Modeling the Routing of an ISP

Bruno Quoitin
(bruno.quoitin@uclouvain.be)
Computer Science & Engineering Department
Université catholique de Louvain, Belgium

This is a joint work with
Sebastien Tandel, **Steve Uhlig**
and **Olivier Bonaventure**

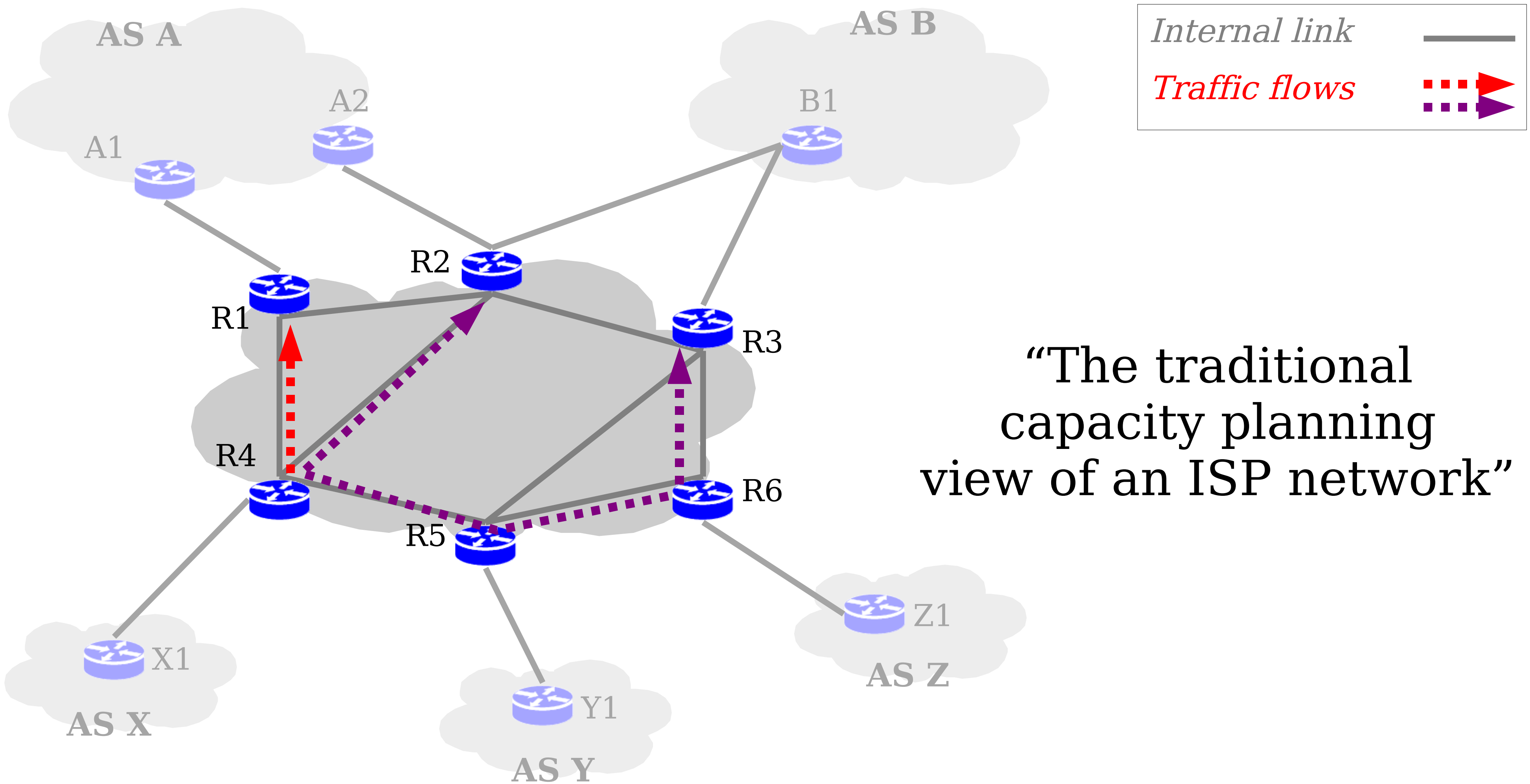
Agenda

- **Motivation**
- **C-BGP: a BGP routing solver**
- **Case study**
  - **Scenario 1: peering placement**
  - **Scenario 2: all single-link failures**
- **Conclusion**

