

Upgrade of the SPring-8 Control Network for Integration of XFEL

T. Sugimoto, M. Ishii, T. Ohata, T. Sakamoto, and R. Tanaka
Japan Synchrotron Radiation Research Institute (JASRI/SPring-8)

Overview

- Motivation
- System requirements and considerations
- Solutions to requirements
 - Segregation of physical network
 - Interconnection of physical network
 - Logical addressing to physical network
 - Segmentation of network
 - Redundancy for non-stop operation
- Refurbishment
- Summary

Motivation

- In 2011, XFEL facility will be established in SPring-8 site.
- For XFEL control system, we use same control architecture of SPring-8.
- New control system must be ready in 2010, when early commissioning (RF aging, etc.) is started.
- Previous control network of SPring-8, was no longer suitable to integrate new control system.
- We considered system requirements, planned to upgrade, and performed refurbishment of the control network.

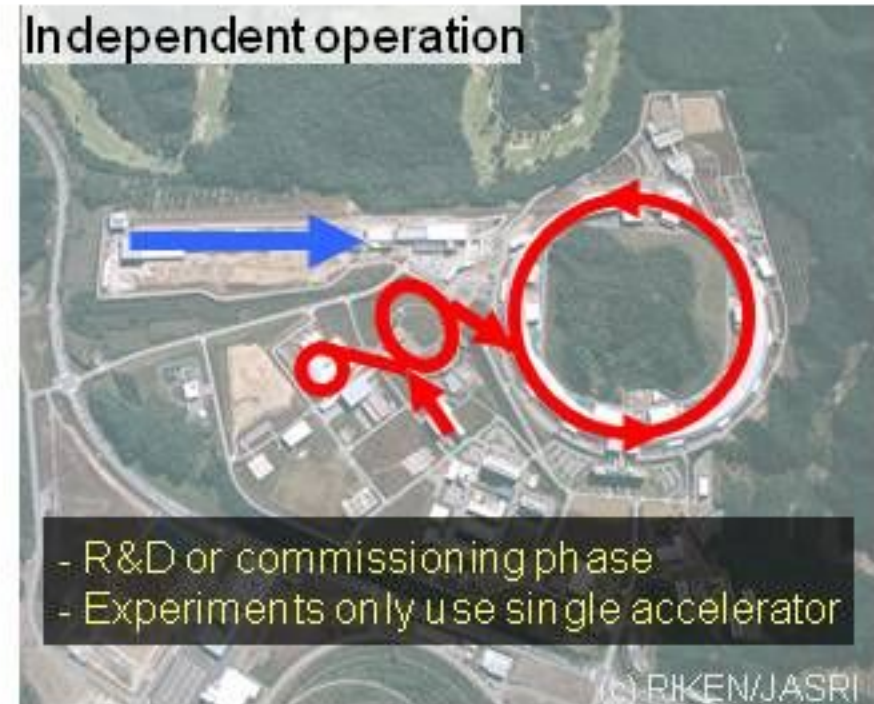


System requirements and considerations

- Flexibility
- Scalability
- Stability
- Availability
- Management

System requirements and considerations

- **Flexibility**
 - Two accelerator complexes must be operated **both disconnected and connected mode**.
 - independent operation
 - combination operation
- Scalability
- Stability
- Availability
- Management



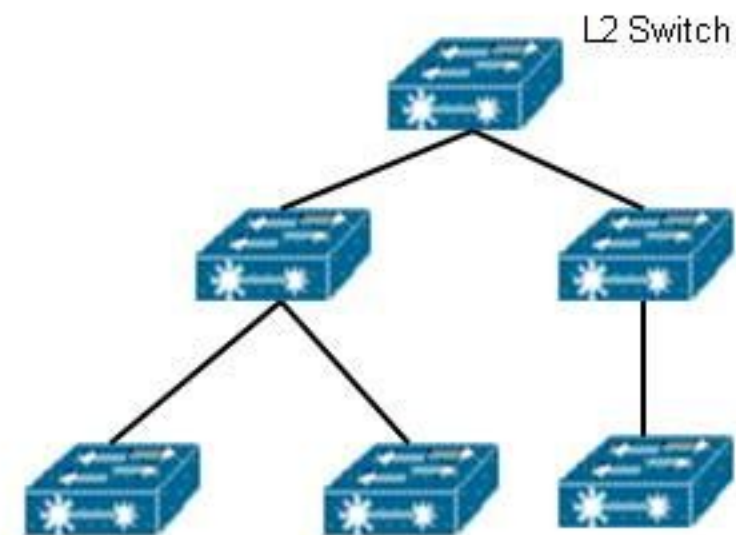
System requirements and considerations

- **Flexibility**
 - Two accelerator complexes must be operated **both disconnected and connected mode**.
 - independent operation
 - combination operation
- Scalability
- Stability
- Availability
- Management



System requirements and considerations

- Flexibility
- Scalability
 - Since reliability of routers was low in 1990s, we have adopted L2 topology instead of L3 topology.
 - We now suffer from IP address exhaustion. (/21: up to 2000 nodes)

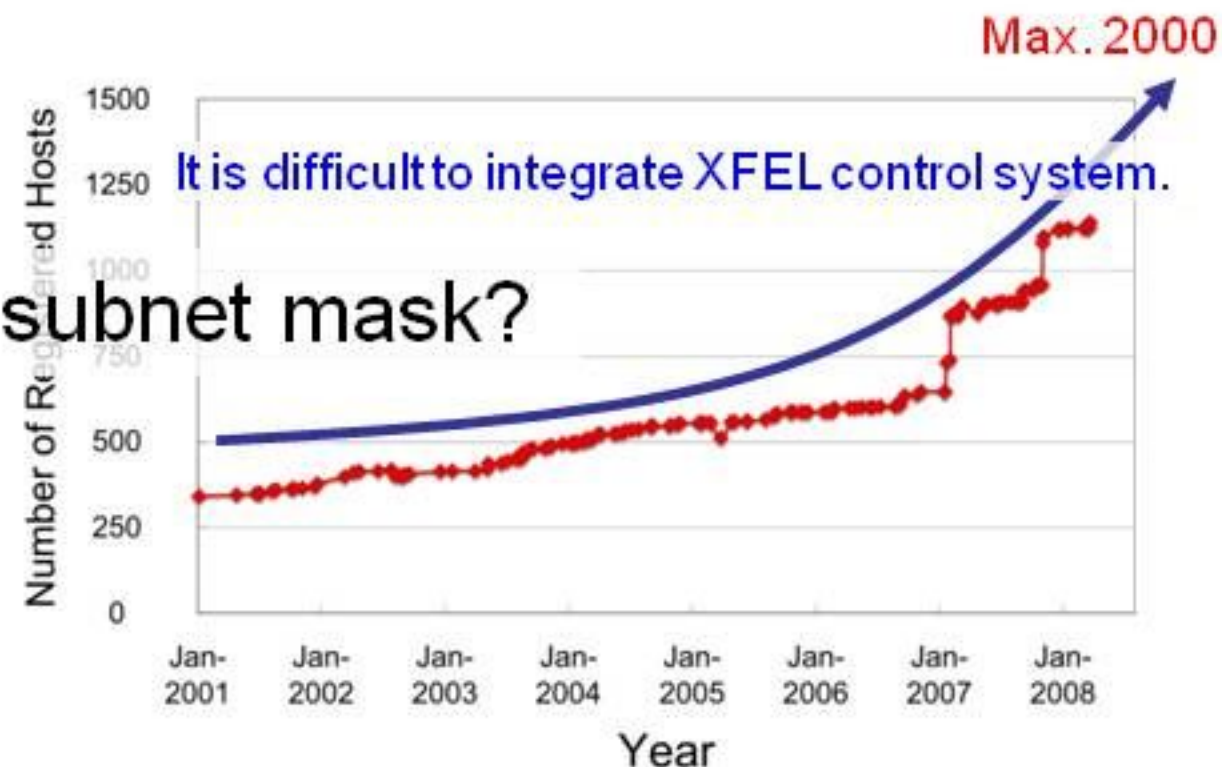


No router was required in the network.

- Stability
- Availability
- Management

Can we expand subnet mask?

