Implementation and Evaluation of Network Management System to Reduce Management Cost Caused by Server Virtualization

Masahiro Yoshizawa, Toshiaki Tarui, and Hideki Okita

Hitachi, Ltd., Central Research Laboratory

Presented at 2nd Workshop on Data Center - Converged and Virtual Ethernet Switching (DC CAVES)
September 6, 2010, Amsterdam, The Netherlands

Objective

- Reduce **operation time** and **operator errors** caused by server virtualization to reduce management cost of data centers.

    1. Develop **an integrated network management system (NMS)** to manage both physical and virtual network elements.

    2. Define **an information model** to represent logical network layers that can be used to develop the integrated NMS.

1

Outline

1. Management Cost Caused by Server Virtualization

2. Related Work

3. Proposal of Virtual Network Management Information Model (VNMI Model)

4. Implementation of Network Management System (NMS) based on VNMI Model

5. Experiments

6. Evaluations

7. Conclusions and Future Work

2

In server virtualization environment, operation time and operator errors are increased by:

1. Different management interfaces between physical network elements and virtual network elements

2. Increasing frequency of changing network configurations

1-2. Virtual Switch

- Virtualization technologies (e.g., VMware, Virtage) make virtual network elements, such as virtual switches.

Example: Virtage (Hitachi's hardware partitioning system)
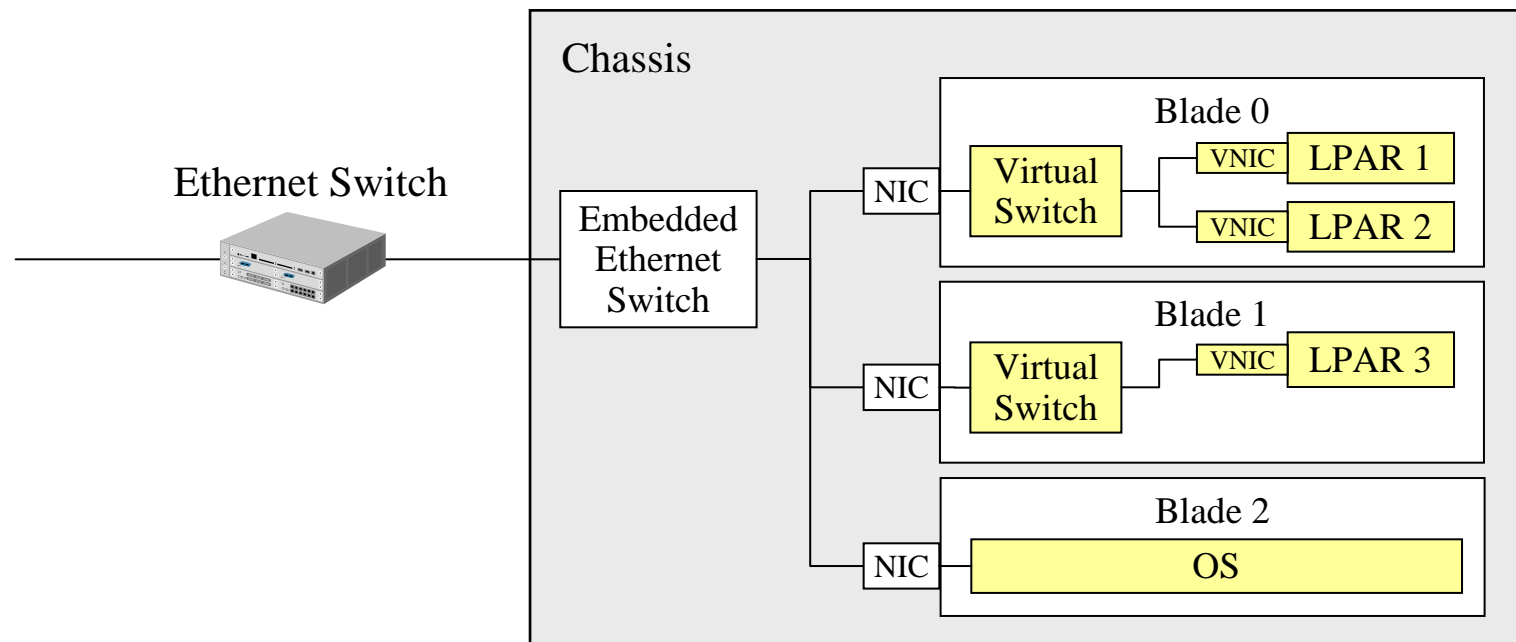
(*) VNIC: Virtual NIC
(**) LPAR: Logical PARtition