Objectives

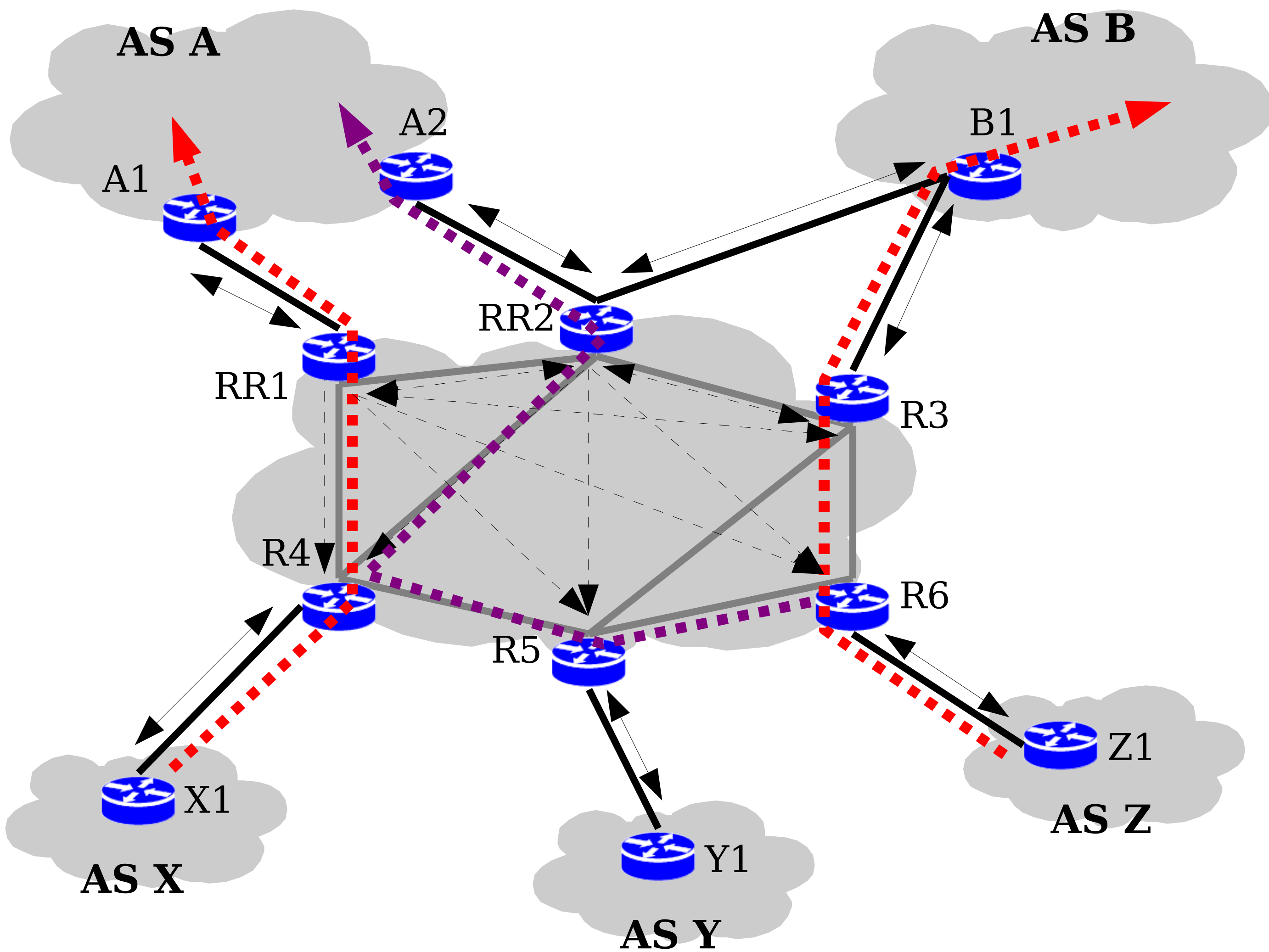
- **Why modeling ?**
  - **What would happen to your interdomain traffic if...**
    - a link is failing ?
    - a router is under maintenance ?
    - a BGP peering is being shutdown ?
    - a new route filtering policy is planned ?
    - a new peering is established at an IXP ?
  - **How would you optimize your interdomain routing for...**
    - performance ?
    - cost ?
    - reliability ?

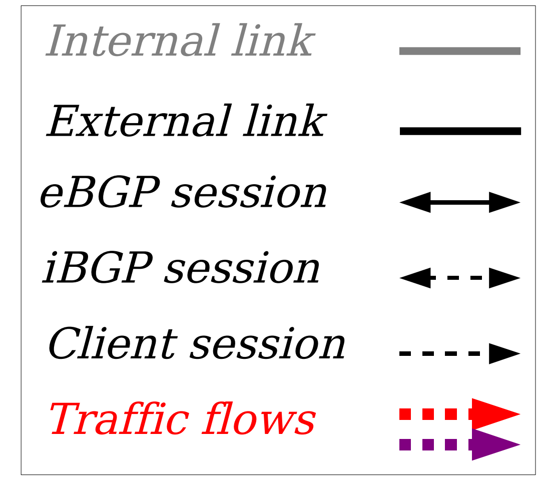
ISP Model



"The traditional capacity planning view of an ISP network"

ISP Model





Reality has:
- Transit traffic
- Multiple egresses
- iBGP topology
- Route-reflectors
- Routing policies
- 215,000 destinations (and counting)
- IGP/BGP interaction
- ...

