INTER-PROVIDER QUALITY OF SERVICE

BRUCE DAVIE
Cisco Fellow
bsd@cisco.com

MPLS 2004

Cisco Systems

Agenda

- Motivation
- QoS Deployment in Single Provider VPNs
- Inter-provider QoS Challenges
- Service Specification Issues
- Mechanisms for Service Delivery
- Measurement and Monitoring
- Conclusions

Scope of the problem

- Inter-provider QoS issues are also relevant to the public Internet, independent of MPLS
- MPLS VPNs a driving application for inter-provider QoS
  - Higher customer expectations for QoS
  - QoS deployment becoming the norm in *single* provider VPNs
  - Don’t need to solve all the inter-organizational issues
    - e.g. Yahoo’s idea of “critical” traffic may differ from mine

→ Moving from single-provider VPN QoS to multi-provider VPN QoS an appealing incremental deployment strategy

Motivation

- QoS a key requirement for many VPN users
- Providers want to increase value by delivering QoS
- Some MPLS-VPN providers looking to increase "footprint" through peering
    - QoS-enabled peering the logical next step
- IPSEC the “easy” way to do inter-provider VPNs today, but unlikely to offer inter-provider QoS anytime soon
- Inter-provider QoS could speed adoption of MPLS VPNs
    - Some global companies see it as *the* reason for MPLS

Motivation (cont.)

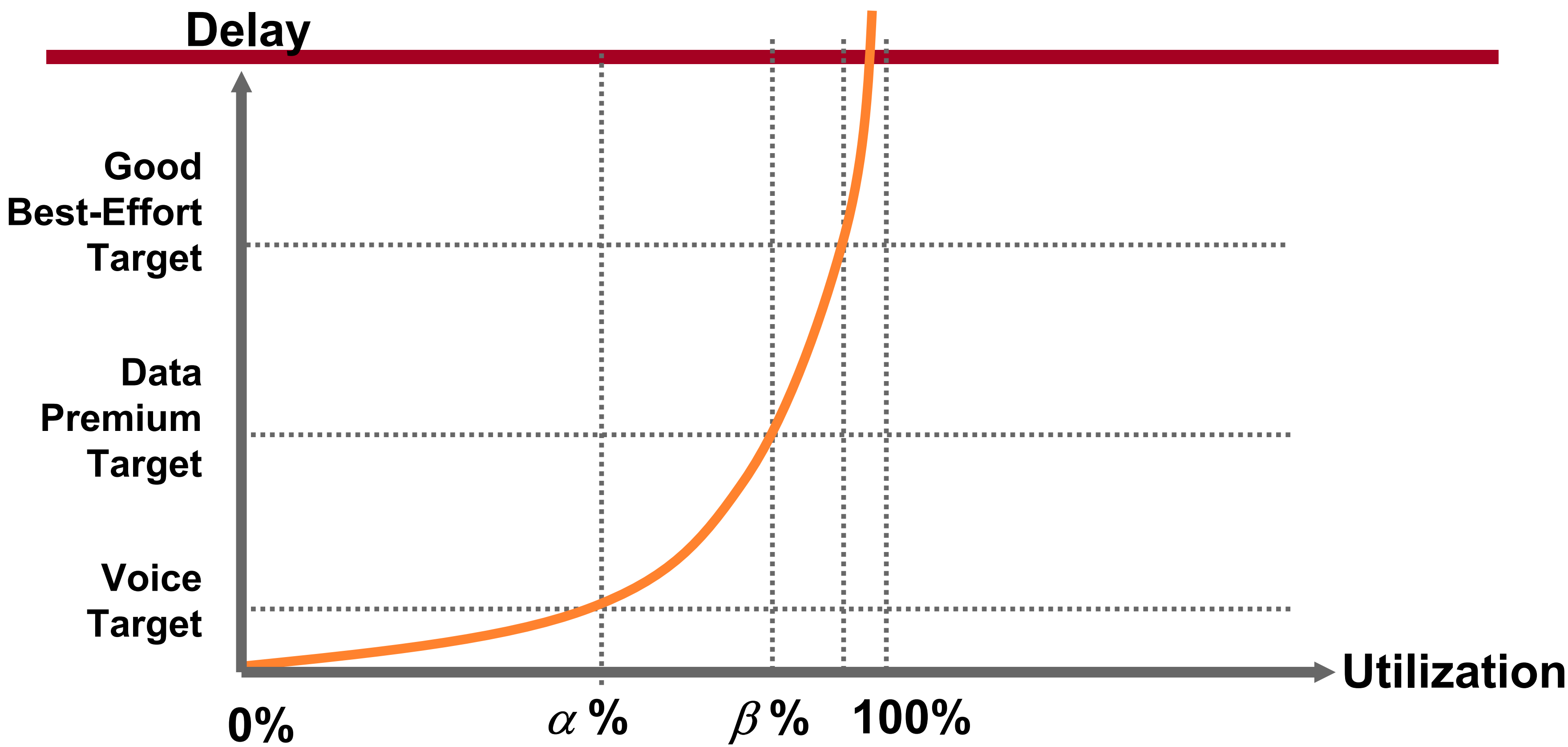
- By enabling interprovider QoS, providers could grow overall market for IP/MPLS-based services
  - Enable migration of more critical applications to IP/MPLS
- Analogy: today's multi-provider Internet has much higher utility than "closed" packet networks of the past
- Ignoring interprovider QoS has two risks:
  - Third parties will meet the need using overlays
  - Tragedy of the Commons - lower overall utility as each provider pursues his own local optimum

