MPLS tutorial

Roopa Prabhu—Cumulus Networks

Feb 12th, 2016
Agenda

- Brief introduction to MPLS
- History
- Linux kernel MPLS LSR support
- MPLS LSR multipath support
- Linux MPLS IP tunnels
- Futures
What is MPLS (multiprotocol label switching) ?

- MPLS is a protocol-independent transport
- Packets are assigned labels
- Packet forwarding decisions are made solely on the basis of labels
- MPLS operates at a layer between data link layer (layer 2) and network layer (layer 3) and is referred to 2.5 layer
- Mainly used in service provider networks: Can carry IP, ATM, frame relay traffic
History

- MPLS evolved as a solution to integrate IP over ATM
- The success of MPLS is a result of the fact that it enables the network to carry all kinds of traffic
Benefits

● unified network infrastructure: many technologies can be transported over it

● the service provider needs only one unified network infrastructure to carry all kinds of customer traffic
Routers in a MPLS domain

- **First router (Label Edge router or LER)**
  - IP routing lookup
  - Attaches labels
  - Forwards based on label
- **Future routers (Label switching routers or LSR)**
  - Use label to route
- **Final destination router (Label Edge router or LER)**
  - Removes label
  - Packet is delivered using normal routing
MPLS Router Domains

IP → LER → LSR → LSR → LER → IP

Push label → swap label → swap label → Pop label
MPLS Deployment example
MPLS label stack entry

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ Label
|                Label                        | Exp |S|   TTL   | Stack
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ Entry

Label: Label Value, 20 bits
Exp: Experimental Use, 3 bits
S: Bottom of Stack, 1 bit
TTL: Time to Live, 8 bits
Linux MPLS Label switching router
MPLS Label switching router

- **CONFIG_MPLS, CONFIG_MPLS_ROUTING**
  - driver: net/mpls/af_mpls.c

- **UAPI:**
  - new route netlink attributes: RTA_NEWDST, RTA_VIA
MPLS LSR: add mpls route

- enable mpls on a network interface
  - `echo 1 > /proc/sys/net/mpls/conf/swp1/input`

- `iproute2`:

  
  ```
  $ip -f mpls route add 100 as 200/300 via inet 10.1.1.2 dev swp1
  $ip -f mpls route show
  100 as to 200/300 via inet 10.1.1.2 dev swp1
  ```
MPLS LSR: add mpls multipath route

```
$ip -f mpls route add 100 \
   nexthop as 200 via inet 10.1.1.2 dev swp1 \
   nexthop as 300 via inet 10.1.1.6 dev swp2

$ip -f mpls route show
100
   nexthop as to 200 via inet 10.1.1.2  dev swp1
   nexthop as to 300 via inet 10.1.1.6  dev swp2
```
Linux MPLS Label Edge router
Light weight tunnel infrastructure (LWT)

- Attach tunnel attributes to routes
- New tunneling abstraction API:
  - Register/Unregister lwtunnel encap ops
  - lwtunnel ops
    - parse/dump encap attributes into per route nexthop lwtunnel state
    - Redirect output to tunnel output handler during packet forwarding
Light weight tunnel infrastructure (LWT) Contd..

- **CONFIG_LWTUNNEL**
  - net/core/lwtunnel.c

- New netlink attributes:
  - RTA_ENCAP, RTA_ENCAP_TYPE

- iproute2:
  - $ip route add help
  -[..snip..]
  - NH := [ encap ENCAPTYPE ENCAPHDR ] [ via [ FAMILY ] ADDRESS ]
    [ dev STRING ] [ weight NUMBER ] NHFLAGS
  - ENCAPTYPE := [ mpls ]
  - ENCAPHDR := [ MPLSLABEL ]
MPLS tunnels using LWT infrastructure

- Implements MPLS ingress LER function
- CONFIG_MPLS_IPTUNNEL
- MPLS ip tunnel driver
  - net/mpls/mpls_iptunnel.c
MPLS tunnel routes: iproute2

$ip route add 10.1.1.0/30 encap mpls 200 via inet 10.1.1.1 dev swp1
$ip route show
10.1.1.0/30  encap mpls 200 via 10.1.1.1 dev swp1

$ip route add 10.1.1.0/30 nexthop encap mpls 200/300 via 10.1.1.1 dev swp1 |
    nexthop encap mpls 700/800 via 40.1.1.2 dev swp3
$ip route show
10.1.1.0/30
    nexthop encap mpls 200/300 via 10.1.1.1  dev swp1
    nexthop encap mpls 700/800 via 40.1.1.2  dev swp3
MPLS tunnel routes: iproute2 (IPV6)

$ ip -6 route add 2001:10:2::/64 encap mpls 300/400 dev swp1

$ip -6 route show
2001:10:2::/64  encap mpls 300/400 via 2001:10:3::/64 dev swp1

$ip -6 route add 2001:10:2::/64 nexthop encap mpls 200 via 2001:10:3::/64 dev swp1 \nexithop encap mpls 700 via 2001:10:4::/64 dev swp3

$ip -6 route show
2001:10:2::/64
  nexthop encap mpls 200 via 2001:10:3::/64 dev swp1
  nexthop encap mpls 700 via 2001:10:4::/64 dev swp3
Next things

- MPLS stats (in progress)
- MPLS L3-VPN support
- MPLS switchdev offload support
kernel versions

- MPLS LSR support: v4.1
- lwt and MPLS ip tunnel: v4.3
- MPLS multipath: v4.5
References

Questions
Bringing the Linux Revolution to Networking

Thank You!