

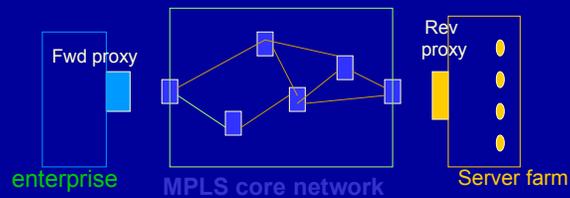
MPLS Based Web Switching

Arup Acharya, Anees Shaikh,
Renu Tewari, Dinesh Verma

IBM TJ Watson Research Center

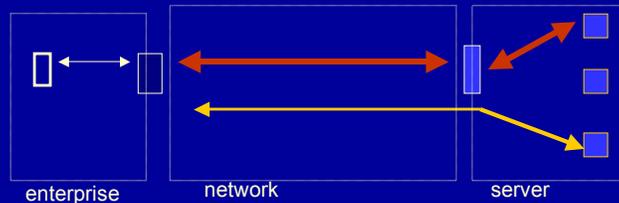
Two Observations/ Trends

- Expected use of **MPLS** in core networks
 - Global Crossing, AT&T Canada have deployed MPLS
 - Use of MPLS as the control plane in optical networks
 - MPLS supported by most router vendors (Cisco, Juniper, Nortel,....)
 - ◆ MPLS switches (L2/ L3) will be “commodity” ?
 - ◆ MPLS/IP in storage area networks
- Specialised **L4-L7 web switches** in front of server farms for
 - Alteon(Nortel), F5, Arrowpoint (Cisco), Foundry
- Can we leverage MPLS for web switching?



Different types of web switching

- ◆ **Content routing**: partitioned content, need to route request to the right server
- ◆ **Load balancing**: replicated content
- ◆ **Affinity**: direct http requests within a session to the same server
- ◆ **Service differentiation**: handle requests differently depending on point of origin/enterprise (e.g., predetermined SLAs)

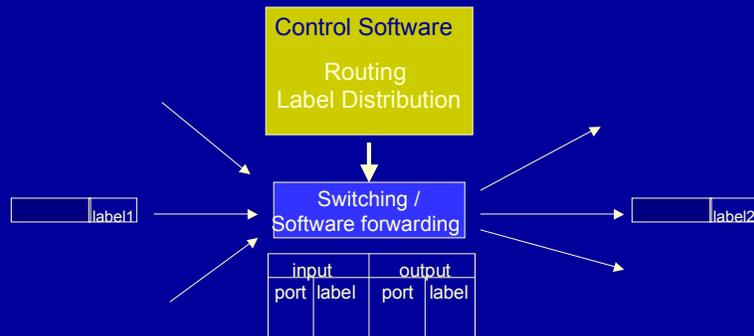


WWW dispatchers : two basic approaches

- ◆ **TCP/IP header for dispatch decisions (layer 4)**
 - Limited functionality : load balancing, affinity (per src addr)
 - High performance (specialized hardware implementations)
 - IBM ND, Cisco LD
- ◆ **Application-layer headers**
 - Requires TCP termination/splicing (layer 7)
 - Content routing, affinity, load balancing ...
 - Bottleneck : TCP connection termination
- ◆ **Our goal : combine L7 flexibility with L4 performance**
 - Eliminate TCP termination bottleneck
 - Use standard hardware instead of specialized switches
 - Leverage use of MPLS in core networks
 - Better price/performance ratio

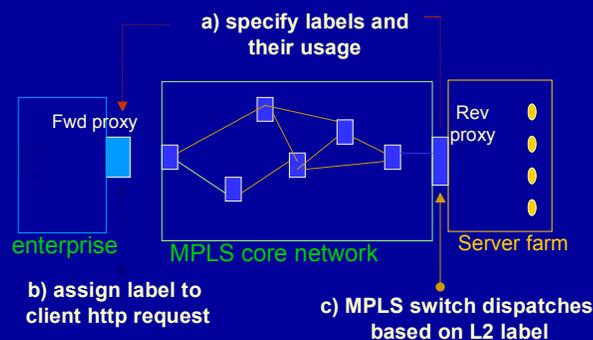
MPLS basic operation

- Label placed between IP and MAC addresses
 - ◆ Labels defined for PPP, ethernet, ATM, Frame relay
 - ◆ Pkt forwarding based on label lookup, not IP address
- Signaling protocols to distribute labels, create a label-switched path
- *Switching hardware is agnostic of label semantics*



