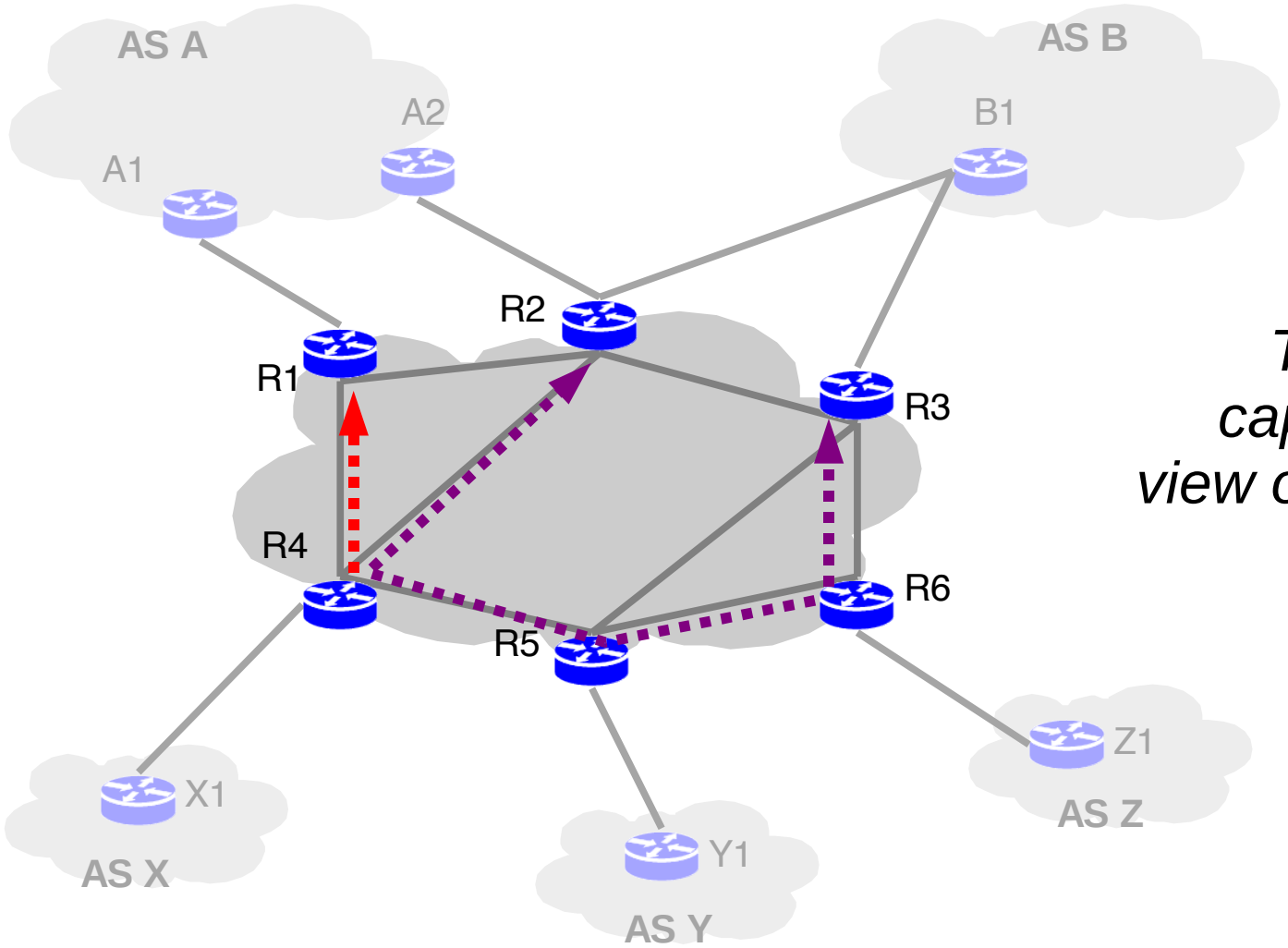
Modeling the Routing of an ISP

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...

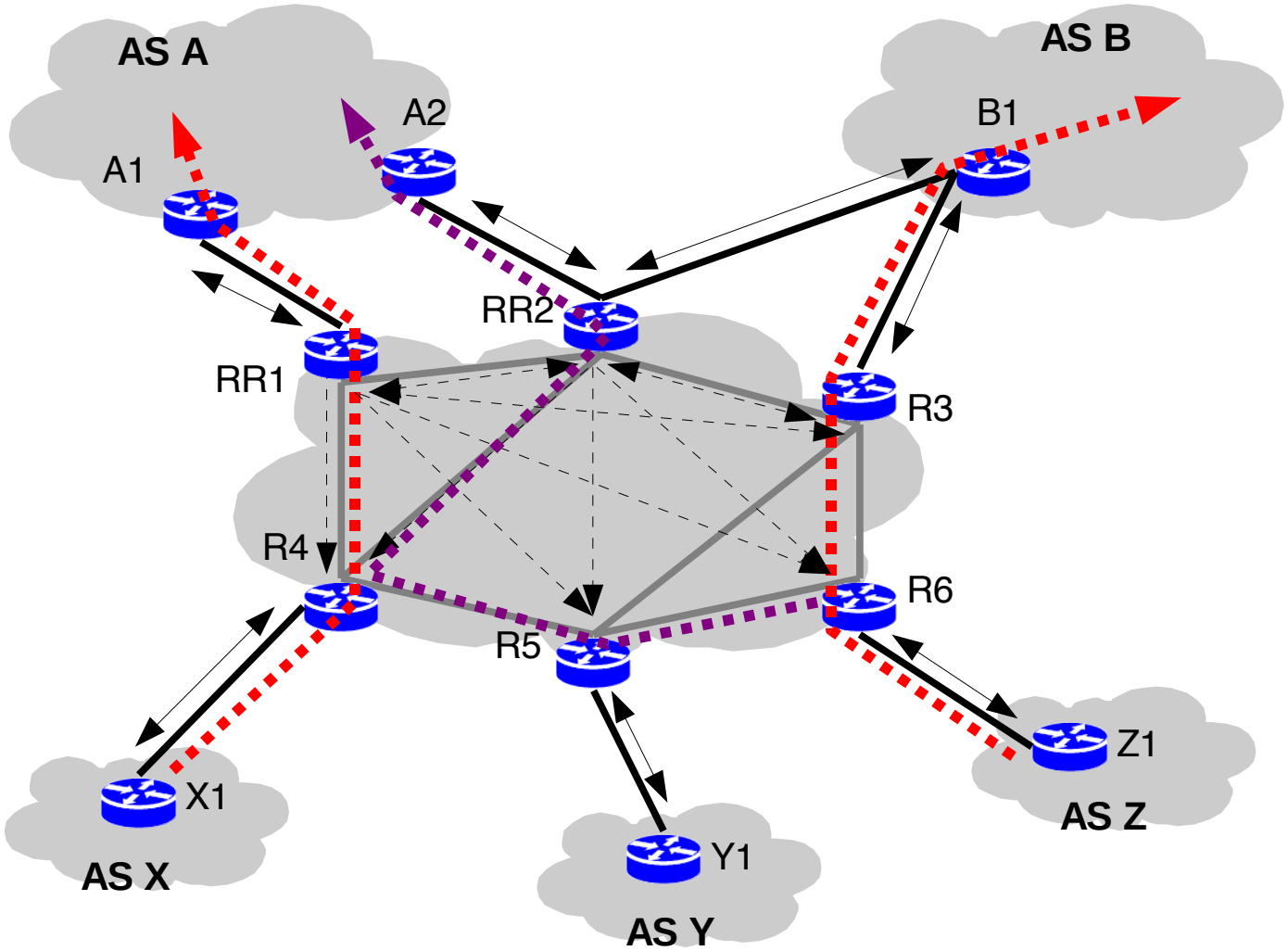
What would happen to my interdomain traffic if...

- 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
- ...

