



Redesign Belnet Network Explained

Grégory Degueldre
Stefan Gulinck



Agenda

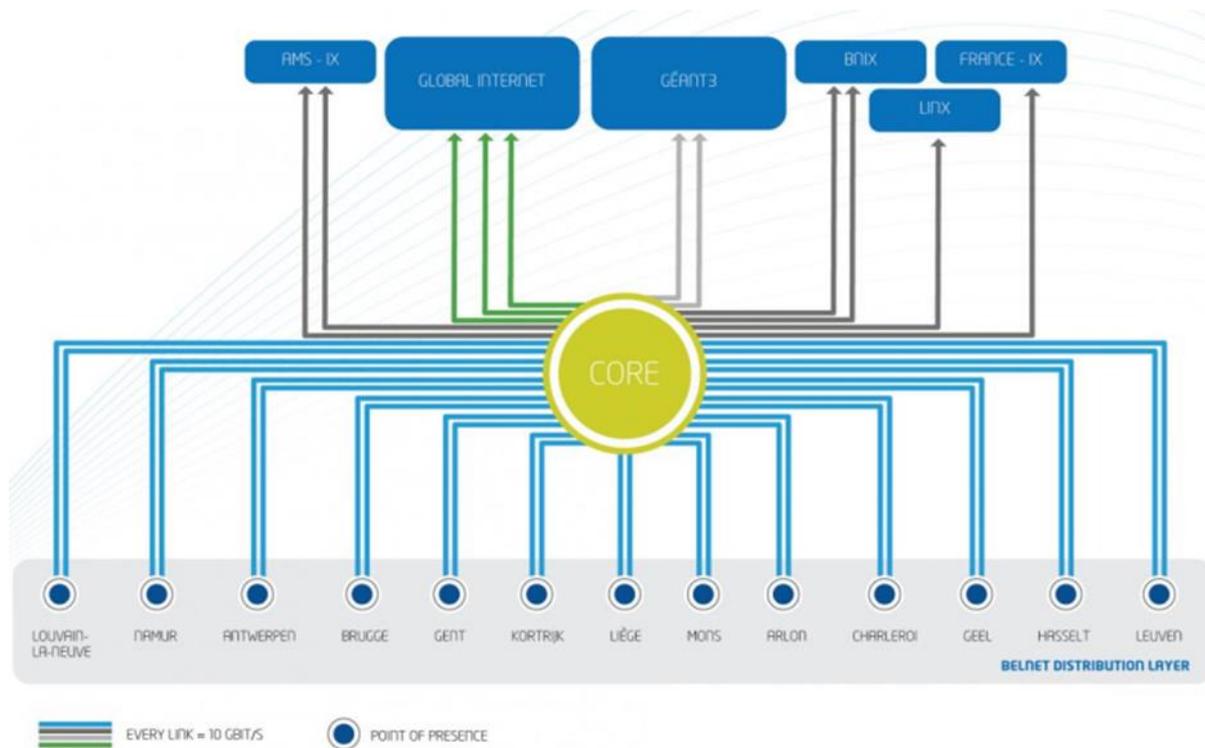


- History of the Belnet network topology
 - Situation as-is
- Driving factors (issues and incidents)
- Actions taken
- Redesign