Single Provider QoS: Current Status

Diffserv deployment in MPLS VPNs

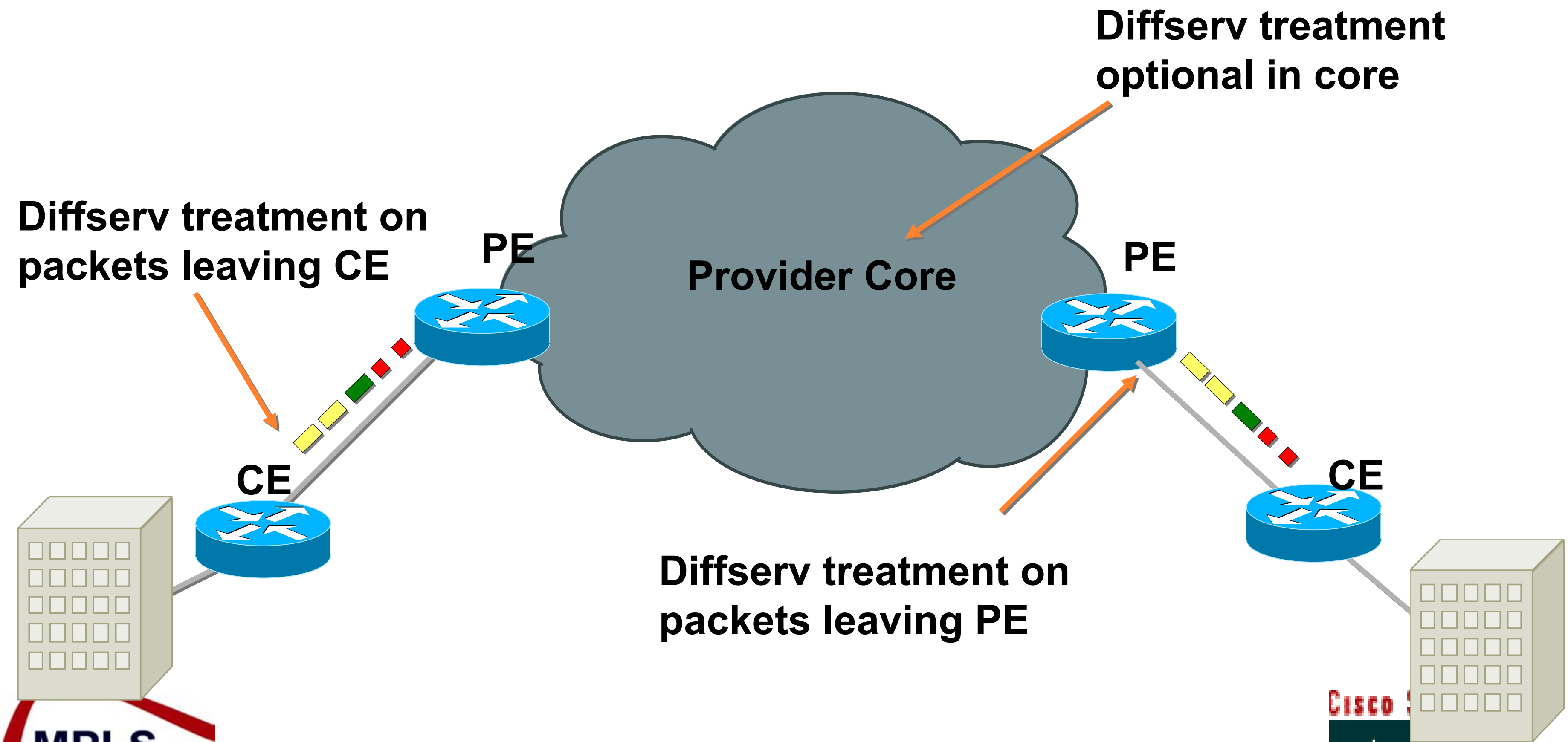
- Edge bandwidth is expensive enough to warrant complexity of Diffserv – don't want to buy enough BW to provide voice-like QoS to all traffic
  - No need to run bulk data at 30% utilization
  - Relative burstiness of data makes matters worse
- Service often competes with/replaces Frame Relay (w/ CIR), so QoS expected
- 200+ providers running RFC2547 VPNs – the majority offer Diffserv-based QoS
- Not much incremental cost to do Diffserv in the core once implemented on edge

Delay/Load Tradeoff



If I Can Keep EF Traffic < α%, I Will Keep EF Delay Under *M1* ms
If I Can Keep AF1 Traffic < β%, I Will Keep AF1 Delay Under *M2* ms

Diffserv on the edge



Diffserv in single provider VPNs

- Customers typically subscribe to 2-4 service classes
  - Number of classes and their definitions vary among providers
- Examples:
  - "VOIP" class offers low latency, low loss up to a rate limit
  - Usually implemented as EF using priority queue (LLQ)
  - "Premium data" class offers bounded latency, low loss, with ability to burst above agreed rate
  - Implemented as AFx with class-based queue
  - Best effort

Interprovider QoS: Challenges

Inter-provider QoS: Issues for Providers

- For providers with global reach, inter-provider QoS opens door to competition
- Lack of customer ownership
- Settlements/revenue recognition
- Finger pointing when SLAs are not met
- Concern about commoditization
- Data/topology hiding

Interprovider Routing Issues

- Customers can't easily predict which ISPs their traffic will traverse
  - With whom do I negotiate my SLA, to whom do I complain?
  - What speed of light delay will I see?
  - Will all SPs in path deliver QoS?