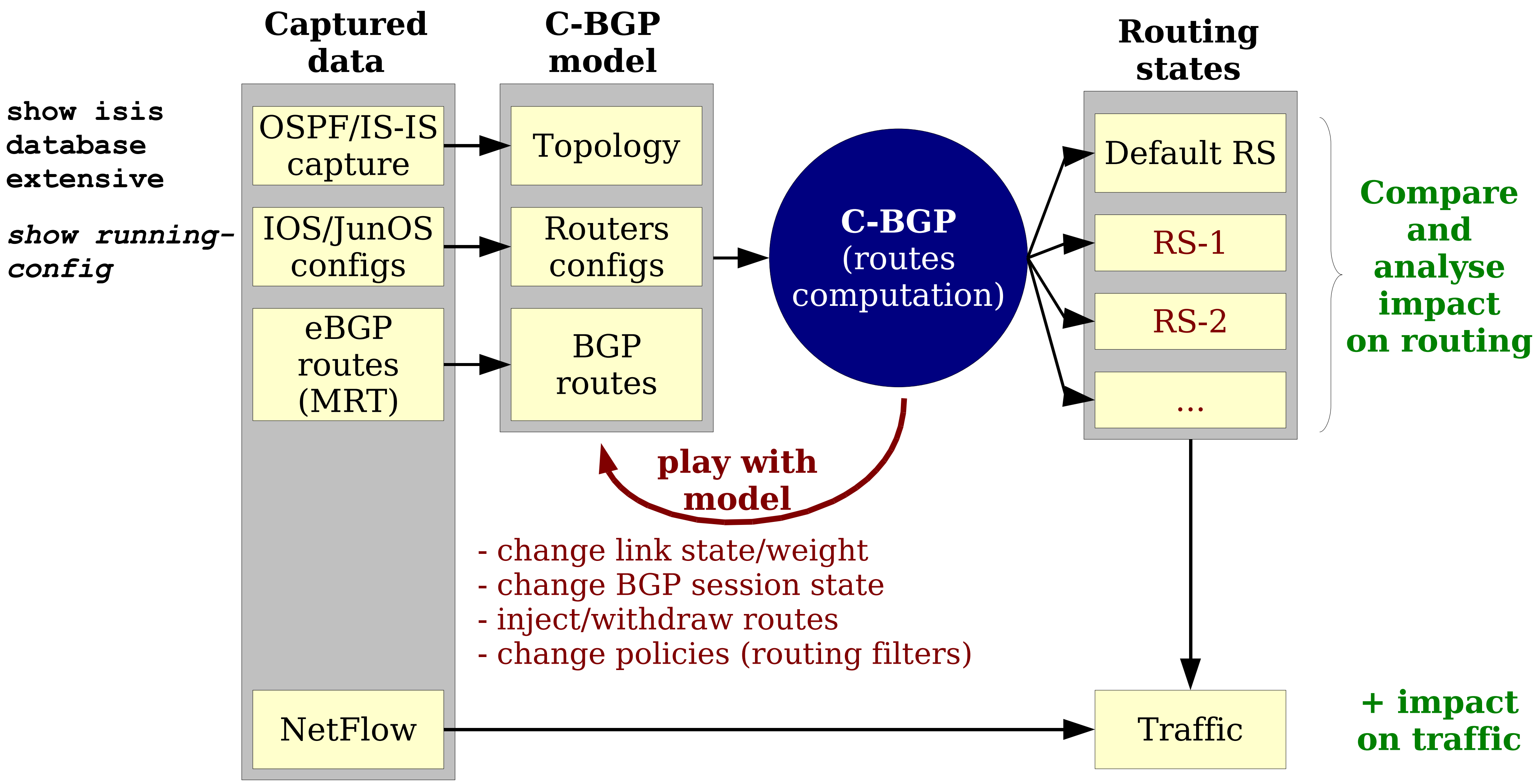
C-BGP

http://cbgp.info.ucl.ac.be

Part of the TOTEM toolbox

- **Purpose**
  - Compute outcome of BGP routing (steady-state) based on topology and routers configuration
- **Features**
  - Complete decision process
  - Versatile route filters
  - iBGP hierarchy (route-reflectors)
  - IGP model
  - Reads BGP dumps in MRT format
  - Configuration in CISCO-like language (scripts available to convert from IOS/JunOS configs)
  - Supports large-scale topologies
  - Open Source, LGPL
- **Applied on** Abilene, GEANT, a French Tier-1

