Metric-Based Traffic Engineering: Panacea or Snake Oil? A Real World Study.

Arman Maghbouleh
am@cariden.com

NANOG 27
Phoenix, Arizona

Where is Reality?

Fortz et al.

"[W]e can find [OSPF] weight settings …[that] get within a few percent of the best possible with general routing,including MPLS."

– (IEEE 2002)

Lorenz et al.

"Source invariant routing can be significantly worse than than per-flow routing."

– (DIMACS 2001)

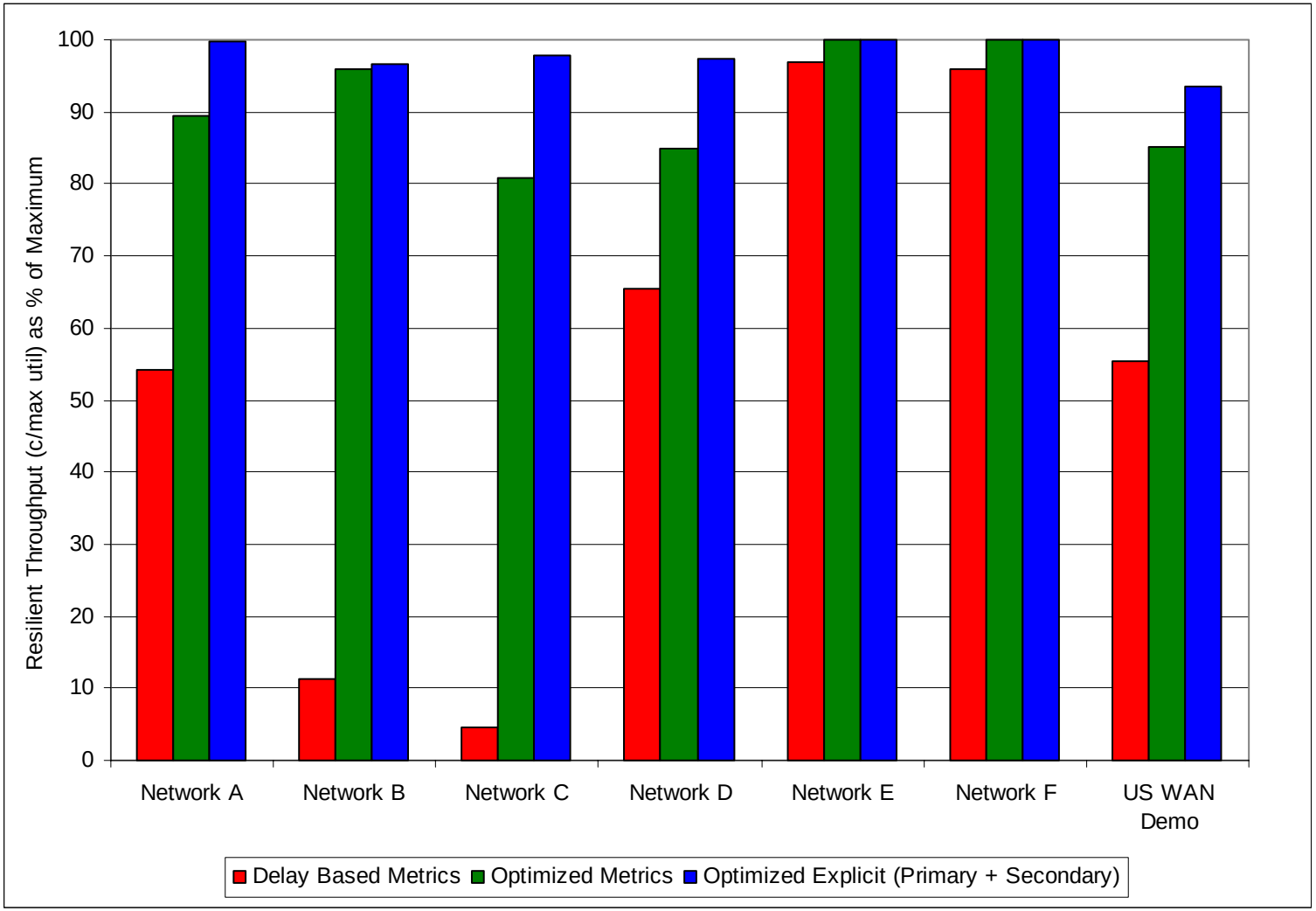
"[W]eight setting for OSPF cannot replace MPLS as a traffic engineering tool."

