

# SIMULATING THE BGP WITH OPNET GURU 10.5

Studienarbeit

Author:

Alexander Probst  
aprobst@uni-koblenz.de

Tutors:

Prof. Christoph Steigner  
Dipl.Inf. Harald Dickel

Universität Koblenz-Landau  
Campus Koblenz  
Fachbereich 4  
Informatik

12.07.2006

**Abstract**

The high cost of routing infrastructure makes checking theories about larger nets a very difficult and expensive task. One possible approach to fight this problem is the use of virtual instead of physical infrastructure. OPNet's IT Guru software is a suite designed to simulate large nets and present relevant information. This allows validating extensive changes before actually implementing them on a productive network or testing theories without the overhead of a physical infrastructure.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	What is OPNet IT Guru 10.5? . . . . .	5
1.1.1	Strengths of OPNet IT Guru . . . . .	5
1.1.2	Weaknesses of IT Guru . . . . .	6
1.2	A short introduction to the Border Gateway Protocol . . . . .	8
1.3	Definiton of Emulation and Simulation . . . . .	10
<b>2</b>	<b>How to use OPNet IT Guru efficiently</b>	<b>11</b>
<b>3</b>	<b>Examples</b>	<b>13</b>
3.1	Simple Net . . . . .	13
3.2	Black Holes in Synchronization . . . . .	17
3.3	Shutting down a physical link during simulation . . . . .	21
3.4	The original task . . . . .	21
3.5	Implementing Routing Policies . . . . .	23
<b>4</b>	<b>Final Words</b>	<b>25</b>
<b>5</b>	<b>Appendix: Content of the CD-ROM</b>	<b>26</b>
	<b>Literature</b>	<b>27</b>

## List of Figures

1	The global statistics dialogue . . . . .	7
2	BGP Routing Table Export . . . . .	8
3	Display space is wasted . . . . .	9
4	Enabling the export of routing tables . . . . .	11
5	Duplicating Scenarios easily . . . . .	12
6	Network Difference Reports . . . . .	12
7	Rapid Configuration - Star Topology . . . . .	13
8	Configuring Background Load . . . . .	14
9	"Simple Net" topology . . . . .	14
10	The Confederation Configuration dialogue . . . . .	15
11	Topology of "SimpleNetConfed" . . . . .	15
12	Routing Table of "withoutRedistribution" . . . . .	16
13	Routing Table of "withRedistribution" . . . . .	16
14	Topology of "SimpleNetRouteReflector", scenario "RR" . . . . .	16
15	Routing Table of Router A with Route Reflection . . . . .	17
16	Routing Table of Router A without Route Reflection . . . . .	17
17	Configuring Route Reflection on the Reflector . . . . .	18
18	Topology of "Blackout" . . . . .	18
19	Manipulating the Start Time of Routing Protocols . . . . .	19
20	Start Times of RIP, BGP in "WithSynchronization2" . . . . .	20
21	Convergence times and activity in "Blackout:WithSynchronization2" . . . . .	20
22	Topology of "Policies" . . . . .	23
23	Routing Table of "Router A" in "Local Pref" . . . . .	24
24	Routing Table of "Router A" in "AS-Path" . . . . .	24
25	Routing Table of "Router A" in "MED" . . . . .	25
26	Routing Table of "Router A" in "EBGP vs. IBGP" . . . . .	25
27	Routing Table of "Router A" in "IGP" . . . . .	25

# 1 Introduction

## 1.1 What is OPNet IT Guru 10.5?

The IT Guru software is a modular suite able to simulate entire networks with up to several dozen nodes. This includes all layers of the OSI reference model[7], from physical links up to application demands. Its primary function, according to OPNet's web site [10] is the support of network planning groups and application developers. IT Guru is able to simulate large amounts of traffic and give detailed reports at the end of a simulation, including routing tables for a multitude of routing protocols at several points of time, network load reports and convergence times. IT Guru projects consist of easily created and compared scenarios. For each scenario, different data and network topologies can be analyzed. OPNet IT Guru offers up to four simulation modes: Discrete Event Simulation (DES), Flow Analysis, ACE Quick Predict, and Hybrid Simulation, which is a derivation of DES. In this paper, due to the license restrictions, only DES is examined; the other simulation types were not available. DES is a packet-based simulation, and therefore best suited for researching protocol behaviour and application performance.

### 1.1.1 Strengths of OPNet IT Guru

There are several reasons for using OPNet IT Guru.

- The greatest strength of IT Guru is its usability. The integrated graphical user interface speeds up the creation of networks and helps to visualize them. Parts of larger nets can be grouped easily to sub-nets, allowing for a visual classification of network nodes. Since this grouping is only an illustration, no information is lost, but a better overview for humans is gained. The graphical display also improves communication, especially with non-technical personnel.
- IT Guru supports a large number of different simulated devices and new ones can be included by the "Device Creator" tool. This allows for a great variety in simulation. The configuration of these devices includes excessive details, which allows for an almost identical reproduction of the real devices. However, radically new methods cannot be simulated.
- Another strength is speed. While emulation needs time – in fact, most emulators do not run in real-time – IT Guru can simulate a network's behaviour over twenty-four hours in a couple of minutes, depending on the simulated network. This facilitates long-term testing within a short period of time before implementing any changes. It also allows a "Trial-and-Error" approach, changing small things with each run to achieve a certain goal. Given skilled IT Guru users, it is possible to create entire networks ad-hoc, integrating non-technical personnel to a high degree into the design process.
- An interesting feature is the Virtual Command Line Interface, which gives the user access to a shell on a simulated router. This shell accepts certain Cisco commands like "synchronization" (which toggles BGP synchronization) or "auto-summary". The instruction set, however, is far from complete. Important commands like "show ip bgp" are not included, and as an additional drawback, it cannot be accessed during simulation.

- OPNet IT Guru features a fairly sophisticated recovery function. Projects and scenarios which due to external or internal influences have been corrupted can often be restored. This function is, however, not perfect, and back-up copies are still recommended.

### 1.1.2 Weaknesses of IT Guru

IT Guru 10.5's approach of simulation not only harbors advantages, but also several limiting drawbacks:

- The simulation is a black box approach. There is no possibility to trigger events during a simulation except the manipulation of protocol start time. For example, the breakdown of the the main physical link between two servers after 20 simulated minutes cannot be simulated, which makes testing back-up solutions or alternate routing tables impossible. This function cannot be reproduced by any means the simulator has to offer; even flooding the simulated physical link with useless background traffic has no visible effect on the exchange of BGP routes, and there is no way of shutting down routing protocols after a certain amount of time. This makes OPNet IT Guru unsuitable for almost all problems based on dynamic changes during simulation.
- New routing protocols cannot be created nor embedded into new devices. This makes IT Guru completely unusable for all kinds of research based on not already implemented protocols. Instead of advancing this certain branch of research, only the already implemented protocols can be examined and improved.
- Another problem is OPNet's license policy. The license for IT Guru, even if it is an older version, has to be renewed every few months, limiting the use in research. This prohibits later re-examination of results, as already created scenarios cannot be examined with an expired license. However, a so-called "Scenario Web Report" can be created, which includes a graphical representation of all simulated nodes as well as their entire configuration tables. Since IT Guru is not open/shared source software, the source code cannot be examined or changed. This is a significant drawback because it prohibits makeshift bug fixing (also look at 3.4) or plug-ins, like a scripting language which would allow for a more in-depth analysis.
- As a part of the license policy, OPNet's forums, help desk and other additional documentation are only accessible to certain kind of license owners. This prevents for example owners of the "university license" from communicating with other users or OPNet personnel via the web site. This problem has been partially resolved; OPNet now maintains a separate forum for universities. While this still does not allow for a complete exchange of information, it is a first step. [1].
- While simulation itself is very fast ("SimpleNet", scenario 1 runs in under one minute), saving projects takes, depending on its size, up to 5 minutes. During this time span no other actions can be taken. This also happens during the automated back-up, which can lead to a loss of concentration.
- The documentation is insufficient and even erroneous. For example, it claims that the BGP routing table belongs to the "menu-selectable reports" which can be chosen from the "Global Statistics" dialogue.