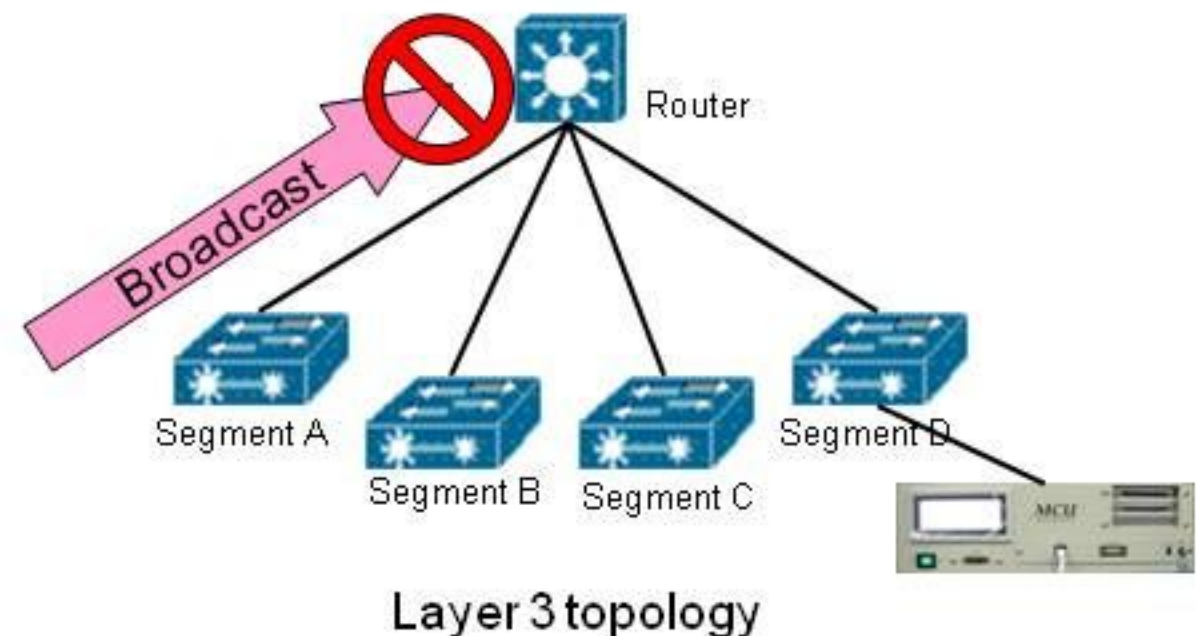
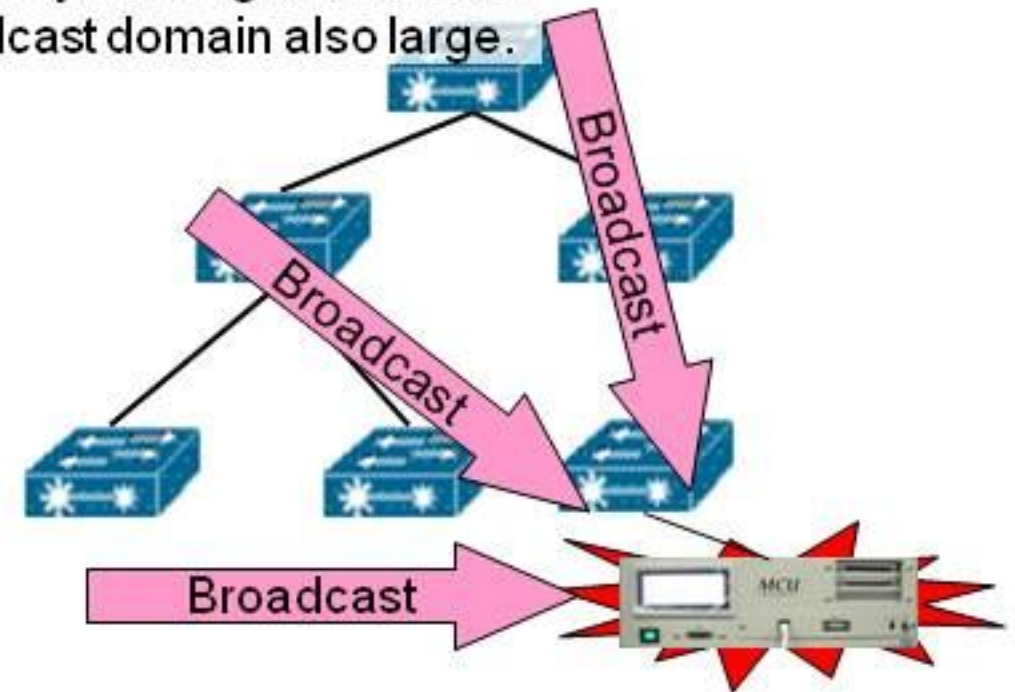
→ **No!**



System requirements and considerations

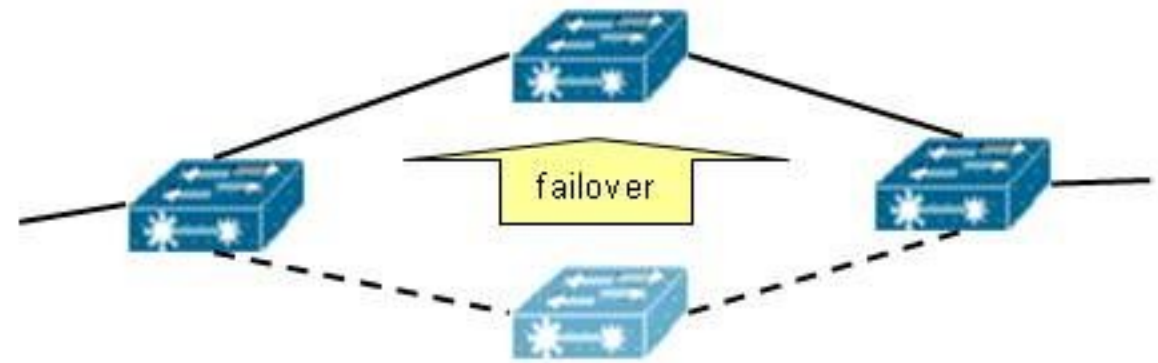
- Flexibility
- Scalability
- **Stability**
 - Some embedded devices are vulnerable against heavy broadcast traffic. [1]
 - Small-segmented L3 topology is worth thinking again.
- Availability
- Management

Large layer 2 segment makes broadcast domain also large.



System requirements and considerations

- Flexibility
- Scalability
- Stability
- **Availability**
 - Non-stop accelerator operation is required.
 - We can obtain more reliable routers than that of 1990s.
- Management



Critical time scale is 1 sec.,
restricted by COD correction intervals.

In 1990s

Single Chassis Router with slow redundant protocol



In 2000s

Multi-chassis router with fast redundant protocol



System requirements and considerations

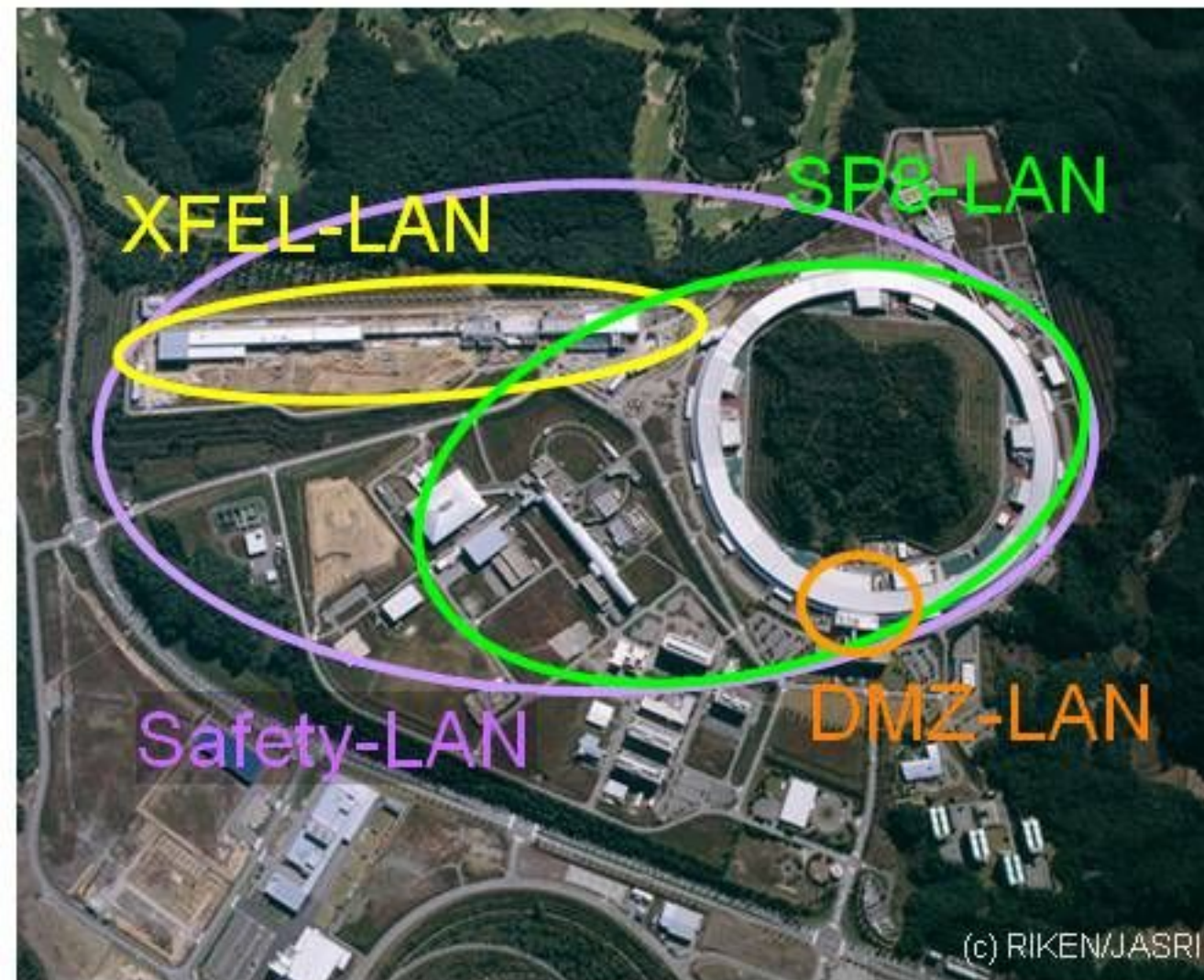
- Flexibility
- Scalability
- Stability
- Availability
- **Management**
 - Simple network architecture makes us to reduce management costs.