– (IETF-RR list 2001)

This Talk

- Six real networks
- Minimize Maximum Utilization
  - UNDER ALL POSSIBLE SINGLE-CIRCUIT FAILURES
- Compare
  - Theoretical-Optimal
  - Optimized Explicit Routing
  - Optimized Metrics
  - Delay-based Metrics

Results (Preview)



Conclusions (Preview)

- Metric optimization close to theoretical optimal
- Limitations are real
  - … but didn't affect the bottom-line significantly
- TE is trivial for some topologies
- Speculation: Effects of limitations is limited because operators already design networks with OSPF limitations in mind

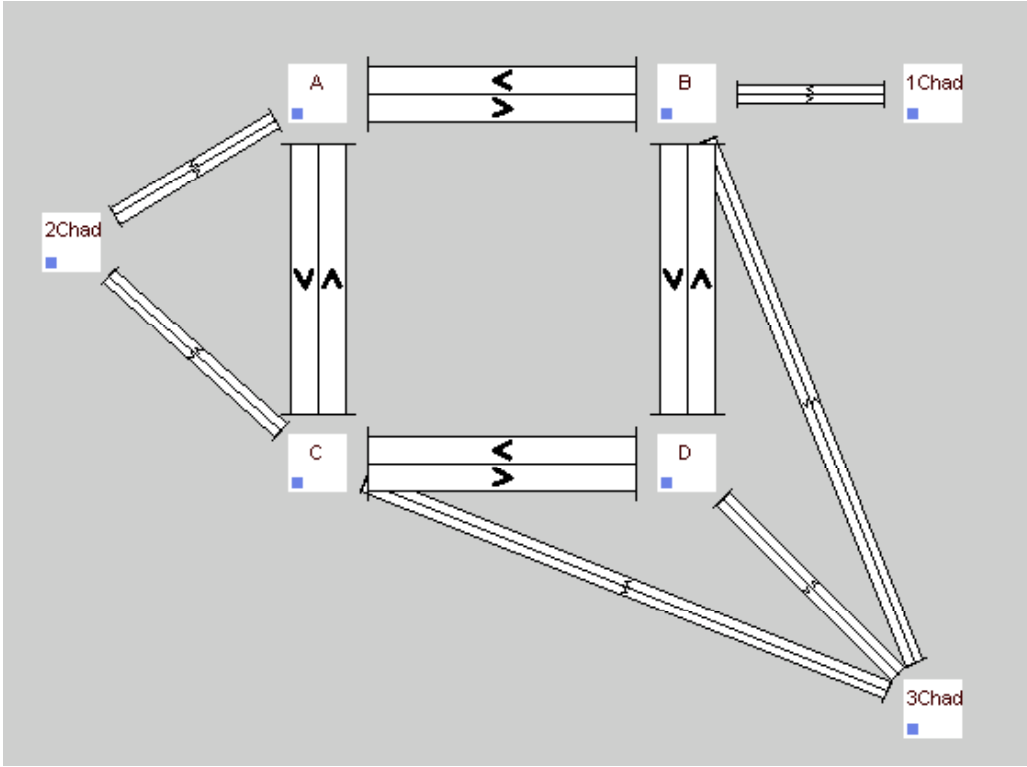
Outline

- Introduction
  - Networks
  - Routing Models
- Results & Analysis
  - Numerical Results
  - Examples of Limitations
- MPLS Notes
- Conclusion

