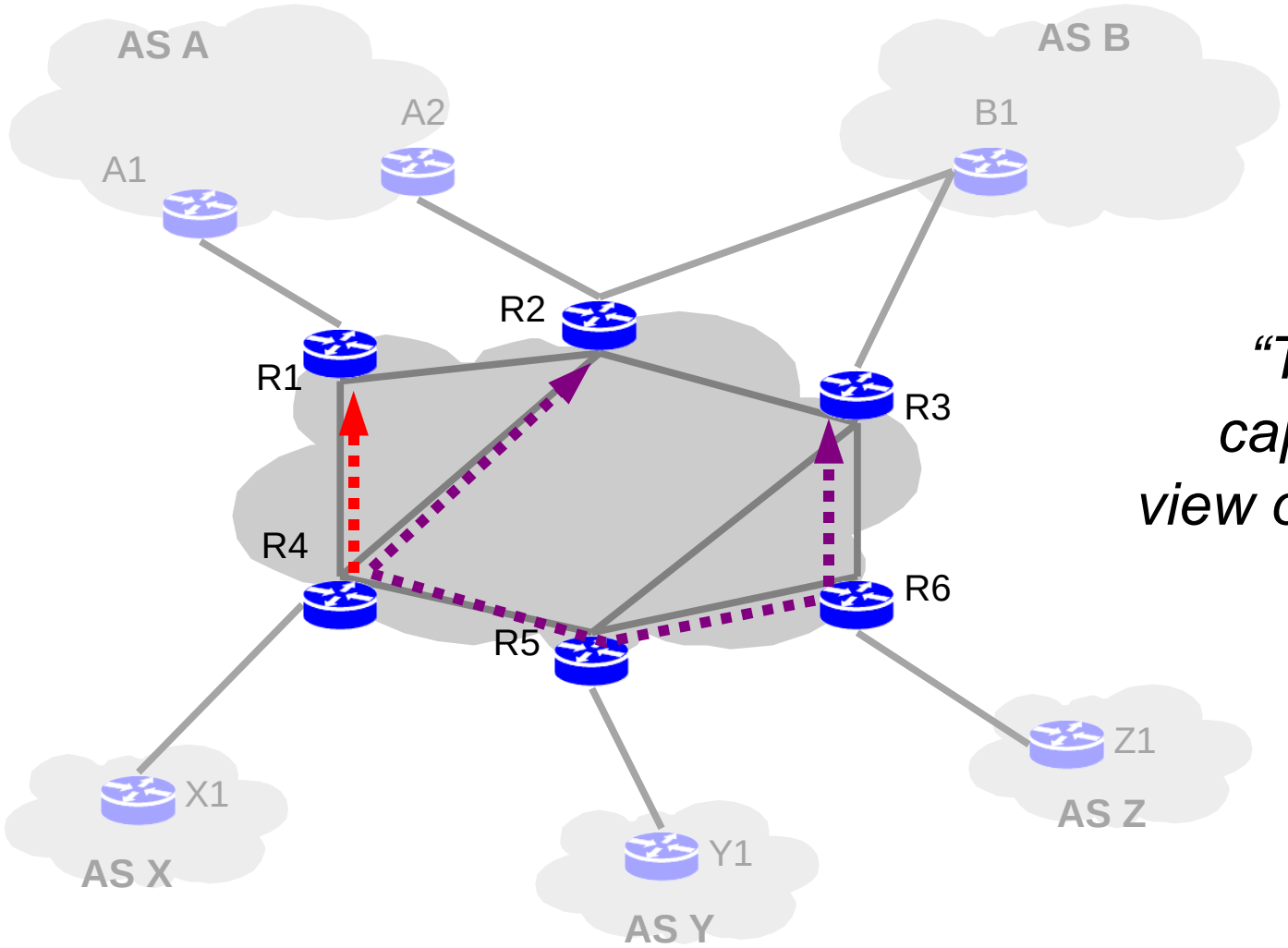
Modeling the Routing of an ISP with C-BGP

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 flows

Can't answer questions such as...

What would happen to my interdomain traffic if...

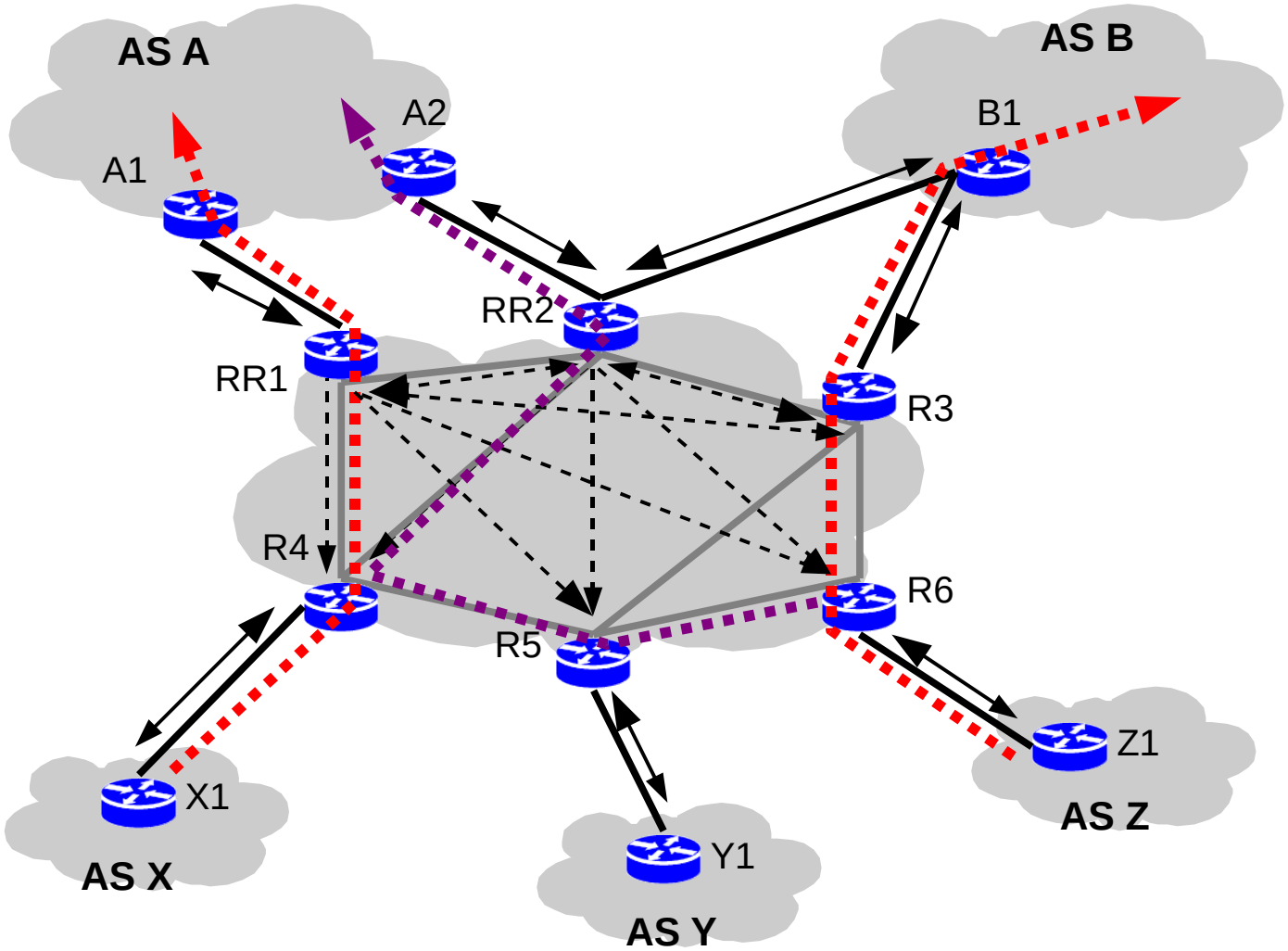
- a link is failing ?
- a router is under maintenance ?
- a BGP peering is being shutdown ?
- a new route filtering policy is planned ?
- a new peering is established at an IXP ?

How would I optimize my interdomain routing for...

- performance ? cost ? reliability ?

How do I compare prospective providers ?