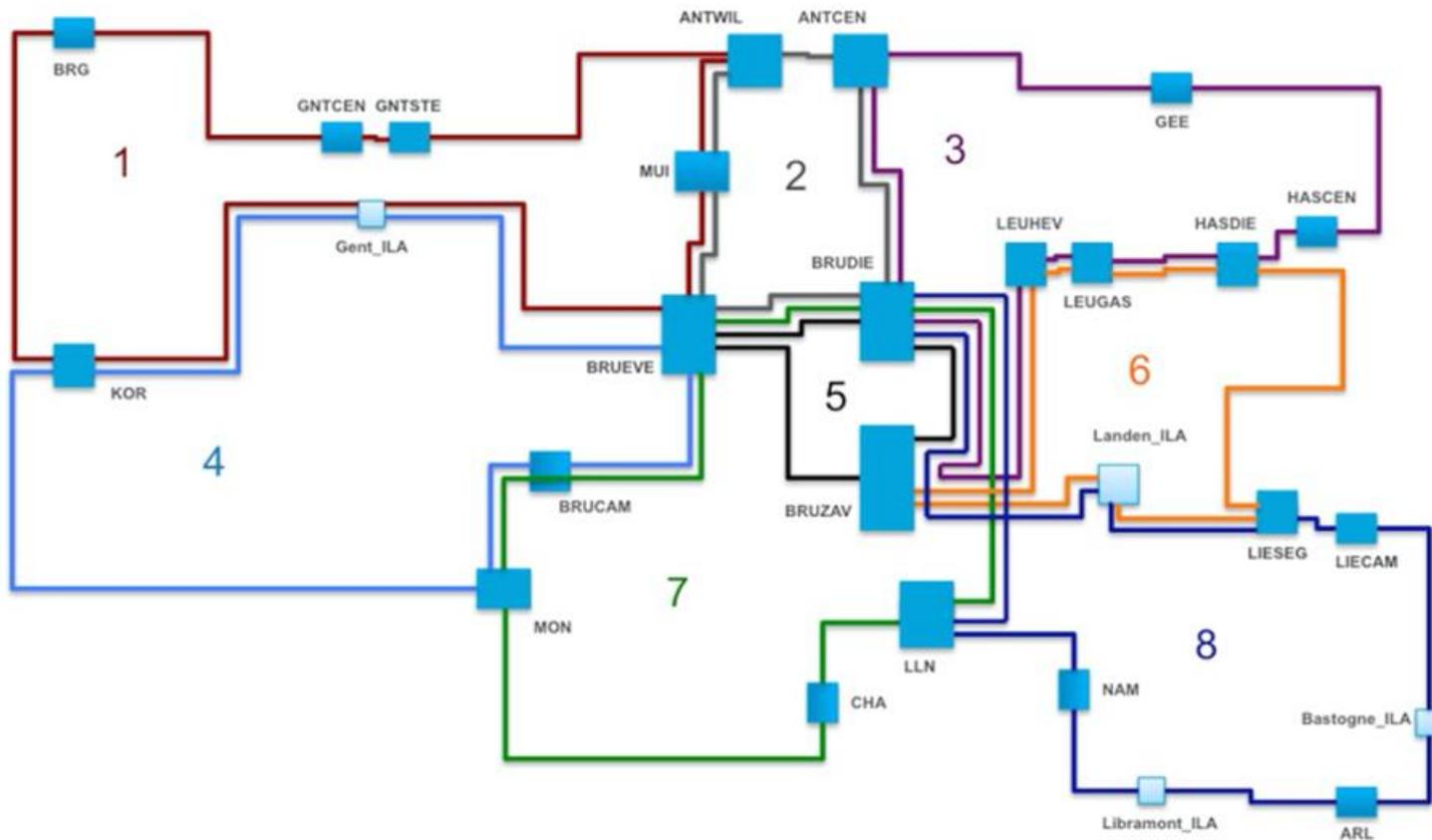
History of the topology



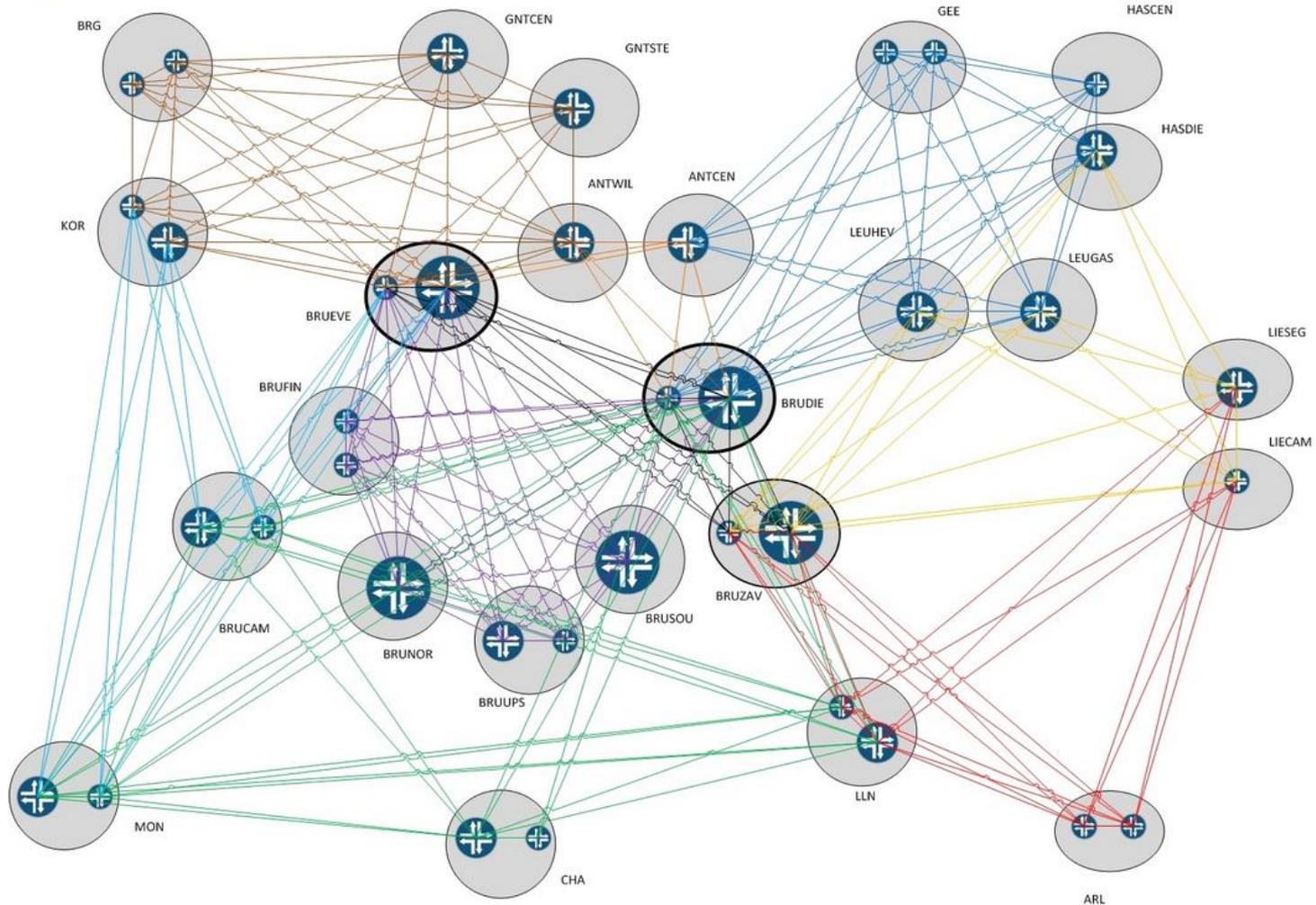
Belnet < 2016



History of the topology



Situation AS-IS



Issues

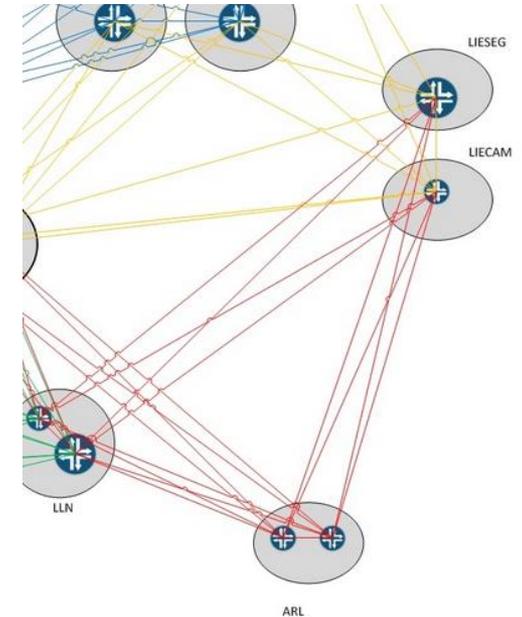


- Roots
 - G8032 bug
 - Ineffective MPLS Fast-Reroute
 - Big increase of traffic on September 2017
 - Bad repartition of bandwidth among the member of a LAG
- Incidents
 - 20/11 : Fiber cut between DC Evere and Zaventem
 - 09-13/12: Card flapping on r1.brueve

Issue 2: Fast-ReRoute (MPLS Redundancy)



- What is FRR ?
 - Redirection sub 50ms on MPLS layer
 - Dispensable with G.8032 but still implemented.
- What's the problem ?
 - Too many VLANs
 - Convergence
 - Path recalculation
 - BGP sessions down with big convergence time
- Work around:
 - BFD timer change to make the recalculation faster.



