MPLS convergence – IGP and BGP Convergence Impacts

-By Susan Hares
skh@nexthop.com

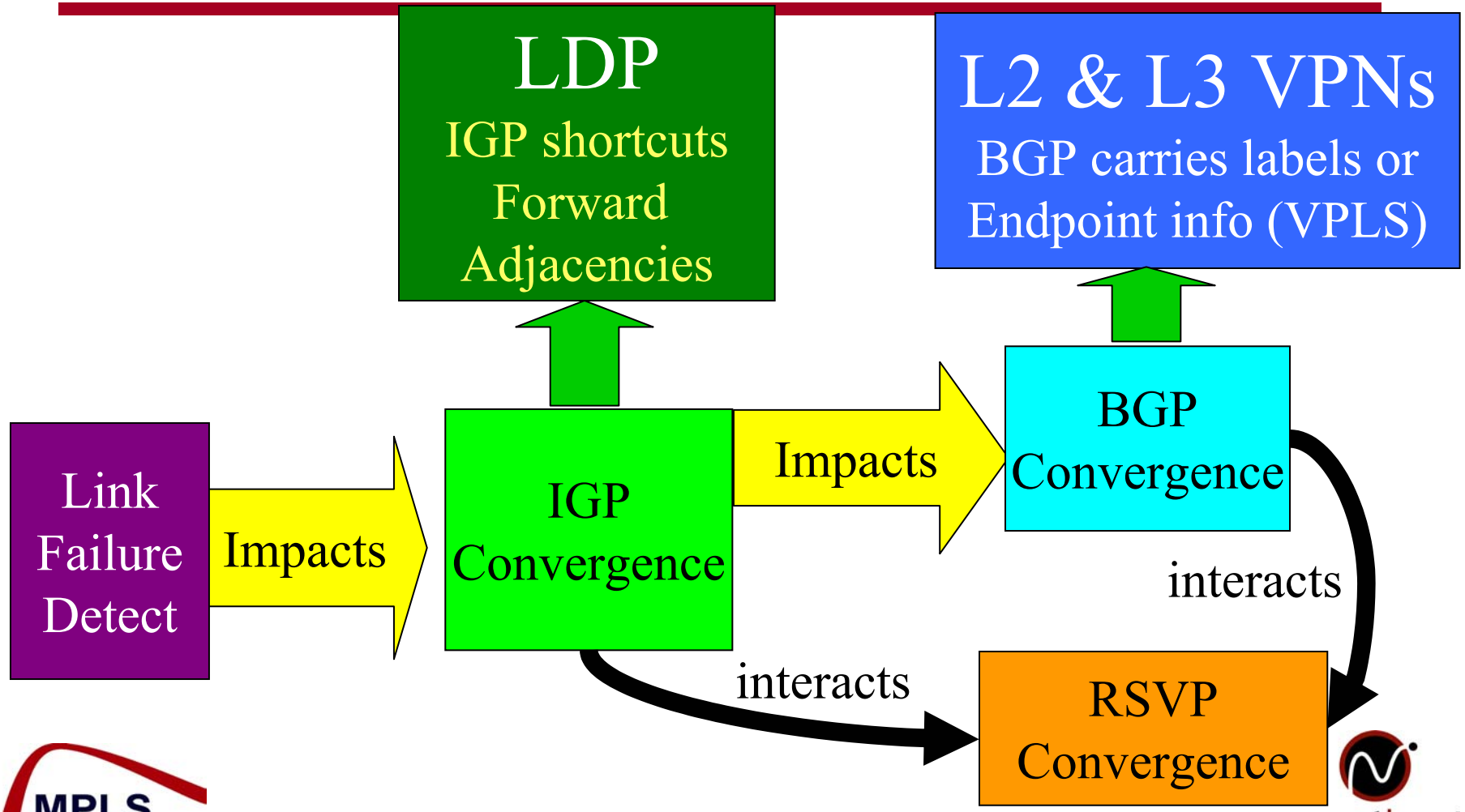
MPLS 2004

nexthop®

Why worry about Performance:

- Reduce Operational expense
  - Reduce downtime
  - Reduce outages due to configuration and provisioning errors
  - Reduce outages due to Security attack
- Increase services
  - VPN services
  - VOIP services
  - Access Services

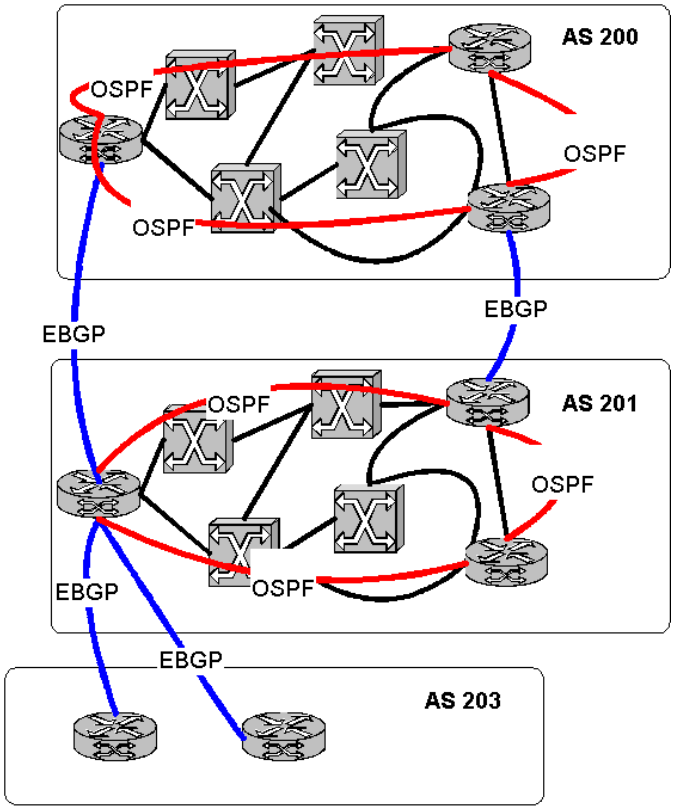
Convergence issues



Caveats on Model

- Model versus Reality
  - Believe the Real Data
  - Real IPv4 Data: Beacon Experiments, Route Views, Telstra Data base
  - Problem: No L3 VPN data recorded separate
- Model tries to match benchmarking with Real data
  - 1st pass based on IGP + BGP models
  - 1st pass – light on the LDP and RSVP models
- Feedback on Model's appreciated

