route instabilities with certain functions. Such as controlling the route and cache validation, BGP and BGP route dampening.

**Controlling route and cache validation** A BGP session is set-up over TCP between two neighbours with an OPEN message. This message contains the BGP version number, its AS number, hold down timer (number of seconds between KEEPALIVE and UPDATE messages) and an identifier. Update messages include various attributes and whenever an administrator changes an attribute or some policy, the complete session is required to be reset before a modification is able to take effect. A reset means that the session is to be broken and restarted.

However, if a session is reset, routing and update messages are interrupted. The consequences of a reset include routing cache invalidation, disappearing routes and route instability cascades throughout the network. If the session is finally restored, the damage is already done.

Cisco Systems added a feature called *soft configuration* which enables administrators to alter attributes, policies and routes on the fly without killing an established session. As a result, the routing caches are not flushed and impact on routing is minimal.

A disadvantage of using soft reconfiguration is that it requires a set of unmodified routes from specified peers to be stored in local memory beforehand. Memory consumption could be high if an administrator decides to deploy a soft reconfiguration with a large number of peers. For each route learned from one of its peers, 250 bytes could be assumed to be required to store it in memory.

**BGP route refresh** To remove the disadvantage of memory consumption during a soft reconfiguration another solution was introduced. This solution, *route refresh capability*, utilizes BGP version's 4 Capabilities Negotiation to enable a dynamic means of requesting that a peer re-advertises all its prefixes learned from a specific peer.

This feature is enabled by default in IOS and should be supported by other BGP peer routers to be able to use the feature. The feature, when using soft reconfiguration, has no overhead of memory and CPU consumption. It also enables the router to iterate over all the prefixes learned from the peer and examine them. The prefixes can be applied to the new policies without a hard reset of the session.

**Route dampening** A mechanism to control route instability, such as route flapping, is route dampening. A flapped route which appears and disappears rapidly causes announcements and withdrawals to be sent repeatedly through the network. The network will receive a large number of traffic which could cause a link's bandwidth to exhaust and a higher CPU utilization of routers.

Dampening falls into two categories: a route is either *behaved* or *ill-behaved*. If a route is referred to as well-behaved it means that this route has shown a high degree of stability over a longer period of time. In contrast, an ill-behaved route is said to experience a high level of instability in a short period of time. If a route misbehaves it should be penalized in such a way that it is proportional to its future instability. Such an unstable route should be suppressed until there is a higher level of confidence that the route has become more stable.

To estimate the future stability of a specific route a route's history should be taken into account. To track its history it is important to count the number of times the route has flapped in a certain interval. If a route flaps it is given a penalty and when it reaches a predetermined threshold the route is suppressed. After suppression a route is still able to accumulate penalties. If a route flaps more frequently in a short period of time the route is suppressed faster.

Similar criteria are deployed to unsuppress a route and be able to announce again. An algorithm is implemented which is able to reduce the penalty exponentially over time. This algorithm is based on a set of predefined parameters by the administrator. In a Cisco environment the following parameters are applied:

- **Penalty** An incremental value which is assigned to a route each time it flaps.
- **Half-life** A configurable value which is the time period that must elapse in order to half the current penalty.
- **Suppress limit** A numeric value which is compared to the penalty. If the penalty's value is greater or equal than the suppress limit the route will be suppressed.
- **Reuse limit** Is also a configurable numeric value and compared against the penalty. However, if the penalty level is less than this limit, a suppressed route which is online will be unsuppressed.
- **Suppressed route** A route which will no longer be announced even when it is up.

- **History entry** A history entry is used to store flap information. For a router to monitor and compute a route's stability, it is crucial to store information regarding a route's flapping time. If a route has become stable again this entry would be useless and could be removed.

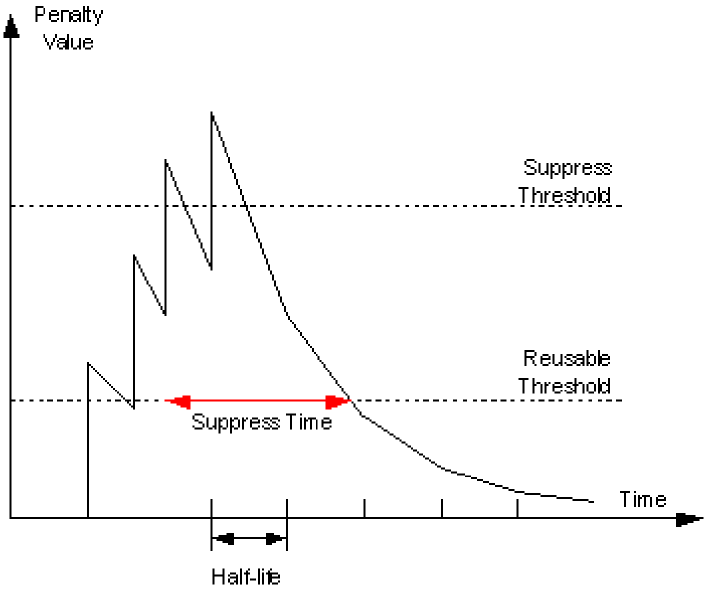

Figure 4: BGP route dampening [29]

Figure 4 reflects the complete process of penalizing a route if it flaps. The penalty's value is reduced exponentially due to parameters such as half-life. This parameter can be altered by the administrator if the route is known to flap frequently. Thus, a longer half-life would be more desirable to stabilize the network. If the value of half-life is set higher the penalty will in effect be reduced more slowly and suppress a route for a longer period of time.

### 2.2.3 Complexity

The Internet has grown in a various number of parameters: number of ASes, Figure 5 illustrates this increasing trend; number of prefixes; size of configurations; more edge than core routers; volume of traffic; number of interconnections; number of routes. However, each AS only has a partial view of the network and the Internet depends on other complex interdependent systems. This lack of information comes from: complexity and scale; information hiding properties of the routing protocol [30, 31, 32], ever changing links and paths between networks and the performance of those paths; security concerns, otherwise there is an improved accuracy for targets for people with malicious intent if they were to know the physical mapping [33]; cost of storing and processing all the data involved; commercial sensitivity, whether how and where networks connect is confidential; lack of good metrics for the networks as a whole [22].

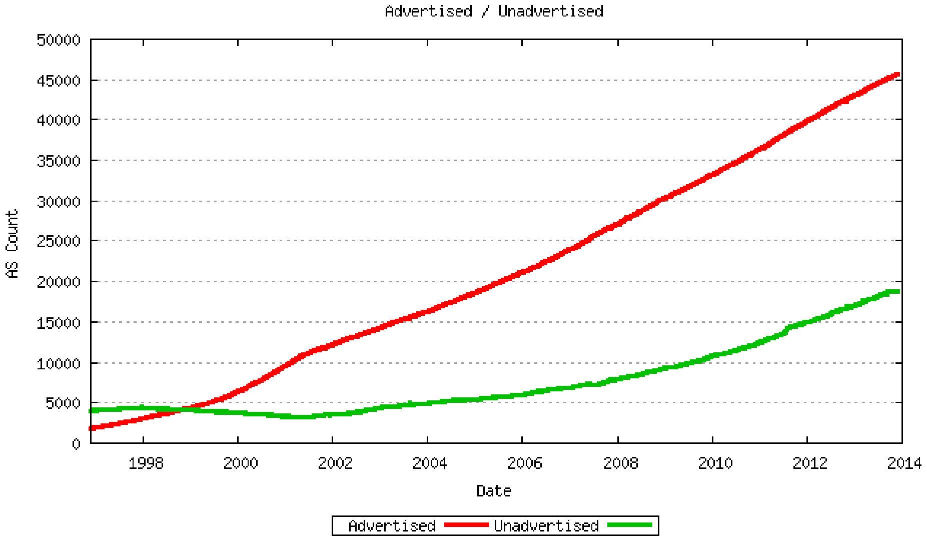

Figure 5: Number of advertised and unadvertised routes [28]

There are various approaches to complexity, such as: layering, object-oriented, structural or shifting complexity by hiding complexity from users and operators. Some define complexity as either: the network as a graph and its routing protocols; the number of spanning trees [34]; linear complexity of the graph's adjacency matrices [35] number of functions, modules or their dependencies and number of execution paths; dependable on configuration, rate of change and dynamic properties [36].

Since network architecture revolves around the concept of fitting the design of a network to its purpose, the question often asked is, ”which network will best fit the requirements”. One part of fitting a specific network design to the requirements is the issue of complexity, approached by different methods, as mentioned in the previous paragraph. An often raised question is, at which point in time would adding a specific protocol, policy or configuration make the network ”too complex”.

## 2.3 Project motivation

Since BGP is a crucial part of today’s Internet it is an important subject to study. This research regards network complexity as a systemic, rather than a component level, problem. Complexity should be measured in terms of all the multiple moving parts involved in the system. Complexity may be more than the complexity of each individual piece. To address systemic level problems two basic approaches exist: interfaces or continuums. By means of interfaces each piece of the system is treated as a *black box* and a complete understanding of the interactions between these black boxes has to be developed. Addressing the problem as a continuum, the impact of a single change or element to the entire system as a set of trade-offs is sought.