Solutions to the requirements

- **Flexibility**
 - We want two control system connected, other time disconnected.
 - Physically segregated and logically flexible design
 - Because one system trouble should not be affected another system.
 - And also, both connect and disconnect mode are required.
- **Scalability**
 - We assigned large address range.
 - /16 address ranges are assigned.
- **Stability**
 - We applied L3 topology with small-segmented network
- **Availability**
 - We use redundant technologies with very fast failover protocol.
- **Management**
 - We should manage all control network by one simple policy.

Segregation of physical networks

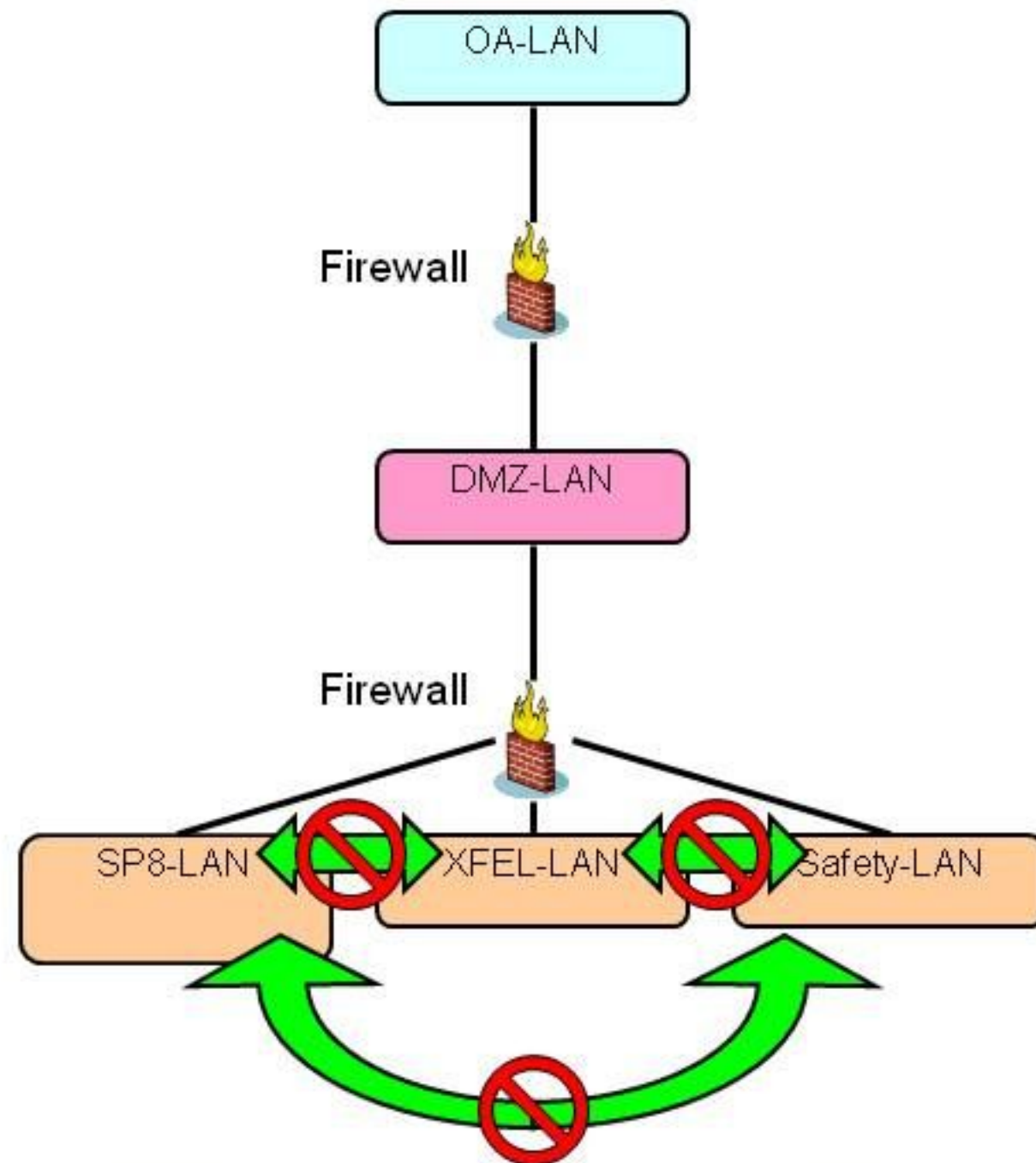
- We redefined network by control systems
 - SP8: SPring-8 control
 - XFEL: XFEL control
 - Safety: radiation monitor, gate control
 - DMZ: program development
- We made physical networks segregated by accelerators.
 - When combination operation is required, we interconnect networks.



(c) RIKEN/JASRI

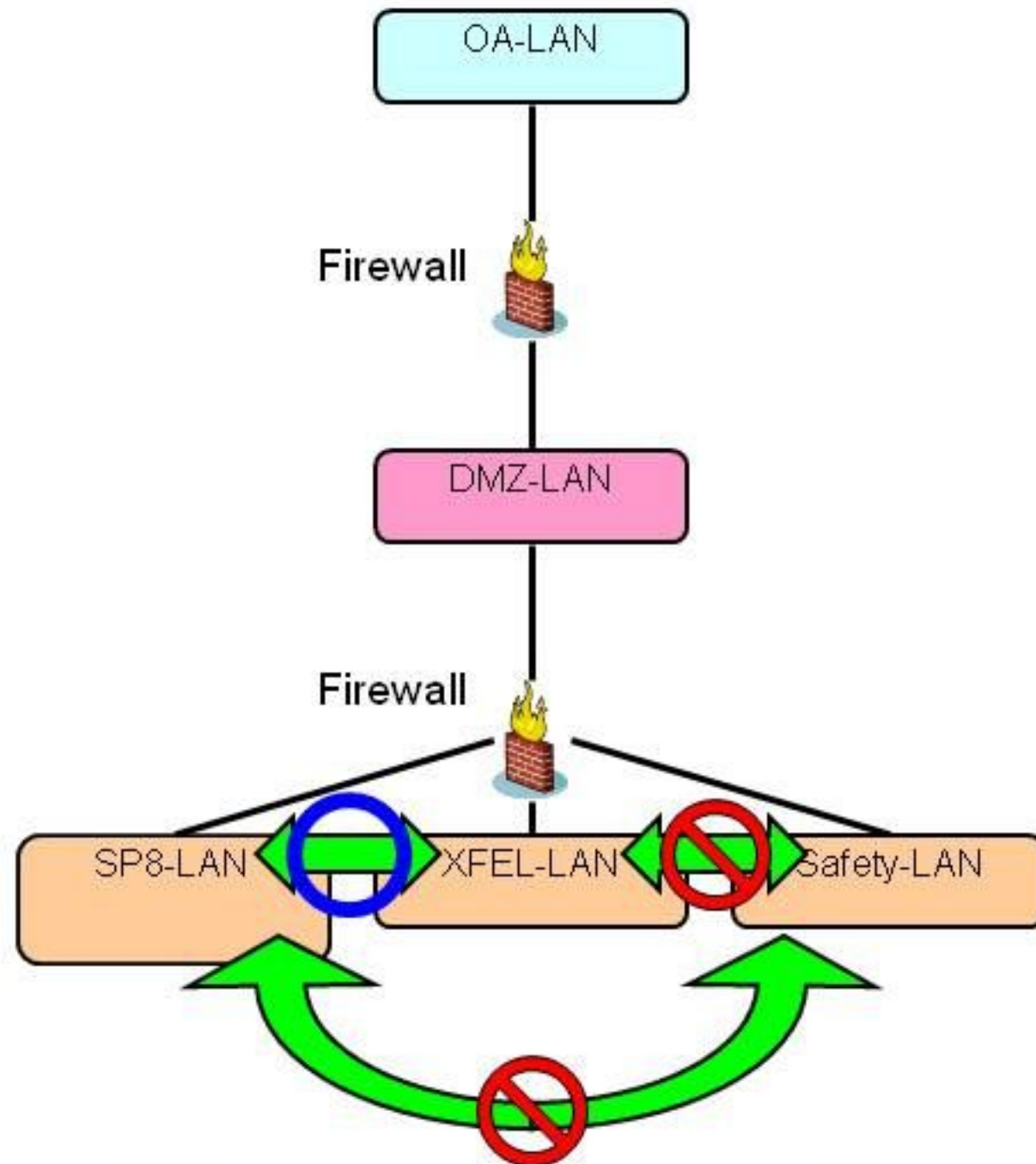
Interconnection of physical network

- When accelerators operated independently, we want to keep independence between physical networks.
 - Multi-port firewalls are used instead of routers.
 - Traffic between physical network are filtered.