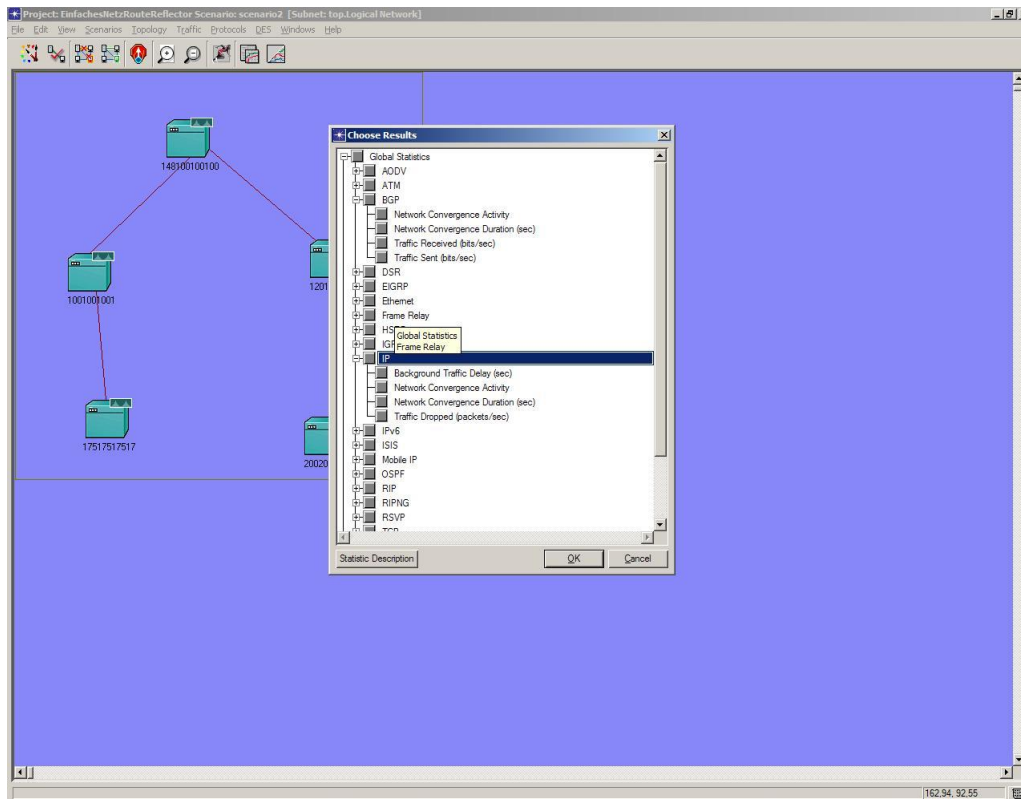


Figure 1: The global statistics dialogue

The screenshot shows this dialogue. As can clearly be seen, there is no entry labeled "BGP routing table". This leads to confusion and prolongs the training period. This specific report can be chosen from editing the simulated object instead:

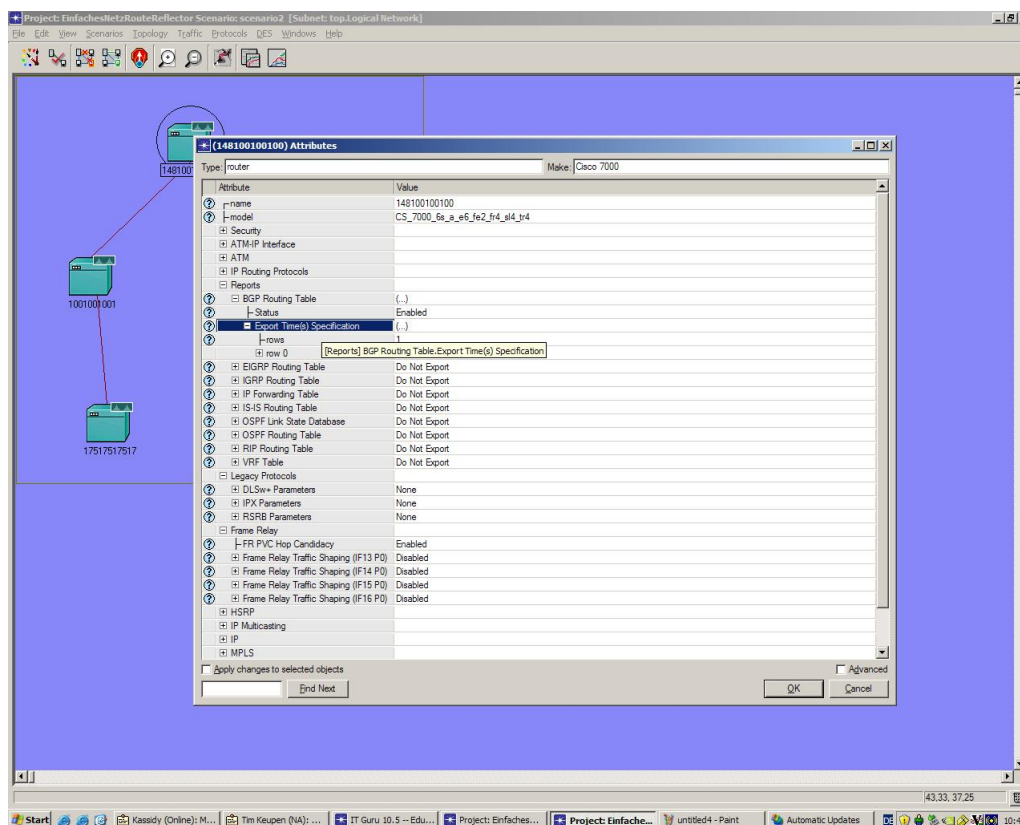


Figure 2: BGP Routing Table Export

- Another problem is the inability of OPNet IT Guru to cope with higher display resolutions. On a 1280x1024 resolution, enabling full screen mode does not enlarge the workspace to more than roughly one quarter of the display. This hinders the management of larger subnets severely and decreases facility of inspection.

This problem can be partially resolved by zooming in; the zoom factor is saved. However, this is not a perfect solution and counterintuitive.

- One other thing is not exactly a drawback, but rather an unfortunate design decision. IT Guru differs in saving projects from the native OS (in this case Microsoft Windows XP Professional). Instead of simply saving a project in a given directory, this directory has to be added to the "model directories". IT Guru also records the name of each project; having a project with the same name in two or more different "model directories" is strictly forbidden, which makes creating back-up copies more difficult. This problem is partly alleviated by the automatic back-ups IT Guru creates after certain intervals of time.

## 1.2 A short introduction to the Border Gateway Protocol

The **Border Gateway Protocol(BGP)** is a protocol for the exchange of network reachability information between **Autonomous Systems(AS)**. An AS is defined as a grouping of "routers under a single technical administration"[18]. In the classical sense, this meant a common IGP and metrics, but in the meantime AS often have several IGPs or are subdivided 1.2. BGP also serves to enforce routing policies. For example, a BGP router could block all traffic coming