The approach described in this research focuses on continuums. In theory, modifying a network to resolve a particular problem (or a class of problems) could add more complexity than is expected or desirable. This might result in an increased likelihood of another class of problems. Discovering different continuums of trade-offs and determining how to measure them has become the key approach in both understanding and measuring systemic complexity.

**Reactivity versus Stability** In previous studies the speed at which the network’s control plane can react to a change in configuration, policies or topology has been a widespread focus [9, 37]. Such convergence in the control plane can be put into four essential categories:

- Detecting the change
- Propagating information about the change
- Determining the (new) best path(s) for destinations after the change
- Changing the forwarding path at each individual element in the network along the altered paths

Each of these categories could be addressed to improve the speed at which the network converges. However, some of these improvements come at the cost of increased complexity.

Changes in the network topology could be detected much quicker if faster echo (or hello) mechanisms, lower layer physical detection or other methods were used. However, these mechanisms can only be implemented at the cost of evaluating and managing false positives and high rates of topology changes. If, for instance, the state of a link in the face of a change can be detected in 10ms, the link could, in theory, change states almost 50 times a second. At this fast rate, tuning a network's control plane to react to such topology changes would be impossible. It is impossible because the network elements involved would have to inject a large number of information into the control plane at the same rate, destabilizing it, and, hence the network itself. Counter solutions in fast down detection techniques include some form of dampening mechanism.

Changes in the network topology have to be propagated throughout the network. This has to be performed so that each network element along the path is able to compute new forwarding tables. Routing information in high speed networks is able to propagate in tens of milliseconds. At this rate the network can propagate multiple changes in a short period of time. However, injecting information in such a short period could lead to overloading the network devices and processes which are participating in the control plane, and, in addition, create destructive positive feedback loops in the network. In order to avoid these consequences, most protocols for the control plane have to regulate the speed at which information about network changes is exchanged between devices. In recent innovations exponential back-off is implemented. These techniques manage the rate at which information is exchanged; the first changes are transmitted rapidly, while subsequent changes will be transmitted with a specified or calculated delay. Such techniques aim to control the destabilizing effects of fast information flows through the control plane. As a result, the added complexity of configuring and managing the rates at which the control plane can propagate information is controlled.

In order to find the best possible path through the network to any given destination all control planes require some form of algorithmic calculation. These algorithms are often lightweight, yet, require some amount of memory and computational power to execute and process all the information. If changes in the network are propagated at a high rate these devices could be overwhelmed. If they are overwhelmed and the running algorithms become processor or memory bound, a device could experience a computational failure altogether. This could, in turn, cause a more widespread network outage. To prevent such a general network outage and computational overloading, control plane protocols are designed to limit how often they can compute the best path through a network by using timers. These timers are also using exponential back-off, meaning the first computation is allowed to run quickly and subsequent computations will be delayed. Managing and configuring these timers is an added source of complexity to the network.

# 3 Methods

Below is a description of how the experiments and simulation are set-up and performed to answer the research questions.

## 3.1 BGP Simulation

In order to perform experiments with the simulator it had to be programmed as a computer program. In 2008, a scalable simulator [14] is developed which is capable of running on a Grid Computing cluster such as the DAS-4 [38] located at the VU. The simulator's implementation has specific requirements and a precise design. The design was mainly focused on scalability due to the size of the topologies to be simulated. The simulator was implemented following the KISS (Keep It Simple Stupid) design principle which dictates as goal that simplicity is key and unnecessary complexity should be avoided. This is achieved by simplifying BGP's protocol. This simplification is one of its key features: instead of focusing on intra-domain communication, both network and protocol are abstracted.

Design requirements were:

- **Scalability** BGP will be simulated on a large scale. A scale such as today's Internet which includes a large number of ASes and interconnections. These properties are achieved by avoiding bottlenecks and make efficient use of the limited resources.

- **Efficiency** For a large network to be simulated every decision was made with efficiency in mind. Such as reducing: number of calculations; messages exchanged; memory used.

- **Relaxed accuracy** Since the model is abstracted from real-world details, hence, accuracy is lost in the results. This is a trade-off that had to be balanced out. BGP's state nor the exact content of messages is necessary. The simulator is focused on generating statistical data and the impact of AS policy changes in the network.

- **Extensibility** The simulator is implemented in such a way that it is highly extensible because the kind of experiments to be performed in the future were not known a priori. Further studies, such as this study, are able to include or implement components which are found to be necessary. For this study, the simulator was slightly altered to accommodate the research questions.

ASes are separated from one another as much as possible. This means that each AS is running in its own BGP process thread and behaves independently from other ASes. They follow the same algorithm and are able to generate and process BGP's principles, such as announcements and withdrawals. BGP's behaviour is thus generalized at a certain level and enables high scalability. Since this comes at a loss of accuracy it is difficult to analyze the behaviour of a single AS. However, global BGP behaviour can be analysed and different trends can be observed.

**Topology** In addition, the simulator is able to take any kind of topology as input. This enables experiments to be performed on different topologies and be able to compare topology properties and their influences on BGP's behaviour. BGP's stability and scalability are issues in two different aspects: increasing routing table size and increase rate of BGP updates, also called churn [4]. In this category the question would be, which topological growth scenarios would lead to higher and faster churn increase for different failure types. This research has used a simple and controllable topology generator [4] which satisfies four basic but fundamental characteristics of the Internet graph. These four characteristics are:

- **Stable topological properties**: ASes in the Internet graph form a hierarchical structure. Customer-provider relationships are formed so that there are normally no provider loops.

- **Power-law degree distribution**: the degree distribution in the Internet topology has shown to follow a truncated power-law [12]. Which means there are few very well-connected ASes while the majority of ASes have only few connections.

- **Strong clustering**: ASes in the Internet are grouped together in clusters, with ASes in the same cluster more likely to be connected to each other. One reason for this clustering is that networks operate in different geographical areas.

- **Constant average path length**: recent measurements show that in spite of tremendous growth in the number of ASes, the AS-level path length has stayed virtually constant at around 4 hops for the last 10 years [39].

In the generated Internet graph each AS is modelled as a single node and connections between two neighbouring ASes as a single logical link. Input parameters to the generator have operational semantics. Instead of specifying abstract graph properties such as clustering coefficient, betweenness or assortativity of the topology, the topology is defined in a more hands-on real-world related manner. For example: how many providers an AS has; how likely it is to peer with other types of ASes; in which region an AS is located; the multi-homing probability of certain types of ASes, type of AS (provider, customer, stub).

## 3.2 Events

In order to analyze BGP's stability and complexity a number of events are implemented to be executed by ASes. These events model different real-world scenarios such as a disconnected link, crashed BGP speaker or flapped route. Each of these scenarios could have a different impact on the stability of the whole network.

These events are:

- **Announcement** At the start of the simulation each AS has a number of prefixes it owns. These prefixes will be announced to their neighbours. However, it is also possible to craft an announcement to be announced by a specific AS at a certain time. These crafted messages are regulated by the coordinator.

- **Withdrawals** When necessary, ASes are able to send withdrawals for no longer reachable ASes. A withdrawal can also be crafted.

- **Log** If a log event is initiated each AS will write specific data measurements, such as received updates and current number of routes, into memory. At the end of the simulation each AS writes the data to a local file. These local dumps are combined to perform in-depth analysis of BGP's behaviour with.

- **Disconnect link / AS** It is possible to drop a link between two ASes or disconnect one AS from all its neighbours. In both cases all the routes and prefixes received through this link will be withdrawn. Withdrawals will be sent to all the other neighbours that are still reachable.

By combining these events it is possible to mimic certain real-world scenarios. These scenarios could help understand how BGP's complexity and stability is formed. The next section will describe how the experiments are set-up.

## 3.3 Experiments

Since the simulator will be running on a Grid cluster the number of compute nodes has to be specified. Currently, 48 or 58 compute nodes (depending on the size of the topology and type of experiment) plus one for the coordinator are assigned for the simulation. A simulation is completed in under a few minutes and is repeated a number of times for higher accuracy.

### 3.3.1 Configuration

The simulator has a number of parameters which define its behaviour:

- **Time scaling** How much faster should the simulation run with respect to real-time. A value of t means that with every second the simulation's time is advanced by t seconds. This value is currently set at 50.

- **MRAI Percentage** The percentage of ASes which have MRAI timers during the simulation. If the percentage is, say 80%, 80% of all the ASes (picked at random) will have a MRAI timer of 30 seconds.

- **Flap distribution** How many ASes should be Cisco-like (as opposed to Juniper-like), when using route flap damping.

- **Flap percentage** How many routers should use route flap damping. Currently, the default is 70%.
- **Internal BGP Max value** This value defines the maximum convergence time for internal BGP and is set at 30 seconds.

### 3.3.2 Dynamics

Experiments would be static if there were no events. An initialization phase is implemented in which ASes announce their prefixes for the first time making the whole network more dynamic. A log event is initiated after a period of time if the network has converged to a *stable state*. After this state has been reached a failure will take place locally in the network. The network will attempt to reach a stable state (converge) again and an after-event log is performed. These logs contain specific data measurements taken during the experiment.

### 3.3.3 Data measurements

In order to observe the effects of network failures on BGP's stability and complexity, data is collected during experiments. The data, which is stored in logs, contains the following measurements:

- **AS characteristics**

  **Degree** Number of neighbours each AS has.

  **Clustering coefficient** How well do the neighbours of an AS know each other. A CC of 1 means that all the neighbours of an AS are also connected with one another. A CC of 0 implies that the neighbours of an AS are not connected with one another.