Internal link
External link
eBGP session
iBGP session
Client session
Traffic flow

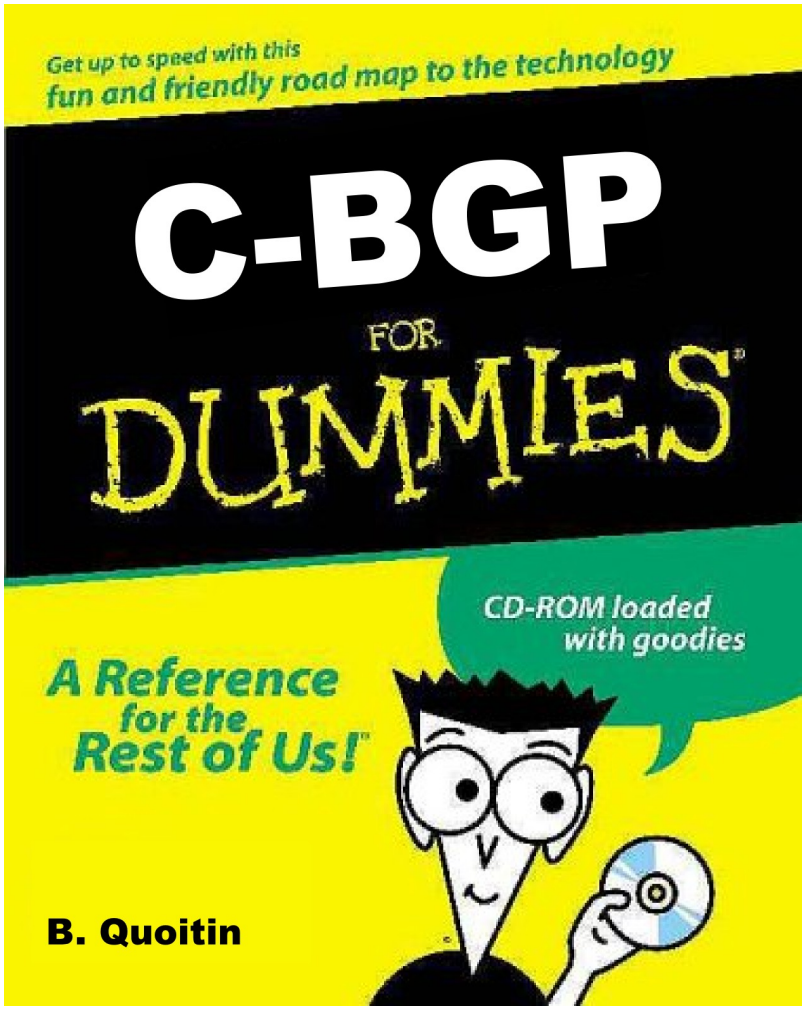
Agenda

- **C-BGP:**
  - **a network topology / config / routing DB**
  - **a BGP routing solver**

- **Case studies**
  - **Scenario 1: peering placement**
  - **Scenario 2: all single-link failures**

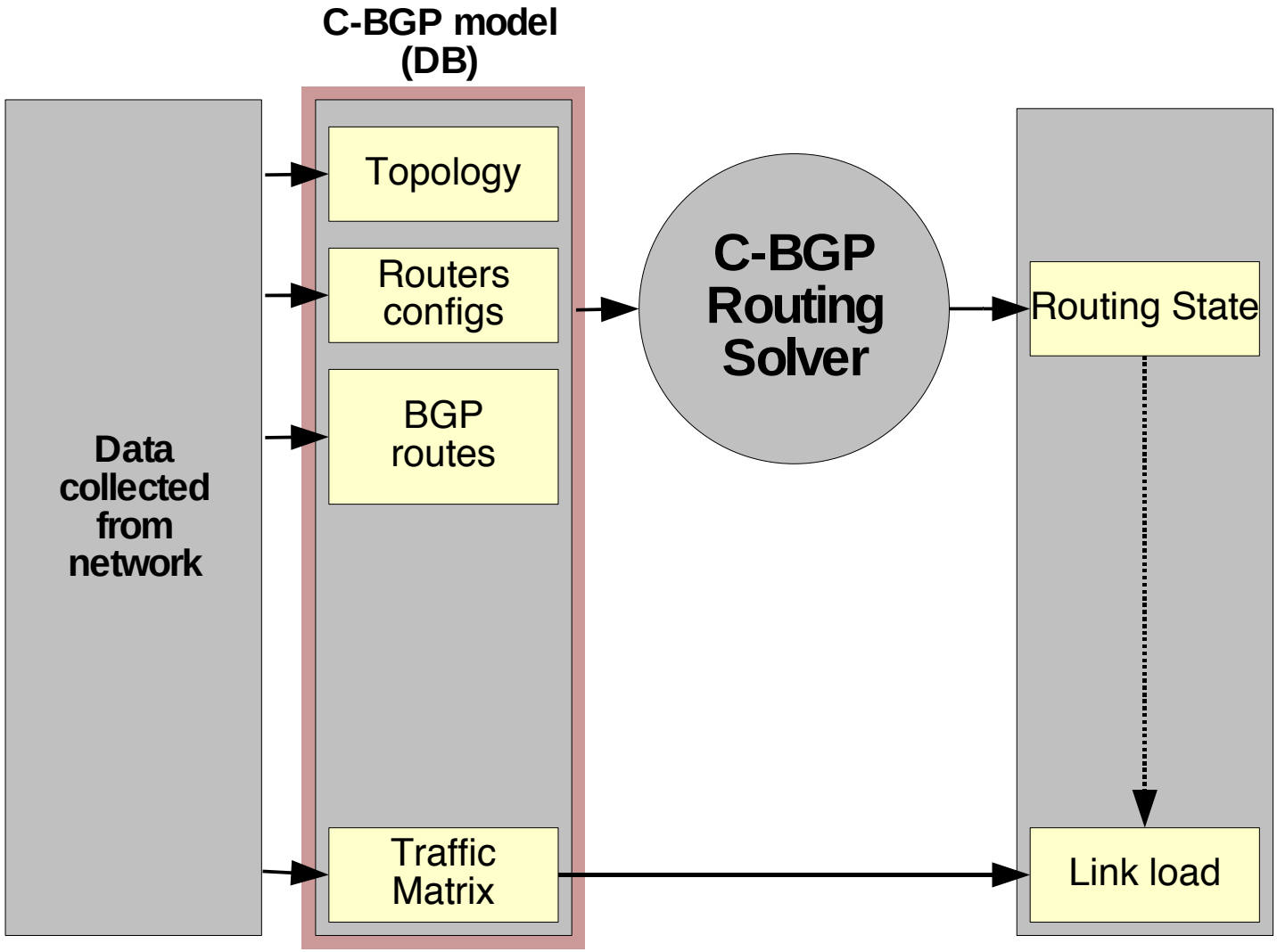
I. C-BGP

C-BGP



- Network topology / configuration DB
- Scriptable BGP Routing solver for large-scale networks
- 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

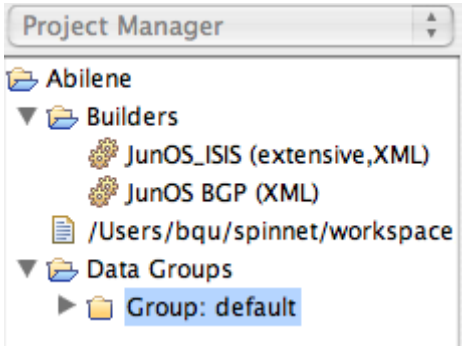
DB: importing...