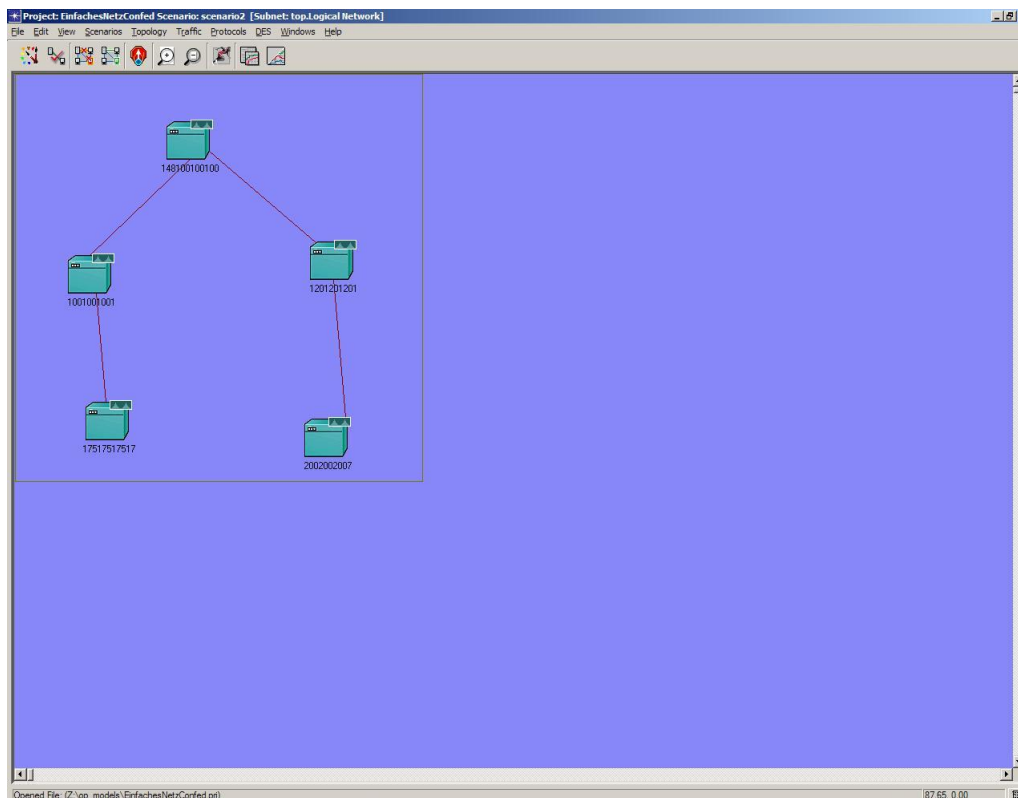


Figure 3: Display space is wasted

from a certain AS, or it could prefer a certain route over another. The BG Protocol is situated on layer 4 of the OSI model, utilizing TCP as underlying transport protocol. Autonomous systems are identified by their **AS Number**.

Contrary to dynamic interior protocols like RIP all neighbors are entered by hand. Route information exchange is complete and the routes are loop-free; each BGP speaker propagates the routes it uses to all of its neighbors. This "hop-by-hop" routing paradigm leads after consolidation to BGP routers which know not every route, but have a valid route for all nets, barring route policies. If BGP is used between different AS, it serves as an Exterior Gateway Protocol (EGP; not to be confused with the actual Exterior Gateway Protocol, which is an implementation of EGP. [8].). If it is used for the exchange of router within an AS, it is referred to as Interior BGP (IBGP).

If BGP has more than one valid route to a target in its routing table, the best route is chosen by the **BGP Route Selection Process**[2]<sup>1</sup>. This selection process, however, can be heavily influenced by the network administration using **routing policies**.

One of the requirements of the IBGP is that all BGP speakers in one AS have to be directly connected as neighbors; they have to be fully meshed. This is very link-intensive; the number of direct connections for  $n$  BGP speakers is  $n \cdot (n-1) / 2$ , leading to high infrastructure cost. There are two ways of solving this problem. One is to use confederations, the other is to use route reflectors.

---

<sup>1</sup>This is a CISCO text; CISCO routers support an additional, optional attribute called weight which is listed in this paper.

A **Confederation** is a super-AS, which consists of several sub-AS. To other AS not within the confederation, the sub-ASs (or **Member-AS**) claim to be only one AS, sending the **AS Confederation Identifier** as their own AS number. To each other, however, they send their **Member-AS Number**. This subdivision removes the need for a full mesh since the member-AS treat each other as different AS; only inside each member-AS the full mesh needs to be upheld. The drawback is the high cost for hardware; each router has to support confederations, which precludes older hardware from being used in such a confederation. [6]

**Route reflectors** work in a different way. A route reflector cluster consists of one route reflector and several clients. The route reflector sends all routes learned by IBGP to its clients. The advantage of this method is that the clients do not need additional software or updates, which lowers hardware cost and facilitates switching from full-mesh to route reflection. However, the route reflector becomes a single point of failure, since the entire cluster is dependent on it. [15]

As of this moment, the most recent version of the BGP protocol is BGP-4, based on IPv4. There are discussions about the future of BGP using IPv6.[18] By the end of the year 2005, BGP routing tables had about 170.000 entries in 26.000 AS [3].

### 1.3 Definiton of Emulation and Simulation

IT Guru has been referred to as a simulator several times until now. Since there are several definitions regarding simulation and emulation, the meaning of **simulator** and, likewise, **simulation** has to be determined. Most of the times, emulation is defined as hardware-based, i.e. a piece of hardware reproducing the functionality of another piece of hardware; for example, a modem behaving like a fax machine. Simulation is in this context defined as a purely software-based solution; for instance, a C-64 simulator running on a PC, or VirtualPC for MAC.[17] This definition, however, is not apt for this paper. VN-UML, which is an alternative to IT Guru, would be considered a simulator, too, since it requires no specialized hardware. Therefore, in the context of this paper, the definitions shall be as followed:

- **Emulation** is a software-based simulation of hardware. This means the virtual hardware is able to communicate with other, not emulated hardware. For example, an emulated network interface card would be able to communicate with a 'real' network interface on the same machine. Also, depending on the kind of emulator, native software written for the emulated hardware can be run on the virtual machine.
- **Simulation** is a software-based simulation of the behaviour of hardware and software inside the simulator. The simulator acts similar to a sandbox [11]; nothing from outside the simulator can interact with the simulated objects.

The main difference between simulation and emulation is the scope. While emulation emulates hardware and allows native software and other machines to interact with an entirely virtual machine, as in Virtual PC, simulation simulates only the behaviour of a system.

While emulation allows more control, such as capturing data with an outside tool, for example using Ethereal in a VNUML/Quagga emulation, or simply stopping the emulation at any point of time, it is also more resource intensive[16]. Simulation is generally faster and allows to simulate even large nets on the same machine. However, due to its black-box approach there is no

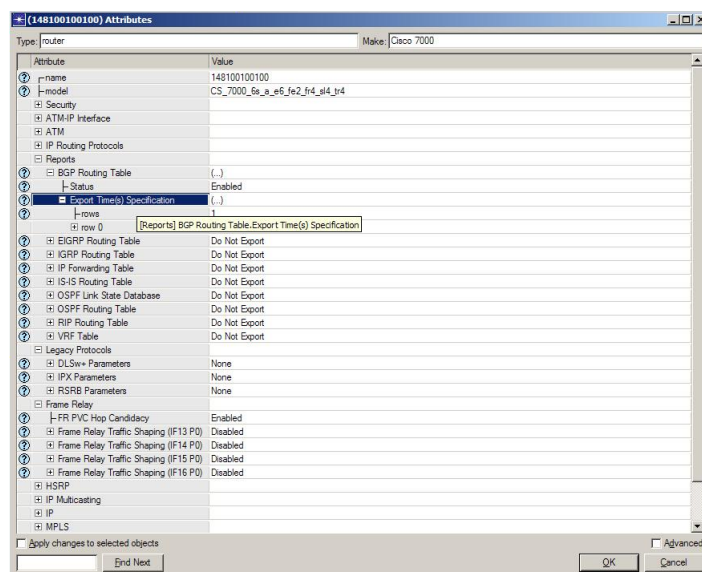


Figure 4: Enabling the export of routing tables

possibility to let outside programs (there are exceptions; compare [12]) or machines communicate with the simulated objects. All means to manipulate the simulation have to be provided by the simulator; this limits visibility and control severely.

As a rule of thumb, emulation tends to be more effective (i.e. more exact, better controllable), while simulation is more efficient.

## 2 How to use OPNet IT Guru efficiently

While IT Guru's documentation is mostly adequate, there are some topics which are either explained falsely or not emphasized enough. This section presents some of these topics.

- Viewing routing and forwarding tables is probably one of the most important features a network simulator can offer. OPNet IT Guru can export these tables for all implemented routing protocols, at any given times, as long as these times are decided before the simulation is run. This option can, contrary to the documentation, be accessed in the 'Reports' sub-menu of the 'Edit Attributes' dialogue of a node running such a protocol.