  **Tier** Which Tier an AS is located in.

- **Measured observations**

  **Received updates** Total number of BGP updates received, both announcements and withdrawals.

  **Received announcement** Total number of received announcements.

  **Received withdrawals** Total number of received withdrawals.

**Total routes** Number of routes an AS has currently installed.

**Total lost routes** Number of routes an AS had to either replace or remove.

**Reachable ASes** Number of ASes that are still reachable through all stored routes.

To give better insight into BGP's dynamics these measurements are used to plot detailed graphs. Such graphs contain, for each Tier, an average number over all the experiments. In each experiment a number of links (1, 4, 8, 16, 32, 64) are disconnected in a topological region. By disconnecting a number of links in a specific region it mimics a catastrophic failure in, for example, a data centre. If more links were to be disconnected a breaking-point might be found with which the network's stability is decreasing. An instability in which more updates have to be propagated and the stress on certain Tiers or ASes would be increased.

### 3.3.4 Set-up

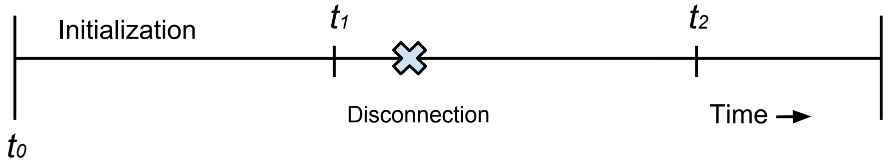

Figure 6: Experimental set-up

In Figure 6 the experimental set-up is displayed. For each experiment the same data measurements are performed at certain intervals. The first measurement point is after the whole network has been booted up, announced their prefixes and has converged to a stable state. A short time later the links are disconnected and the network has to converge again. If the final converging point has been reached the data is stored and processed for analysis. This analysis involves all the data measurements mentioned above and compares the three Tiers with each other. A Tier might be affected in a different manner, for example, by receiving more announcements or withdrawals.

**Below, the experiments to be performed in an Internet-like generated topology [4] are described**

- **Disconnecting random links**

  In the first set of experiments a Tier 1 AS is chosen as the central-AS in each topology. This AS will be the center of an experiment and links, random between different Tiers in its vicinity at max 2 hops distance, will be disconnected.

- **Disconnecting links between Tier 1 ASes**

  For the second set of experiments a Tier 1 AS is chosen again but in these experiments the links to be disconnected are only among Tier 1 ASes.

- **Disconnecting complete ASes**

  In the third set of experiments 1 AS is disconnected and the Area of Impact (AoI) is computed. The AoI is defined as the number of received announcements or withdrawals at each depth with a central-AS as point of view. This means the direct neighbours are at depth 1 and the neighbours of the neighbours (excluding already visited neighbours) at depth 2, etcetera, until all ASes in the network have been visited. By plotting the received number of announcements and withdrawals at each depth the impact on the whole network can be analysed. The impact on the neighbouring ASes compared to those further away, in terms of received update messages, can be shown.

**Below, the experiments to be performed on a real-world Internet topology provided by CAIDA [19] are described**

- **Disconnecting the Internet**

  In the last set of experiments the topology consists of a large number of ASes typical for today's Internet. The number of ASes ranges from 12000 to 42000 ASes and can be found on CAIDA's website [19]. ASes consider their relationships as proprietary information and do not make them public. Therefore, to generate an Internet topology of all ASes, researchers have to rely upon AS relationship inference algorithms. One such algorithm, as defined by CAIDA, examines the customer cone of an AS. This cone consists of the set of ASes that can be reached from each

AS while following only its customer links. The size of the customer cone reflects the number of ASes that pay, either directly or indirectly for transit, and provides a more reliable metric of the size of an AS than its degree.

The experimental set-up for disconnecting the Internet is the same as the second set in which links between Tier 1 ASes will be disconnected. By disconnecting these links it is possible to analyze the impact on a larger topology.

# 4 Results

Below is an extensive description of all the results. Firstly, characteristics of the topologies are described. Secondly, Tier distribution and path diversity are explained. Lastly, results of disconnecting a number of links and a complete AS are given.

## 4.1 Topology characteristics

In order to perform simulations, several topologies were generated with an Internet-like topology generator [4]. These topologies vary in number of ASes (1000, 2000, 4000, 8000) and number of ASes per Tier. However, these topologies have many characteristics in common.

Table 1: Distribution of ASes over all Tiers per topology

| | **1k** | **2k** | **4k** | **8k** |
|---|---|---|---|---|
| Tier 1 | 13 | 13 | 13 | 13 |
| Tier 2 | 44 | 78 | 231 | 675 |
| Tier 3 | 943 | 1909 | 3756 | 7312 |

**Tier distribution** One such common characteristics is the distribution of ASes over the Tiers. In Table 1 the distribution of ASes over all Tiers per topology is given. The number of Tier 1 ASes remains constant at 13. In today's Internet, 13 ASes are considered to be the top Tier ASes. These ASes provide connectivity to all their customers, peer with one another and do not have to pay for transit services. The number of ASes in Tier 2 slowly increases and Tier 3 has the largest portion of ASes. The chance of an AS being in Tier 2 is 30%, otherwise, the AS will be placed in Tier 3. ASes with only one neighbour are always considered to be in Tier 3.

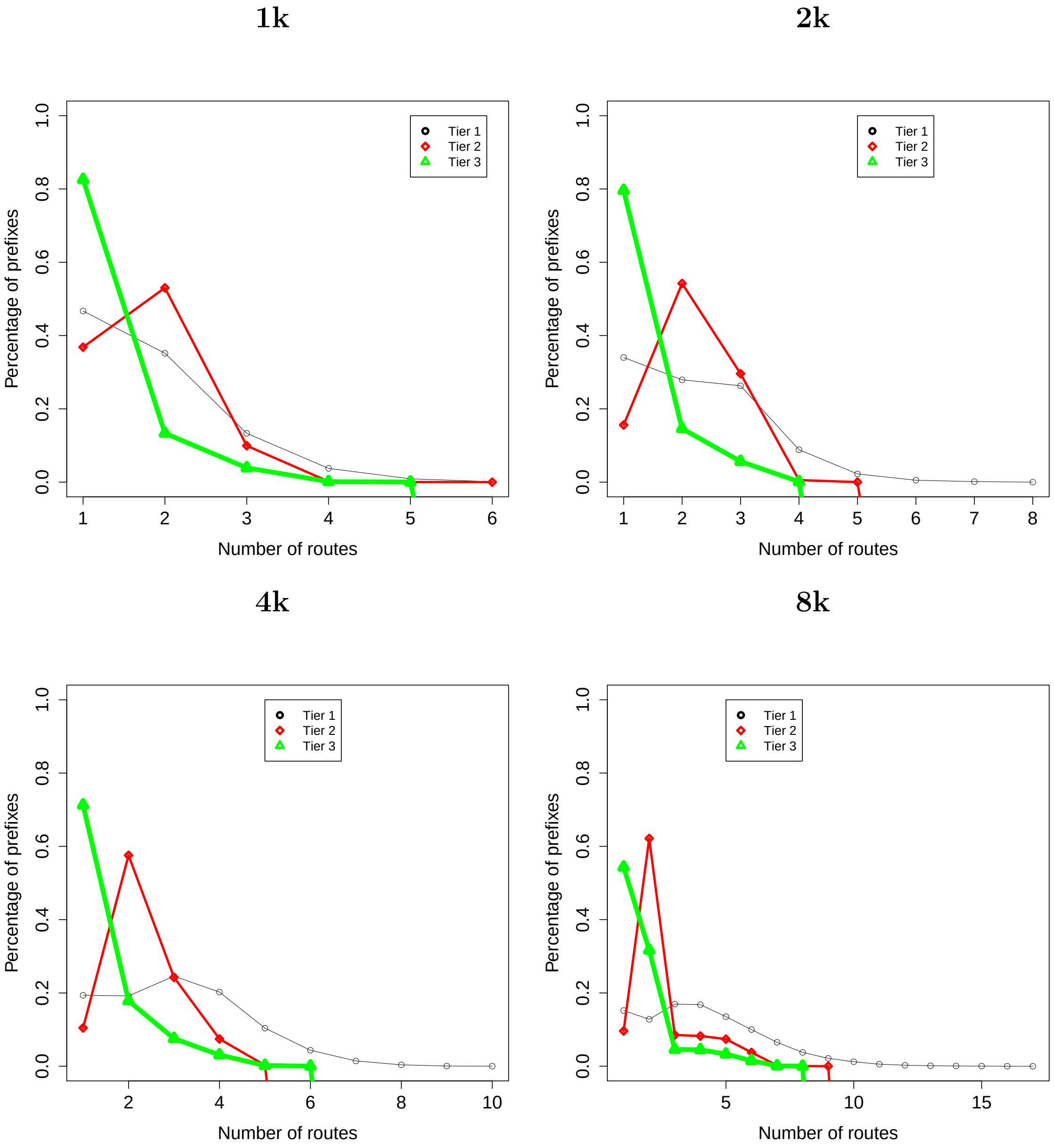

Figure 7: Path diversity versus number of disconnected links

**Path diversity** The path diversity shows the percentage of prefixes for each Tier with a certain number of routes. For example, Tier 3 ASes have 82% of their installed prefixes with only 1 route, while Tier 1 has 48% and

Tier 2 38%.

