UNIVERSITÄT OSNABRÜCK

Real-world ~~Requirements~~ Pitfalls of Segment Routing Traffic Engineering

Timmy Schüller

Osnabrück University

RIPE 78

Reykjavik, May 21, 2019

Segment Routing ..

.. is a Source Routing Architecture

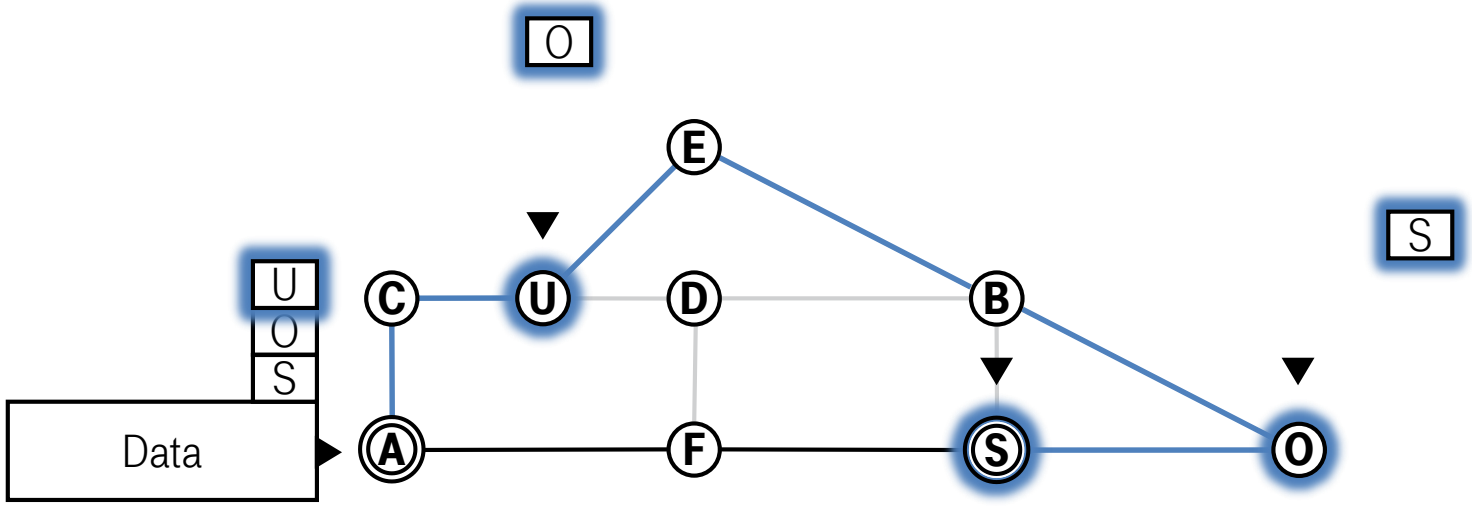
.. uses Node, Adjacency and Service Segment IDs to define “checkpoints”

.. encodes a path as a stack of SIDs

.. consults IGP to reach an SID

→ SR tunnel ≔ concatenation of shortest paths

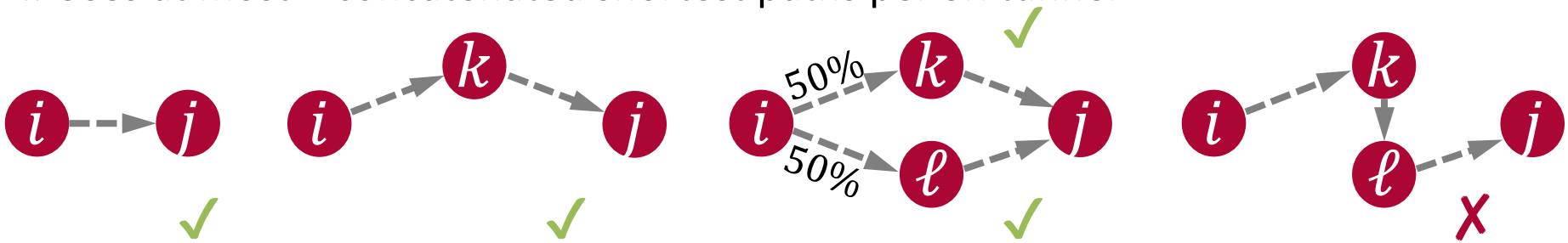
.. can be implemented with MPLS or IPv6 & an IGP extension



2SR ..