Config changed to avoid BGP to flap

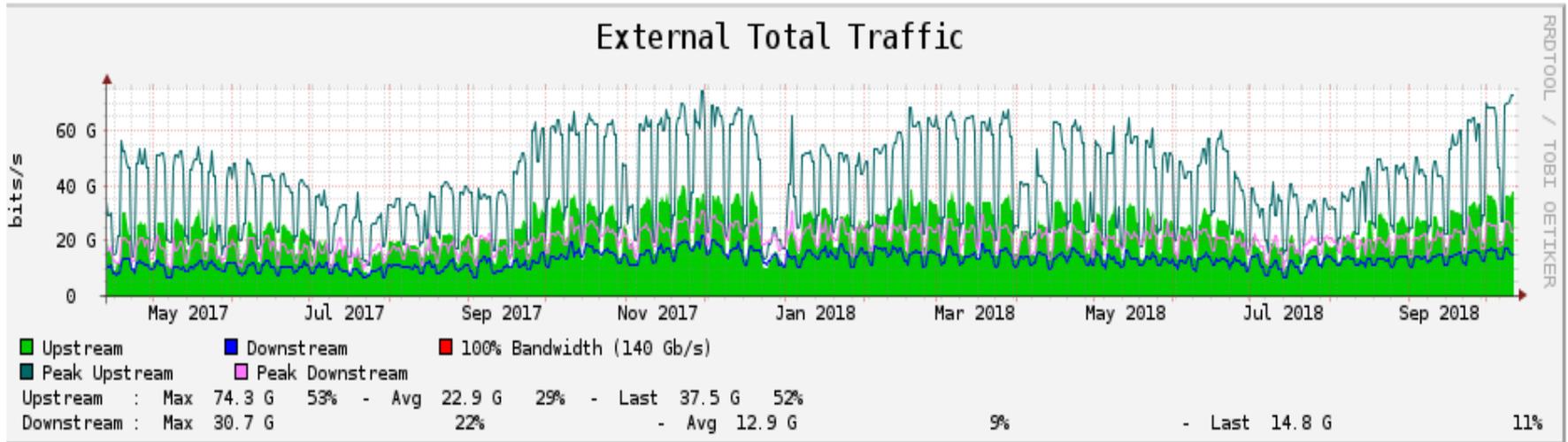
But

Reroute not sub 50ms

Issue 3: Poor hashing algorithm



Graphics* of your connection(s)
External Total - PoP BEL - External



- Yearly traffic increase on backbone
 - Use of cloud services (Office365, etc.)
- Capacity Mgt : issue with order of 100GE cards.
 - Extra ports in LAG

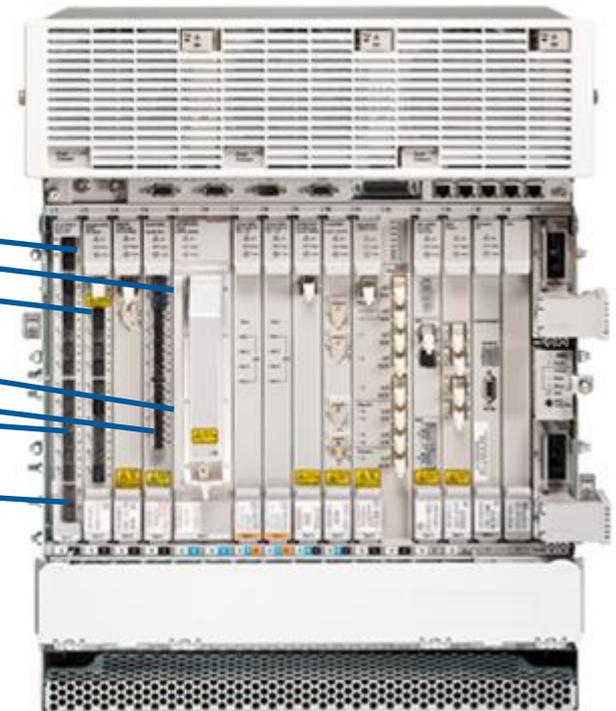
No big deal...

