

Modeling the Routing of an ISP

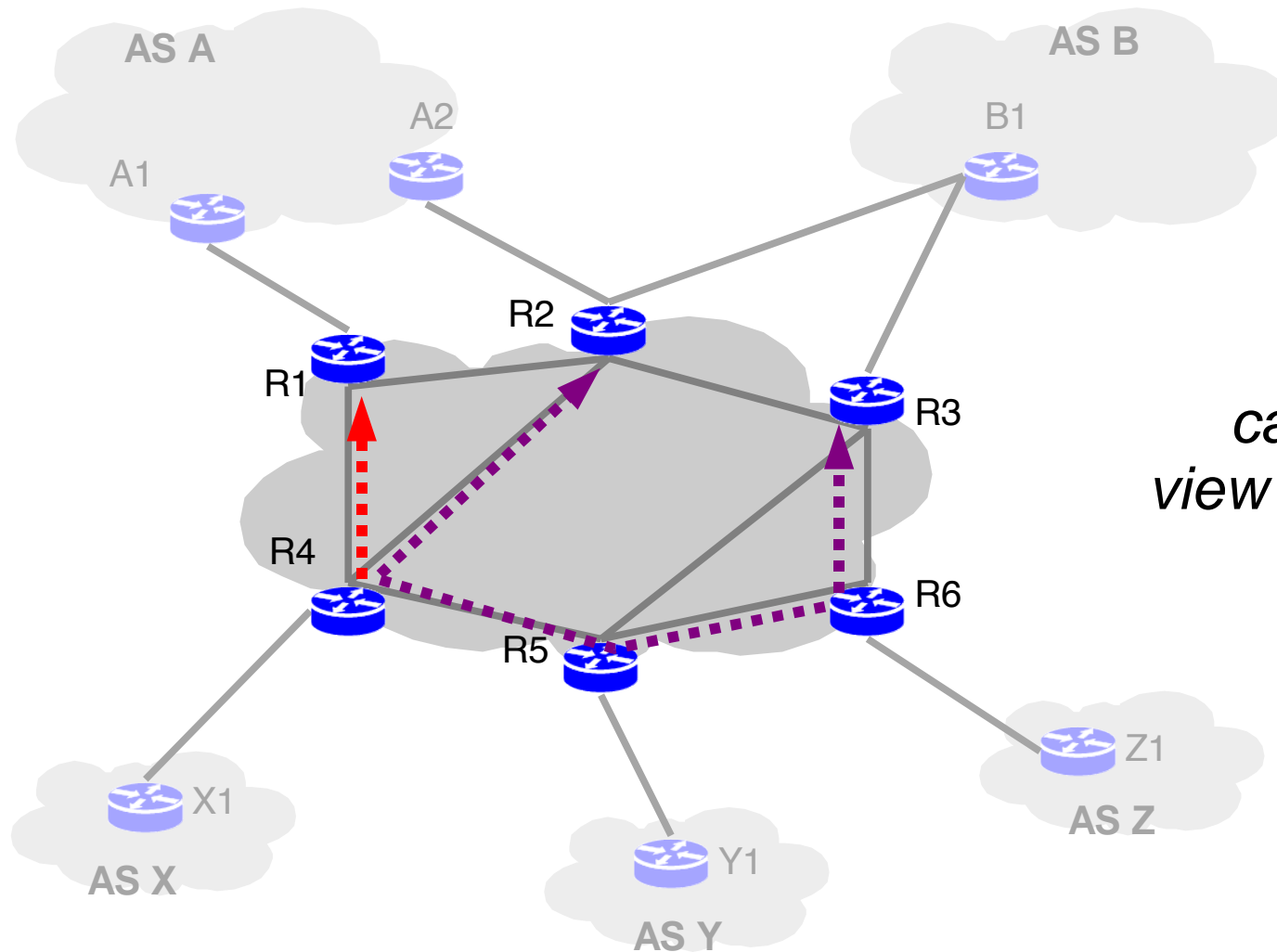
Bruno Quoitin

bruno.quoitin@uclouvain.be

IP Networking Lab (INL)
Computer Science & Engineering Department
Université catholique de Louvain, Belgium



ISP Model



*The traditional
capacity planning
view of an ISP network*

Internal link
Traffic flow



Can't answer questions such as...

What would happen to my interdomain traffic if...

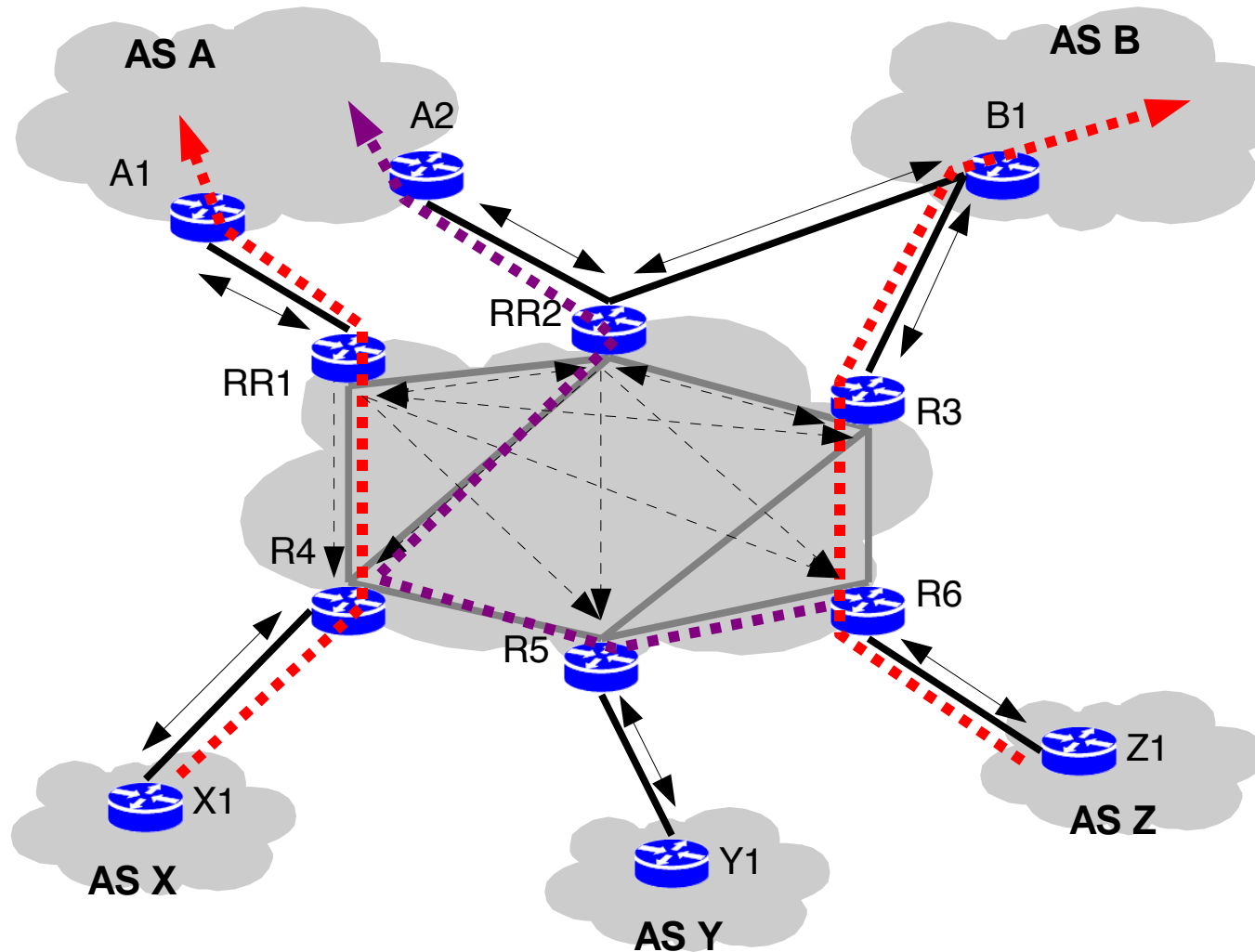
- a link is failing ?
- a router is under maintenance ?
- a BGP peering is being shutdown ?
- a new route filtering policy is planned ?
- a new peering is established at an IXP ?