1-3. Different Management Interfaces

- Management Interface for physical switches
  - Simple Network Management Protocol (SNMP)
  - Device-specific Command-line Interface (CLI)

- Management Interface for virtual switches
  - Remote console of each hypervisor

- Different interfaces require additional operations in each phase of network management (e.g., designing, constructing, and monitoring the network).

- Network configurations need to be changed more frequently in server virtualization environment.

    – This flexibility is regarded as an advantage of server virtualization.

- These operations require changes in configurations of both physical and virtual network elements.

    – Virtual Local Area Network (VLAN)

    – Quality of Service (QoS)

    – Access Control List (ACL)

2-1. Related Work

- ## Improving Functionality of Virtual Switches
  - Cisco and Arista
  - IEEE 802.1Qbg/h (HP and Cisco are leading.)

- ## Standardizing Information Model of Virtual Switches
  - Hypervisor-specific CIMs: VMware, Xen, ...
  - Generic CIM extension
    - W. Fuertes, et al.
    - DMTF

7

2-2. Improving Functionality of Virtual Switches

- Improving management interfaces of virtual switches
  - Nexus 1000V and Arista vEOS support SNMP and a familiar CLI.
  - Operators can manage both switches in a familiar way.
  - Still difficult to configure and monitor multiple virtual switches distributed to data center networks.

- Aggregating functions of virtual switches into fewer physical switches
  - IEEE 802.1Qbg Edge Virtual Bridging (EVB) and 802.1Qbh Bridge Port Extension (standards under discussion)
  - Cisco VN-Link technology (product technology)
  - Operators have to manage fewer physical switches.
  - Technologies – not standardized yet

8

- Hypervisor vendors have defined their Common Information Model (CIM) extension.
  - VMware CIM APIs, Xen-CIM project, ...

- W. Fuertes, et al. [15] defined a generic CIM extension.
  - W. Fuertes, F. Meneses, J. E. López de Vergara, et al., "A Generic Model for the Management of Virtual Network Environments", In Proc. of the 12th NOMS, pp. 813-816, 2010.

- The Distributed Management Task Force (DMTF) is defining standards of the virtual network elements.

(continued)

(continued)

- When models are standardized, it will be easy to build NMSs to support virtual switches on heterogeneous hypervisors.

- However, a simple model for representing an entire virtual network (physical/virtual) has not standardized yet.

- Integrated NMS based on the information model satisfies the following requirements:

#	Goal	Requirements
1	Represent logical network automatically	NMS should create information on logical networks by using configurations of physical network elements.
2	Change network configurations automatically	NMS should create new configurations of physical network elements and operating procedures by operator's request to logical networks.
3	Manage entire data center network	Information model should be used to represent a mix of physical and virtual servers and heterogeneous hypervisors.