Issue 3: Poor hashing algorithm



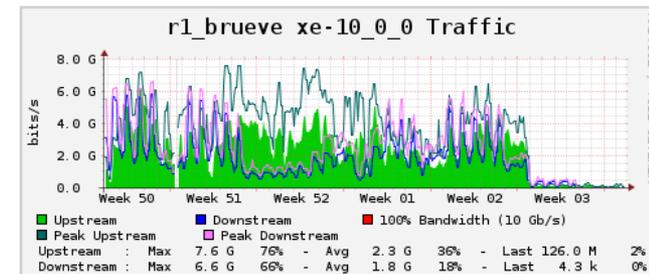
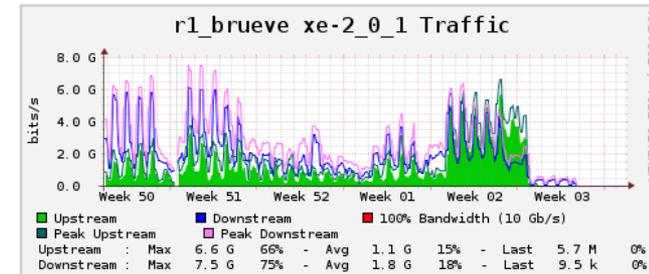
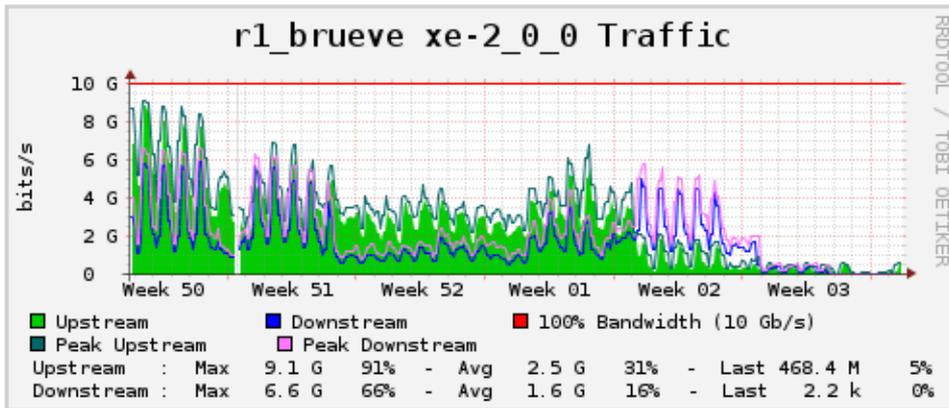
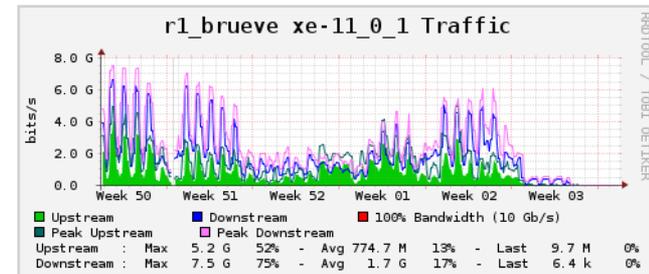
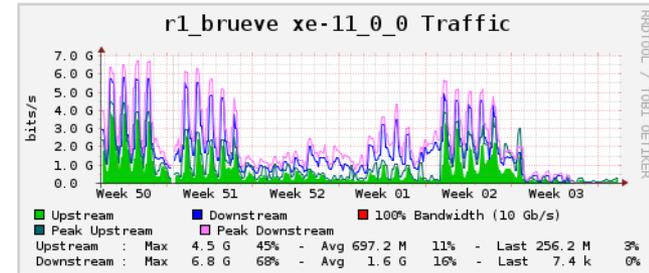
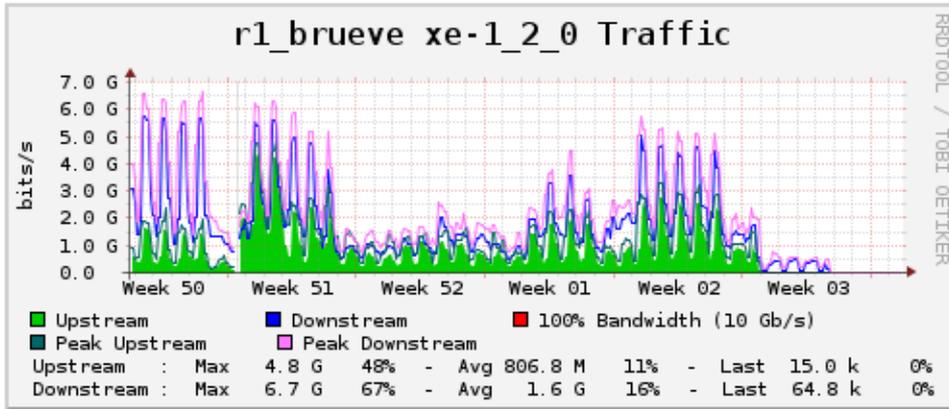
Repartition done by hashing algorithms

ciena : the network specialist



6500 Family
Packet-Optical Platform

Issue 3: Poor hashing algorithm



100GE card in Prod (EVE & ZAV & DIE)