How would I optimize my interdomain routing for...

- performance ?
- cost ?
- reliability ?



ISP Model



Reality has:

- Transit traffic
- Multiple egresses
- iBGP topology
- Route-reflectors
- Routing policies
- 250,000 destinations (and counting)
- interaction w/ IGP
- ...

<i>Internal link</i>	
<i>External link</i>	
<i>eBGP session</i>	
<i>iBGP session</i>	
<i>Client session</i>	
<i>Traffic flow</i>	



Agenda

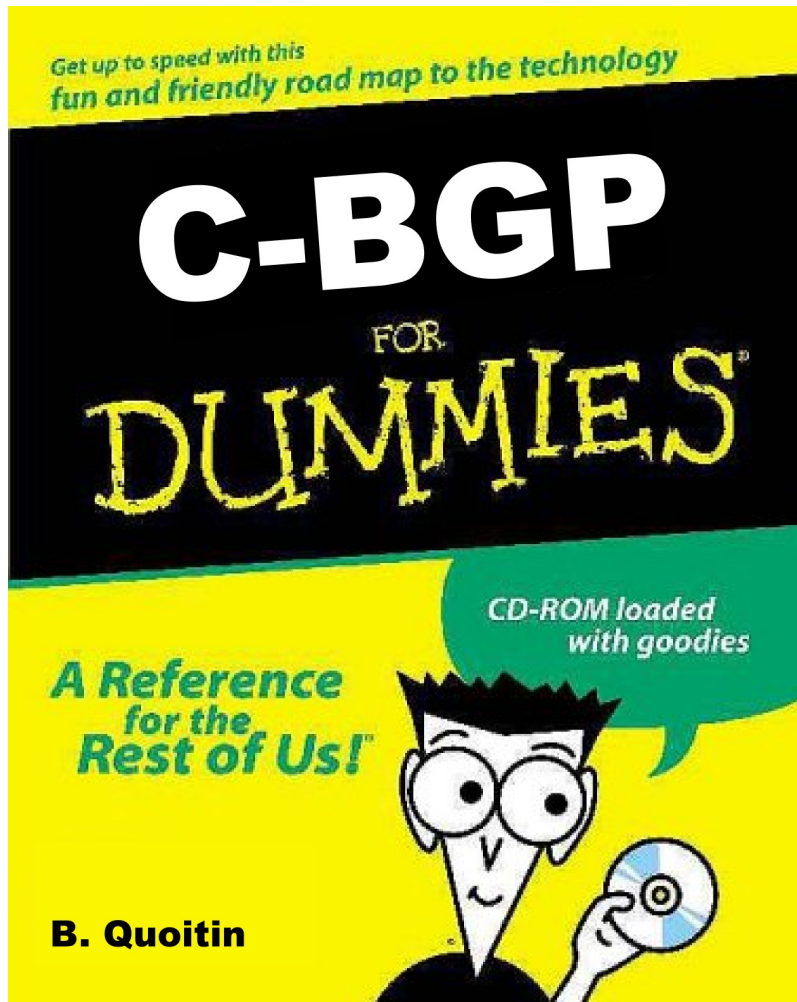
- **C-BGP:**
 - a network topology / config / routing DB
 - a BGP routing solver
- **Case studies**
 - Scenario 1: peering placement
 - Scenario 2: all single-link failures



I. C-BGP



C-BGP



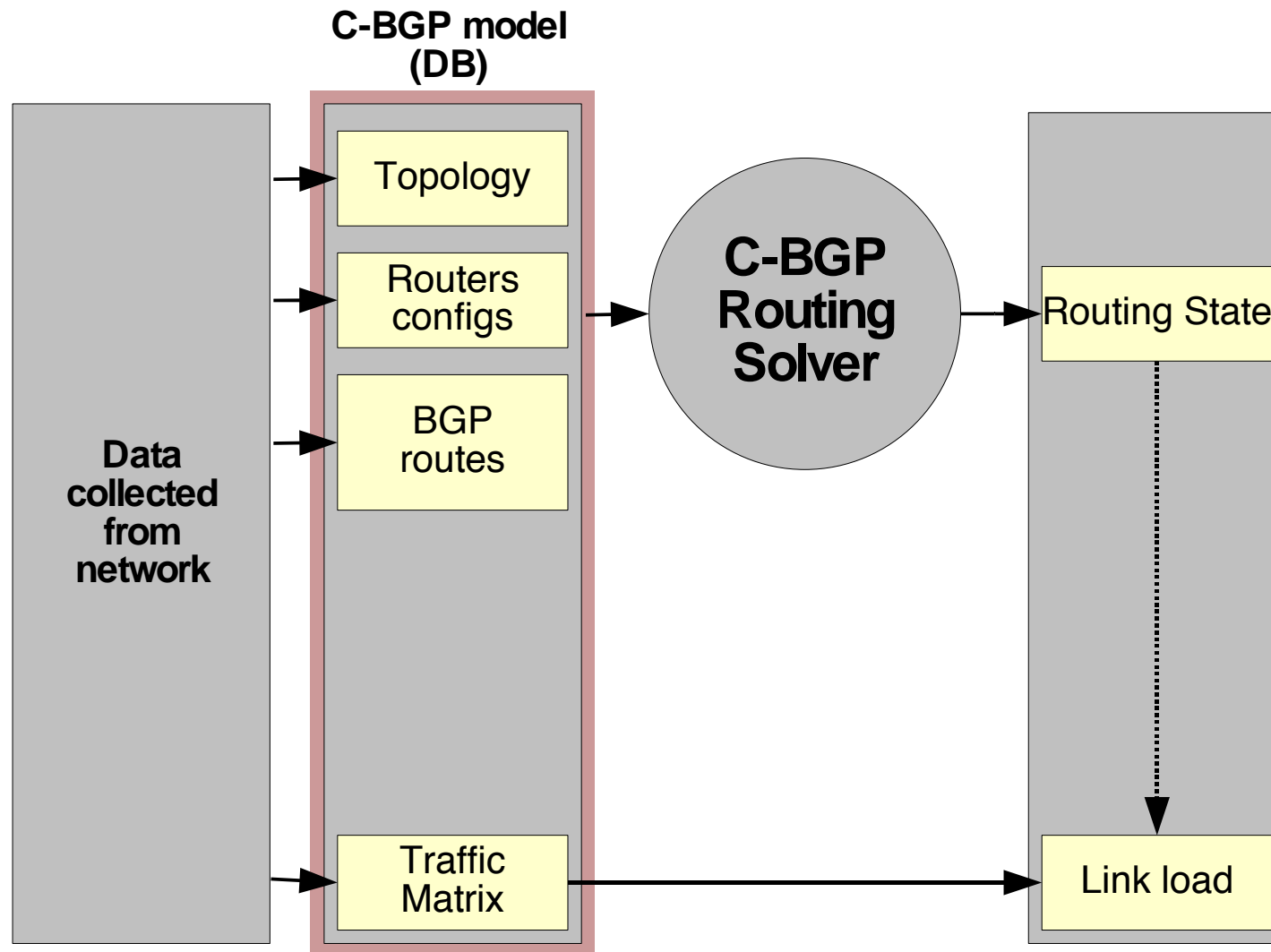
- Network topology / configuration DB
- Scriptable BGP Routing solver for large-scale networks
- developed by INL@UCLouvain
- supported by:



RÉGION WALLONNE



C-BGP Database



C-BGP Database

- **Network topology database (DB)**
 - **Layer-3 model:**
 - nodes: routers and LANs (pseudo-nodes)
 - links: IP links (need to be numbered)
 - link attributes: latency, bandwidth, load
 - **Additional information:**
 - static routes
 - IP tunnels
 - **Large-scale topologies**
 - tried with > 10000 nodes / 60000 links



DB: importing...