Networks

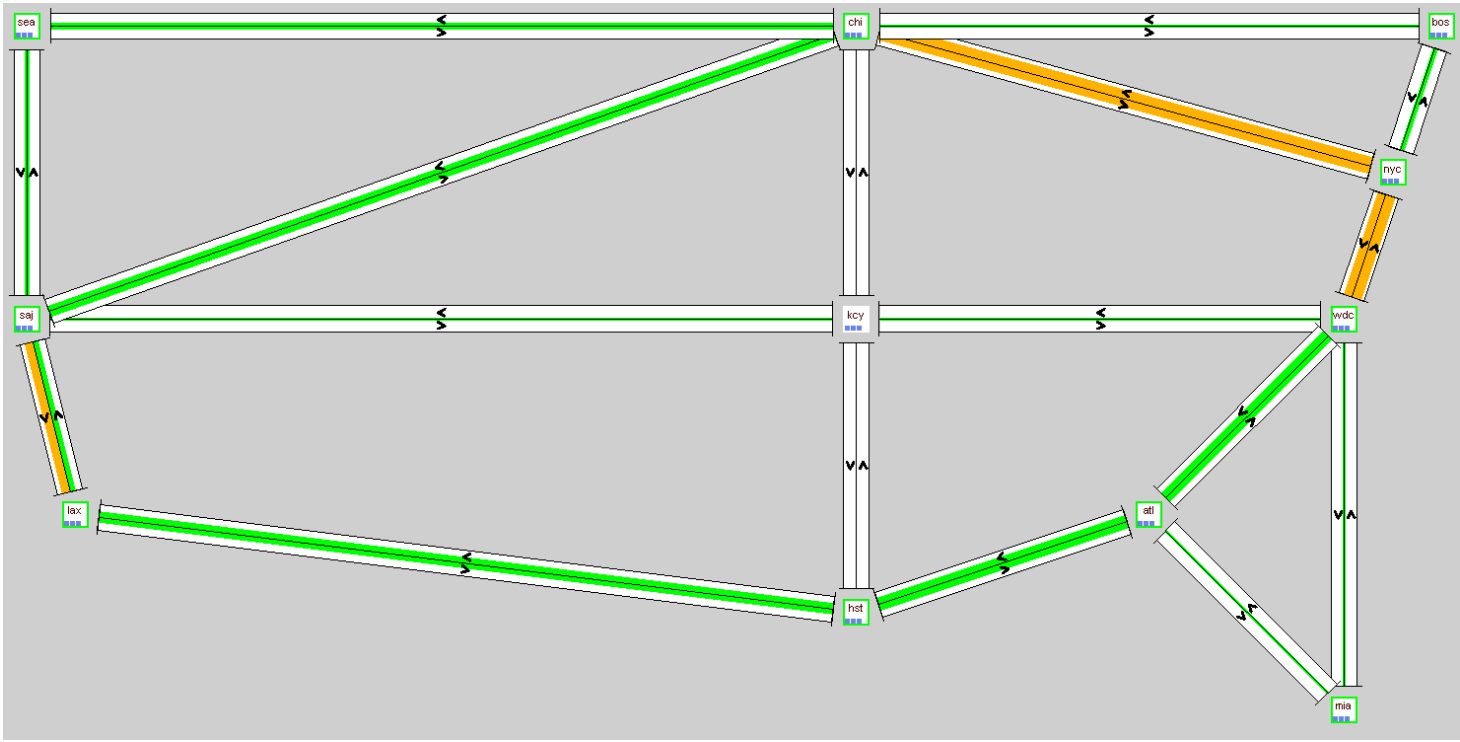
- Tier 1, tier 2, content-delivery network
- Global, U.S., Europe
- Some already deployed MPLS
  (that is, measured traffic matrix versus estimated traffic matrix)
- Five operational, one proposed
- Topologies
  - Modified Hub & Spoke **(see representative plot)**
  - Typical U.S. Meshes **(see representative plot)**
  - Global Mesh **(not reproduced for reasons of confidentiality)**

Modified Hub-&-Spoke



- High capacity simple core
- Peripheral nodes connected
  - Singly, doubly, and infrequently triply

Typical U.S. Backbone



- Three+ paths across country
- Elephants and mice demands

Plot Legend

- White squares represent sites (PoPs)
- Small blue squares represent routers
- Lines are physical links
- Thickness represents capacity
- Color & fill thickness represents utilization
  - (red >90%, orange >75% failure)
- Blue arrows represent paths
  - (solid for normal, dashed for failure)
- X represent failure locations

Global Meshes

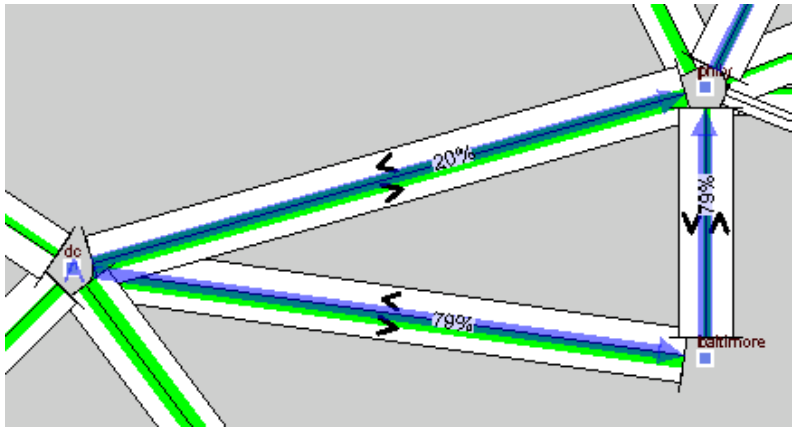
- No prototype shown for reasons of confidentiality
- Combinations of meshes, rings,…
  - Topology bottlenecks across oceans
- Large range of capacities
  - (e.g. OC-3 to OC-192)

Outline

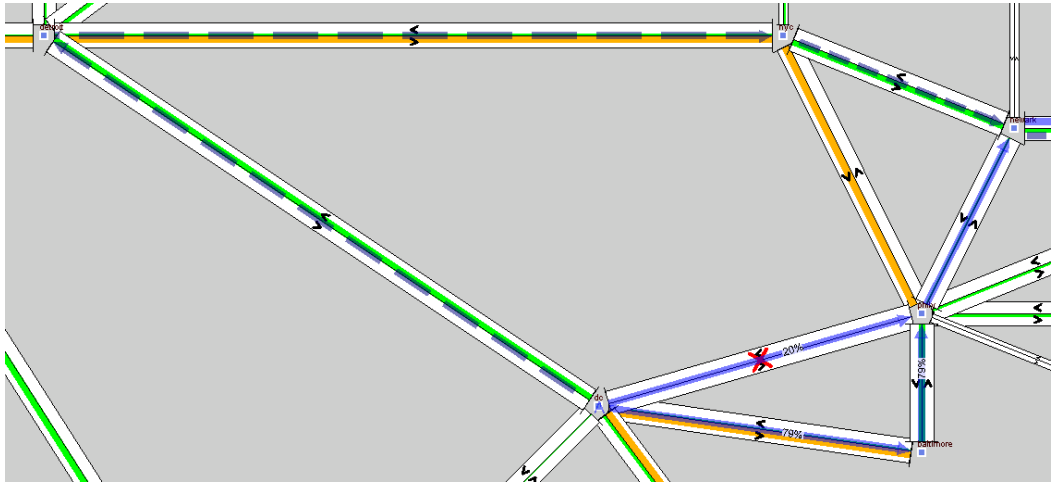
- Introduction
  - Networks
  - Routing Models
- Results & Analysis
  - Numerical Results
  - Examples of Limitations
- MPLS Notes
- Conclusion

