The Future of MPLS

A Service Provider Toolkit

Matt Kolon
MPLS Japan
October 2003
Service Provider Survival for 2004

- The market continually changes what service providers need to offer and support.
- Maintaining service flexibility is crucial.
- Internet access is only one of these services.
- Addressing all services with a single infrastructure is tricky.
- There is no place for technical dogma in engineering or planning.
Example Solutions: VPNs

- Layer 3 – 2547
- Layer 2 – Martini/Kompella
- VPLS
- Subscriber/Site level QOS
- IPSec Termination
A Versatile IP/MPLS Toolkit

• Versatility: ability to support multiple profitable services via a common set of tools
  • Facilitate creation of new services that supplement existing services
  • Reduce the cost of rolling out new services

Service providers need a toolkit that can support expanding existing service offerings with minimum additional effort/ cost.
Toolkit Requirements - Overall

- Simple Versatility
  - Have as few tools as possible…
  - … but no less than needed to support the services
  - A single operational infrastructure and a small set of basic tools mean cost savings in terms of:
    - Educating the NOC staff
    - Educating customers
    - Building tools/expertise to manage services
  - But simplicity is not the ultimate goal, profitability is
    - Simplicity → profitability
Toolkit Requirements - Details

- Support for multi-AS/multi-provider operations
- Multi-vendor interoperability based on open specifications
- Effective QoS Capability
- Internet-size Scalability
- TDM-like Availability
- Operational Economy
  - These are the “cost of entry” for nearly all potential services.
Inter-provider VPNs

- **Application**
  - Allows Service Providers to partner and extend geographic reach of VPN

- **Service Provider Benefits**
  - Standards based approach
  - Explore new markets without the need to build new infrastructure
  - Reach new customer base
  - New revenue stream

- **Enterprise Customer Benefits**
  - Reach - Common national or international networks
  - Single bill
  - Fewer relationships to manage
Other Inter-provider Work Needed!

- Billing
- Accounting
- QoS
- Traffic Engineering
- Settlement / Metering
- Session Negotiation
- Essentially, we are talking about one or more new NNI definitions.
Toolkit Requirements - Details

- Support for multi-AS/multi-provider operations
- Multi-vendor interoperability based on open specifications
- Effective QoS capability
- Internet-size Scalability
- TDM-like Availability
- Operational Economy
  - These are the “cost of entry” for nearly all potential services.
Multi-vendor Interoperability

• What matters:
  • Interoperability
  • Time to market
  • A fitting solution to a well-defined problem

• What does NOT matter:
  • IETF Standards Status
  • Internet Draft + working code is necessary, but also sufficient.
  • Credible vendors and providers understand this!
IETF Standards and time to market

- RSVP TE extensions:
  - November 1998 Internet Draft - December 2001 Proposed Standard
- OSPF TE extensions:
  - Internet Draft since April 1999 – still not an IETF standard
- 2547 VPN (aka BGP/MPLS VPN)
  - Internet Draft since March 2000 – still not an IETF standard
- draft-martini
  - Internet Draft since November 2000 – still not an IETF standard

Going through the IETF standards process takes longer and longer… speeding up the IETF process may take even longer…
Toolkit Requirements - Details

- Support for multi-AS/multi-provider operations
- Multi-vendor interoperability based on open specifications
- Effective QoS capability
- Internet-size Scalability
- TDM-like Availability
- Operational Economy
  - These are the “cost of entry” for nearly all potential services.
Class of Service
Diff-Serv – MPLS CoS Model

**Ingress Edge Function:**
- Upstream flows policed
- Downstream flows shaped
- Aggregate many flows & associate with a Class (marked with DSCP)
- DSCP mapped to EXP bits in MPLS header
- Queue and drop accordingly

**Core Function:**
- Perform the appropriate queuing and dropping based on class (EXP bit marking)

**Egress Edge Function:**
- Queue and drop accordingly based on class (DSCP)

Simple but no guarantees
Class of Service
MPLS Traffic Engineering

MPLS TE - RSVP
u Constraint-based LSP setup - computes a path with sufficient unreserved TE bandwidth
u Allows efficient use of available bandwidth
u Each link has a single bandwidth pool available

No interaction with Diff-Serv - all traffic is treated the same
The Next Step: Diff-Serv Traffic Engineering

- Policing, Shaping plus Admission Control per Class
- Insufficient unreserved TE bandwidth for Premium Class Traffic
- Insufficient unreserved TE bandwidth for Premium Class Traffic

Diff-Serv TE guarantees bandwidth!