11

1) Physical Layer

- Configurations of physical elements in a device-specific data format
- Device-specific operating procedures to change the configurations

2) Virtual Layer

- Configurations of logical elements in a common data format
- Logical operating procedures to change the configurations
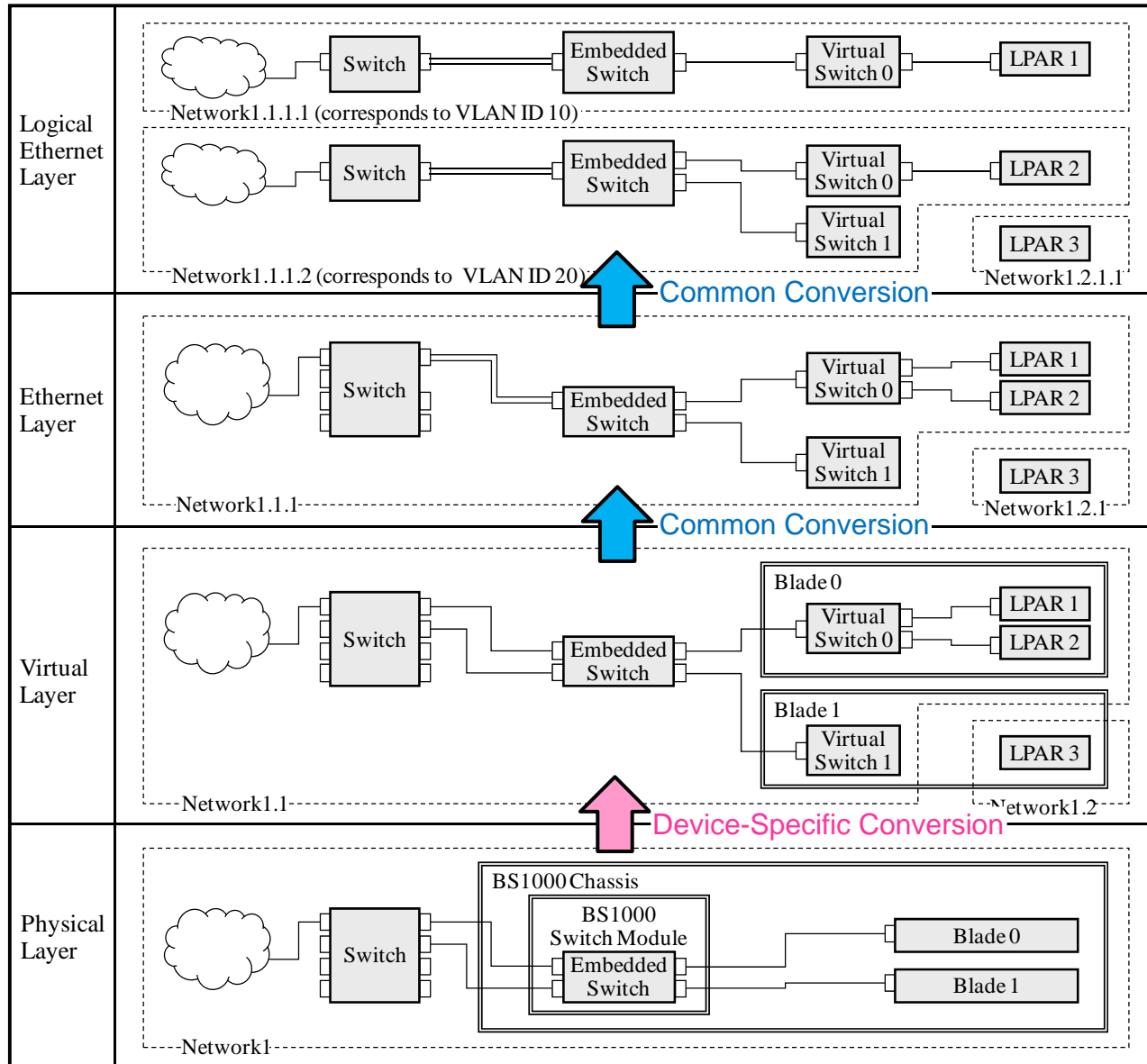
3) Ethernet Layer

- Configurations of logical elements that have Ethernet support
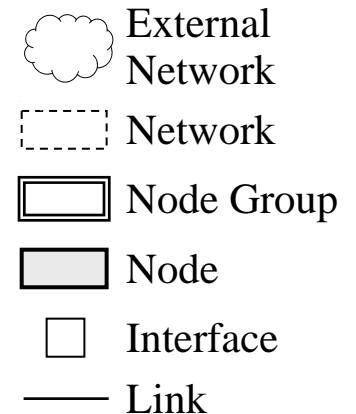
4) Logical Ethernet Layer