Build project from real router/network data

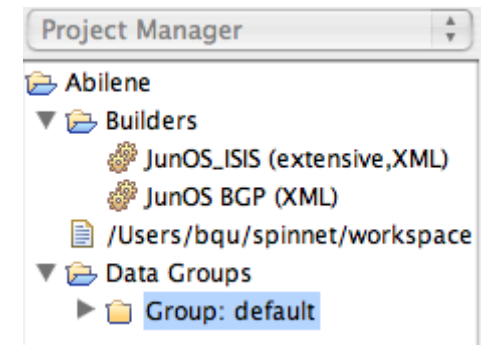


*"show isis
database extensive"*

"show bgp neighbors"



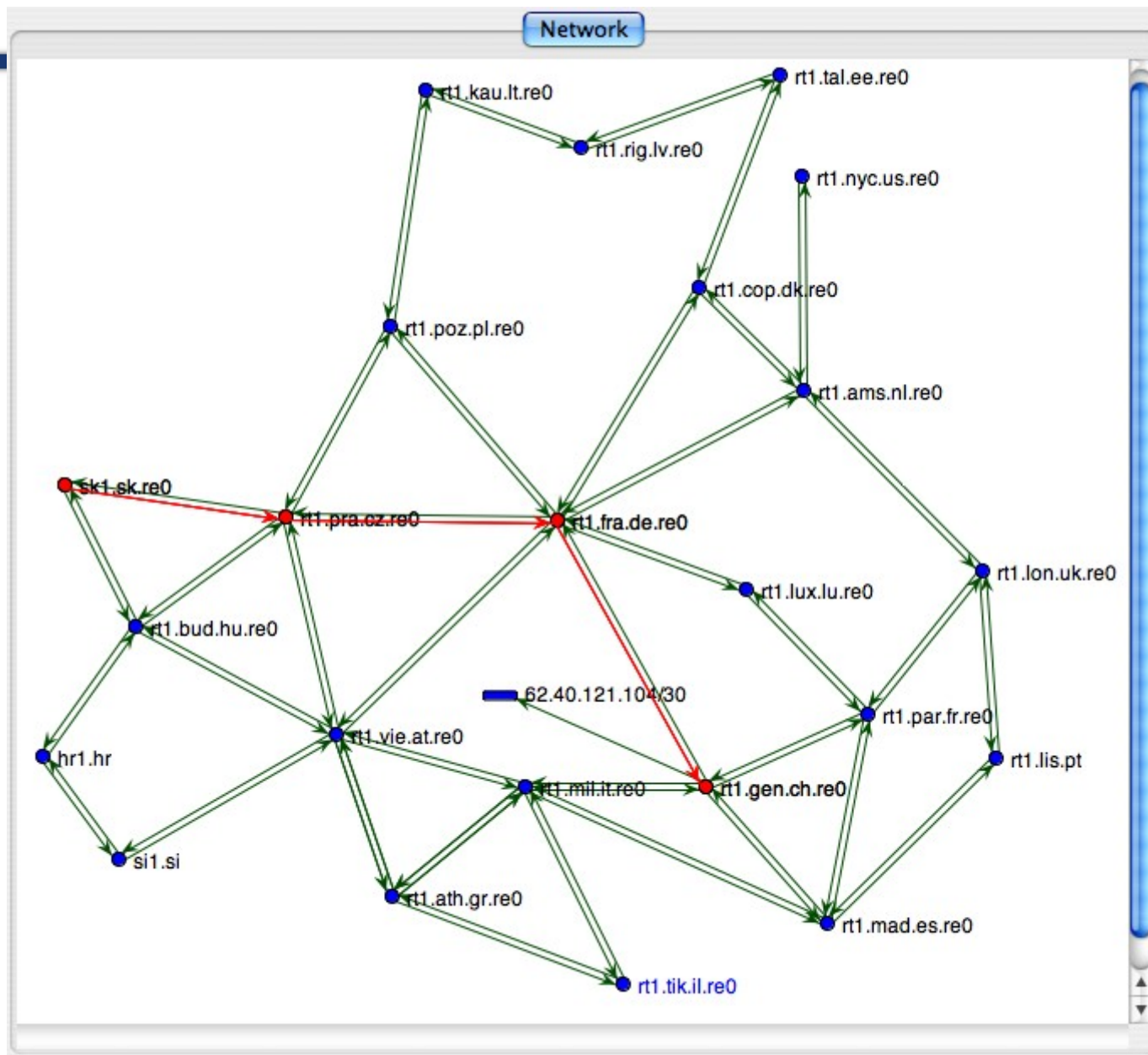
- JunOS IS-IS LSDB (TXT / XML)
- JunOS OSPF LSDB (TXT)
- IOS OSPF LSDB
- IOS IS-IS LSDB
- JunOS BGP neighbors (XML)
- IOS Running config
- JunOS Running config
- ...⁽¹⁾



