IGP Tuning in an MPLS Network.
Experience and Design Issues.

Deutsche Telekom AG, T-Com Headquarters,
Technology Engineering, Internet Backbone Architecture.
Dr. Martin Horneffer

T · · Com

IGP Tuning in an MPLS Network.

Presentation Overview.

1. Motivation –
   Why do IGP-based traffic-engineering in an MPLS-network?

2. Getting the traffic-matrix.
2.1 The LDP based approach.
2.2 Extrapolating missing data.

3 Simulation and Optimization.
3.1 Defining the topology.
3.2 Metric design issues.
3.3 Virtual topology.

4. Summary.
5. Todos and Future Work.

IGP Tuning in an MPLS Network.
Motivation.

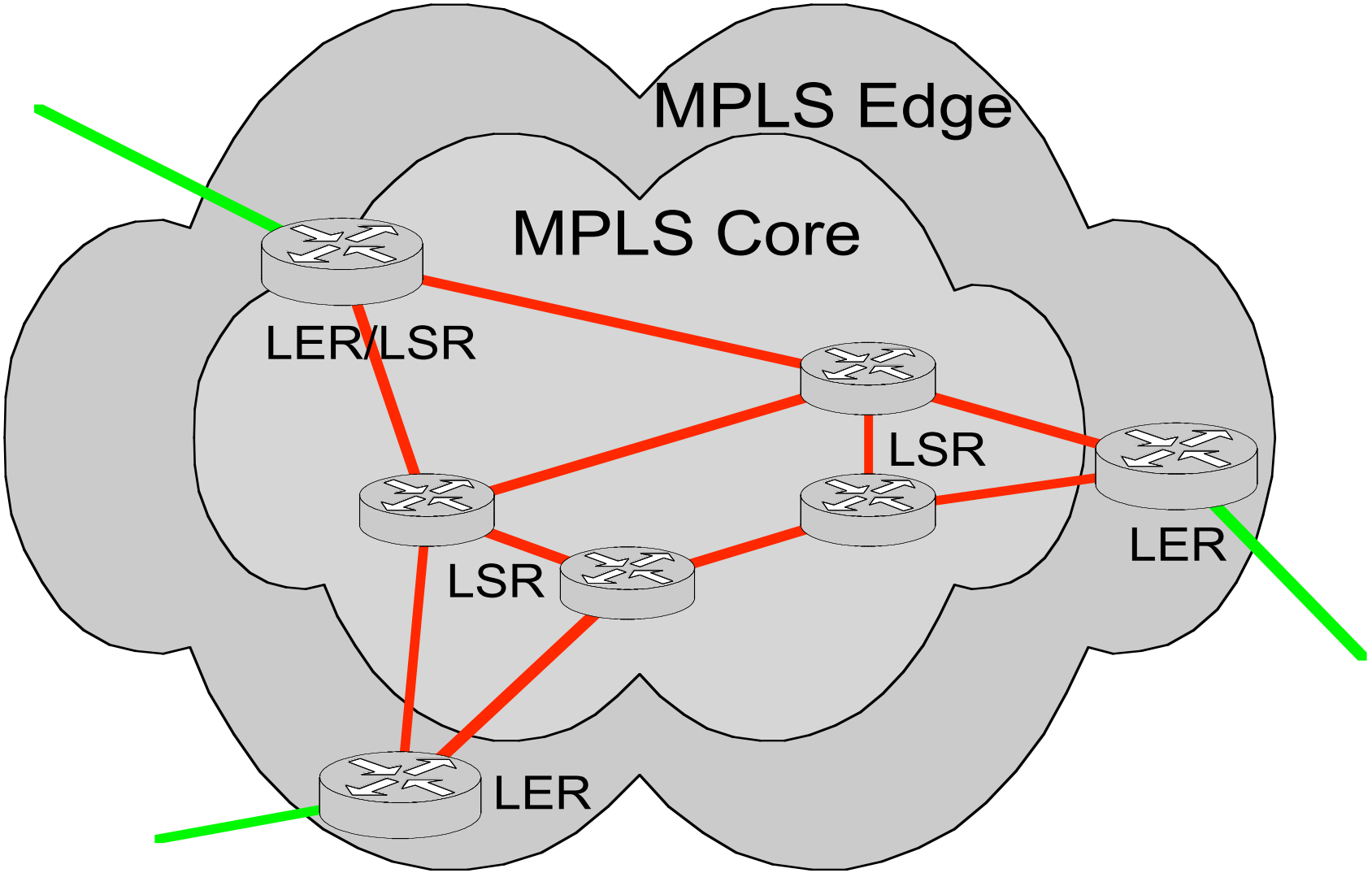
Our network design (AS 3320): MPLS and LDP for all traffic.

