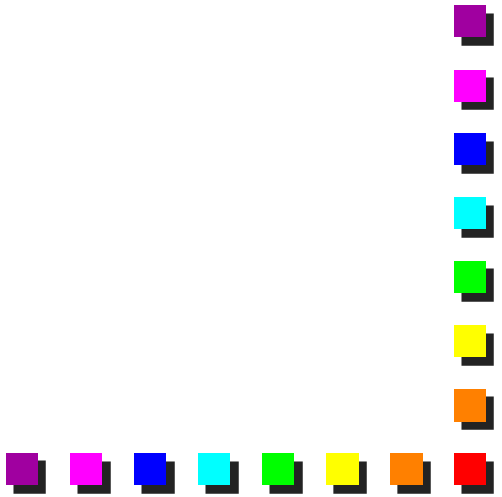
The Network Layer in Local Area Networks

Fulvio Risso

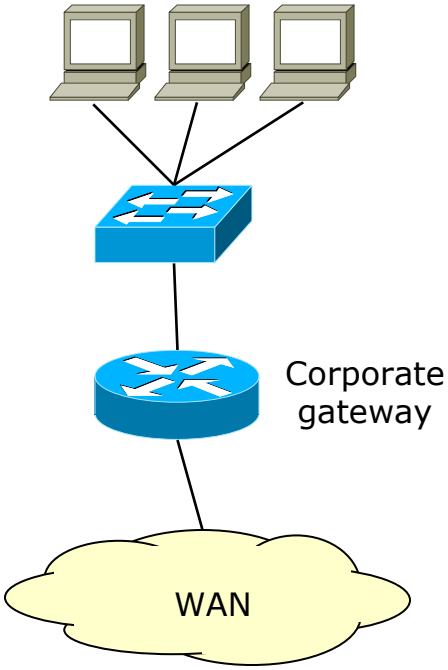
Politecnico di Torino

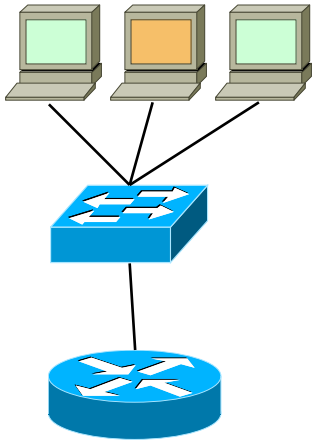


Copyright notice

LANs and Routers

- Routers are a fundamental part of a LAN
  - We cannot imagine a network without access to the Internet and/or some other corporate networks across WAN
  - We cannot image corporate networks without VLANs





One-arm router

The fundamental question in LAN design

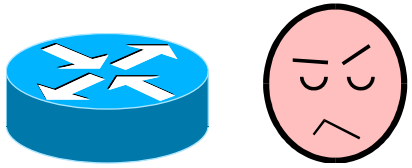
- Do we have to use routers only for VLAN interconnection and to connect to the Internet?
- (or) we have to put more routers in our networks?

- In that case, where do we have to stop L2 and start L3?
  - At the edge of the corporate network (exit gateway)
  - In the core of the corporate network (backbone)
  - In the distribution of the corporate network (distribution)
  - In the access (e.g. all hosts within different /30 networks), with not L2 at all

Main strength of L2 vs. L3

- Simple
  - To produce → very cheap
  - To use → plug and play deployment
- High performance
- Seamless mobility at L3
- Transparent

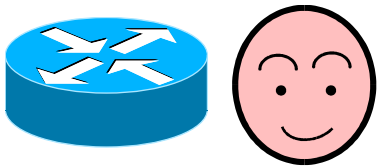
- Complex
  - Processing (extract layer 3 packets, update them, build a new layer 2 frame)
  - Longest prefix matching
  - Sophisticated routing protocols
- Slower processing speed
- More expensive
- No mobility
- No transparency
  - Hosts need to be configured



L2 vs. L3: Security

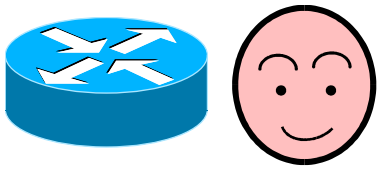
- No network isolation
  - ARP spoofing attack, MAC flooding attack, …
- Filtering database populated automatically
- ACLs need to be defined "per MAC"
- Not enough intelligence to operate "per TCP/UDP port", "per application", etc.

- Network isolation for better security
- IP addresses configured by the network manager
- ACL can be defined "per network"
- Usually more intelligence also at transport and application layer



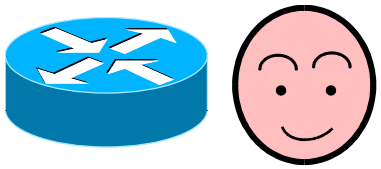
L2 vs. L3: Scalability – Addressing

- Flat addressing
  - Each reachable host must first be explicitly listed in the filtering database
- Hierarchical addressing
  - A single entry can forward data to multiple destinations
    - E.g., destinations with the same network prefix
  - Ensures a limited increase in the size of the forwarding table when the network grows
    - The granularity of the aggregation grows



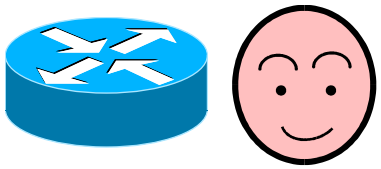
L2 vs. L3: Scalability – Broadcast traffic

- The broadcast traffic grows with the number of hosts
- Broadcast storm vulnerability
  - Closed paths (meshes)
  - Interface malfunctions
- Routers do not forward broadcast traffic



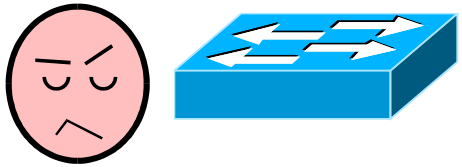
L2 vs. L3: Scalability – Network size

- Limited network diameter
  - IEEE 802.1D suggests no more than 7 cascading switches
  - A fine tuning of parameters allows a greater number
  - IEEE 802.1s to ensure functionality on big networks
- No limits in the network diameter
  - Multiple routing domains, if needed

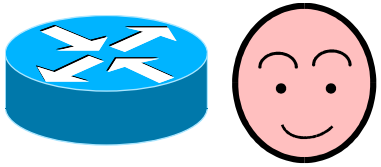


L2 vs. L3: Network paths

- Limited routing capabilities
  - All traffic is forwarded using the same spanning tree
    - Suboptimal paths
    - The usage of network resources is not fair
      - The links that are outside the spanning tree are not used!



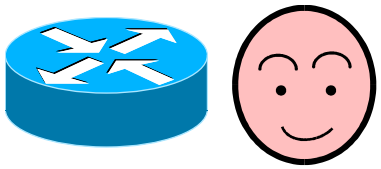
- Multiple forwarding trees are used
  - Multipath routing available too



L2 vs. L3: Transient

- Slow failure recovery
  - It may take about a minute
  - Better with IEEE 802.1w (Rapid Spanning Tree Protocol)
- Possible flooding of data traffic during transient
  - Artificial ageing of filtering database entries

- Multiple forwarding trees are used
  - Multipath routing available too
- Network outage limited to a subset of the network or to a subset of destination
  - The others are usually marginally affected by the outage



L2 vs. L3: which one is the better?

- It depends…
  - Simplicity and costs are usually much more important than other parameters
- Some common choices:
  - Access at L2 (with VLANs)
  - Backbone still L2
    - Sometimes L3
  - L3 for the exit gateway and VLAN interconnection
    - Sometimes for the backbone; unusual in the access
  - L4-7
    - Exit gateway (for protection)
    - Data-center (for load balancing, QoS, etc).

Routers and performance

- A router is slow
- A switch is fast

- What would happen to the performance of my network if I rely on L3 devices?

- Answers (in the past)
  - Use L2 where you can, and limit L3 as much as possible
  - Let's develop new technologies which are faster
    - Asynchronous Transfer Mode

Mixing L2 and L3 with the network

- Three problems
  - Router redundancy
    - Routers are not transparent; L2 switches are
  - Performance
    - We may want routers as fast as L2 switches
  - Network optimization
    - Data may cross the same link twice

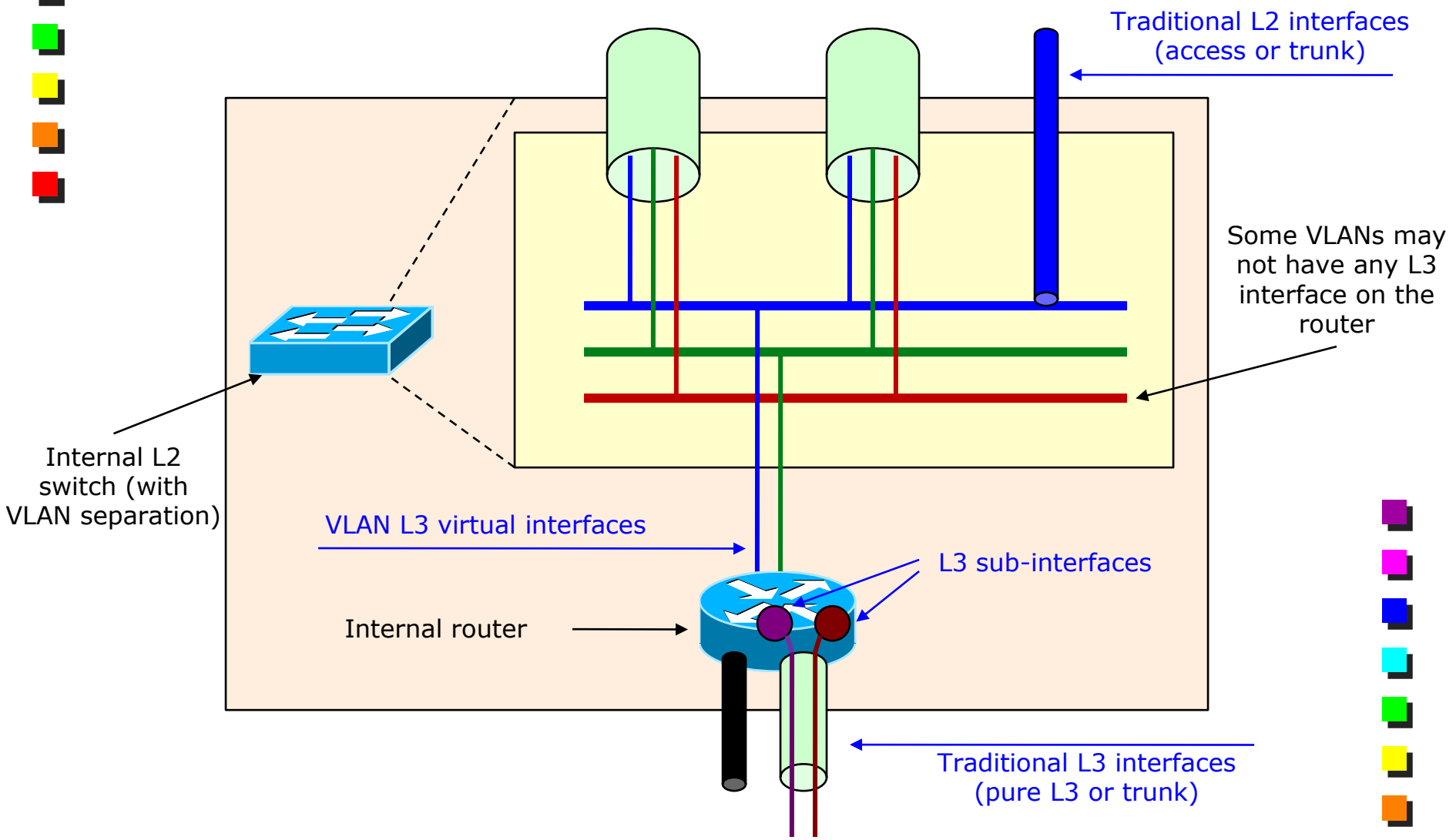
Layer 3 Switch

- New term invented by the marketing at the end of 90s
  - Switch reminds something that goes faster and faster
- Layer3 switch = router
- Differences mainly in how features are implemented
  - L3 switch is usually a pure hardware device
    - Although currently also high-end routers are pure hardware
  - Usually more oriented to corporate needs
    - No sophisticated routing protocols (e.g. BGP)
    - Access Control Lists (ACL) in hardware (for security)
    - Additional features
      - Content accelerators
      - Firewalls
    - Limited set of network interfaces

Multilayer switch (1)

- Device that integrates both L2 and L3 capabilities in the same box
  - First multilayer was a switch and a router (on two different cards) in the same chassis
  - Possible because processing capabilities are order of magnitude better than 10 years ago
- Customer can buy a multilayer switch and the configure the interfaces in L2 or L3 mode according to its needs
- L2 and L3 is the most common form of multilayer

Multilayer switch (2)