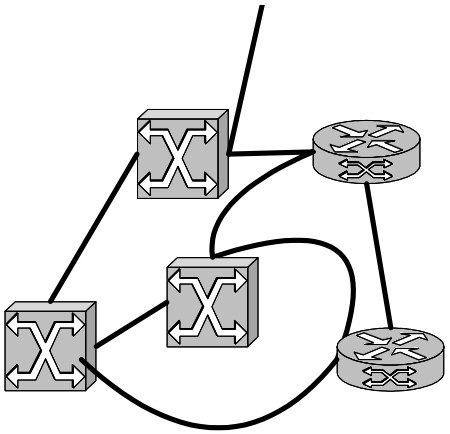
MPLS Model of Convergence



- Models for Convergence
  - Link, IGP, BGP, LDP, RSVP-TE
  - Route Traffic models
- Benchmark
  - Single Node vs Multiple Nodes
- Results

Link Failures Times

Link Failure Detect



- Model Equation: D = Minimum(3*Link-timer, S-out)
  - **Best: Hardware detection**
  - **Planned: Link-Timer : "alive" timer for the link**
  - **S-out = 3 times Link-Timer**
- Physical link
  - Sonet:**Link Time up/down = 10 ms**
    - **Improvements if change to 1 ms**
  - Ethernet, Wireless: Use of Signaled Link Down
- IP layer - BFD light weight Hello

IGP Convergence times

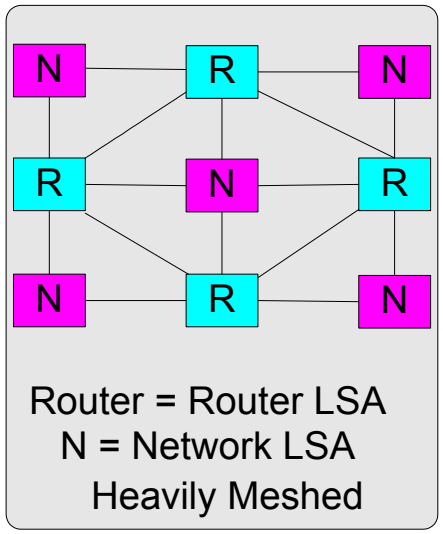
$$LoC(p) = D + O + QSP + (h * F) + SPF(n) + RIB(p) + FIB(p) + DD + CRR$$

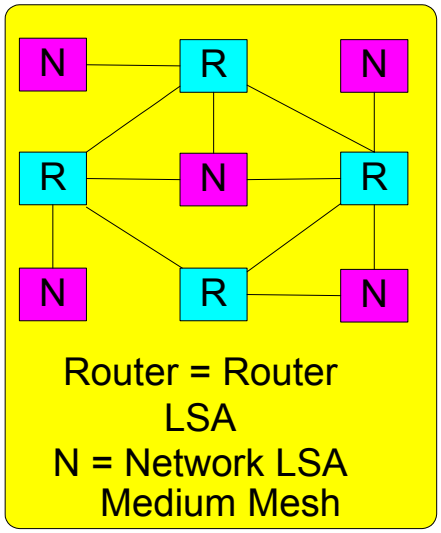
- D: Failure is detected
- O: New LSP is originated
  - Insertion of External routes occurs here
- QSP: cumulative queuing, serialization, propagation,
- H*F: LSP is flooded up to rerouting node
- SPF: Shortest path calculation
- RIB: RIB is updated
- FIB: FIB is updated
- DD: LCs are updated
- CRR: BGP recursion is fixed