Build project from real router/network data

"show isis database extensive"

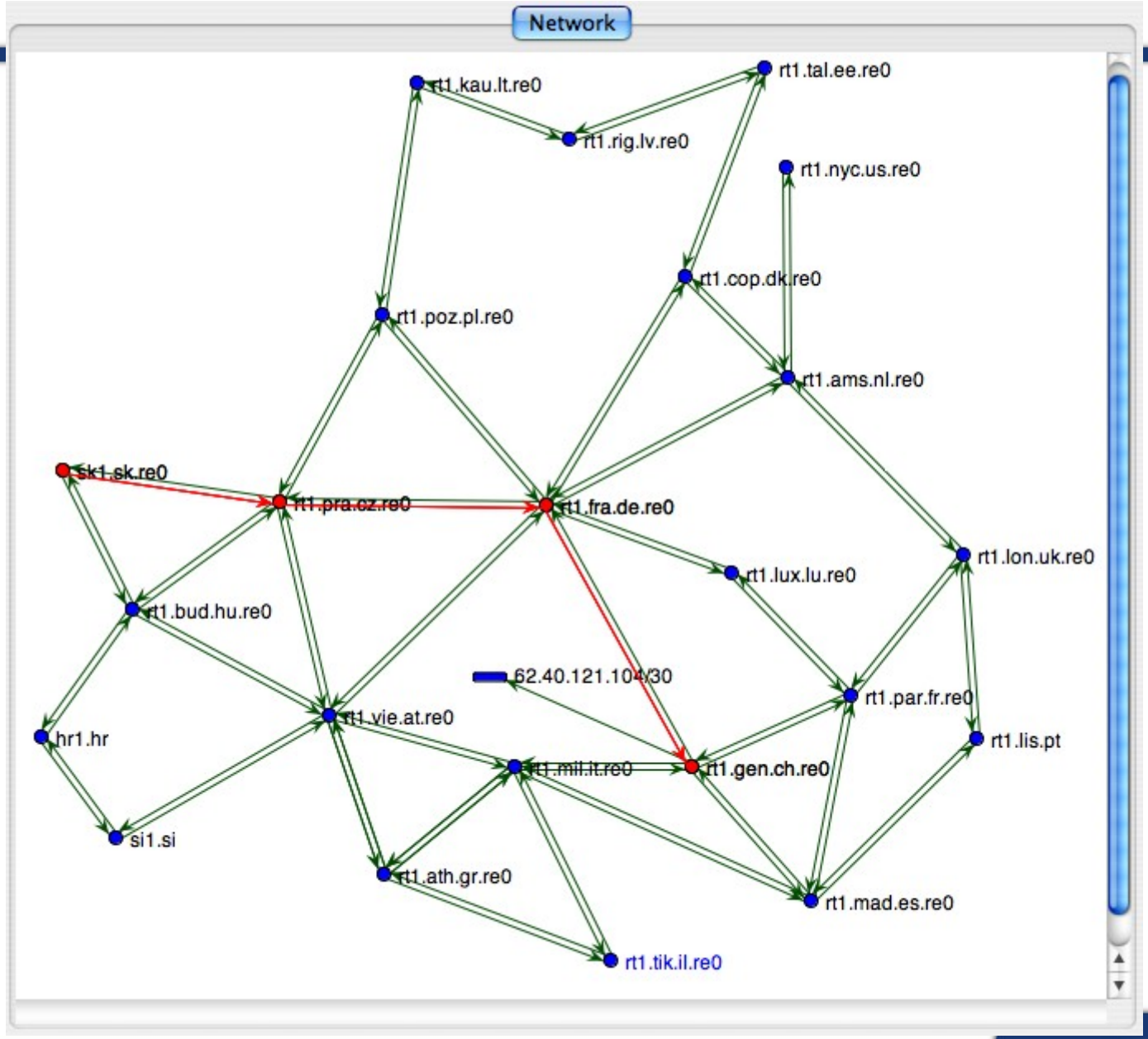
"show bgp neighbors"

- 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)

