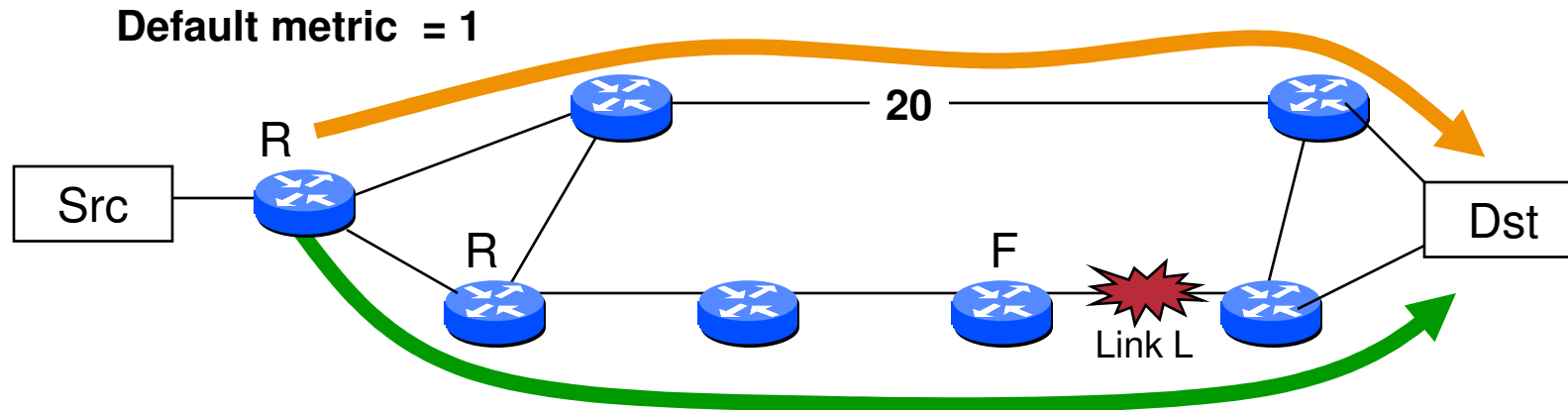


# **Fast IGP Convergence**

**John Evans – [joevans@cisco.com](mailto:joevans@cisco.com)**

# Convergence



- Assume a flow from Src to Dest
- T1: when L dies, the best path is impacted
  - loss of traffic
- T2: when the network converges, a next best path is computed
  - traffic reaches the destination again
- Loss of Connectivity: T2 – T1, called “convergence” hereafter
- Analyzed for streams going to IGP and BGP learned prefixes