- LSR (Label Switch Routers):
  - Forward traffic only based on labels.
  - Limited routing table, no BGP.
- LER (Label Edge Routers):
  - ingress LER pushes label for egress LER.
  - Full routing table (public IPv4, IPv6, VPN, ...).
- Mixed LSR/LER: some routers have both functions.

➔MPLS for forwarding, plain IGP for routing!

IGP Tuning in an MPLS Network.

MPLS Design.



IGP Tuning in an MPLS Network.
Why do IGP-based traffic-engineering in an MPLS-Network?

"If you have an MPLS-network, you can easily do traffic-engineering by RSVP-based explicit routing."

Did some research:

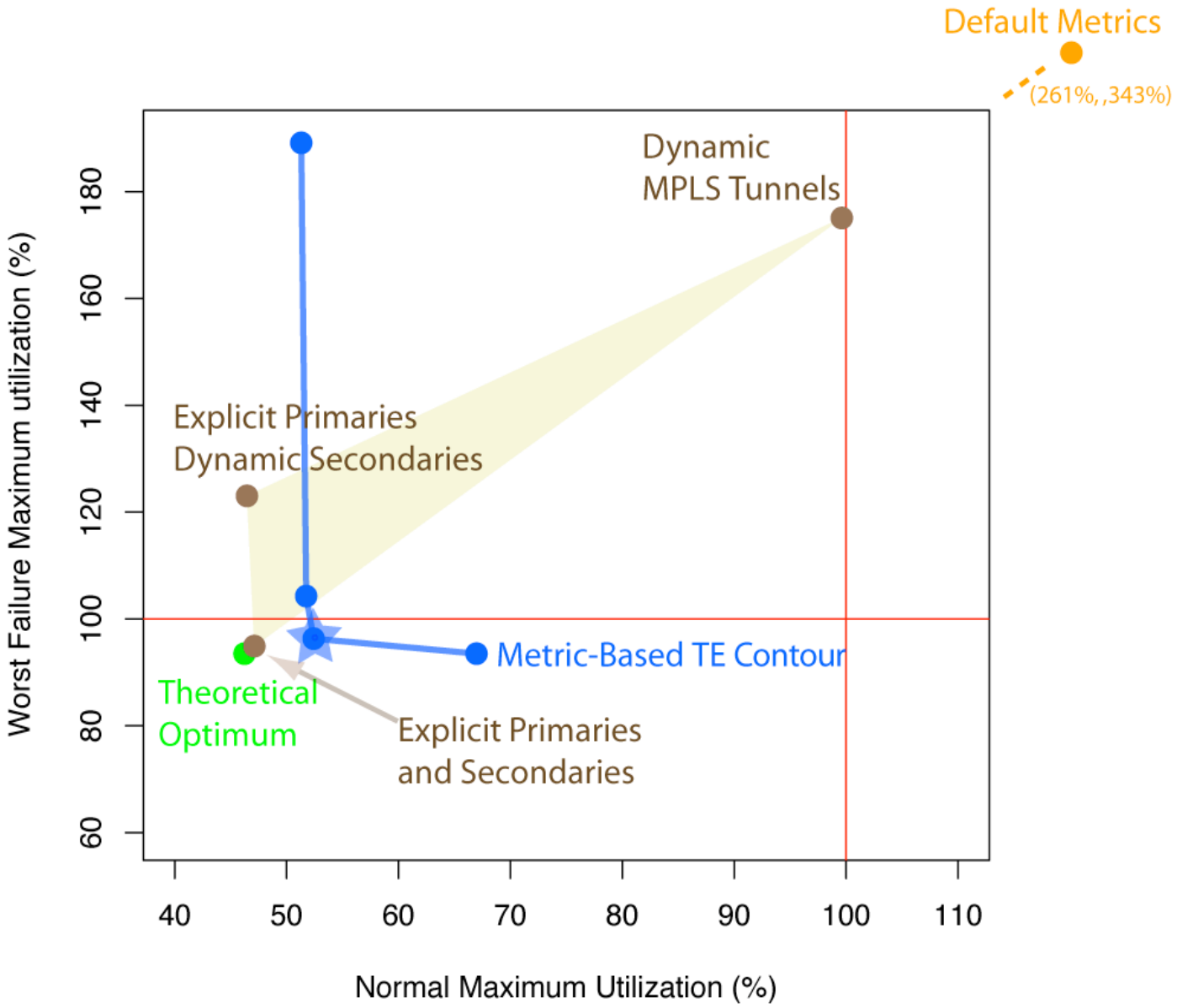
- Got traffic matrix first.
- Tried different designs, algorithms and tools on our actual topology and traffic matrix.