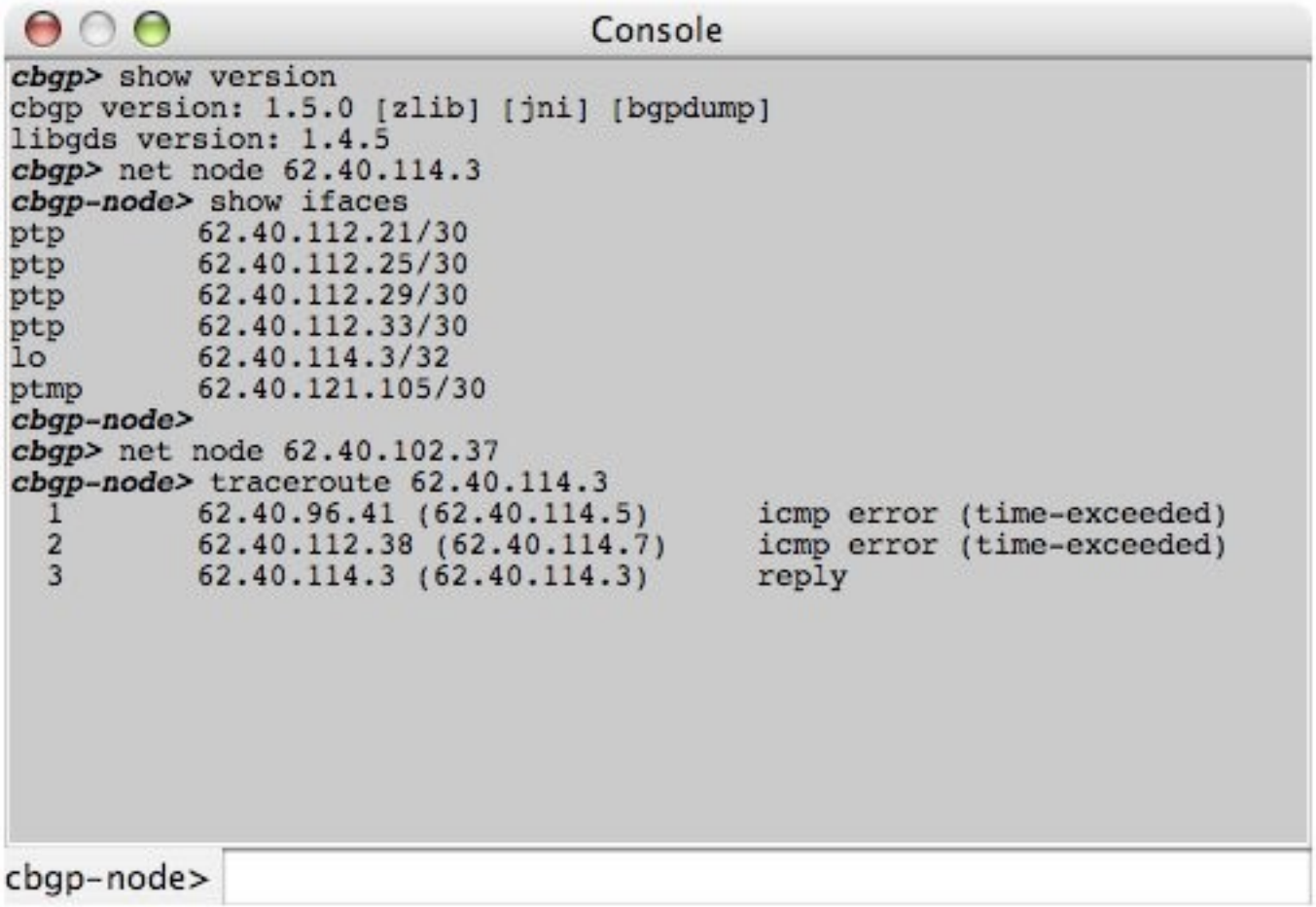


(1) other vendor formats can be easily added

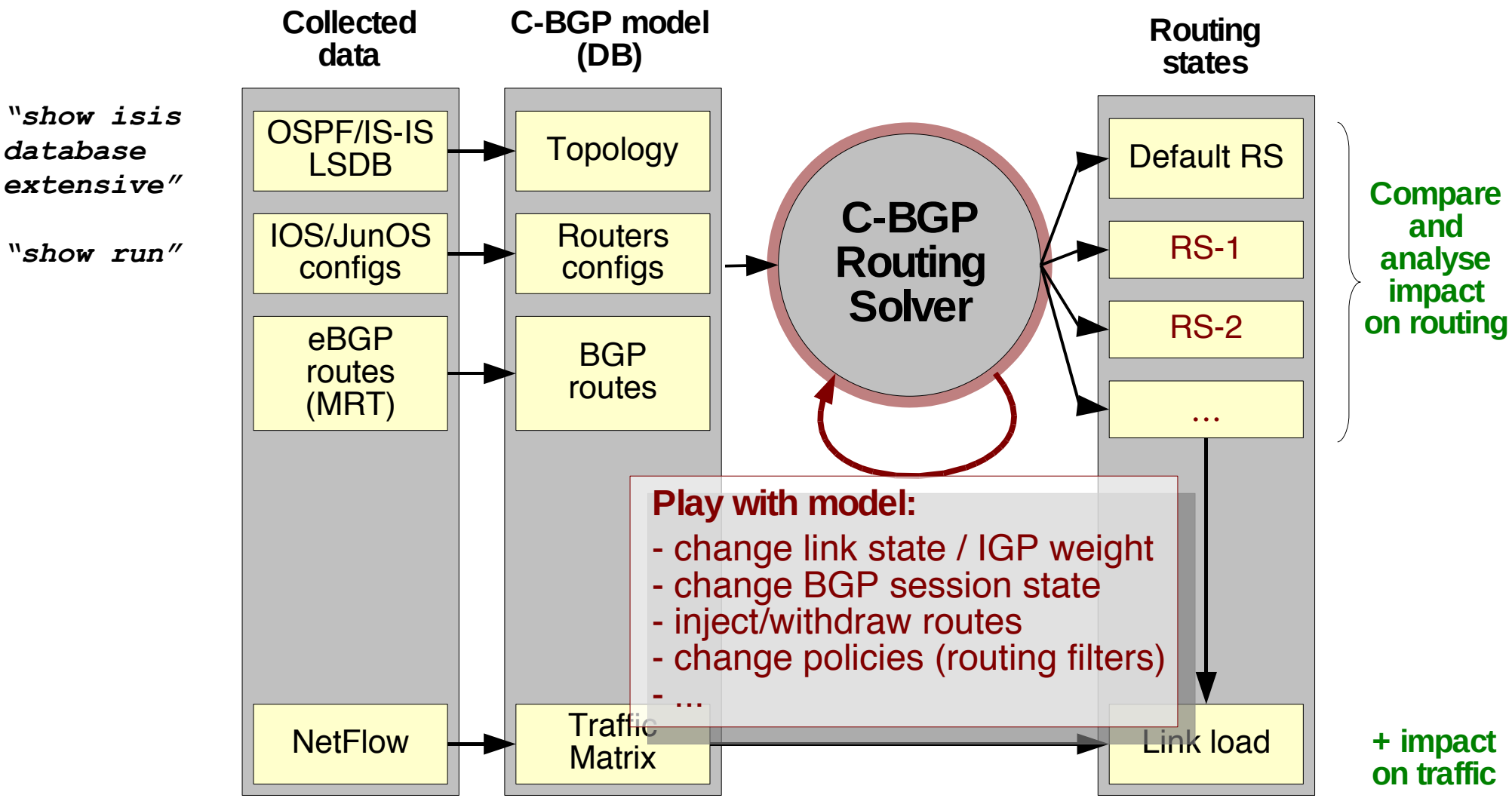
DB: network visualization



DB: network queries (CLI)