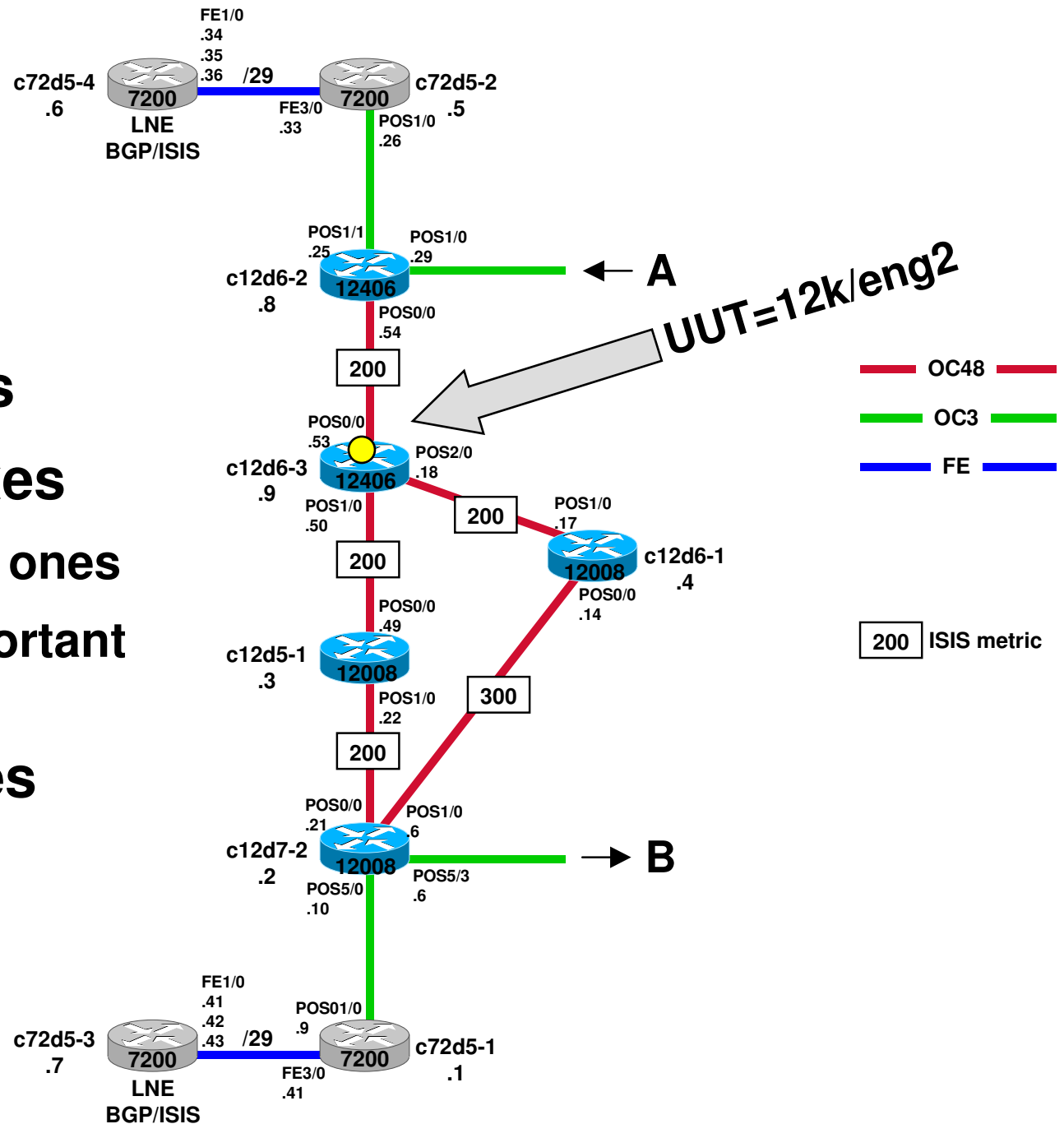
# Objective

---

- **Sub-second for**
  - the first 500 IGP Prefixes
  - all BGP prefixes whose next-hop is within the first 500 IGP prefixes assuming the BGP routes are stable
- **IGP: ISIS**
  - also applicable to OSPF

# Lab Setup

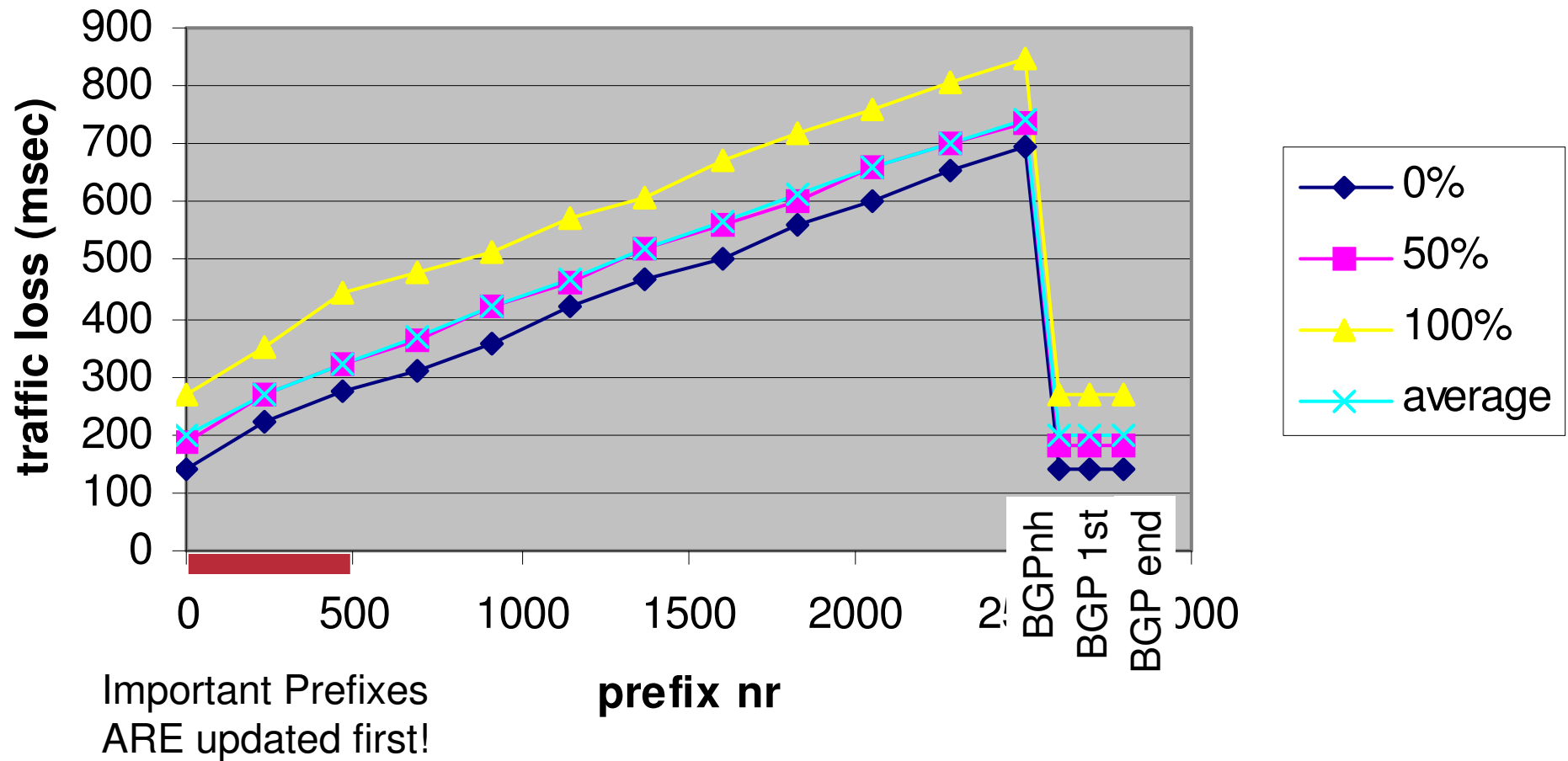
- Pre 12.0(27)S
- 1000 ISIS nodes
- 2500 ISIS prefixes
  - 500 important ones
  - 2000 non-important ones
- 160k BGP routes
- No flap
- POS