Service Concatenation Challenges

- What service is obtained when the services of several providers are concatenated?
- Lack of common service definitions makes this especially hard
    - e.g., if one provider measures jitter over a month and one measures over 5 minutes, how is jitter defined when these two providers are concatenated?
- Even consistent services can be hard to concatenate

An analogy: Octane ratings

- Three aspects of standardization
  - A technical definition of the octane rating scale
  - A defined test to measure octane rating
  - Standardized "tiers" of octane (regular, mid-grade, super)
- Still some opportunities for differentiation
  - e.g. go beyond the minimum requirement for "super" with 94-octane

Service Specification Issues

- Need to define the performance metrics (e.g. jitter, loss) consistently (cf. Octane)
- Need a small set of common services that can be concatenated across providers (cf. Regular, Mid, Super)
  - Providers free to offer additional services
- Specifications must be consistent at a detailed level
  - E.g. same delay percentiles, same averaging intervals
- Axiom: Implementation mechanisms must be left to the providers
  - Anything from simple overprovisioning to DS-TE

Example service definition - Telephony

- Packets marked as “EF”
- PE polices EF from customer using a token bucket
  - rate & burst are negotiated SLA parameters
- For “in-contract” traffic, provider commits to
  - Loss < x% in any y minutes
  - Mean delay < D in any y minutes
  - 99th percentile delay $< D_1$ in any y minutes
  - 99.9th percentile delay $< D_2$ in any y minutes

Concatenation challenges

- Many aspects of service need to be consistent among providers (e.g. measurement interval “y”)
- Adding delay percentiles is pessimistic
  - Sum the 99th percentile across 3 providers and you have the 97th percentile
- Uncertainty about number of providers in path
  - Sum of mean delays may get too big

Routing

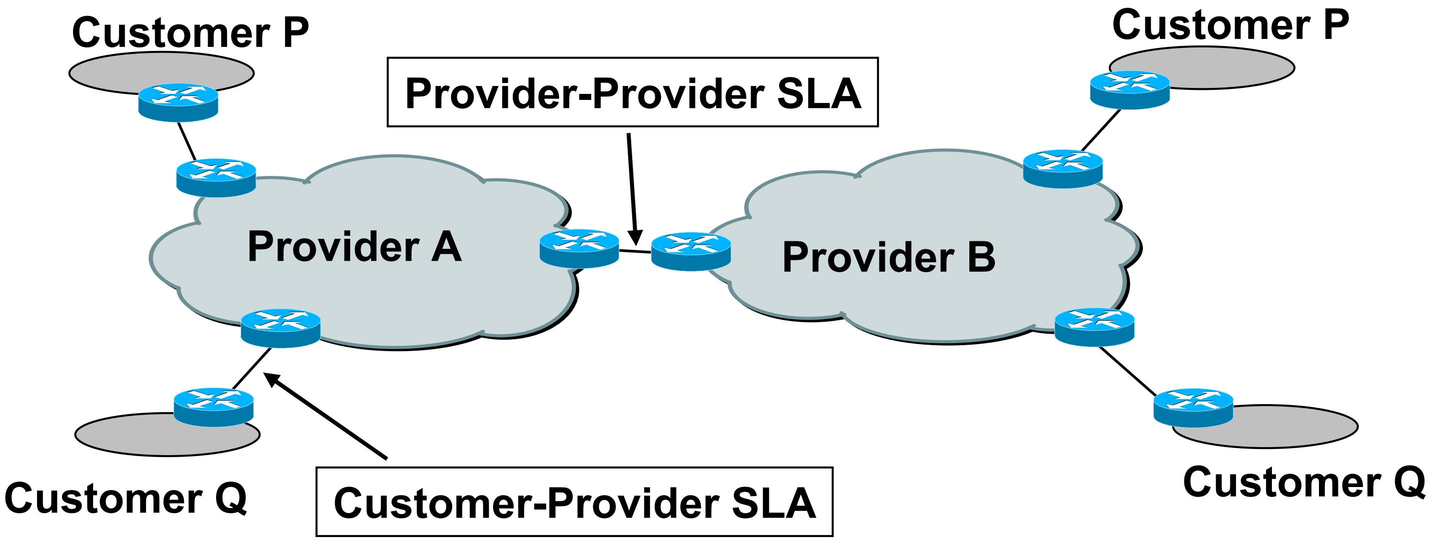
- Problem: how to ensure that traffic needing QoS traverses providers who can deliver QoS?
- Possible approach:
  - Define a small set of services
  - Use a small number of BGP community attributes to indicate which services an AS supports
  - Route QoS-enabled traffic to QoS-enabled ASes
  - Note: may require QoS-aware forwarding

Implementation Mechanisms

Service Delivery Mechanisms

- Overprovisioning
- Basic Diffserv (or MPLS-Diffserv as in RFC3270)
- MPLS TE
- Diffserv-aware TE (DS-TE)
- Aggregate RSVP (RFC 3175)
- Inter-AS TE

Concatenated SLAs



- Provider A aggregates traffic from end customers and negotiates an SLA with Provider B much like any other customer of B
- Note that any provider may choose to deliver his SLAs using Diffserv PHBs, overprovisioning or other methods