➔ Optimization of explicit routing had somewhat better results, but the difference to SPF-based routing was small.

➔ Explicit routing promised to introduce new risks and pitfalls.

➔ SPF-based routing is working and is well-understood by operations.

Network Utilization Study



IGP Tuning in an MPLS Network.
Getting the Traffic Matrix.

The LDP based approach:

- Poll MPLS counters (or "LDP statistics") and actual routing information from LSR.
- Have to know topology, e.g. from router configuration.
- Calculate traffic matrix for all FEC (forwarding equivalence class) between all LSR.

➔ Get the exact traffic matrix without having to manipulate the routing!

(s.a. S. Schnitter, T-Systems; M. Horneffer, T-Com. Traffic Matrices for MPLS Networks with LDP Traffic Statistics. Proc. Networks2004, VDE-Verlag 2004.)

IGP Tuning in an MPLS Network.
Getting the Traffic Matrix.

LDP statistics on JUNOS:

```
user@router> show ldp traffic-statistics
FEC                  Type      Packets       Bytes  Shared
 62.225.16.134/32    Transit   1236933   103984630  No
                     Ingress         0           0  No
[..]
user@router> show route table inet.3
[..]
62.225.16.134/32    *[LDP/9] 06:08:27, metric 1
                     > via so-0/0/0.0, Push 39
[..]
```

IGP Tuning in an MPLS Network.
Getting the Traffic Matrix: Issues.

LDP statistics on IOS:

```
router# show mpls forwarding-table
Local  Outgoing   Prefix           Bytes tag  Outgoing  Next Hop
tag    tag or VC  or Tunnel Id     switched   interface
26     26         62.225.24.184/30 0          PO10/0    point2point
       46         62.225.24.184/30 0          PO13/1    point2point
27     Pop tag    62.225.17.203/32 56529738   PO2/0     point2point
[..]
```

- No "ingress" in IOS!
  Cisco routers only count/report MPLS-packets here,
  no feasible way to count IP-packets per FEC (as of 2004).
- Sources are wrongly seen at next MPLS hop.

IGP Tuning in an MPLS Network.
Getting the Traffic Matrix: Issues.

Problem only affects routers of mixed LSR/LER type.

- There are only a few routers of this type in our network, but they contribute an important proportion of the traffic.
- The amount of missing traffic sourced by a router is known from interfaces counters.