ISP Model



Reality has:

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

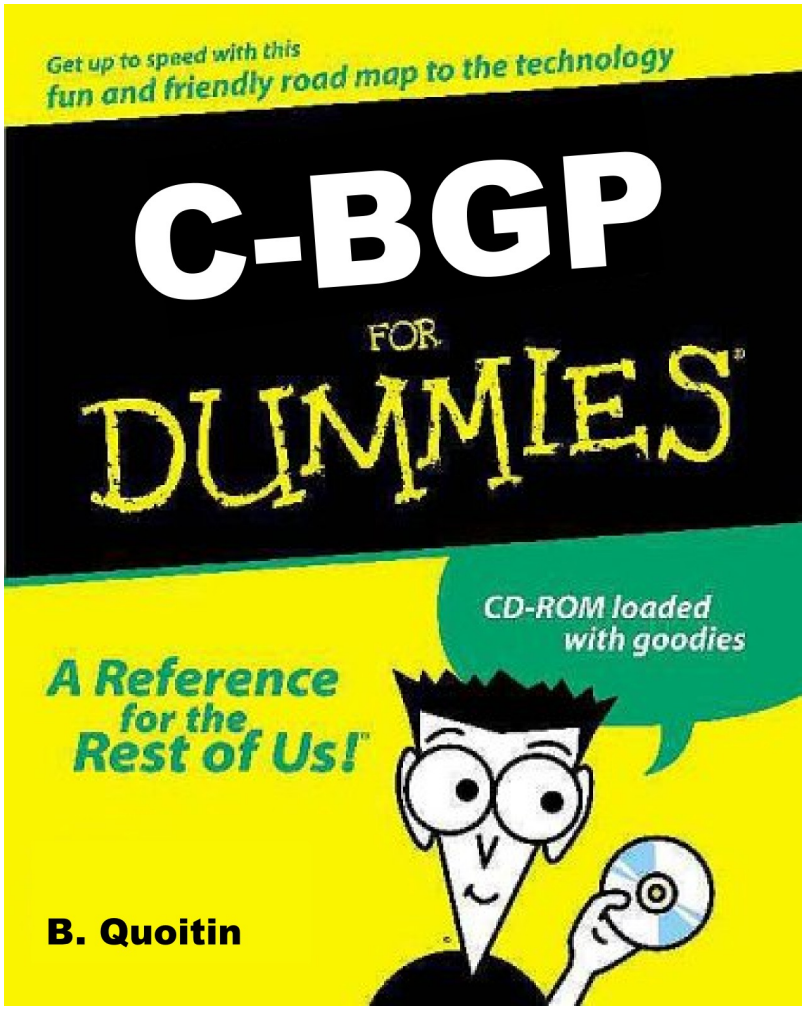
Internal link
External link
eBGP session
iBGP session
iBGP RR-client session
Traffic flows

Agenda

- **C-BGP (backend)**
  - a network topology / config / routing DB
  - a BGP routing solver
- **SPINNET (frontend)**
  - network data parsers (model builders)
  - network topology visualization
- **Case study**
  - Peering placement

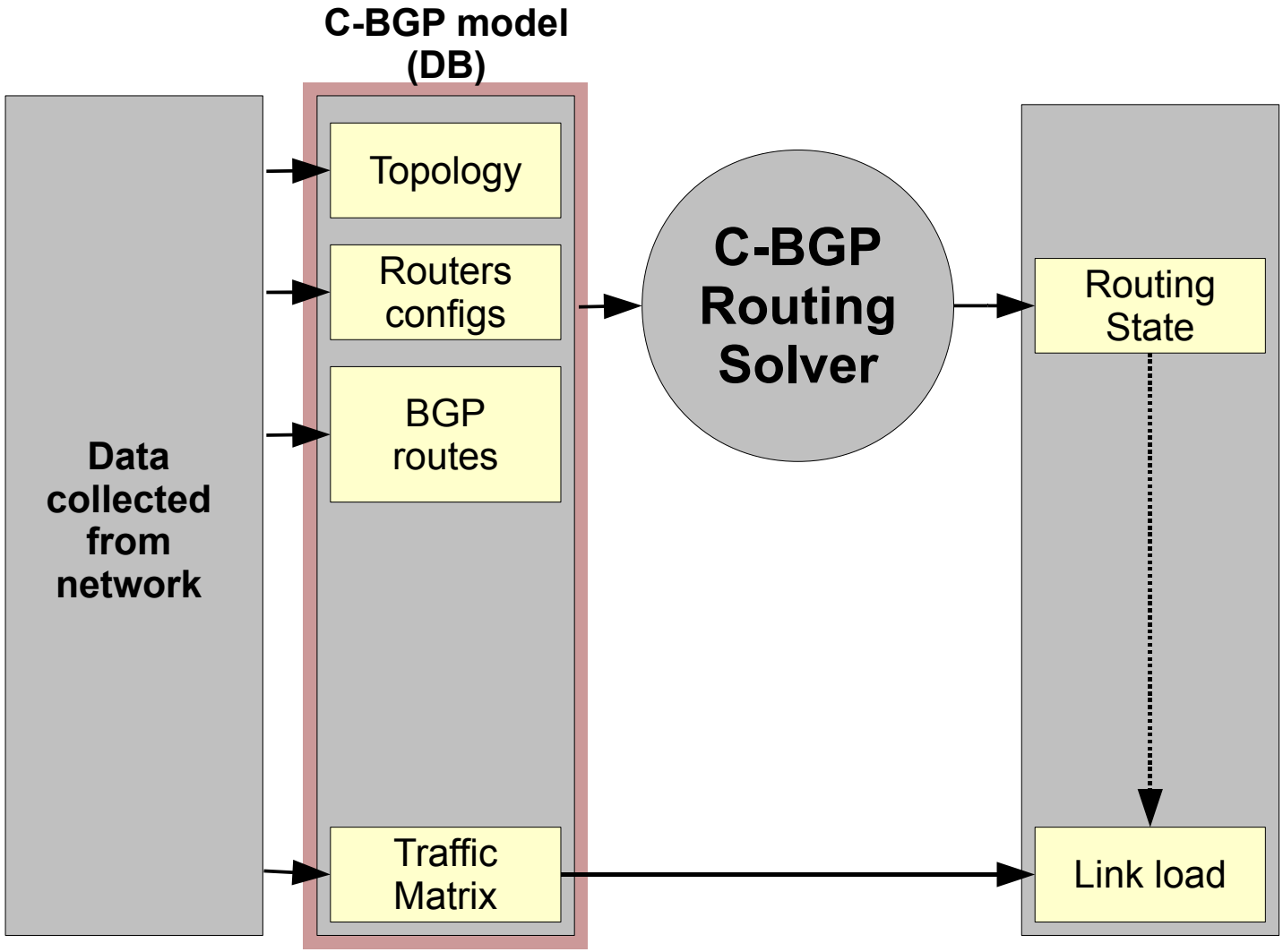
I. C-BGP

C-BGP



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

C-BGP Database



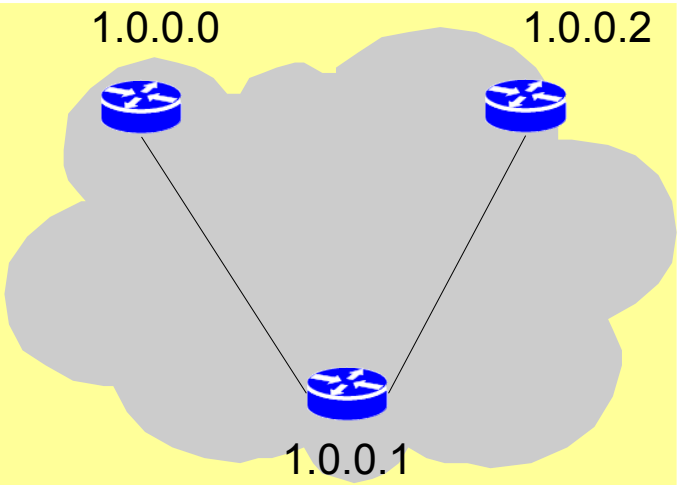
C-BGP Database

- **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**

C-BGP Script Example

Topology

```
net add node 1.0.0.0
net add node 1.0.0.1
net add node 1.0.0.2
net add link 1.0.0.0 1.0.0.1
net add link 1.0.0.1 1.0.0.2
```



Note: c-bgp v2.0.0 is required.