Dominate variables With high speed/bandwidth links

References: www.nanog.org/mtg-0202/ppt/cengiz.pdf; mtg-0010/ppt/cengiz.pdf

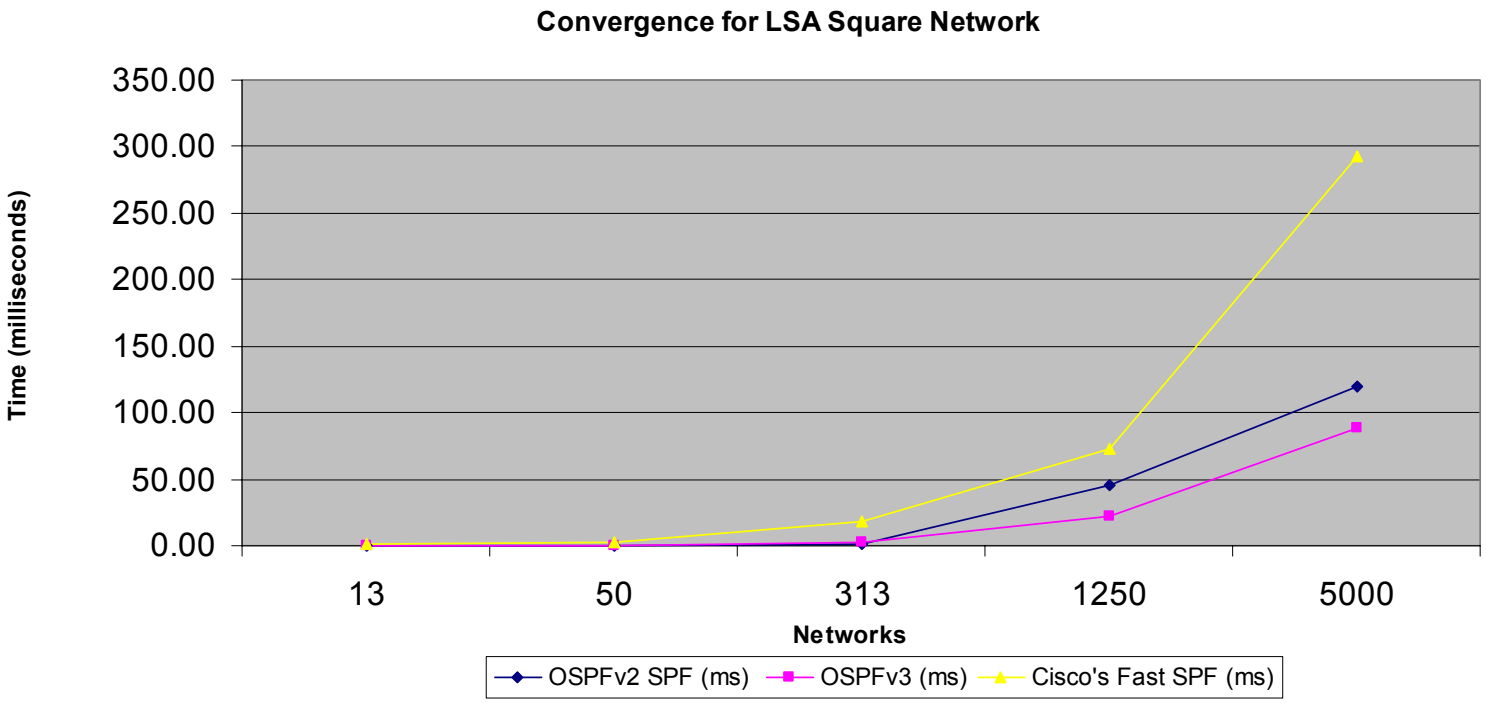
IGP Topologies (OSPF, ISIS)





- Types of topology
  - Lightly connected (1-2 links/ router)
  - Medium connected (2-3 links/router
  - Highly connected (4-5 links/router)
    - Least Squares Test topology

OSPF Convergence time

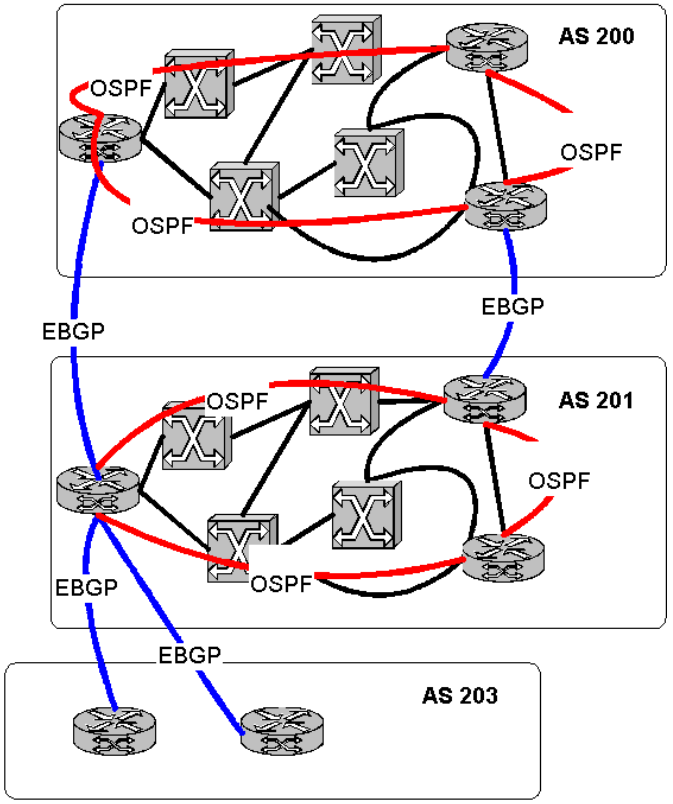


Caveat: Cisco utilized medium mesh; GateD utilizes heavily meshed

OSPF External Routes caveat

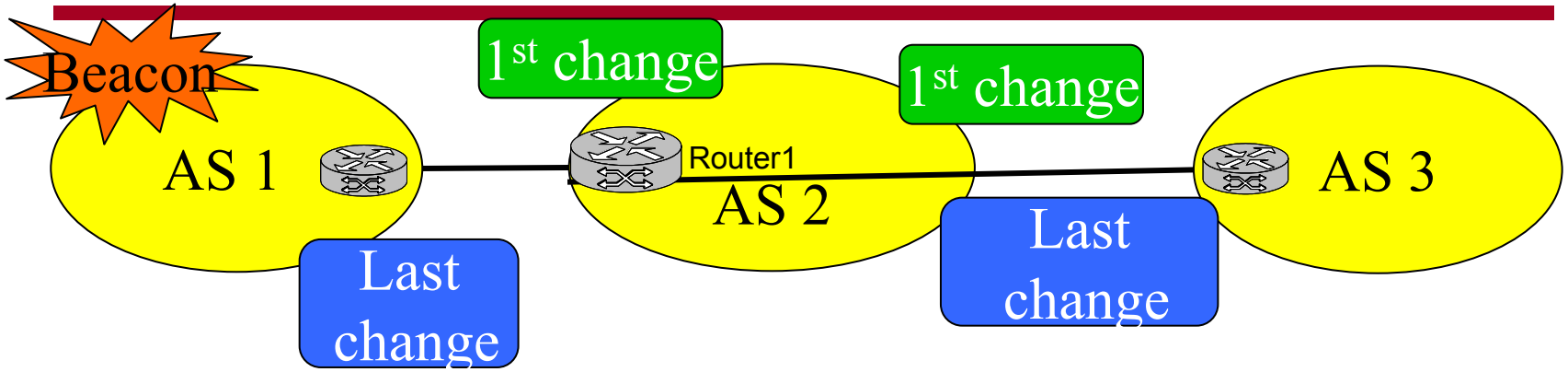
- The deployment of the OSPF code may be in a single area or in multiple areas.
  - Model equation gives single Area equation (most used)
- A more complex mode adds:
  - Inter-Area SPF computation
  - Intra-Area Database Lookup
  - Intra-Area SPF calculations for AS-External (type-5) and NSSA Computations
  - Summary Computations for an Area Border Router [Summary network (type-3) and Summary ASBR (type-4)]
- Benchmarking may give this
  - With internal sensors, or
  - For simple situations in black box