Path diversity could be an indicator of the robustness of an individual AS. An AS could be said to be robust if it knows a large percentage of prefixes with multiple (back-up) routes. If one route fails, the AS could decide to use a back-up route and all would be well.

With an increasing number of ASes and links the percentage of prefixes with multiple routes also increases. The total number of routes is topological dependant but the percentages show similarities over all topologies. Tier 1 ASes commonly have a long tail because these ASes are well-connected. The total number of prefixes with more than 1 route increases to around 80%. Tier 2 ASes, located between Tier 1 and 3, seem to have very few prefixes with only 1 route (10-40%) and many prefixes with 2 routes. Only a few prefixes in Tier 3 have more than 1 route because most of the ASes only have 1 neighbour from which they could have received their prefixes.

## 4.2 Disconnecting random links

In the following set of experiments links were disconnected at random.

Figure 8 contains the average route length. This average is calculated over all installed prefixes and their current route. Figure 9 shows the average number of received announcements. Figure 10 displays the average received withdrawals. These plots are for topologies with 1000, 2000 and 4000 ASes in which links are disconnected between random ASes at max 2 hops distance.

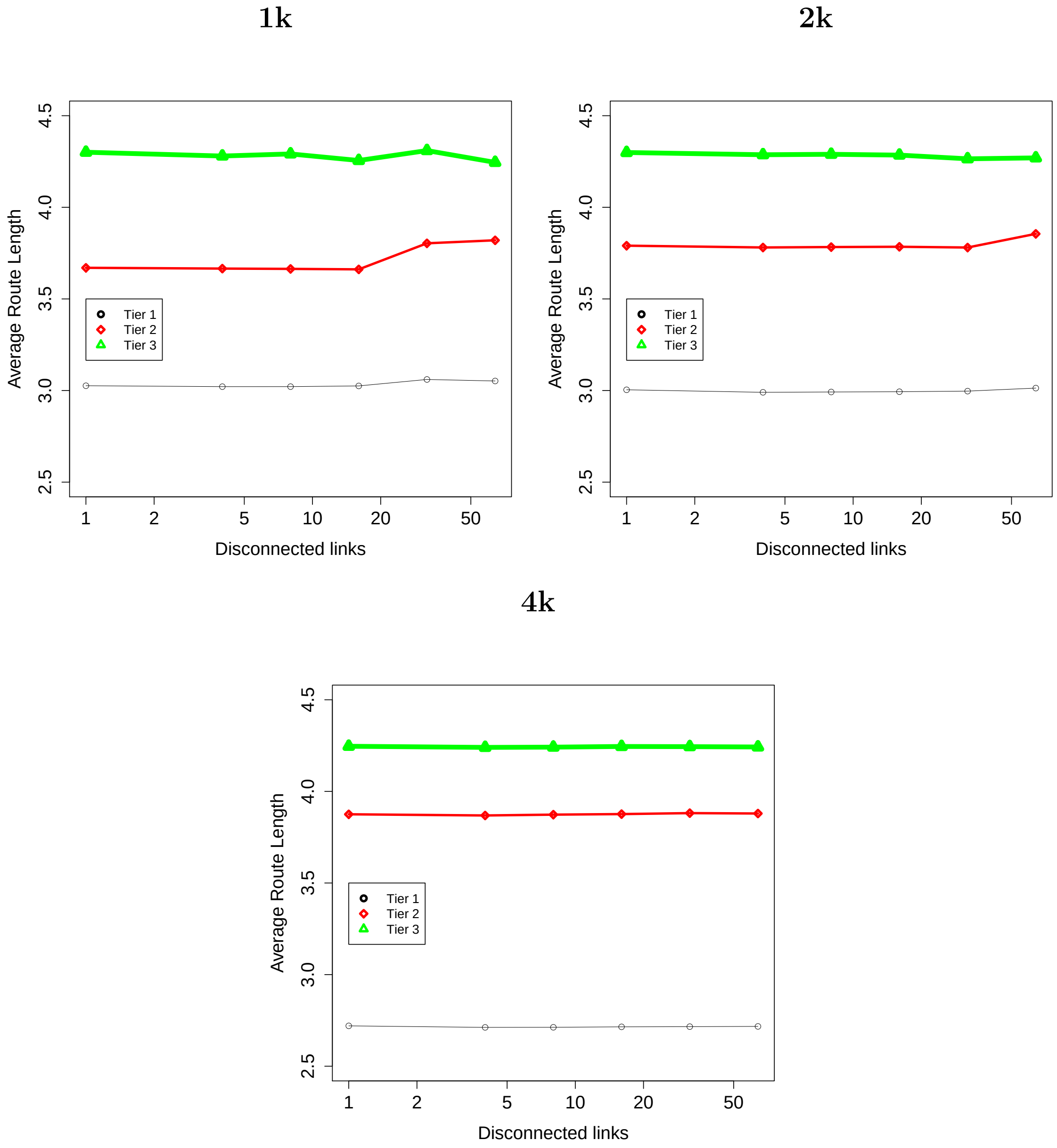

Figure 8: Average route length versus number of disconnected links

The average route length for each Tier remains on average the same for all topologies: 3.034 for Tier 1; 3.715 for Tier 2; 4.28 for Tier 3. These numbers show that the average route length is distinct for each Tier. This phenomenon could be explained by characteristics of the topology. Tier 1 ASes are located in the center of the topology which commonly have a higher degree, are better connected and have, on average, shorter paths. Only in the topology of 1000 ASes seem to be a significant change of average route length for both Tier 2 and 3. In contrast, the average route length remains almost constant for both 2000 and 4000 and might be an indication of robustness. In case of failures, there will be no decrease in connectivity and traffic could still flow efficiently.

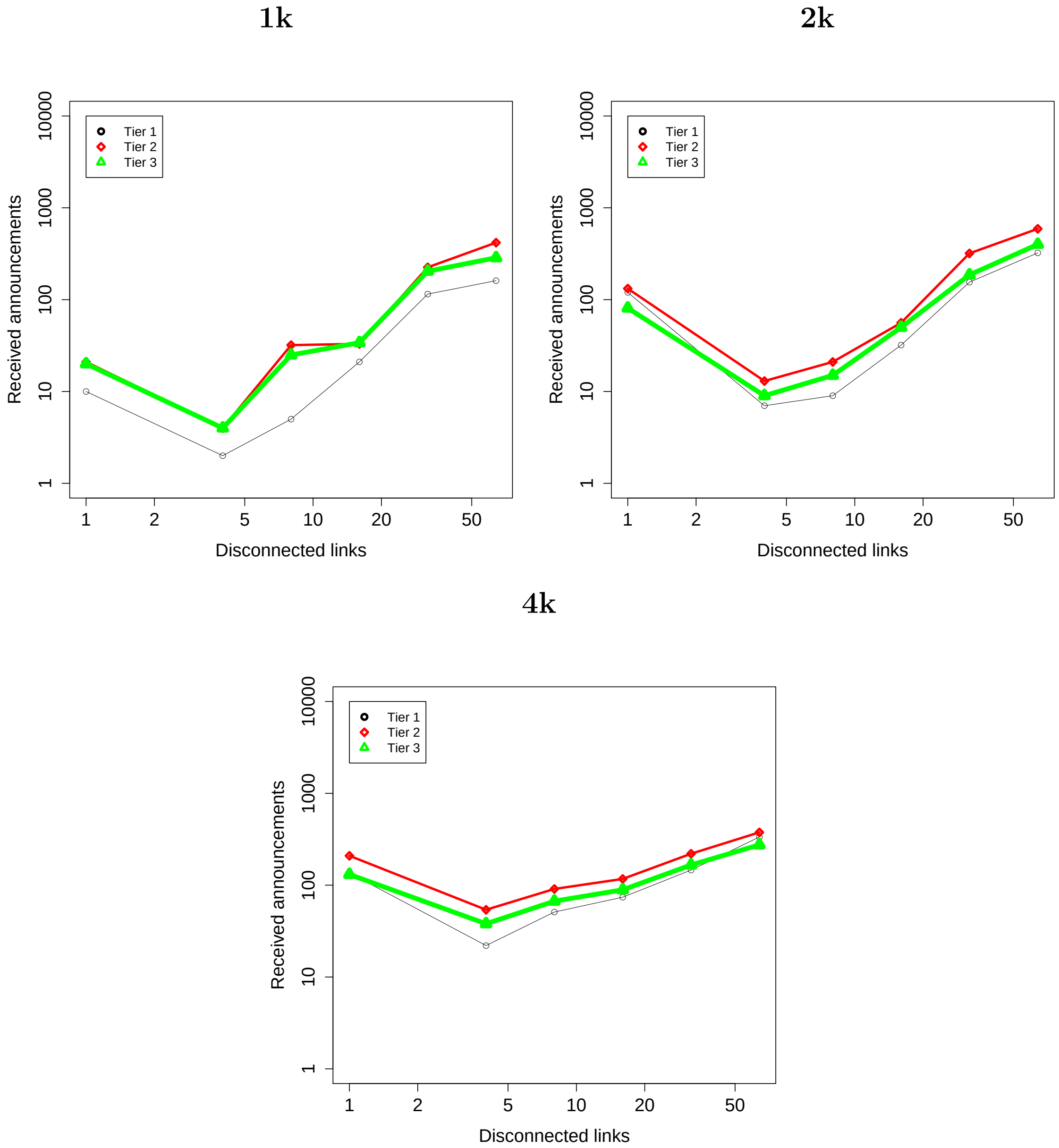

Figure 9: Average received announcements versus number of disconnected links

The number of announcements sent gradually increases by the number of disconnected links. Yet, a consistent drop from 1 to 4 disconnected links could be seen since the 1 disconnected link is between the targeted AS and a direct neighbour. In contrast, 4 to 64 disconnected links are located farther away from the center AS.