Overprovisioning

- A fine solution for many providers today
- Potentially costly in terms of fiber & routers
  - Need to engineer the network to meet the most stringent needs of any class, for all offered traffic
  - Some argue that this is offset by lower opex
- May be risky in event of link/node failure
- Lack of differentiation among classes

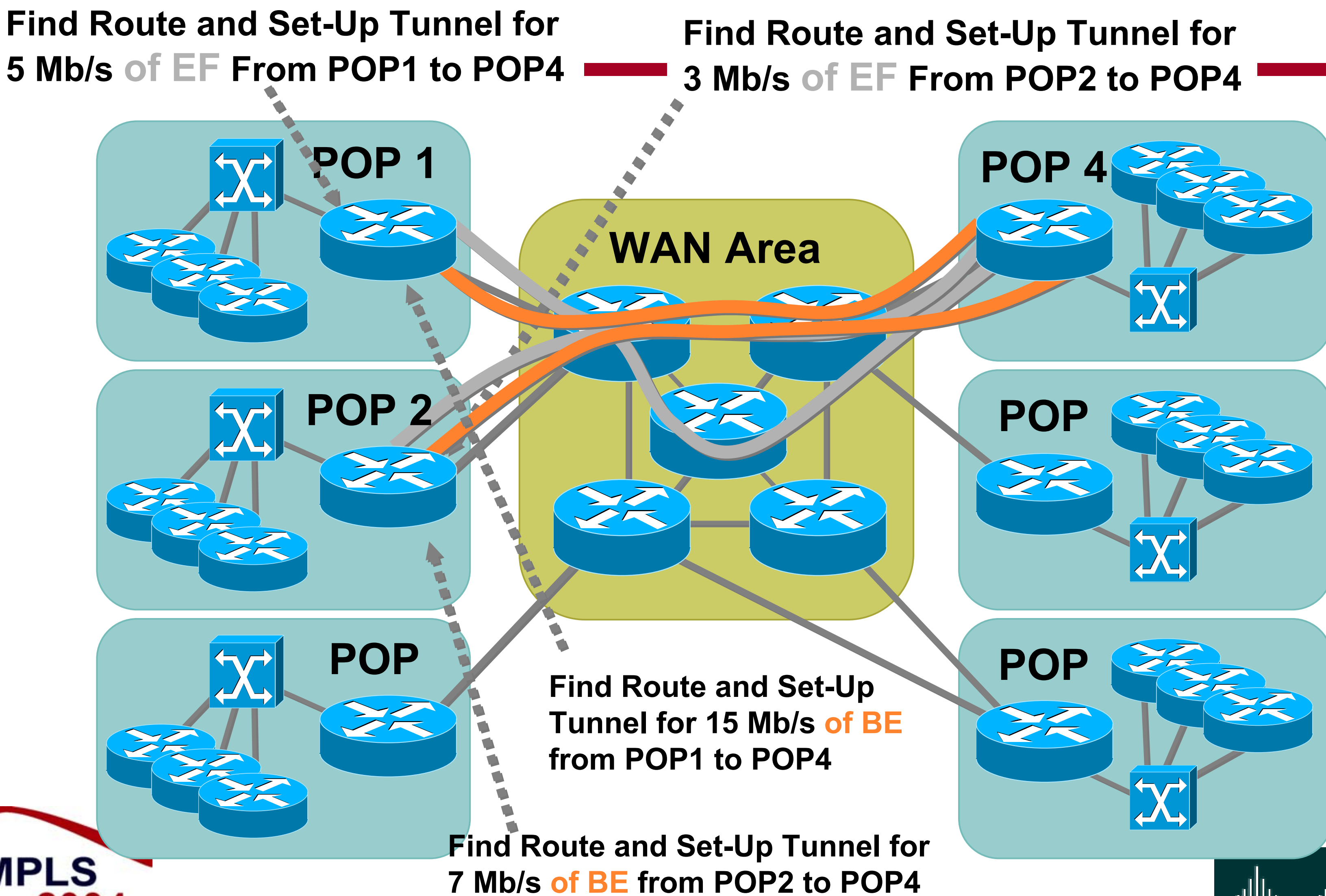
Diff-Serv-Aware TE (DS-TE)

- DS-TE is more than MPLS TE + MPLS Diff-Serv
- DS-TE makes MPLS TE aware of Diff-Serv:
  - DS-TE engineers separate LSPs for different QoS classes
  - DS-TE takes into account the bandwidth available to each class (e.g. to queue)
  - DS-TE takes into account separate engineering constraints for each class
    - e.g. I want to limit voice traffic to 70% of link max, but I don't mind having up to 100% of BE traffic
    - e.g. I want overbooking ratio of 1 for voice but 3 for BE
  - DS-TE may take into account different metrics (e.g. delay)
- DS-TE ensures specific QoS level of each Diff-Serv class is achieved

Diffserv-aware Traffic Engineering

- Provider builds a traffic matrix for “premium” traffic
  - Based on measurements, SLAs, growth projections, etc.
- DS-TE provides means to engineer paths for that traffic independent of best effort and with different utilization targets