BGP Convergence



- BGP is Path Vector
  - rumor protocols with paths
- BGP is Policy Protocol
  - Amount policy matters in convergence
- BGP distributes routes from IGPs and Labels from MPLS
  - Redistribution times matter
- BGP is widely deployed
  - Data of BGP routes patterns matter
  - Processing patterns matter: Fractal or Markov

Multi-AS Convergence for Beacon with multiple routes



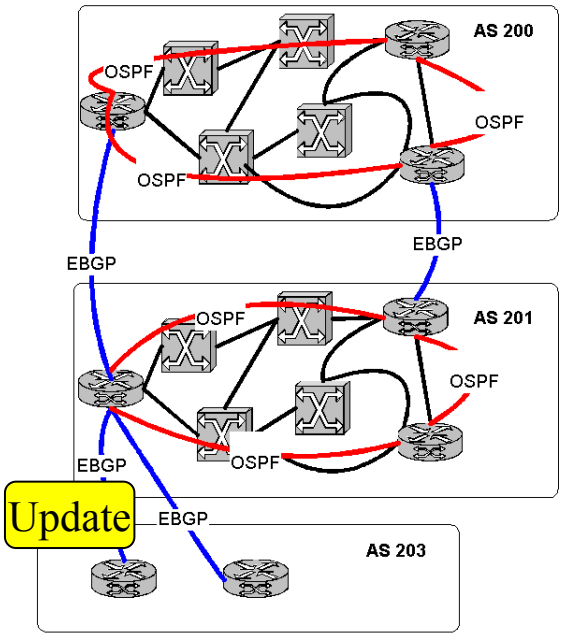
Convergence Terms as seen by AS 2 Router 1

- T0 – 1st change received from AS 1
- T1 – 1st change sent to AS 3
- T2 - Last change received from AS 1
- T3 – Last change received from AS 1

Beacon Experiments are collecting data !!

Received E-BGP updates

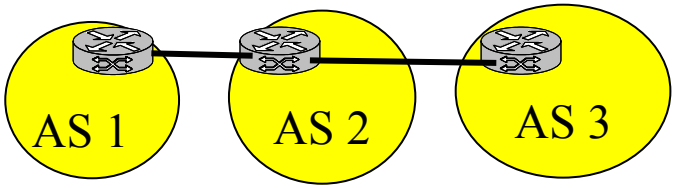
E-BGPec(p) = Ur + (Tr + QSP + RS-Policy + PV+ RIB)(rts) + IBGPec+ EBGP-lpec + EBGP-ASpec + FIB + CRR



- **Ur: Update is received**
- **Tr – transmission of Update**
- **QSP: cumulative queuing, serialization, propagation**
- **RS-Policy: Route Selection Policy**
- **PV: Path Vector Calculation**
- **RIB: RIB is updated**
- **DS-Policy: Distribution Policy**
- **IBGP-ec – IBGP convergence based on exterior routes**
- **EBGP-lpec – Convergence EBGP Peers for BGP peer local to receiving BGP router**
- **EBGP-ASpec – Convergence of EBGP Peers for BGP peers within AS**
- **FIB: FIB is updated**
- **CRR – BGP Route Recursion is check**

E-BGP with no IBGP peers

E-BGPec(p) = $Ur_{(T2\text{-}To)}$ + (Tr + QSP + RS-Policy + PV+ RIB)(rts)
+ nEBGP-lpec (DS-Policy + RIB + Uo + Tx) +
nEBGP-Asp-ec(Uo + Tr + QSP + RS-Policy) + FIB + CRR



- **Ur: A sequence of New Updates are received**
- **Uo – update originated**
- **Tr – processing reception of the New Updates with Route(s)**
- **QSP: cumulative queuing, serialization, propagation**
- **RS-Policy: Route Selection Policy**
- **PV: Path Vector Calculation**
- **RIB: RIB is updated**
- **DS-Policy: Distribution Policy**
- **TX – Update is transmitted**
- **FIB: FIB is updated**
- **CRR – BGP Route Recursion is check**

I-BGP convergence: transit

BGPlec(p) = **Ur** + N-Peers (Tr + QSP + RS-Policy + PV + RIB + DS-Policy+ RIB + Tx) + FIB + CRR

- **D: Failure is detected**
- **LoC(p): IGP convergence**
- **Uo: New Update is originated**
- **Uor: New Update is originated based on E-BGP received**
- **Tr – transmission of Update**
- **Tx – Update is transmitted**
- **QSP: cumulative queuing, serialization, propagation**
- **RS-Policy: Route Selection Policy**
- **DS-Policy: Route Distribution Policy**
- **PV: Path Vector Calculation**
- **FIB: FIB is updated**
- **RIB: RIB is updated**
- **CRR – Route Recursion is check**

•Worse Case = Sequential N-Peers

• Best case = 1 ½ peers

I-BGP convergence: Originated

BGPoc(p) = **D + LoC(p) + Uo** + N-Peers (Tr + QSP + RS-Policy + PV + RIB + DS-Policy+ RIB + Tx) + FIB + CRR