C-BGP



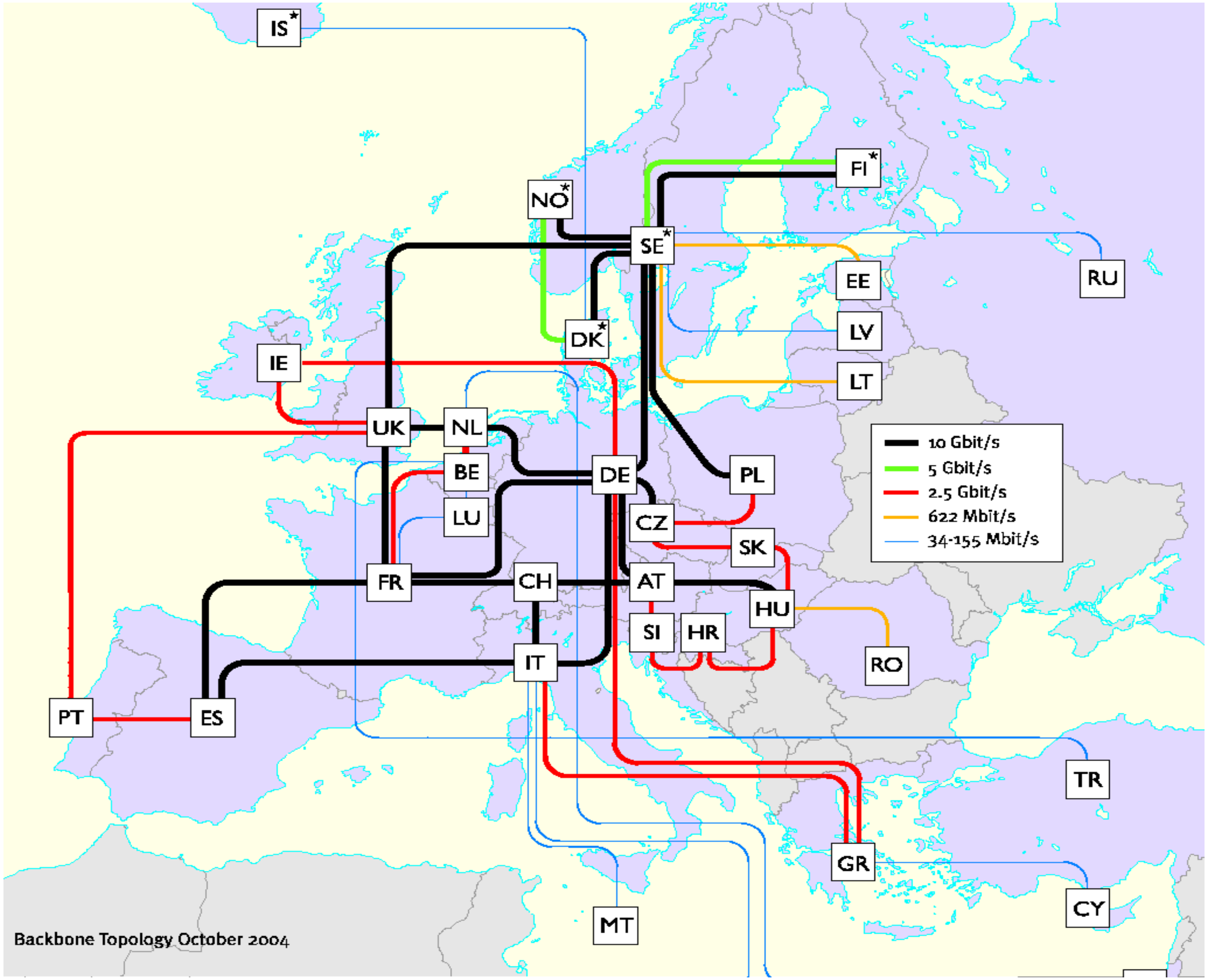
Captured data
C-BGP model
Routing states
show isis database extensive
show running-config
OSPF/IS-IS capture
Topology
IOS/JunOS configs
Routers configs
eBGP routes (MRT)
BGP routes
C-BGP (routes computation)
Default RS
RS-1
RS-2
...
Compare and analyse impact on routing
play with model
- change link state/weight
- change BGP session state
- inject/withdraw routes
- change policies (routing filters)
NetFlow
Traffic
+ impact on traffic

C-BGP example

```
cbgp> bgp router 198.32.12.9
cbgp-router> debug dp 214.3.50.0/24
Debug Decision Process
----------------------
AS11537, 198.32.12.9, 214.3.50.0/24

[ Current Best route: ]
*> 214.3.50.0/24        198.32.12.25    100     0       668 1503        i

[ Eligible routes: ]
*> 214.3.50.0/24        198.32.12.25    100     0       668 1503        i
*  214.3.50.0/24        198.32.12.137   100     0       668 1503        i
*  214.3.50.0/24        198.32.12.153   100     0       668 1503        i
*  214.3.50.0/24        198.32.12.169   100     0       668 1503        i

[ Highest LOCAL-PREF ]
[ Shortest AS-PATH ]
[ Lowest ORIGIN ]
[ Lowest MED ]
[ eBGP over iBGP ]
[ Nearest NEXT-HOP ]
[ Lowest ROUTER-ID ]
*> 214.3.50.0/24        198.32.12.25    100     0       668 1503        i

[ Best route ]
*> 214.3.50.0/24        198.32.12.25    100     0       668 1503        i
```

Abilene model
Show BGP routing decisions

Case study: GEANT (AS20965)