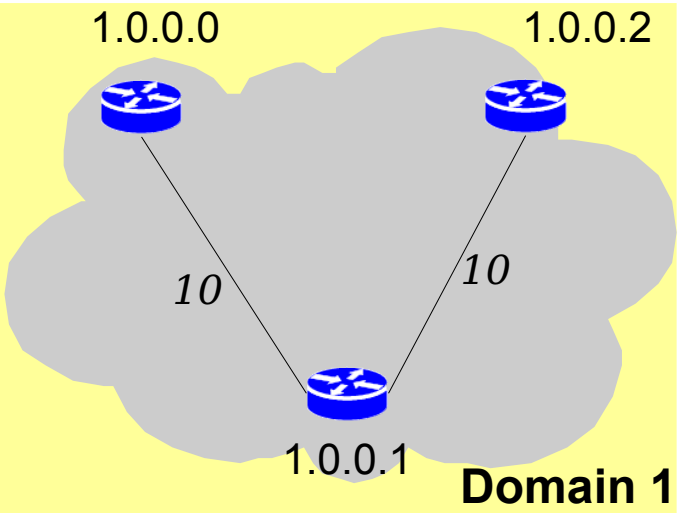
C-BGP Script Example

Topology

```
net add node 1.0.0.0
net add node 1.0.0.1
net add node 1.0.0.2
net add link 1.0.0.0 1.0.0.1
net add link 1.0.0.1 1.0.0.2
```

IGP

```
net add domain 1 igp
net node 1.0.0.0 domain 1
net node 1.0.0.1 domain 1
net node 1.0.0.2 domain 1
net link 1.0.0.0 1.0.0.1 igp-weight --bidir 10
net link 1.0.0.1 1.0.0.2 igp-weight --bidir 10
net domain 1 compute
```



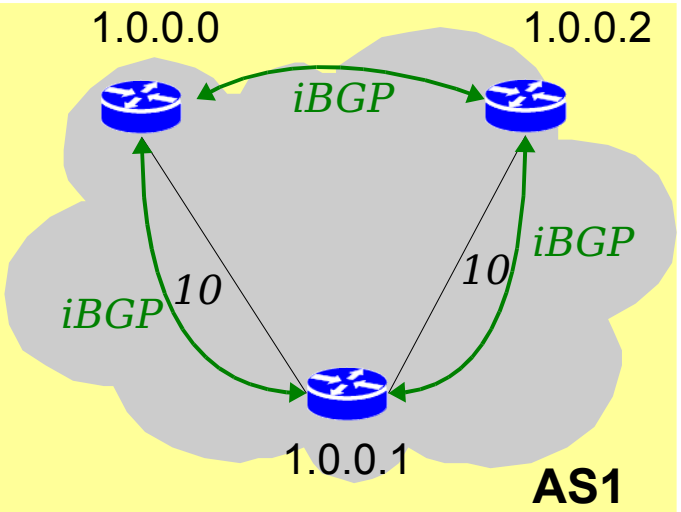
Note: c-bgp v2.0.0 is required.

C-BGP Script Example

Topology
```
net add node 1.0.0.0
net add node 1.0.0.1
net add node 1.0.0.2
net add link 1.0.0.0 1.0.0.1
net add link 1.0.0.1 1.0.0.2
```

IGP
```
net add domain 1 igp
net node 1.0.0.0 domain 1
net node 1.0.0.1 domain 1
net node 1.0.0.2 domain 1
net link 1.0.0.0 1.0.0.1 igp-weight --bidir 10
net link 1.0.0.1 1.0.0.2 igp-weight --bidir 10
net domain 1 compute
```

BGP
```
bgp add router 1 1.0.0.0
bgp add router 1 1.0.0.1
bgp add router 1 1.0.0.2
bgp domain 1 full-mesh
```



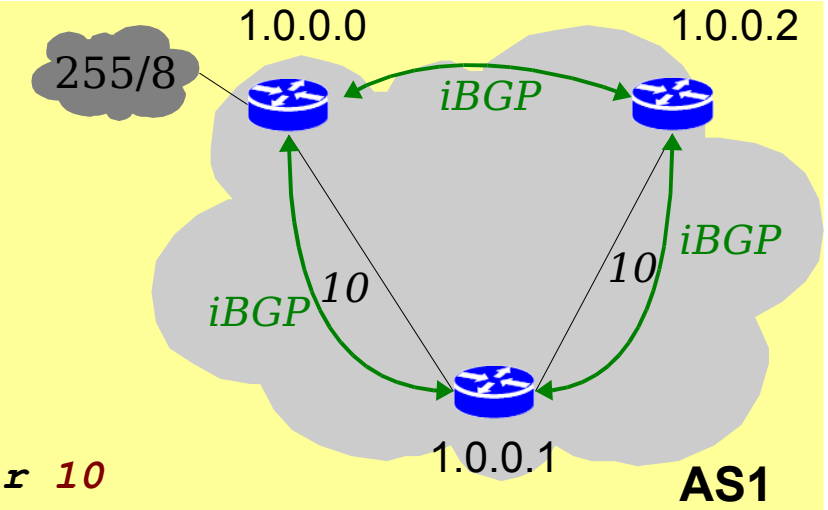
Note: c-bgp v2.0.0 is required.

C-BGP Script Example

Topology
```
net add node 1.0.0.0
net add node 1.0.0.1
net add node 1.0.0.2
net add link 1.0.0.0 1.0.0.1
net add link 1.0.0.1 1.0.0.2
```

IGP
```
net add domain 1 igp
net node 1.0.0.0 domain 1
net node 1.0.0.1 domain 1
net node 1.0.0.2 domain 1
net link 1.0.0.0 1.0.0.1 igp-weight --bidir 10
net link 1.0.0.1 1.0.0.2 igp-weight --bidir 10
net domain 1 compute
```

BGP
```
bgp add router 1 1.0.0.0
bgp add router 1 1.0.0.1
bgp add router 1 1.0.0.2
bgp domain 1 full-mesh
bgp router 1.0.0.0 add network 255/8
sim run
```



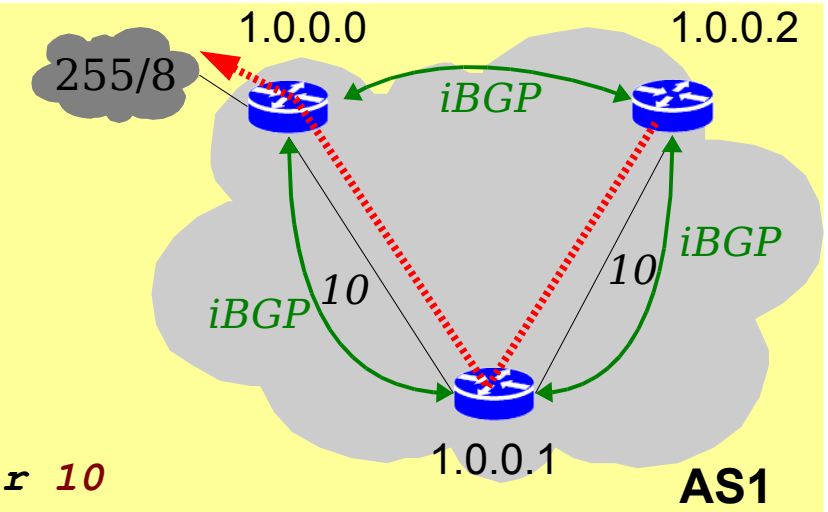
Note: c-bgp v2.0.0 is required.

C-BGP Script Example

Topology
```
net add node 1.0.0.0
net add node 1.0.0.1
net add node 1.0.0.2
net add link 1.0.0.0 1.0.0.1
net add link 1.0.0.1 1.0.0.2
```

IGP
```
net add domain 1 igp
net node 1.0.0.0 domain 1
net node 1.0.0.1 domain 1
net node 1.0.0.2 domain 1
net link 1.0.0.0 1.0.0.1 igp-weight --bidir 10
net link 1.0.0.1 1.0.0.2 igp-weight --bidir 10
net domain 1 compute
```

BGP
```
bgp add router 1 1.0.0.0
bgp add router 1 1.0.0.1
bgp add router 1 1.0.0.2
bgp domain 1 full-mesh
bgp router 1.0.0.0 add network 255/8
sim run
net node 1.0.0.2 record-route 255.0.0.0
```