- **D: Failure is detected**
- **LoC(p): IGP convergence**
- **Uo: New Update is originated**
- **Uor: New Update is originated based on E-BGP received**
- **Tr – transmission of Update**
- **Tx – Update is transmitted**
- **QSP: cumulative queuing, serialization, propagation**
- **RS-Policy: Route Selection Policy**
- **DS-Policy: Route Distribution Policy**
- **PV: Path Vector Calculation**
- **FIB: FIB is updated**
- **RIB: RIB is udpated**
- **BGP – Route Recursion is check**

•Worse Case = Sequential N-Peers

• Best case = 1 ½ peers

LDP Convergence- IGP/LDP node

LDPoc(p) = D + LoC(p) + LO + (h * (LDP-node))*Path + LIB + LFIB

LDP-node = Lr + LSP-RS-Policy + LD + LSP-DS-Policy + Lx + LIB + LFIB ***(preliminary)***

- D: Failure is detected
- LoC(p): IGP convergence [From OSPF]
- LO: New Update to Labels are originated
- h – hops for LDP distribution from ingress to egress
- Path – Pathways factor
- LDP-node – per node distribution
- Lr – Label update received
- Lx – Label update sent
- LIB – label information base updated
- LFIB – label forwarding base updated

•Worse Case = All paths sequential

• Best case = 1 or 1½ paths

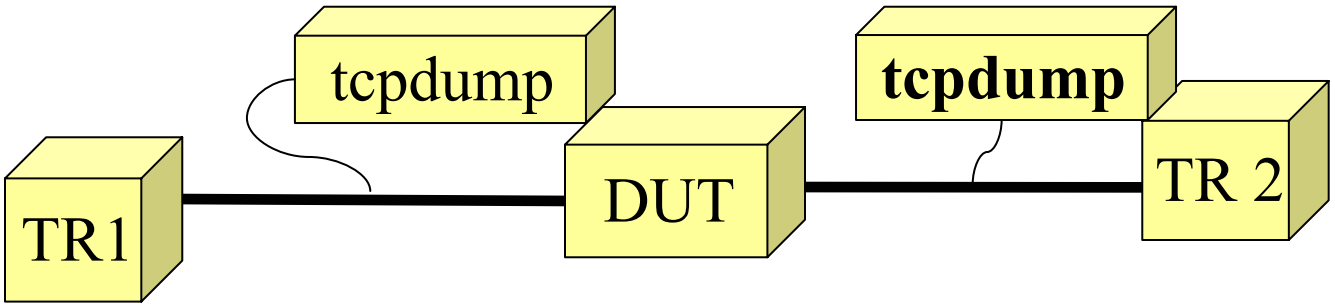
RSVP-TE Convergence times

$$LoC(p) = RR + RRO + QSP + (h * (rspv\text{-}node)) + LIB + LFIB$$

- RR: Reservation Requested
- RRO: New RSVP Request
  - Insertion of External routes occurs here
- QSP: cumulative queuing, serialization, propagation,
- H*F: LSP is flooded up to rerouting node
- SPF: Shortest path calculation
- RIB: RIB is updated
- FIB: FIB is updated
- DD: LCs are updated
- CRR: BGP recursion is fixed

Benchmarking Performance

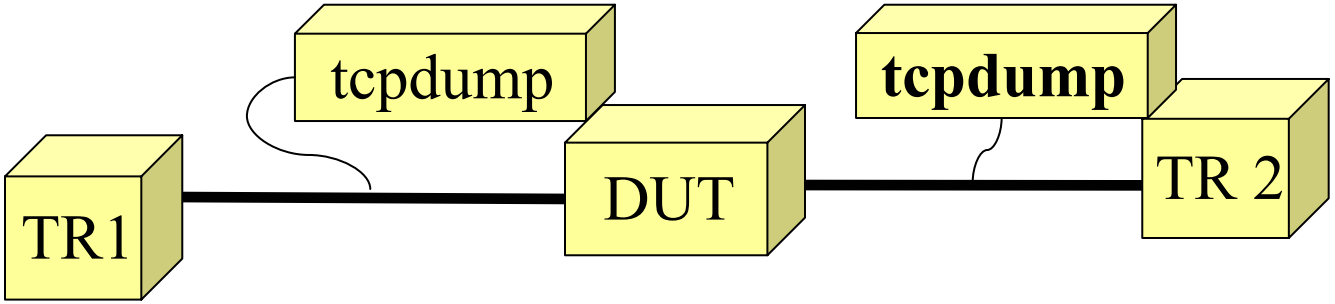
Benchmarking: Test topologies



•IGP

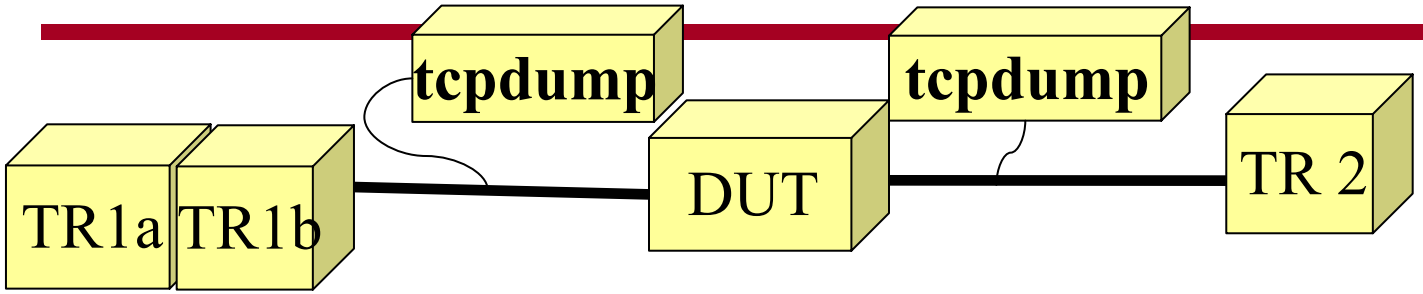
- Data stream: Generated OSPF LSA/ISIS LSP Sent to node
- Internal measures: SPF convergence, inter-area,