C-BGP Routing Solver



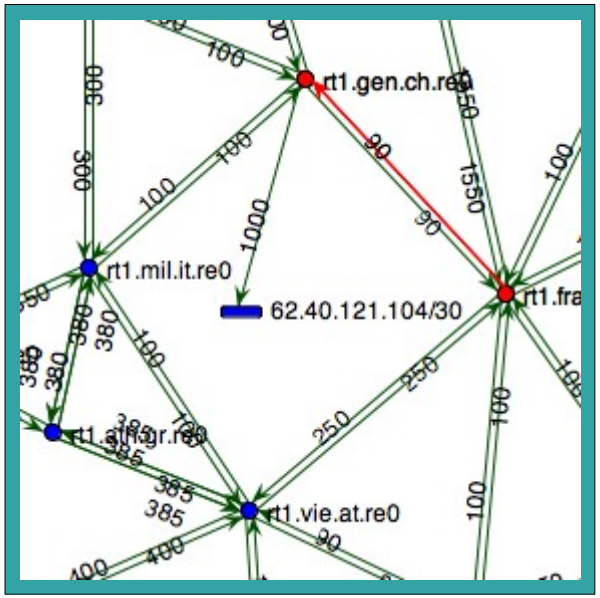
Routing Solver

- **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

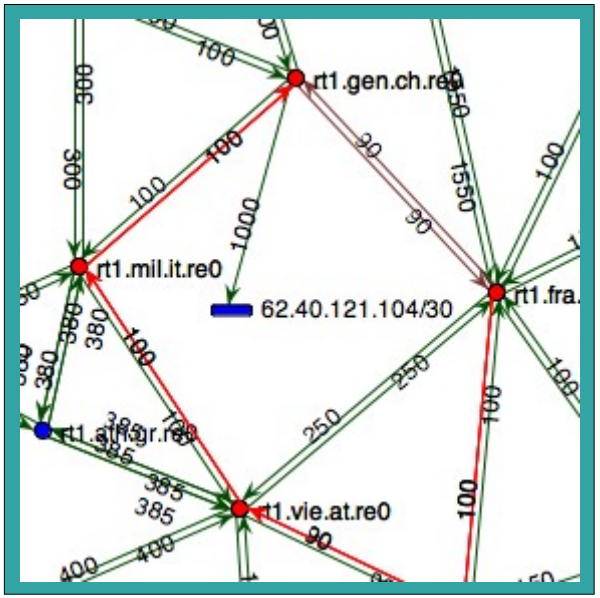


1). **Fail link** between Frankfurt and Geneva

2). Recompute routes

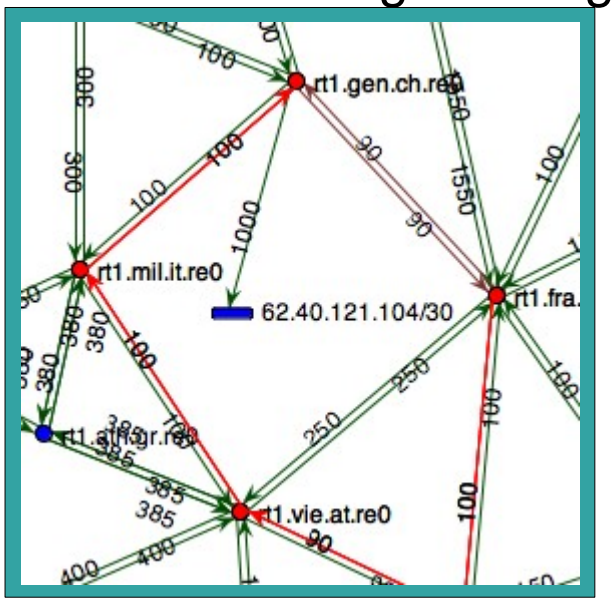
3). Trace route

After failure



What-if scenarios

Before IGP weight change

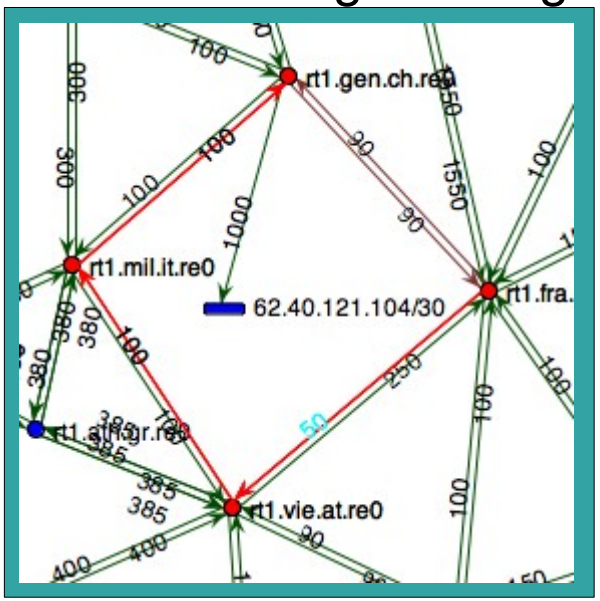


1). Change **IGP weight** of link between Frankfurt and Vienna

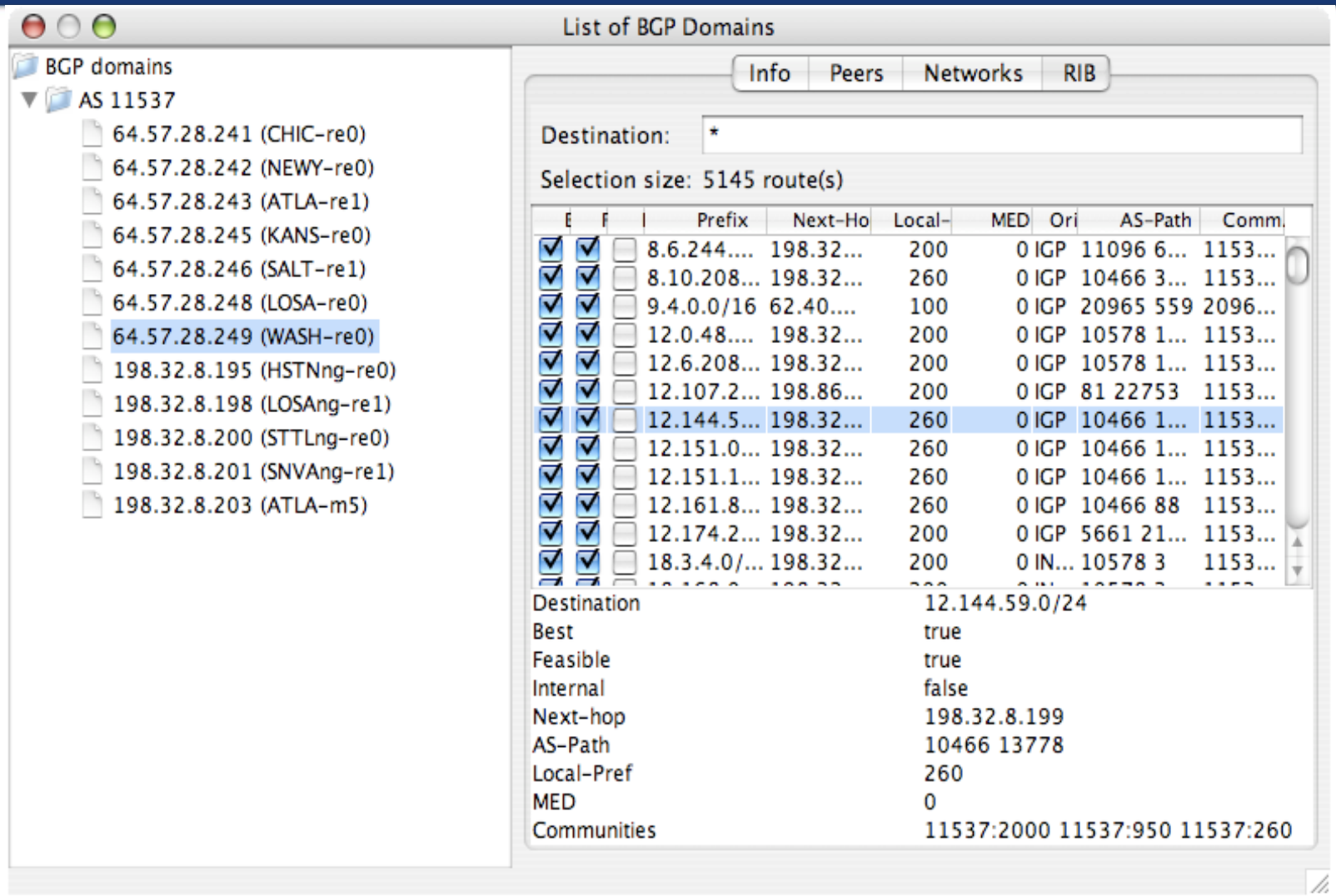
2). Recompute routes

3). Trace route

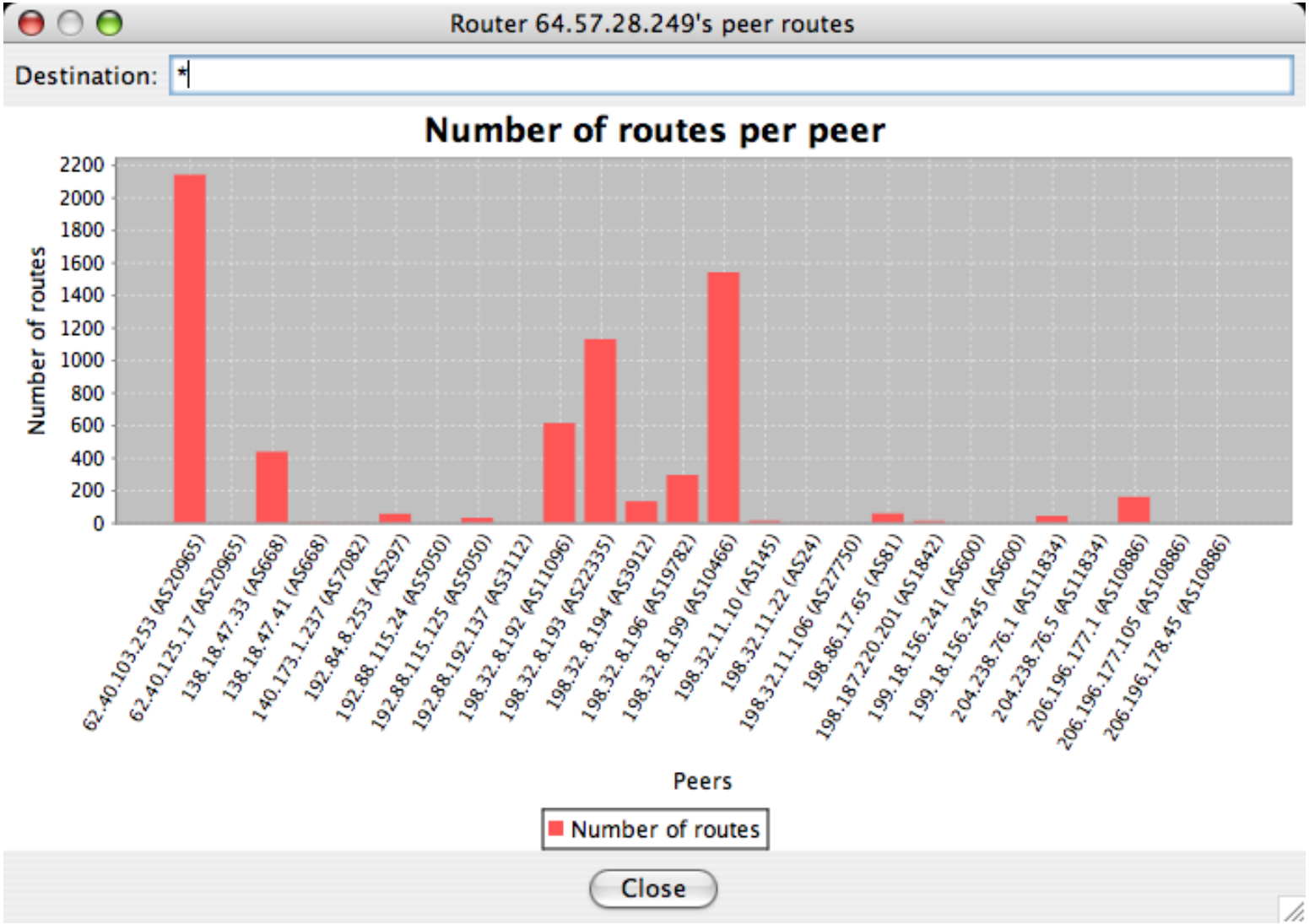
After IGP weight change