Theoretical Optimal

- Result of multicommodity flow optimizations
- No shortest-path limitation
  - I.e., possibly source-based routing



Arbitrary Splits of Demands

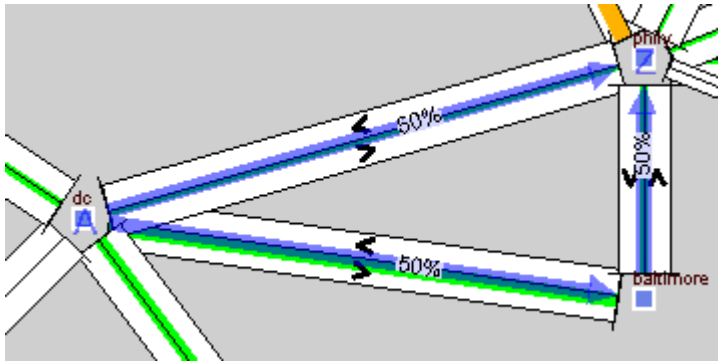


Routing changed on failure

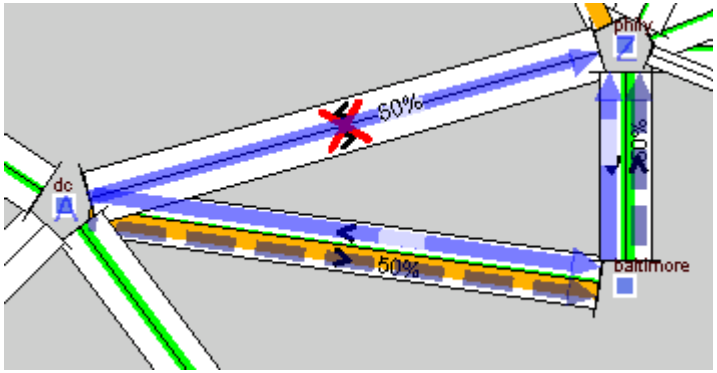
*Real case used with permission.

Shortest-Path Metric Routing

- OSPF, IS-IS
- 1/n Equal-Cost Multipath
- Single set of metrics for all failures