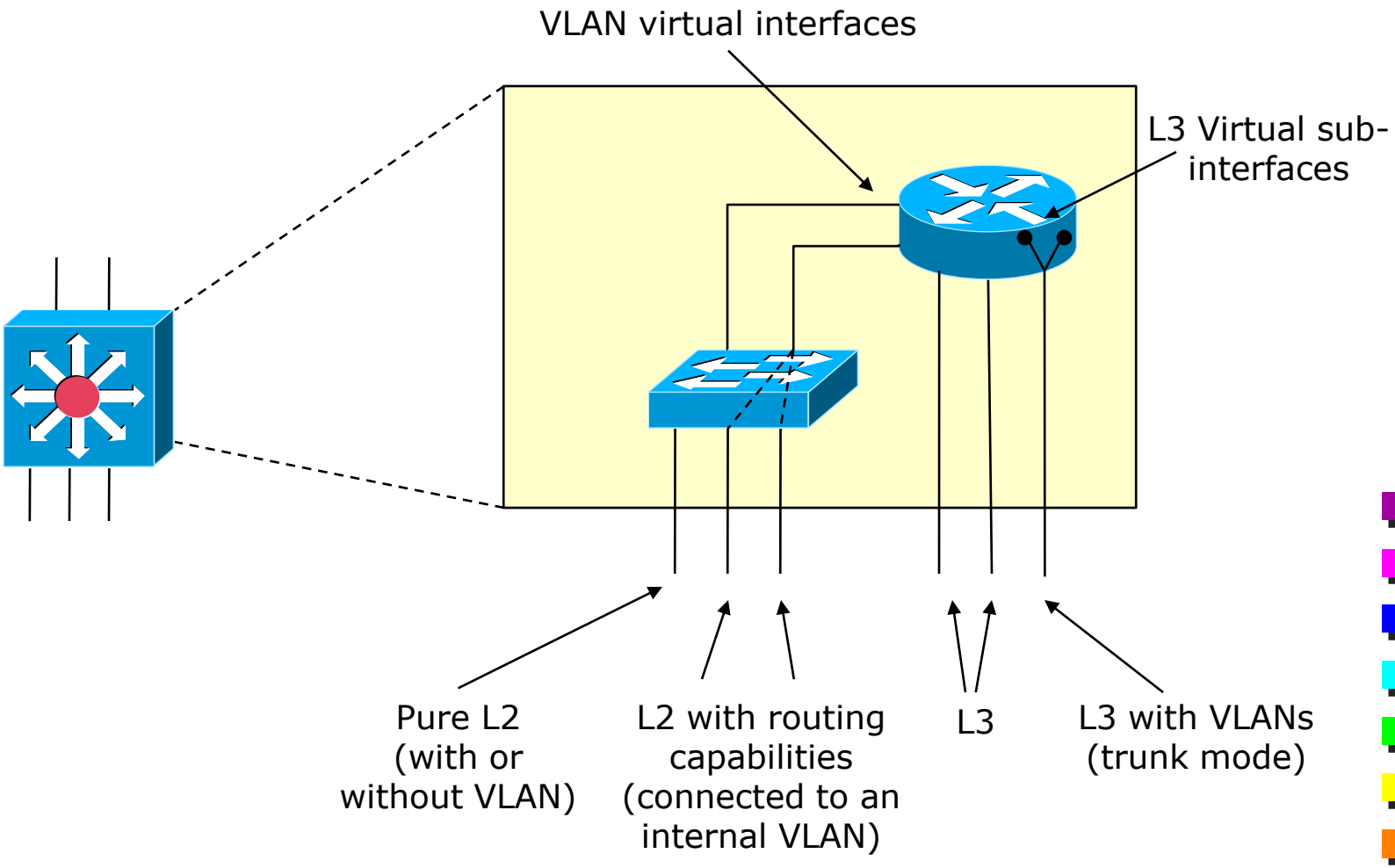
Interfaces in a L2-L3 switch (1)

- Four kinds of interfaces can be configured
- 1) “Traditional” L2
  - Pure L2, either in access or trunk mode
  - It may be further “associated” to a L3 VLAN interface (that will ac as one-arm router)
- 2) “Traditional” L3
  - Traditional routed interface (connected to a switch in access mode)
  - Routed interface connected in “trunk” mode
    - In this case it originates logical sub-interfaces in order to handle VLANs

Interfaces in a L2-L3 switch (2)

- 3) Virtual VLAN interface (L3)
  - Virtual interface
  - Connects the "internal router" to the "internal switch", and acts as one of the interfaces of the one-arm router (for VLAN interconnection)
- 4) Sub-interface for VLANs (L3)
  - Virtual interface at L3
  - Terminates a VLAN onto a sub-interface
- May not be all available at the same time
- Each type of interface accepts specific configuration commands
  - E.g., L2 interfaces cannot accept an IP address, while L3 interfaces do

Interfaces in a L2-L3 switch (3)



Interfaces in a L2-L3 switch (4)

- L2 interface
  - STP active

```
!
 interface fastethernet0
!
```

- L2 interface connected to a VLAN
  - Equivalent to a pure L2 interface
  - Optionally, it may be associated to a virtual VLAN interface (if routing capabilities are required)
  - STP active

```
!
 interface fastethernet0
   switchport access vlan 2
!
```

Interfaces in a L2-L3 switch (5)

- VLAN interfaces (virtual interfaces within the switch)
  - Equivalent to a pure L3 interface (although virtual)
  - No spanning tree active

```
!
 interface vlan2
   ip address 1.1.1.1 255.255.255.0
!
```

Interfaces in a L2-L3 switch (6)

- Pure L3 interface
  - Looks like a L2 link connected to an host
  - No spanning tree active
    - BPDU may be received, but the interface cannot go in blocking state
    - An RSTP switch connects to it through an Edge port
  - Some devices have L3 as default, other have L2 as default (use "*no switchport*" command)

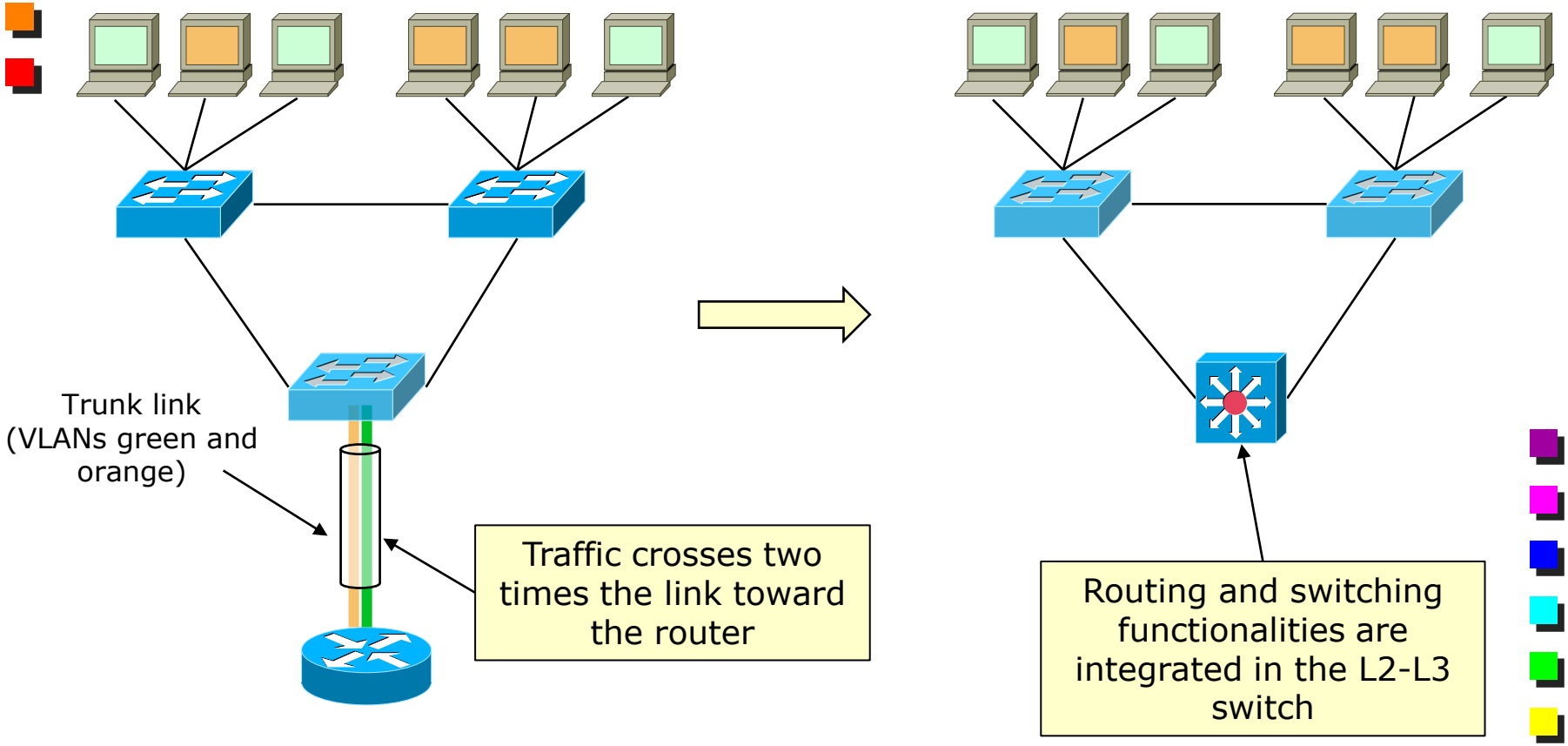
```
!
 interface fastethernet0/0
   ip address 1.1.1.1 255.255.255.0
!
interface fastethernet1/0
   no switchport
   ip address 2.2.2.2 255.255.255.0
!
```

Interfaces in a L2-L3 switch (7)

- Pure L3 interface in trunk mode
  - Plus sub-interfaces for VLAN handling

```
!
 interface fastethernet0/0
   no ip address
!
interface fastethernet0/0.10
   encapsulation dot1q 10
   ip address 10.10.10.1 255.255.255.0
!
interface fastethernet0/0.20
   encapsulation dot1q 20
   ip address 20.20.20.1 255.255.255.0
!
```

Network optimization with a multilayer



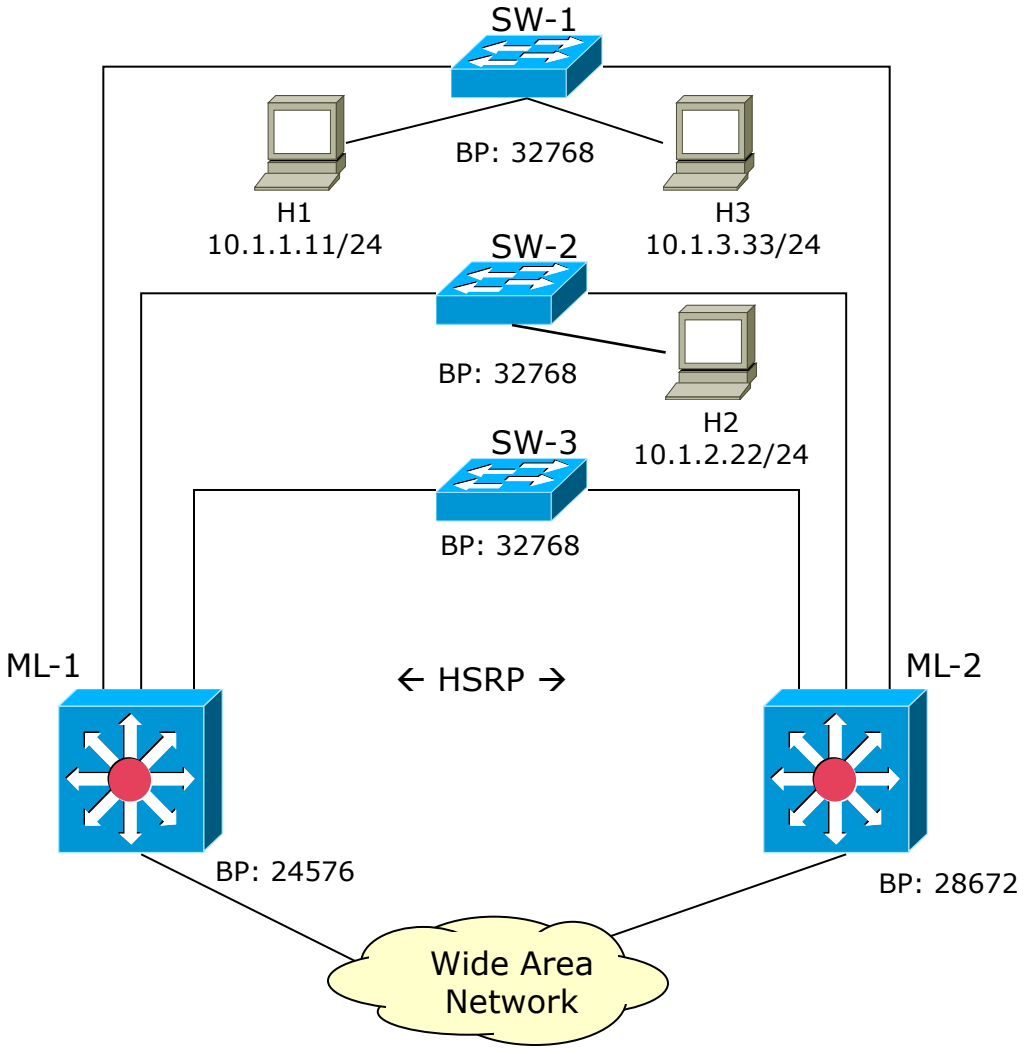
First configuration: L2 or L3?