- Same as Ethernet Layer, but with different network representation
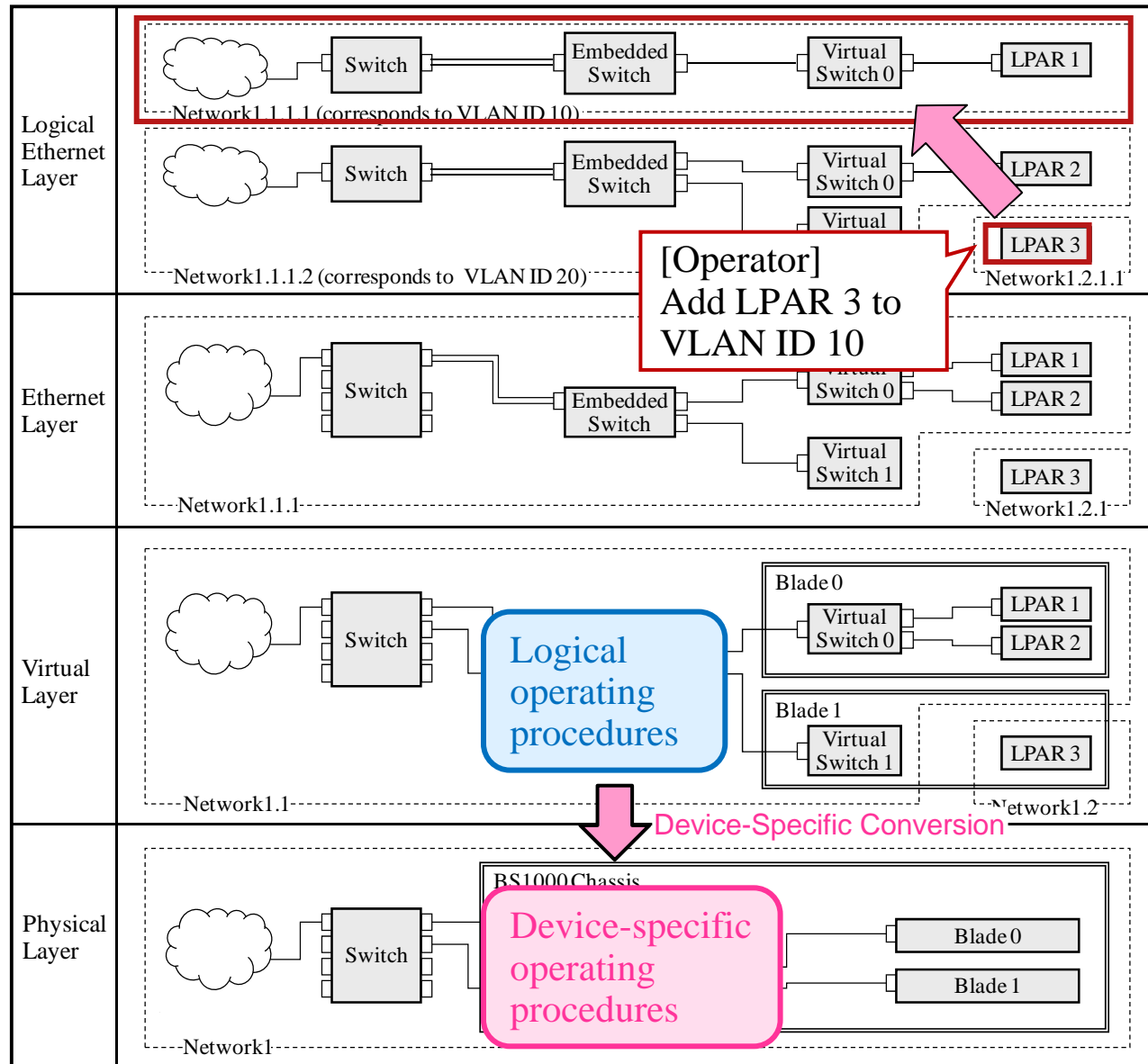- Logical ethernet network = VLAN