Interconnection of physical network

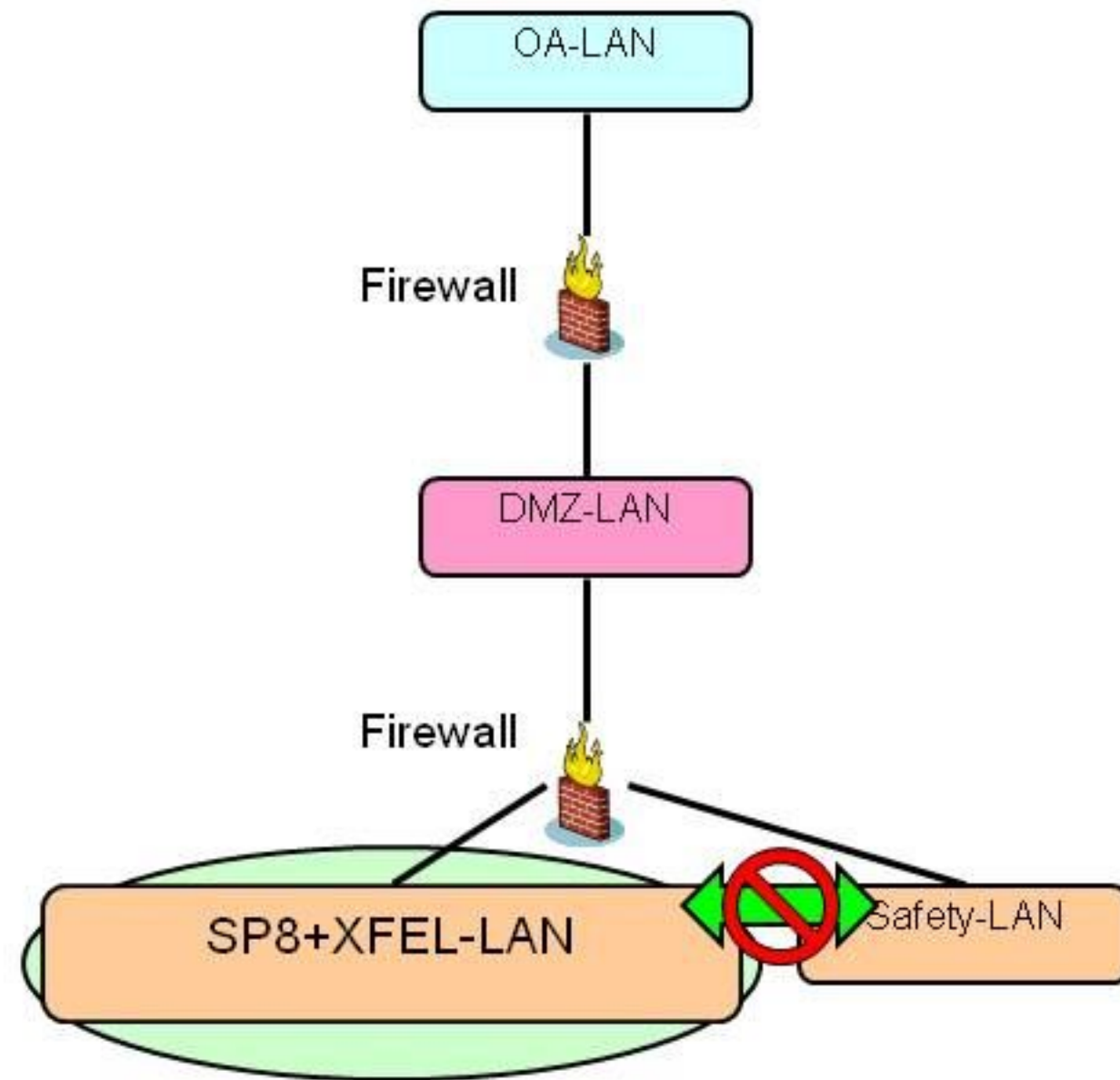
- When accelerators operated independently, we want to keep independence between physical networks.
- When combination operations of accelerators are performed, we should archive control systems united.
 - We only change firewall policy.



Interconnection of physical network

- When accelerators operated independently, we want to keep independence between physical networks.
- When combination operations of accelerators are performed, we should archive control systems united.
 - We only change firewall policy.
 - We consider two physical network as one logical network.

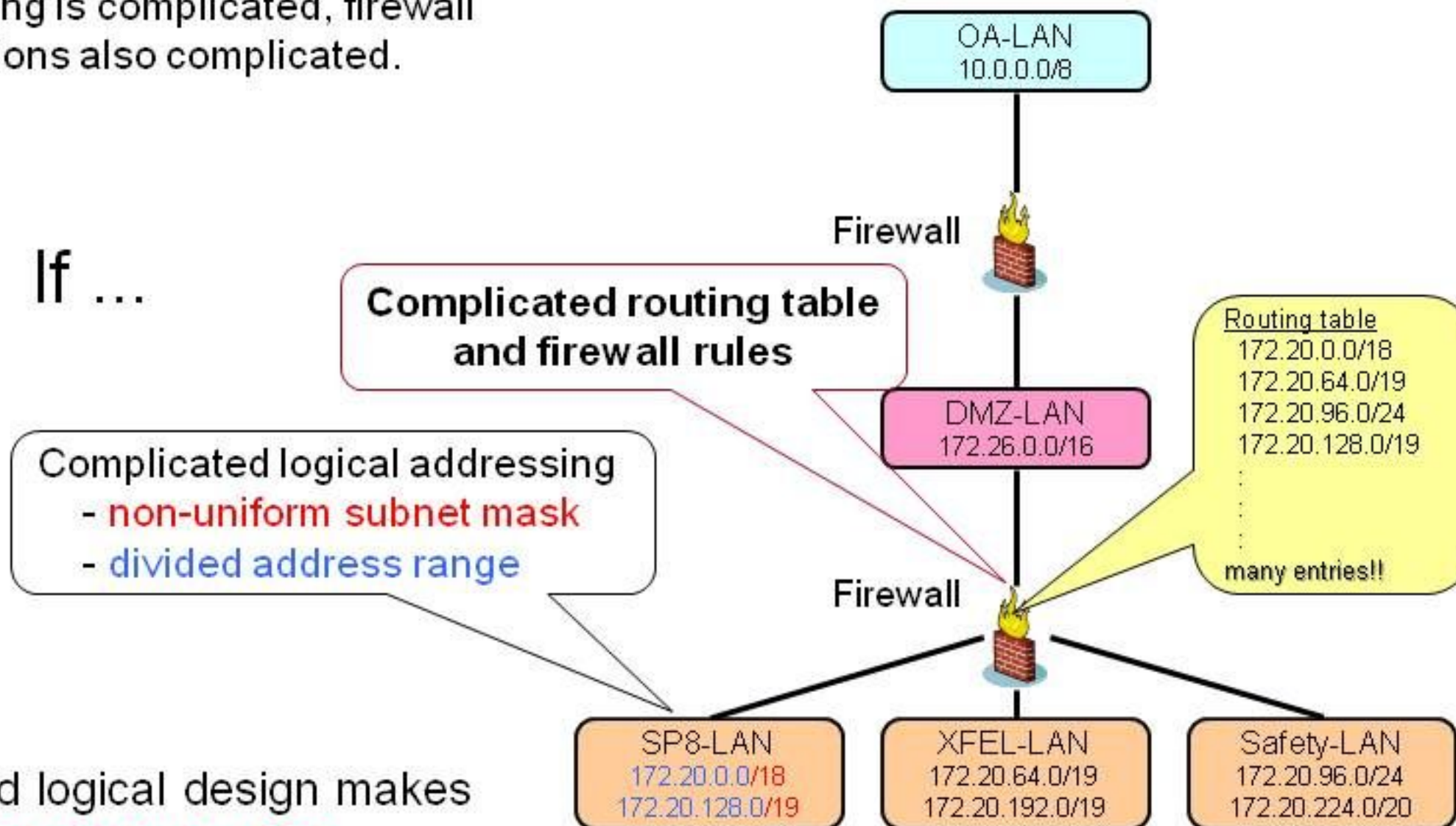
Then, two accelerators can run cooperated.