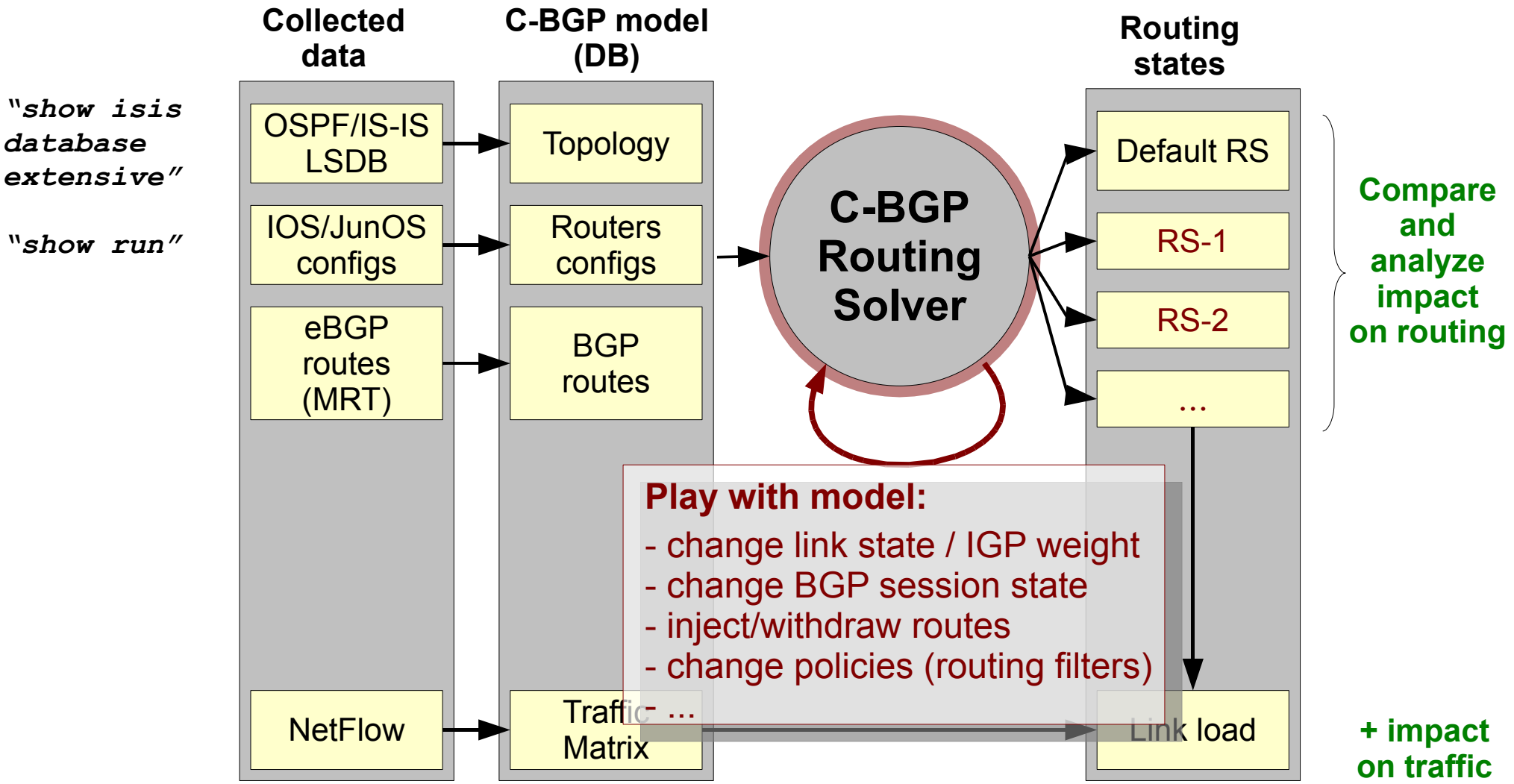

```
1.0.0.2 255.0.0.0     UNREACH 3       1.0.0.2 1.0.0.1 1.0.0.0
```

Note: c-bgp v2.0.0 is required.

C-BGP 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)
  - **multiple BGP domains (AS)**

C-BGP Routing Solver



II. SPINNET

DB: importing...

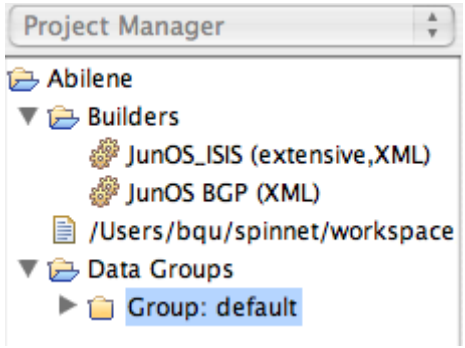
Build project from real router/network data

"show isis database extensive"

"show bgp neighbors"

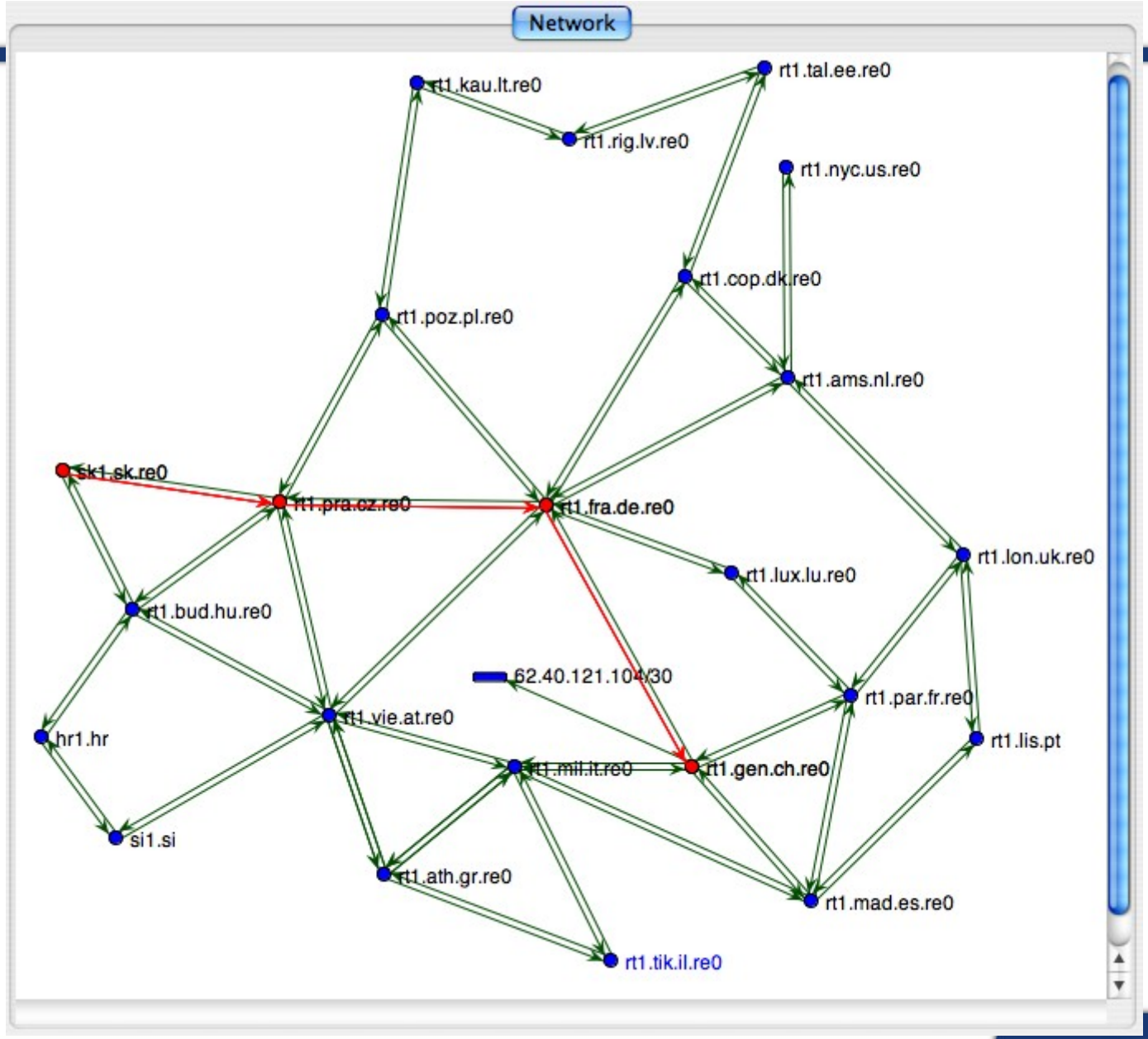
"show ip ospf database router/network/..."