The supported routing tables can be exported at any point of time during the simulation. This facilitates, for example, the comparison between behaviour before and after a certain protocol has been started 3.2.

- One of OPNet IT Guru's best features is the separation between projects and scenarios. One project consists of one or more scenarios; these scenarios, however, do not have to be even remotely similar. The scenario menu offers several useful commands. For example, an existing scenario can be duplicated [2] which facilitates constructing several similar network topologies with only small alterations greatly. A scenario can be compared by creating a Network Difference Report, even to a scenario of a totally different project.2

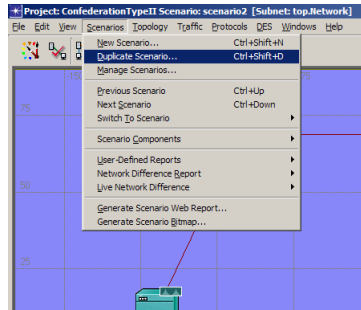


Figure 5: Duplicating Scenarios easily

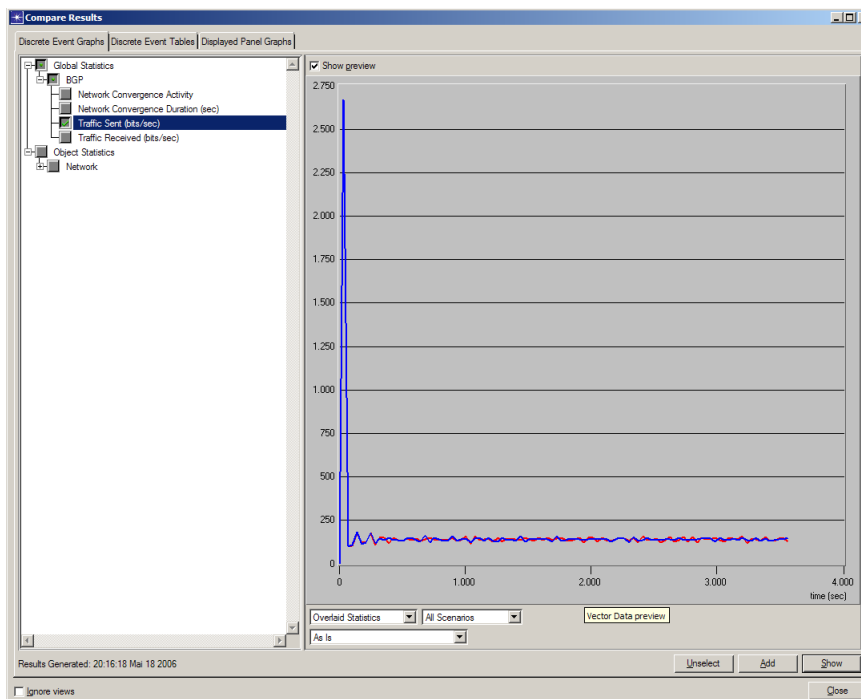


Figure 6: Network Difference Reports

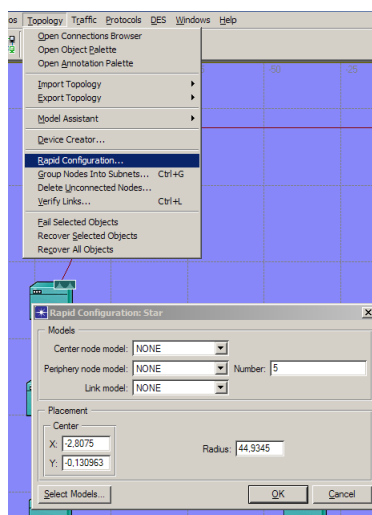


Figure 7: Rapid Configuration - Star Topology

- The DES log is a very helpful tool for discovering errors. It notes at which node, if applicable, an unexpected event occurred, which protocol caused the message, the simulated point of time the message was created and whether the event represents an symptom or a concrete error. The complete event message features, besides a more extensive description of the event, possible causes as well as most of the time very exact and specific suggestions how to solve the problem. Sometimes, the suggestions are either wrong, or caused by intrinsic behaviour of the simulator; for example connecting IBGP routers via the "Protocols->BGP->Configure IBGP peers" command leads to a DES log entry in which the lack of an IP address is noted. Therefore, the user needs to act at one's own discretion.
- A useful tool for the fast creation of simple topologies is the so-called "Rapid Configuration". The supported topologies are star, full mesh, partial random mesh, bus, ring, tree and the unconnected net. This covers all the basic topologies.
- Assigning background traffic to a physical link can be of interest. For example, if a new application or protocol is implemented in an already existing net, it can be easier to abstract from the actual details of the net's traffic and generalize it as a constant background traffic. This can be done easily by setting the "Background Load" attributes of a physical link.

### 3 Examples

The functioning of OPNet IT Guru can be best explained using examples.

#### 3.1 Simple Net

These nets were specifically designed to show OPNet IT Guru's abilities with normal AS, member-ASs and route reflectors.

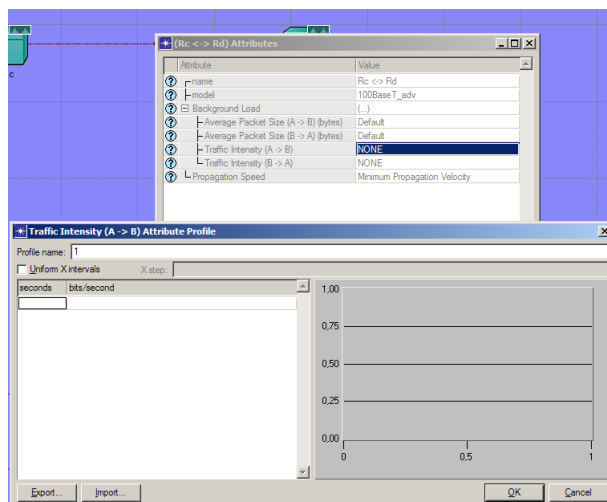


Figure 8: Configuring Background Load

Table 1: Configuration of "SimpleNet"

Name	AS	NRI
1001001001	1	100.100.100.1
1201201201	1	120.120.120.1
1751751757	17	175.175.175.7
2002002007	7	200.200.200.7

### Project SimpleNet

The SimpleNet project is a simple net which serves only to show correct usage of IT Guru and how to simulate simple Autonomous Systems. This first example uses only one scenario. It consists of four routers, each of which propagates a certain net via BGP Network Reachability Information(NRI); the propagated net corresponds to the name of the router (i.e. the NRI about net 120.120.120.1/8 originates from router 1201201201). The underlying IGP is RIP, but since Redistribution is disabled, the routes it found have no effect on the BGP routing table. The exact NRI and AS of each router can be seen in table 3.1. The BGP routing table of all four routers contains at the end of the DES routes for all four advertised nets, thus proving that the simulator to works correctly.

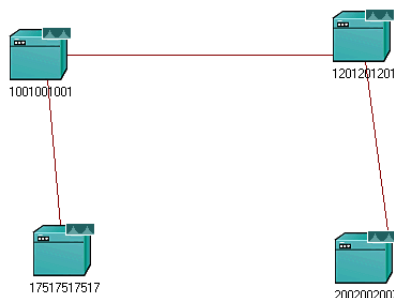


Figure 9: "Simple Net" topology

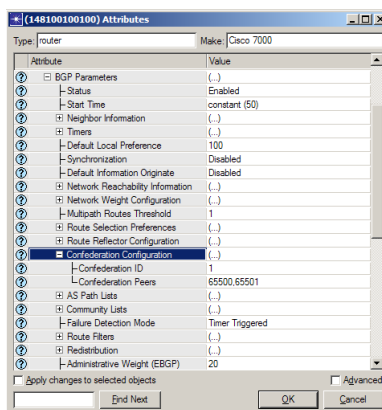


Figure 10: The Confederation Configuration dialogue

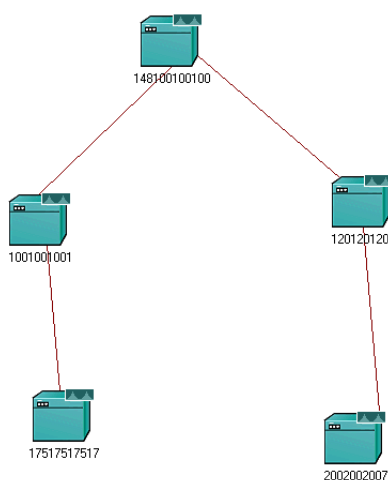


Figure 11: Topology of "SimpleNetConfed"

### Simple Net with Confederations

The next project is called "SimpleNetConfed". It illustrates a single AS Confederation consisting of five routers which are separated into several member-ASs. The topology can be seen in fig. 3.1. Confederations are configured in the 'Edit Preferences' dialogue of each router.