Logical addressing to physical network

- Logical addressing is very important.
 - If addressing is complicated, firewall configurations also complicated.

If ...

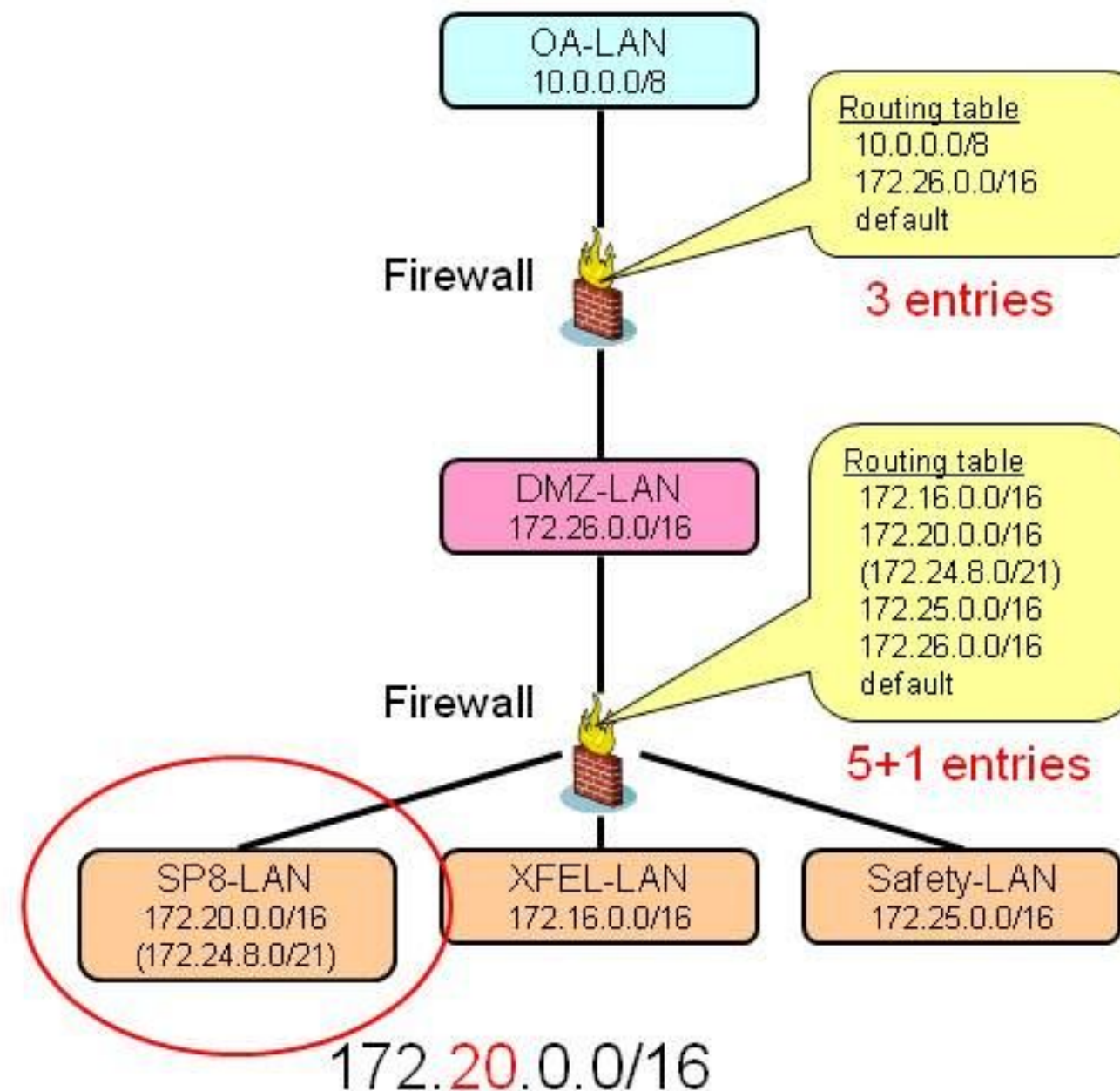


Complicated logical design makes management cost very high.

Logical addressing to physical network

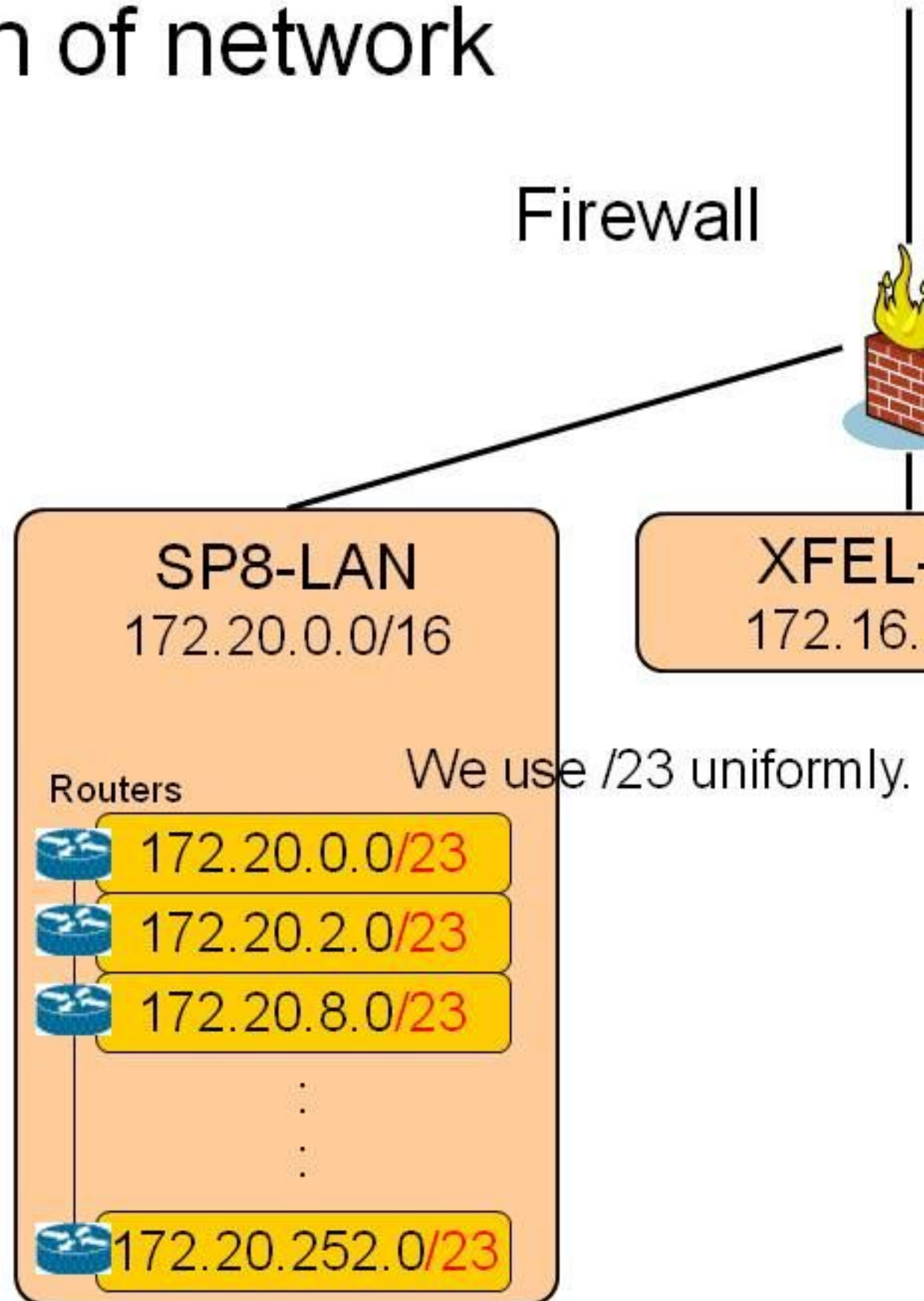
- Logical addressing is very important.
- We should assign IP address range and routing table **with simple policy**.
 - We assigned /16 address ranges to each LAN.
 - 2nd octet indicates physical network.
 - 16: XFEL-LAN
 - 20: SP8-LAN (24: old SP8-LAN)
 - 25: Safety-LAN
 - 26: DMZ-LAN
 - Then, routing table is very simple.
 - We should manage a few entries.
 - No dynamic routing is necessary.