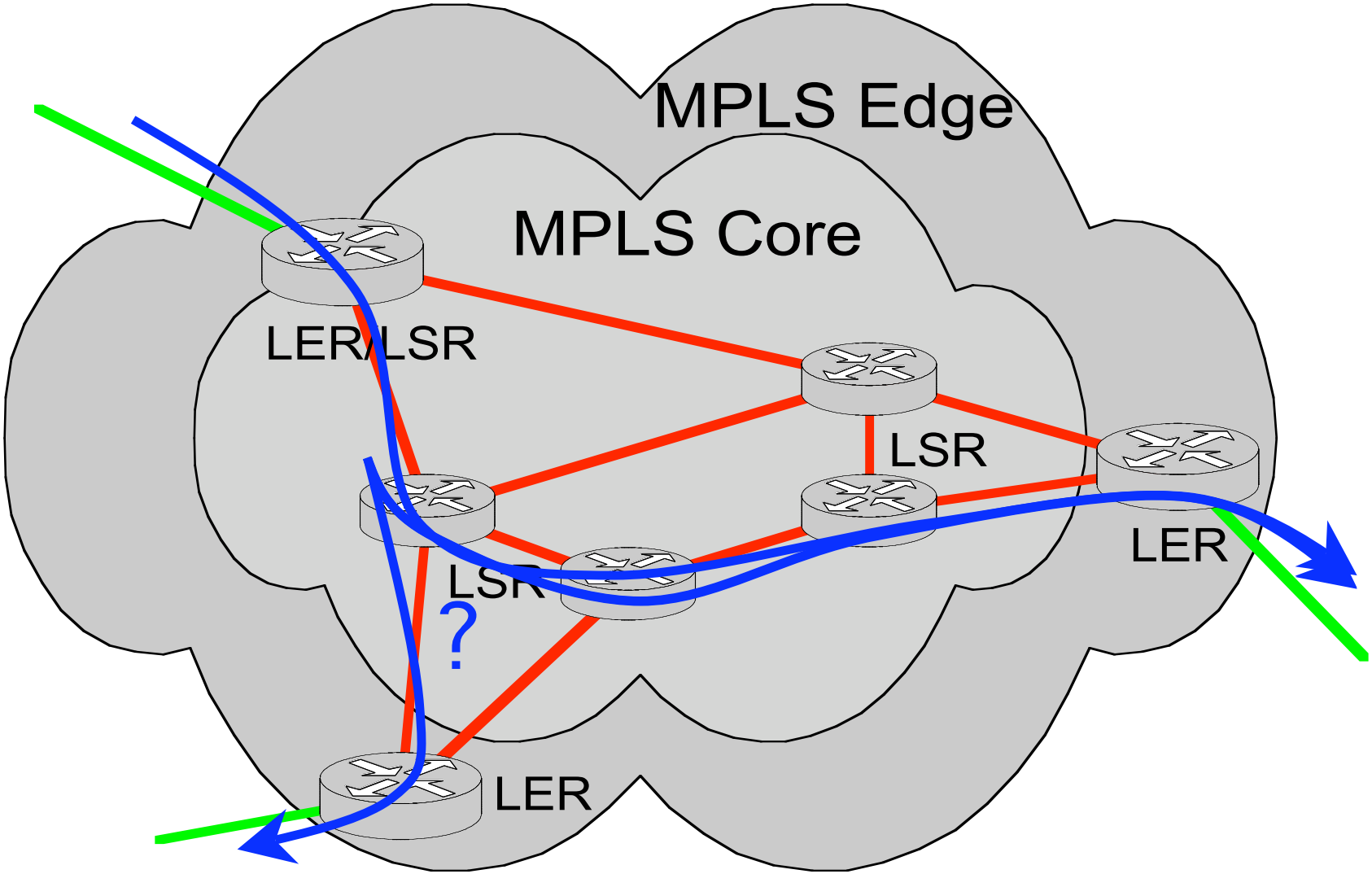
Two methods for estimating the missing traffic (i.e. guess the distribution of locally sourced traffic on possible destinations):

- Use traffic from locally connected LER.
- Use the distribution of the opposite direction.