# Remote noLB – ISIS

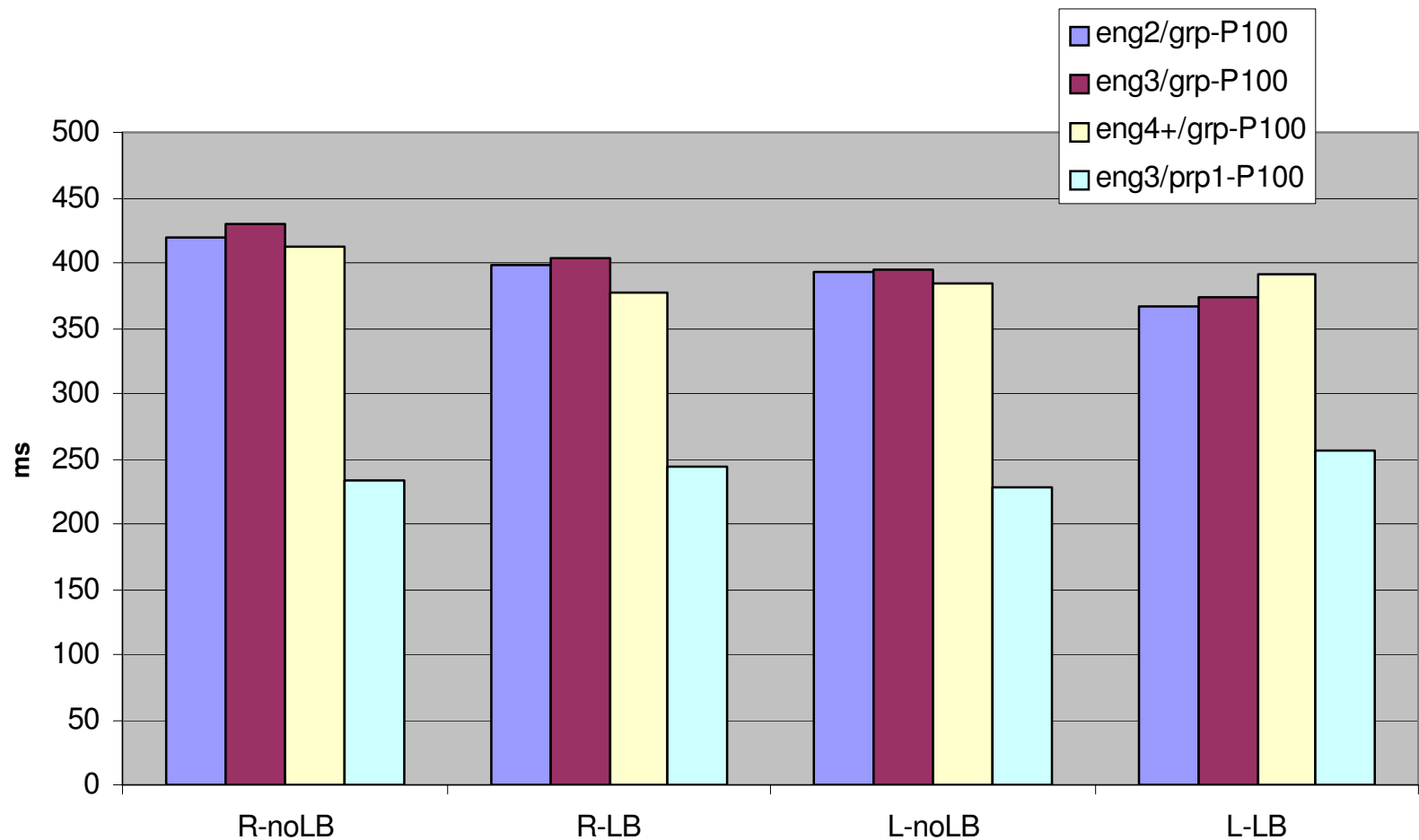
c12k- -eng2- -pr50-lc50-ipc20-bgp160-remote-nolb

Agilent measurements



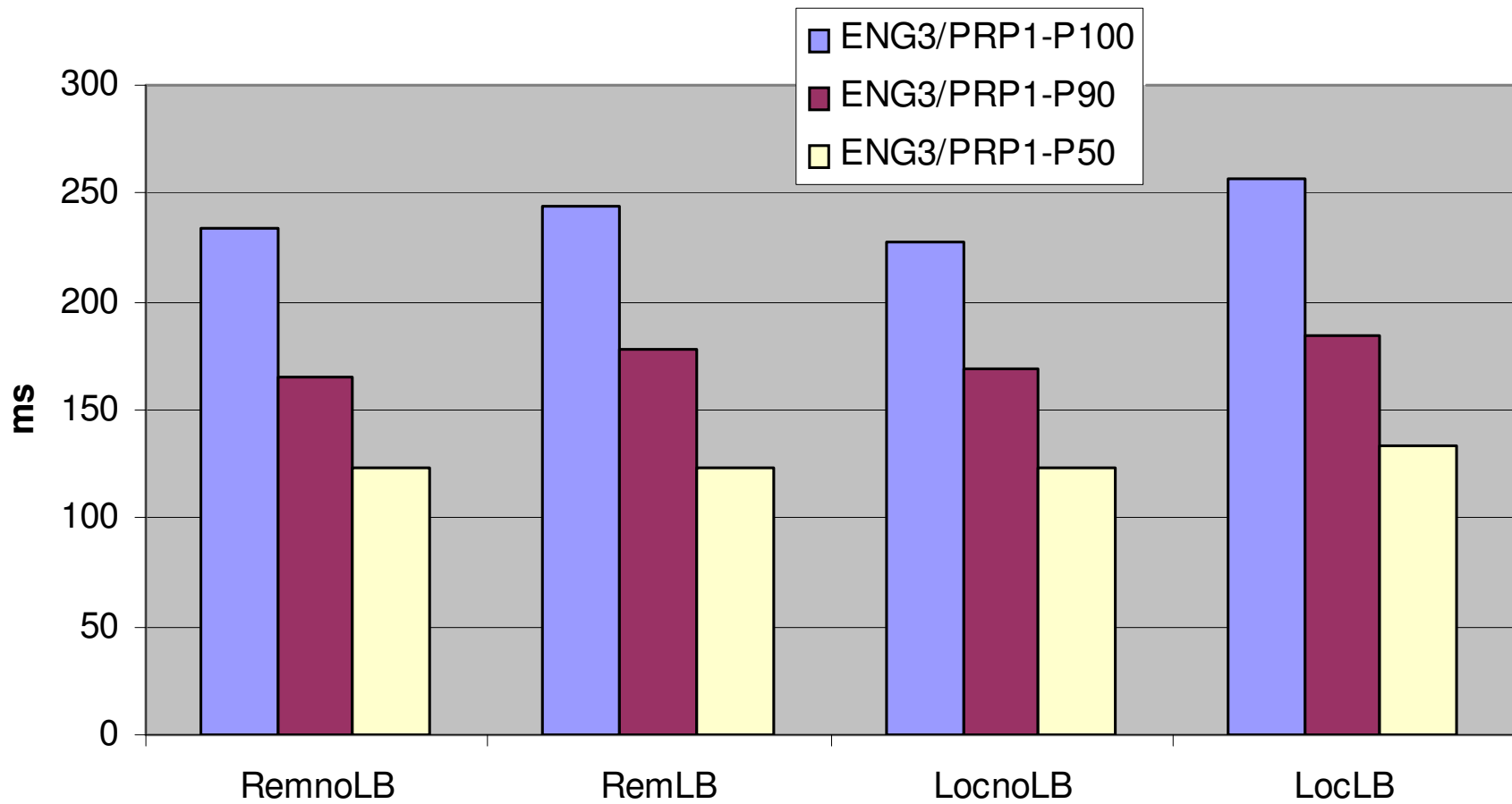
# Black-Box measurements: Max(Pref #500)