But

Still NOK for other POPs

08/11/2018 | Redesign Belnet Network Explained



Incident 1: Fiber cut Evere - Zaventem



- 20/11/2017 : Fiber cut
- Impact: Saturation on bruzav impacting nearly all Belnet customers.
- Reactions:
 - New direct optical links between brueve and bruzav routers to offload the LAG.
 - Duplicated VLAN and MPLS path to increase the chance of a better repartition.

Bought some time waiting for the 100GE

Incident 2: Card flapping at brueve



- 9/12 – 13/12
- Flap of fpc (Juniper card)
 - Impact:
 - Backbone instability for all customers
 - Instability for customers connected on that specific fpc
 - Reactions:
 - Shutdown of the interface from the LAG => stable again but intensification of the issue of LAG repartition
 - All component have been replaced (fpc/mic/XFP/SFP)

Conclusion



- The situation is complex and is the result of a lot of design choices and workaround for encountered bugs/issues.
- Belnet has done a lot of things to improve the network and to diminish the impact during incident but there is still to be done
- Murphy hasn't help us a lot as everything that could go wrong has gone wrong.

Actions taken



- Redesign of the Network as a Project
 - **Project brief is approved as P1**
- COS → Class of service. Guarantee access to network management when things go A-wire
- Further upgrade 100GE card
 - On r1.brudie (central ring)
 - Redundancy on all three routers of central ring
- Redistribute transit routers more over the network
- We've abandoned G8032

Still To do...



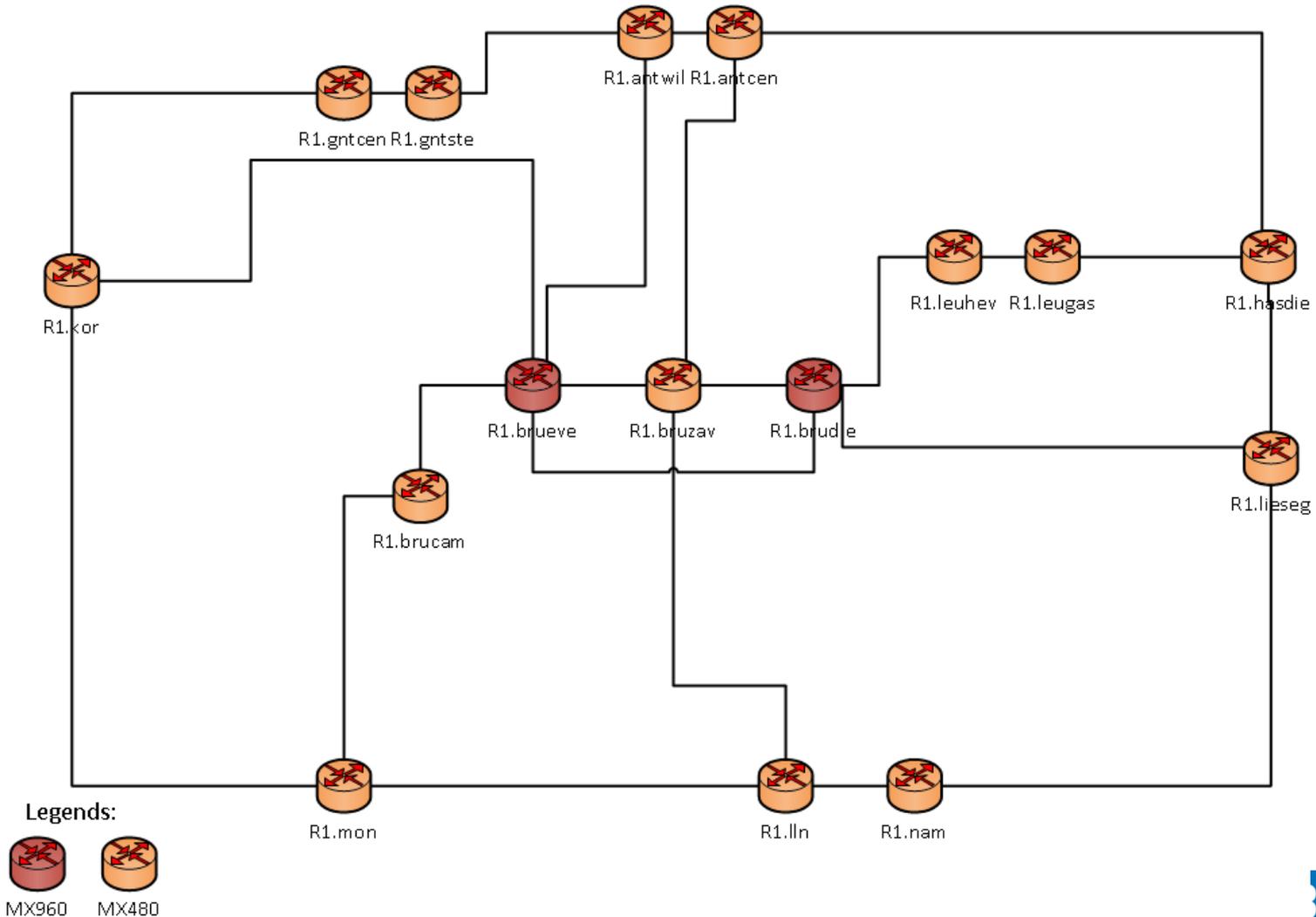
- Redesign Network and make it more robust and resilient.
 - Simplified network
 - Fast recovery and fast convergence
 - Better managed network for capacity management
- Solve Hashing issue
 - Testing and chasing third party to have a better hashing algorithm, i.e. 5-tuple hashing