Routing State Queries



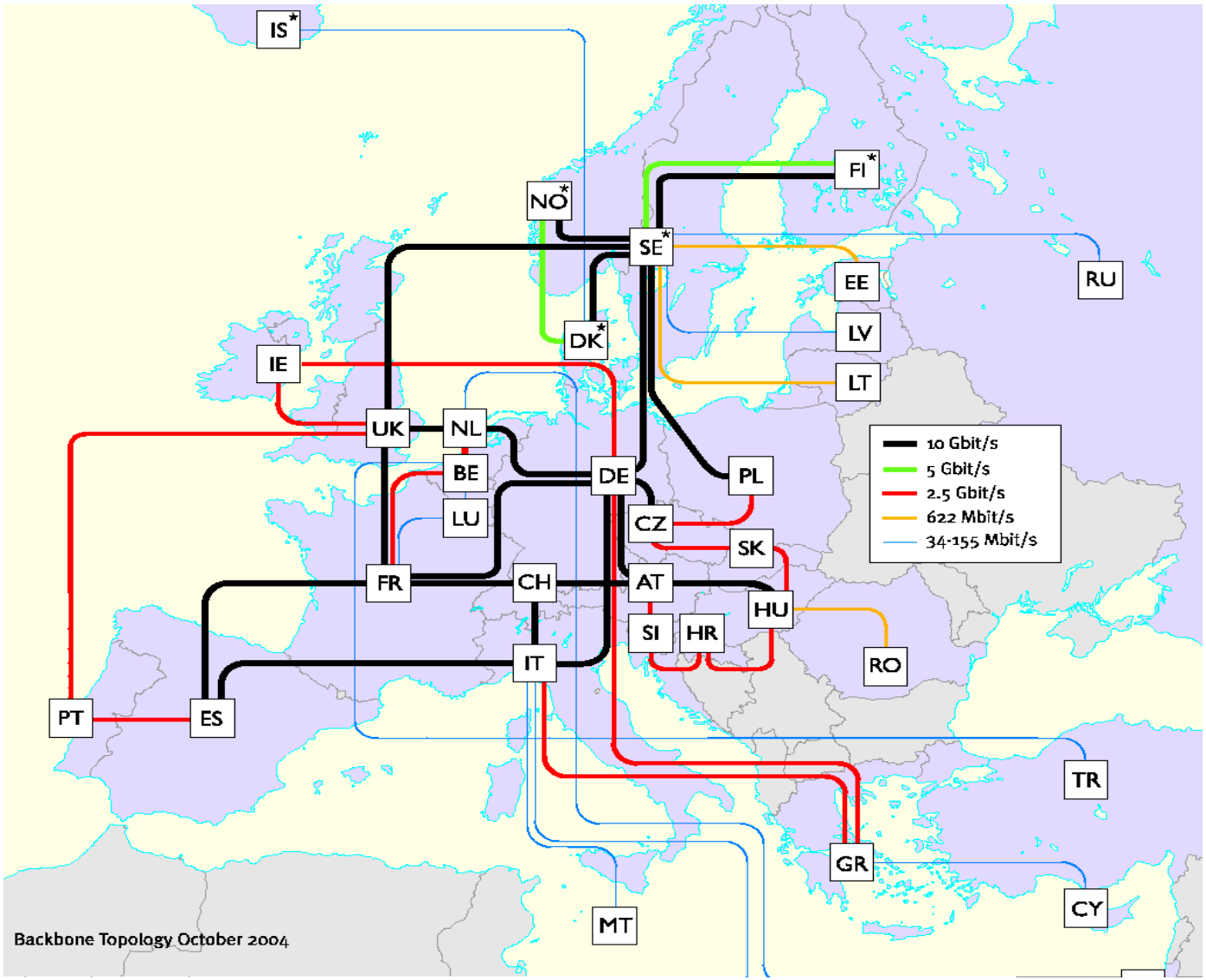
Routing State Queries



II. Case studies

Anonymous
French Tier-1

Case study: GEANT (AS20965)



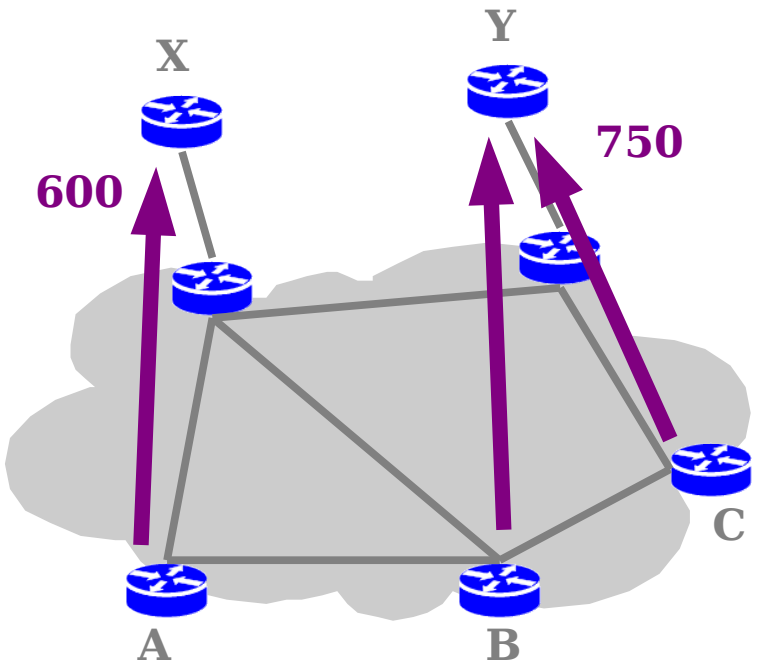
Case study: GEANT (AS20965)

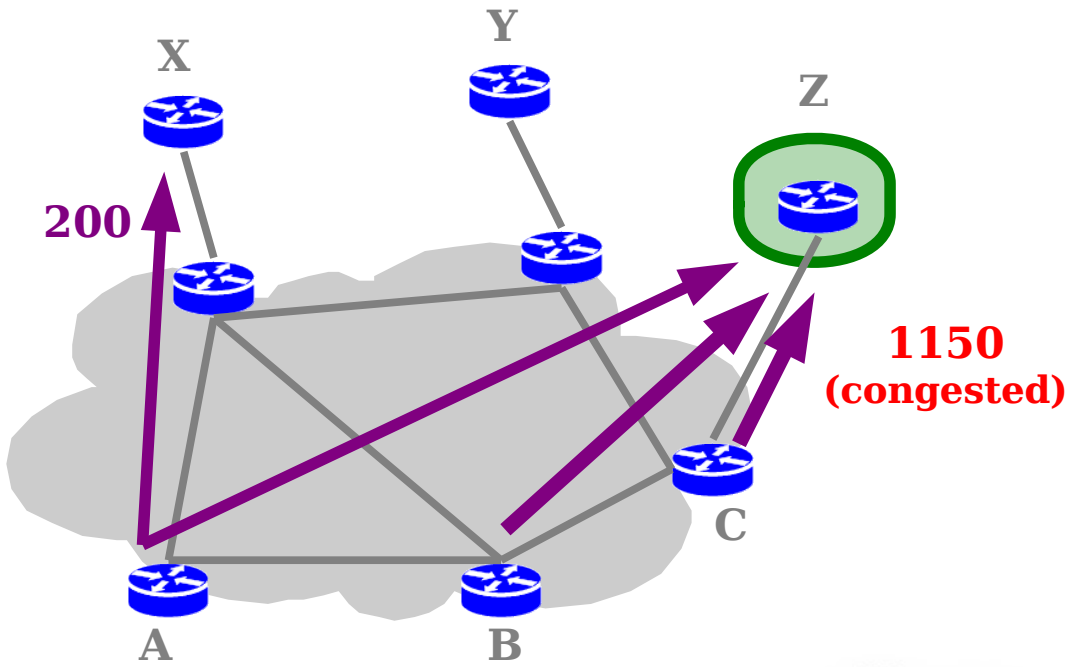
- **Topology**
  - Obtained from IS-IS trace, cross-checked with map
  - 23 nodes, 38 core links, 53 edge links (6 with upstreams)
- **Routing data**
  - Collected using Zebra in the iBGP (only best eBGP)
  - 640,897 eBGP routes
    - 150,071 prefixes (clustered in 406 groups)
- **Traffic data**
  - 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