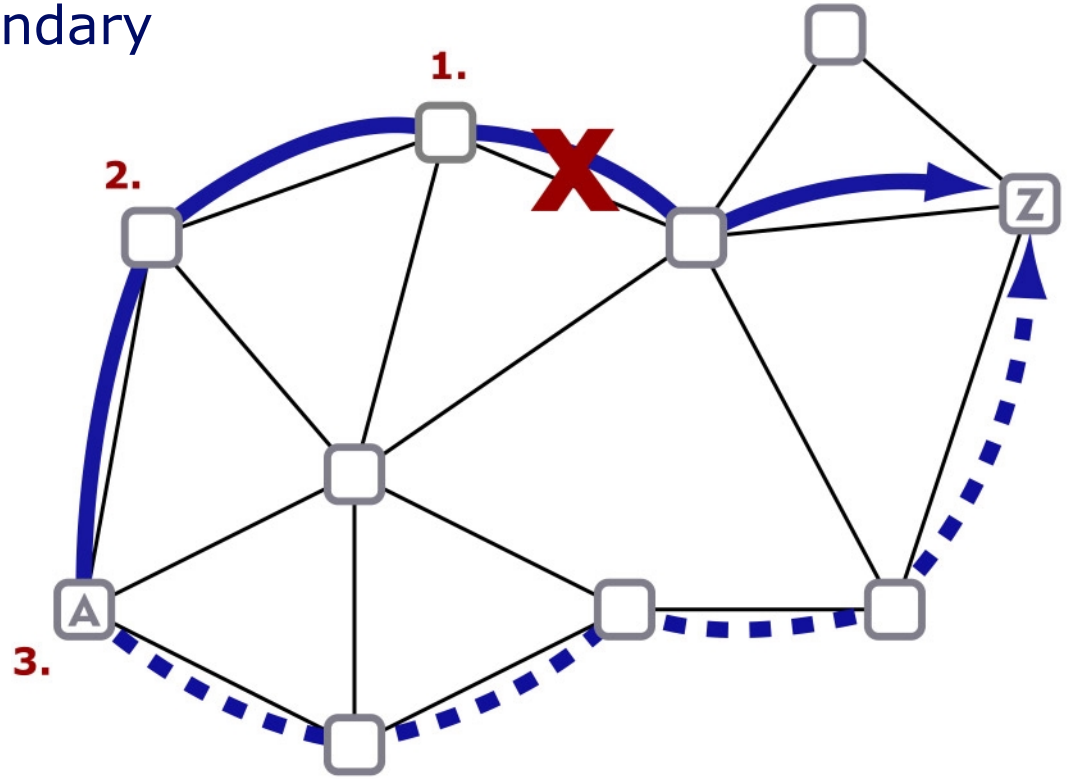
Equal splits on ECMP



Metrics not change after failure

Explicit Routing

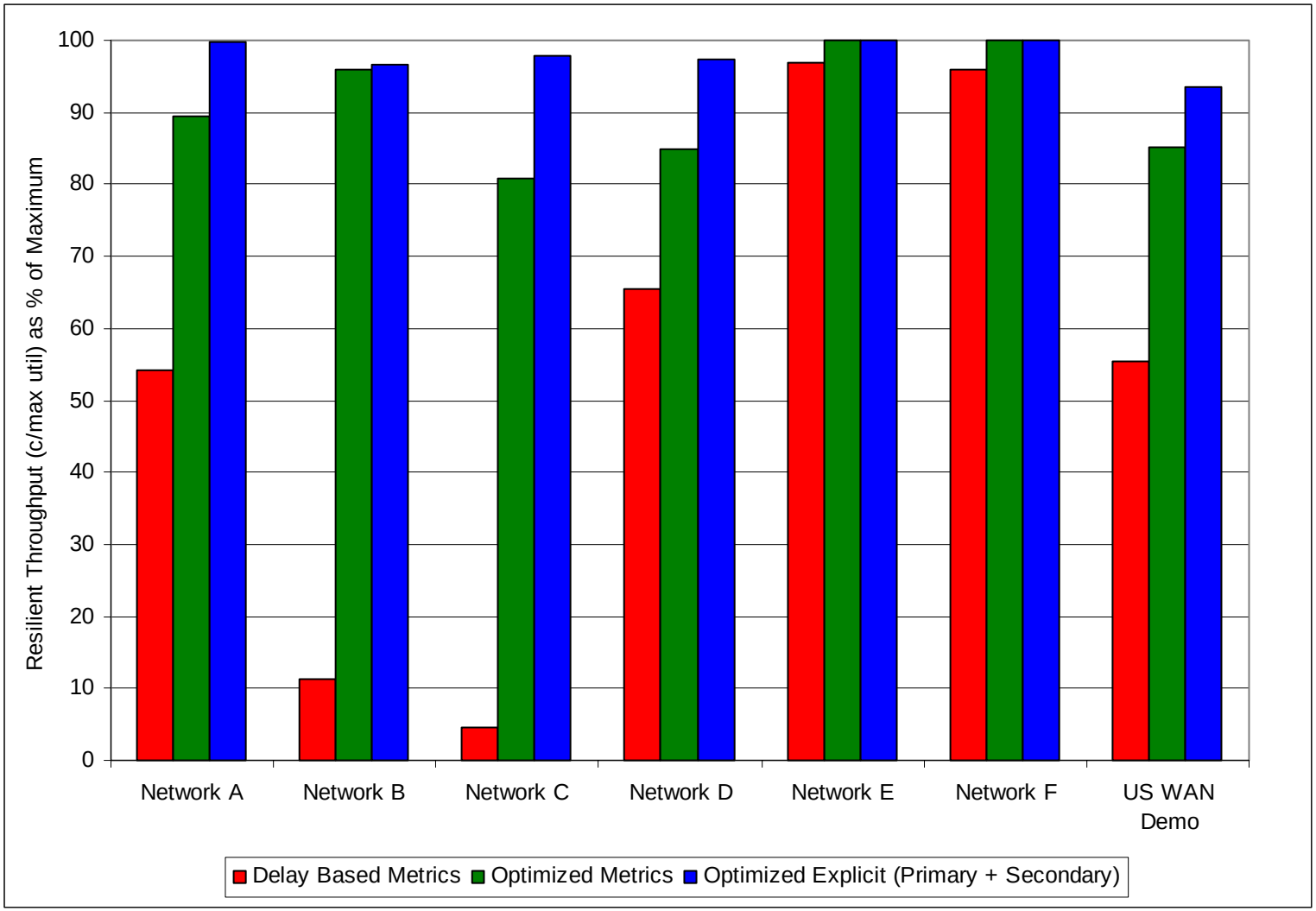
- A primary and secondary path for each source-destination pair
  - Link-diverse secondary