Worst-Case Convergence for prefix #500 for 100 iterations

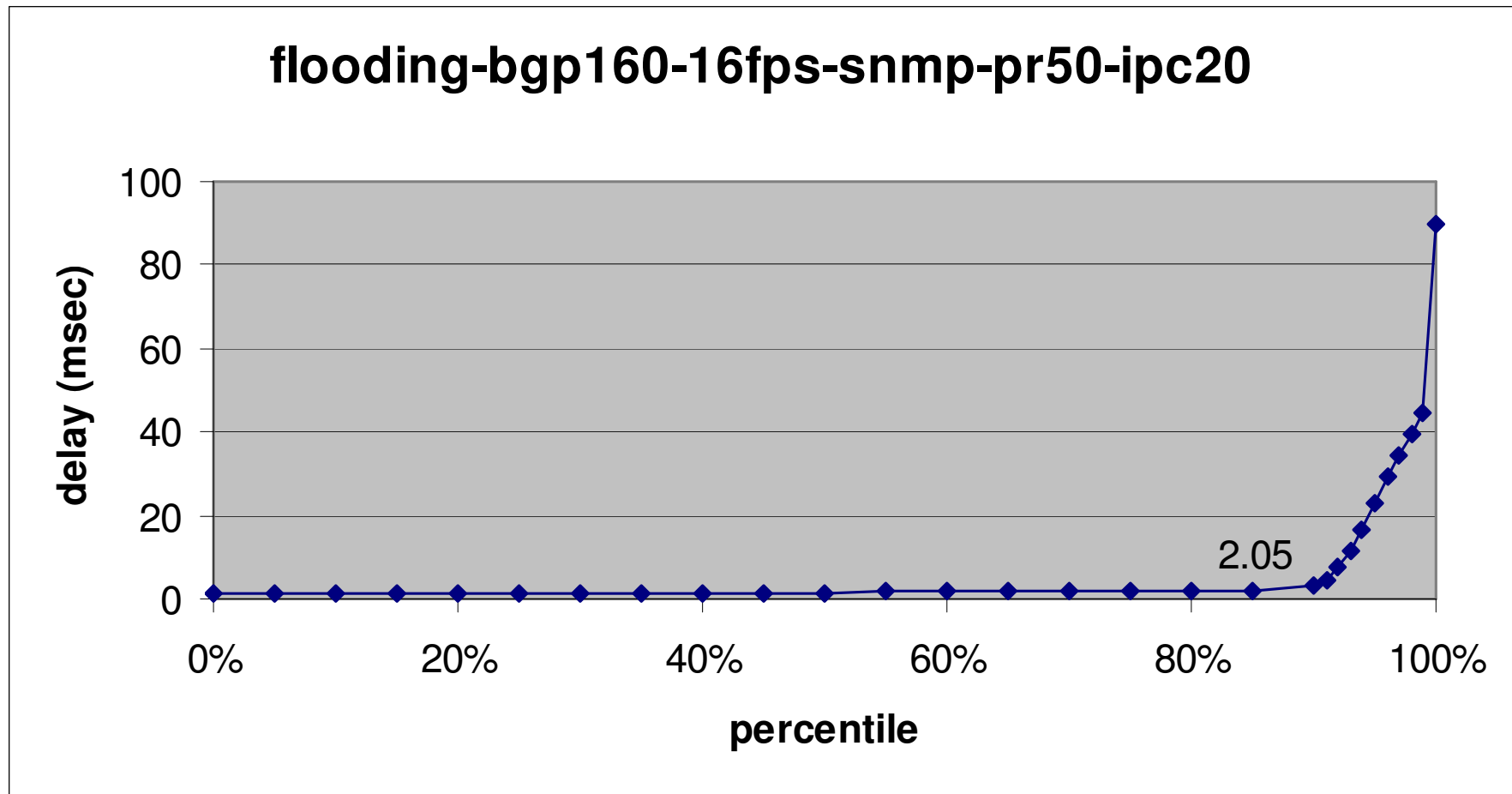


# Max vs average for 500 first prefixes

Worst-case vs Average - Eng3/PRP1 - 100 iterations



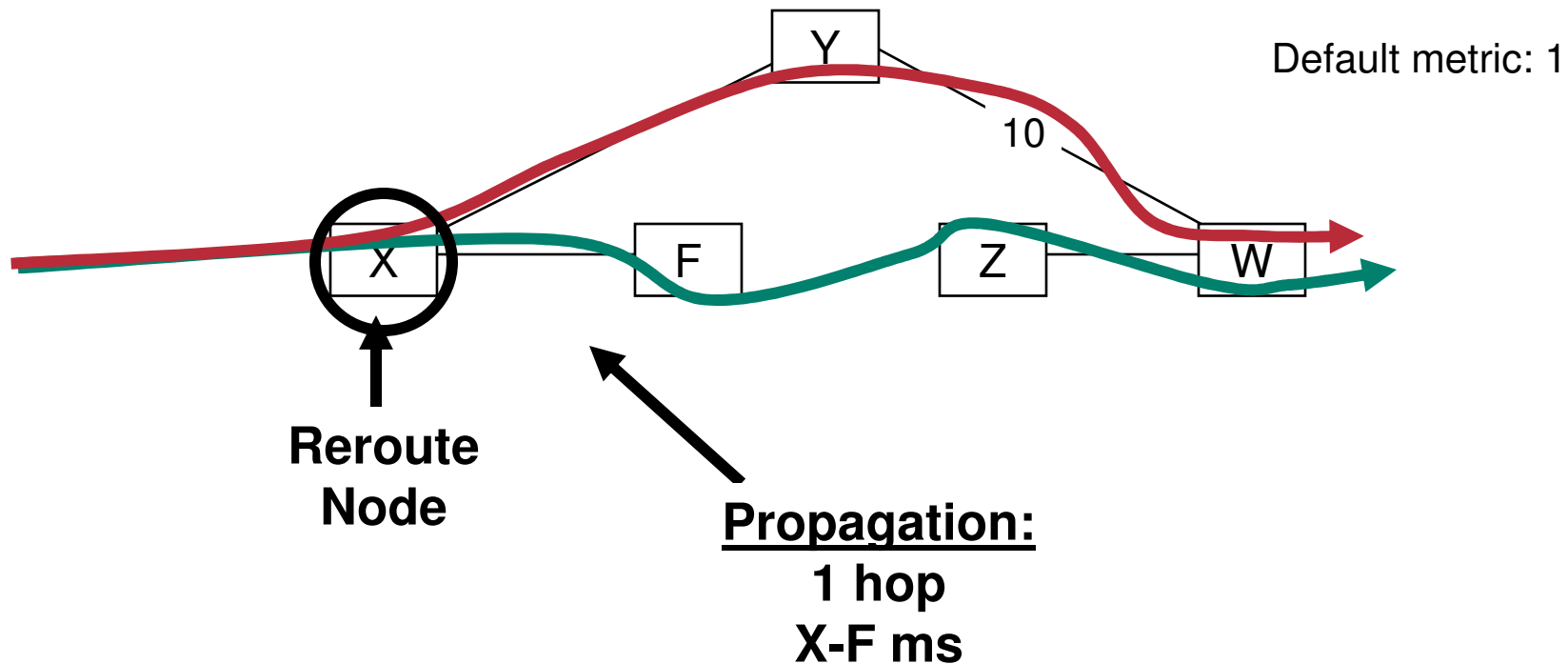
# Flooding impact



- **Flooding occurs before SPF**

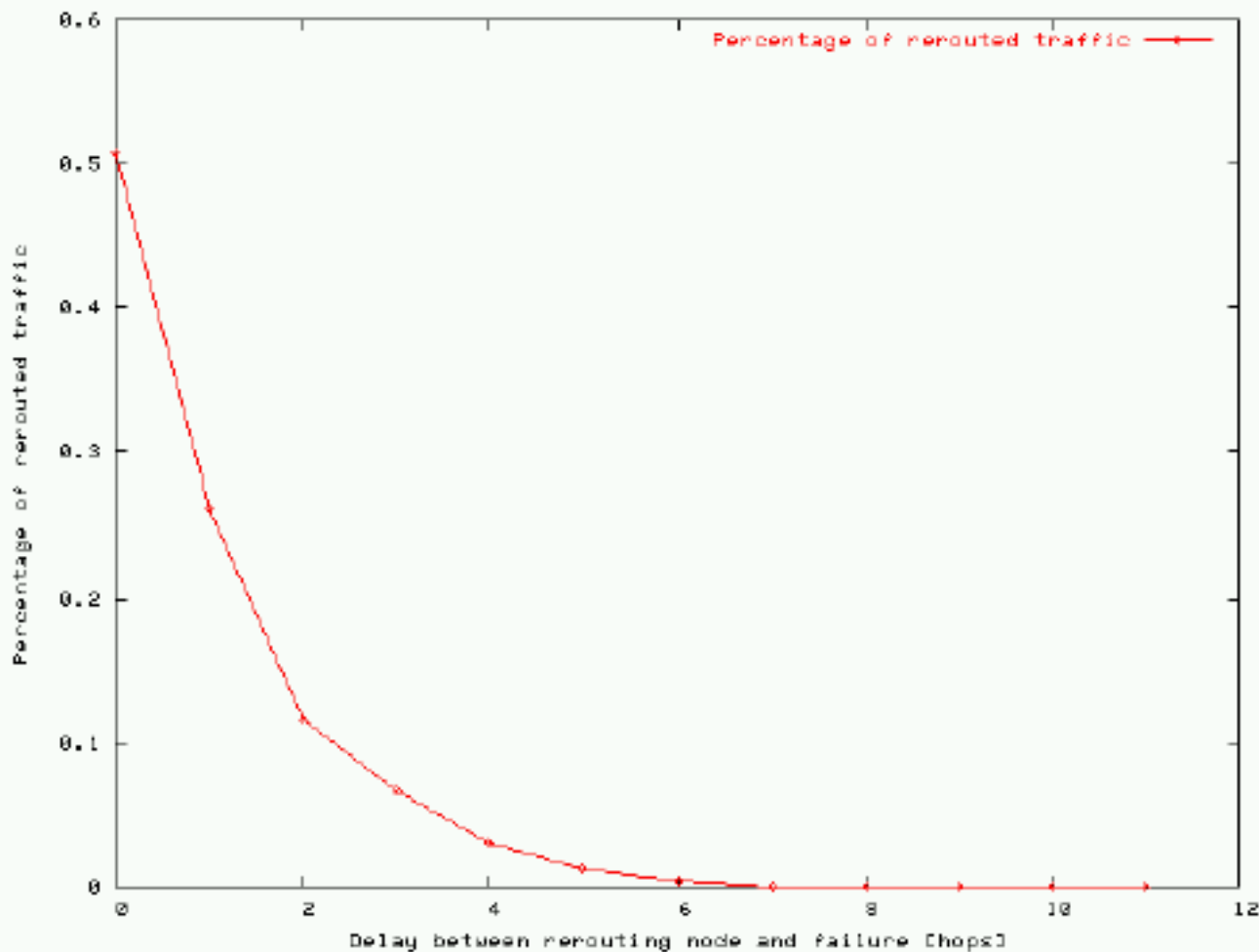