➔ Select the method that fits best for each affected router.

IGP Tuning in an MPLS Network.

Getting the Traffic Matrix: Issues.



IGP Tuning in an MPLS Network.
Defining the Topology.

Cannot include all routers, have to select a suitable sub-topology:

- Scalability issues of simulation/optimization algorithms. Work for <100 routers and a few thousand demands, but not for hundreds of routers and tens of thousands of demands.
- Edge and regional routers may not contribute much to traffic-engineering.

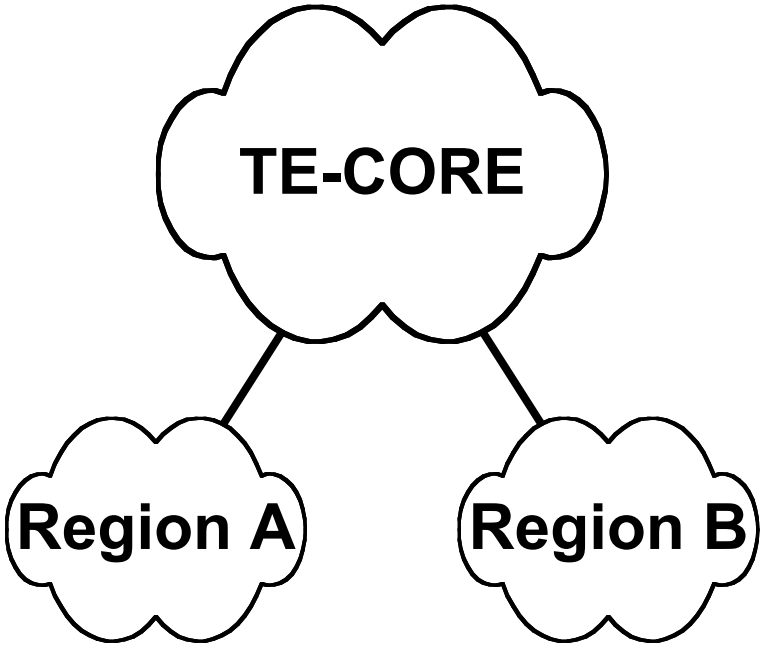
But need to include:

- Long and expensive links (e.g. international, inter-continental).
- All routers and links that are significant for distributing the traffic among the long and expensive links.
- Suitable representation of edge/regional routers.

IGP Tuning in an MPLS Network.
Topology and IGP Metric Optimization.

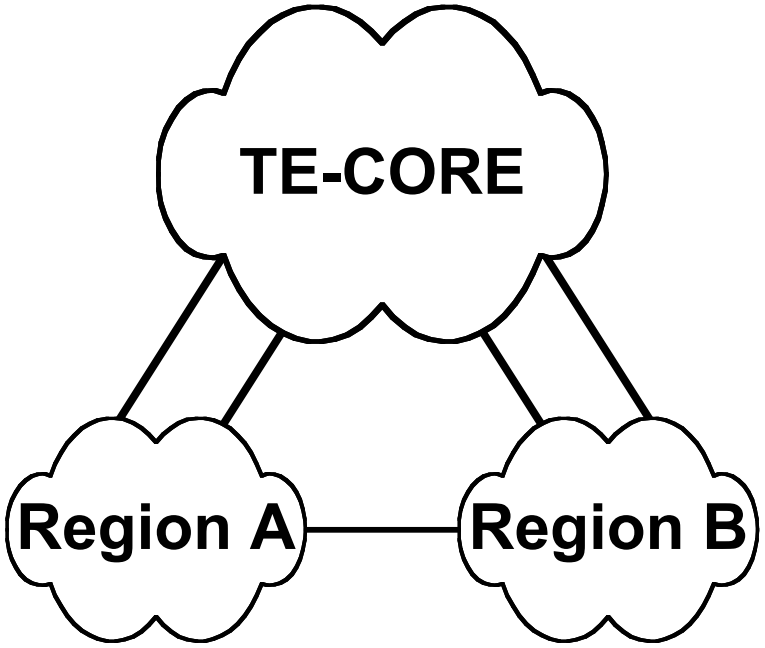
Ideal topology:

- Regions connected to TE-core by one link.
- No other links between regions.