⁽¹⁾ other vendor formats can be easily added



DB: network visualization

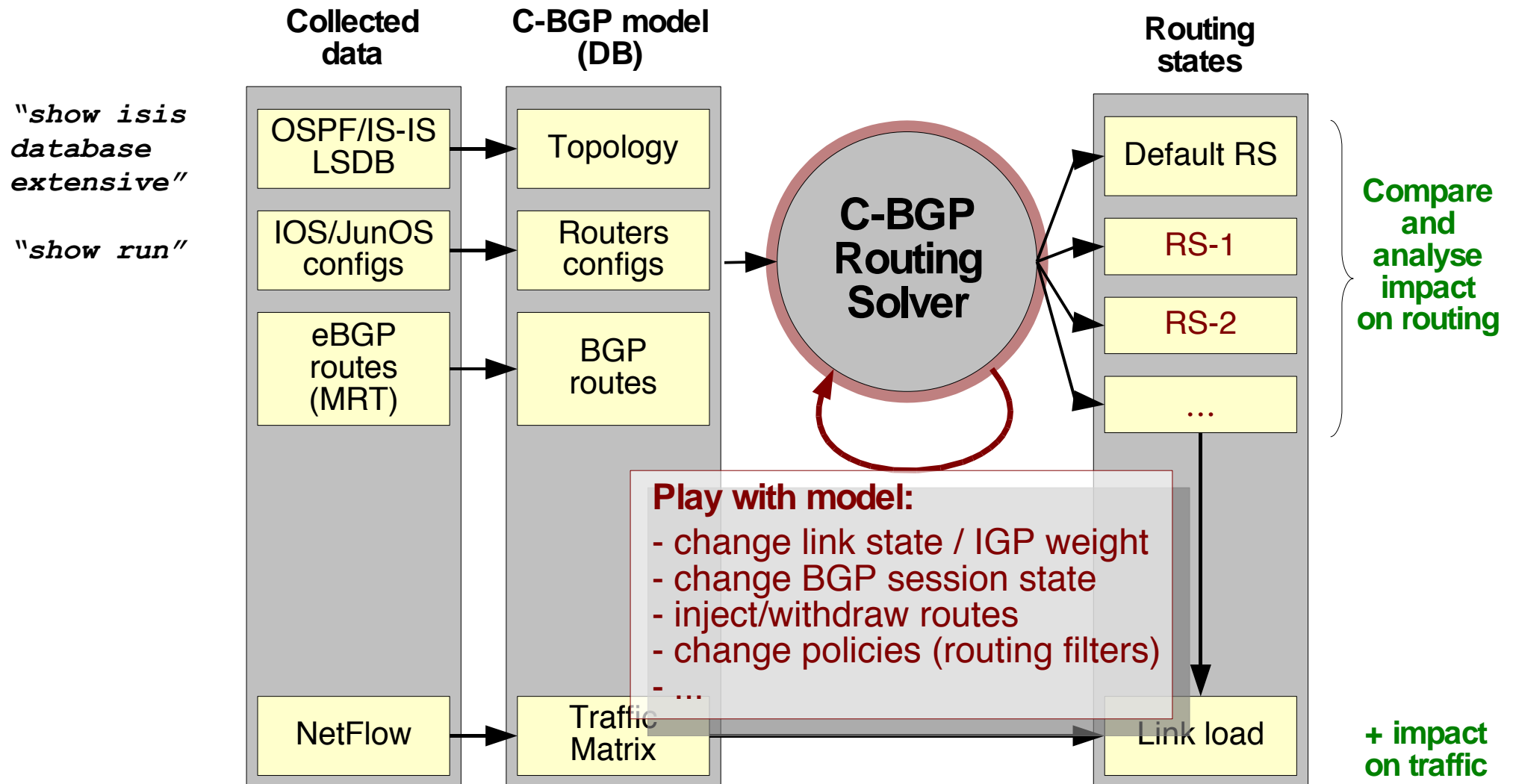


DB: network queries (CLI)

```
Console
cbgp> show version
cbgp version: 1.5.0 [zlib] [jni] [bgpdump]
libgds version: 1.4.5
cbgp> net node 62.40.114.3
cbgp-node> show ifaces
ptp      62.40.112.21/30
ptp      62.40.112.25/30
ptp      62.40.112.29/30
ptp      62.40.112.33/30
lo       62.40.114.3/32
ptmp     62.40.121.105/30
cbgp-node>
cbgp> net node 62.40.102.37
cbgp-node> traceroute 62.40.114.3
 1      62.40.96.41 (62.40.114.5)      icmp error (time-exceeded)
 2      62.40.112.38 (62.40.114.7)    icmp error (time-exceeded)
 3      62.40.114.3 (62.40.114.3)     reply
cbgp-node>
```



C-BGP Routing Solver



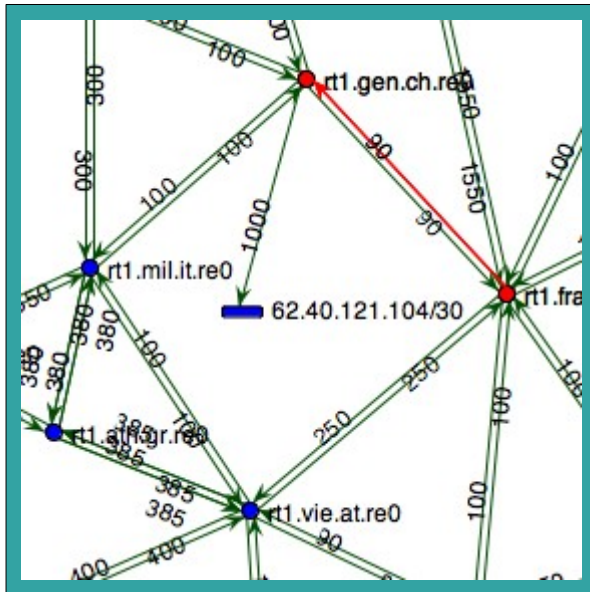
Routing Solver

- **IGP model**
 - static, centralized computation
 - support for ECMP, single area (currently)
- **BGP model**
 - compute steady-state outcome of BGP convergence
 - full decision process
 - versatile route filters
 - iBGP hierarchy (route-reflectors)
 - reads BGP table dumps and BGP message traces in MRT format (draft-ietf-grow-mrt-07.txt)

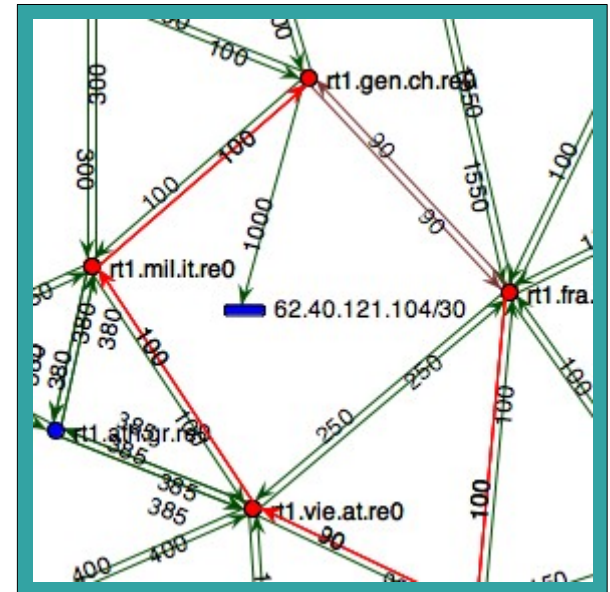


What-if scenarios

Before failure



After failure

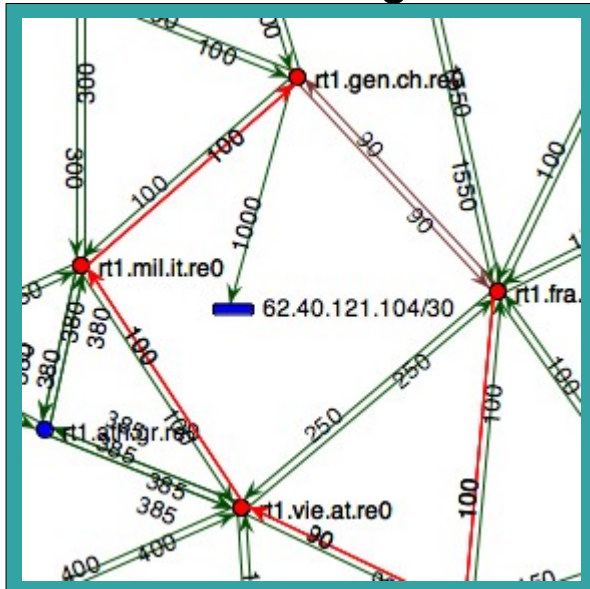


- 1). **Fail link** between Frankfurt and Geneva
- 2). Recompute routes
- 3). Trace route

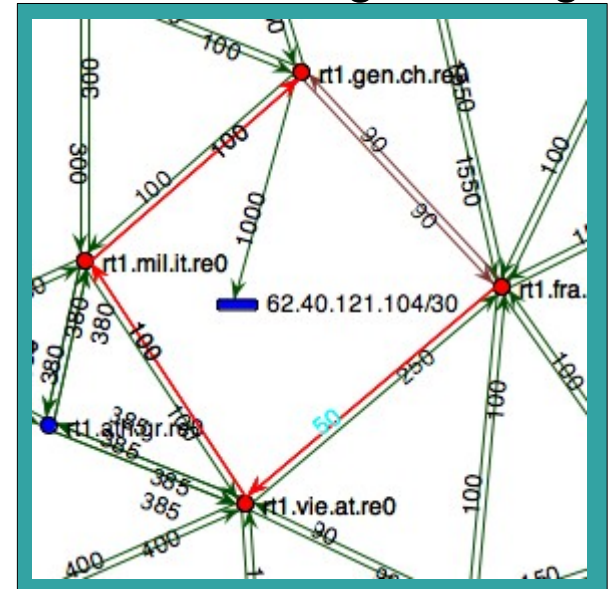


What-if scenarios

Before IGP weight change



After IGP weight change



- 1). Change **IGP weight** of link between Frankfurt and Vienna
- 2). Recompute routes
- 3). Trace route



Routing State Queries

The screenshot shows a network management interface titled "List of BGP Domains". On the left, a tree view shows "BGP domains" expanded to "AS 11537", listing various IP ranges and their corresponding regions (e.g., CHIC-re0, NEWY-re0, ATLA-re1, KANS-re0, SALT-re1, LOSA-re0, WASH-re0, HSTNng-re0, LOSAng-re1, STTLng-re0, SNVAng-re1, ATLA-m5). The "WASH-re0" entry is highlighted.