Per-Class Traffic Engineering

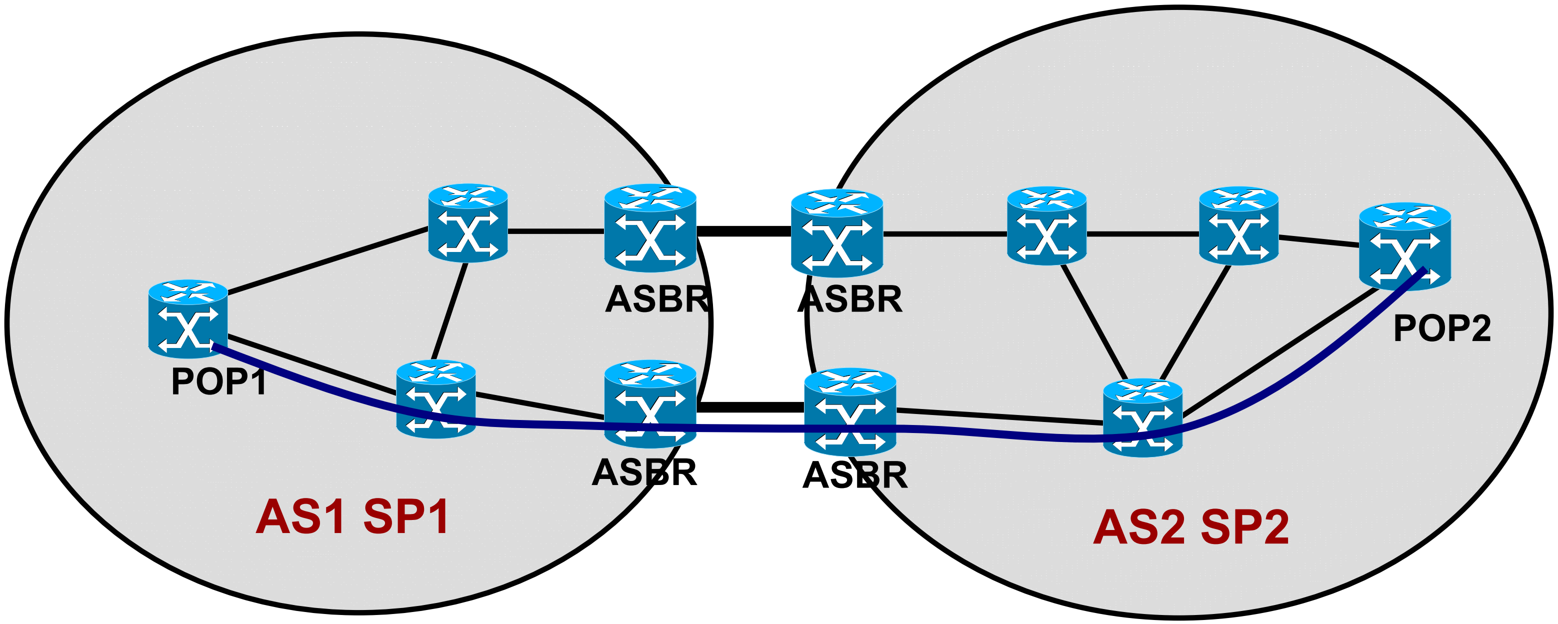


MPLS 2004

Aggregate RSVP

- Defined in RFC 3175
- Provides somewhat similar capabilities to DS-TE
  - Allocate resources along a path for some traffic aggregate (e.g. all EF traffic from pop A to pop B)
  - But only along the shortest path
  - Naturally works across area and AS boundaries, if all providers support it