Real topology:

- Regions connected to TE-core by more than one link.
- Maybe even links between regions.



IGP Tuning in an MPLS Network.
Contradicting Requirements on the IGP Metric Design.

Traffic-Engineering: Requirements for IGP metric optimization

- Metrics in TE-core can be anything; as low or as large as deemed right by the optimization algorithm.
- Metrics in the regions have to be large – larger than any possible metric in the TE-core.

➔ Predictable routing and stable traffic matrix in the TE-core.

RFC 3345: Requirements for iBGP route reflection

- Inter-Cluster links must have a higher IGP metric than the intra-Cluster links.

➔ Avoid persistent route oscillation.

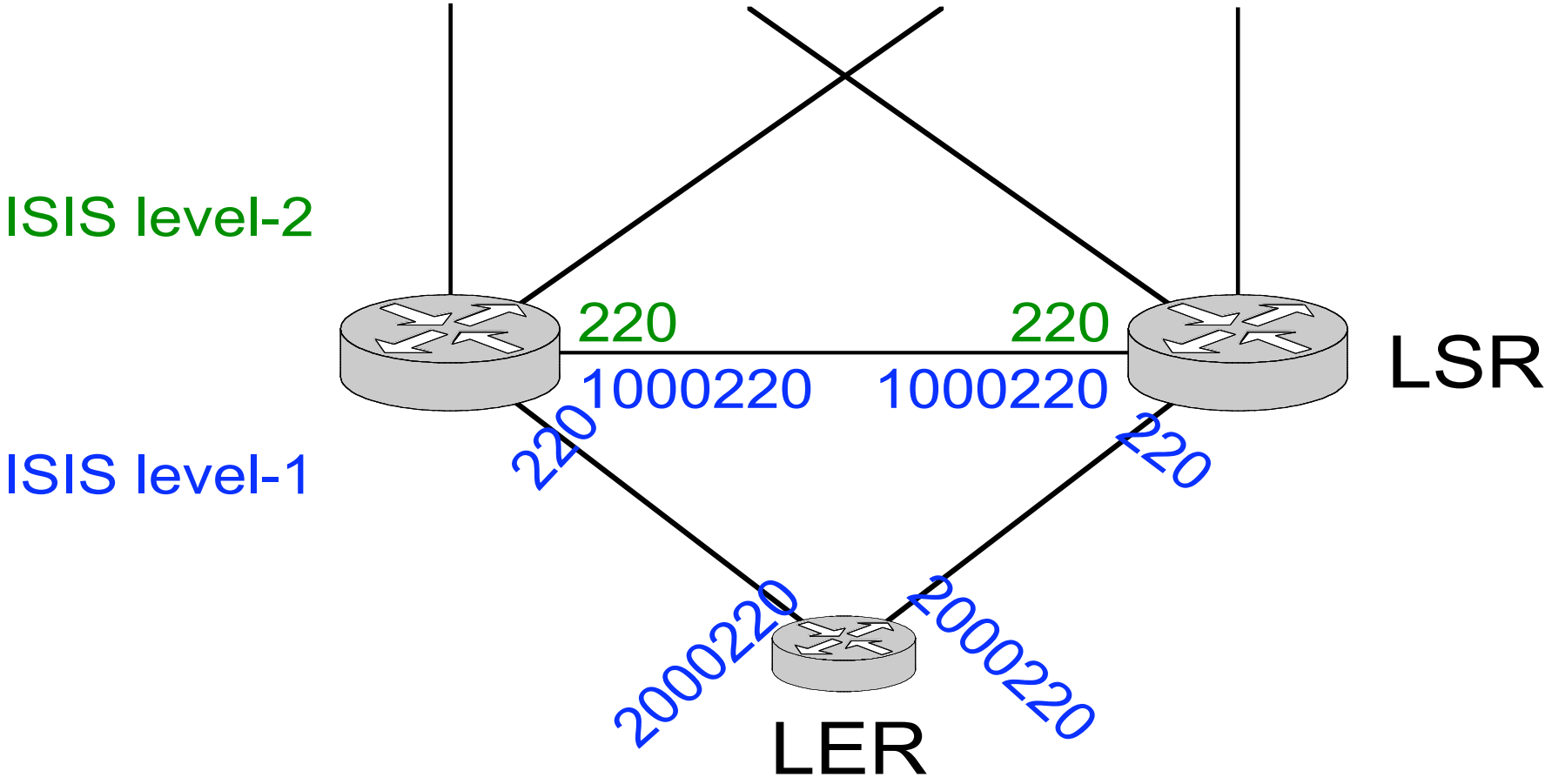
IGP Tuning in an MPLS Network.
IGP Metric Design Issues.