Tier 2 received most of the announcements and for both 32 and 64 disconnected links the number of received announcements grows rapidly. An overall trend is that Tier 2 ASes receive more announcements. This could be explained by means of the hierarchical structure of the topology: Tier 2 ASes are located in between Tier 1 and 3 ASes and could thus receive more announcements.

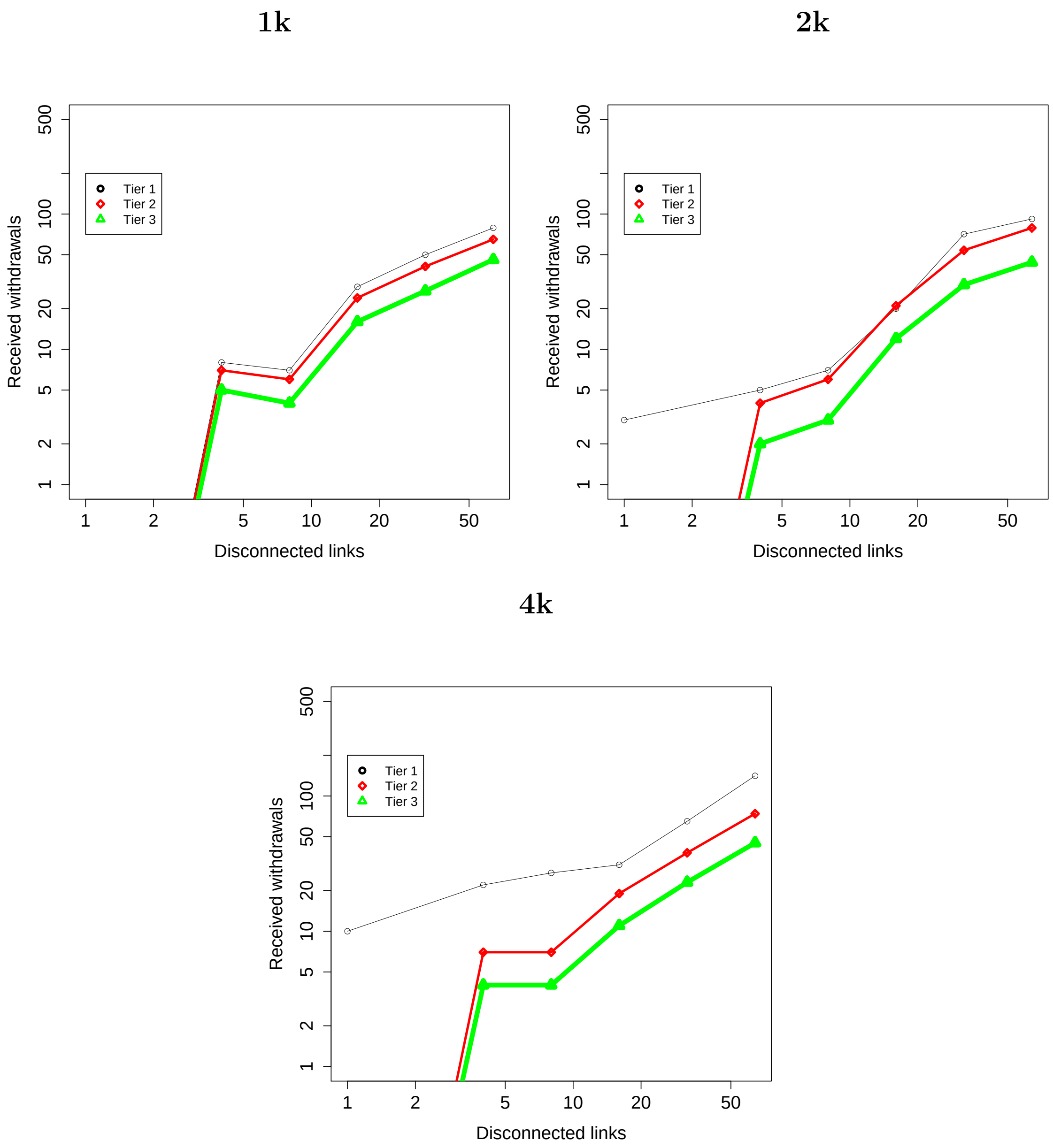

Figure 10: Average received withdrawals versus number of disconnected links

In case of withdrawals, more are sent during 16, 32 and 64 links being disconnected. This could mean that more alternative paths disappear, resulting in a withdrawal. Tier 1 ASes have a tendency to receive more withdrawals and act as a catch-all because of their higher degrees and path diversity.

Since the links are chosen at random, there is a possibility that links are disconnected between Tier 1 - 2 or Tier 1 - 3. This would not make the experiment local but more widespread and will result in more parameters.

## 4.3 Disconnecting links between Tier 1 ASes

Below are the results for the second set of experiments in which links are disconnected between Tier 1 ASes. By disconnecting links between Tier 1 ASes various random parameters are eliminated and focus is switched to regional events. Such regional events could be a catastrophic disaster or power-failure in a data center. A failure on a regional scale could possibly impact the global behaviour of the network.

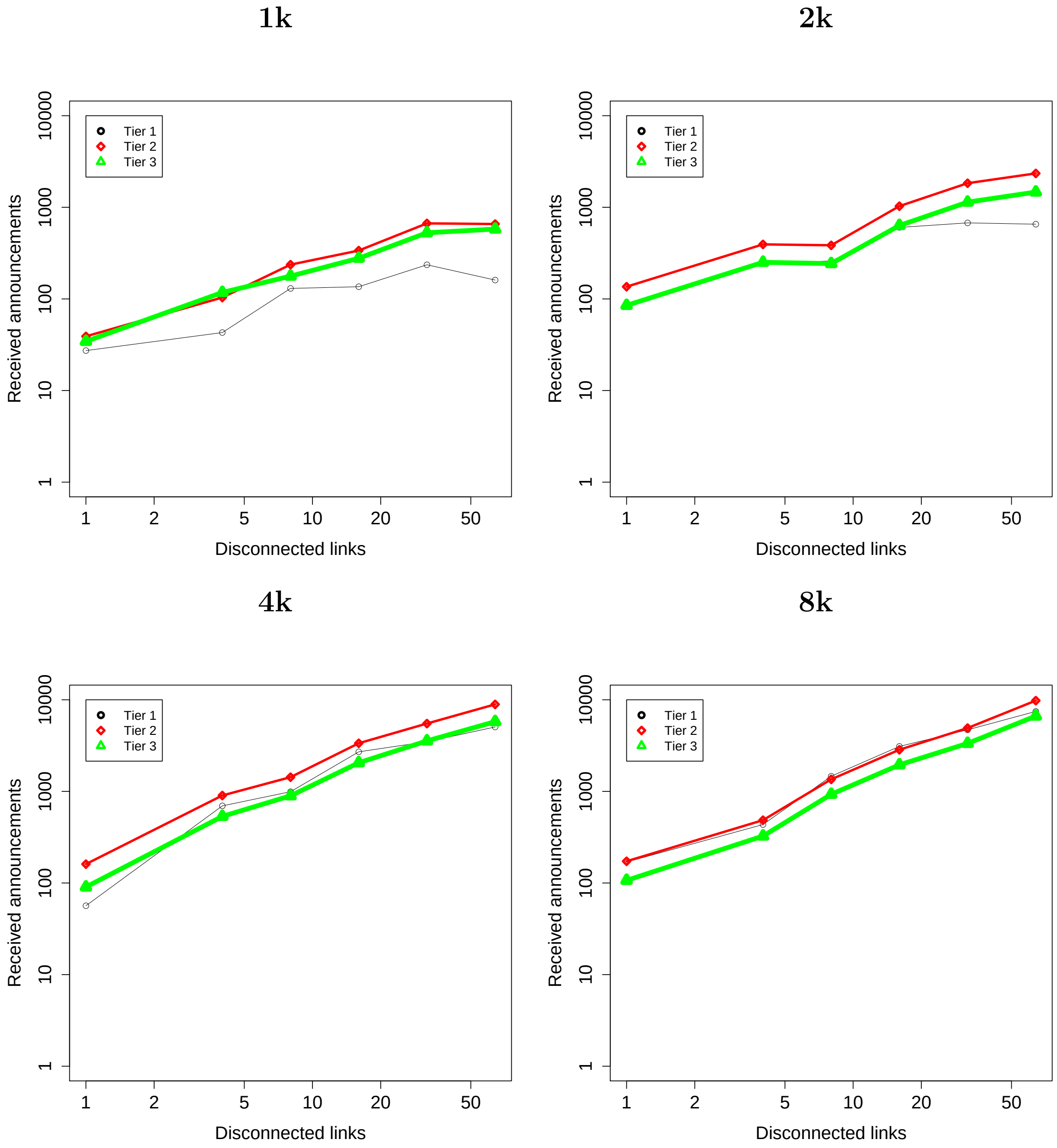

Figure 11: Average received announcements versus number of disconnected links

In Figure 11 the average received announcements is shown. The number of messages received gradually increases for each of the topologies with a max of almost 10000. For each Tier the number of received announcements is almost equal. Since more links are disconnected, more announcements have to be sent. All these announcements have to be propagated to each AS in the network. Each AS will know about the changes and decide to route traffic to more preferred routes.