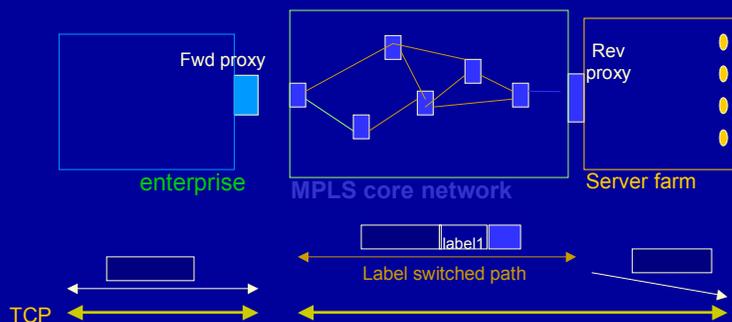
Proposed Approach

- Map L4-L7 semantics onto MPLS labels and use commodity MPLS switches for web-switching (in MPLS, L3/routing semantics applied to L2 labels)
 - No TCP termination at reverse proxy and yet, achieve content routing, load balancing, affinity
 - Out of path return allowed
- Requires participation of client-side proxy in web-switching decisions



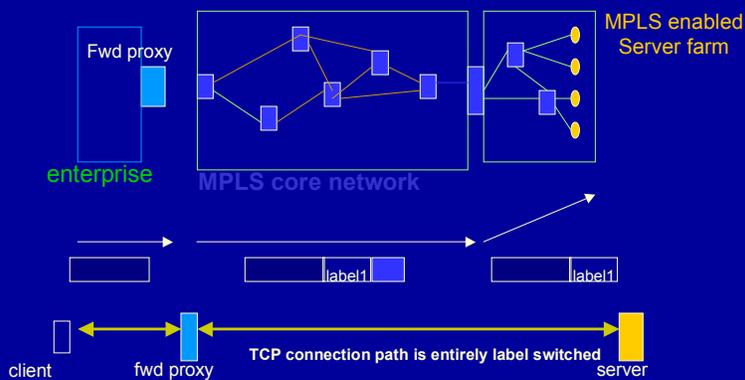
MPLS label stacking

- Outer label used for routing in the core
- Inner label used for web switching
- Reverse proxy communicates semantics of (inner) label to fwd proxy
 - Example: content based routing
 - Forward proxy assigns inner label based on URL
 - Label1 → www.cnn.com/headlinenews/
 - Label2 → www.cnn.com/fn/
- Reverse proxy routes to the right server based on inner label



Server farm supports MPLS

- If the server farm network is also MPLS enabled, then label switched path from forward proxy all the way to server
 - Dispatcher is a MPLS switch
 - Inner label allows desired L4-L7 functionality (content routing,.....)

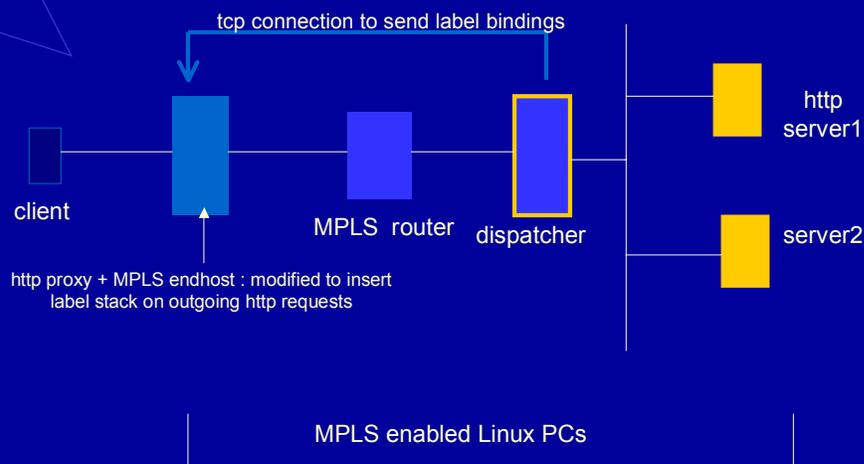


Label Distribution

- ◆ Control connection between dispatcher and proxy
 - Content based routing : URL⇒label mappings
 - Load balancing : set of labels, proxy round-robins client requests across labels (could also send weights+labels)
 - Affinity : set of labels, proxy assigns same label to all client requests within a given time T
 - Service differentiation : proxy receives sets of labels, one set per class (gold/.../..), assumes pre-defined service agreement
- ◆ Dispatcher populates label table at layer2

Prototype

- ◆ Objectives : # of simultaneous http requests scales linearly with #servers, (i.e. the dispatcher is not a bottleneck), and yet support content routing,



Deployment Scenarios

- ISP offering Web hosting service (proxies and server farms are controlled by a single entity)
- Web hosting provider has SLAs with particular clients

- If proxy and dispatcher are in different MPLS domains, the solution will depend on (yet undefined) inter-domain MPLS standards
- Alternative to MPLS labels : use IP addresses or port# ?
 - Conceptually yes, but does not leverage commodity layer2 switching hardware

THANKS!

Contact : arup/tewarir/dverma@us.ibm.com