Standard STP (no PVST)

3 IP networks

- 10.1.1.0/24
- 10.1.2.0/24
- 10.1.3.0/24

HSRP for Default Gateway redundancy

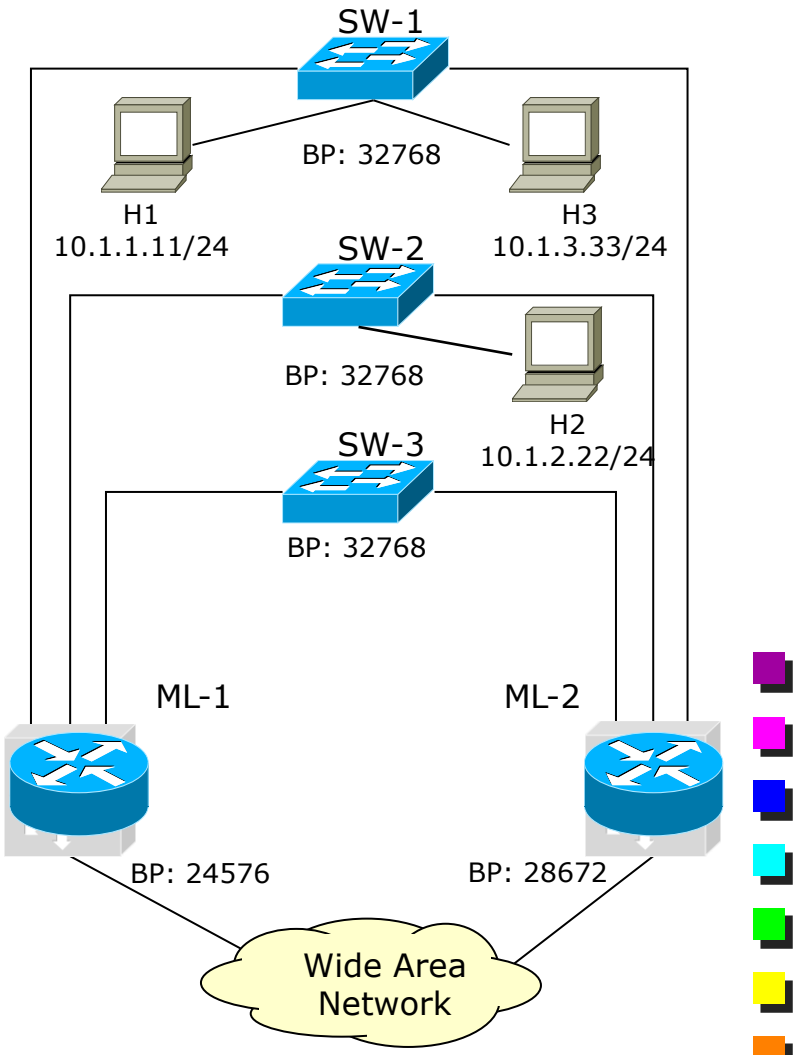


Does the multilayer operate at L2 or L3?

Let's see what happens with both options in the next slides

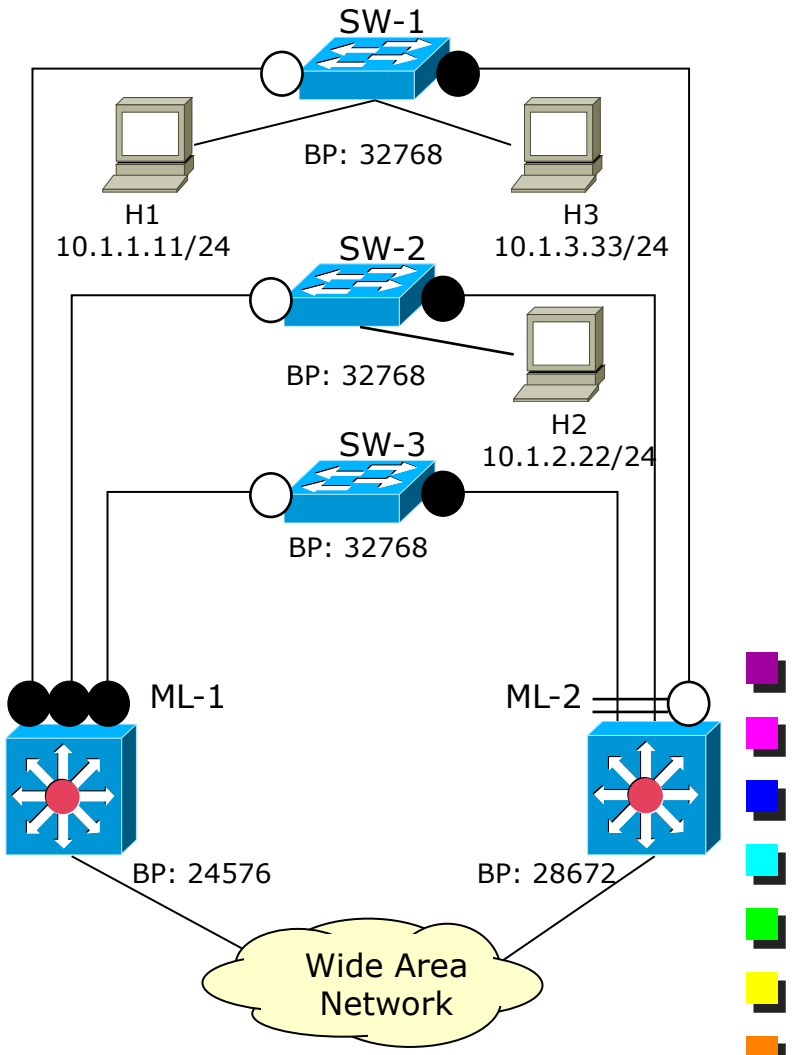
First configuration: L2 or L3? → L3

- Upper interfaces operating at L3
  - Each interface is associated to a specific L3 network
  - IP networks are statically associated to each floor switch (SW-x)
  - I.e., cannot distribute an IP network (e.g., 10.1.1.0/24) across multiple floor switches
  - VLANs are valid only within each floor switch
- Lower interfaces at L3 are fine
- Three distinct instances of STP
  - Bridge Priorities on ML-x are useless
- Not very flexible



First example: L2 or L3? → L2

- Upper interfaces operating at L2
  - IP addresses are associated to the internal VLAN interface of the multilayer
  - Possibility to distribute VLANs and IP networks across all the floor switches
- Single instance of STP
- Lower interfaces at L3 are fine



First configuration: L2 or L3? The final outcome

Single STP

3 VLANs for connecting the end users

All links betweens switches are in trunk mode

ML-1 Configuration

VLAN 1 address: 10.1.1.11/24

VLAN 2 address: 10.1.2.11/24

VLAN 3 address: 10.1.3.11/24

HSRP Group 1 (active): 10.1.1.3

HSRP Group 2 (active): 10.1.2.3

HSRP Group 3 (active): 10.1.3.3

ML-2 Configuration

VLAN 1 address: 10.1.1.2/24

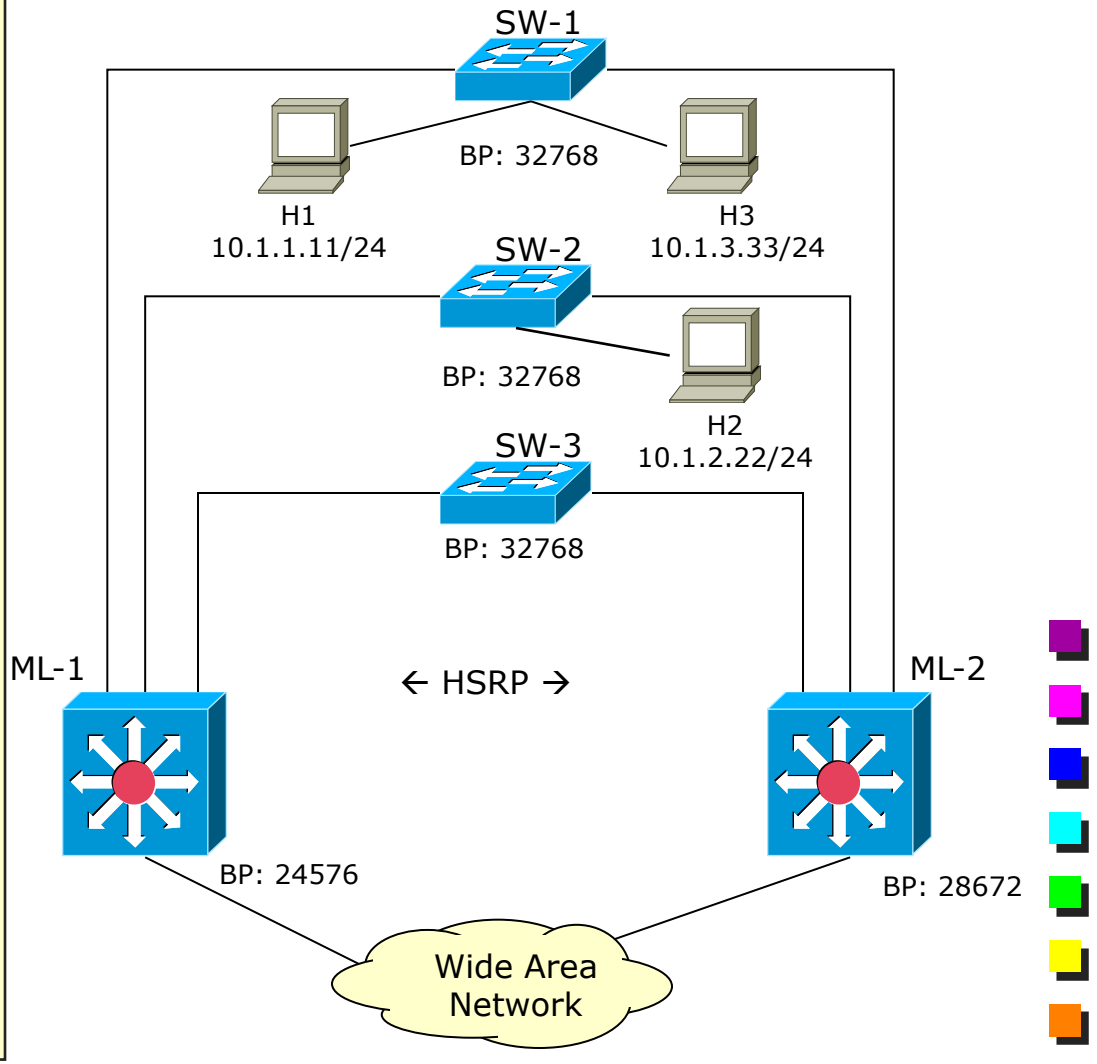
VLAN 2 address: 10.1.2.2/24

VLAN 3 address: 10.1.3.2/24

HSRP Group 1 (standby): 10.1.1.3

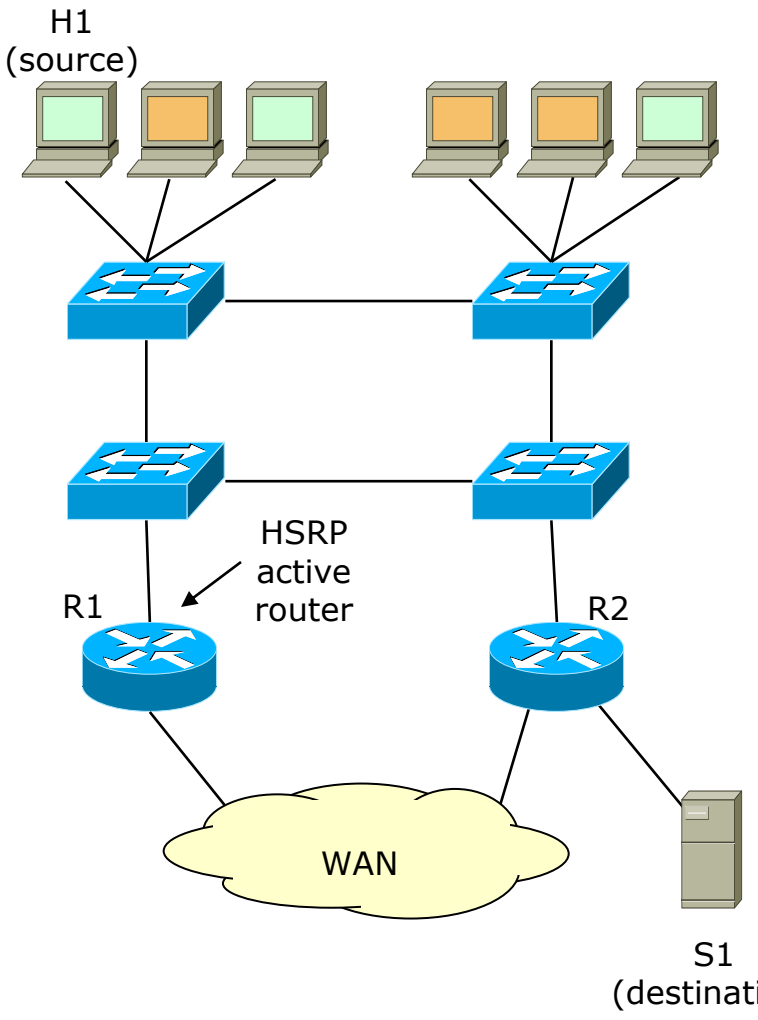
HSRP Group 2 (standby): 10.1.2.3

HSRP Group 3 (standby): 10.1.3.3



Additional L3 traffic within the LAN (1)

- L3 traffic from R1 and R2
  - R1 is the active router
  - The best path toward the destination involves router R2
  - R1 will forward that traffic to R2
- Do we want that traffic goes through the LAN?
- Other possible traffic: routing protocols between R1 and R2
  - Propagating in the WAN the list of IP networks present in the current campus
  - Receive from the WAN the list of IP networks available and decide which egress router is the best for reaching that network



Additional L3 traffic within the LAN (2)