ISIS level-1/level-2 boundary may help:

- Traffic from level-2 cannot pass through level-1-only links and go back to level-2.
- But route-leaking still causes interference.
- And route-leaking is strictly required for BGP-free MPLS routing in the core.

IGP Tuning in an MPLS Network.
IGP Metric Design Issues.

Careful IGP metric design needed:



IGP Tuning in an MPLS Network.
Virtual Topology.

Simulation is used for normal case as well as for failure scenarios.

- Edge and regional routers must be represented in a way that shows the correct routing for the normal case and at least roughly for failure cases.

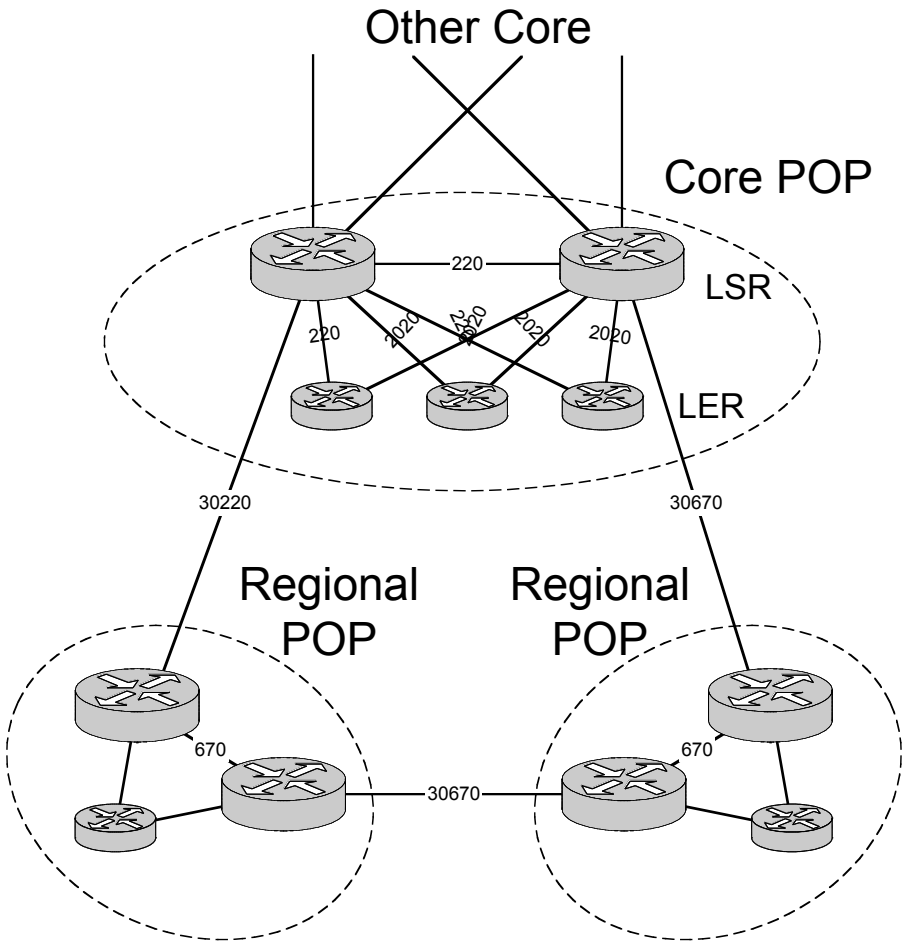
In our actual network each core POP has two LSR that connect to several regional POPs and many local edge routers.