Redesign



- Issues:
 - Hashing
 - Fast Reroute
 - Fast route convergence
 - QoS matching
- Manageability:
 - Readability of Network
 - Capacity Plan
 - Monitoring
- Cost
- IP Topology
 - Full-meshed
 - Ring
 - Star
- Transport Technology
 - Layer 1 (OTN)
 - Layer 2 (ELINE)
 - Layer 2 (ELAN)
- Onion vs Flat
 - Flexibility vs convergence

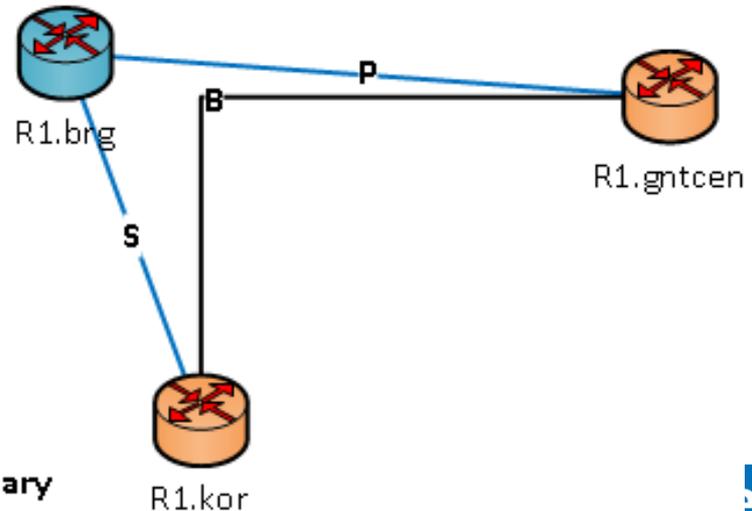
L2 Topology backbone (TO-BE)



Onion Approach



- Full routing table not on MX104 anymore
- (+) Better convergence time for BGP update
- (+) Memory usage on MX104
- MX104 will receive default route from two MX480/MX960
- (-) Less good decision about traffic routing
- (-) May require migration of customers with full routing table



P = primary
S = secondary
B = Backbone

Capacity study



- BRUSSELS (BRUDIE, BRUEVE, BRUZAV): 200Gbps
- 40Gbps:
 - ANTCEN
 - ANTWIL
 - BRUCAM
 - HASDIE
 - LEUHEV
 - LEUGAS
 - LLN
- 20Gbps: all others



Thank you for your attention





.be