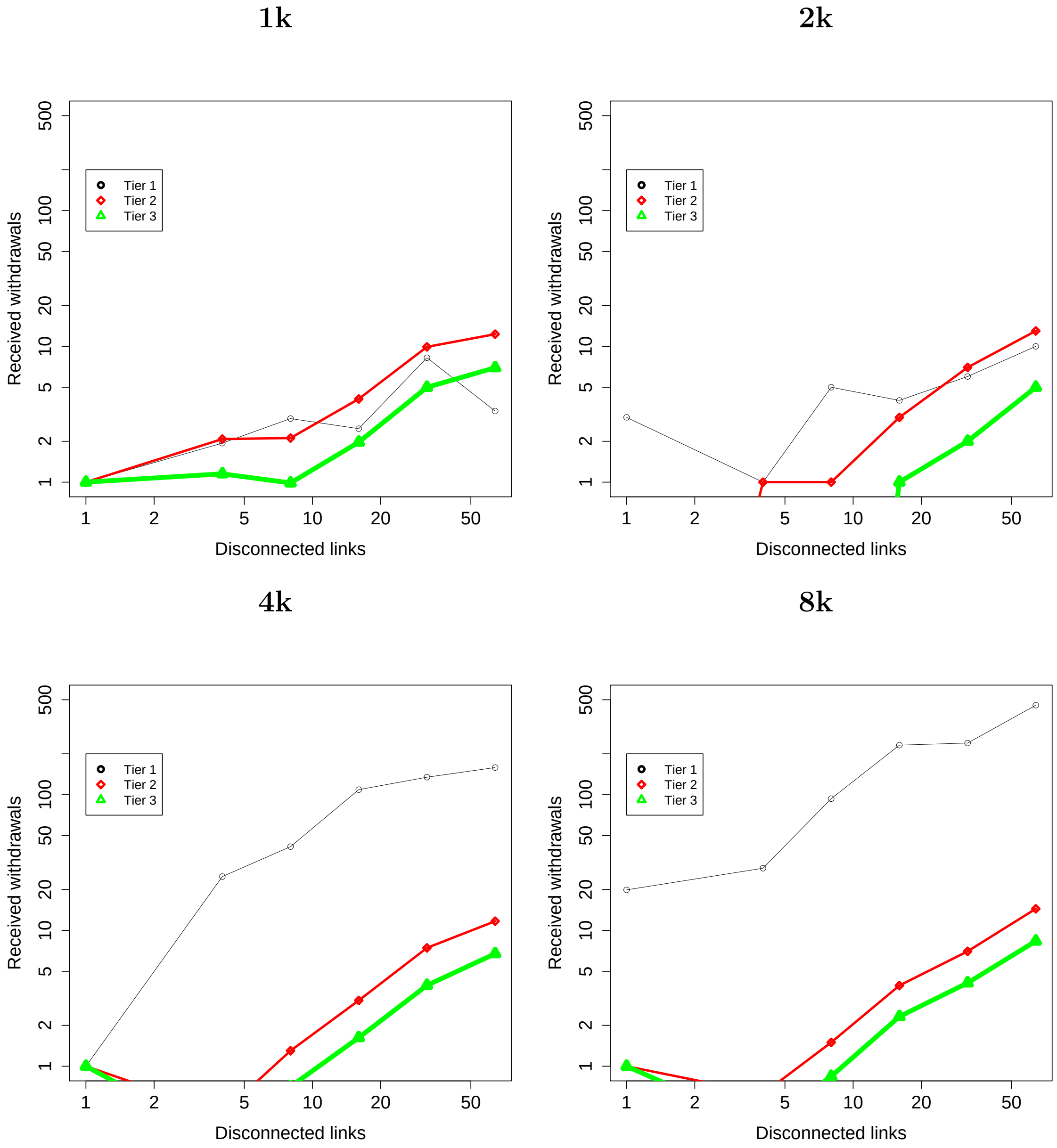

Figure 12: Average received withdrawals versus number of disconnected links

In Figure 12 the average received withdrawals is shown. Both 1000 and 2000 have on average lower than 20 received withdrawals. This means that even with a larger number of disconnected links only few withdrawals are exchanged.

However, in topologies with 4000 and 8000 ASes a distinct trend could be seen. This trend shows that Tier 1 ASes receive more withdrawals, around 100-500. An explanation for this phenomenon can be found in the path diversity and topological structure. The path diversity shows that more than 80% of all the known prefixes by Tier 1 ASes have more than 1 route. If one prefix would have to be withdrawn, there is a high chance its route could be replaced with an alternative route. This could also be an indication of the network's stability. In this case, Tier 1 ASes know alternative routes and propagate this knowledge to both Tier 2 and 3.

## 4.4 Disconnecting complete ASes

In the following experiments one Tier 1 AS was disconnected from all its neighbours. This means that each neighbour would have to withdraw all received prefixes received from this AS. If there are alternative routes to the otherwise lost prefixes it will result in an announcement. In Figure 13 the average received announcements and withdrawals and path length are shown for each topology.

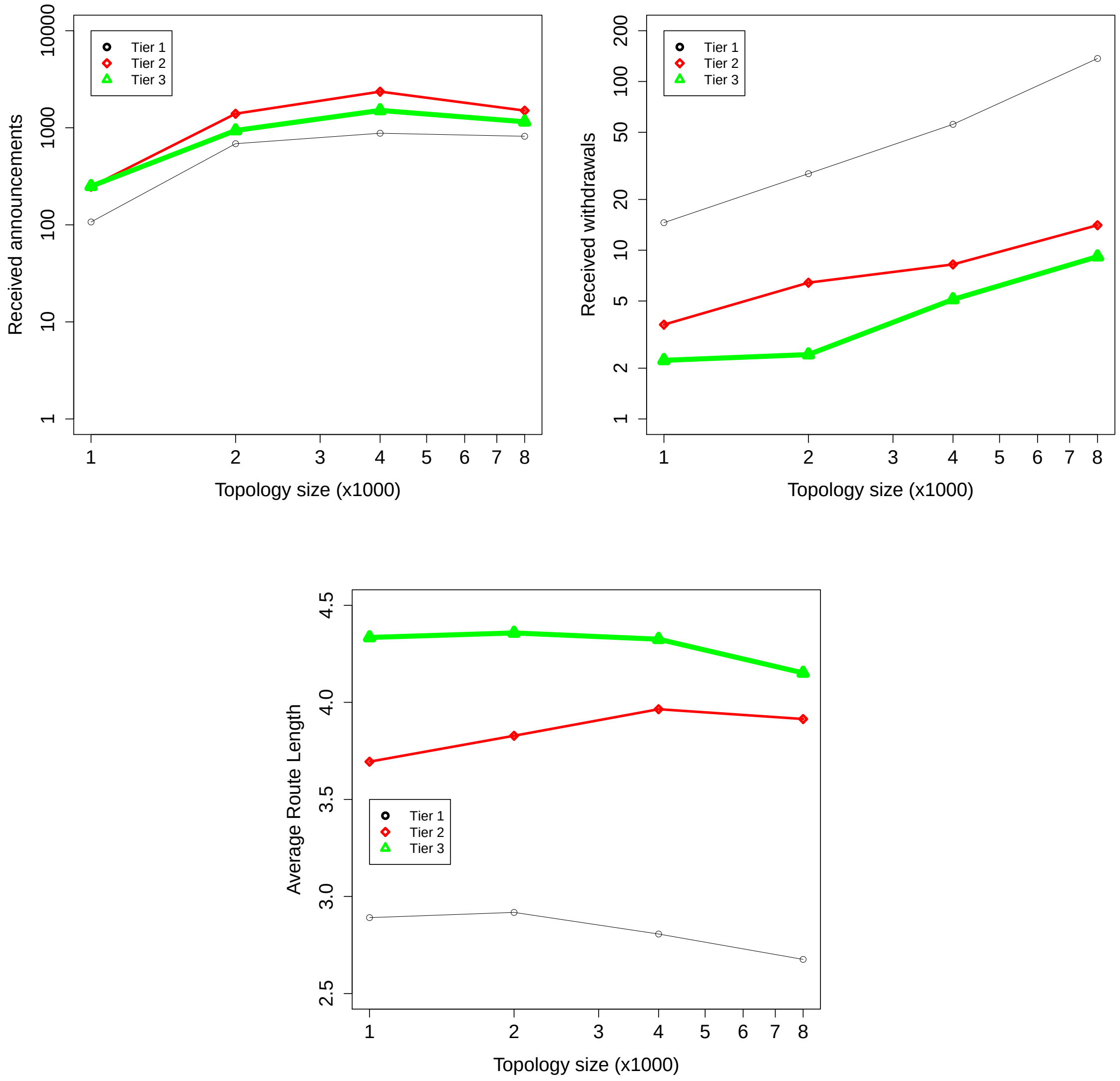

Figure 13: Averages for disconnecting nodes in each topology

The underlying question in this set of experiments is concerned with the network's stability in the face of a complete AS failure, as well as an increasing number of ASes. In case of 1000 ASes there are on average 100-250 announcements received and for larger topologies this number increases to approximately 1000 messages. Since this number of messages remains the same for the other topologies, there might be a maximum to the number of messages received. However, for the average number of withdrawals an increasing trend can be seen. This trend shows that the 12 remaining ASes in Tier 1 receive most of the withdrawals. Since these ASes are located closer to the disconnected AS there is a higher probability they receive more withdrawals overall. This trend could also be seen in the experiments with disconnected links between Tier 1 ASes.