- JunOS IS-IS LSDB (TXT / XML)
- JunOS OSPF LSDB (TXT)
- IOS OSPF LSDB
- IOS IS-IS LSDB
- JunOS BGP neighbors (XML)
- IOS Running conf g
- JunOS Running conf g
- … [1]

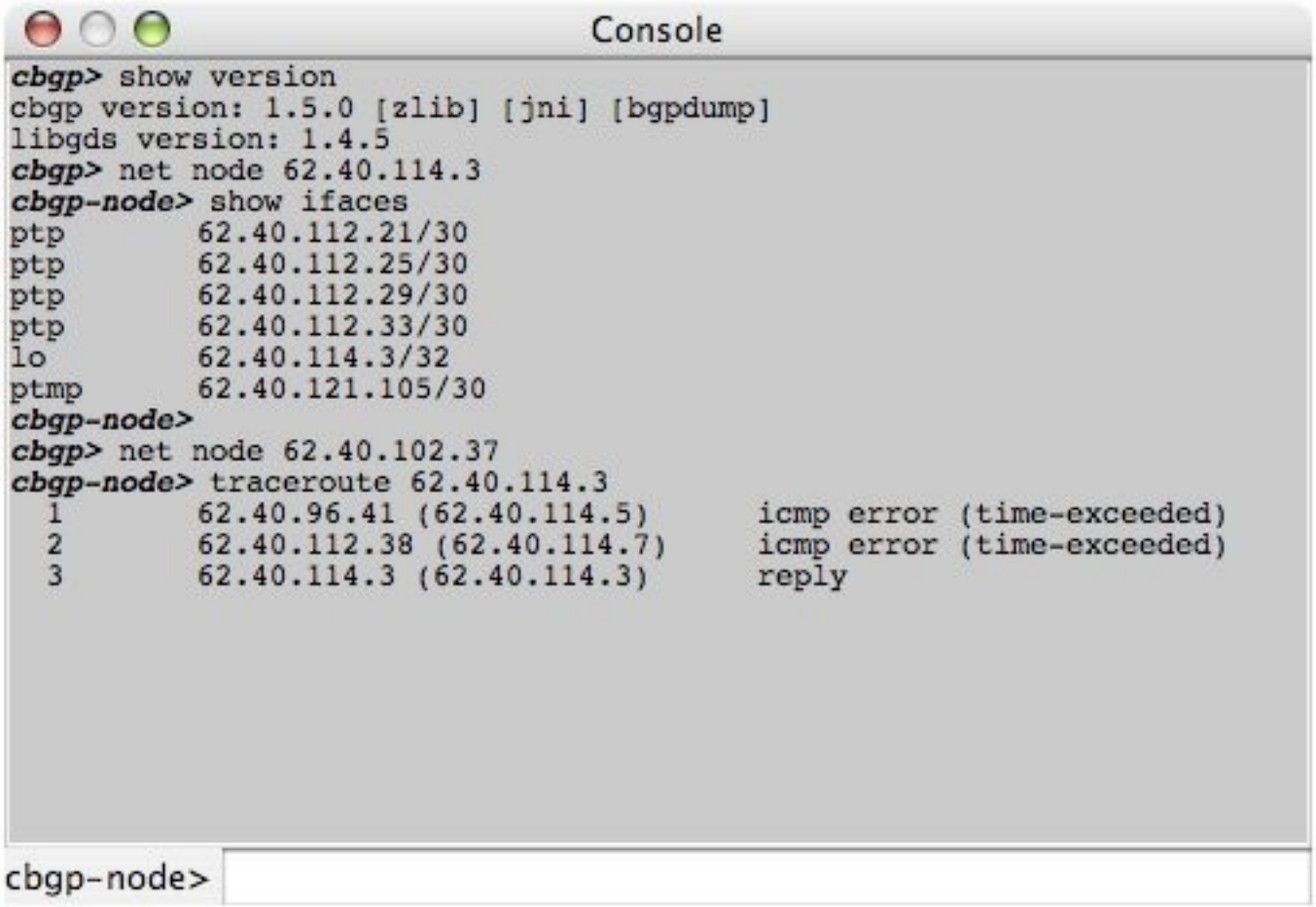


[1] other vendor formats can be easily added

DB: network visualization

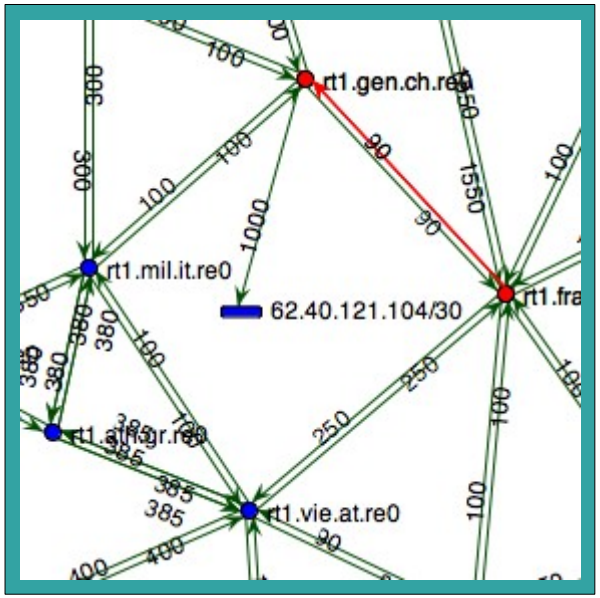


DB: network queries (CLI)