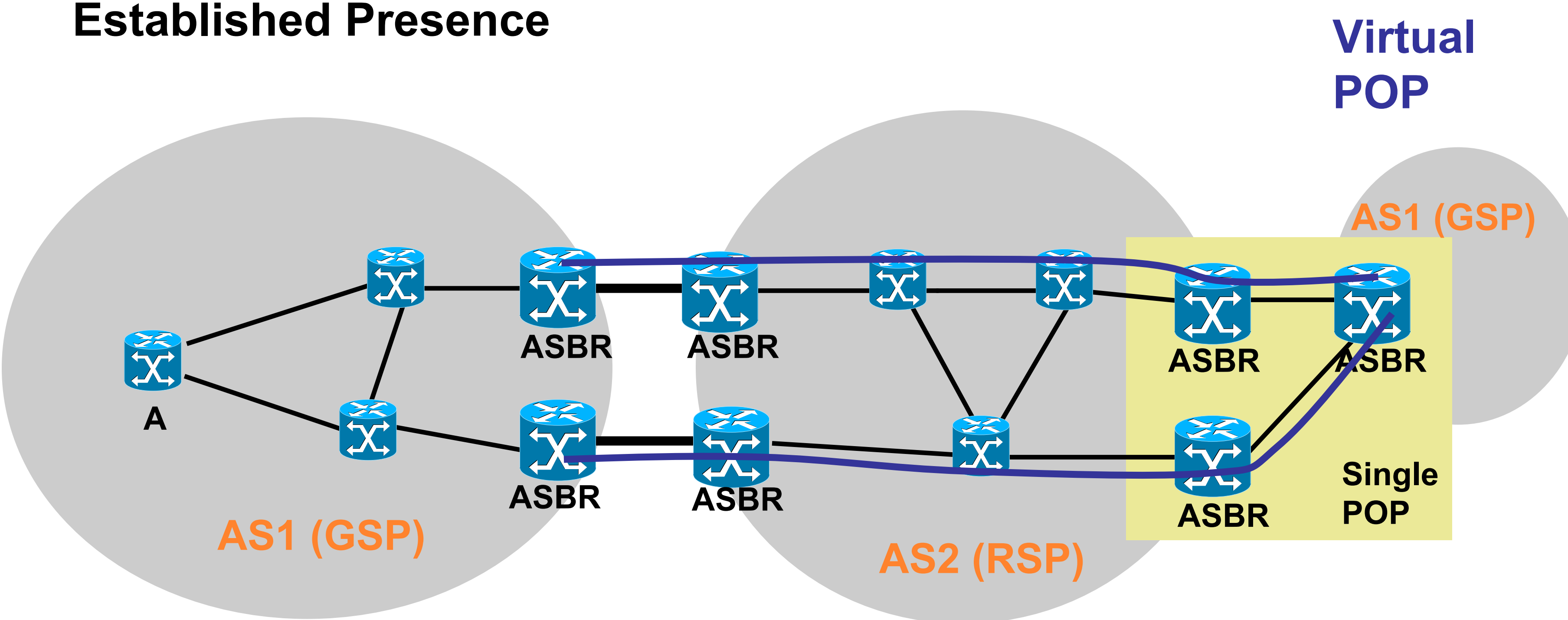
Inter-AS Traffic Engineering



- When providers have suitable knowledge of traffic matrices, inter-AS TE LSPs can be used to provide pop-pop guarantees
- Potential n^2 problem for widespread use

Deployment Scenarios: Extended/Virtual POP

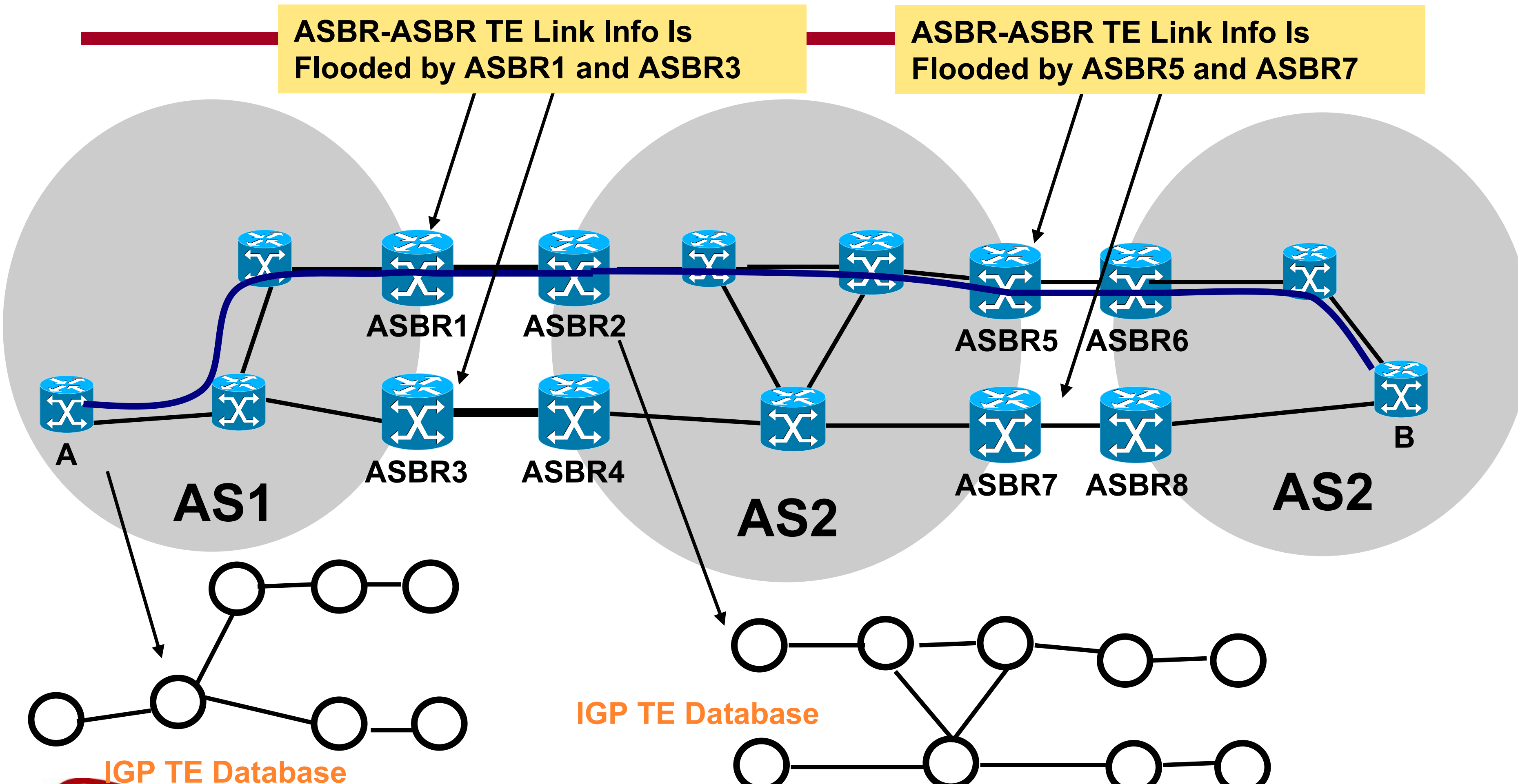
A Global Service Provider (GSP) Expands Its Reach in a Region where a Regional Service Provider (RSP) Has Already Established Presence



Per-AS Path Calculation

- One approach: use of “loose” route
- Headend calculates path only as far as it can “see”
  - i.e. to an egress ASBR
- ASBR expands the route—calculates path to next ASBR
  - And so on until destination is reached
- Main problem: headend must choose an ASBR
  - Doesn’t know enough to pick “best” ASBR
  - There may be no valid path from that ASBR to destination
  - This problem is repeated at each ASBR along path
  - No guarantee that path found is the shortest