On the right, the "Info" tab is selected. The "Destination" field is set to "*", and the "Selection size" is 5145 route(s). Below this is a table of routing information:

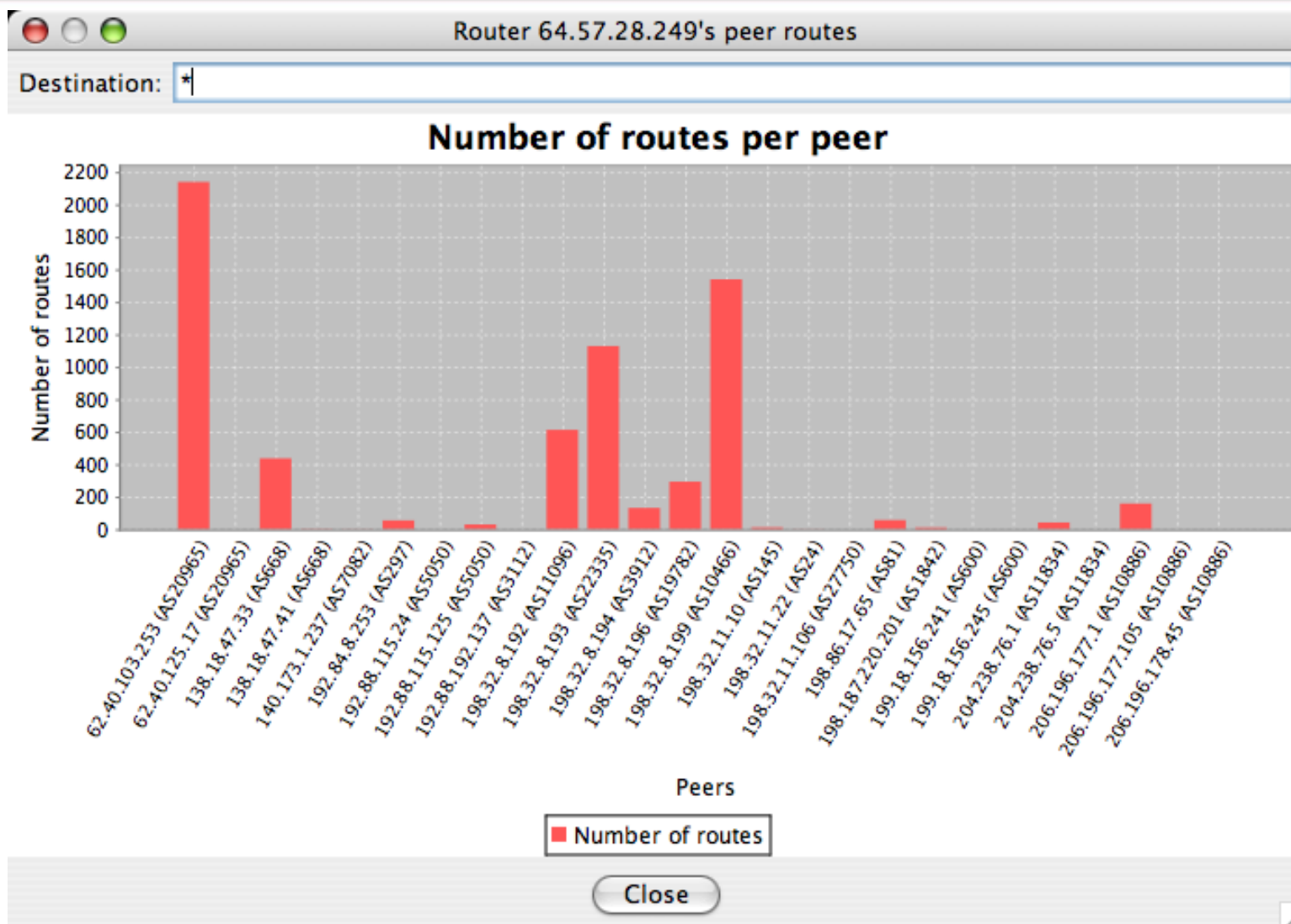
E	F	I	Prefix	Next-Ho	Local-	MED	Ori	AS-Path	Comm.
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	8.6.244....	198.32...	200	0	IGP	11096 6...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	8.10.208...	198.32...	260	0	IGP	10466 3...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	9.4.0.0/16	62.40...	100	0	IGP	20965 559	2096...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.0.48....	198.32...	200	0	IGP	10578 1...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.6.208...	198.32...	200	0	IGP	10578 1...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.107.2...	198.86...	200	0	IGP	81 22753	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.144.5...	198.32...	260	0	IGP	10466 1...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.151.0...	198.32...	260	0	IGP	10466 1...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.151.1...	198.32...	260	0	IGP	10466 1...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.161.8...	198.32...	260	0	IGP	10466 88	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	12.174.2...	198.32...	200	0	IGP	5661 21...	1153...
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	18.3.4.0/...	198.32...	200	0	IN...	10578 3	1153...

Below the table, a summary for the selected route (12.144.59.0/24) is shown:

Destination: 12.144.59.0/24
Best: true
Feasible: true
Internal: false
Next-hop: 198.32.8.199
AS-Path: 10466 13778
Local-Pref: 260
MED: 0
Communities: 11537:2000 11537:950 11537:260



Routing State Queries



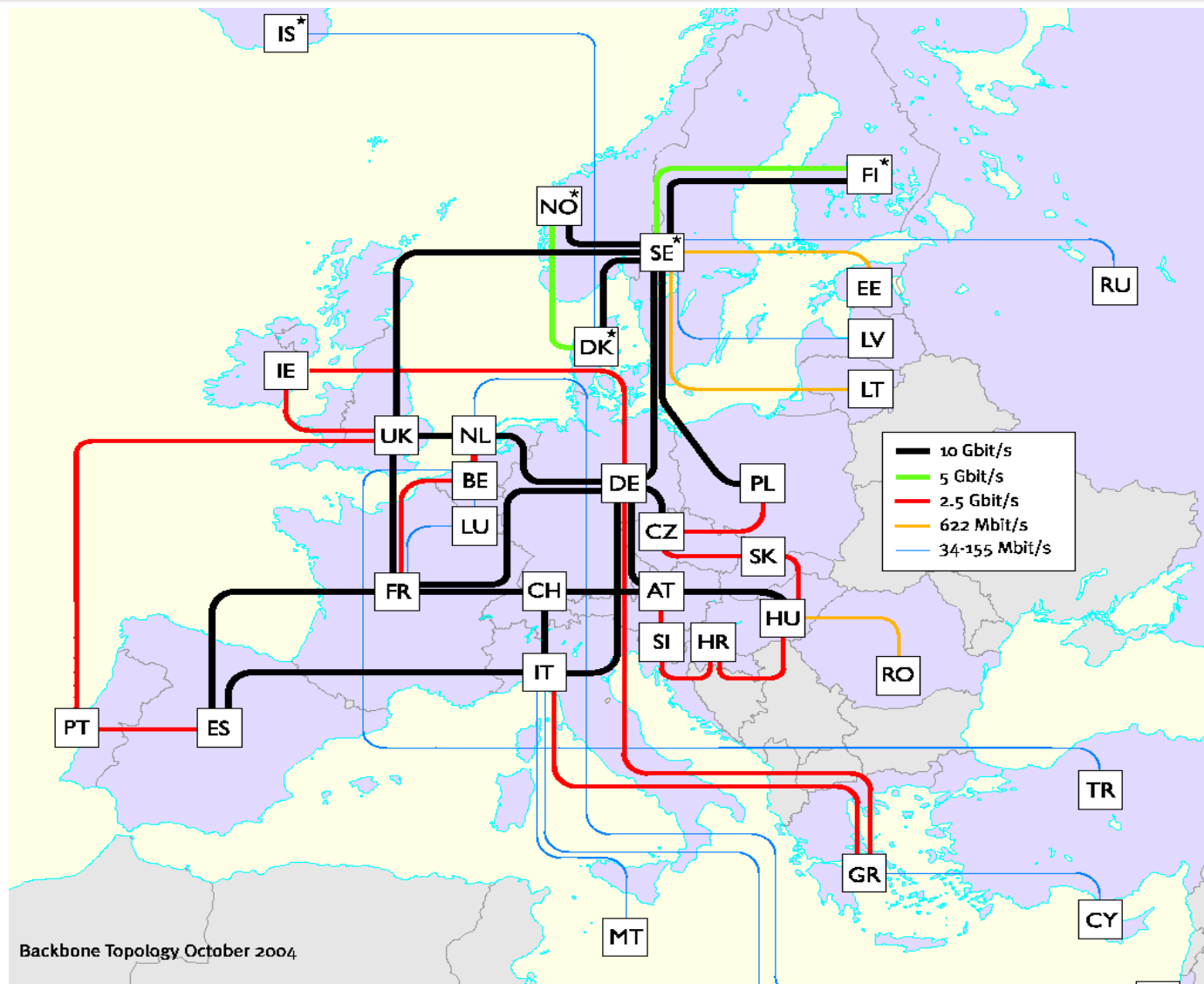
II. Case studies



*Anonymous
French Tier-1*



Case study: GEANT (AS20965)