What-if scenarios

Before failure

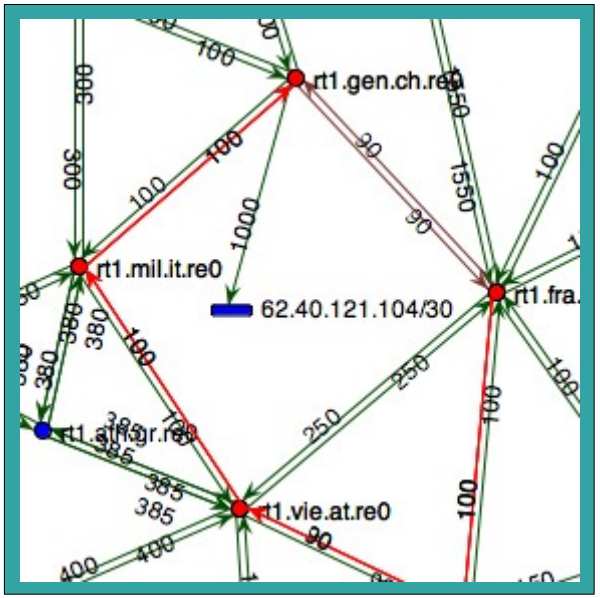


1). **Fail link** between Frankfurt and Geneva

2). Recompute routes

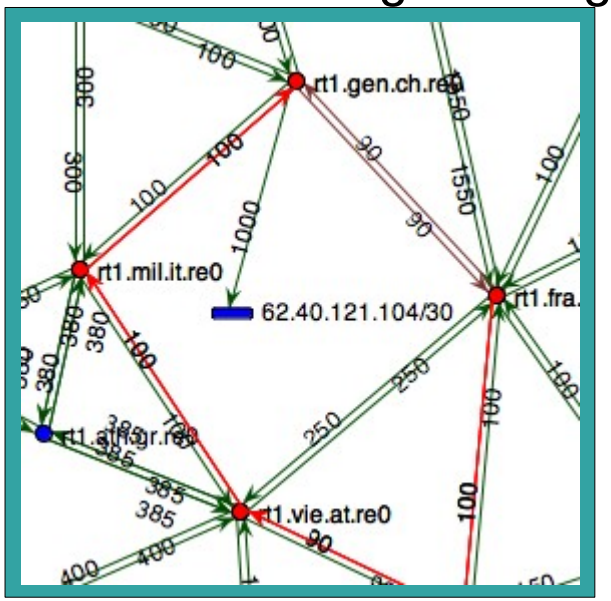
3). Trace route

After failure



What-if scenarios

Before IGP weight change

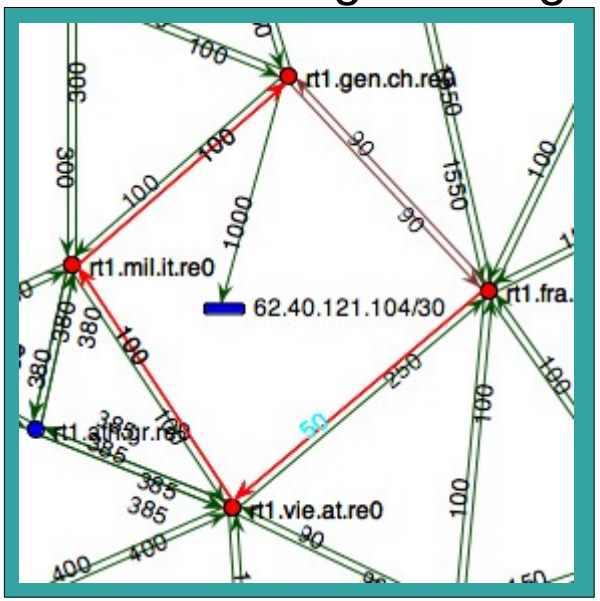


1). Change **IGP weight** of link between Frankfurt and Vienna

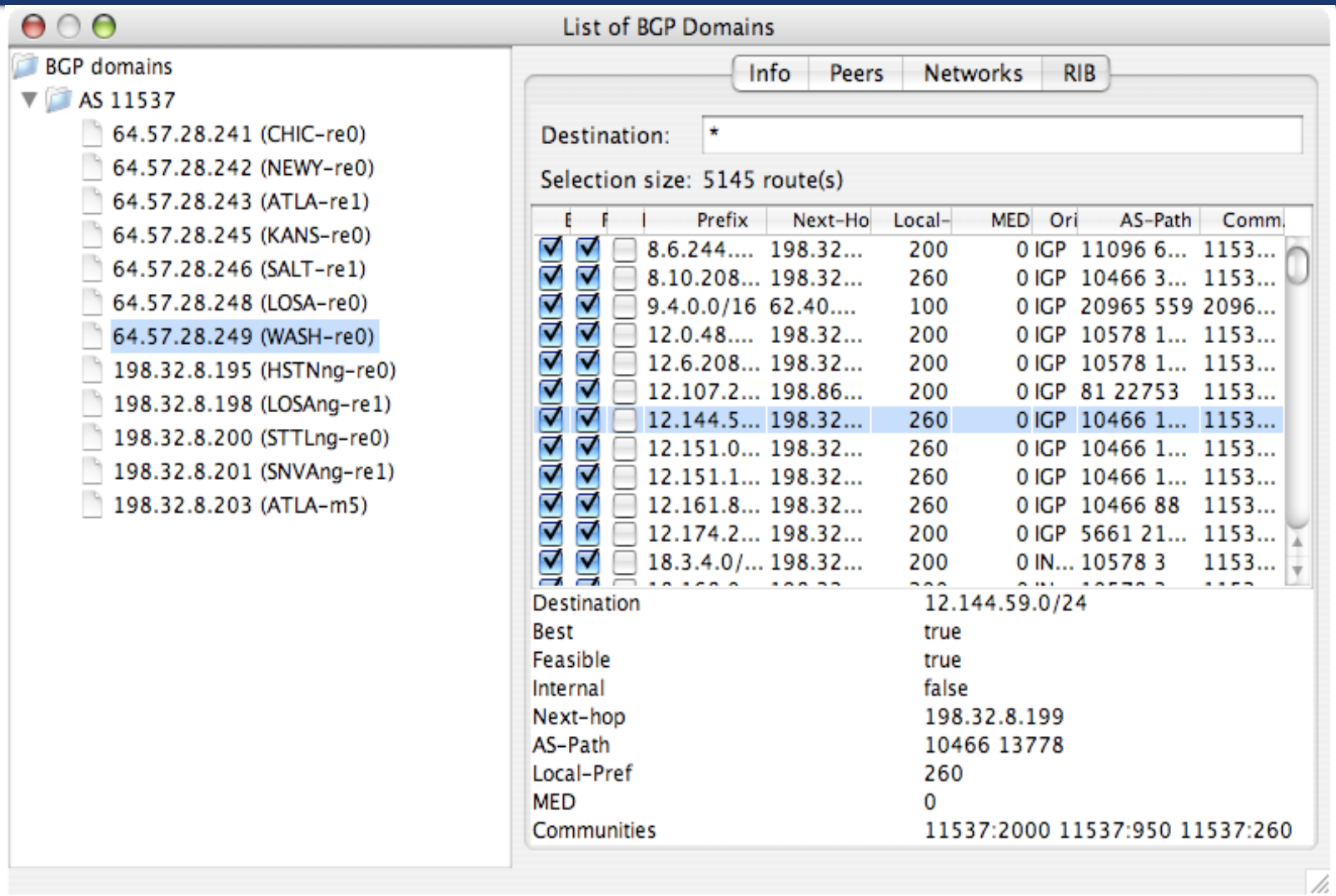
2). Recompute routes

3). Trace route