- Connects Diff-Serv classes with RSVP traffic engineering database
- Admission Control per Class based bandwidth pools
- Sessions refused to maintain guaranteed bandwidth for traffic class
- Constraint Based LSP setup - per Class with different bandwidth constraints
- LSP path setup now governed by unreserved bandwidth per class
Toolkit Requirements - Details

- Support for multi-AS/multi-provider operations
- Multi-vendor interoperability based on open specifications
- Effective QoS capability
- Internet-size Scalability
- TDM-like Availability
- Operational Economy
  - These are the “cost of entry” for nearly all potential services.
Scalability in the proper context: RSVP as an example

Outside of TE context:

• “supporting numerous small reservations on a high-bandwidth link may easily overtax the routers and is inadvisable” (RFC2208)
• “the resource requirements for running RSVP on a router increases proportionally with the number of separate sessions” (RFC2208)

In the context of TE:

• We are not dealing with small reservations, as one TE LSP aggregates many micro flows
• Total number of RSVP sessions (total number of TE LSPs) is bounded by the number of edge routers of a service provider, and is independent of the total volume of traffic carried by the service provider
More than one way to scale

Goal: limit the volume of information that a single box must handle

- By information aggregation/abstraction via hierarchical routing
  - used for scaling the Internet service (CIDR)
- By partitioning the problem into a set of independent sub-problems
  - used for scaling 2547 VPNs, BGP autodiscovery/signaling for VPLS, and L2 VPNs services
Toolkit Scalability: BGP

- Excellent example of appropriate protocol re-use
  - Began life as an inter-domain routing protocol
  - Has assumed additional service control roles
    - 2547bis, VPLS, kompella-draft, etc.
- Leverages knowledge and infrastructure already paid for by service providers.
- Keeps the toolkit manageable, without limiting flexibility or performance.
Toolkit Requirements - Details

• Support for multi-AS/multi-provider operations
• Multi-vendor interoperability based on open specifications
• Effective QoS capability
• Internet-size Scalability
• TDM-like Availability
• Operational Economy
  • These are the “cost of entry” for nearly all potential services.
Reliability: Toolkit High Availability

• Goal: minimize service disruption due to faults
  • handle both control plane and data plane faults
• Control Plane graceful restart and hitless switchover for handling faults in the control plane
  • Requires the ability to preserve/maintain data plane in the presence of control plane faults
• MPLS Fast Reroute for handling faults in the data plane
  • More than one kind of FRR, but that’s another story...
Graceful Restart

- Separate control and data planes
- If router’s control plane fails, data plane can keep forwarding packets
- Neighbors hide failure from all others routers in the network
- Other routers are never made aware of failure
- When router recovers, neighbors sync up without disturbing forwarding.
Hitless RE Switchover

- Protects against Single Node Hardware Failure
- Primary (REP) and Secondary (RES) utilize keepalive process
  - Automatic failover to RES
  - Synchronized Configuration
- REP and RES share:
  - Forwarding info + PFE config
- REP failure does not reset PFE
  - No forwarding interruption
  - Only Management sessions lost
  - Alarms, SNMP traps on failover

Not the Same As Protocol Stateful Mirroring
MPLS Fast Reroute

- Single user command at head end to enable Fast Reroute.

- Fast reroute is signaled to each LSR in the path.
- Each LSR computes and sets up a detour path that avoids the next link and next LSR.
- Each LSR along the path uses the same route constraints used by the head-end LSR.
Toolkit Requirements - Details

- Support for multi-AS/multi-provider operations
- Multi-vendor interoperability based on open specifications
- Effective QoS capability
- Internet-size Scalability
- TDM-like Availability
- Operational Economy
  - These are the “cost of entry” for nearly all potential services.
Economical Ops: Resource sharing

- By competent system design
- By competent software engineering
- Examples:
  - ability to limit memory usage by individual 2547 VPN customers by imposing an upper bound on the number of routes in a VRF
  - ability to limit CPU usage by individual 2547 VPN customers by limiting the rate of control (routing) traffic on the CE-PE link
  - ability to limit bandwidth usage by individual 2547 VPN customers
  - ability to use different routing protocol process for each service
- There is an existence proof (products on the market) that it is possible to build multi-service platforms that provide adequate resource sharing
Economical Ops: Helping Staff

• Introduce new operational practices carefully
• Tools and software can ease new procedures
• Train, train, train!
• Certification
• All of these lead to better customer service
The Toolkit meets requirements

- By taking advantage of the commonalities between multiple services
- By using tools that are general and easily extendable
- By using tools that have appropriate scaling properties
- By using tools that can operate across multiple ASs/service providers
- By using tools that are applicable to a wide variety of multiplexing and switching technologies (e.g., routers, SONET/SDH Cross Connects, Optical Cross Connects, etc.)
- By documenting specifications in Internet Drafts, and by using multi-vendor interoperability testing events
Simple Versatility

- Small set of tools, many possible uses:
  - IP/{G}MPLS data and forwarding plane
    - Diffserve QoS capability
  - BGP with Multi-Protocol Extensions
  - OSPF/ISIS with TE extensions
  - RSVP-TE
  - LDP
  - Graceful restart for all protocols (BGP, OSPF/ISIS, RSVP-TE, LDP)
Summary

IP/MPLS Toolkit:

Enabling providers to deliver as many profitable services as possible with as few tools as possible
Thank you!

matt@juniper.net