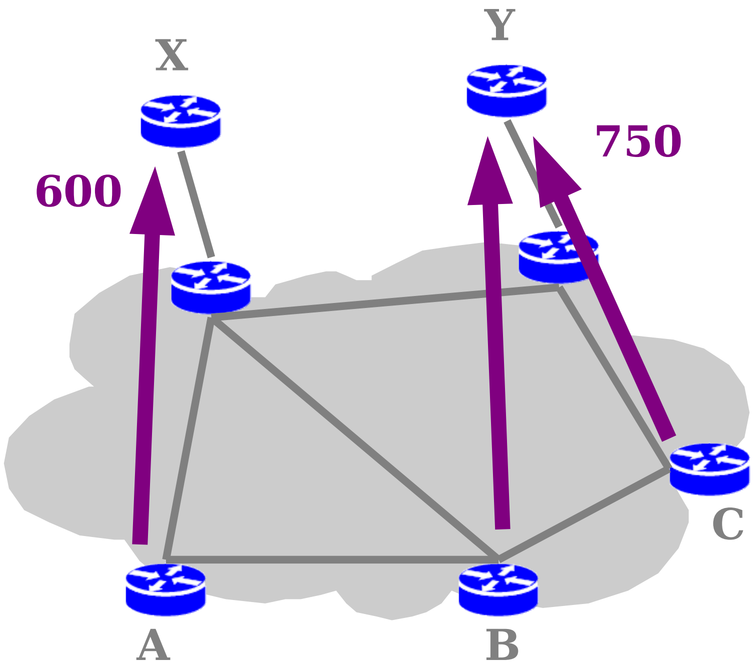
Case study: GEANT (AS20965)

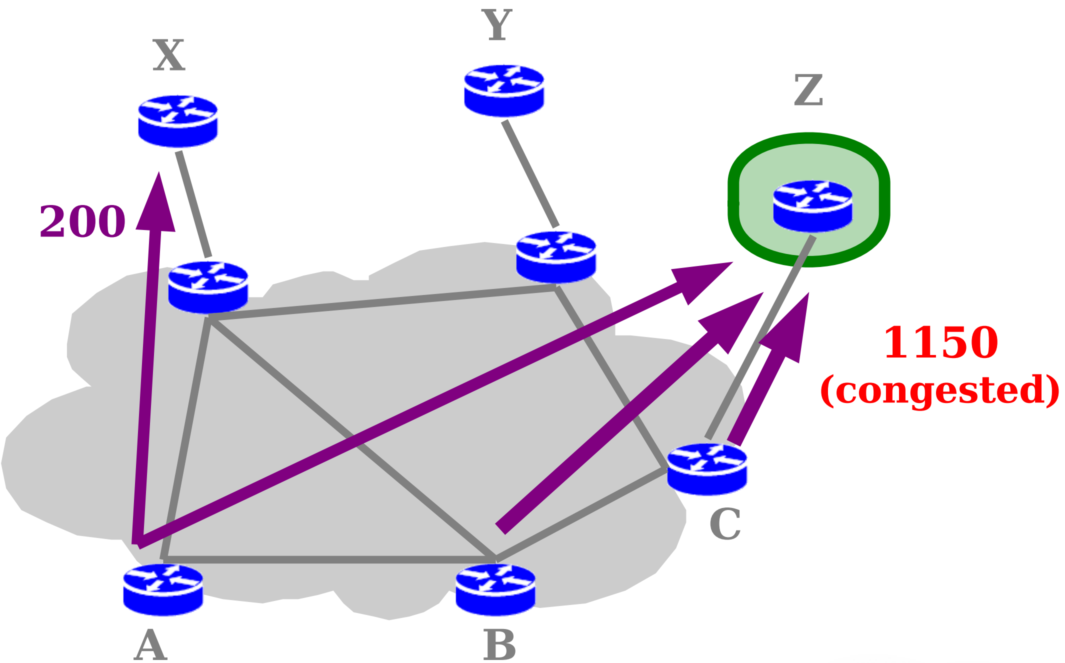
- **Topology**
  - Obtained from IS-IS trace, cross-checked with map
  - 23 nodes, 38 core links, 53 edge links (6 with upstreams)
- **Routing data**
  - Collected using Zebra in the iBGP (only best eBGP)
  - 640,897 eBGP routes
    - 150,071prefixes (clustered in 406 groups)
- **Traffic data**
  - NetFlow collected on all external interfaces
  - Sampling rate: 1/1000
  - About 150 GB per month
  - Aggregated in /24 src/dst prefixes (scripts available)