The average route length shows different trends for each Tier. Tier 1's average route length decreases with larger topologies. In comparison, this number increases for Tier 2 ASes but shows a minor decline for a topology of 8000. These differences could be related to the structural characteristics of the topology. In topologies with a larger number of ASes there are more links overall. Furthermore, BGP prefers routes which are shorter and more optimal.

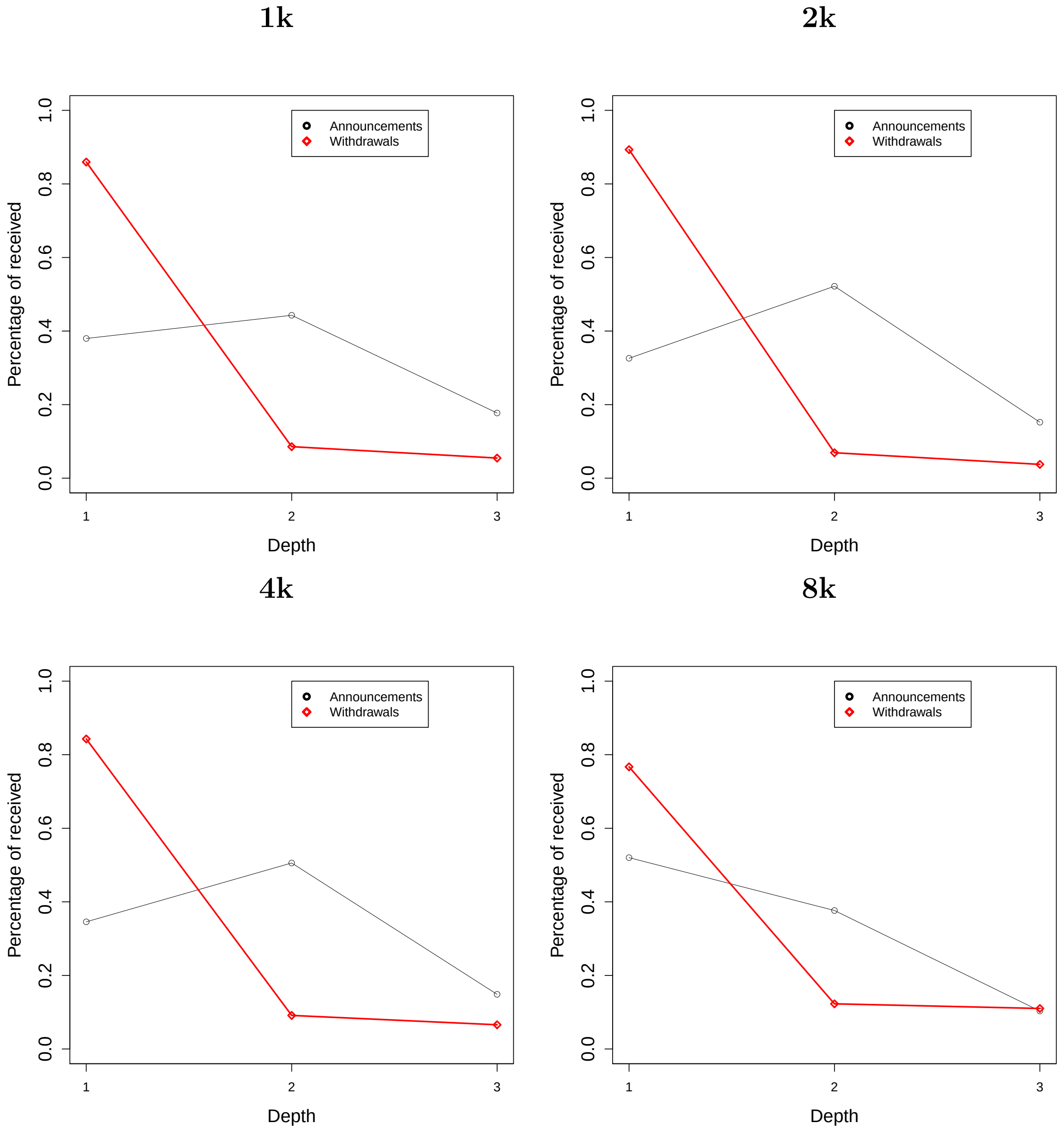

Figure 14: Area of Impact in each topology

Figure 14 depicts the area of impact for all the topologies and shows the percentage of received announcements and withdrawals at each depth. At depth 1 are the direct neighbours of the disconnected AS. These direct neighbours receive around 80% of all the withdrawals and less than 10% is received at depth 2. Since more than 60% (about 50% for a topology of 8000 ASes) of all the announcements is received at depth 2 and 3, most of the withdrawals received at depth 1 are converted to announcements. This trend can be seen in each topology.

## 4.5 Disconnecting the Internet

Unfortunately, due to limited resources on the DAS-4 [38] this last set of experiments could not be performed.

# 5 Conclusion

This study focused on BGP's stability and complexity and investigated how the network would react to different types of failures.

## 5.1 Summary

In summary, BGP is de facto routing protocol for today's Internet and is an important subject to study. Autonomous Systems (ASes) are different regions with administrative control and are interconnected with either provider, customer or peering relations. More interconnections are added on a daily basis and various sources of instability in the network exist. In case of route flapping, a BGP-speaker advertises a destination network in quick sequence. The convergence time is the time it costs for the whole network to reach a stable state in which each router maintains the same topological information. With an increasing number of Autonomous Systems memory in routers becomes insufficient. Routers have to hold over 400.000 prefixes and this number shows an increasing trend. Network upgrades and routine maintenance are required from time to time but could lead to instability in the network.

## 5.2 BGP's Complexity

BGP's complexity is discussed and shown to be multi-fold. Not only does the number of ASes increase, as well as configuration sizes of routers, volume of traffic, number of interconnections, routes and edge routers. Due to the limited view of routers and dynamic behaviour of BGP, another layer of complexity is added in which properties of the network could be hidden. How and where networks connect is confidential and sensitive information.

The complexity of BGP could be defined as all the moving parts of the network. With each new link or AS added, the overall complexity increases. Each AS maintains a partial view of the complete network and attempts to converge to a stable state. This stable state might not be reached at all if new connections are added or if failures keep occurring.

## 5.3 Topology

A scalable, efficient, accurate and extensible simulator is used for simulating real-world scenarios. Synthetic topologies with different sizes are generated following an Internet-like structure: customer-provider-peer relations exist; few ASes have a high number of neighbours and a large number of ASes have only few neighbours; core ASes are clustered together; the average path length is kept at a constant rate. Each AS is modelled as a single node and connections as a single logical link. The simulator is capable of sending announcements, withdrawals, disconnect links and nodes completely. However, these topologies are very limited and the depth from a center AS is max 3. This means that the edge of the network could be reached within 3 steps.

## 5.4 BGP's Instability

To investigate BGP's instability two failure types were deployed in the simulator. These two types are: disconnecting links between Tier 1 ASes and disconnecting a Tier 1 AS completely from its neighbours. The experimental set-up consists of an initialization phase in which each AS announces their prefixes. If the network has converged to a stable state each AS computes a log, including number of received announcements and withdrawals. After the first log a number of links are disconnected. The network attempts to converge again and an aftermath log is taken. With these logs statistical analysis could be performed.

## 5.5 Path Diversity

Path diversity has shown that the number of routes per prefix might depend on the Tier the AS is located in. A large percentage of prefixes received by

Tier 3 routes only have one route. In addition, Tier 1 ASes have more than 50% of their prefixes with more than 1 route. Path diversity might be an indicator for the network's robustness. If one AS has a large percentage of prefixes with more than 1 route the AS could be said to be stable. If one prefix has to be withdrawn there might be a back-up path which could be installed instead.

If more links are disconnected both the average number of received announcements and average number of received withdrawals show an increasing trend. However, most of the withdrawals are received by Tier 1 ASes. An explanation for this phenomenon is based on the path diversity: these ASes have more alternative paths and are able to act as a catch-all. This phenomenon might be an indicator for the network's stability: if an AS is said to be stable it might have a higher path diversity than other ASes. If a large portion of the ASes have a significant path diversity the complete network could be said to have a high degree of stability.

By disconnecting a Tier 1 AS completely, the same trends for average received announcements and withdrawals could be seen. Yet, there might be a limit to the average number of received messages.