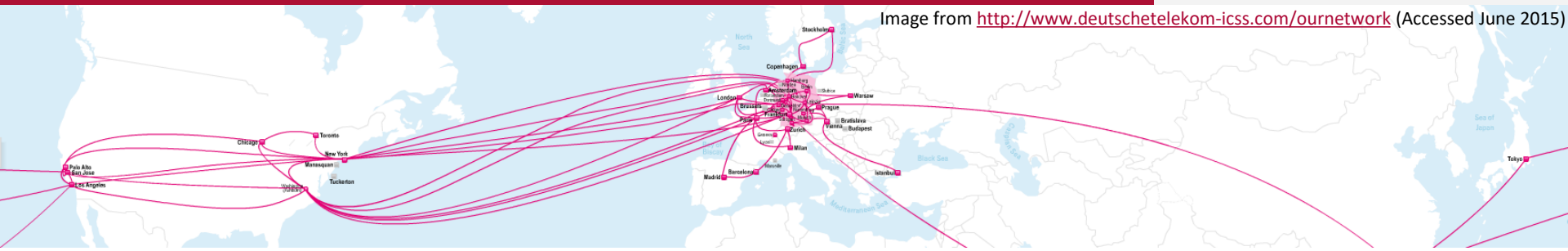
.. minimizes max. link utilization

.. Uses at most 2 concatenated shortest paths per SR tunnel



Minimize max. utilization

Image from http://www.deutschetelekom-icss.com/ournetwork (Accessed June 2015)

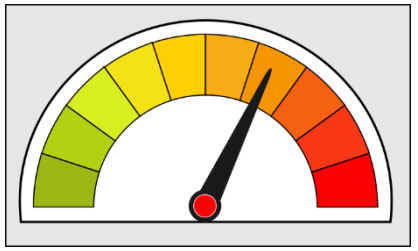


Input

$$\begin{aligned} &\min\ \theta \\ &\text{s.t.} \quad \sum_k x_{ij}^k \geq t_{ij} \qquad \forall (ij) \\ &\sum_{ij}\sum_k g_{ij}^k(e)\, x_{ij}^k \leq \theta\, c(e) \qquad \forall e \\ &x_{ij}^k \geq 0 \qquad \forall (ij) \end{aligned}$$