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

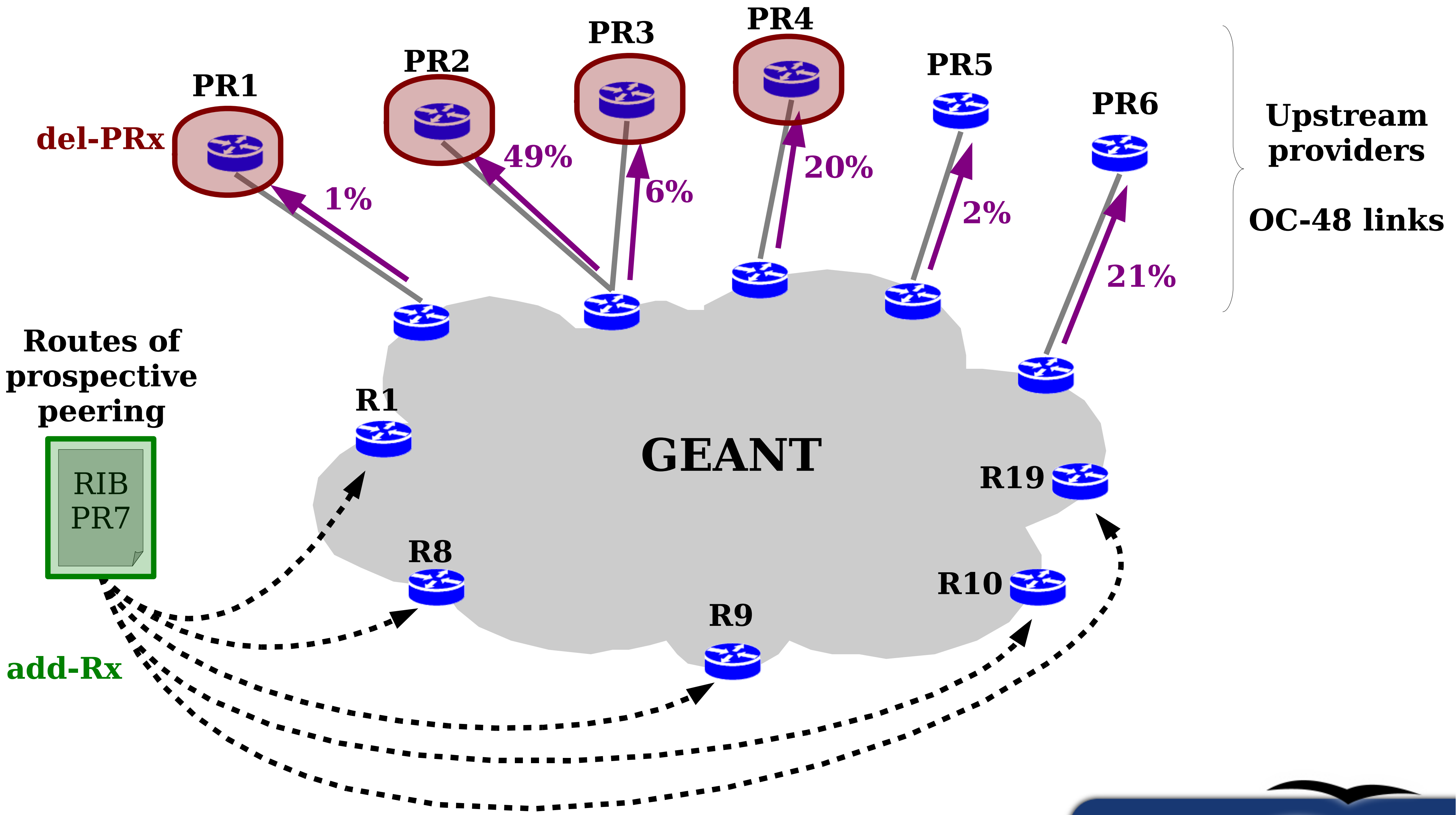




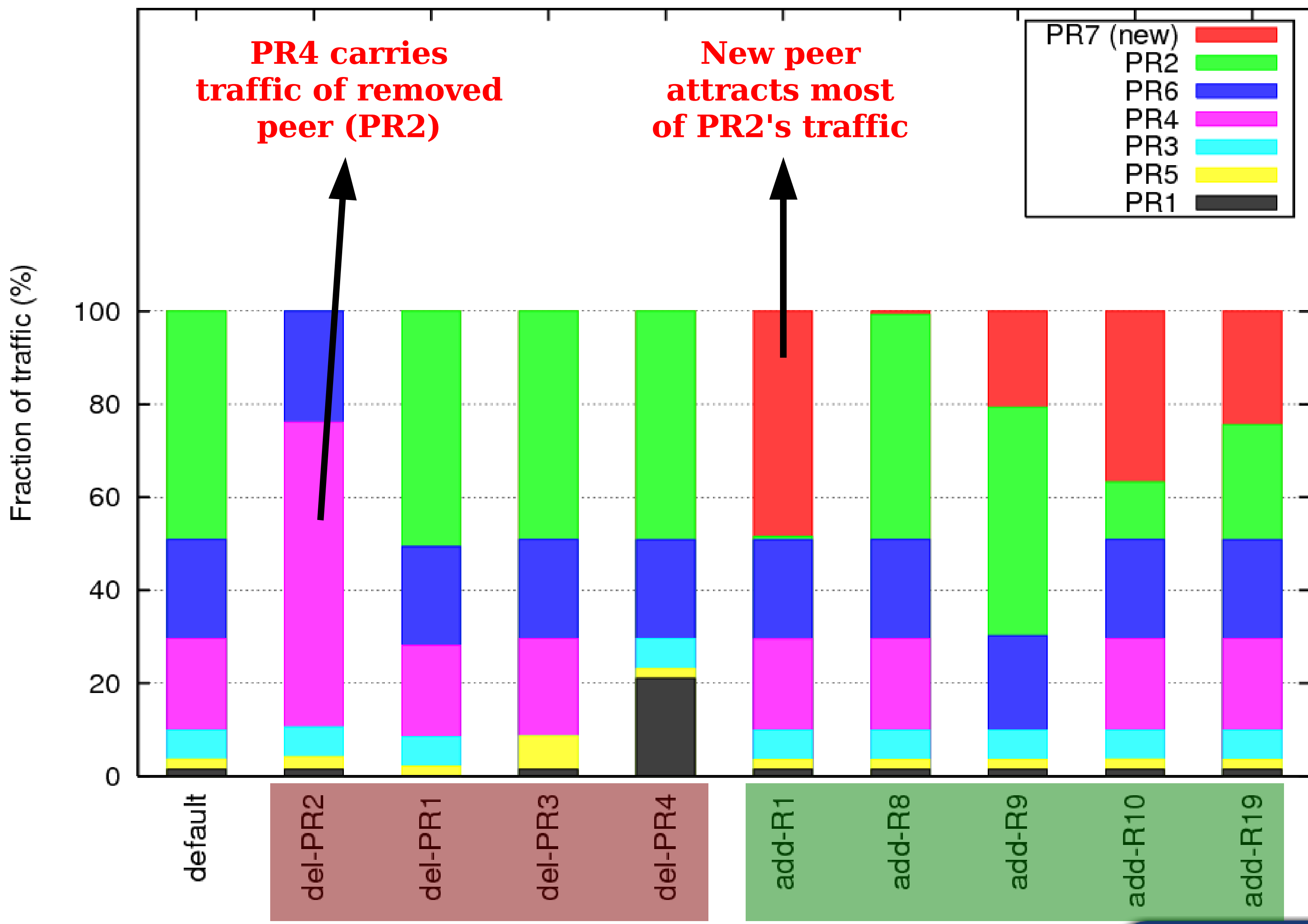
1st Scenario: peering placement

- **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



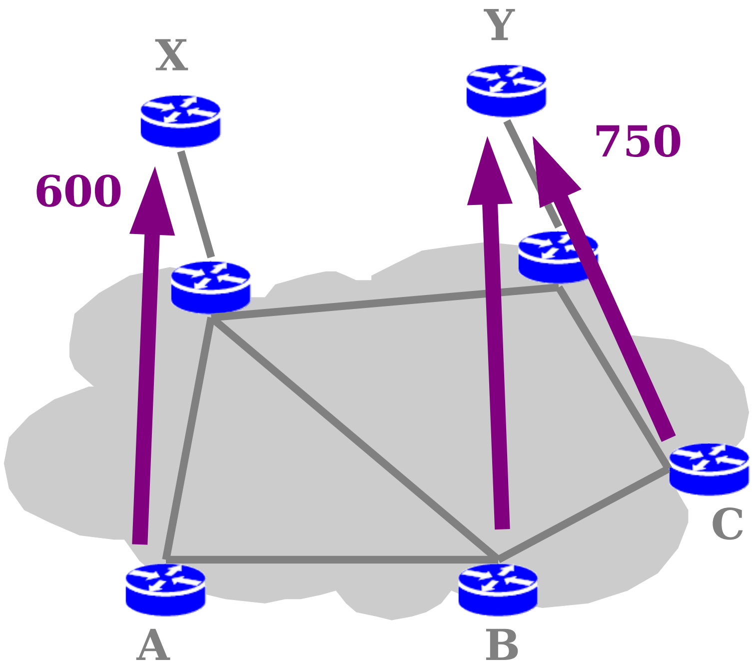
1st Scenario: peering placement

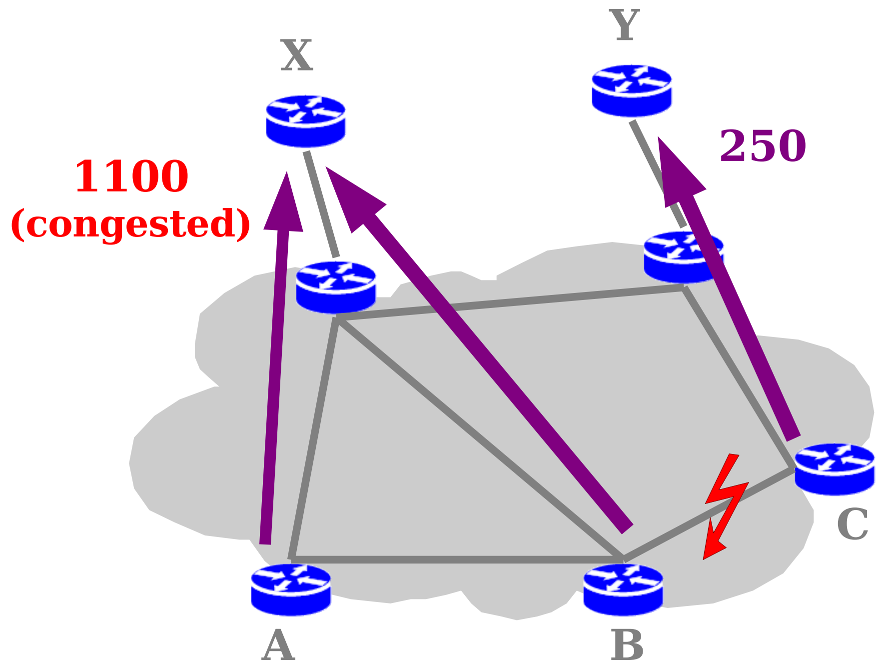


2nd Scenario: link failures

- **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

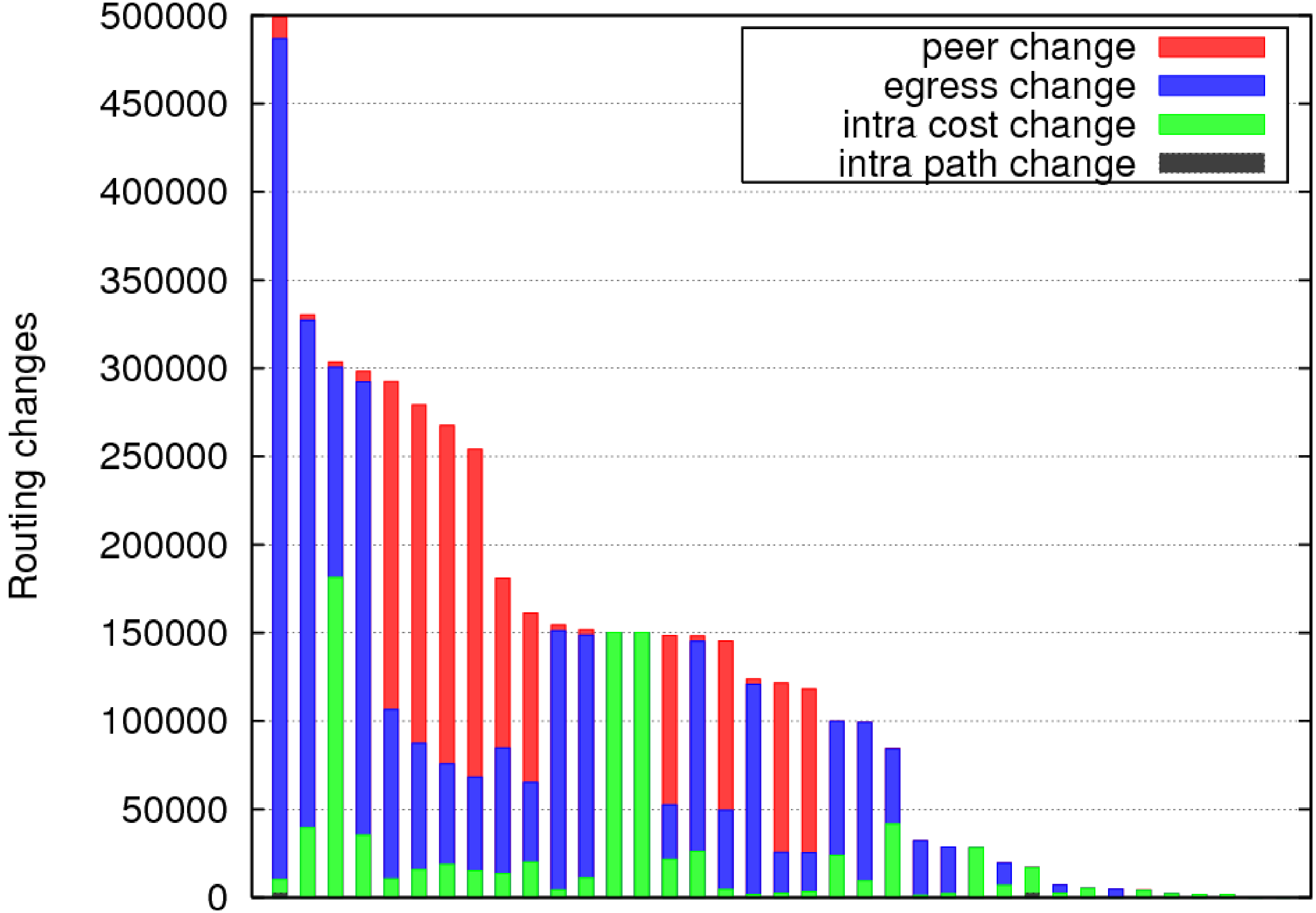




2nd Scenario: link failures

- **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**
  - Useful to predict **impact of events** on service & can be used as a **capacity planning** tool (if TM available)
  - Successfully applied to Abilene/Geant and a French T-1
  - Capacity planning tools focus on intradomain only
    - our experiments show that most routing changes are egress/peer changes ⇒ **taking BGP into account is not an option !**

- **Further work**
  - **Inbound traffic:** introduce neighbor ISPs in the model (already possible in C-BGP)
  - Computation of **failover matrices** *[Telkamp et al]*
  - MPLS/BGP **VPNs**

Thanks for your attention !

C-BGP

http://cbgp.info.ucl.ac.be

Contact information

Bruno Quoitin
Computer Science and Engineering Department
Université catholique de Louvain, Belgium
E-mail: bruno.quoitin@uclouvain.be

I will be in the room for further details, demos, ...

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