Outline

- Introduction
  - Networks
  - Routing Models
- Results & Analysis
  - Numerical Results
  - Examples of Limitations
- MPLS Notes
- Conclusion

Results


Resilient Throughput (c/max util) as % of Maximum
100
90
80
70
60
50
40
30
20
10
0
Network A
Network B
Network C
Network D
Network E
Network F
US WAN Demo
Delay Based Metrics
Optimized Metrics
Optimized Explicit (Primary + Secondary)

Results (in text)

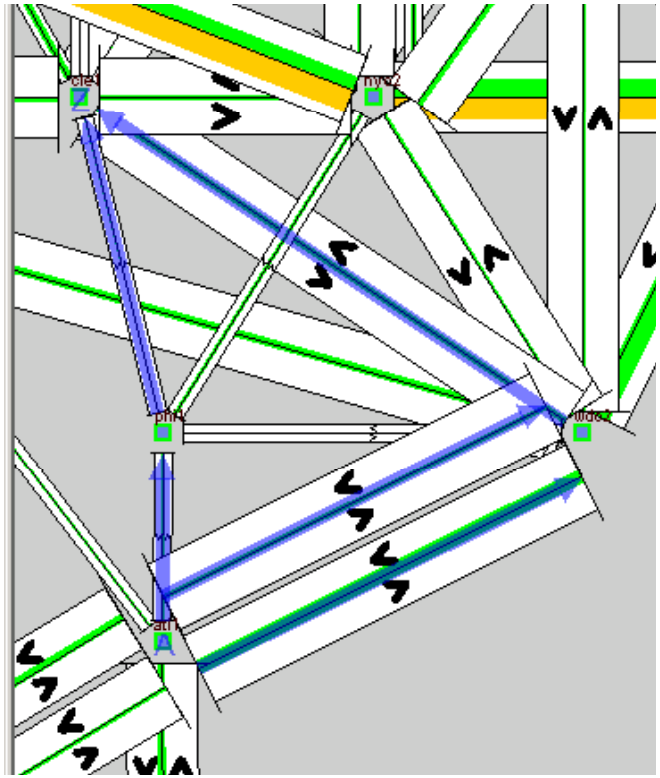
- Can optimize SPF metrics within 80%-95% of maximum theoretical efficiency
- … trivially at 100% for simple topologies
- Explicit routing around 95%
- Metrics based on delay perform poorly with wide range of capacities

Analysis

- ECMP 1/n split inappropriate
  - Parallel links different capacities
  - Approximately parallel links
    - See the “Dissimilar Parallels” and “Tri-ECMP” examples
- One set of metrics limits failure response
  - See the “Escape Failure” example