Output

Max. util. = 64%

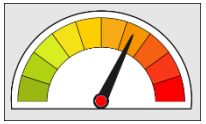


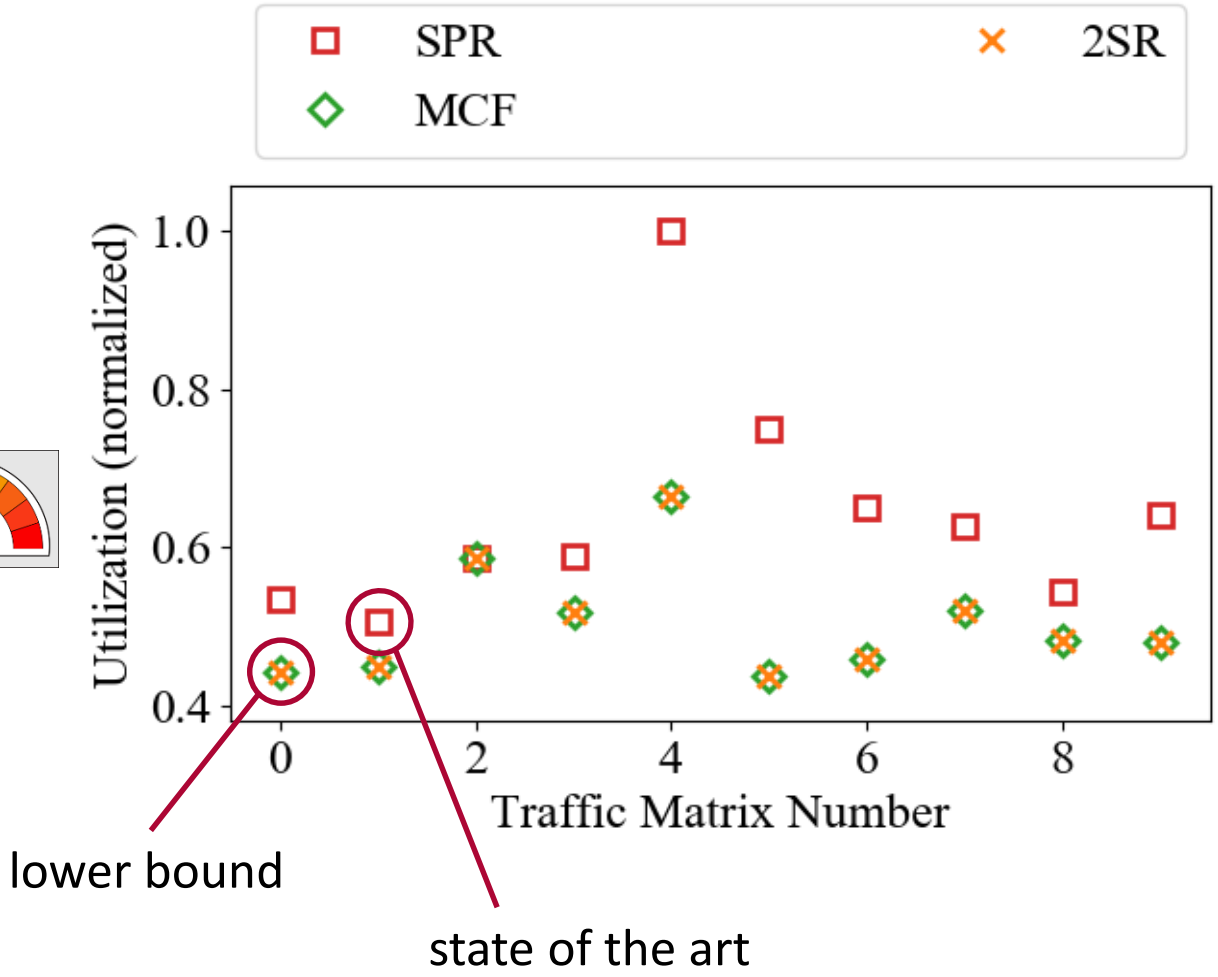
Source	Destination	Traffic
RKV	FRA	12
OS	RKV	7
FRA	RKV	42
etc.		

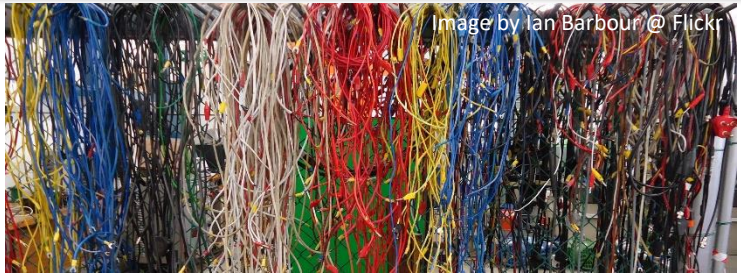
Use SR Paths:
RKV → OS → FRA
FRA → OS → RKV
etc.

2SR formulation from Bhatia et al.: *Optimized Network Traffic Engineering using Segment Routing.* IEEE INFOCOM, 2015