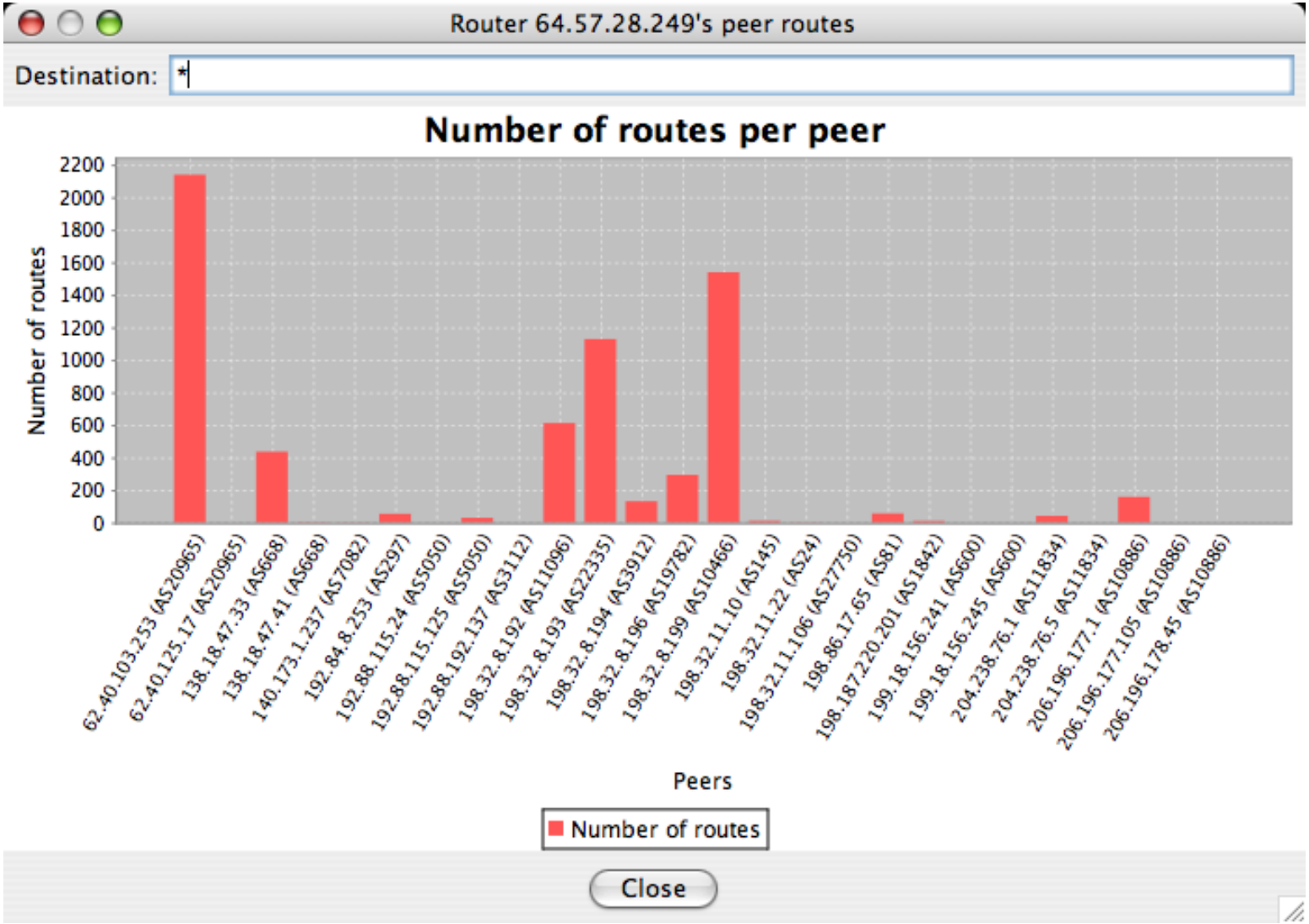
After IGP weight change



Routing State Queries



Routing State Queries



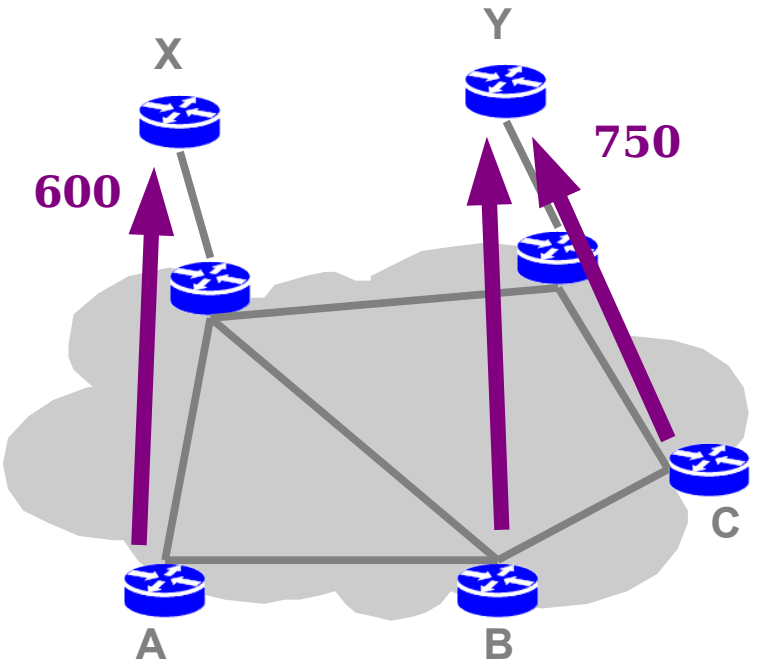
III. Case study

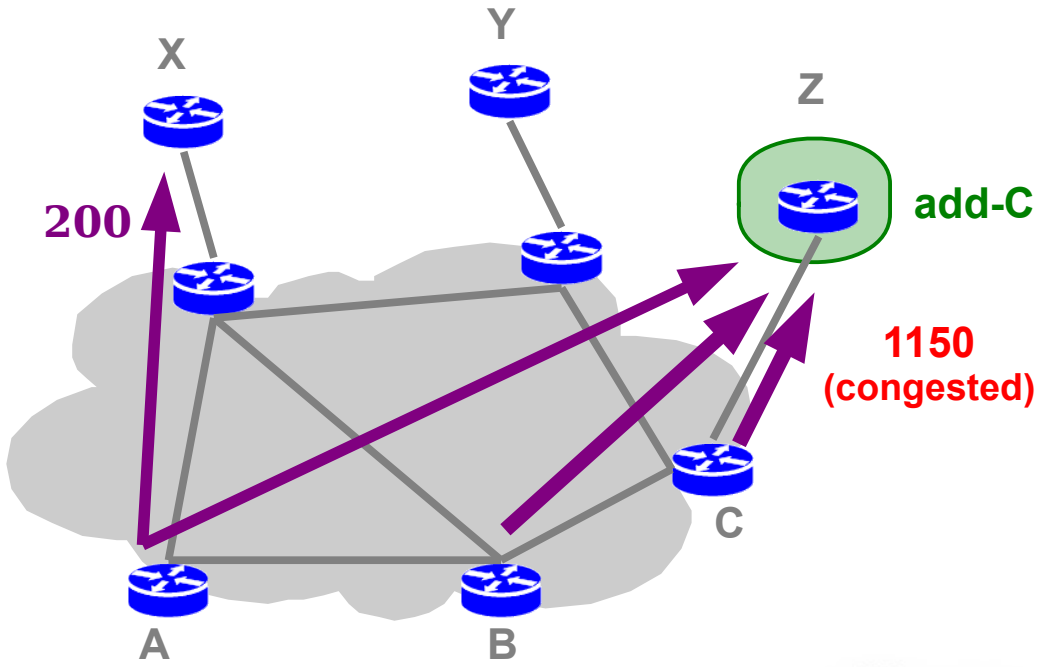
"Anonymous French Tier-1"

Peering placement

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

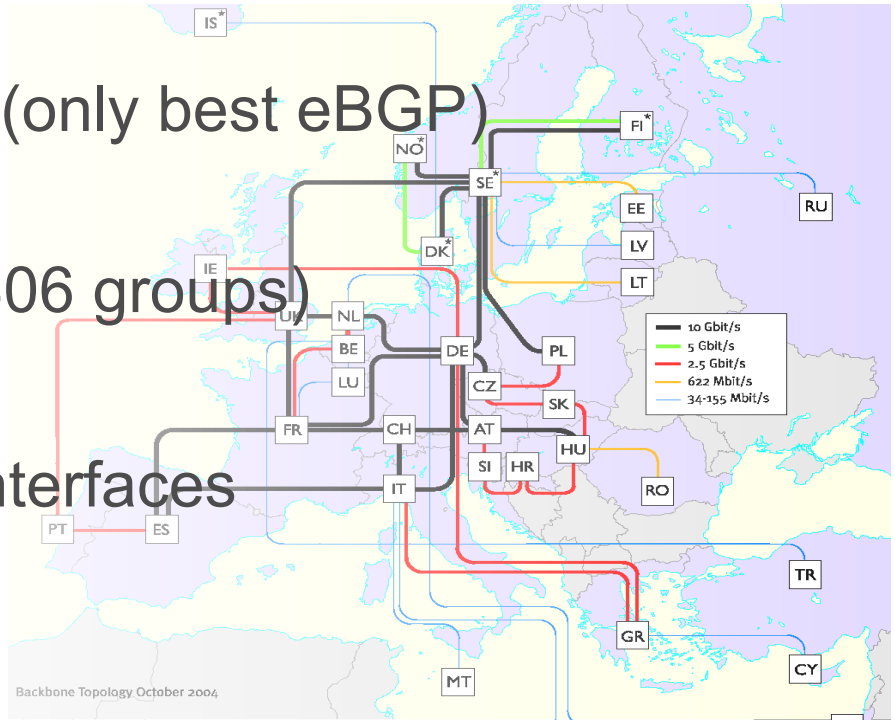
	X	Y
A	600	0
B	0	250
C	0	500



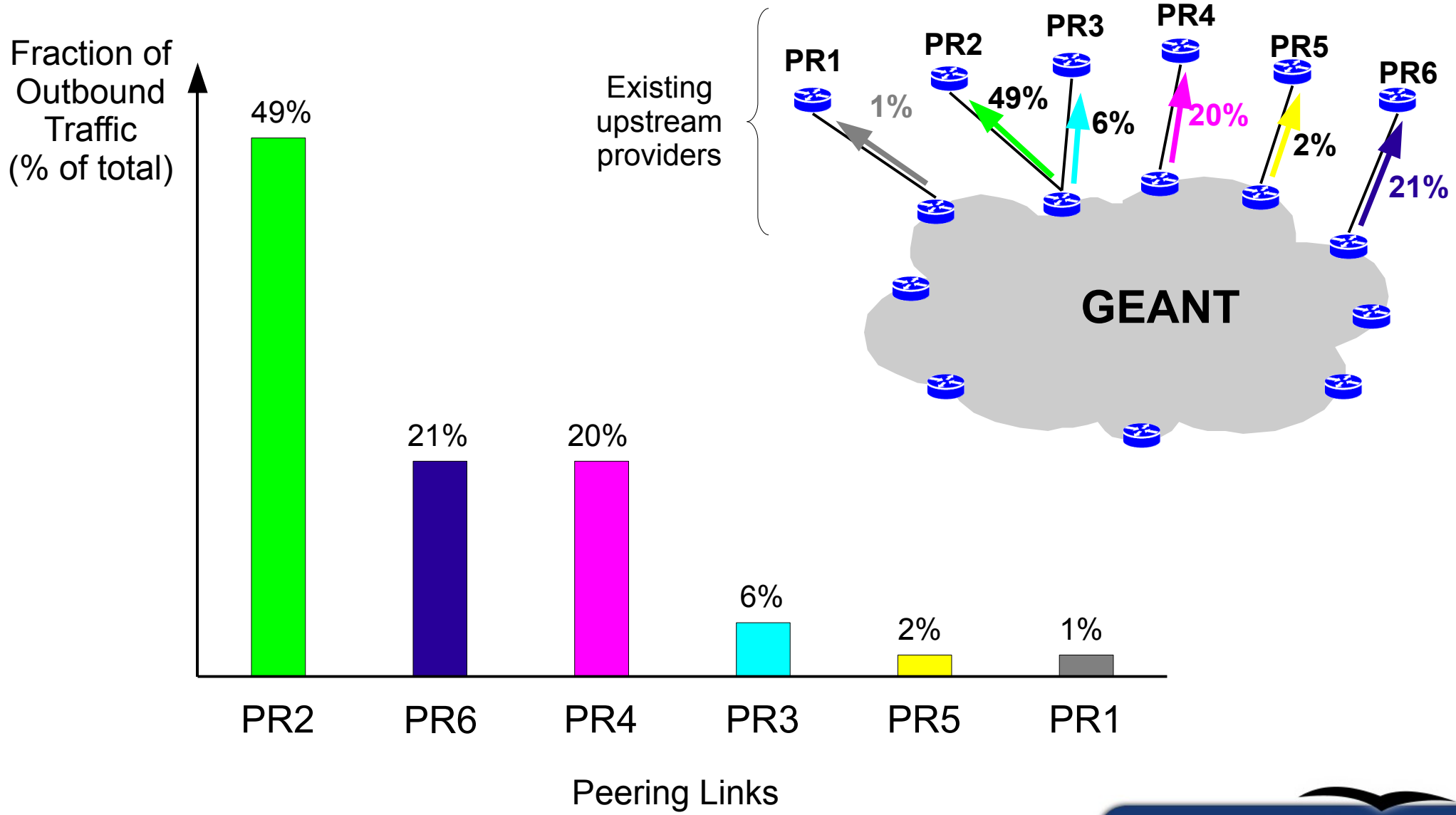


Case study: GEANT (AS20965)