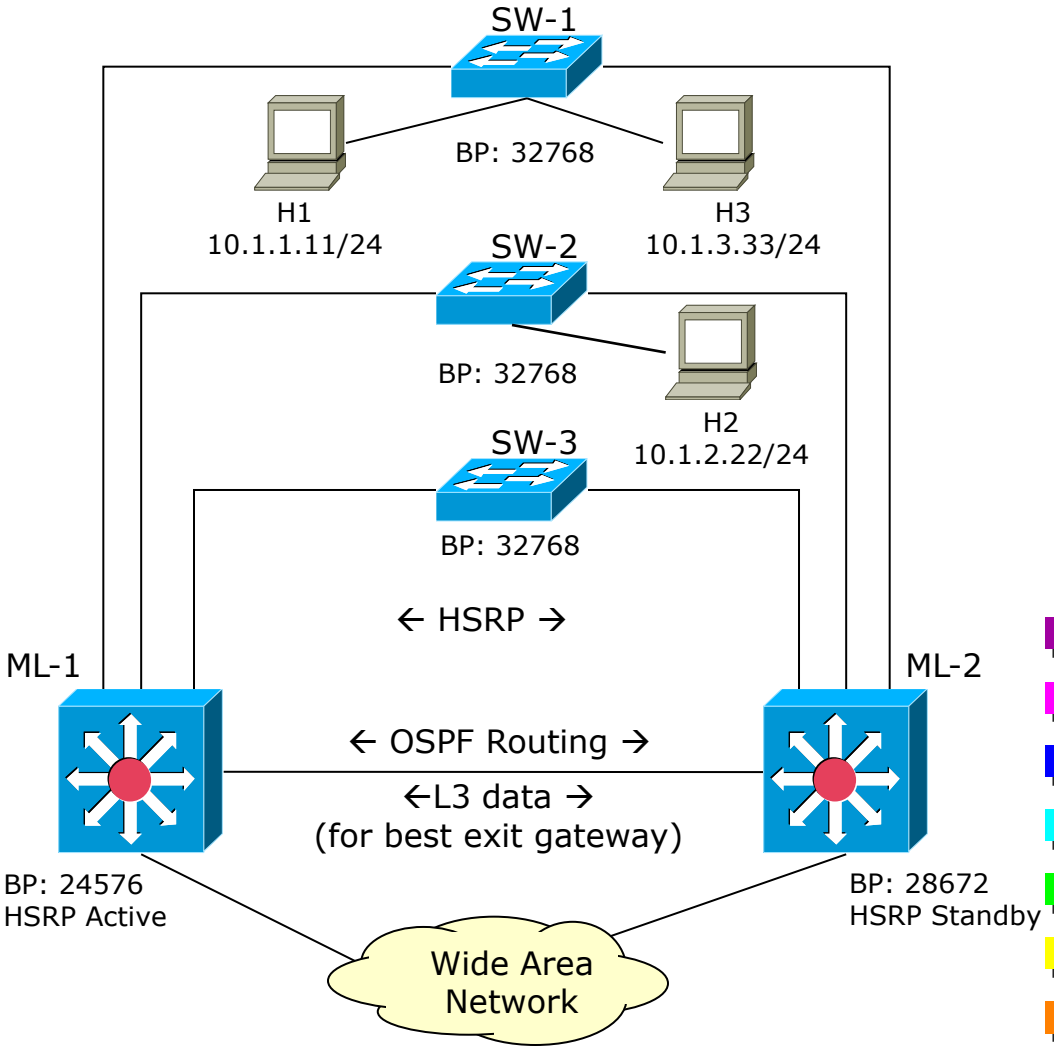
- A specific L3 link has been added between routers
  - Allows routing protocols to exchange routing messages
  - Allows exchanging packets at L3 between the two routers in order to select the best exit gateway

- In principle, routers can use one of the other VLANs also for the routing traffic
    - E.g. a set of routes from ML-1 to one of the real IP addresses of ML-2
  - Discouraged; local hosts will receive messages from the routing protocols and can potentially intercept traffic between routers

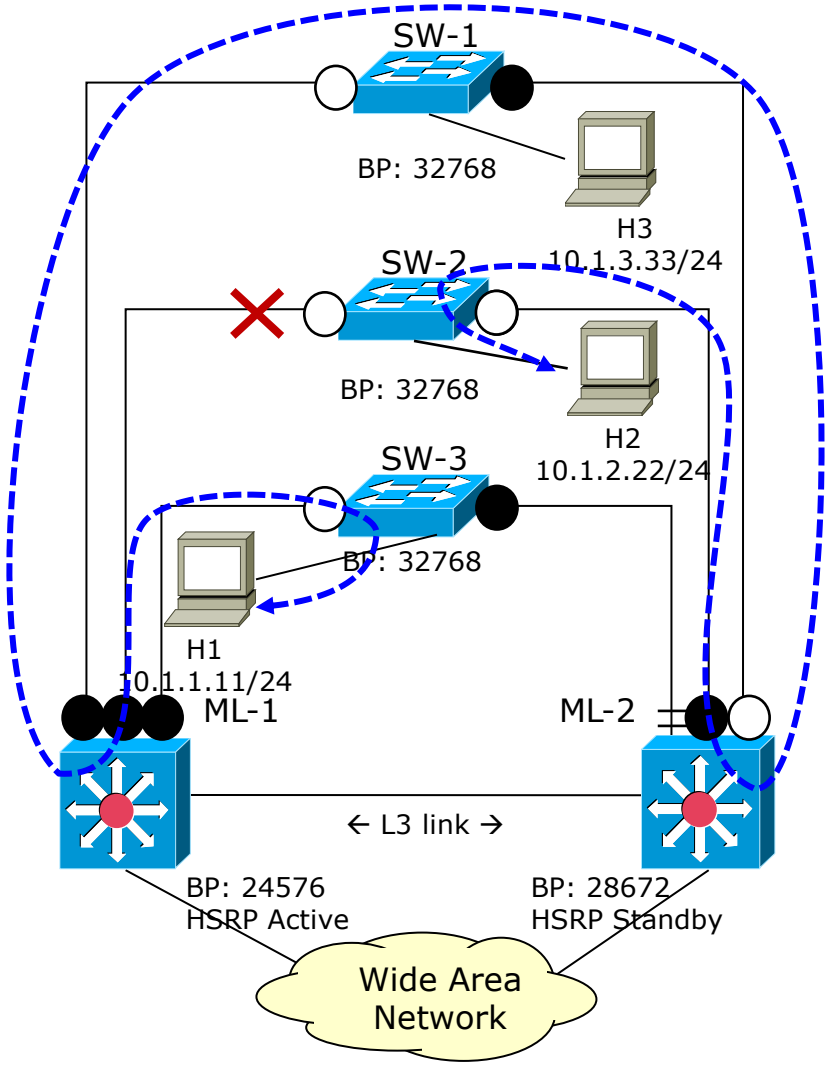
Second config: a better L2-L3 network (1)

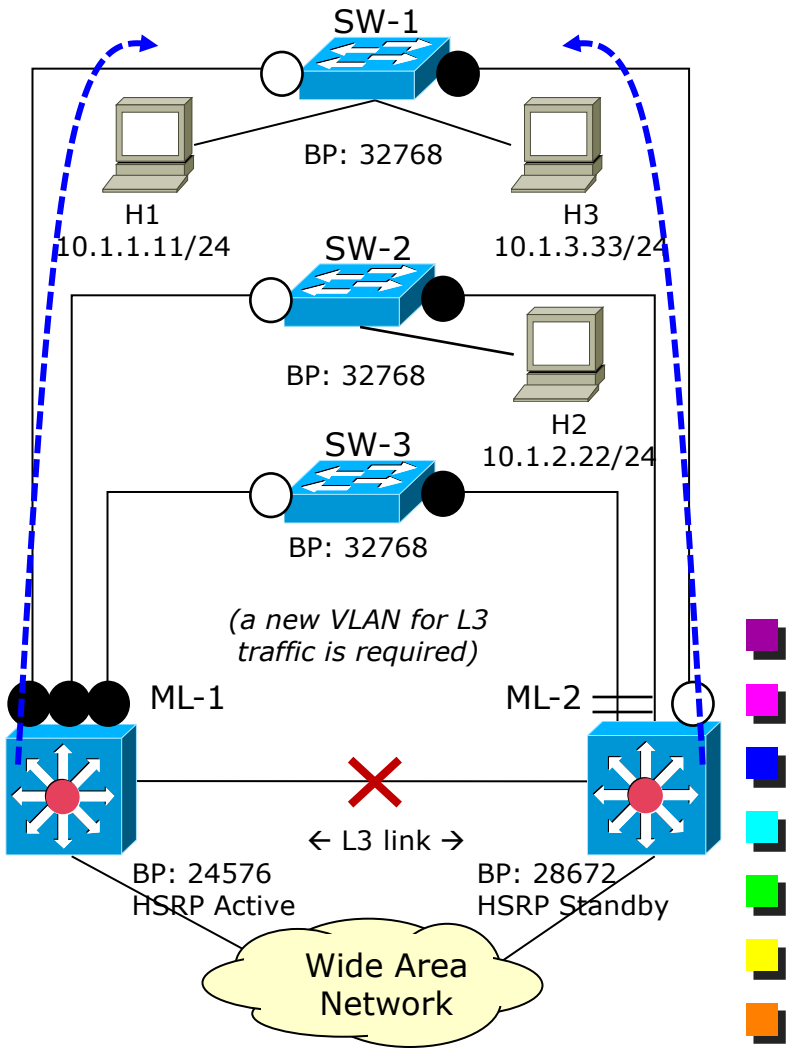
Previous configuration, plus one additional link for transporting L3 traffic (routing and data toward the best exit gateway)

New link configured at L3



Second config: a better L2-L3 network (2)





Second config: a better L2-L3 network (3)

- Still troubles in case of some failures
- Fault 1: a link from ML-1 (root bridge) to SW-x fails
  - Un-optimized path
  - L3 links do not participate to the spanning tree, and cannot be used to “re-protect” the L2 network
- Fault 2: the link between ML-1 and ML-2 fails
  - VLANs must be configured in order to allow also “pure” L3 traffic to be exchanged
  - Routing protocols on that interface must be configured as well
    - Are we sure that L3 protocols will select the direct link?
    - Routers may have two equivalent paths (e.g., both operating at 1Gbps) between each other, hence only one will be used

Second config: lesson learned

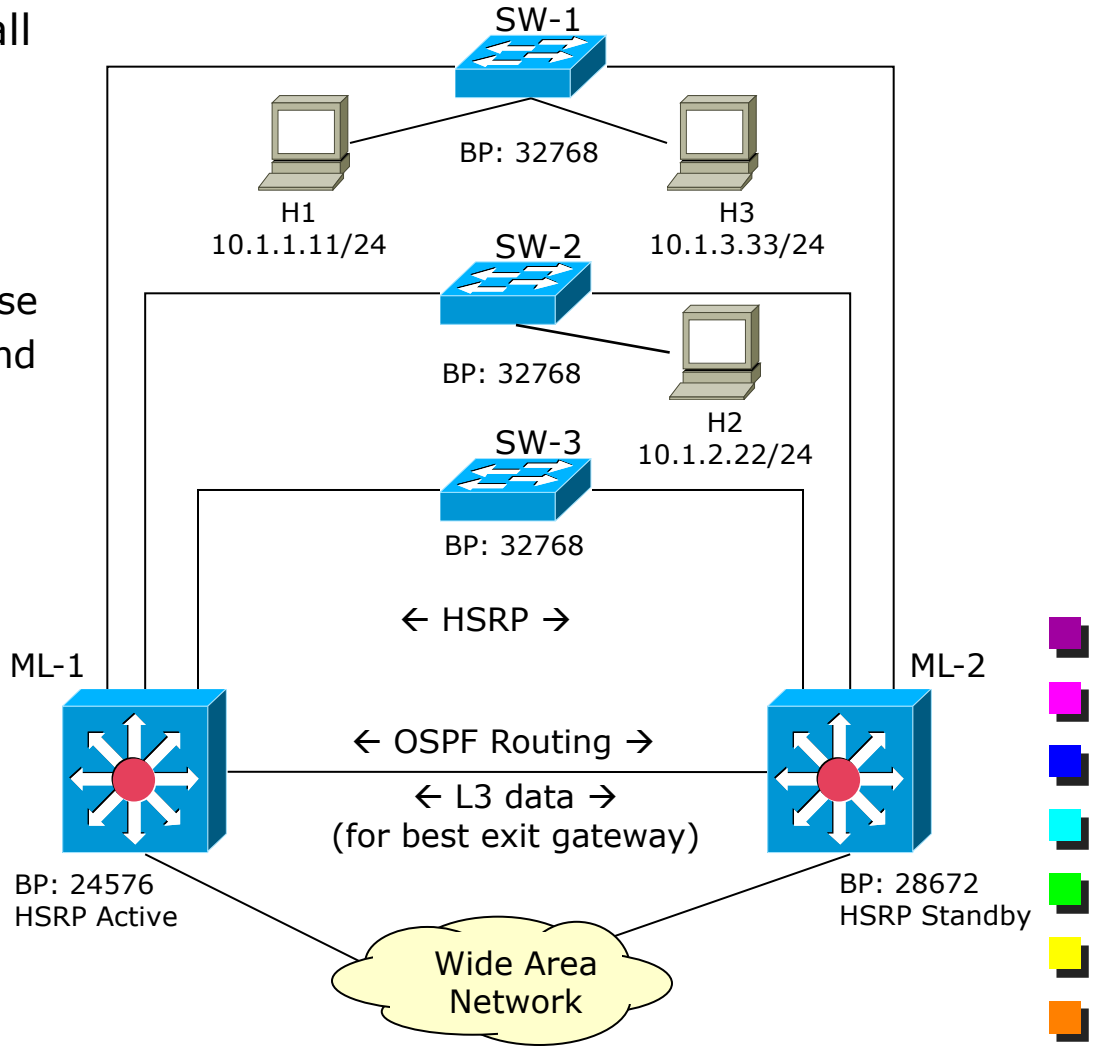
L2 and L3 operate independently.

A fault in L2 network cannot be automatically re-protected by the L3 network.

A fault in L3 network cannot be automatically re-protected by the L2 network.