Note:
/16 address range (65000 nodes) is enough to solve IP address exhaustion.
(Old network is /21 range, up to 2000 nodes)



Segmentation of network

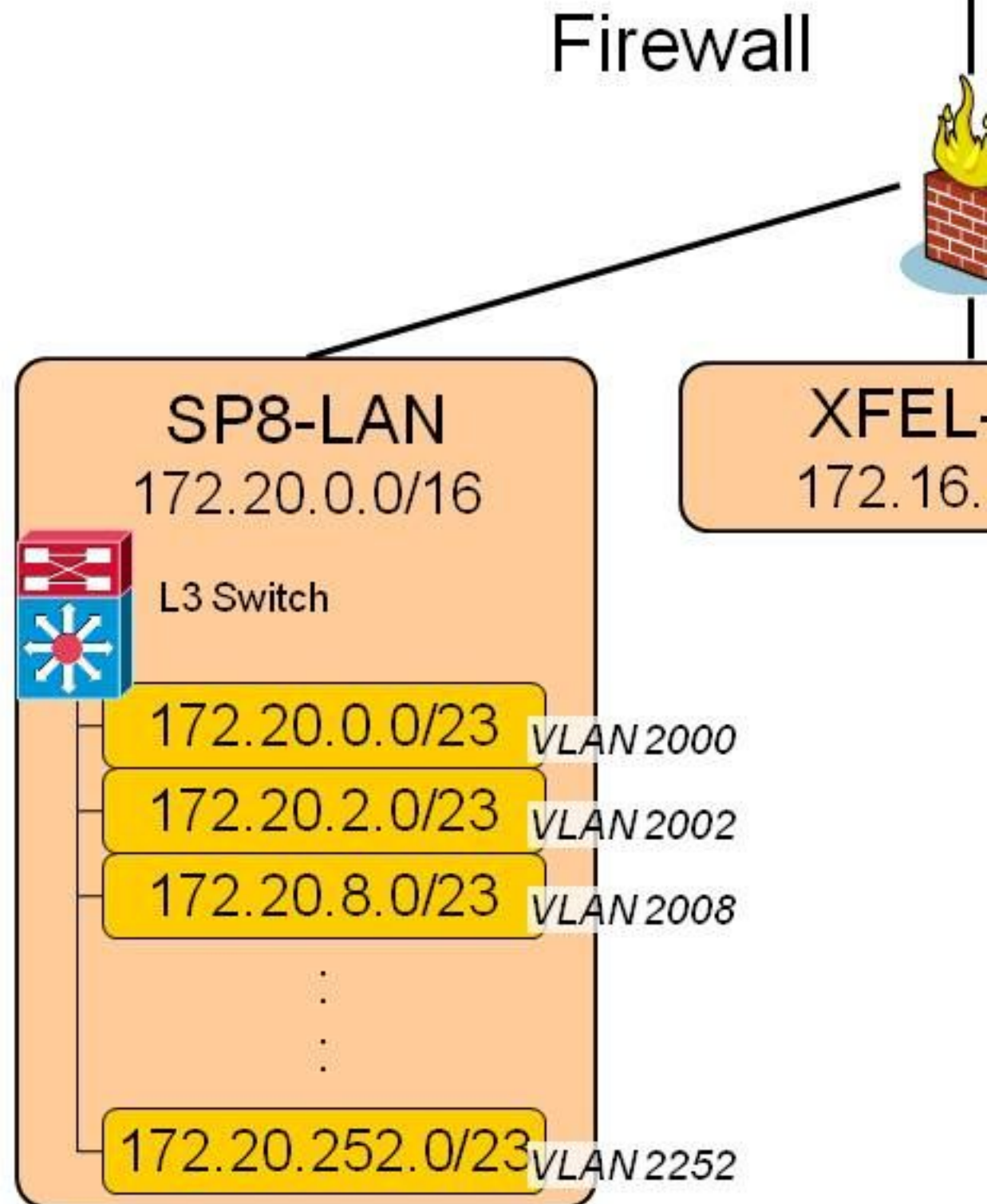
- Each LAN should be segmented into small network.
 - Because segmentation benefits some embedded devices suffered from heavy broadcast loads.



Segmentation of network

- Each LAN should be segmented into small network.
 - Because segmentation benefits some embedded devices suffered from heavy broadcast loads.
- We use L3 switch for the purpose of inter-segment routing.
 - Using L3 switch, we could reduce number of managed routers.
 - We assigned 802.1Q VLAN ID to each network using simple relational expression.

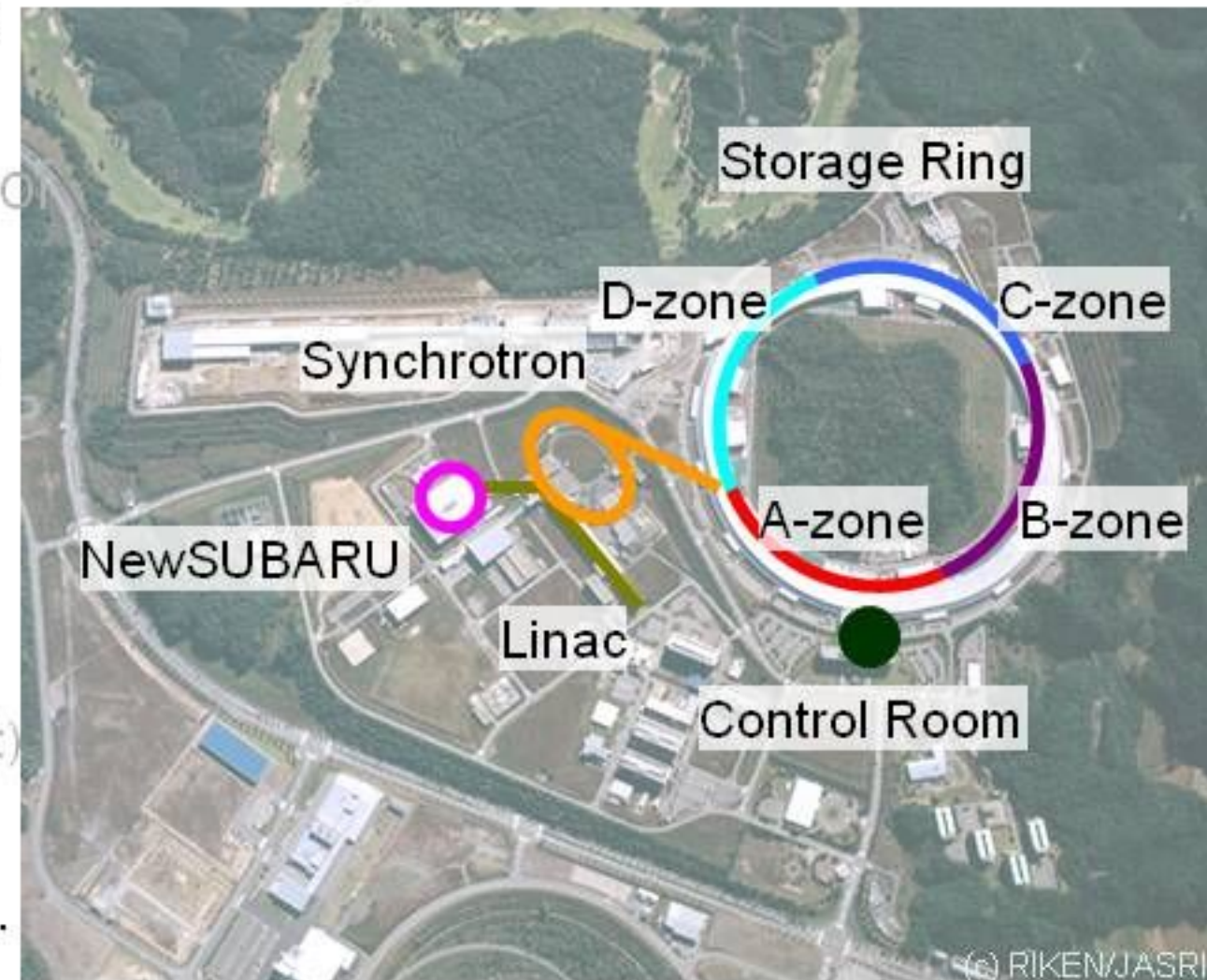
$$\text{VLAN ID} = (\text{2nd octet}) \times 100 + (\text{3rd octet})$$



Segmentation of network

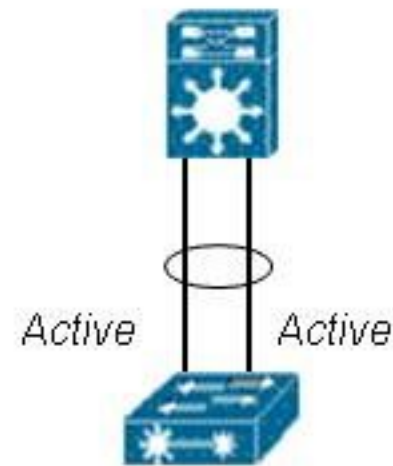
- Each LAN should be segmented into small network.
 - Because segmentation benefits some embedded devices suffered from heavy broadcast loads.
- We use L3 switch for the purpose of inter-segment routing.
 - Using L3 switch, we could reduce number of managed routers.
 - We assigned VLAN ID to each network using simple relational expression.
$$\text{VLAN ID} = (\text{2nd octet}) * 100 + (\text{3rd octet})$$
- Then, network is segmented for example Linac, Synchrotron, ...