The underlying Interior Gateway Protocol (IGP) is in this example RIP. The project itself consists of two scenarios. In "withoutRedistribution" the BGP table is built only from BGP NRI. Each router lists one network reachable due to its NRI. In "withRedistribution", route redistribution is active on all routers, so RIP entries are inserted into the BGP routing table. This leads to a considerably larger routing table, since not only routers and nets, but also individual network adapters are included. Again, this proves that the simulator is working correctly and as expected.

### Simple Net with Route Reflection

The third of the "SimpleNet" projects, "SimpleNetRouteReflector" illustrates the usage of Route Reflectors. The simulated net consists of five routers, whose NRI and AS can be seen in table 3.1. Routers A, B and C constitute the AS 1. The IGP is RIP.

The "NoRR" scenario exemplifies a net without Route Reflection, in which routers A, B and C

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin	Community	Route Dis
200.200.200.7/24	CONFED EBGP	141.86.76.1	Logical Network.2...	IF1	0	100	0	7	EGP		
175.175.175.17/16	CONFED EBGP	178.178.178.7	Logical Network.1...	IF2	0	100	0	(65500) 17	EGP		
120.120.120.1/8	CONFED EBGP	97.97.97.1	Logical Network.1...	IF1	0	100	0		IGP		
100.100.100.1/8	CONFED EBGP	23.23.23.23	Logical Network.1...	IF2	0	100	0	(65500)	IGP		
148.100.100.100/16	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP		

Figure 12: Routing Table of "withoutRedistribution"

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin	Community	Route Dis
175.175.175.17/16	CONFED EBGP	178.178.178.7	Logical Network.1...	IF2	0	100	0	(65500) 17	EGP		
200.200.200.7/24	CONFED EBGP	141.86.76.1	Logical Network.2...	IF1	0	100	0	7	EGP		
120.120.120.1/8	CONFED EBGP	97.97.97.1	Logical Network.2002002007		0	100	0		IGP		
100.100.100.1/8	CONFED EBGP	23.23.23.23	Logical Network.1...	IF2	0	100	0	(65500)	IGP		
148.100.100.100/16	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP		
192.0.0.0/24	RIP	97.97.97.1	Logical Network.1...	IF1	10	100	100		Incomplete		
192.0.2.0/24	RIP	23.23.23.23	Logical Network.1...	IF2	10	100	100		Incomplete		
192.0.0.0/24	RIP	97.97.97.1	Logical Network.1...	IF1	10	100	100		Incomplete		
141.86.0.0/16	RIP	97.97.97.1	Logical Network.1...	IF1	10	100	100		Incomplete		
192.0.1.0/24	RIP	23.23.23.23	Logical Network.1...	IF2	10	100	100		Incomplete		
178.178.0.0/16	RIP	23.23.23.23	Logical Network.1...	IF2	10	100	100		Incomplete		
170.160.0.0/16	RIP	23.23.23.23	Logical Network.1...	IF2	10	100	100		Incomplete		
192.0.4.0/24	Direct	192.0.4.1	Logical Network.1... LB0		10	100	100		Incomplete		
23.0.0.0/8	Direct	23.23.23.22	Logical Network.1...	IF2	10	100	100		Incomplete		
97.0.0.0/8	Direct	97.97.97.2	Logical Network.1...	IF1	10	100	100		Incomplete		

Figure 13: Routing Table of "withRedistribution"

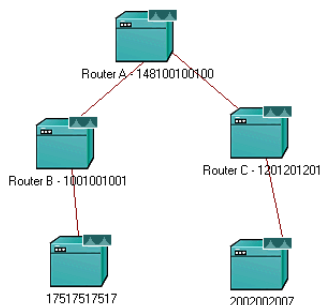


Figure 14: Topology of "SimpleNetRouteReflector", scenario "RR"



Table 2: Configuration of "SimpleNetRouteReflector"

Name	AS	NRI
Router A - 148100100100	1	148.100.100.100
Router B - 1001001001	1	100.100.100.1
Router C - 1201201201	1	120.120.120.1
2002002007	7	200.200.200.7
175.175.175.7	17	175.175.175.7

The screenshot shows the 'View Results' window for Router A. The routing table is displayed with the following data:

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin	Community	Route Distinguisher
200.200.200.7/24	RR IBGP	141.86.76.1	Logical Network.2...	IF1	0	100	0	7	EGP		
175.175.175.17/16	RR IBGP	178.178.178.7	Logical Network.1...	IF2	0	100	0	17	EGP		
120.120.120.1/8	RR IBGP	97.97.97.1	Logical Network.R...	IF1	0	100	0		IGP		
100.100.100.1/8	RR IBGP	23.23.23.23	Logical Network.R...	IF2	0	100	0		IGP		
148.100.100.100/16	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP		

Figure 15: Routing Table of Router A with Route Reflection

are directly linked. Scenario RR shows the same net, without the direct connection between routers B and C. This is replaced by a Route Reflection cluster, with router A as the reflector, and Routers B and C as clients. The routing tables show the correctness of the simulation. In the first scenario they depict Router A having routes to all propagated nets. The second scenario shows the same end result except that in this case all routes are learned via Route Reflection and are accordingly marked.

### 3.2 Black Holes in Synchronization

The BGP synchronization rule affects so-called "transit-AS", ASs which transport traffic from one AS to another. Synchronization demands that all IBGP routers of the transit-AS have to learn any IBGP route (i.e. routes which are learned via IBGP, and not from other ASs) via IGP, too, before they propagate them to external peers. [14] Synchronization is not needed if either all routers in the transit-AS run BGP, or if the AS is not a transit-AS, but a stub (an AS which is directly connected to only one other AS).

One problem that can occur with synchronization is the so-called 'black hole' phenomenon. If BGP is up and running and external routing requests come in before the internal routing table is built, synchronization can lead to the dropping of transit traffic if the IBGP routers on the route have not yet learned this route via an IGP. [13]

The screenshot shows the 'View Results' window for Router A. The routing table is displayed with the following data:

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin	Community	Route Distinguisher
175.175.175.17/16	IBGP	178.178.178.7	Logical Network.1...	IF2	0	100	0	17	EGP		
200.200.200.7/24	IBGP	141.86.76.1	Logical Network.2...	IF1	0	100	0	7	EGP		
120.120.120.1/8	IBGP	97.97.97.1	Logical Network.R...	IF1	0	100	0		IGP		
100.100.100.1/8	IBGP	23.23.23.23	Logical Network.R...	IF2	0	100	0		IGP		
148.100.100.100/16	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP		

Figure 16: Routing Table of Router A without Route Reflection

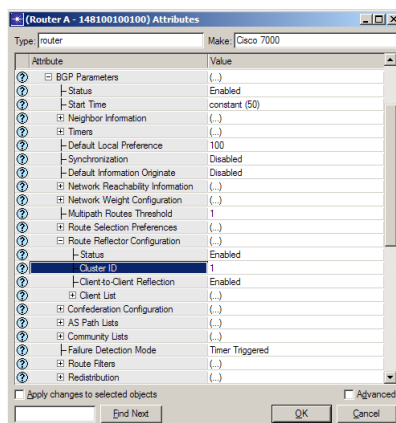


Figure 17: Configuring Route Reflection on the Reflector

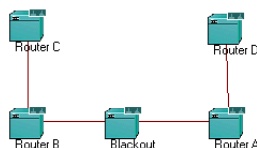


Figure 18: Topology of "Blackout"

The project "Blackout" [3.2] simulates this in four scenarios:

**Scenario WithSynchronization**

The scenario "WithSynchronization" demonstrates a case in which synchronization is advisable. The exact configuration can be seen in 3.2. AS 100 serves in this case as Transit AS for ASs 300 and 400. The underlying IGP is RIP. After a DES has been run, Router C has complete routing information; this is the expected behaviour. The BGP routing table converged at 1 minute 10 seconds into the simulation. This was measured in Routers A and C.