Third config: an even better network (1)

- Better choice: configure all the link of the LAN at L2
- All links can be used to reprotect each other
  - Not optimized path in case the link between ML-1 and ML-2 fails

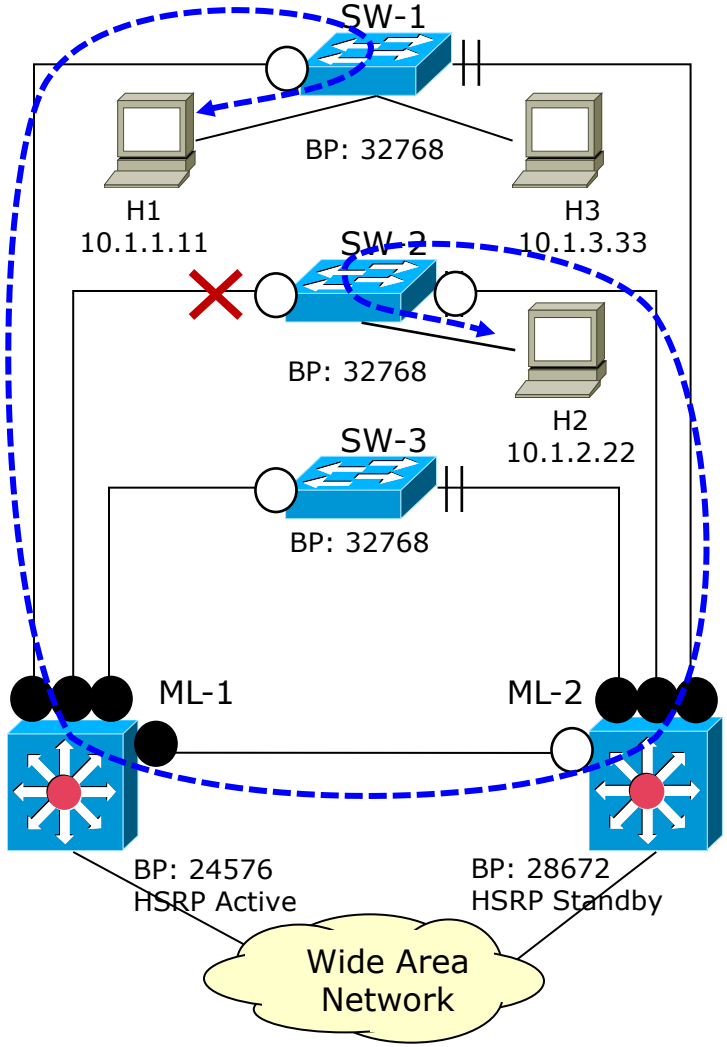


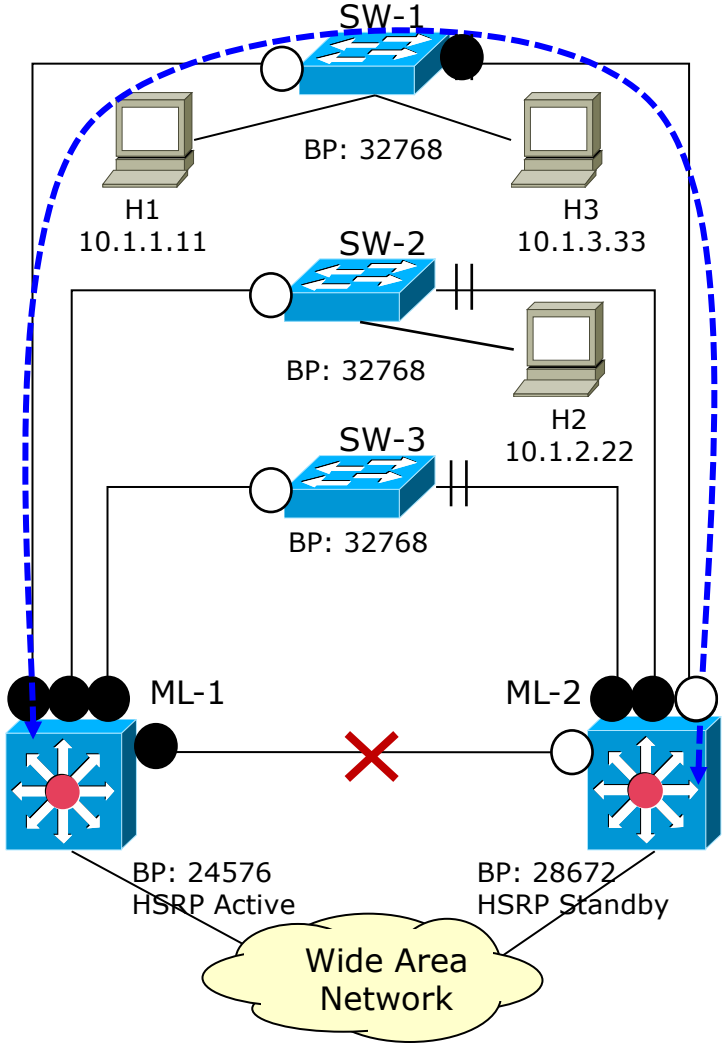
Single Spanning Tree

Four VLANs

- 3 for data (with HSRP)
- 1 for L3 data (OSPF + data toward the best exit gateway)

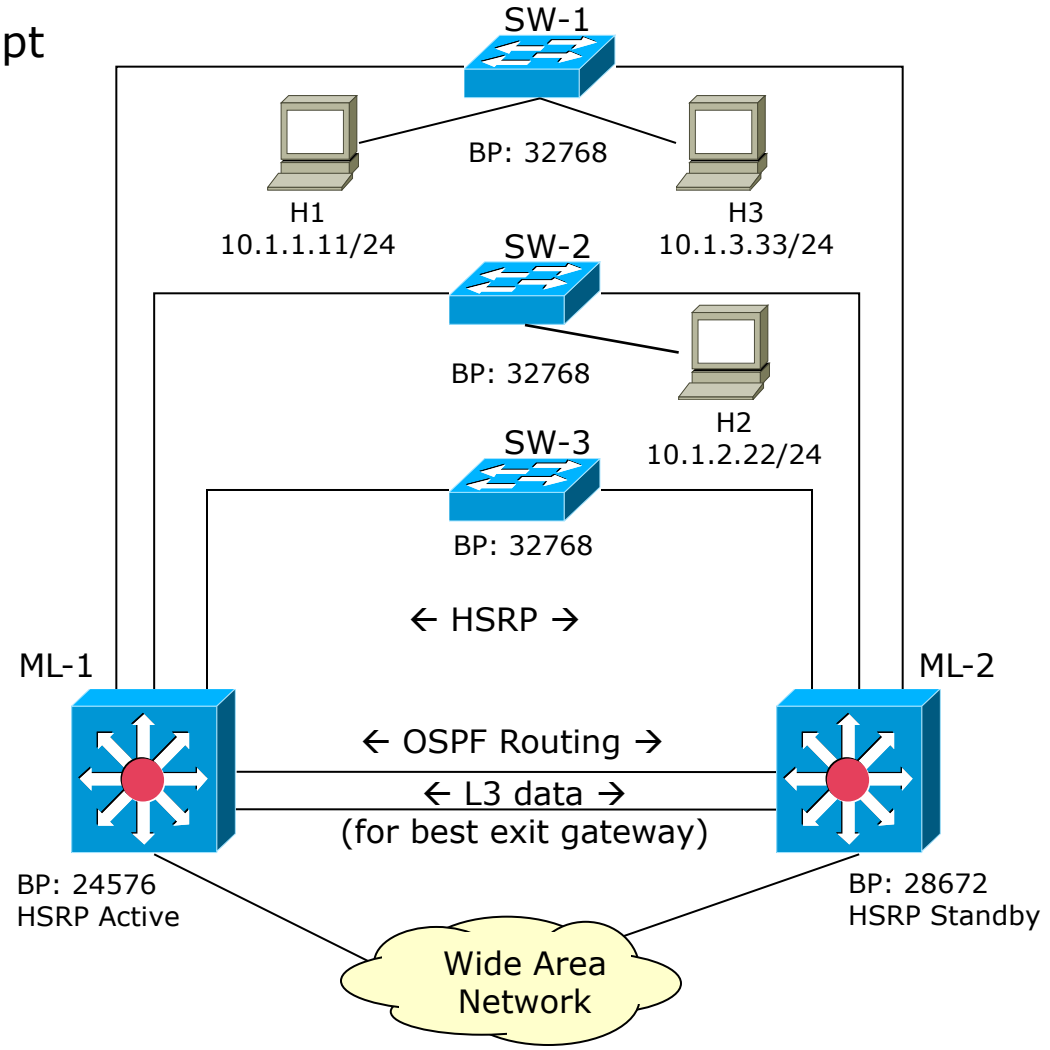
Third config: an even better network (2)





Fourth config: the "definitive" network (1)

- All links configured (except the WAN links) as L2
- Two redundant links between multilayers