Case study: GEANT (AS20965)

- **Topology**

- Obtained from IS-IS trace, cross-checked with map
- 23 nodes, 38 core links, 53 edge links (6 with upstreams)

- **Routing data**

- Collected using Zebra in the iBGP (only best eBGP)
- 640,897 eBGP routes
 - 150,071 prefixes (clustered in 406 groups)

- **Traffic data**

- NetFlow collected on all external interfaces
- Sampling rate: 1/1000
- About 150 GB per month
- Src. / dst. aggregated in /24 prefixes

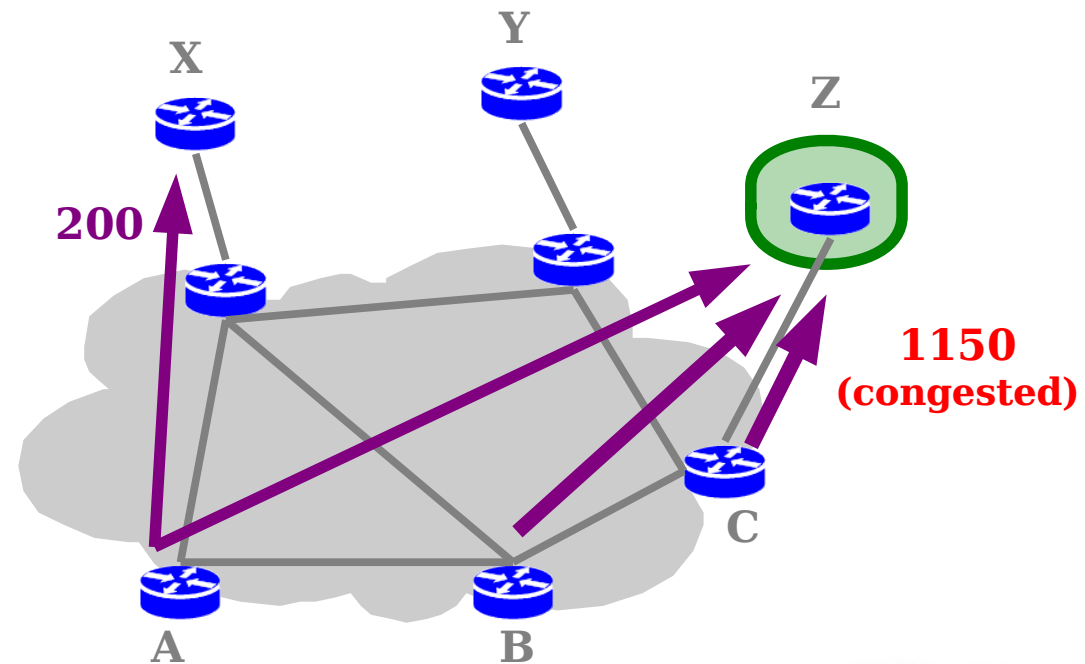
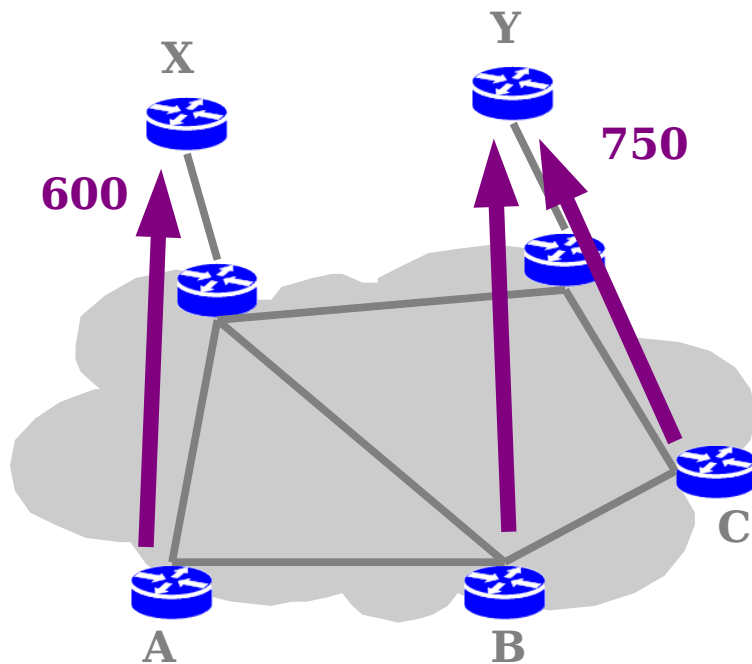


1st Scenario: peering placement

- **Example**

- 2 upstream providers, 1Gbps links
- Peer with new provider Z in C

	X	Y
A	600	0
B	0	250
C	0	500



1st Scenario: peering placement

- **Objective**

- Investigate addition/removal of peerings
- Goal: better balance traffic load, reduce peering cost, ...

- **Methodology**

- **Scenario add-Rx**

- Consider a prospective peering *PR* (full RIB)
- Inject routes of *PR* at router *Rx*