**Scenario WithoutSynchronization**

Scenario "WithoutSynchronization" has the same topology. The only difference is that synchronization is disabled. Interestingly, BGP route table convergence time measured in Routers A and C is 1 minute 10 seconds, identical to that of scenario "WithSynchronization".

Table 3: Configuration of "Blackout"

Name	AS	NRI
Router A	100	100.100.100.1
Router B	100	120.120.120.1
Blackout	100	175.175.175.7
Router C	300	200.200.200.7
Router D	400	000.000.000.0

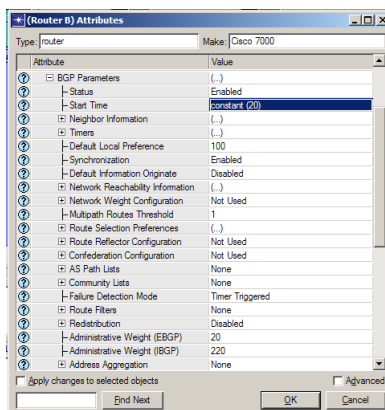


Figure 19: Manipulating the Start Time of Routing Protocols

Table 4: Project "Synchronization" Results

Scenario	Synchron. status	Synchron rec?	convergence time
WithSynchronization	enabled	yes	0
WithSynchronization2	enabled	no	0
WithoutSynchronization	disabled	yes	0
WithoutSynchronization2	disabled	no	0

### Scenario Without Synchronization2

Scenario "WithoutSynchronization2" features again the same topology. However, not only is synchronization disabled, but the start time for RIP is set at 70 seconds. The BGP starting time has been set at 20 seconds. 3.2At this point of time, in the former examples the BGP routing table has already converged. This time, the BGP routing table converges at 20 seconds for the first time, but only BGP is active and router Blackout runs no BGP. The BGP routing table reflects this: at 50 seconds, Router A's routing table contains only the net the Router itself propagates and the net the external Router C propagates. All traffic from Router C to Router D or similar is dropped. At 183 seconds, the routing table has converged a second time after RIP has started and it contains the same routing information as in "WithoutSynchronization". The DES log notes the dropping of several packets, because the IP routing table has no routing information for the target. This is the expected behaviour.

### Scenario WithSynchronization2

Scenario "WithSynchronization2" features again the same topology. The BGP start time is 20 seconds, while RIP start time is set at 70 seconds (compare 3.2). The main difference is the fact that synchronization is active. BGP convergence time is first at 20 seconds, almost as soon as BGP is started. However, RIP is still deactivated. After RIP is started, the BGP routing table converges a second time – like in "WithoutSynchronization2" – at 185 seconds, two seconds later than in "WithoutSynchronization2". This supports the theory that disabling Synchronization leads to a faster routing table convergence. [13]

The results of this experiment are resumed in table 3.2

⊖	BGP Parameters	(...)
⊖	Status	Enabled
⊖	Start Time	constant (20)
⊖	Neighbor Information	(...)
⊖	Timers	(...)
⊖	Default Local Preference	100
⊖	Synchronization	Enabled
⊖	Default Information Originate	Disabled
⊖	Network Reachability Information	(...)
⊖	Network Weight Configuration	Not Used
⊖	Multipath Routes Threshold	1
⊖	Route Selection Preferences	(...)
⊖	Route Reflector Configuration	Not Used
⊖	Confederation Configuration	Not Used
⊖	AS Path Lists	None
⊖	Community Lists	None
⊖	Failure Detection Mode	Timer Triggered
⊖	Route Filters	None
⊖	Redistribution	Disabled
⊖	Administrative Weight (EBGP)	20
⊖	Administrative Weight (IBGP)	220
⊖	Address Aggregation	None
⊖	Auto-Summarization	Enabled
⊖	EIGRP Parameters	(...)
⊖	IGRP Parameters	(...)
⊖	IS-IS Parameters	(...)
⊖	OSPF Parameters	(...)
⊖	Start Time	constant (70)
⊖	Stop Time (seconds)	65.0

Figure 20: Start Times of RIP, BGP in "WithSynchronization2"

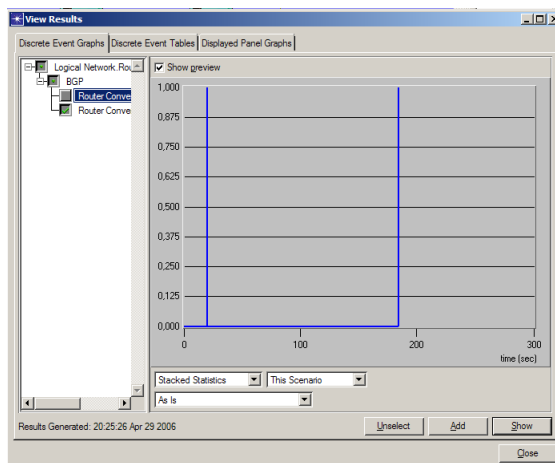


Figure 21: Convergence times and activity in "Blackout:WithSynchronization2"

Table 5: Configuration of "ShutDown"

Name	AS	NRI
Router A	1	12.0.0.0
Router B	1	None

### 3.3 Shutting down a physical link during simulation

Another important feature would be the breakdown or explicit wanted shut-down of a physical link anytime during the simulation. In 1.1.2 it has been stated that even overcrowding the physical link with virtual background traffic does not hinder the propagation of BGP routes. The project "ShutDown" has been created to demonstrate this phenomenon. "ShutDown" consists of two scenarios; Both scenarios consist of two routers linked with a 10BaseT-physical link. Both routers are BGP neighbors. Router A propagates the net 12.0.0.0/8 via NRI and is in AS 1. Router B does not propagate any information and is in AS 1 as well. The scenario "NoBackgroundTraffic" has no background traffic and serves as a control sample, the second one, "BackgroundTraffic", has 120 Gigabit/second of virtual background traffic allotted to the physical link, exceeding by far its capabilities. After running the DES on "NoBackgroundTraffic", however, Router B has a valid route for the net 12.0.0.0/8 in its routing table. This is not the expected behaviour, and disables this approach for overcoming this flaw.

### 3.4 The original task

The original task of this paper was to simulate a constant BGP route oscillation in an inter-network. Since BGP updates are propagated throughout the entire internet, such a constant oscillation would cause significant overhead [4]. If this oscillation could be simulated with the IT Guru software, possible solutions could be tested easily and without the need to acquire and configure real hardware. This would greatly reduce the costs of researching the problem. The first example given in the paper consisted of eight routers, grouped in 4 different Autonomous Systems.

The oscillation occurred because one of the systems was divided in either two route reflection-clusters or two member-ASs of a single confederation. The routers linking the sub-AS chose because of this lack of full routing knowledge different routes and overrode each other frequently because of a lack of full visibility. Some routers have no total knowledge of all available access points to the neighboring AS. This leads to conflicts regarding the Multi-Exit Discriminator attribute.

However, the consistent route oscillation could not be simulated. While most of the behaviour was simulated correctly, no oscillation occurred; the routing tables converged after a short time. Neither the route reflection project nor the confederation scenarios behaved as described in the RFC3345.