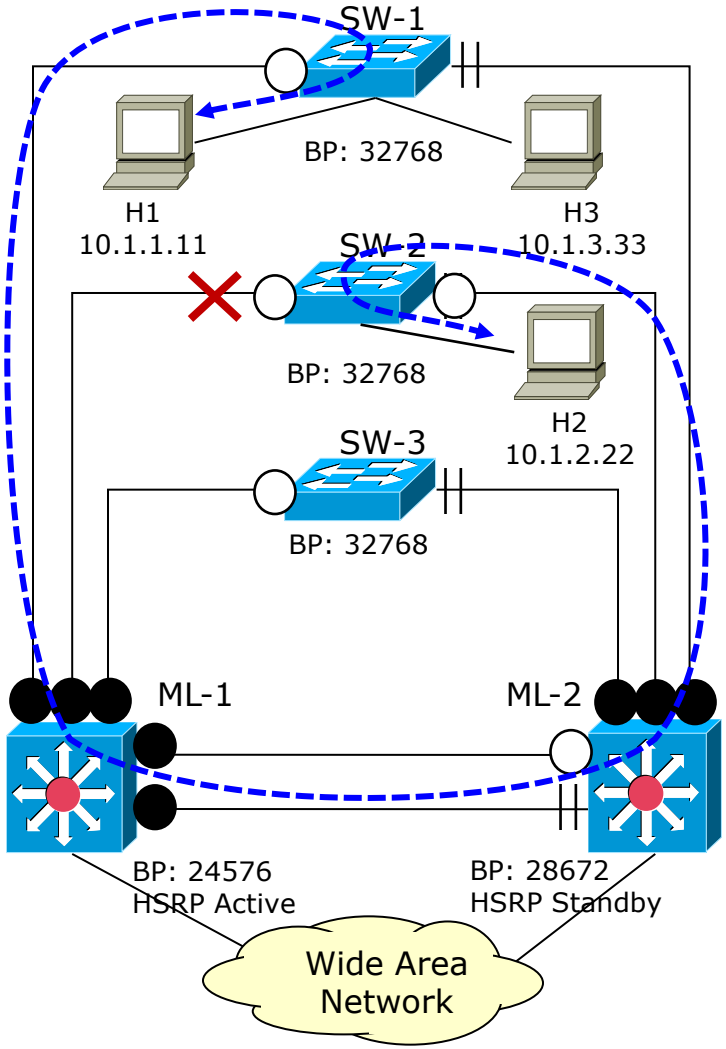


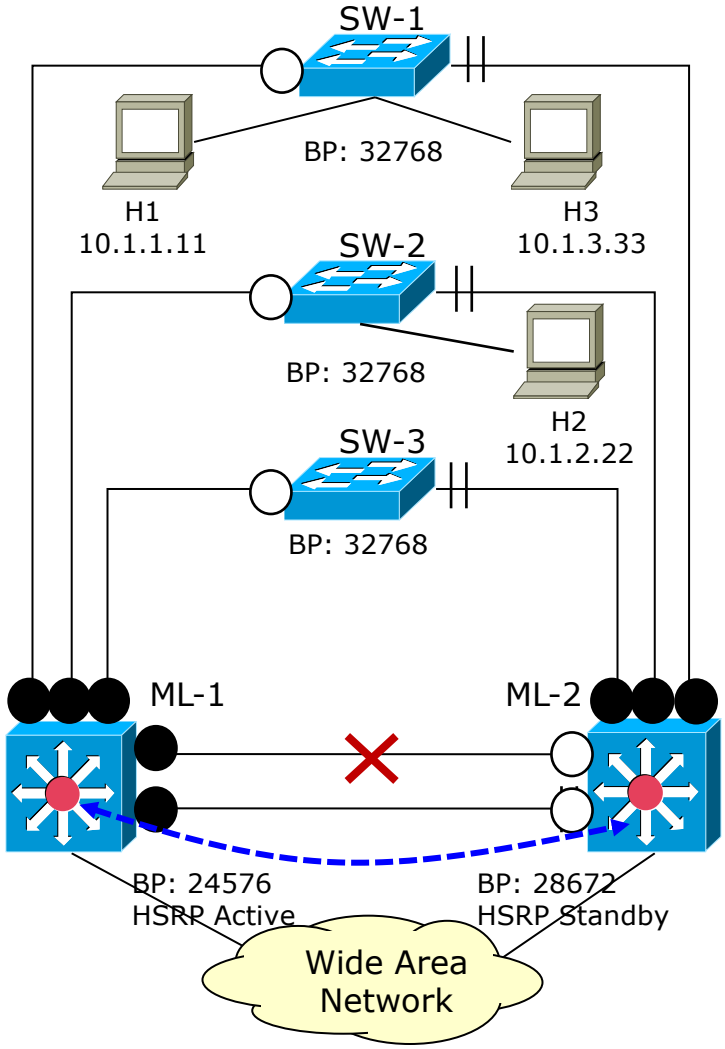
Single Spanning tree

Four VLANs

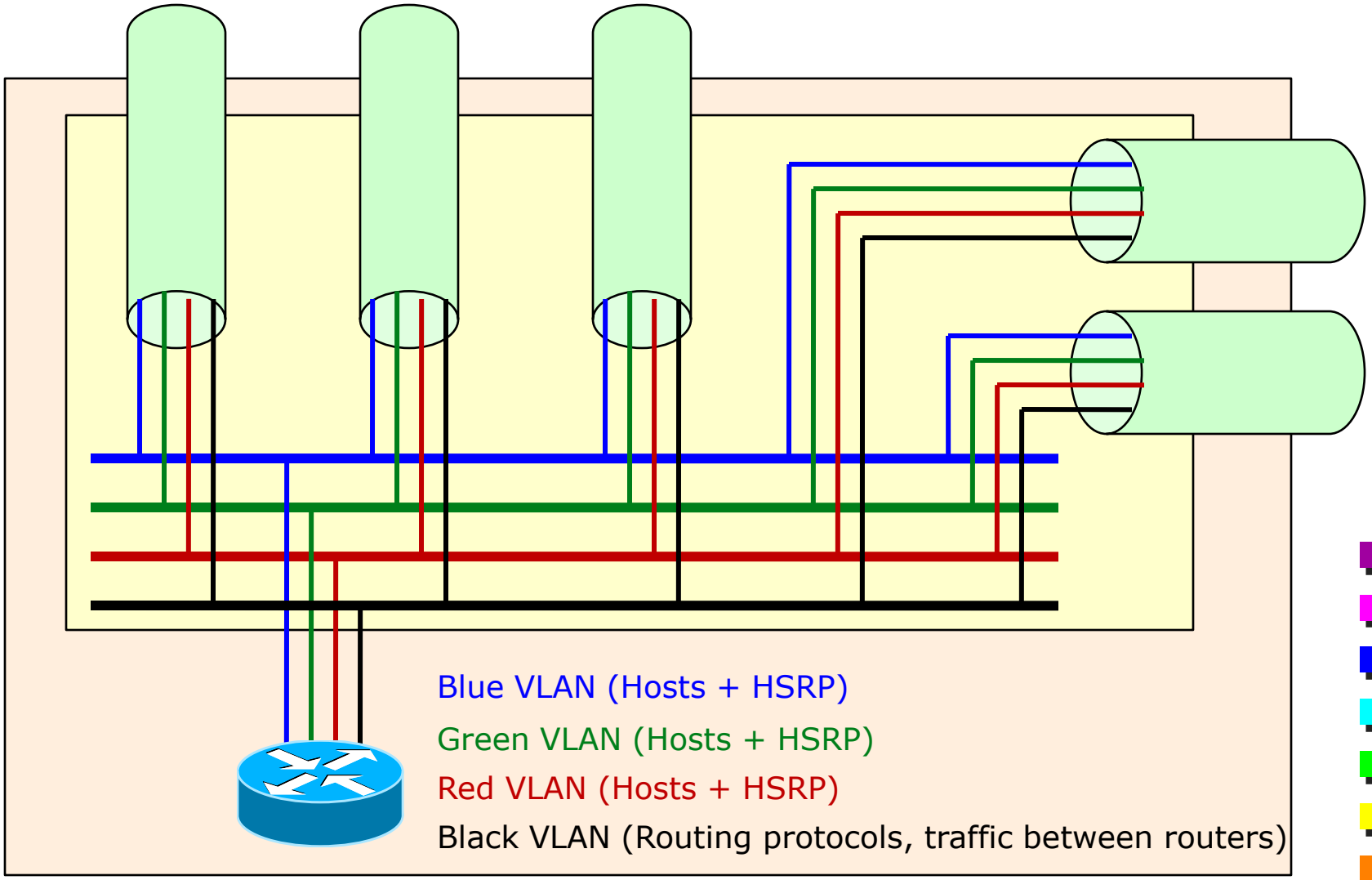
- 3 for data (with HSRP)
- 1 for L3 data (OSPF + data toward the best exit gateway)

Fourth config: the "definitive" network (2)



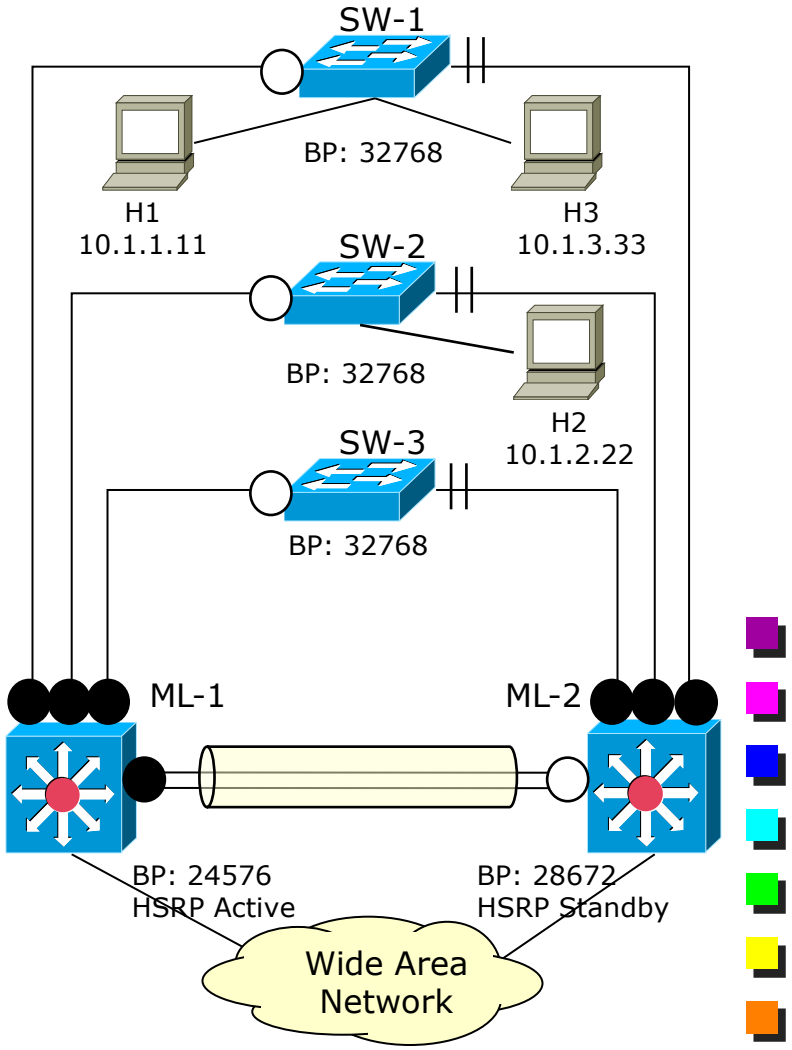


Fourth config: logical view of the L2-3 switch

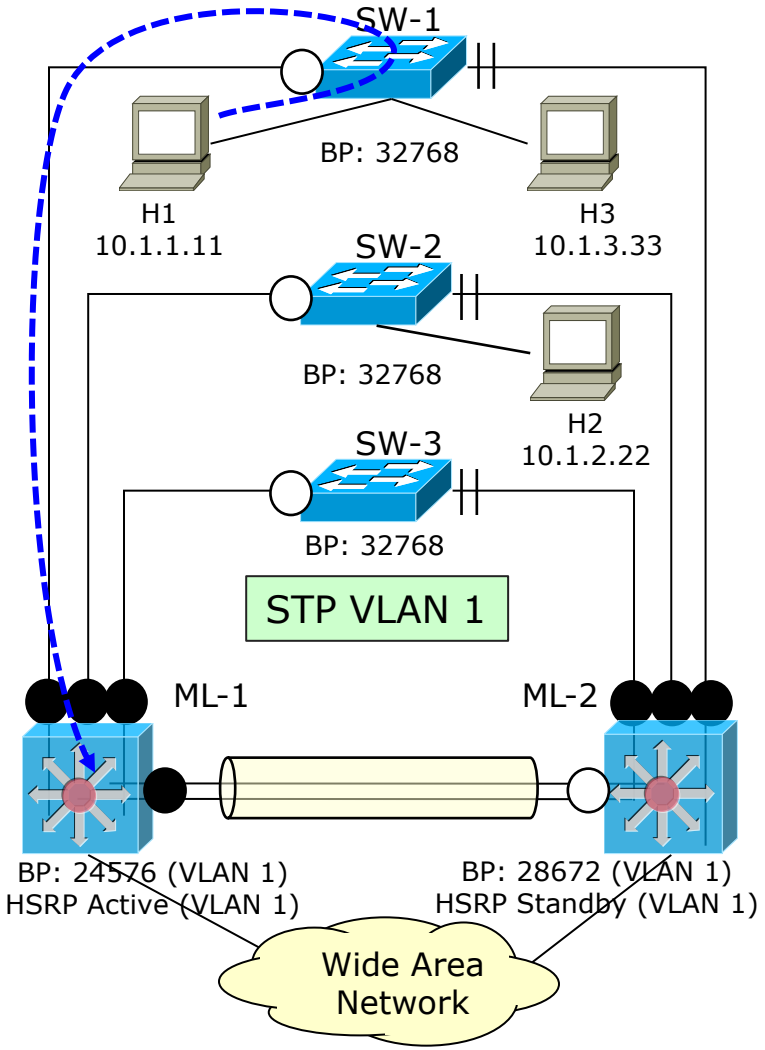


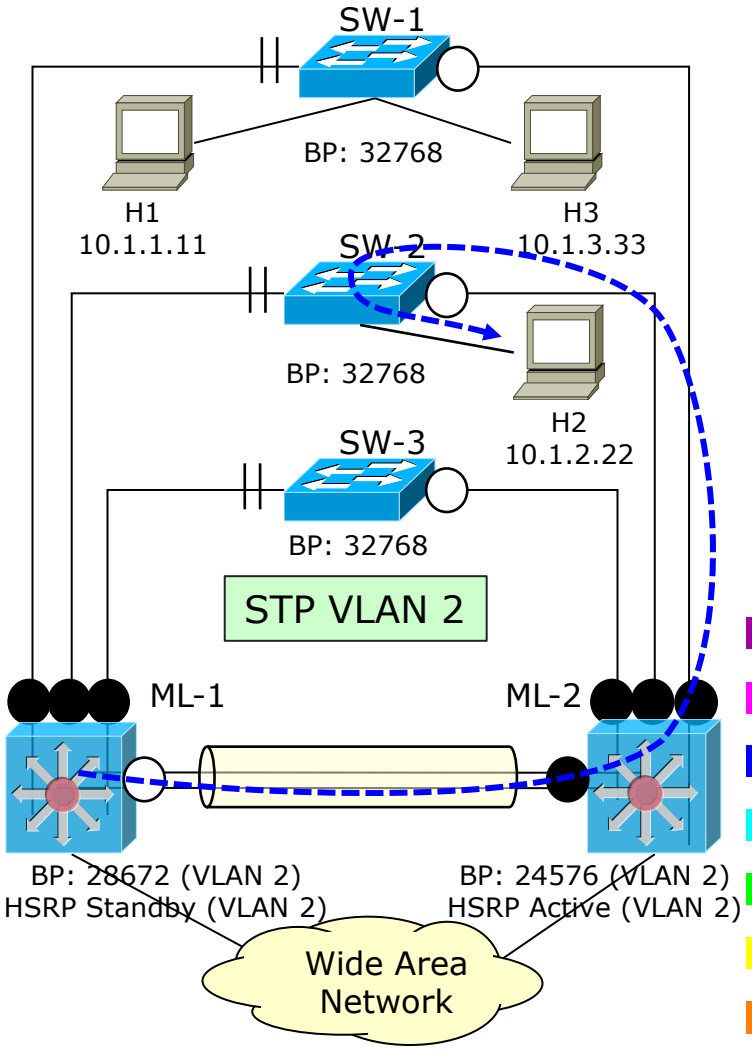
Fourth config: a final improvement

- Links between ML-1 and ML-2 in Link Aggregation
  - More bandwidth
  - No STP transient in case one link fails