# Propagation distance - analysis



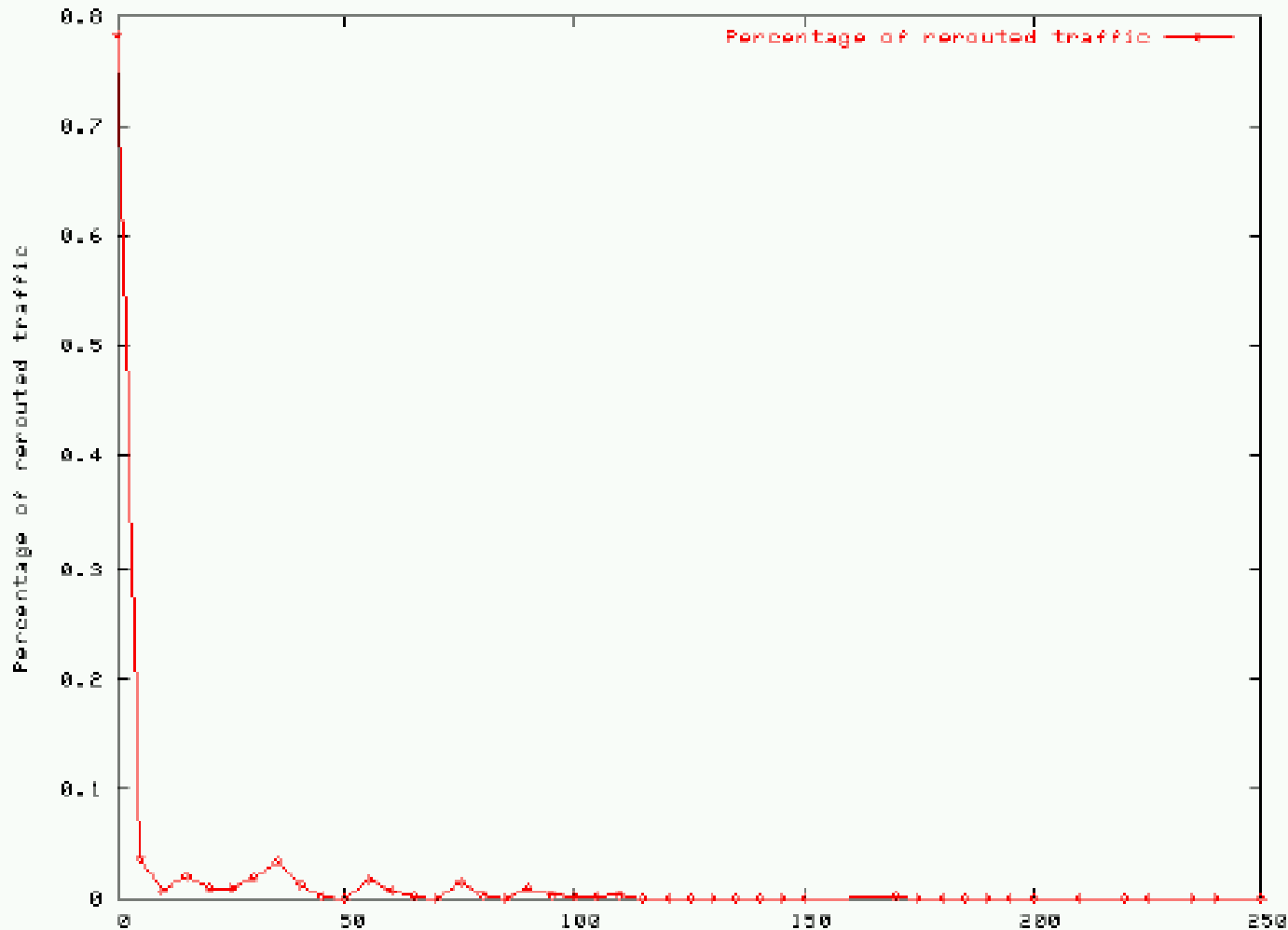
- **R: point where the old and new paths diverge**
  - this is a worst-case estimation of P!