3-3. Creation of Logical Networks

Data of each layer can be created from the data of the lower layer

Logical Ethernet Layer

Network1.1.1.1 (corresponds to VLAN ID 10)
- Switch — Embedded Switch — Virtual Switch 0 — LPAR 1

Network1.1.1.2 (corresponds to VLAN ID 20)
- Switch — Embedded Switch — Virtual Switch 0 — LPAR 2
- Virtual Switch 1 — LPAR 3 (Network1.2.1.1)

Common Conversion

Ethernet Layer

Network1.1.1
- Switch — Embedded Switch — Virtual Switch 0 — LPAR 1 / LPAR 2
- Virtual Switch 1 — LPAR 3 (Network1.2.1)

Common Conversion

Virtual Layer

Network1.1
- Switch — Embedded Switch — Blade 0: Virtual Switch 0 — LPAR 1 / LPAR 2
- Blade 1: Virtual Switch 1 — LPAR 3 (Network1.2)

Device-Specific Conversion

Physical Layer

Network1
- Switch — BS1000 Chassis — BS1000 Switch Module — Embedded Switch — Blade 0 / Blade 1

Legend
- External Network
- Network
- Node Group
- Node
- Interface
- Link

3-4. Creation of Network Configuration

HITACHI
Inspire the Next
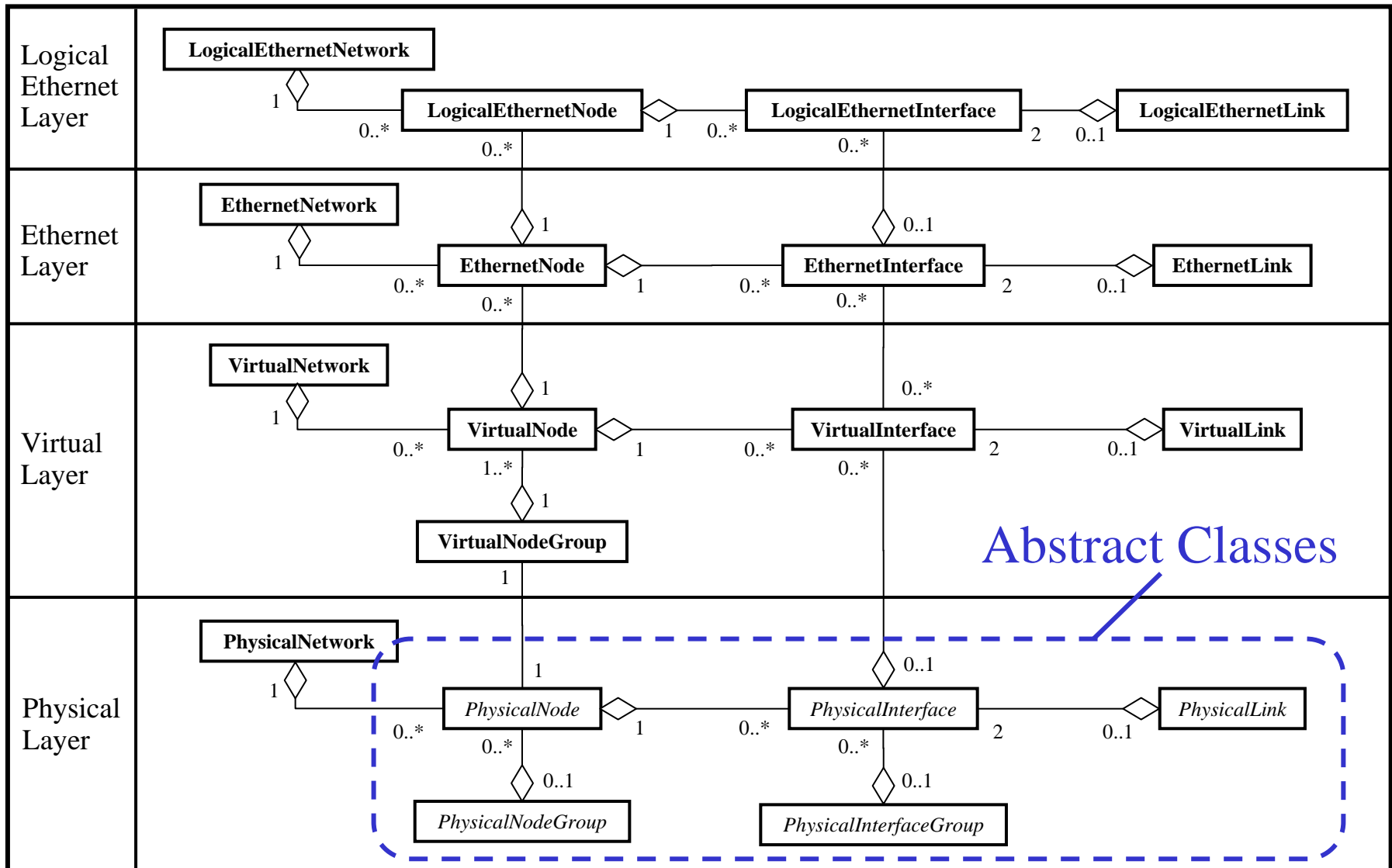
- Child classes of *PhysicalNode* class include device-specific configurations and conversion processes.