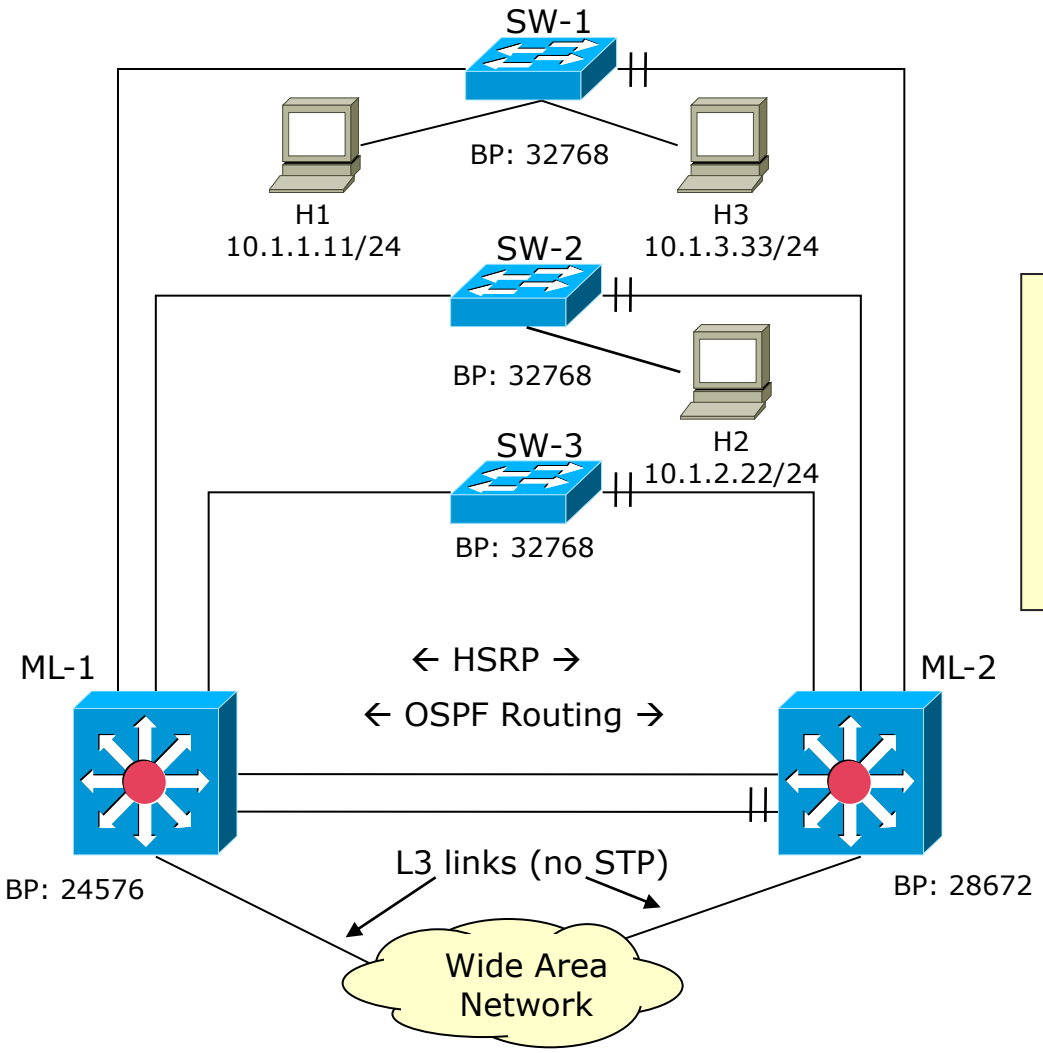


Fifth config: Per-VLAN STP





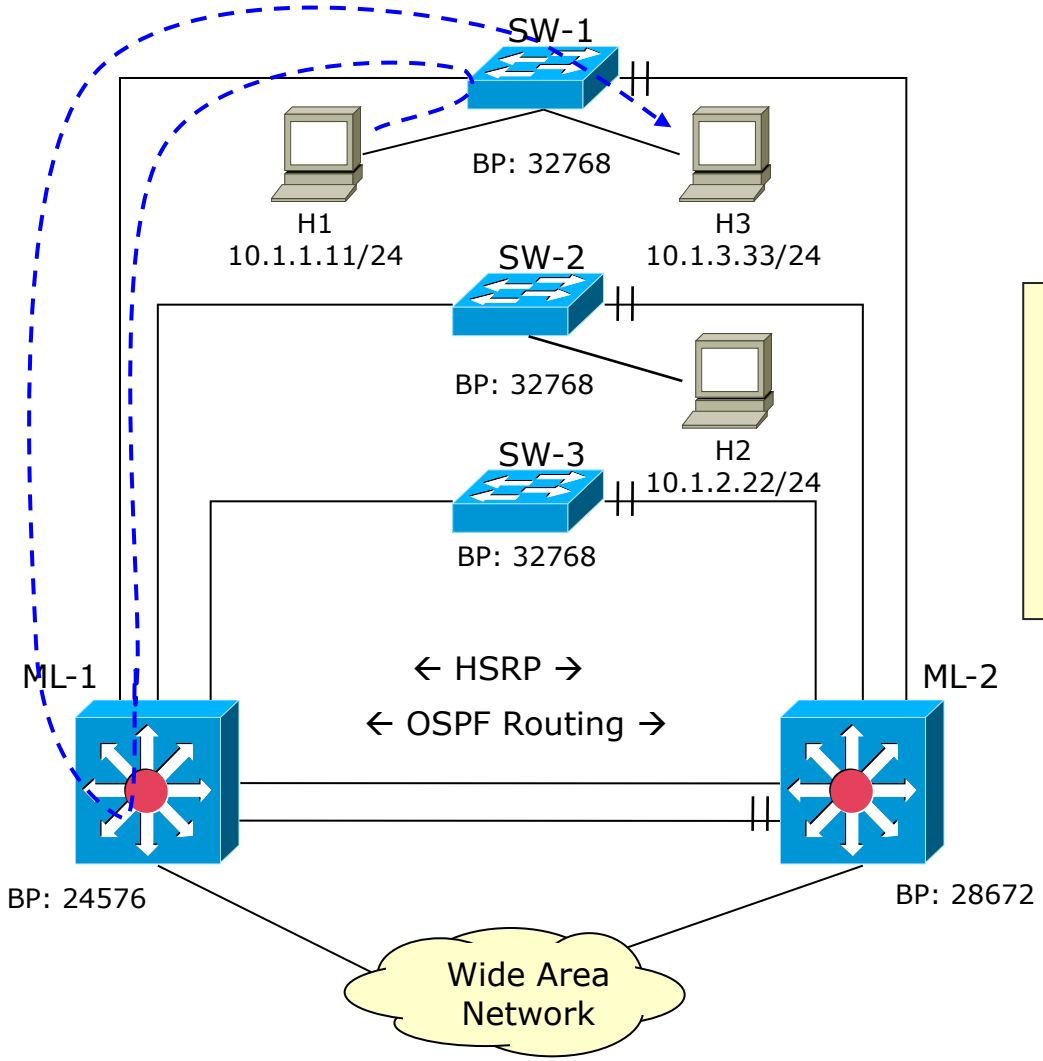
Example 1: forwarding at L2



Lesson learned

In case of forwarding at L2, only the physical topology (i.e., the outcome of the STP) matters

Example 2: forwarding at L3

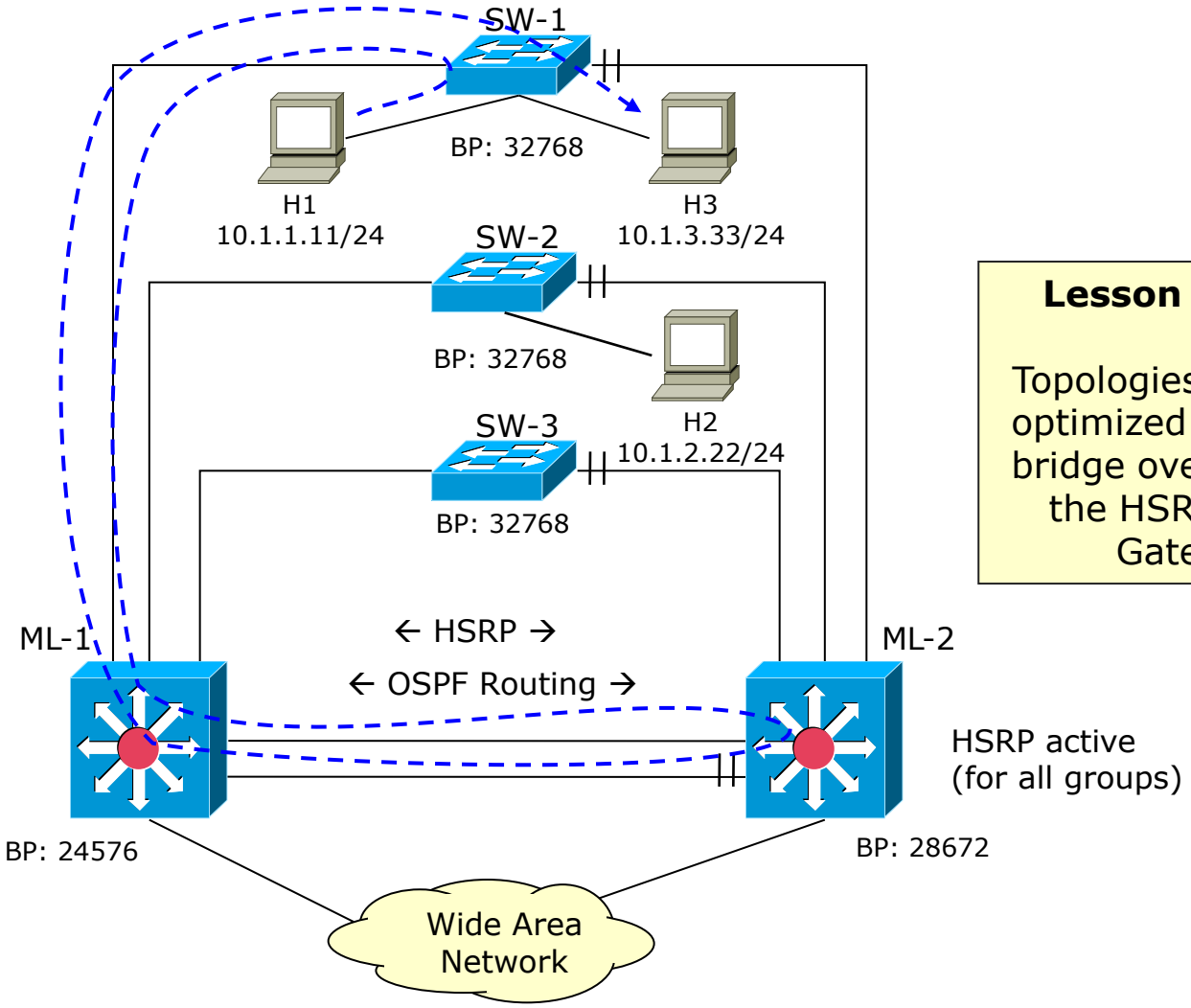


HSRP active
(for all groups)

Lesson learned

In case of forwarding at L3, we care about the STP physical topology and the HSRP Active Gateway

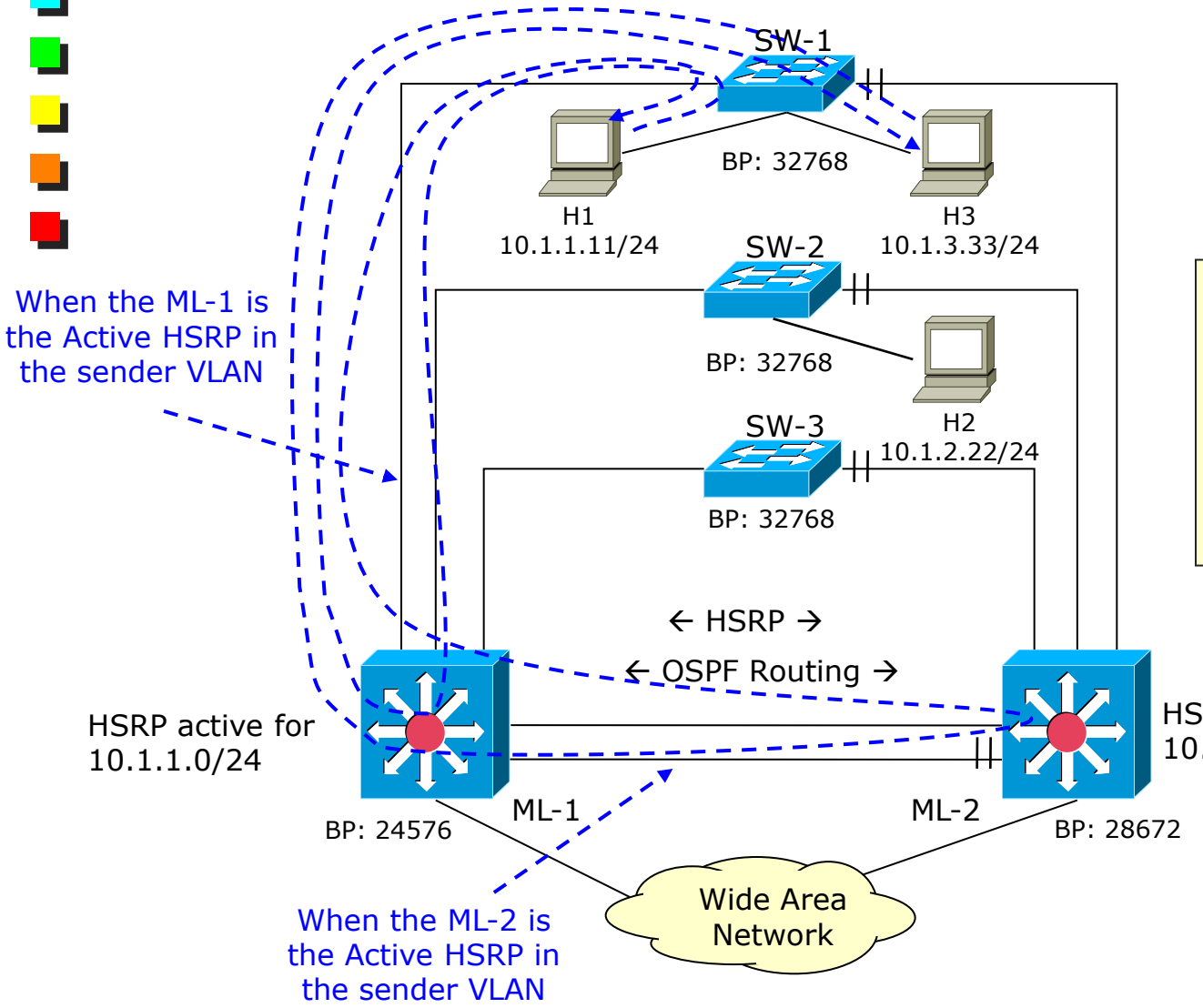
Example 3: forwarding at L3 (root != active)



Lesson learned

Topologies are more optimized if the root bridge overlaps with the HSRP Active Gateway

Example 4: forwarding at L3 (multiple Active Gateways)