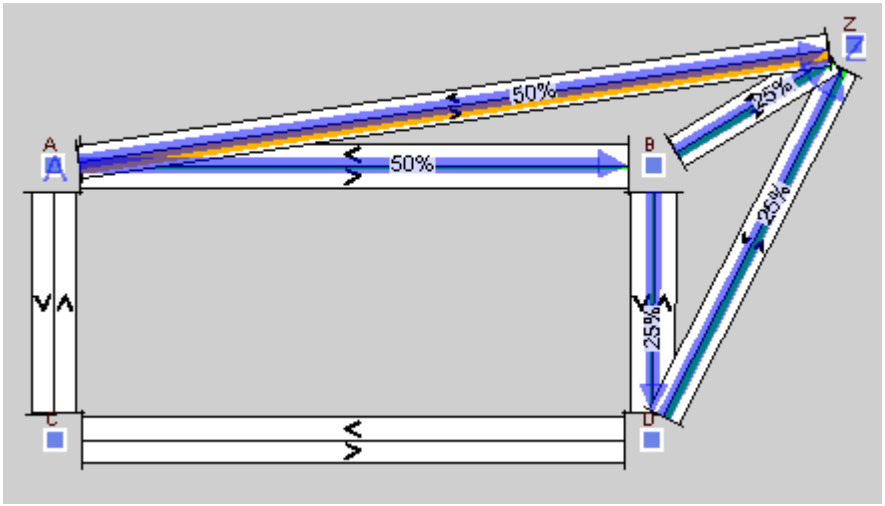
"Dissimilar Parallels" Example

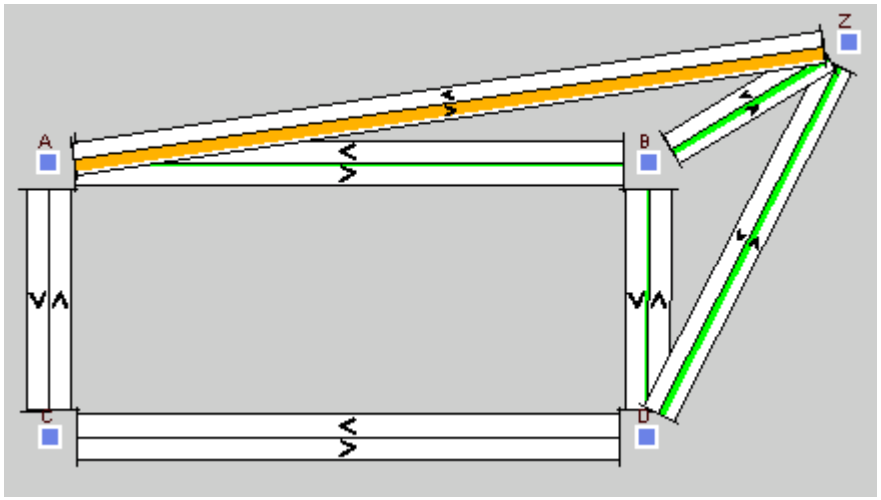
- ECMP for
  ATL-DC-CLE (2xOC-48)
  ATL-Philly-CLE (OC-12)
- If bottleneck region, 1/n split not advisable
- In practice, not see OC-12 in parallel with OC-48 in bottleneck regions



*Real case used with permission.

Tri-ECMP Example

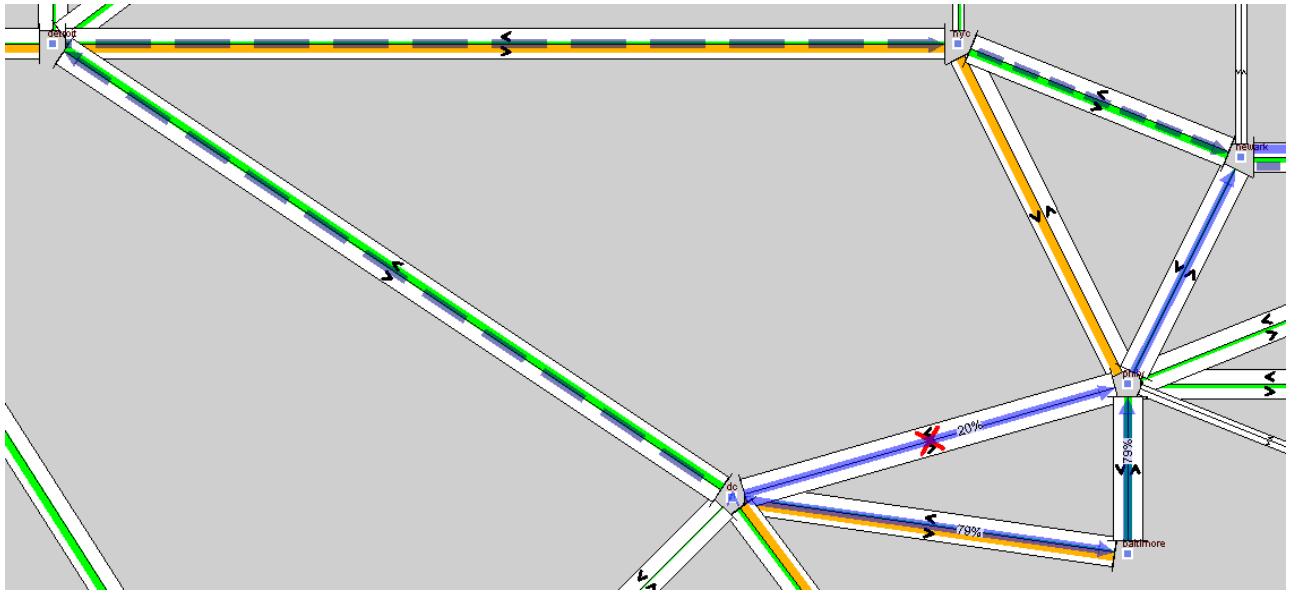




- 50/50 split at source
  the second 50% then split in 2
- Top link overloaded
- Would want 1/3, 2/3 split at source
- 1/n ECMP results in uneven load-balancing

*Example from Lorenz et al. 2001. Not from a real network.

"Escape Failure" Example



- Plot shows theoretically optimal
- Bottleneck in DC-Philly-Baltimore triangle
- Want load-balancing under normal
- Need to get out of congested region when link is down
- Can't do that with OSPF
  - All traffic would've moved to remaining ECMP

*Real case used with permission.