PhysicalNode

+ *virtualize() : VirtualNodeGroup*
    To convert the physical node into a group of virtual nodes
+ *applyProcedures(vProcs : List<VirtualProcedure>) : List<PhysicalProcedure>*
    To change the configurations of the physical node, and create physical procedures
+ *createConfigurationsElement(document : Document) : Element*
    To create an XML element from the configurations of the physical node
+ *parseConfigurationsElement(Element confElem) : void*
    To create the physical node from a specified XML element

VirtageServerA	VMwareServerB	EthernetSwitchC

1. Create VirtualNodeGroup objects by *virtualize()* method of each physical node object.

2. Guess connections between VirtualNode objects from data of the physical layer, and connect them.

Virtual Layer

Physical Layer

17
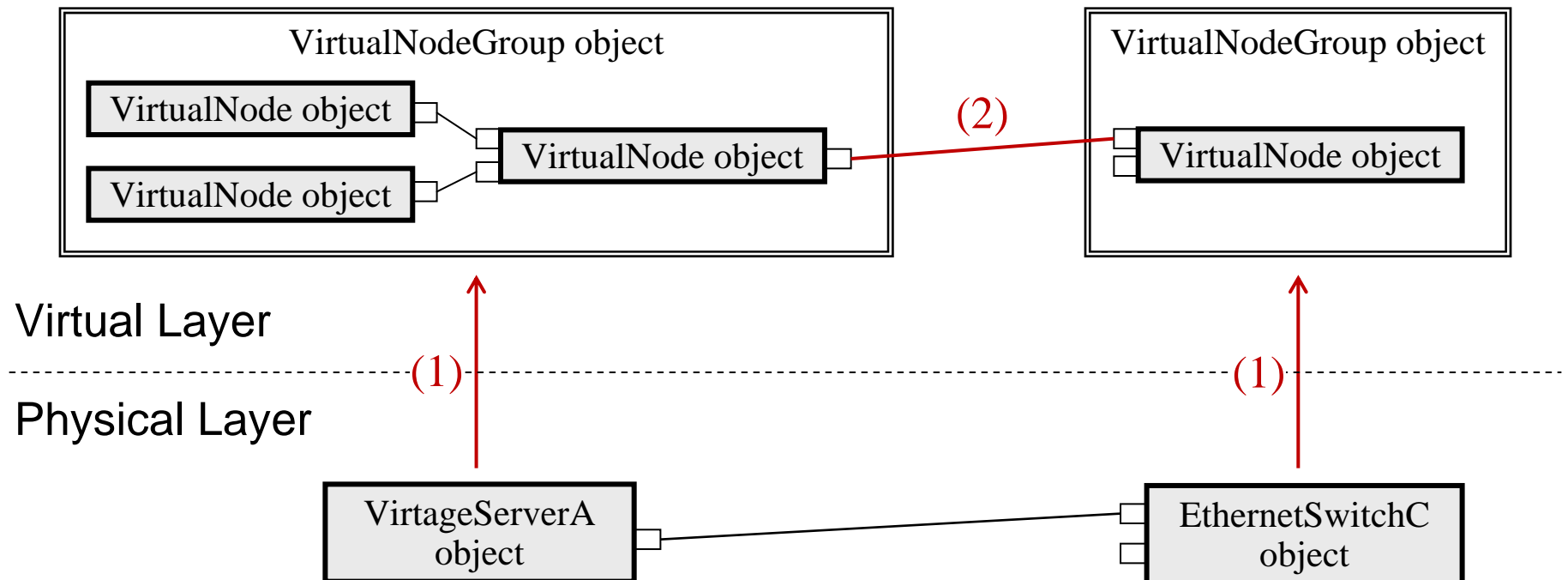
5-1. Experiments

- Prototype NMS - developed based on information model; experiments conducted

- Experiments - targeted at design phase of network management

- Operating procedures to change VLAN configurations - created by six examinees with and without prototype; examiner measured the time required to create them

18

5-2. Prototype

- Prototype with a command-line interface
  - Basic commands
    - Commands to load and save data of the physical layer
  - Show commands
    - Commands to show data of each layer
  - Design commands