- **Scenario del-PRx**

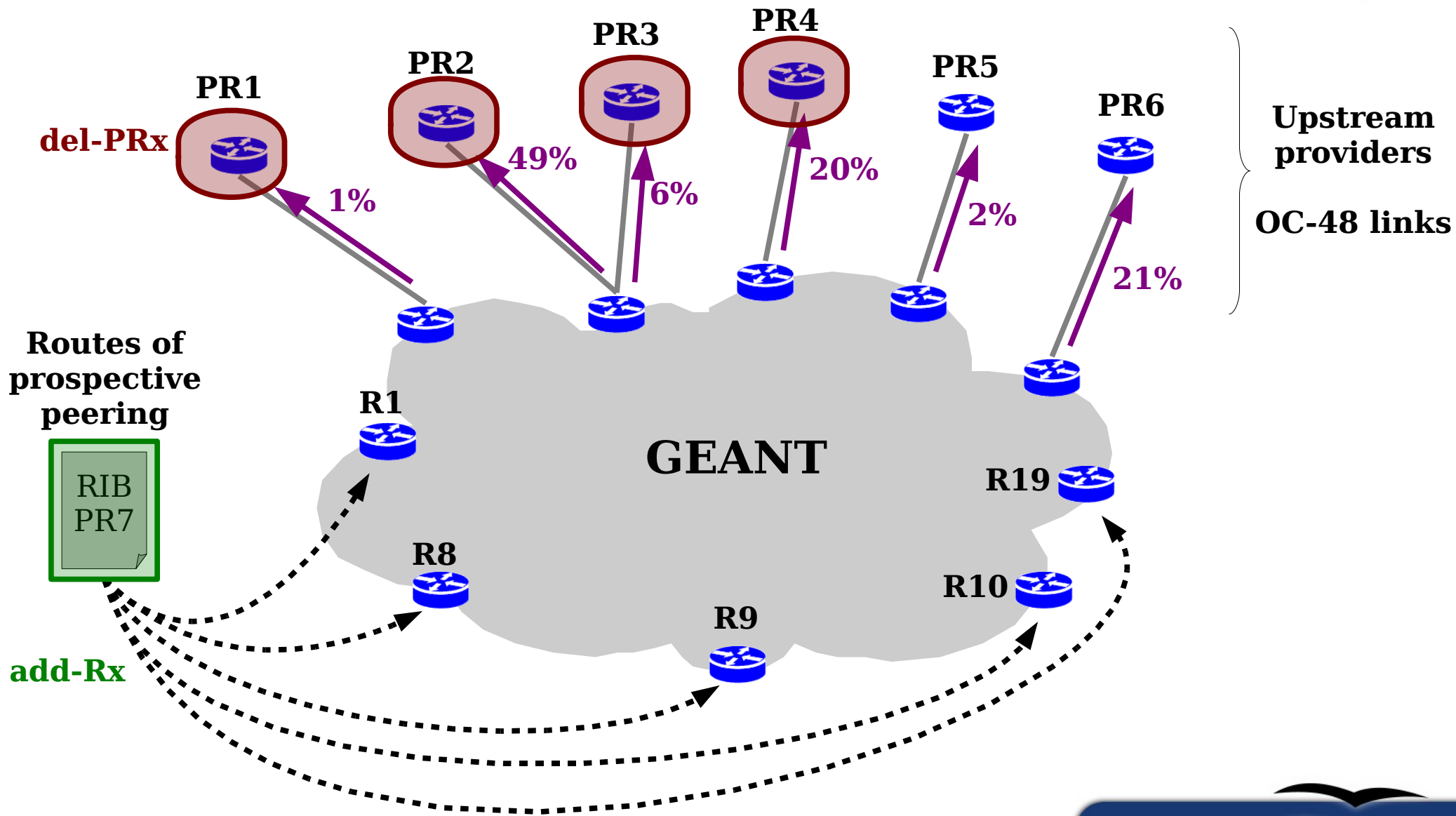
- Remove the routes learned from an existing peer *PRx*

- **Metric**

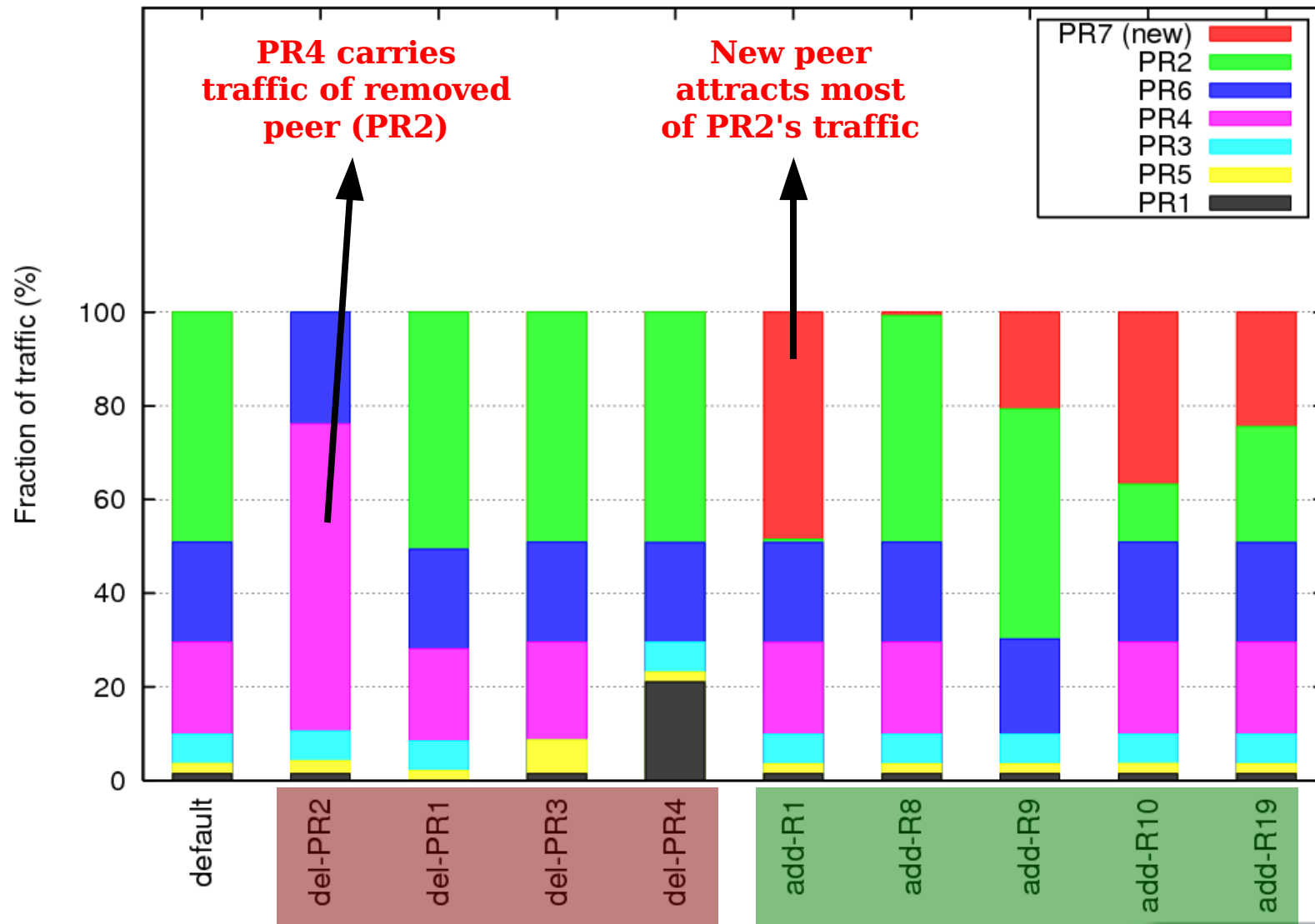
- distribution of traffic among peering links
(here: 6 most important links, OC-48 with upstream providers)