# P: Propagation in number of hops



Worldwide ISP with traffic matrix – summary for the failures of the 340 most loaded links. Pessimistic definition of R

# P: Propagation in ms (light speed)



Worldwide ISP with traffic matrix – summary for the failures of the 340 most loaded links. Pessimistic definition of R

# Conclusion

---

- **Sub-Second objective is realistic**
  - conservative
- **Technology has significantly improved**

**Why is it possible?**

# Components contributing to loss of connectivity

- **D: Failure is detected**
- **O: New LSP is originated**
- **QSP: cumulative queueing, serialization, propagation**
- **h\*F: LSP is flooded up to rerouting node**
- **SPT: SPT is updated**
- **RIB: RIB/FIB is updated**
- **DD: LC's are updated**
- **BGP recursion is fixed**

**LoC(p) =**

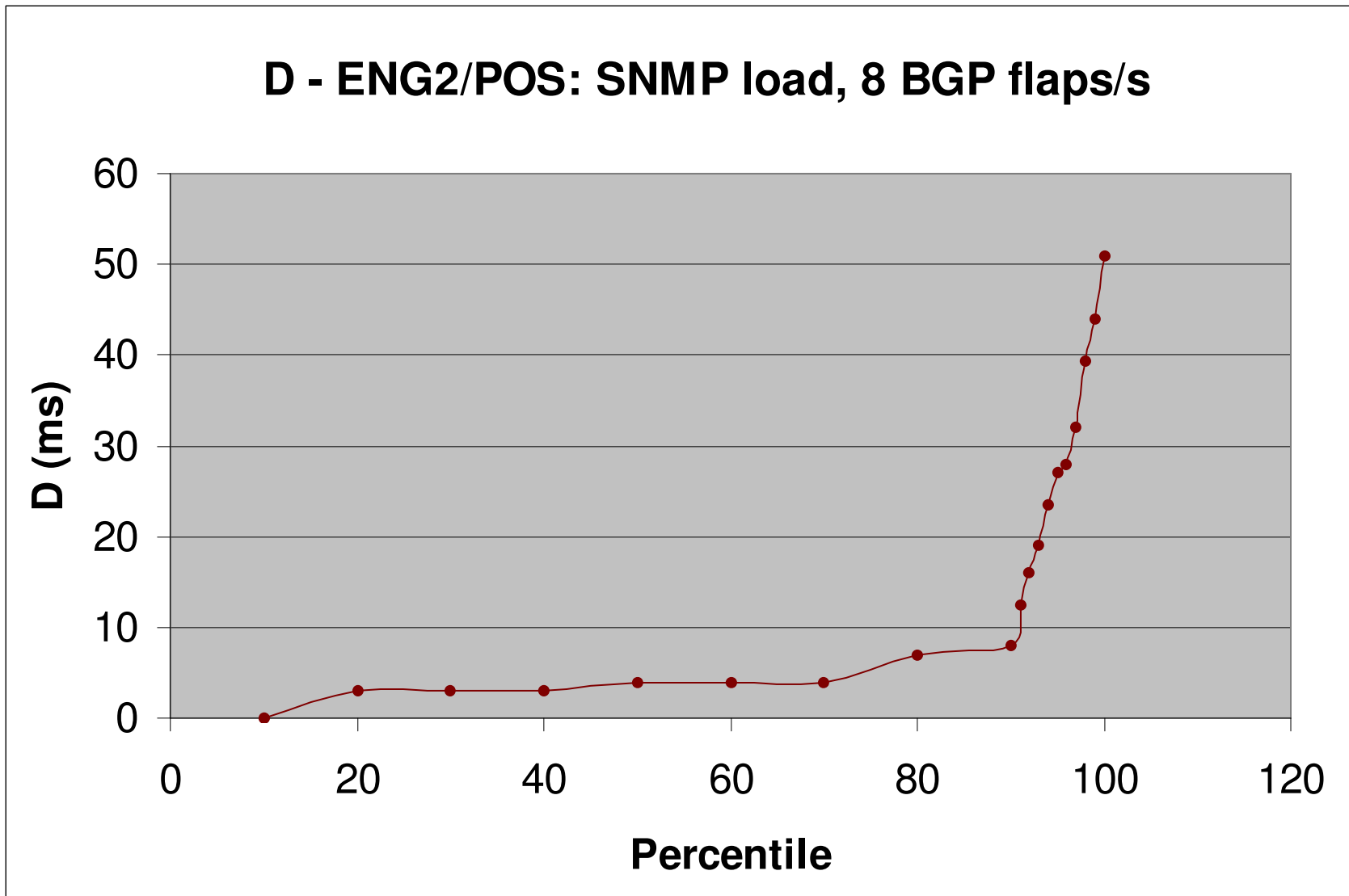
**D + O + QSP + (h \* F) + SPF(n) + Rib(p) + DD + CRR**

# D: POS – excellent for Convergence

---

- **Very fast Link Failure detection**
  - no need for fast IGP hello's
- **Various timers to order protection techniques**
  - SONET/Optical protection
- **Native anti-flap property**
  - down info is signalled very fast
  - up info is confirmed for 10s before relaying to intf.