## 5.6 Area of Impact

The Area of Impact (AoI) illustrates the average number of received announcements and withdrawals at each depth with the disconnected AS as centre point. Direct neighbours receive 80% of all the withdrawals while less than 10% is received at depth 2. The AoI might be an indication of the stability of the whole network and how widespread a failure could be. A desirable effect would be a decline in number of messages with increasing depth. A small AoI would indicate that a failure has been local and the network remains stable. By remaining stable in the face of failures the network's scalability increases as well.

## 5.7 BGP's Stability

BGP's Stability could be defined by means of both the Path Diversity and Area of Impact. The network could be said to be stable if a large portion of all ASes have a significant Path Diversity. This means that many ASes should have a sufficient number of alternative paths to replace withdrawn routes. The Area of Impact would have to be local in order to maintain a stable network. If a failure occurs, only a few ASes would notice it and the network's stability is guaranteed.

### 5.8 Future research

This study forms a basis for future research into BGP's complexity and stability and how local changes could affect the whole network. The network, if faced with failures, reacts differently depending on the number of ASes. Tier 1 ASes, in a topology of 4000 and 8000 total ASes, for example, receive most of the withdrawals. This, in turn, could result in higher stability for the other Tiers due to their higher path diversity. However, a higher strain would be put on the Tier 1 ASes to process all the update messages.

During this research many different parameters were revealed which could be investigated. To reduce the scope, focus was placed on using one particular topology setting but with different sizes. Future research could dive into even larger topologies; different topology settings such as higher clustering in the core or at the edge; and disconnect ASes in either Tier 2 or 3; include more prefixes per AS to model the Internet in more detail. Larger topologies could be extracted from CAIDA [19] and used as input for the simulator. The simulator would then be able to simulate real Internet-like topologies and scenarios. In addition, BGP's stability might be sensitive to certain topological properties which could be revealed during future experiments.

It is important to ensure future studies of BGP and the Internet's stability. With an ever increasing number of ASes, connections, traffic and users, connectivity has to be guaranteed. If the increasing strain on the existing infrastructure induces failures, the overall stability could be reduced.

## 6 Acknowledgements

Hereby I would like to thank: Benno Overeinder for his dedication, feedback and patience; Spyros Voulgaris for his useful insights covering large scale computing, distributed systems and network topologies; NLnetLabs for providing a great working atmosphere and experience; Femke for her growing support during this project.

# References

[1] M. Schuchard, C. Thompson, N. Hopper, and Y. Kim, “Taking Routers Off Their Meds: Unstable Routers and the Buggy BGP Implementations That Cause Them,” tech. rep., University of Minnesota, Nov 2011.

[2] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang, “BGP routing stability of popular destinations,” in *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pp. 197–202, ACM, 2002.

[3] A. Elmokashfi, A. Kvalbein, and C. Dovrolis, “On the scalability of BGP: the roles of topology growth and update rate-limiting,” in *Proceedings of the 2008 ACM CoNEXT Conference*, p. 8, ACM, 2008.

[4] A. Elmokashfi, A. Kvalbein, and C. Dovrolis, “On the scalability of BGP: the role of topology growth,” *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 8, pp. 1250–1261, 2010.

[5] A. Szekeres, B. Overeinder, and G. Pierre, “Multi-path inter-domain routing: The impact on BGPs scalability, stability and resilience to link failures,” Master’s thesis, Vrije Universiteit Amsterdam, Aug 2011.

[6] X. Sun, S. G. Rao, and G. G. Xie, “Modeling complexity of enterprise routing design,” in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, pp. 85–96, ACM, 2012.

[7] K. Butler, P. McDaniel, and W. Aiello, “Optimizing BGP security by exploiting path stability,” in *Proceedings of the 13th ACM conference on Computer and communications security*, pp. 298–310, ACM, 2006.

[8] M. Nicholes and B. Mukherjee, “A survey of security techniques for the border gateway protocol (BGP),” *Communications Surveys & Tutorials, IEEE*, vol. 11, no. 1, pp. 52–65, 2009.

[9] M. Schuchard, A. Mohaisen, D. Foo Kune, N. Hopper, Y. Kim, and E. Y. Vasserman, “Losing control of the internet: Using the data plane to attack the control plane,” in *Proceedings of the 17th ACM conference on Computer and communications security*, pp. 726–728, ACM, 2010.

[10] K. Butler, T. R. Farley, P. McDaniel, and J. Rexford, “A survey of BGP security issues and solutions,” *Proceedings of the IEEE*, vol. 98, no. 1, pp. 100–122, 2010.

[11] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, “The flattening internet topology: Natural evolution, unsightly barnacles or contrived collapse?,” in *Passive and Active Network Measurement*, pp. 1–10, Springer, 2008.

[12] M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” in *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 251–262, ACM, 1999.

[13] J. Wu, H.-z. Deng, and Y.-j. Tan, “Spectral measure of robustness for Internet topology,” in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, vol. 6, pp. 50–54, IEEE, 2010.

[14] M. Wojciechowski, B. Overeinder, G. Pierre, M. Van Steen, and J. Mincer-daszkiewicz, “Border gateway protocol modeling and simulation,” Master’s thesis, Vrije Universiteit Amsterdam, Jul 2008.

[15] H. Sam, *Internet Routing Architectures, 2/E.* Pearson Education India, 2008.

[16] C. Labovitz, G. R. Malan, and F. Jahanian, “Internet routing instability,” *Networking, IEEE/ACM Transactions on*, vol. 6, no. 5, pp. 515–528, 1998.

[17] Y. Rekhter and T. Li, “A border gateway protocol 4 (BGP-4),” 1995.

[18] G. H. Tony Bates, Philip Smith, “CIDR REPORT.” `http://www.cidr-report.org/as2.0/`, July 2013.

[19] “The CAIDA AS Relationships Dataset.” `http://www.caida.org/data/active/as-relationships/`.

[20] “Classless inter-domain routing (CIDR): an address assignment and aggregation strategy,”

[21] Y. Wang, M. Schapira, and J. Rexford, "Neighbor-specific BGP: more flexible routing policies while improving global stability," in *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, pp. 217–228, ACM, 2009.

[22] C. Hall, R. Anderson, R. Clayton, E. Ouzounis, and P. Trimintzios, "Resilience of the Internet Interconnection Ecosystem," in *Economics of Information Security and Privacy III*, pp. 119–148, Springer, 2013.

[23] M. Yannuzzi, X. Masip-Bruin, and O. Bonaventure, "Open issues in interdomain routing: a survey," *Network, IEEE*, vol. 19, no. 6, pp. 49–56, 2005.

[24] B. Quoitin, C. Pelsser, L. Swinnen, O. Bonaventure, and S. Uhlig, "Interdomain traffic engineering with BGP," *Communications Magazine, IEEE*, vol. 41, no. 5, pp. 122–128, 2003.

[25] Y. R. Yang, H. Xie, H. Wang, A. Silberschatz, A. Krishnamurthy, Y. Liu, and L. E. Li, "On route selection for interdomain traffic engineering," *Network, IEEE*, vol. 19, no. 6, pp. 20–27, 2005.

[26] T. G. Griffin and B. J. Premore, "An experimental analysis of BGP convergence time," in *Network Protocols, 2001. Ninth International Conference on*, pp. 53–61, IEEE, 2001.

[27] T. G. Griffin, F. B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Transactions on Networking (ToN)*, vol. 10, no. 2, pp. 232–243, 2002.

[28] G. Huston, "BGP Table." `http://bgp.potaroo.net`, Oct. 2013.

[29] L. H3C Technologies Co., "BGP Introduction." `http://www.h3c.com`, Oct. 2013.

[30] B. Augustin, B. Krishnamurthy, and W. Willinger, "IXPs: mapped?," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pp. 336–349, ACM, 2009.

[31] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet optometry: Assessing the broken glasses in internet reachability," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pp. 242–253, ACM, 2009.

[32] H. Chang and W. Willinger, "Difficulties measuring the Internet's AS-level ecosystem," in *Information Sciences and Systems, 2006 40th Annual Conference on*, pp. 1479–1483, IEEE, 2006.

[33] L. Blumenfeld, "Dissertation could be security threat," *Washington Post*, vol. 8, 2003.

[34] R. Grone and R. Merris, "A bound for the complexity of a simple graph," *Discrete mathematics*, vol. 69, no. 1, pp. 97–99, 1988.

[35] D. L. Neel and M. E. Orrison, "The linear complexity of a graph," *the electronic journal of combinatorics*, vol. 13, no. 1, p. R9, 2006.

[36] M. H. Behringer, "Classifying network complexity," in *Proceedings of the 2009 workshop on Re-architecting the internet*, pp. 13–18, ACM, 2009.

[37] S. Vissicchio, *Governing Routing in the Evolving Internet.* PhD thesis, Universitadegli Studi di Roma Roma Tre, Dottorato di Ricerca in Ingegneria, Sezione Informatica ed Automazione, XXIV Ciclo, 2012.

[38] "Distributed ASCI Supercomputer 4." `http://www.cs.vu.nl/das4`.

[39] A. Dhamdhere and C. Dovrolis, "Ten years in the evolution of the Internet ecosystem," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pp. 183–196, ACM, 2008.