10 Snapshots of a Tier 1 ISP Backbone from 2018



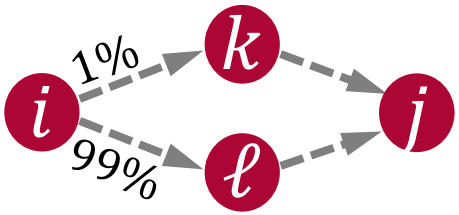




Minimize max. utilization

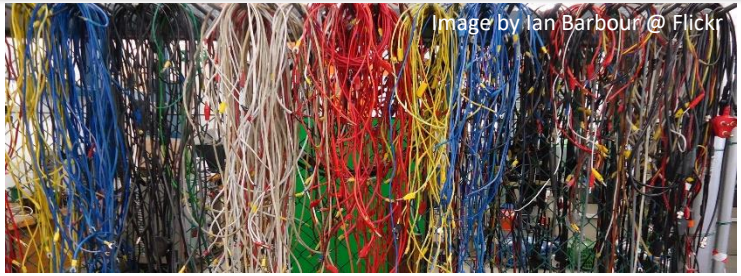
Keep # of SR tunnels low

Limit split factors



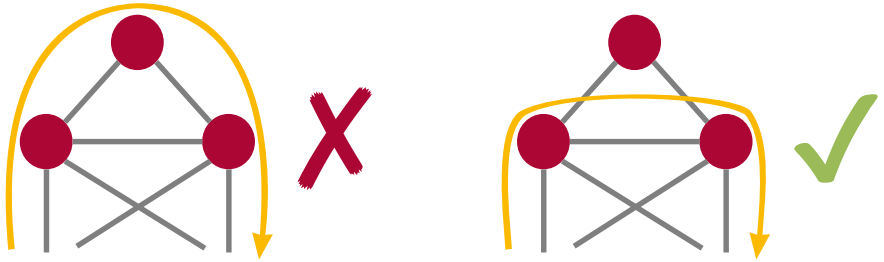
$$
\begin{aligned}
\min \quad & \theta' \; \frac{1}{2\,\lambda\,\theta} + \sum_{k \neq j} u_{ij}^{k} \\
\text{s.t.} \quad & \sum_{k} u_{ij}^{k} = 1 && \forall ij \\
& \sum_{ij} \sum_{k} g_{ij}^{k}(e)\, u_{ij}^{k}\, t_{ij} \leq \theta'\, c(e) && \forall e \\
& \theta' \leq \lambda\theta \\
& u_{ij}^{k} \in \{0,1\} && \forall ijk
\end{aligned}
$$

The Tunnel Limit Extension (TLE)

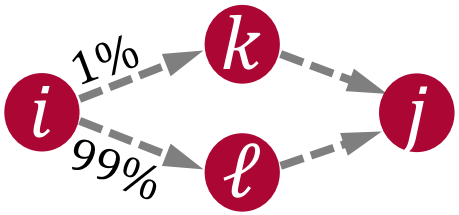


Minimize max. utilization

Keep # of SR tunnels low



Limit split factors



Avoid transit on edge nodes

Step 1: Define nodes to be blacklisted (e.g. '*Edge')

Step 2: Automatically reorder list of nodes:

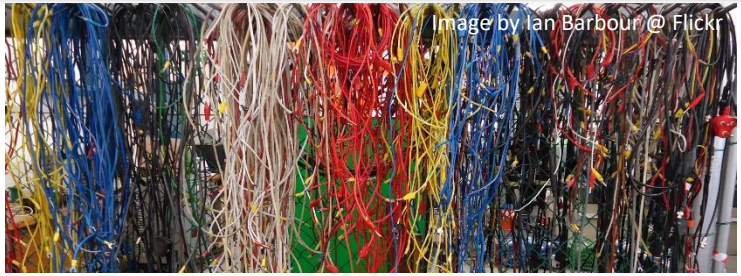
Osnabrück A	Osnabrück A
Osnabrück B	Osnabrück B
Osnabrück Edge	Reykjavik A
Reykjavik A	Reykjavik B
Reykjavik B	Frankfurt A
Reykjavik Edge	Frankfurt B
Frankfurt A	Osnabrück Edge
Frankfurt B	Reykjavik Edge
Frankfurt Edge	Frankfurt Edge

Nodes that may be used as checkpoints

Nodes that must only be source and/or destination

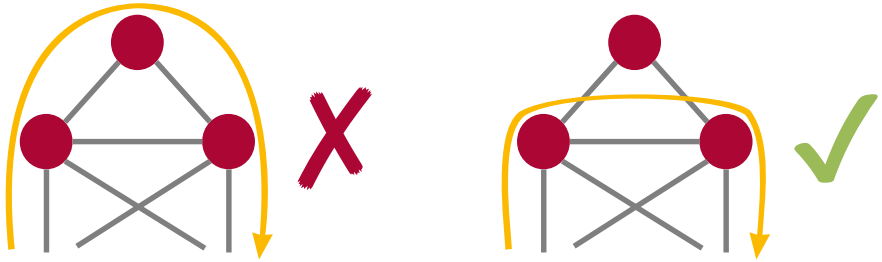
Step 3: Add constraint to optimization: $u_{ij}^{k} = 0 \qquad \forall ijk \mid k \geq B \wedge k \neq j$

With this modular design, any node may be blacklisted!