1st Scenario: peering placement



1st Scenario: peering placement

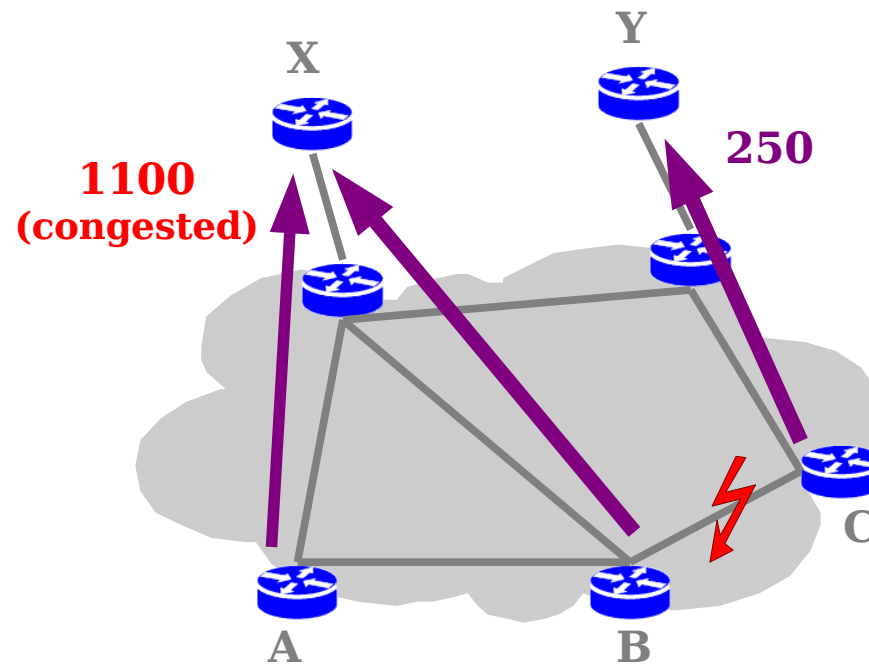
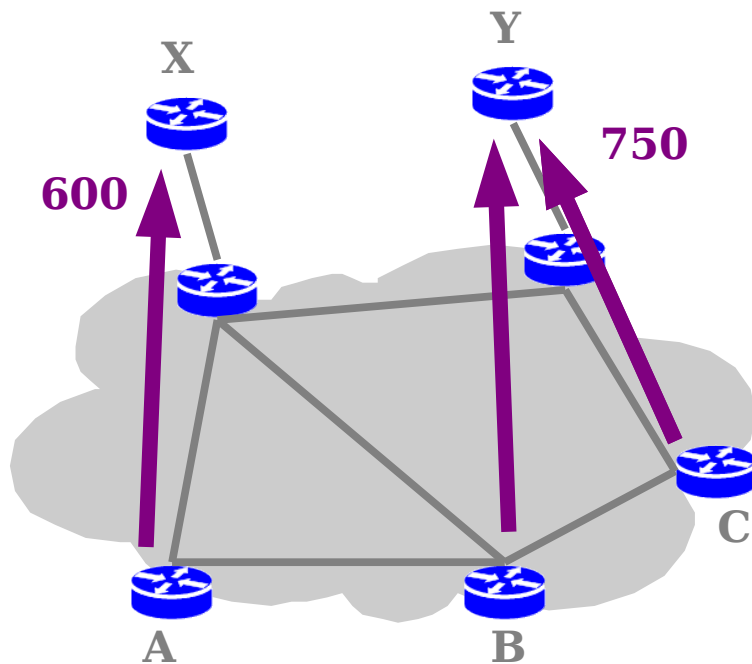


2nd Scenario: link failures

- **Example**

- Traffic to upstream X and Y
- 1 Gbps links
- Internal link failure: $B \leftrightarrow C$

	X	Y
A	600	0
B	0	500
C	0	250



2nd Scenario: link failures

- **Objectives**

- Study impact of **single-link internal** failures on routing
- Consider all interdomain routes

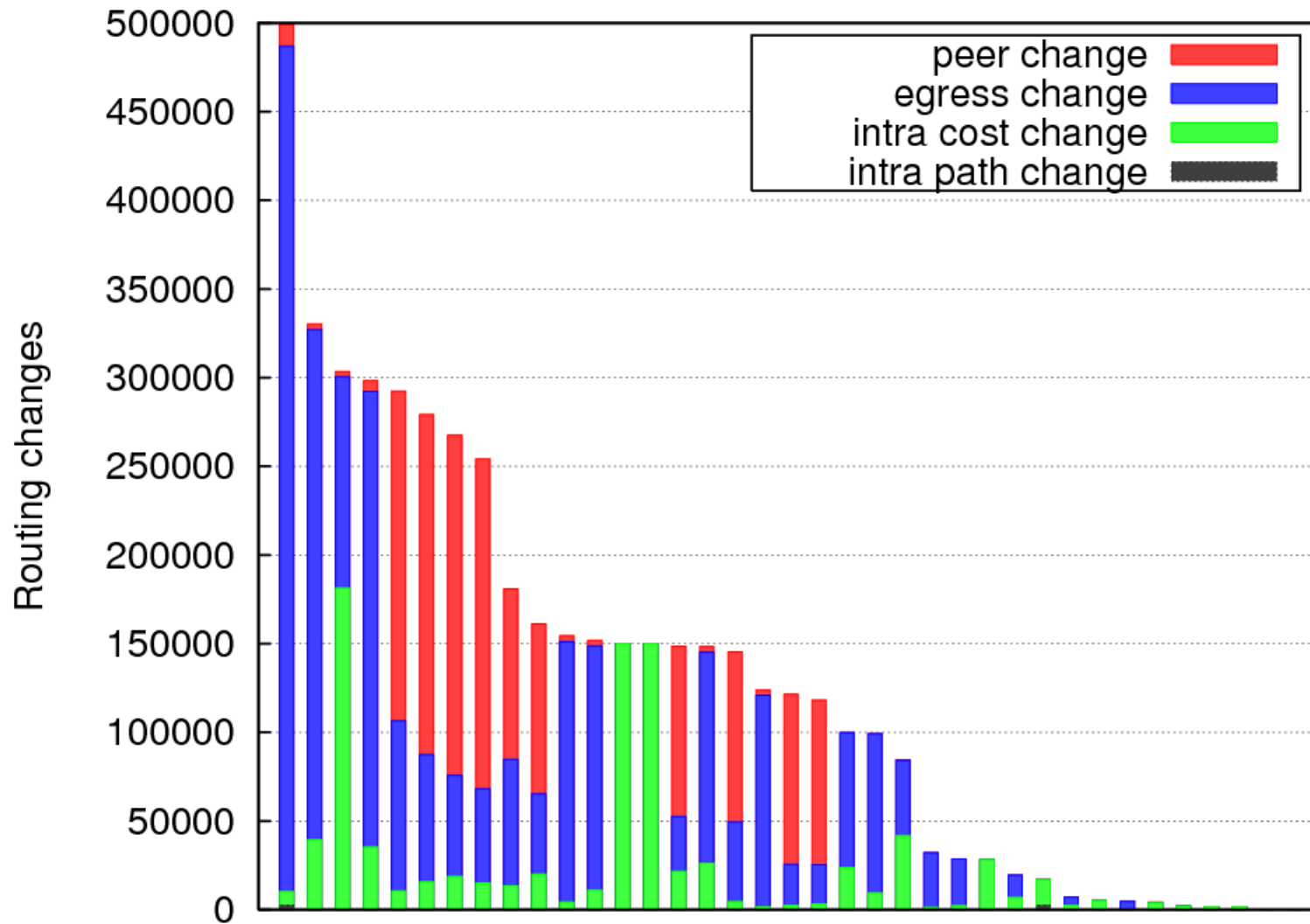
- **Methodology**

- **Classification of routing changes**

- Prefix reachability
- Peer change: neighbor AS has changed
- Egress change: same AS, egress router changed
- Intra cost change: same egress, IGP cost changed
- Intra path change: same egress, same IGP cost, path changed (only when ECMP is allowed)



2nd Scenario: link failures



**Most changes
are interdomain
changes !!!**



Conclusion

- **Modeling the routing of an ISP is complex !**
 - Many parameters and data sources are involved.
- **Tools such as C-BGP**
 - Helps **understand & visualize** routing protocol interaction (IGP / BGP) on large networks with many destinations
 - Useful to predict **impact of events / choices** on routing
 - Can be used as/with a **capacity planning** tool (if TM available) to predict the impact on link load



Thanks for your attention !

Visit our website:
<http://cbgp.info.ucl.ac.be>

The C-BGP core is released under the LGPL license.

The C-BGP core has been tested under the following platforms:



FreeBSD®



References

- ***Modeling the Routing of an ISP Network***, B. Quoitin and S. Uhlig, IEEE Network, Vol 19(6), November 2005.
- ***Semi-automatic AS-wide converter for C-BGP***, S. Tandel. Available from <http://alumni.info.ucl.ac.be/standel/bgp-converter>
- ***Providing public intradomain traffic matrices to the research community***, S. Uhlig, B. Quoitin, S. Balon and J. Lepropre, ACM SIGCOMM Computer Communication Review, Vol 36(1), January 2006.
- ***The Interaction of IGP Weight Optimization with BGP***, S. Cerav-Erbas, O. Delcourt, B. Fortz and B. Quoitin, In Proceedings of ICISP'06, p. 9, August 26 - 29, 2006.
- ***Network-Wide Prediction of BGP Routes***, N. Feamster and J. Rexford, IEEE/ACM Transactions on Networking, April 2007.
- ***TOTEM toolbox***. Available from <http://totem.run.montefiore.ulg.ac.be>