# D



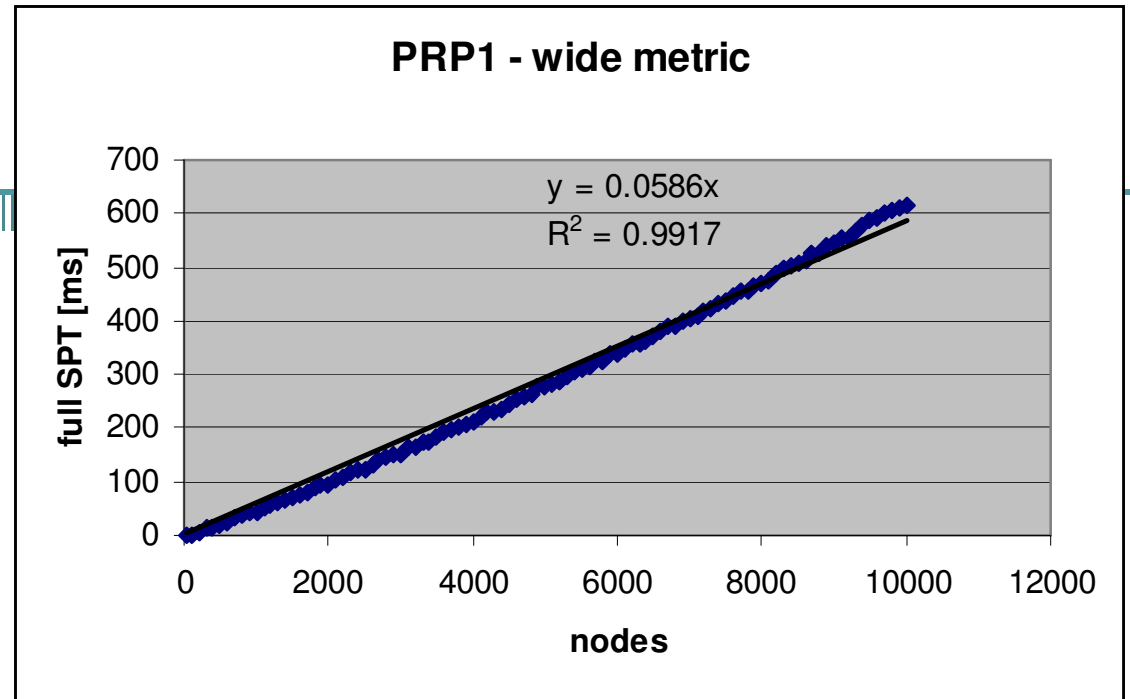


# Probability of the worst-case

---

- **D: there are two detection points**
  - WC must occur at the same time on 2 points
- **F: there are many flooding paths**
  - WC must occur at each hop for the same LSP along all possible paths
- **unlikely**

# SPT computation

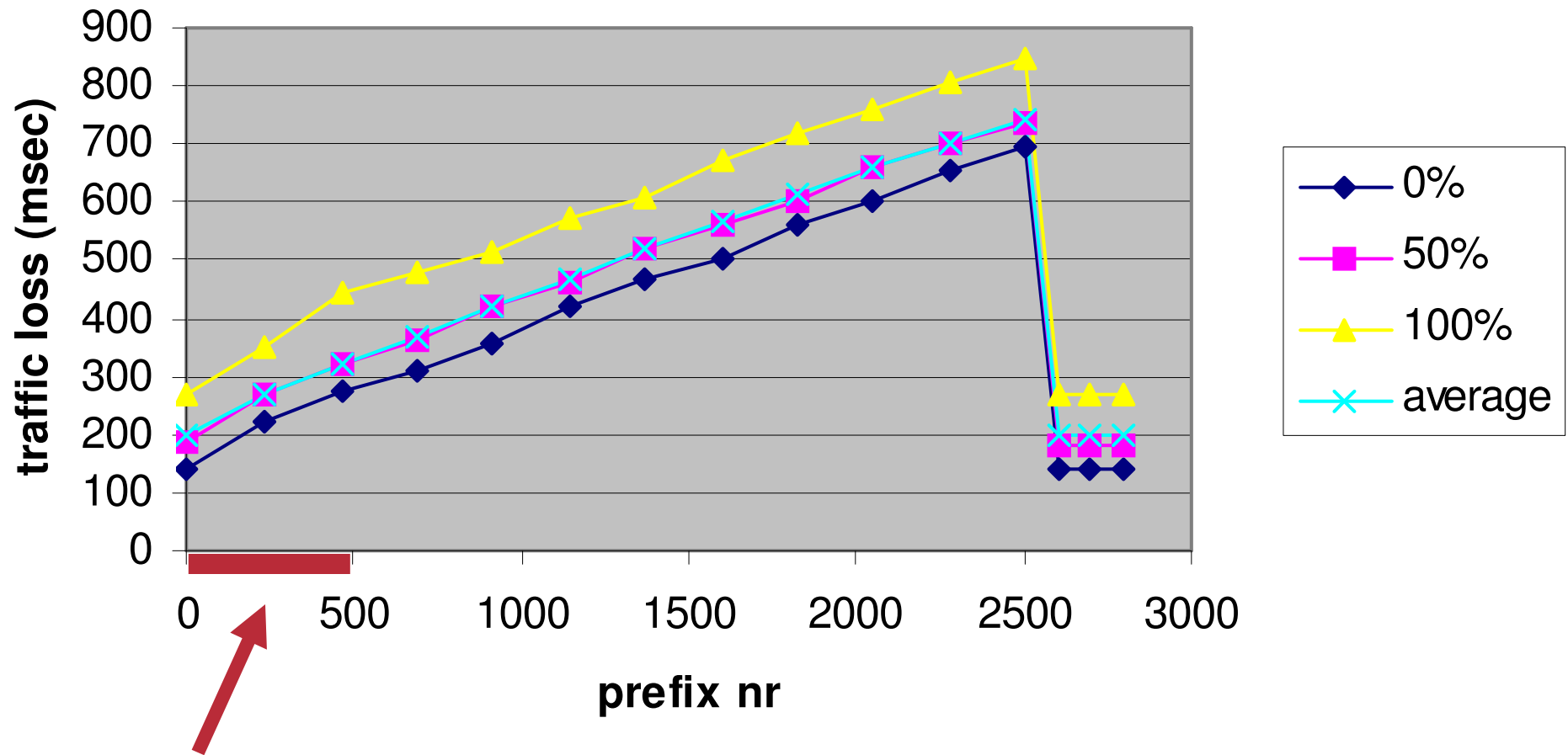


- Full SPT (wide metric): 600 nodes => 35 ms
- Incremental-SPF benefits come on top of this
  - roughly: only the nodes impacted by the failure do matter as opposed to all the nodes of the topology for a 'normal' SPF

# RIB update – prioritized update

c12k-sprint5-pr50-lc50-ipc20-bgp160-remote-nolb

Agilent measurements



# Conclusion

---

- **Sub-Second objective is realistic**
  - conservative
- **Technology has significantly improved**