Segmentation of SP8-LAN

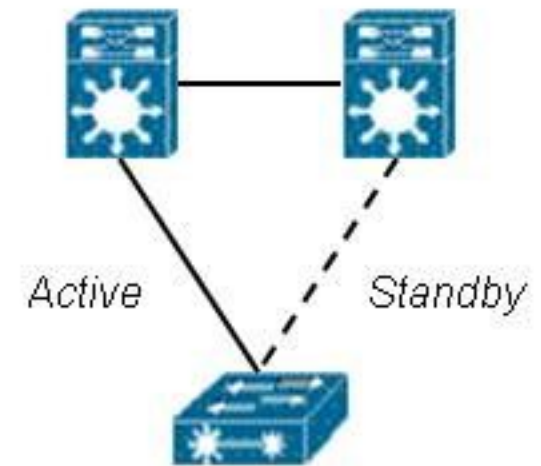


Redundancy for non-stop operation

- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).



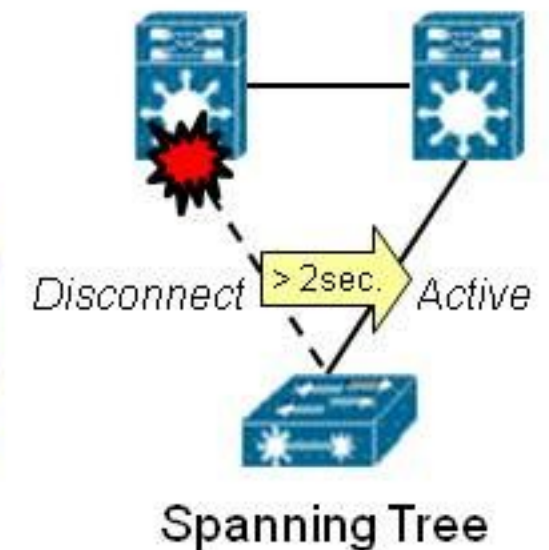
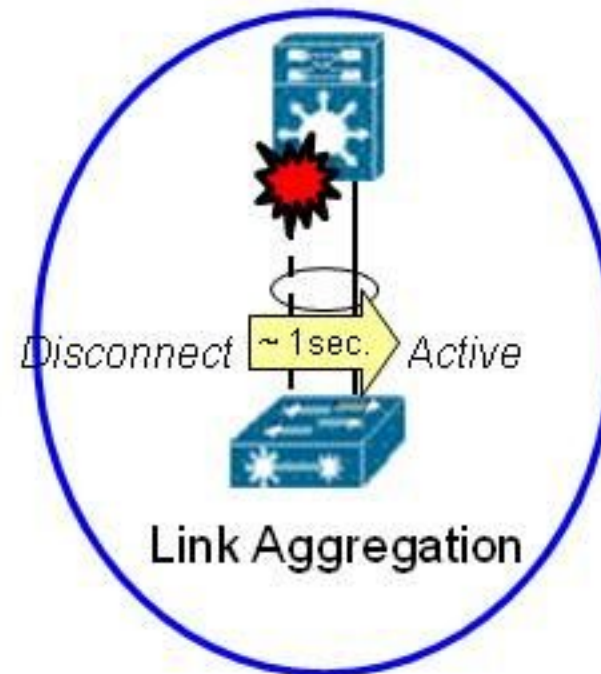
Link Aggregation



Spanning Tree

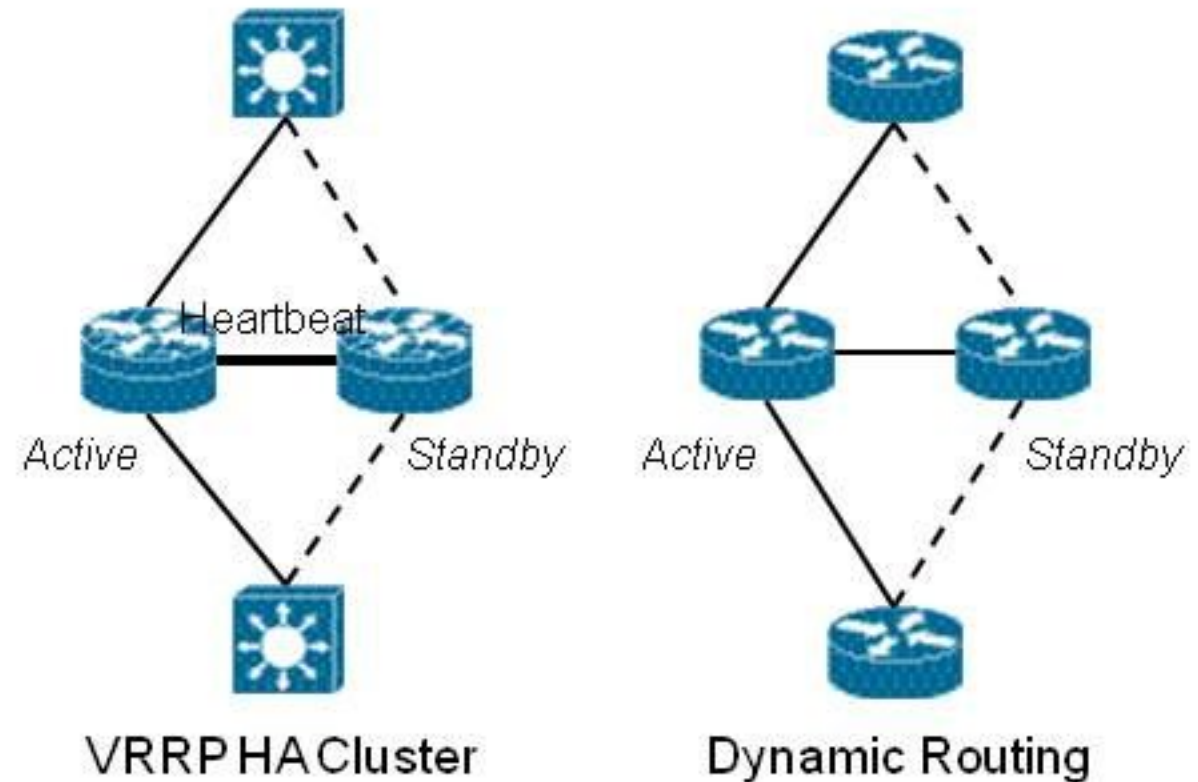
Redundancy for non-stop operation

- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).