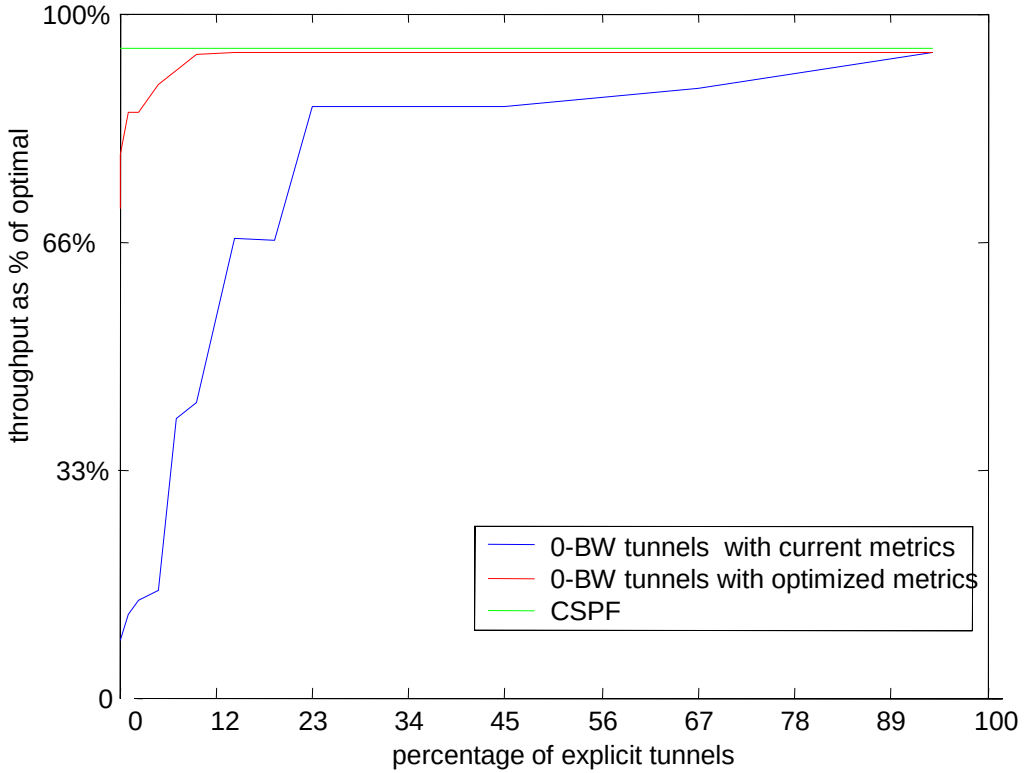
Outline

- Introduction
  - Networks
  - Routing Models
- Results & Analysis
  - Numerical Results
  - Examples of Limitations
- MPLS Notes
- Conclusion

MPLS Bashing?

- No.
- MPLS provides multiple routing options
  - Constraint-Based Routing
    - Very efficient (in our experience)
    - May be an operational pain (non-deterministic etc.)
  - LDP, RSVP with 0-bandwidth are SPF
    - Can take advantage of scalability of metric-based TE
  - Explicit Routing
    - Offline strategic, online tactical
  - Hybrids
- Other Features
  - Traffic matrix available out-of-the-box
  - Hot standby, Fast reroute

Hybrid: SPF-Explicit



X-axis tunnels were explicitly routed from largest to smallest.

- Metric optimization + explicit routes as needed

  "We expect this is not an unreasonable approach."
  -Randy Bush
  - **Also: Ben-Ameur et al. France Telecom, draft-wang-te-hybrid-approach-00.txt**
- Few tunnels explicit if start with good metrics

***Real case used with permission.**

Outline

- Introduction
  - Networks
  - Routing Models
- Results & Analysis
  - Numerical Results
  - Examples of Limitations
- MPLS Notes
- Conclusion

Conclusions (Preview)

- Metric optimization close to theoretical optimal
- Limitations are real
  - … but didn't affect the bottom-line significantly
- TE is trivial for some topologies
- Speculation: Effects of limitations is limited because operators already design networks with OSPF limitations in mind

References

- B. Fortz, J. Rexford, and M. Thorup, "Traffic Engineering With Traditional IP Routing Protocols" in IEEE Communications Magazine, October 2002.
- D. Lorenz, A. Ordi, D. Raz, and Y. Shavitt, "How good can IP routing be?", DIMACS Technical Report 2001-17, May 2001.
- Cariden "IGP Traffic Engineering Case Study", Cariden Technologies, Inc., October 2002.
- B. Fortz and M. Thorup, "Internet traffic engineering by optimizing OSPF weights" in Proceedings of IEEE INFOCOM, March 2000.
- B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS weights in a changing world" IEEE Journal on Selected Areas in Communications, volume 20, pp. 756-767, May 2002.
- L. S. Buriol, M. G. C. Resende, C. C. Ribeiro, and M. Thorup, "A memetic algorithm for OSPF routing" in Proceedings of the 6th INFORMS Telecom, pp. 187188, 2002.
- M. Ericsson, M. Resende, and P. Pardalos, "A genetic algorithm for the weight setting problem in OSPF routing" J. Combinatorial Optimization, volume 6, no. 3, pp. 299-333, 2002.
- W. Ben Ameur, N. Michel, E. Gourdin et B. Liau. Routing strategies for IP networks. Telektronikk, 2/3, pp 145-158, 2001.

Acknowledgments

- Andrew Lange (Cable & Wireless/Exodus)
- Thomas Telkamp (Global Crossing)
- Clarence Filsfils (Cisco)
- Jennifer Rexford (AT&T Labs)