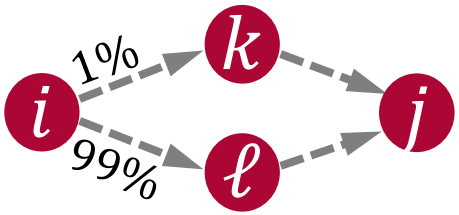


Minimize max. utilization

Keep # of SR tunnels low

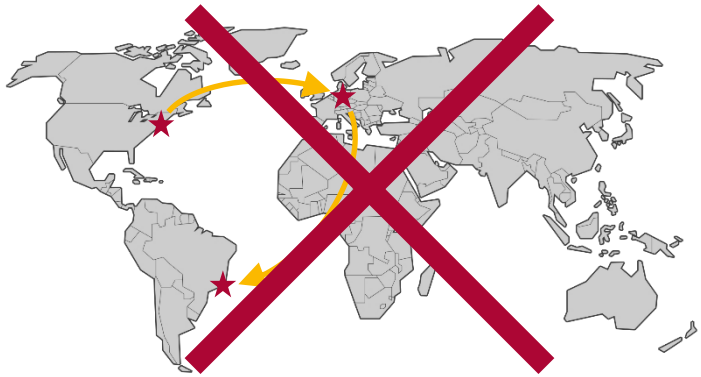


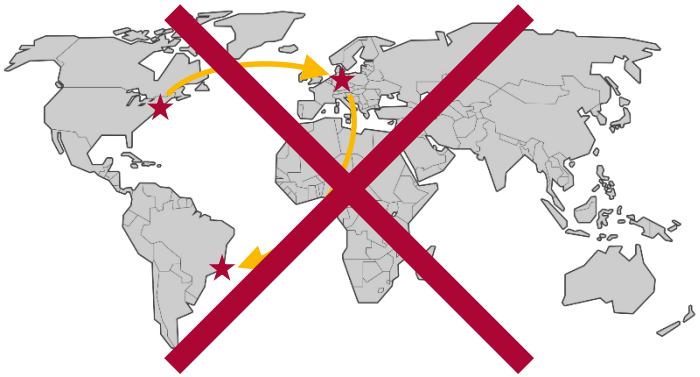
Limit split factors



Avoid transit on edge nodes

Avoid bypassing latency requirements





Intracontinental Delay Constraint:

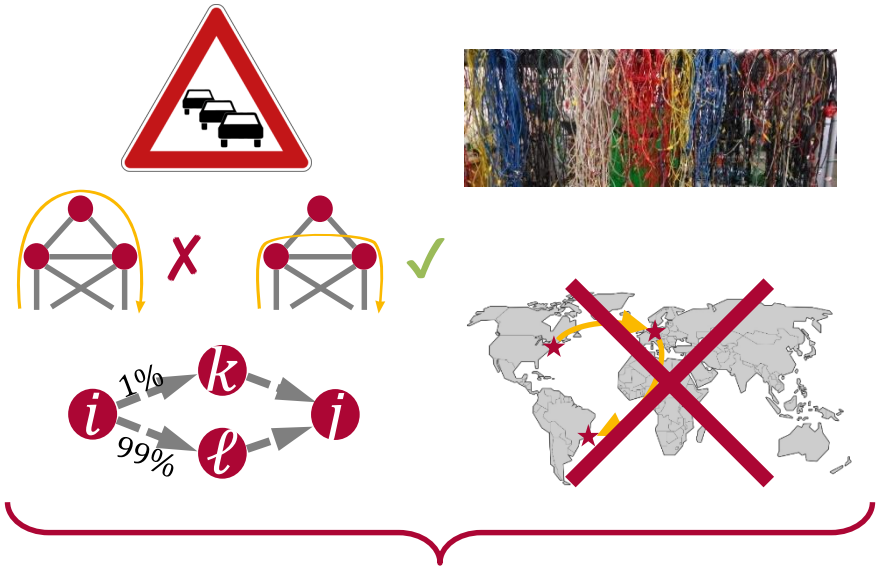
$$u_{ij}^{k} = 0 \qquad \forall ijk \mid continent(i) = continent(j) \wedge continent(k) \neq continent(i)$$

Intranational Delay Constraint:

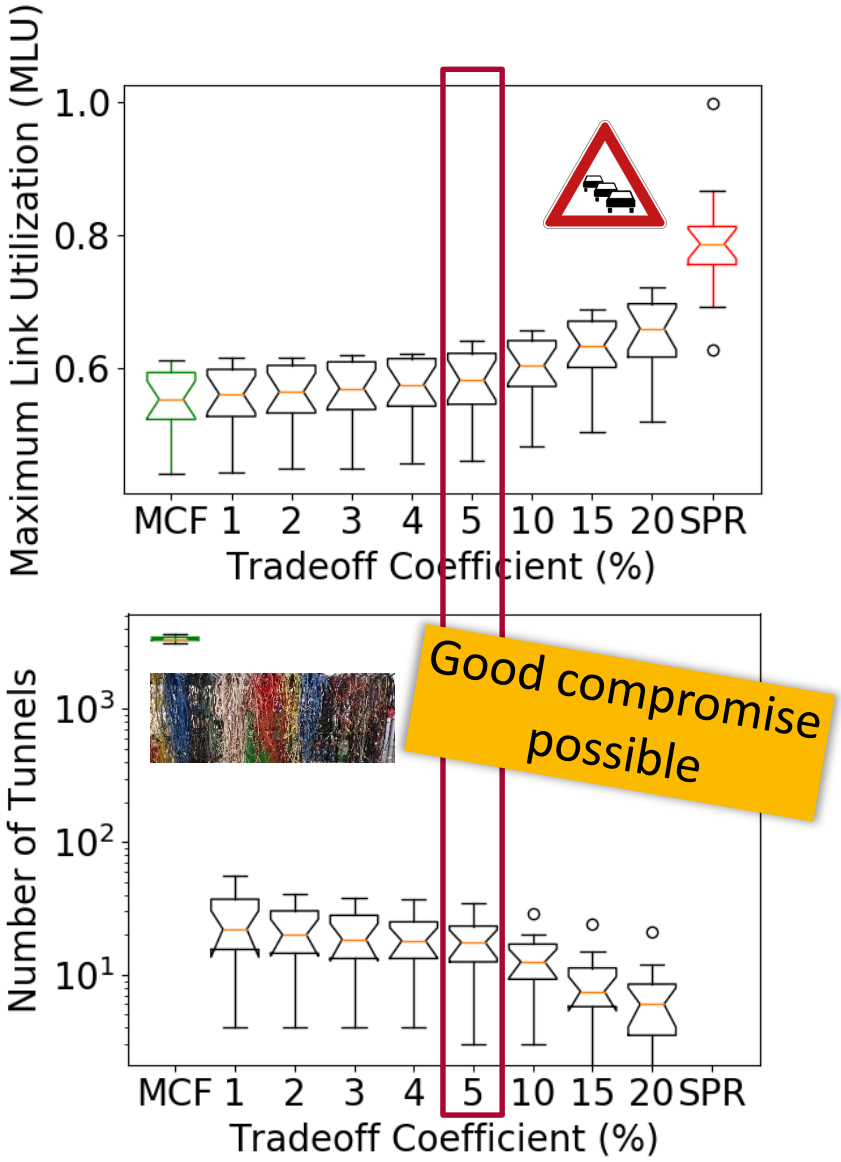
$$u_{ij}^{k} = 0 \qquad \forall ijk \mid nation(i) = nation(j) \wedge nation(k) \neq nation(i)$$

Intra-Site Delay Constraint:

$$u_{ij}^{k} = 0 \qquad \forall ijk \mid site(i) = site(j) \wedge site(k) \neq site(i)$$



$$
\begin{aligned}
&\min\ \theta \\
&\text{s.t.} && \sum_k x_{ij}^k = t_{ij} && \forall ij \\
&&& \sum_{ij}\sum_k g_{ij}^k(e)\, x_{ij}^k \le \theta\, c(e) && \forall e \\
&&& x_{ij}^k = 0 && \forall ijk \mid k \ge B \wedge k \ne j \\
&&& x_{ij}^k = 0 && \forall ijk \mid k \notin T(LCA(i,j)) \\
&&& x_{ij}^k \ge 0 && \forall ijk \\
&\min\ \theta' \ \frac{1}{2\,\lambda\,\theta} + \sum_{k \ne j} u_{ij}^k \\
&\text{s.t.} && \sum_k u_{ij}^k = 1 && \forall ij \\
&&& \sum_{ij}\sum_k g_{ij}^k(e)\, u_{ij}^k\, t_{ij} \le \theta'\, c(e) && \forall e \\
&&& \theta' \le \lambda\theta \\
&&& u_{ij}^k = 0 && \forall ijk \mid k \ge B \wedge k \ne j \\
&&& u_{ij}^k = 0 && \forall ijk \mid k \notin T(LCA(i,j)) \\
&&& u_{ij}^k \in \{0,1\} && \forall ijk
\end{aligned}
$$