Redundancy for non-stop operation

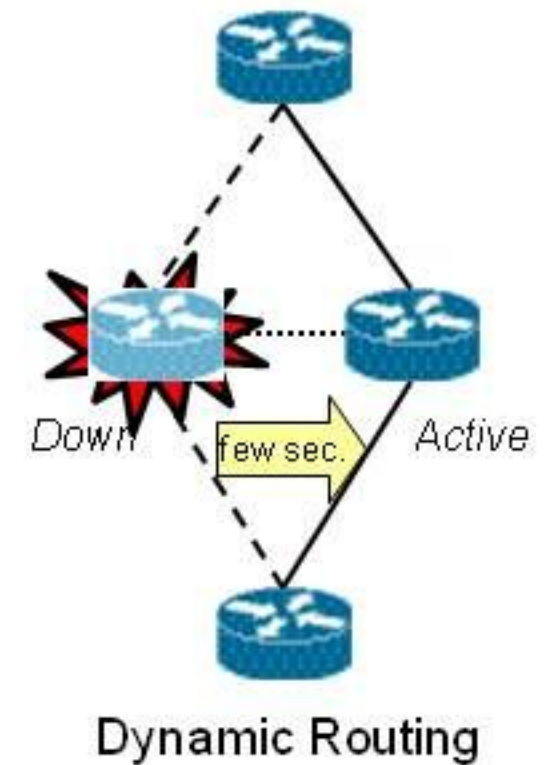
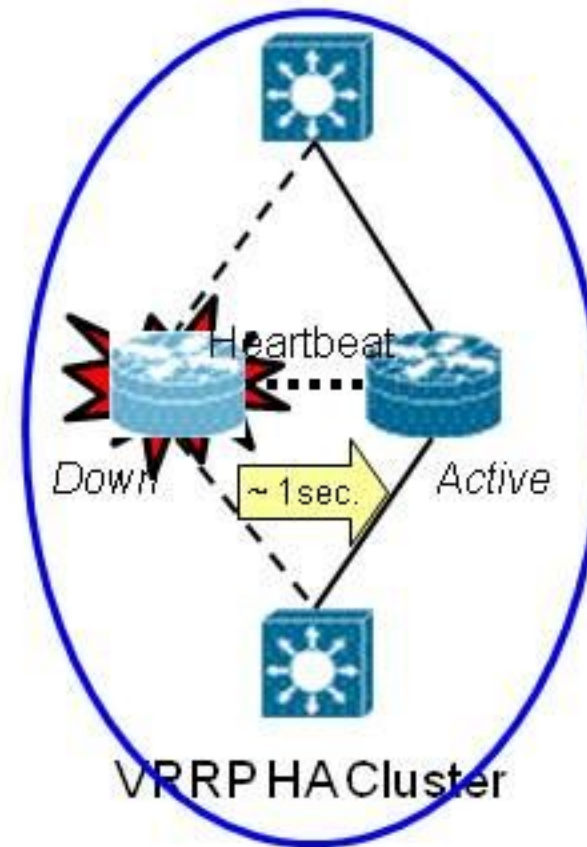
- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).
 - For layer 3, VRRP HA clustering (~ 1 sec.) is preferred rather than dynamic routing such as RIP, OSPF, BGP (> few sec.).



These technologies are already used at core-distribution network switches.

Redundancy for non-stop operation

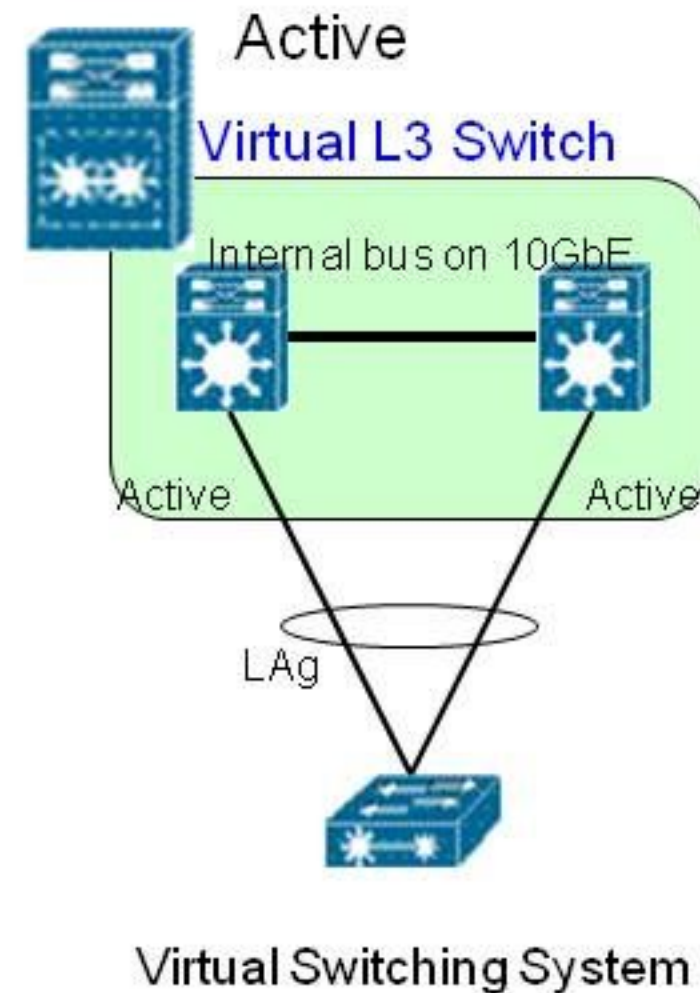
- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).
 - For layer 3, VRRP HA clustering (~ 1 sec.) is preferred rather than dynamic routing such as RIP, OSPF, BGP (> few sec.).



These technologies are already used at core-distribution network switches.

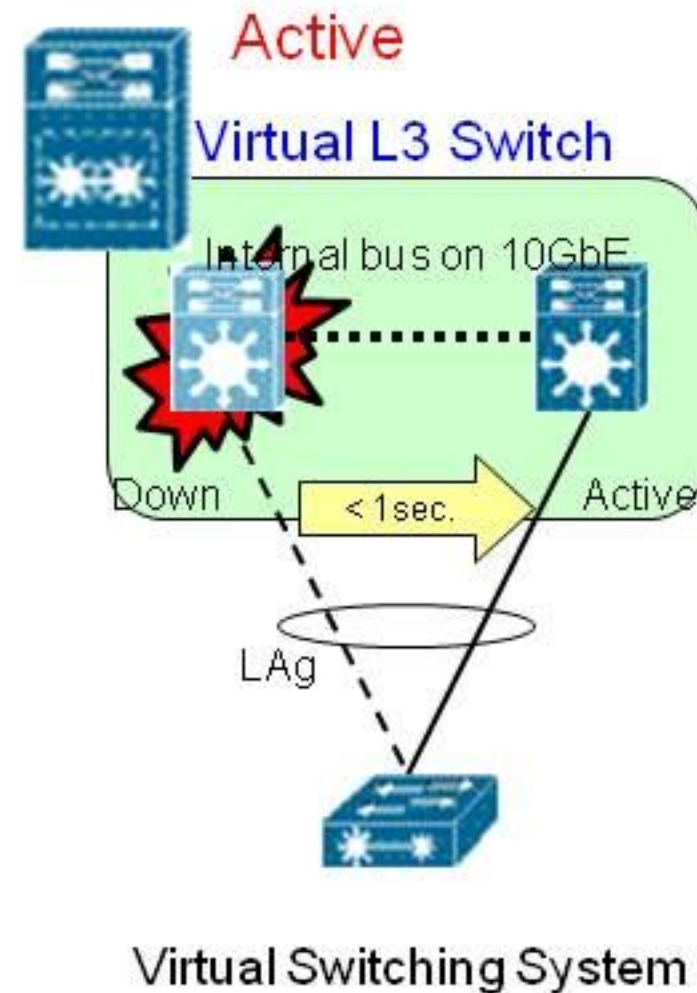
Redundancy for non-stop operation

- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).
 - For layer 3, VRRP HA clustering (~ 1 sec.) is preferred rather than dynamic routing such as RIP, OSPF, BGP (> few sec.).
- We have another plan to introduce Cisco VSS (< 1 sec.) for both L2 and L3 redundancy.



Redundancy for non-stop operation

- We must archive continuous network operation with < 1 sec. failover time.
 - For layer 2, LAg (~ 1 sec.) is preferred rather than STP or RSTP (> 2 sec.).
 - For layer 3, VRRP HA clustering (~ 1 sec.) is preferred rather than dynamic routing such as RIP, OSPF, BGP (> few sec.).
- We have another plan to introduce Cisco VSS (< 1 sec.) for both L2 and L3 redundancy.



Refurbishment of control network

- We performed upgrade of control network in summer 2009.
 - It took 4 days (50 man-days).
 - More than 1000 nodes were reconfigured:
 - IP address / default route
 - Hostname and domain name (with new naming conventions)
 - Name resolution (DNS and NIS)
 - File system (NFS)
 - other configurations
- 1 month has passed since autumn machine time started.
 - No trouble has been reported.
- Now, control network is ready to integrate XFEL control system.

Summary

- In 2011, XFEL will be established, and its control system must be ready in 2010.
- There are many requirements to integrate new control system.
 - Keywords: flexibility, scalability, stability, availability, management
- **We considered solutions to requirements**
 - Segregation of physical network
 - Interconnection of physical network
 - Logical addressing to physical network
 - Segmentation of network
 - Redundancy for non-stop operation
- We successfully refurbished control network, and now the network is ready to integrate XFEL control system.

Acknowledgements

- We thank our colleagues for supporting this work.
 - SPring-8 Controls and Computing Division:
 - A. Yamashita, T. Masuda, Y. Furukawa, T. Matsushita, R. Fujihara, T. Hamano, T. Hirono, M. Kago, K. Kawata, M. Koderu, T. Matsumoto, M. Takeuchi, M. Yamaga
 - SPring-8 Safety Division:
 - O. Nakatani
 - RIKEN XFEL Control System Group:
 - T. Fukui
 - SPring-8 Accelerator Division
 - LASTI, University of Hyogo