Benchmarking: Test topologies



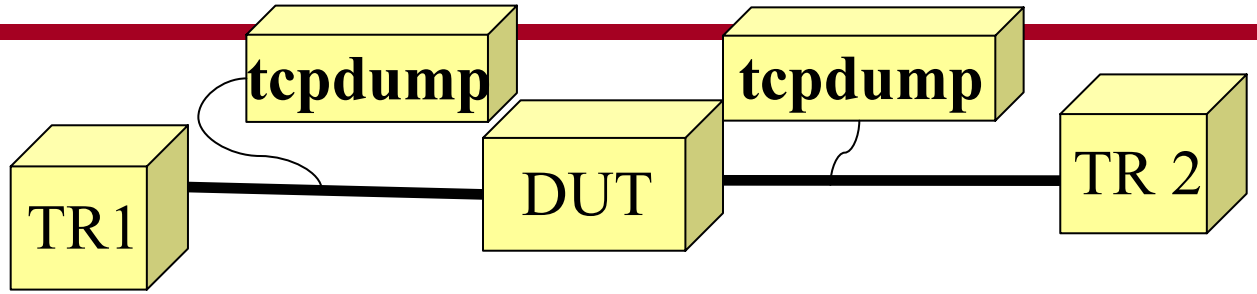
- BGP
  - Data stream parameters: Prefix mixtures, # of AS in Path
    - No substitute for the Real data mixtures
    - Models: Fractal (blast) or Markov arrival rates
  - Policy parameters: RS-Policy and DS-Policy in term of ACLs or route maps
  - TCP Parameters
  - BGP protocol: off WRD, IBGP synchronization, Smoothing, Min-Route Advertisement time
  - Measurements: Blackbox (tcpdump) versus Internal

Benchmarking: Test topologies



- L3 VPN
  - Data stream: BGP test patterns with Label insertion based on translation
  - Measurement: blackbox (tcpdump) vs internal
- LDP
  - Data stream:
    - Generated OSPF LSA/ISIS LSP inserted into Node
    - LDP generates label per prefix
  - Internal measures: SPF convergence, inter-area,

Benchmarking: Test topologies

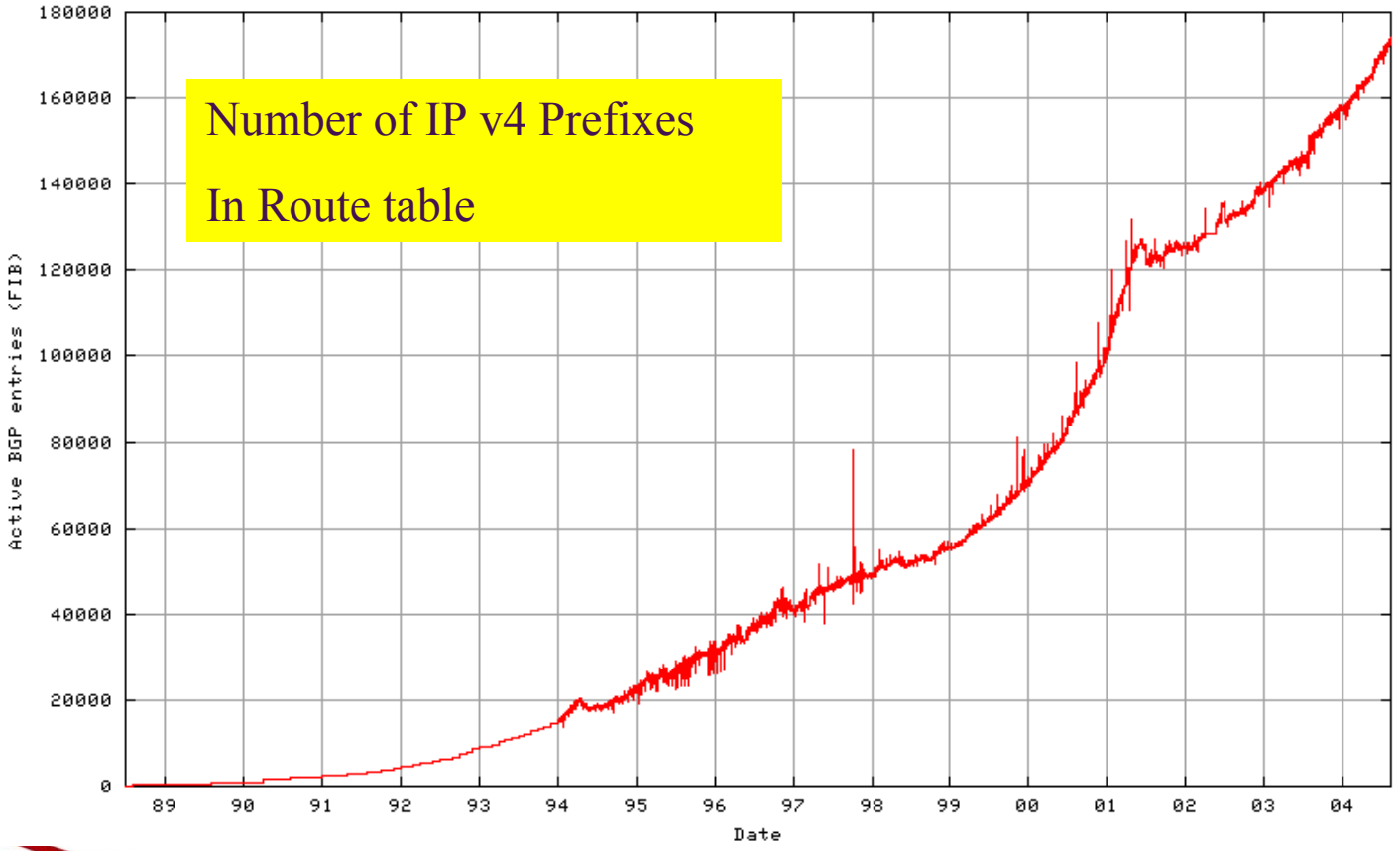


- **RSVP**
  - **Data stream: Generated Requests based on Traffic manager model**
  - **External measurement: TCP dump or TR2**