These are represented in the simulation by "virtual routers":

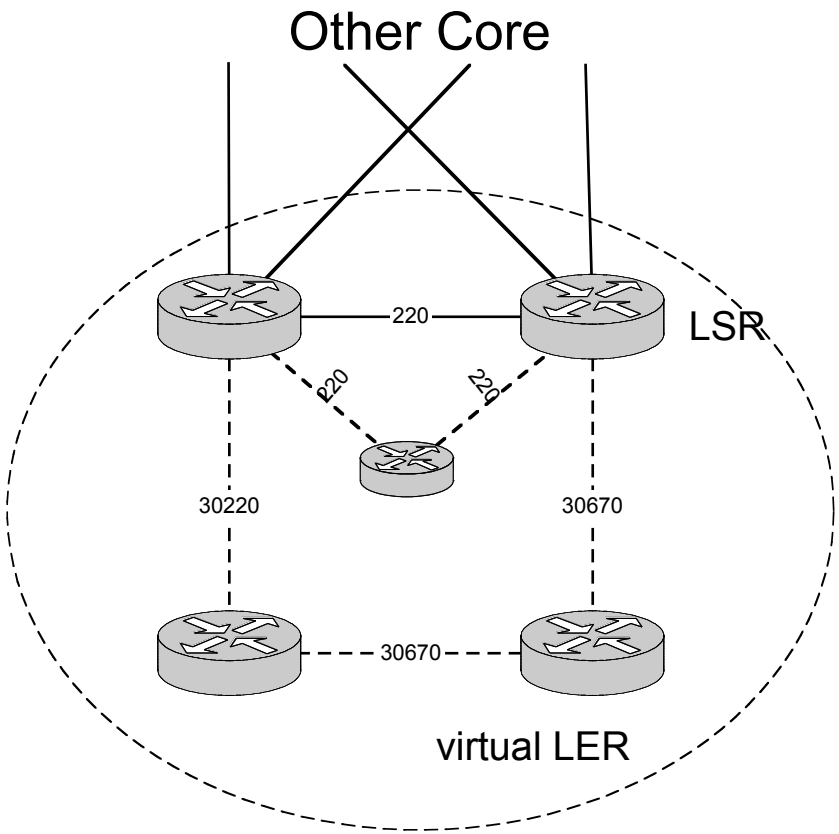
- one symmetrically connected edge router and
- two asymmetrically connected routers.

IGP Tuning in an MPLS Network.
Virtual Topology.

Real topology:



Virtual topology:



IGP Tuning in an MPLS Network.
Summary.

- MPLS can help traffic-engineering, even if no RSVP is used and routing is strictly SPF-based: counters per FEC.
- The sub-topology for traffic-engineering has to be carefully selected.
- The metric design for the whole topology has to cope with traffic-engineering as well as with iBGP route reflection (if used).
- Good (although not perfect) optimization is possible with very little additional CAPEX and OPEX.
- Simulation can verify that SLAs are not compromised by a more efficient use of the network.

IGP Tuning in an MPLS Network.
Todos and Future Work.

- Further improve simulation and optimization accuracy:
  - Usable per-FEC counters for IP packets in IOS.
  - Optimize virtual topology.
  - Include redundant peerings in simulation.

- Class-based traffic engineering:
  - Forwarding decision also based on precedence/ experimental bits.
  - Use multi-topology ISIS or mixed IGP/explicit scenario.