- **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



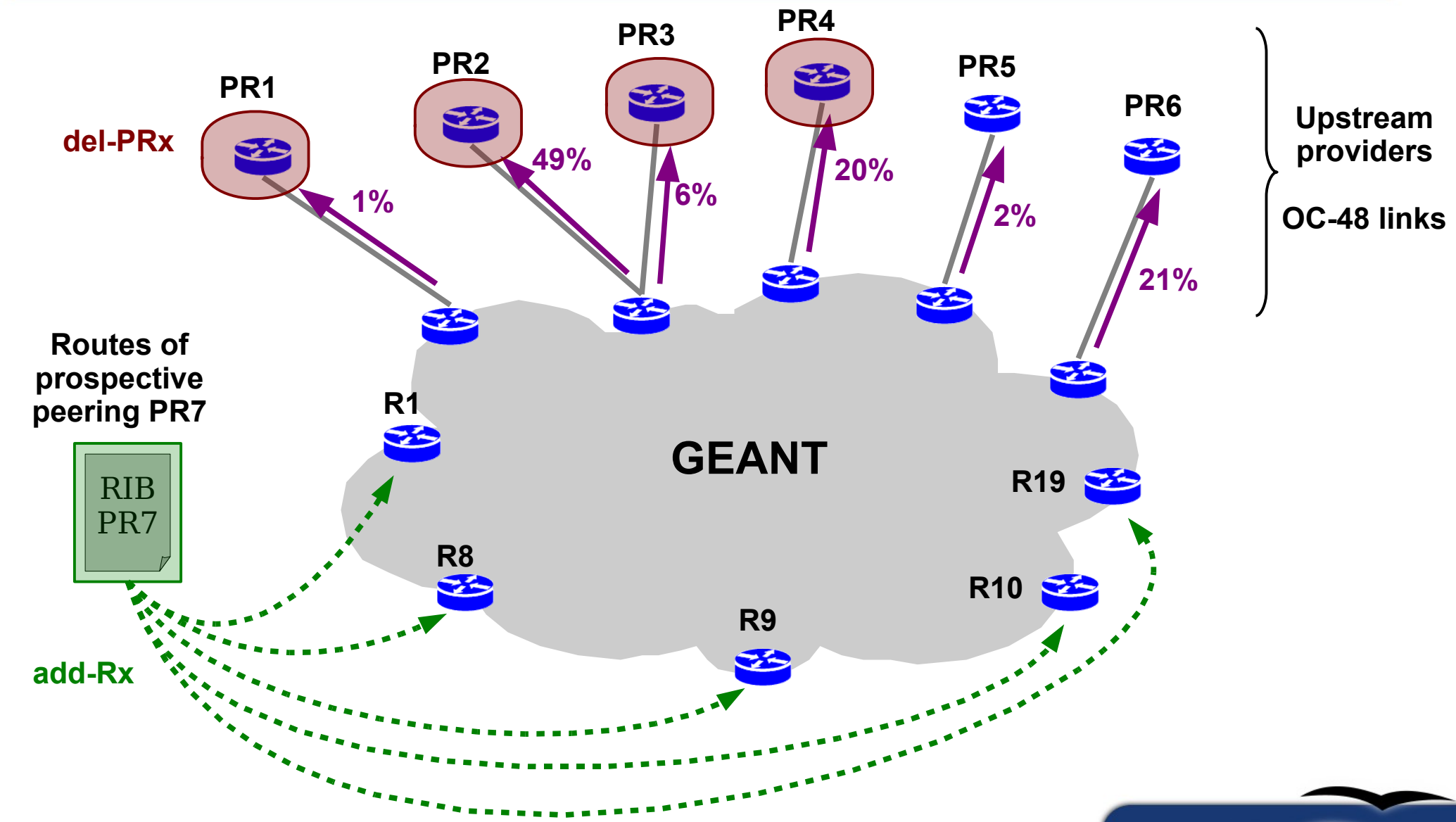
Peering placement



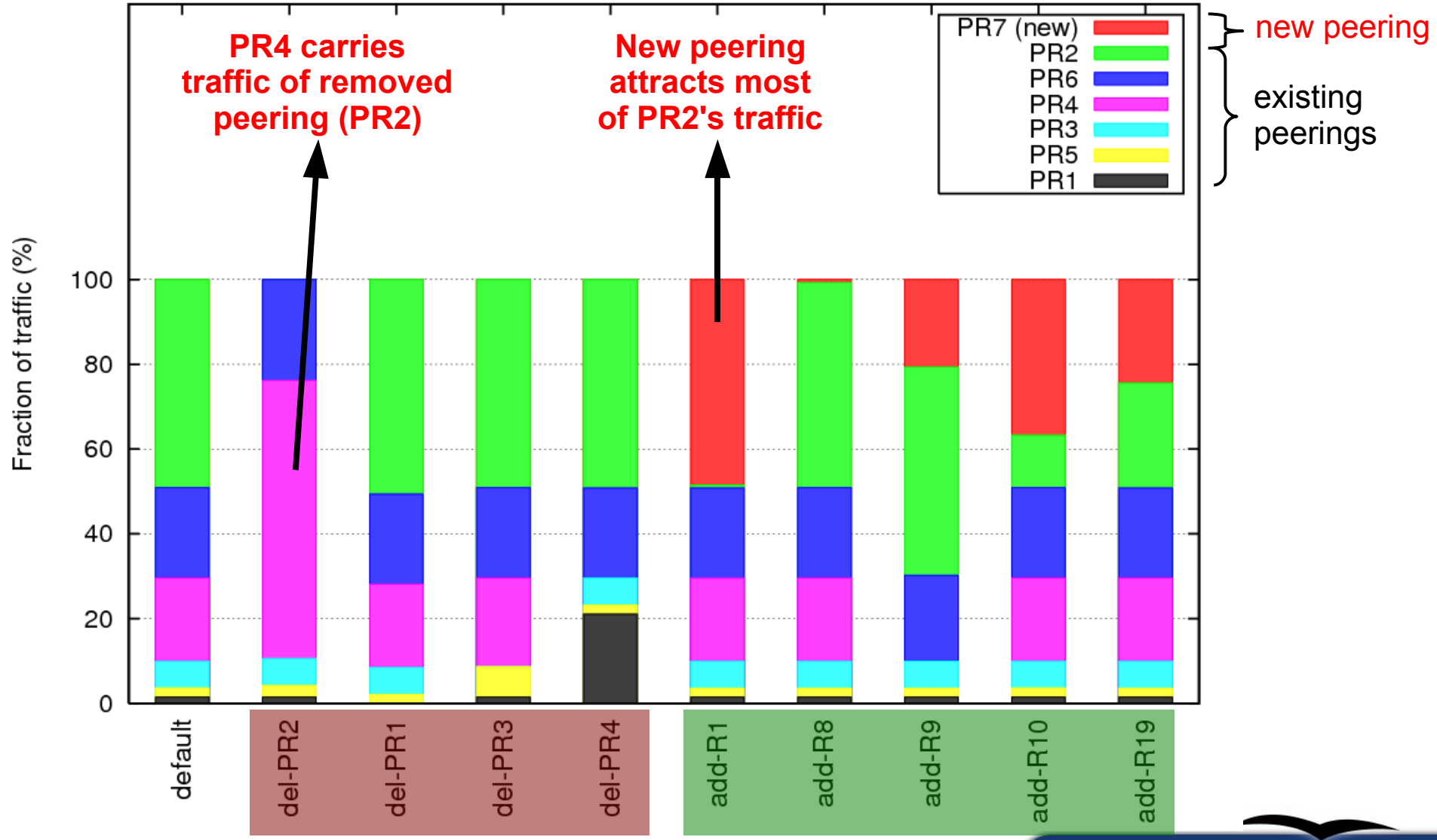
Peering placement

- **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)

Peering placement



Peering placement



Conclusion

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



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