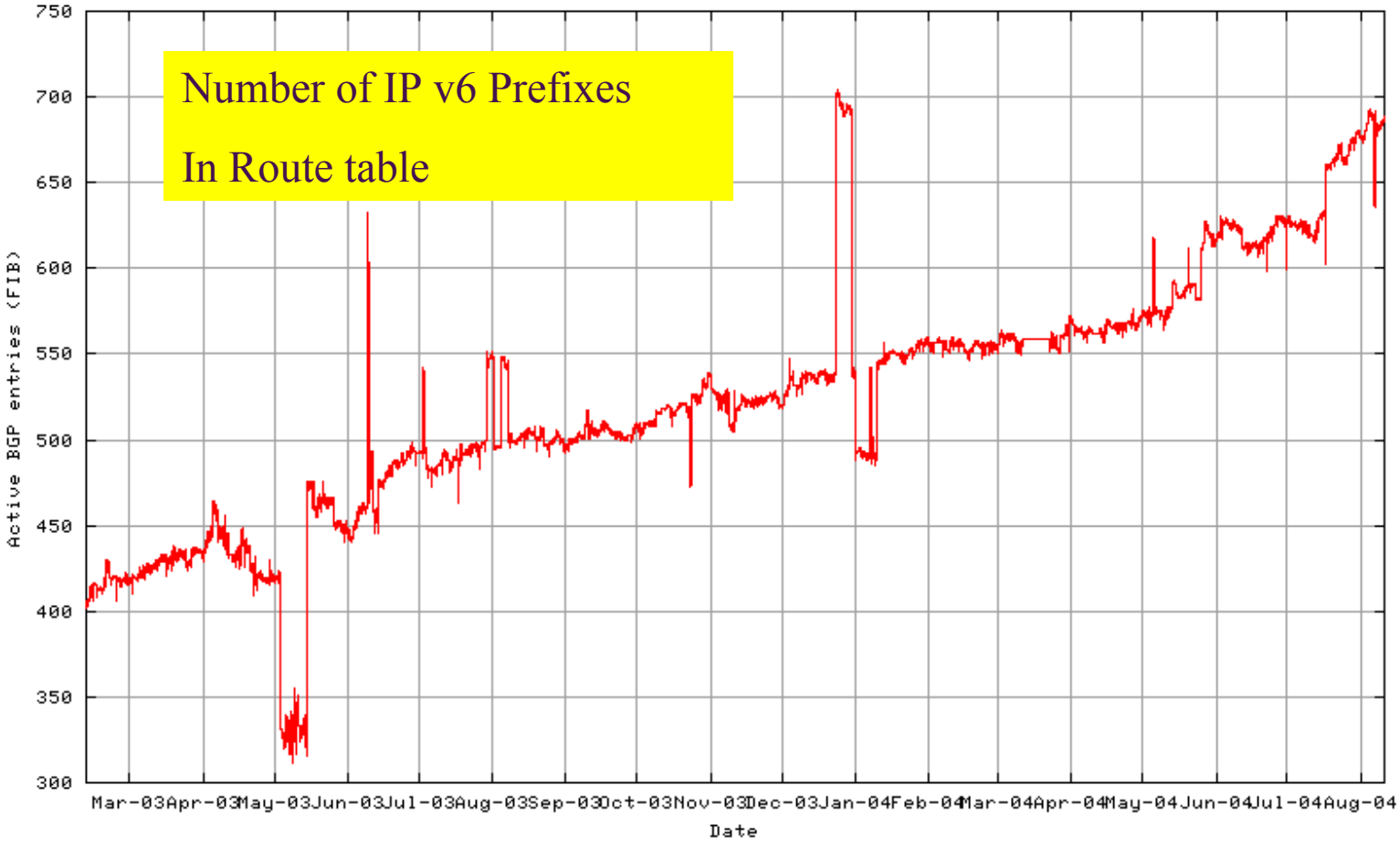
Q&A

MPLS 2004

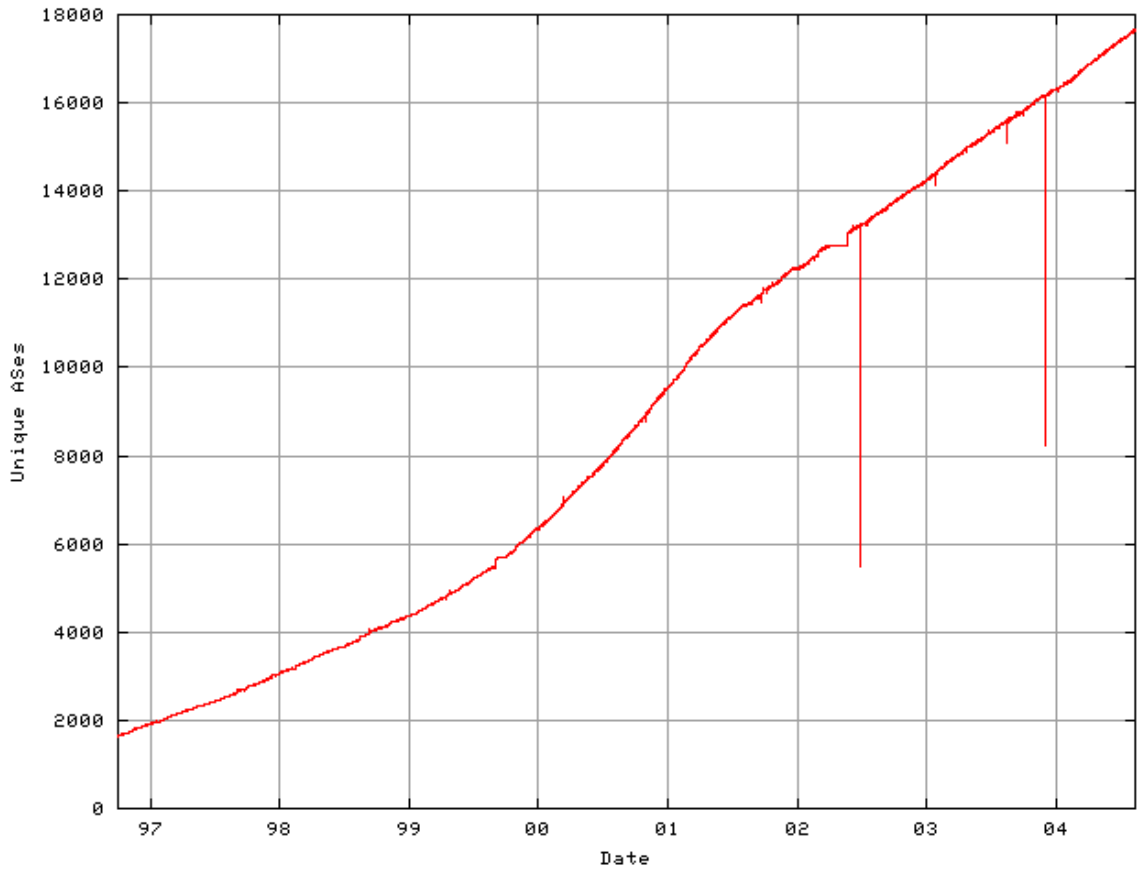
The Internet Continues to Grow



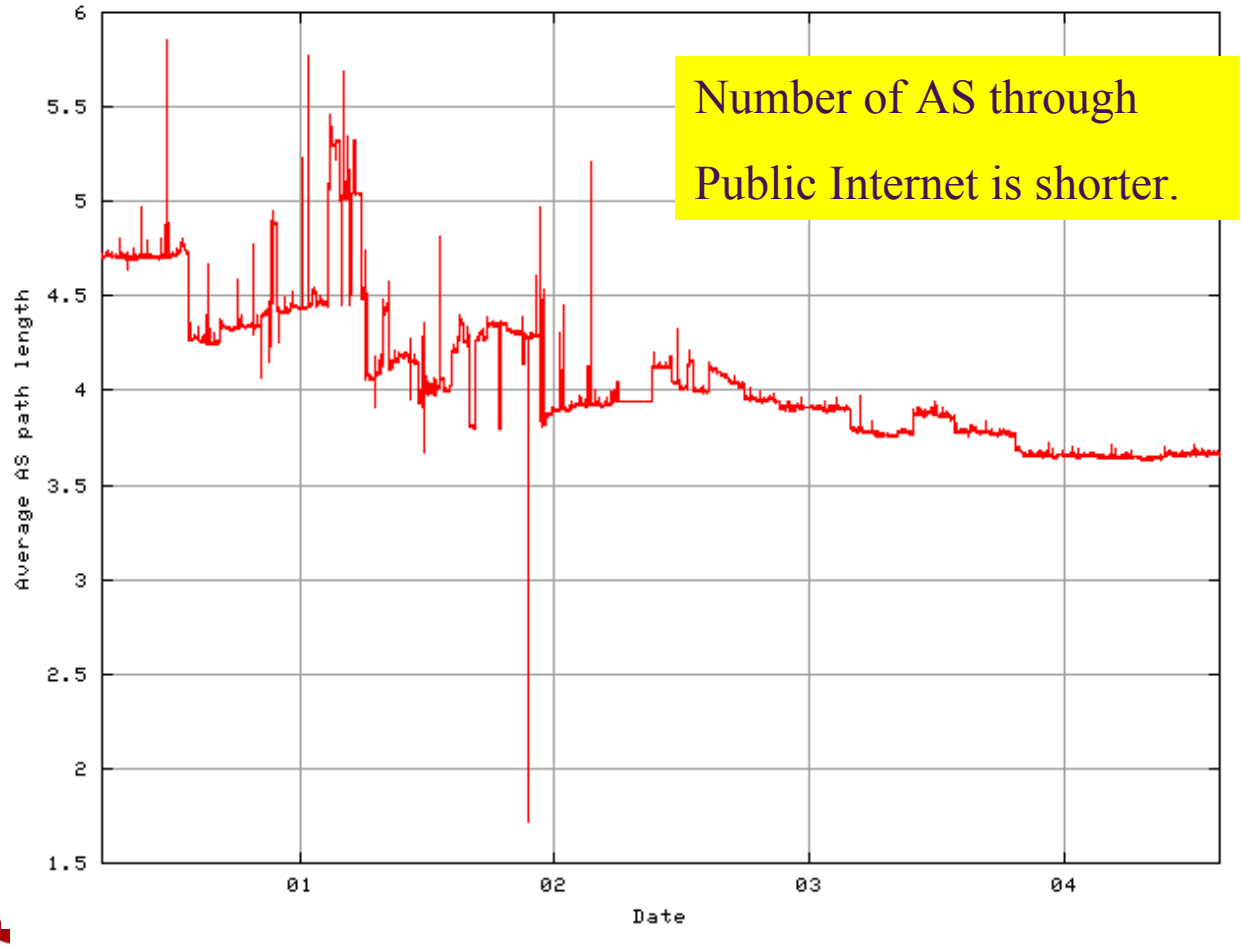
IP v6 is emerging



The Number of Unique Networks (AS) is growing



Networks are getting more connected



Less network prefixes per AS