The project "RouteReflectionTypeI" is a simulation of the "Route Reflection and Type I Churn" scenario, as explained in RFC3345 [4]. The basis of "Type I Churn" is

Table 6: Configuration of "Type I Churn"

Name	AS	NRI	RR Cluster
Ra(RR)	1	11.0.0.0	1
Rb(C)	1	12.0.0.0	1
Rc(C)	1	13.0.0.0	1
Rd(RR)	1	14.0.0.0	2
Re(C)	1	15.0.0.0	2
AS 6	6	16.0.0.0	None
AS 10	10	17.0.0.0	None
AS 100	100	100.100.100.0	None

Table 7: Configuration of "Type II Churn"

Name	AS	NRI	Sub-AS
Ra	1	11.0.0.0	65501
Rb	1	12.0.0.0	65501
Rc	1	13.0.0.0	65500
Rd	1	14.0.0.0	65500
Re	1	15.0.0.0	65502
Rf	1	16.0.0.0	65502
Rg	1	17.0.0.0	65502
AS200	200	21.0.0.0	None
AS300	300	31.0.0.0	None
AS400	400	10.0.0.0	None

- a single-level Route Reflection or AS Confederations design is used in the network AND
- the network accepts the non-transitive BGP MED attribute from two or more ASs for a single prefix and the MED values are unique.

Each router propagates a certain virtual net via NRI, which can be seen in table 3.4. The underlying IGP is RIP.

By checking the number of BGP updates, BGP traffic and convergence time it can be proven that no constant oscillation occurs. Also, the routing tables show no differences even if they are exported at different points of time. The route exchange is complete, and each router chooses a best path.

The project "ConfederationI" is based on the same topology; however, instead of route reflection, the AS is a confederation divided in two sub-AS - 65000 (which replaces route reflection cluster 1) and 65001 (which replaces route reflection cluster 2). Otherwise, the projects are completely identical.

The project "ConfederationII" follows the scenario "AS Confederations and Type II Churn." This topology is different from the "Type I Churn" used before. In this project, the confederation is divided in three sub-AS with a total number of eight routers, and 3 additional AS (simulated by a single router each.). The route oscillation should occur between routers Rc, Rd and Re.

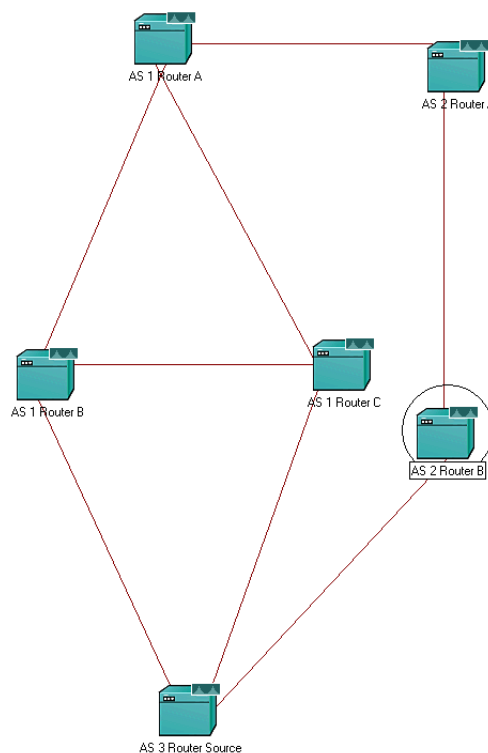


Figure 22: Topology of "Policies"

Table 8: Configuration of "Policies"

Name	AS	NRI
AS 1 Router A	1	11.0.0.0
AS 1 Router B	1	12.0.0.0
AS 1 Router C	1	13.0.0.0
AS 2 Router A	2	21.0.0.0
AS 2 Router B	2	22.0.0.0
AS 3 Router Source	3	31.0.0.0

### 3.5 Implementing Routing Policies

Very important for the network administration of an AS is the enforcement of routing policies. Depending on contracts with other ASs a wrong decision regarding whether traffic is sent over AS A or AS B can be very cost intensive or even actively harmful - routing traffic over a transit AS which drops all traffic coming from your AS would be disastrous. Therefore, a commercial network simulator has to be capable of simulating this. The project "Policies" will demonstrate IT Guru's abilities in this area.

The project is divided in N scenarios. All share the same basic topology 3.5. All routers run BGP. The underlying routing protocol is RIP. The NRI and AS of each router can be seen in table 3.5.

Not all decision criteria are examined here. Criteria such as 'oldest path' or 'Lowest router id' are in this project ignored. Also, "Lowest Origin" is already demonstrated in the scenario

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Orign	Community	Route Distinguisher
31.0.0.1/8	EBGP	113.0.0.2	Network:AS 2 Rou...	IF3	0	500	0	2 3	EGP		
21.0.0.0/8	EBGP	113.0.0.2	Network:AS 2 Rou...	IF3	0	500	0	2	EGP		
13.0.0.0/8	IBGP	112.0.0.2	Network:AS 1 Rou...	IF2	0	100	0		IGP		
12.0.0.0/8	IBGP	111.0.0.2	Network:AS 1 Rou...	IF1	0	100	0		IGP		
22.0.0.0/8	EBGP	113.0.0.2	Network:AS 2 Rou...	IF3	0	500	0	2	EGP		
11.0.0.0/8	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP		

Figure 23: Routing Table of "Router A" in "Local Pref"

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Orign
21.0.0.0/8	EBGP	113.0.0.2	Network:AS 2 Rou...	IF3	0	100	0	2	EGP
31.0.0.1/8	IBGP	131.0.0.2	Network:AS 3 Rou...	IF2	200	100	0	3	EGP
13.0.0.0/8	IBGP	112.0.0.2	Network:AS 1 Rou...	IF2	0	100	0		IGP
12.0.0.0/8	EBGP	111.0.0.2	Network:AS 1 Rou...	IF1	0	100	0	4	EGP
22.0.0.0/8	EBGP	113.0.0.2	Network:AS 2 Rou...	IF3	0	100	0	2	EGP
11.0.0.0/8	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP

Figure 24: Routing Table of "Router A" in "AS-Path"

"Original", and therefore without a scenario of its own.

### Local Preference

The first attributes to be checked is the local preference (**LocalPref**). The path with the highest LocalPref is chosen. The default LocalPref is 100.

The best way to change the LocalPref – or any of the other BGP attributes – of a route is to create a route map and apply this map to a BGP neighbor in the "Edit Attributes->IP Routing Protocols->BGP Parameters->Neighbor Information->row [x]->Routing Policies" dialogue of the affected router. The "direction" of the route map declares whether the changes are applied to incoming or outgoing traffic.

In this scenario, "AS 1 Router A" has a single routing policy established, called "Route Map 1". This policy sets the local preference of all routes coming from "AS 2 Router A" to 500, which means these routes are more preferable. This is evident in the routing table of "AS 1 Router A": All propagated nets except those from its own AS are reached via "AS 2 Router A". This is the expected behaviour. The routing table of "AS 1 Router A" 3.5 demonstrates this.

### AS-Path

Another criteria is the length of the AS-Path, which denotes all AS on the route to this destination. The shortest AS-Path is seen as the best one. The AS Number of "AS 1 Router B" will be changed to "4" to better demonstrate this selection. Additionally, "AS 1 Router C" has a route map applied to incoming routes from "AS 3 Router Source" which changes their MED to 200. This serves only to prove which route is selected.

The routing table of "AS 1 Router A"3.5 shows the selected route for "31.0.0.0". Of course – as according to [2] – the route with the shorter AS Path is selected – (3) over (3 4), as can be seen at the MED applied to the route.

### Multi-Exit Discriminator

The Multi-Exit Discriminator becomes important if there is more than one route from one AS to another. In this topology, AS 1 is connected to AS 3 via "AS 1 Router B" as well as "AS 1 Router C". If all other attributes are equal, the MED is the tie-breaker. If this is be the case, the route with the lower MED is preferred.



Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin
21.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	2	EGP
31.0.0.1/8	IBGP	131.0.0.2	Network AS 3 Rou...	IF2	5	100	0	3	EGP
13.0.0.0/8	IBGP	112.0.0.2	Network AS 1 Rou...	IF2	0	100	0		IGP
12.0.0.0/8	IBGP	111.0.0.2	Network AS 1 Rou...	IF1	0	100	0		IGP
22.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	2	EGP
11.0.0.0/8	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP

Figure 25: Routing Table of "Router A" in "MED"

Destination	Source Protocol	Next Hop Address	Next Hop Node	Outgoing Interface	MED	Local Preference	Weight	AS Path	Origin
21.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	5 2	EGP
31.0.0.1/8	IBGP	122.0.0.2	Network AS 3 Rou...	IF1	0	100	0	3	EGP
13.0.0.0/8	IBGP	112.0.0.2	Network AS 1 Rou...	IF2	0	100	0		IGP
12.0.0.0/8	IBGP	111.0.0.2	Network AS 1 Rou...	IF1	0	100	0		IGP
22.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	5	EGP
11.0.0.0/8	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP

Figure 26: Routing Table of "Router A" in "EBGP vs. IBGP"

The scenario "MED" uses two route maps. "AS 1 Router B" applies a MED of 200 to all incoming routes from "AS 3 Router Source". "AS 1 Router C" applies a MED of 5 to all routes from "Source". The result is as expected: The route with the lower MED is preferred, as can be seen in the routing table of "AS 1 Router A"3.5.

### EBGP versus IBGP

In case of another tie, routes learned via EBGP should be preferred to routes learned via IBGP. For this scenario, the topology has to be slightly changed. "AS 2 Router A"’s AS Number is changed to 5. This is necessary so that the AS-Path length for "AS 2 Router B" is the same, whether (3 2) or (5 2). The routing table of "AS 1 Router A"3.5 shows the preferal, regardless of lower-ranking attributes such as IGP metric for next hop, the route via (5 2); this is the expected behaviour, since (5 2) is learned via EBGP, unlike (3 2), which is learned only from IBGP.

### IGP metric

Another criteria is the IGP metric to the next hop. If all other criteria are equal, the route with the lowest IGP metric is chosen. To demonstrate this, Routers "AS 1 Router B" and "AS 1 Router C" use route maps, changing the IGP metric of the next hop for outgoing BGP routes to "AS 1 Router A" to 200 for "B" and to 5 for "C". Not surprisingly, the routing table’s 3.5 next hop entry for "31.0.0.0" is "AS 1 Router C". This is the expected behaviour.

## 4 Final Words

The IT Guru software allows for rapid creation of large nets. These nets can be configured using the graphical user interface, which in theory should simplify usage. However, due to the inexact documentation and complex and sometimes counterintuitive menus, the familiarization time is still too long. Once the user has accustomed to the program, the many possible options allow

21.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	2	EGP
31.0.0.1/8	IBGP	131.0.0.2	Network AS 3 Rou...	IF2	5	100	0	3	EGP
13.0.0.0/8	IBGP	112.0.0.2	Network AS 1 Rou...	IF2	5	100	0		IGP
12.0.0.0/8	IBGP	111.0.0.2	Network AS 1 Rou...	IF1	200	100	0		IGP
22.0.0.0/8	EBGP	113.0.0.2	Network AS 2 Rou...	IF3	0	100	0	2	EGP
11.0.0.0/8	BGP	Unresolved	Unresolved	Unresolved	0	100	32768		IGP

Figure 27: Routing Table of "Router A" in "IGP"

for a detailed and fast simulation, as long as the simulator itself works correctly.

The lack of documentation, newsgroups or similar convenient ways of exchanging information with other users, at least for the free licenses, hinders the process of learning to use OPNet IT Guru. This is a severe disadvantage in comparison to open source simulators and emulators, like VN-UML.

### **Outlook**

There are several possible avenues future research can take. A new version of IT Guru has been launched recently, which may or may not be able to simulate the routing oscillation. Also, this new version may also have an improved user interface, or a scripting language. In these cases, a thorough investigation could be of interest.

Another possible avenue could be the examination of BGP simulation/emulation, especially of the problem of route oscillation, with other software. One such paper is "BGP-Oszillation mit dem VNUML-Simulator" [5], which employs Quagga/VN-UML to emulate this specific problem. Using this emulator, the author could recreate the route oscillation within the Route Reflector scenario. The emulation of oscillation in the "Confederation" scenarios, however, did not work as predicted. Another well-known open source simulator is NS-2 [[9]], for which this problem has not yet been researched.

Also, BGP Simulation is but a small part of IT Guru's abilities. It could be interesting to research the correctness of other protocols like IGPs (RIP or OSPF) or analyze traffic flows caused by different protocols. Newly designed devices, for example new network nodes like routers could be simulated in OPNet IT Guru. Using the Device Creator, the peculiarities of these nodes could be reproduced. The limitations of the Device Creator, and how to circumvent them, may be also worth a look.

Another strength of the simulator is the support of new, user-created application profiles. Existing nets could be tested on their capability of meeting the requirements of new programs. Given certain application profiles, it could be researched how an existing net could be extended so it would satisfy the new requirements.

## **5 Appendix: Content of the CD-ROM**

The accompanying CD-ROM contains in the folder "projects" all project files for the related examples in this paper. Each project is saved in an individual sub-folder named for it. In the folder "paper" an electronic version of this paper can be found as an PDF. This PDF features Hyperlinks within the document itself.

## References

- [1] OPNet IT Guru University Teaching License. Technical report, OPNet Technologies, Inc., 1984. Last checked on May 19th, 2006.
- [2] BGP Best Path Decision algorithm. <http://www.cisco.com/warp/public/459/25.shtml#background>. Last checked on May 19th, 2006.
- [3] BGP Statistics. <http://bgp.potaroo.net/>. Last checked on May 4th, 2006.
- [4] J. Scudder D. McPherson, P. Traina. RFC3065 - Autonomous System Confederations for BGP. Technical report, AOL Time Warner, TCB, Cisco Systems, 2002. Last checked on May 19th, 2006.
- [5] P. Drekiš. BGP-Oszillation mit dem VNUML-Simulator. Master's thesis, Universität Koblenz-Landau, 2006.
- [6] D. McPherson et al. RFC3345 - Border Gateway Protocol (BGP) Persistent Route Oscillation Condition. Technical report, Juniper Networks, Amber Networks, Cisco Systems, 2001. Last checked on May 18th, 2006.
- [7] International Organization for Standardization. The OSI reference model. [http://en.wikipedia.org/wiki/OSI\\_model](http://en.wikipedia.org/wiki/OSI_model). Last checked on May 19th, 2006.
- [8] D.L. Mills. Exterior Gateway Protocol: Formal Specification. Technical report, Network Working Group, 1984. Last checked on May 19th, 2006.
- [9] NS-2 - The Network Simulator. <http://www.isi.edu/nsnam/ns/>. Last checked on April 29th, 2006.
- [10] OPNet Technologies Website. <http://www.opnet.com/>. Last checked on May 1st, 2006.
- [11] Definition: Sandbox. <http://de.wikipedia.org/wiki/Sandbox>. Last checked on May 2nd, 2006.
- [12] Network Associates Network Analyzer/IT Guru compability. <http://www.opnet.com/products/modules/mvi.html>. Last checked on May 23rd, 2006.
- [13] BGP Case Studies: Synchronization. [http://www.cisco.com/en/US/tech/tk365/technologies\\_tech\\_note09186a00800c95bb.shtml#synch](http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a00800c95bb.shtml#synch). Last checked on May 30th, 2006.
- [14] BGP Synchronization Rules. [http://www.911networks.com/bgp\\_synchronisation\\_rules](http://www.911networks.com/bgp_synchronisation_rules). Last checked on April 29th, 2006.
- [15] R. Chandra T. Bates. RFC1966 - BGP Route Reflection. Technical report, Cisco Systems, 1996. Last checked on May 15th, 2006.
- [16] VN User Mode Linux. <http://jungla.dit.upm.es/~vnuml/>. Last checked on May 19th, 2006.
- [17] Microsoft Virtual PC. <http://www.microsoft.com/windows/virtualpc/default.mspx>. Last checked on April 29th, 2006.

- [18] S.Hares Y. Rekhter, T. Li. RFC4271 - A Border Gateway Protocol 4. Technical report, IBM, Cisco Systems, 2006. Last checked on May 19th, 2006.