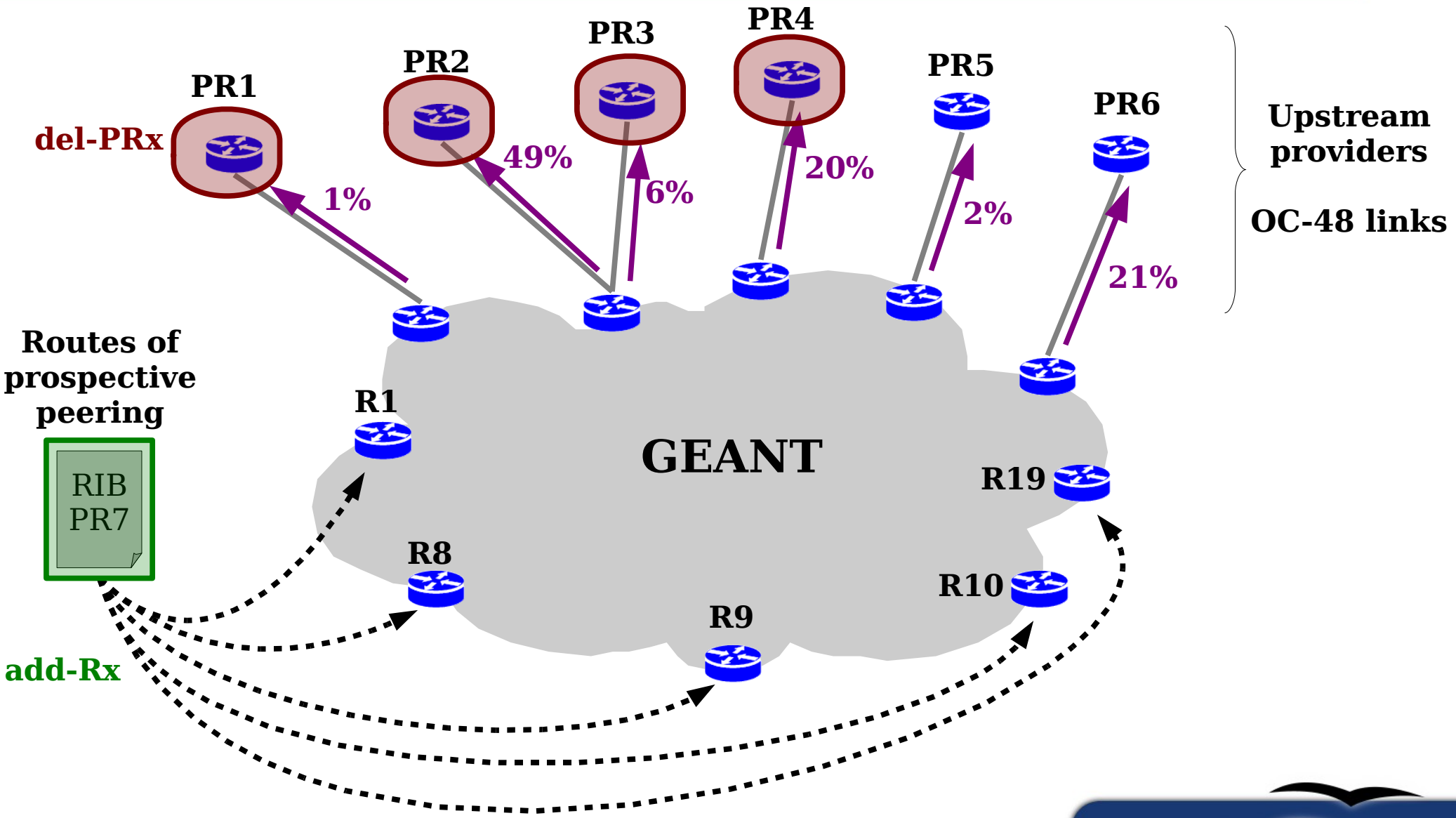




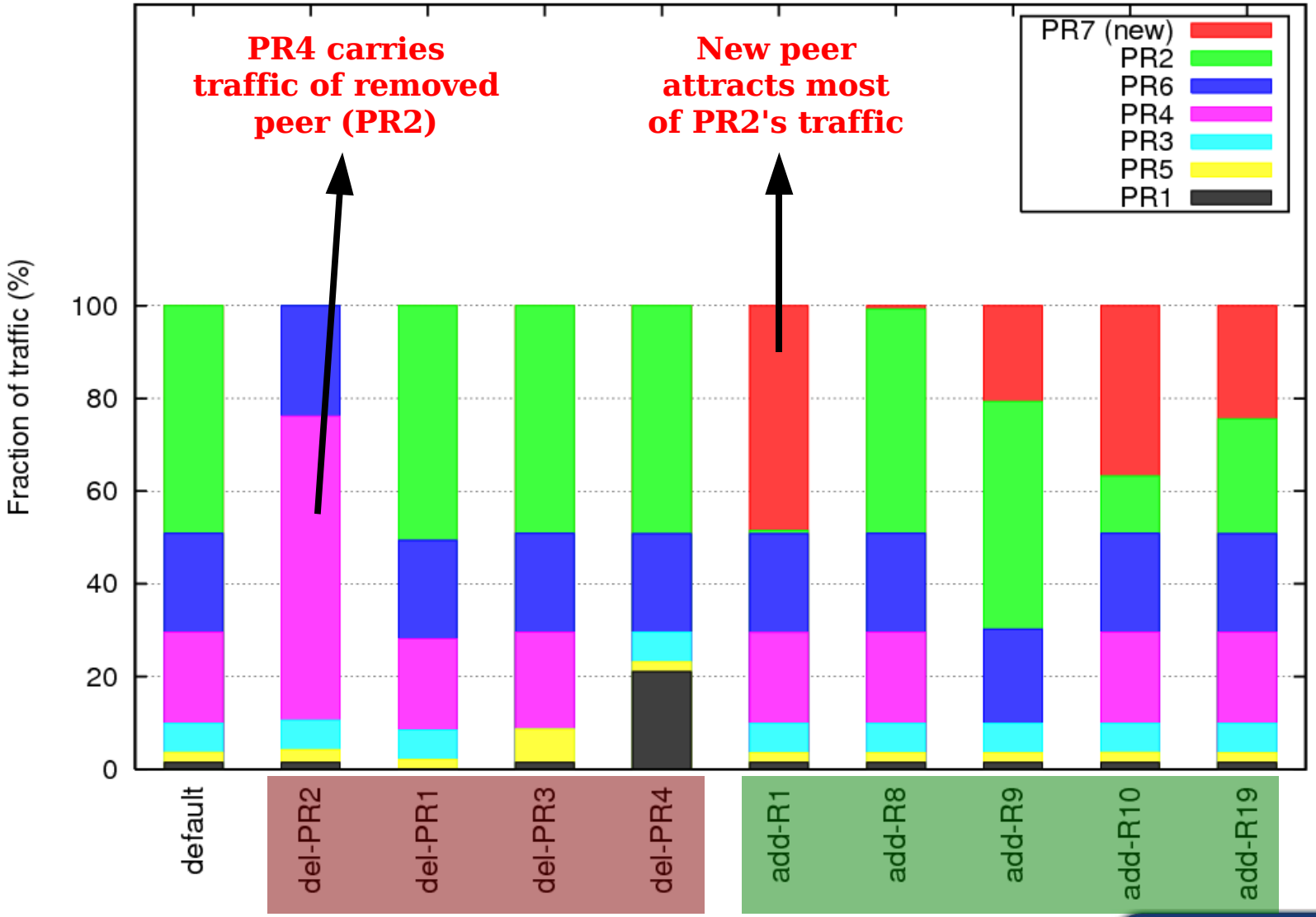
1st Scenario: peering placement

- **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



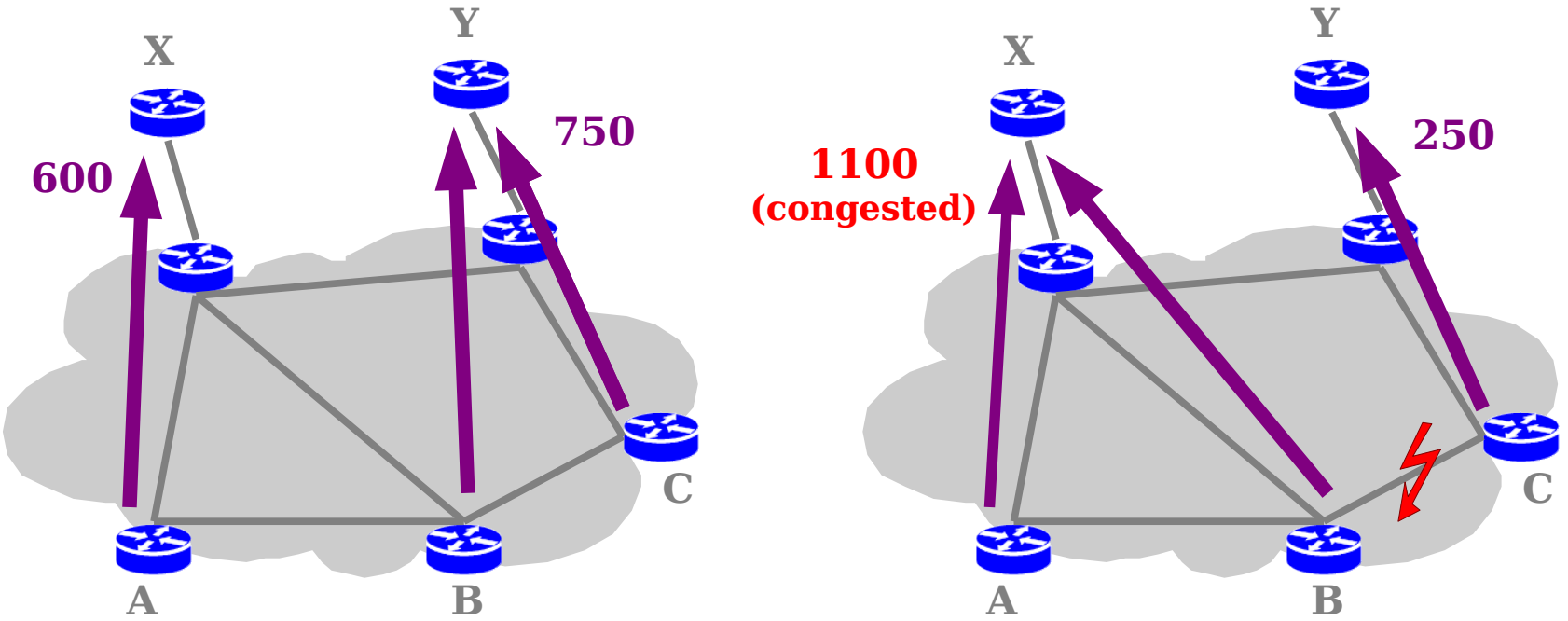
1st Scenario: peering placement



2nd Scenario: link failures