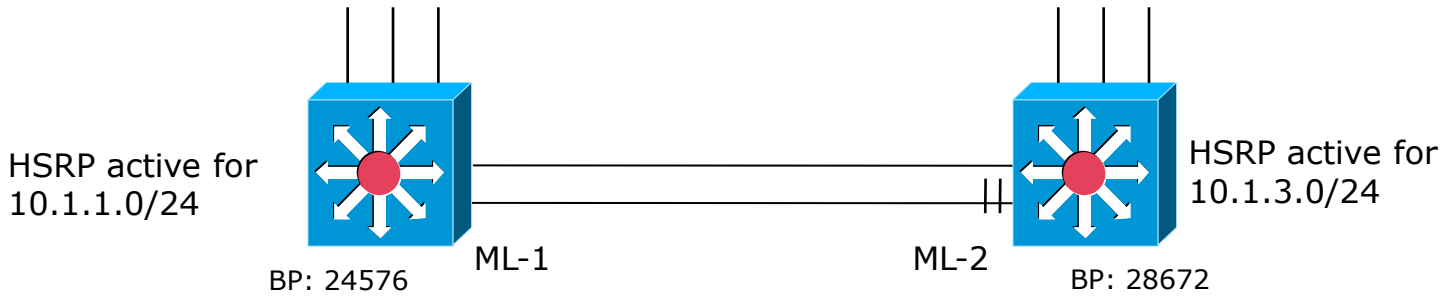


Lesson learned

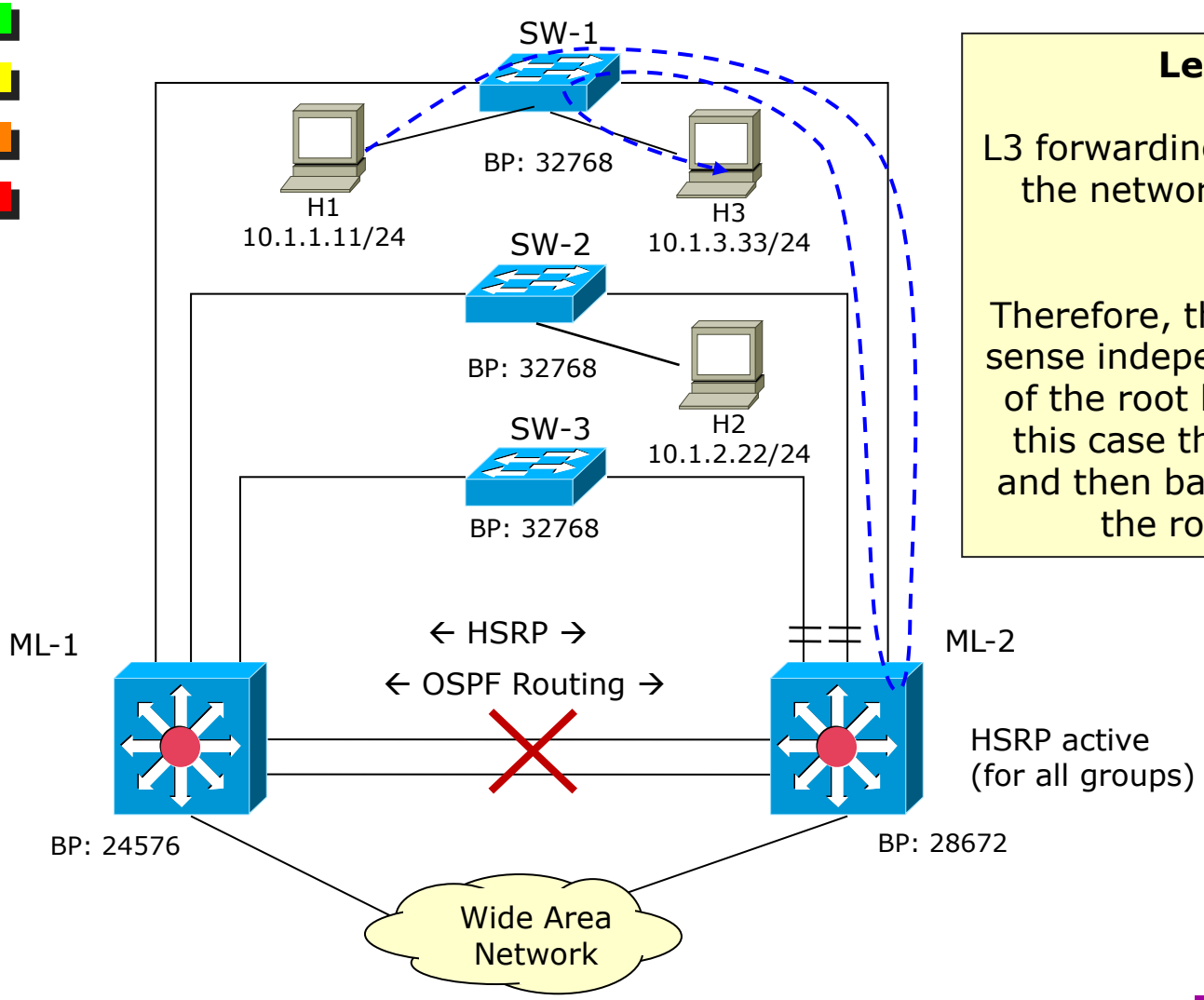
In case of multiple HSRP Active Gateway for different groups,we may have asymmetric paths

Example 4: HSRP Active and Routing Packets

- If a router is HSRP active:
  - Answers to ARP requests for that IP network
  - Sends traffic to that IP network using the Virtual MAC address
- HSRP does have any influence in case a router receives traffic that has to be routed into another IP network
  - The router is still a router
- In our example, ML-1 will forward IP traffic to network 10.1.3.0/24 even if act as stand-by router for those IP addresses



Example 5: fault between ML-1 and ML-2

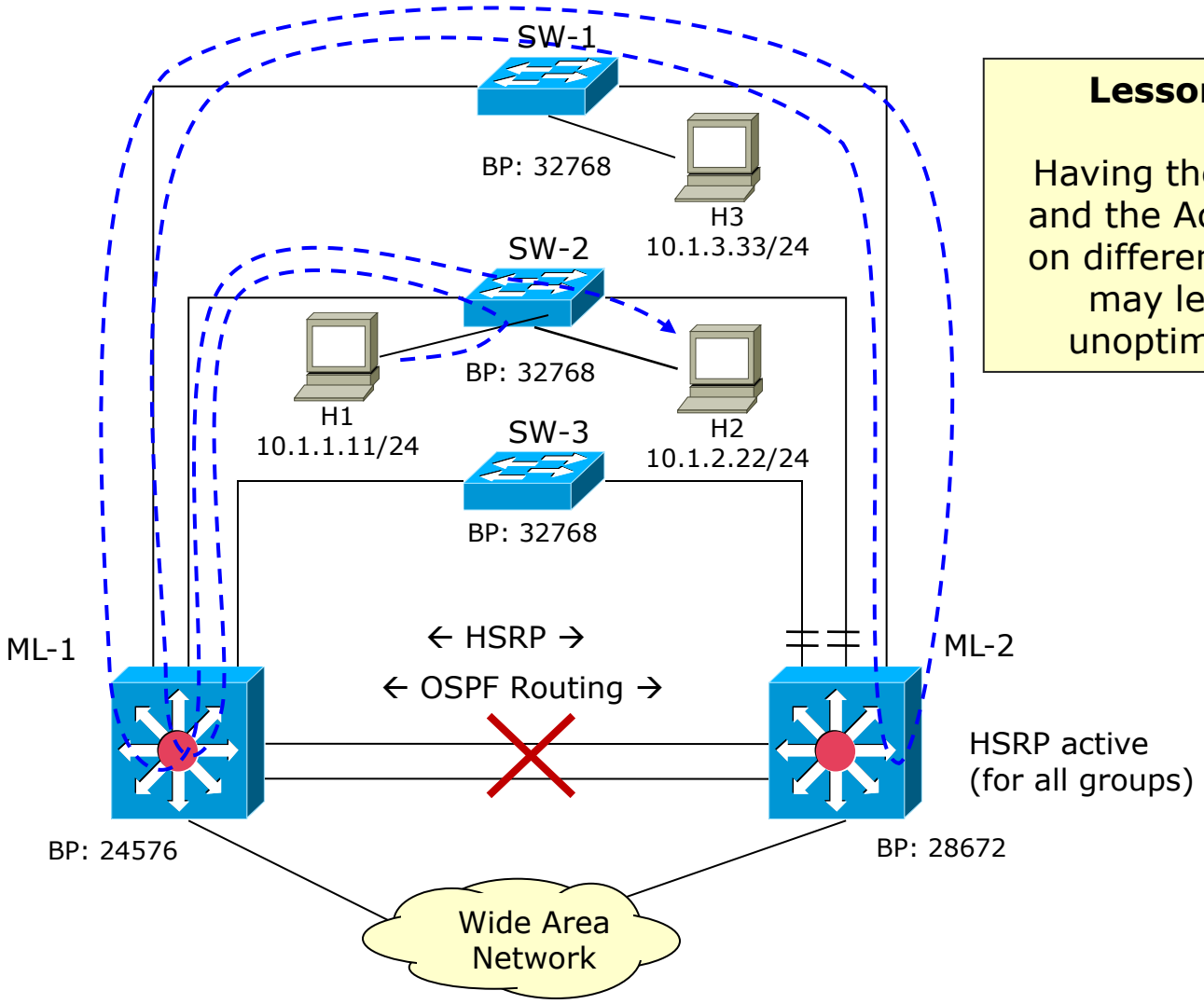


Lesson learned

L3 forwarding takes the best path on the network that results from L2 STP.

Therefore, the path at L3 is in some sense independent from the location of the root bridge. For example in this case the packet goes to ML-2 and then back to the user although the root bridge is ML-1.

Example 6: same fault, different hosts



Lesson learned

Having the Root Bridge and the Active Gateway on different machines it may lead to very unoptimized paths.

How to trace frames on L2-L3 networks

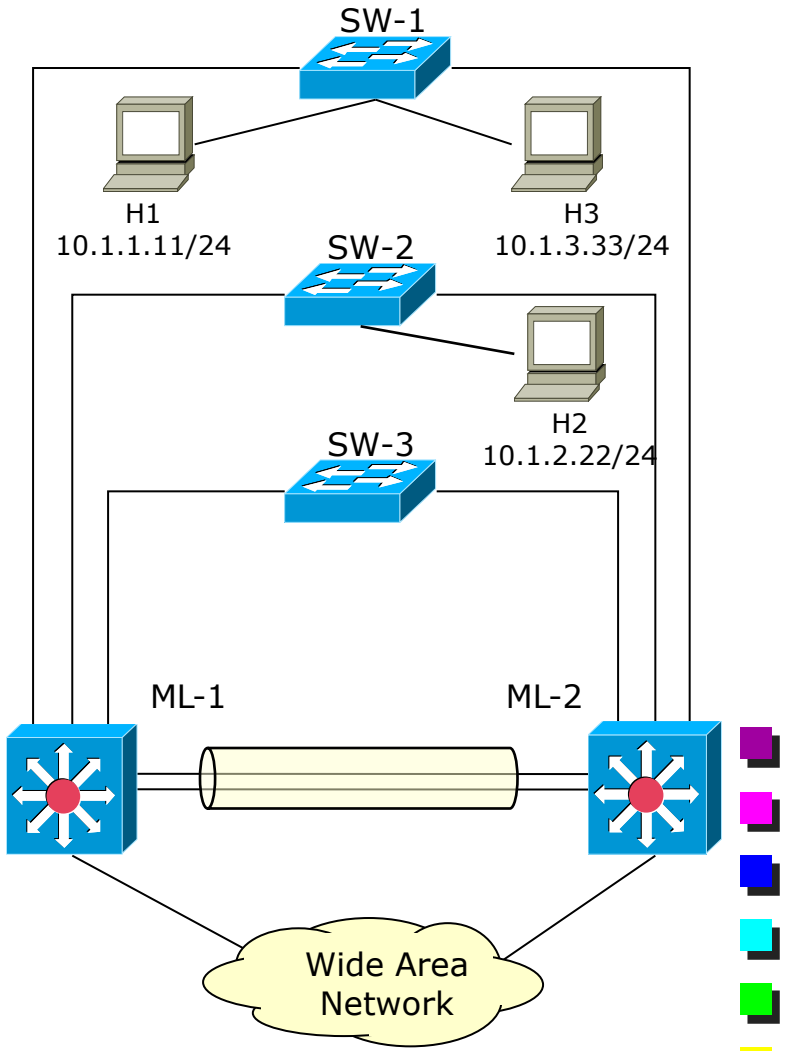
- Explode the multilayer switch in order to detail which links are L2 and which are L3
- Run the STP on the L2 network
  - Remember that each port has its own status
    - In this slide show only the most significant port have been associated with the status… but be careful!
- Determine the active HSRP router within each VLAN
- Determine the path of the packet at L3

L4-7 Multilayer switches (1)

- Available devices that are able to work L4-7
  - Often only on some functionalities
- Different intelligence on different linecards
  - Some may be L2 only, but other cards can provide more L4-L7 intelligence (if needed)
  - E.g., firewall, application-layer balancer, etc.
- No separate boxes required
  - Simple management (one box does all)
  - Lower costs
  - More Flexibility
- Oriented primarily to the enterprise market

L4-7 Multilayer switches (2)

- Multilayers ML-1 and ML-2 are good candidates for having L4-L7 functionalities
  - Most of the traffic (either internal or toward the Internet) is concentrated here
  - Good place to enforce security policies, balance traffic across servers, etc.



Conclusions

- L2-L3 switches are definitely mainstream
  - Much more flexibility in defining where to handle traffic at L2 or L3
    - It's a matter of a software configuration
- Most campus networks are almost entirely L2 (with VLANs)
- L3 at the edge (e.g. servers, data centers) and for connection toward the WAN
- Multilayer (L2-L3) in the core in order to speed-up VLAN interconnection
- Important to engineer the L3 network taking into account the actual L2 topology (STP)