Loose Hop Expansion



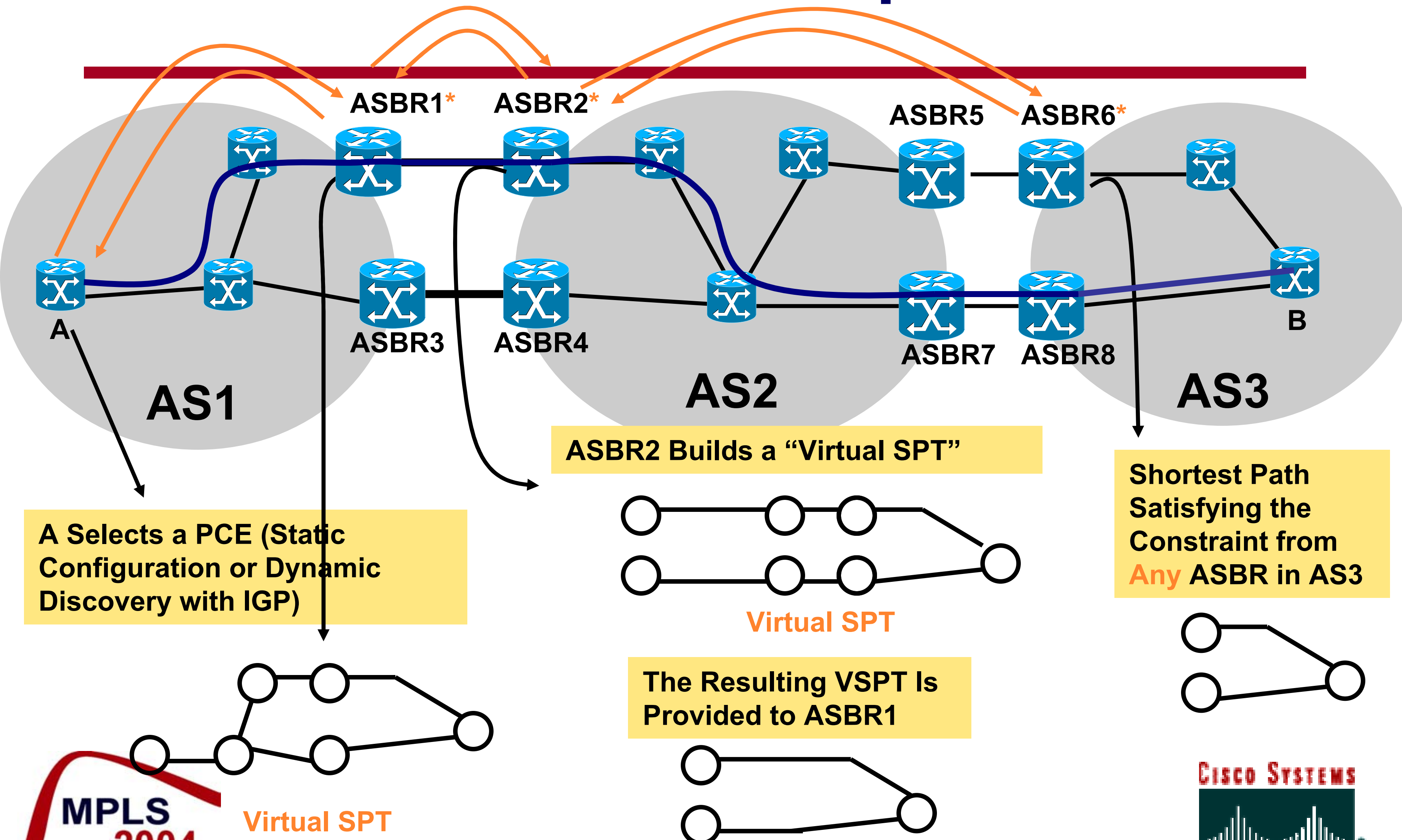
MPLS 2004

Cisco Systems

Distributed Path Computation

- Key idea: use a “path computation element” (PCE) in each AS
- PCEs communicate with each other to gather information about the topology and resources along a sequence of ASes
- PCE for each AS calculates a set of shortest paths from all its ingress ASBRs to the destination
- Each PCE reports only those paths that meet the constraints to the next AS
- Able to calculate shortest path that meets the constraints
  - Limitation: Topology at the AS level must be a tree

Distributed Path Computation



MPLS 2004

Cisco Systems

Comparison of Approaches

PER AS PATH CALCULATION

- No impact on routing or signaling scalability
- Minor protocol extensions
- Doesn't find shortest path in general
- May fail to find paths that exist

DISTRIBUTED PCE APPROACH

- No impact on routing or signaling scalability
- More complex protocol extensions and need for PCEs
- Will find shortest path in general
- Will find a path if one exists

Bottom Line: Two Valid Approaches, Complexity vs. Optimality Tradeoff

Measurement & Monitoring

Importance of measurement

- Customer wants to know if his SLAs are getting met
- Providers need to be able to
  - Monitor the performance of peers
  - Troubleshoot & locate cause of SLA violations
  - Verify customer problems
- Much more demanding than single provider case

Measurement Approaches

- Active
  - Enhanced ping
  - Possible to gather lots of data but at a cost (extra traffic & processing)
- Passive
  - Attempt to infer network behavior by monitoring user traffic