- **Example**
  - Traffic to upstream X and Y
  - 1 Gbps links
  - Internal link failure: B ↔ C

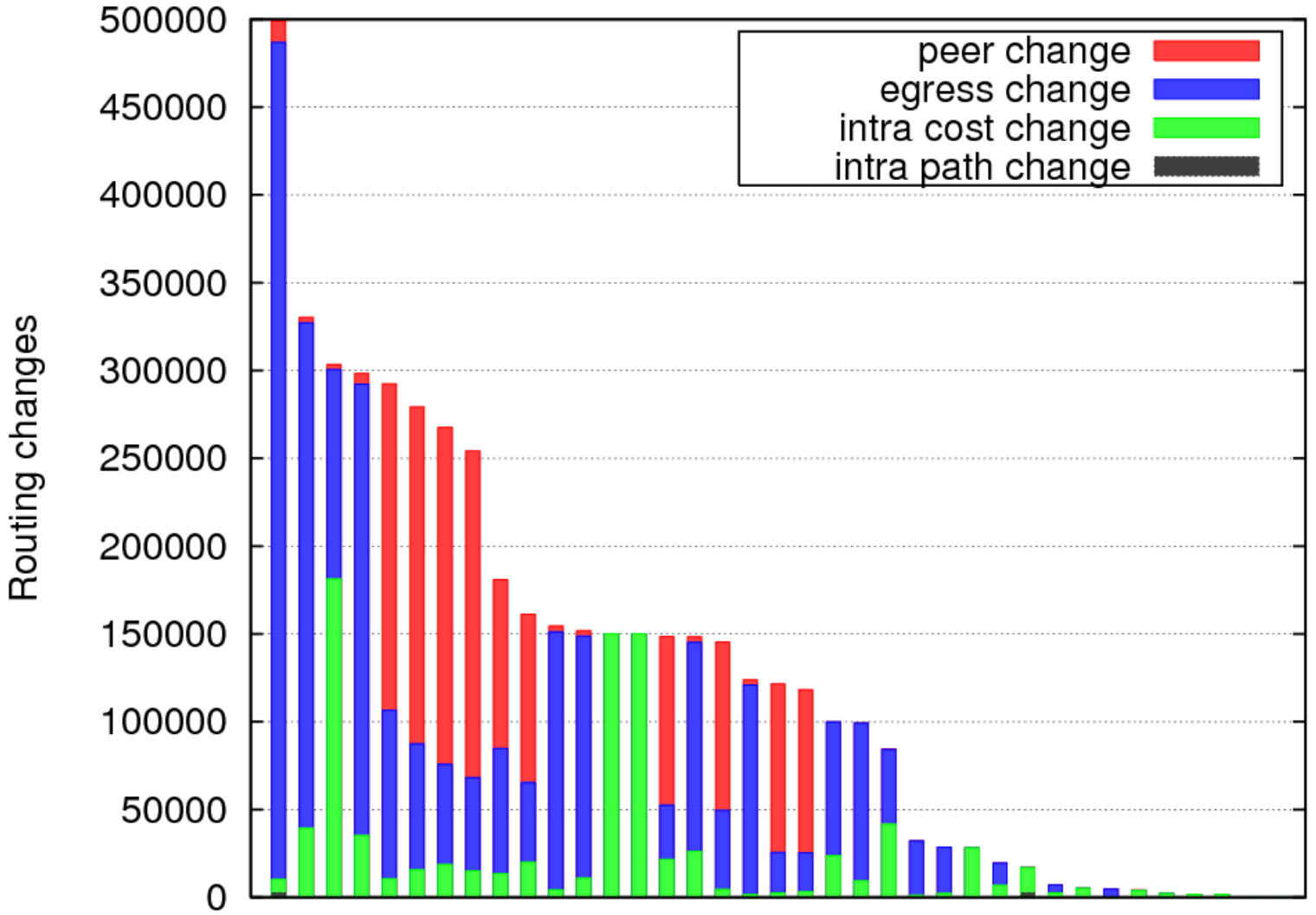
	X	Y
A	600	0
B	0	500
C	0	250



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:



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