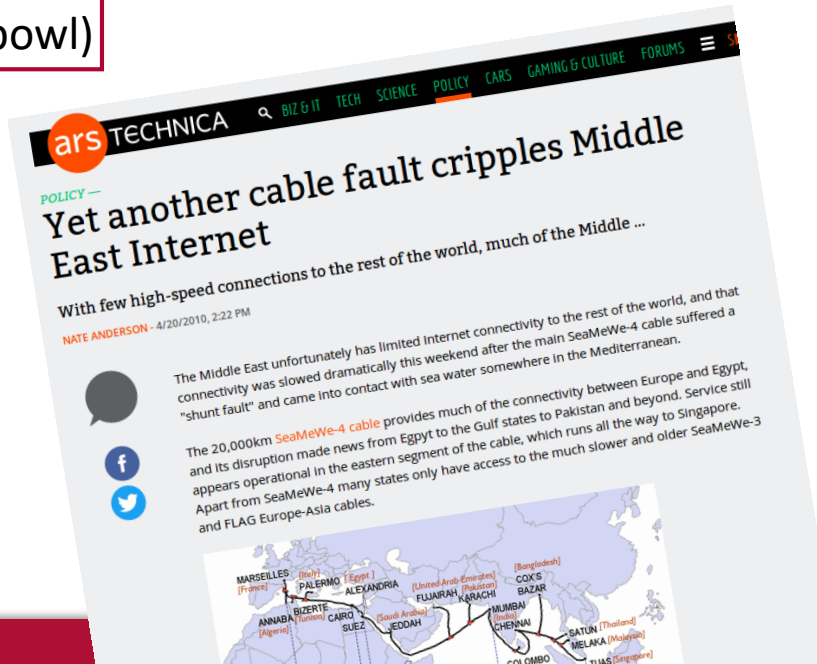


Reconfigure rarely & little

Do we need to re-optimize?

- Time of optimization or deployment of SR tunnels
- Hardware outage
- Irregularities in traffic matrices (e.g. Superbowl)
- True changes of the topology (e.g. adding a new link)

Be resilient against failures


ars TECHNICA

BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE FORUMS

POLICY

Yet another cable fault cripples Middle East Internet

With few high-speed connections to the rest of the world, much of the Middle ...

NATE ANDERSON - 4/20/2010, 2:22 PM

The Middle East unfortunately has limited Internet connectivity to the rest of the world, and that connectivity was slowed dramatically this weekend after the main SeaMeWe-4 cable suffered a "shunt fault" and came into contact with sea water somewhere in the Mediterranean.

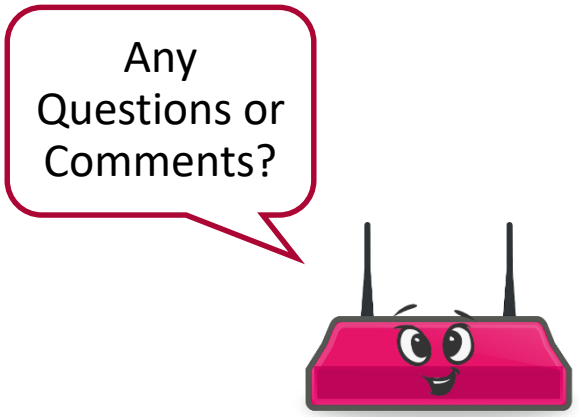
The 20,000km SeaMeWe-4 cable provides much of the connectivity between Europe and Egypt, and its disruption made news from Egpyt to the Gulf states to Pakistan and beyond. Service still appears operational in the eastern segment of the cable, which runs all the way to Singapore. Apart from SeaMeWe-4 many states only have access to the much slower and older SeaMeWe-3 and FLAG Europe-Asia cables.

SR is a powerful Traffic Engineering tool

It has to be carefully fine-tuned towards specific use-case

- Number of SR tunnels
- Traffic splitting
- Special nodes
- Latency
- etc.

Details on this work can be found in: Schüller et al., "*Traffic Engineering using Segment Routing and Considering Requirements of a Carrier IP Network*", IEEE/ACM Transactions on Networking, vol. 26, no. 4, pp. 1851-1864, 2018



Time's up? Find me at
https://sys.cs.uos.de/schueller
or contact me directly:
tschueller@uos.de