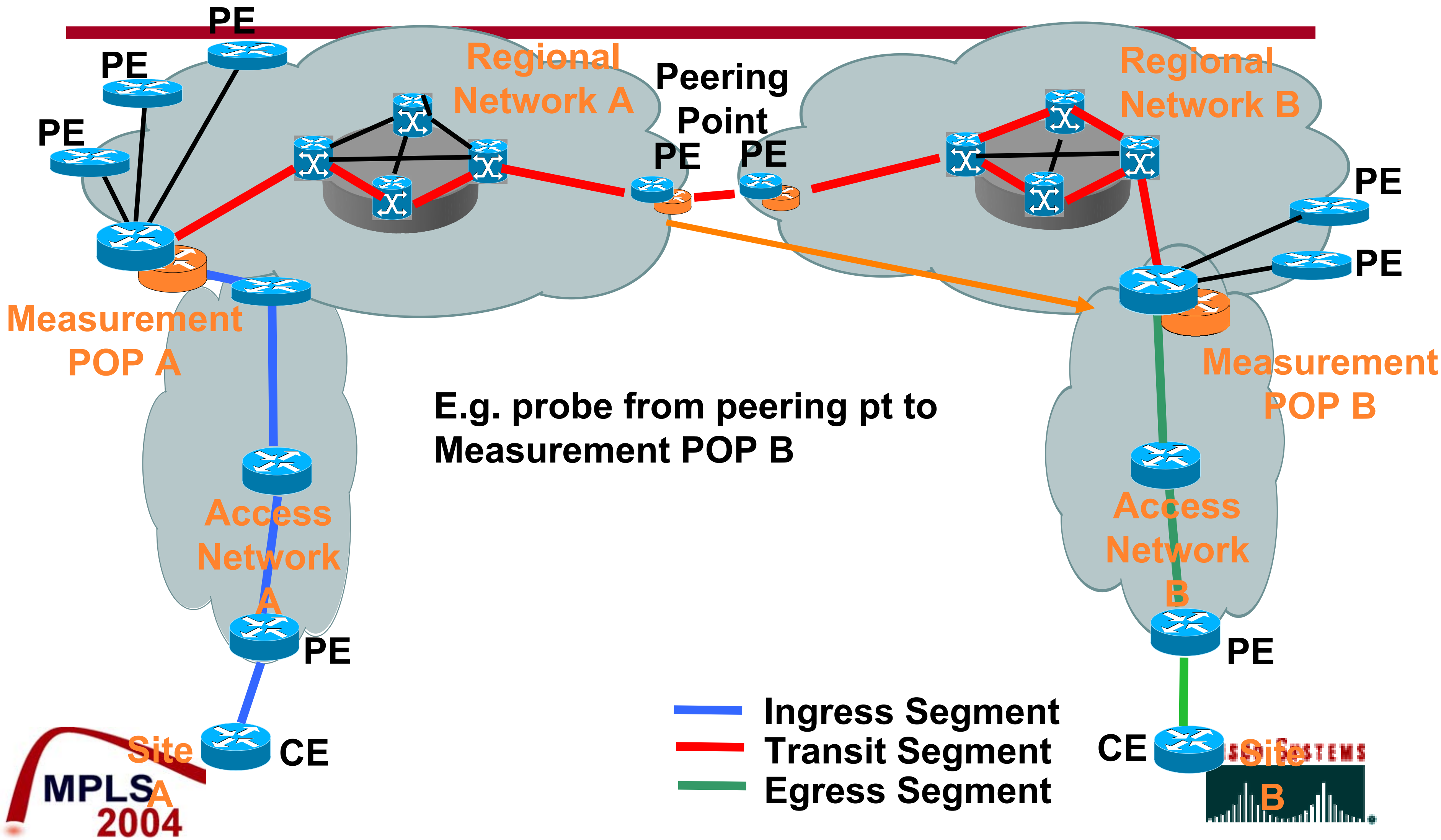
Active Measurement

- Ping enhancements - a standardization opportunity
- Cisco's "Service Assurance Agent" demonstrates one model:
  - UDP packet sent to a "responder"
  - Packet carries timestamp & sequence # for loss, delay, jitter measurement
  - Send probes with range of DSCP values and lengths to assess performance of all traffic classes

Scaling Active Measurement

- Getting a full picture of the network's performance by probing implies heavy probe load
  - Challenge for the probing platforms as well as generating extra traffic
- Possible approaches:
  - Probe less often when all is well, more often when troubleshooting
  - Probe segments of end-to-end path, rather than full mesh of PE-PE probes
  - "Trust but verify" - SP can report his measurements to other SPs, who may initiate probes to verify reported data

Scaling Active Measurement



E.g. probe from peering pt to Measurement POP B

Concluding Remarks

Next Steps

- Would be best to standardize
  - Performance metrics
  - Small set of service classes
  - Measurement techniques
  - Reporting methods
- SPs who want to support interprovider QoS need to
  - Work out bilateral agreements with each other (N^2), or
  - Agree on common approach among themselves (e.g. what measurement & reporting requirements must be met)
- Common approach may be easier for large N

Conclusions

- Inter-provider QoS is needed by customers to reap full benefits of MPLS VPNs
  - Thus, a revenue opportunity for providers
- Depends on some providers backing off from the “single omnipresent provider” approach
- Key technical requirements:
  - Performance metrics
  - Small set of common services offered by multiple providers
  - Routing support
  - Measurement and reporting
  - Freedom to adopt diverse implementation techniques