    - Commands to change the configurations of the physical elements and create the operating procedures

- In the experiments, a design command to connect a virtual node (virtual machine or LPAR) to a logical ethernet network (VLAN) was mainly used.

19

```
EL# load Question2
Configuration data Question2 has been loaded.

EL# lnlist
Question2.1.1.1 [10] [Chassis1/Blade0.LPAR1, Chassis1/Blade0.LPAR2,
 Chassis2/Blade0.LPAR1, Chassis2/Blade0.LPAR2, Chassis2/Blade1.LPAR
1, Chassis2/Blade1.LPAR2]
Question2.1.1.2 [20] [Chassis1/Blade1.LPAR1, Chassis1/Blade1.LPAR2,
 Chassis1/Blade2.LPAR1, Chassis1/Blade2.LPAR2, Chassis1/Blade3.LPAR
1, Chassis3/Blade0.LPAR1, Chassis3/Blade0.LPAR2, Chassis3/Blade1.LP
AR1, Chassis3/Blade2.LPAR1]
Question2.1.1.3 [30] [Chassis2/Blade2.LPAR1, Chassis2/Blade2.LPAR2,
 Chassis2/Blade3.LPAR1, Chassis3/Blade1.LPAR2, Chassis3/Blade2.LPAR
2, Chassis4/Blade0.LPAR1, Chassis4/Blade0.LPAR2, Chassis4/Blade1.L
PAR1, Chassis4/Blade1.LPAR2]
Question2.1.1.4 [40] [Chassis1/Blade3.LPAR1, Chassis2/Blade3.LPAR1,
 Chassis3/Blade2.LPAR2]

EL#
```

```
EL# lnconnect
USAGE: lnconnect lpar_or_basic logical_ethernet [-i switch_to_ignore
...]

EL# lnconnect Chassis4/Blade2.LPAR1 Question2.1.1.1 -i AX-4.0
13 Virtual Procedure(s) and 7 Physical Procedure(s) has been generat
ed.

EL!# show-p-proc
(1) Set VLAN ID 10 to trunk port 0/2 of AX2400S AX-3.
(2) Add VLAN ID 10 to AX2400S AX-2.
(3) Set VLAN ID 10 to trunk port 0/13 of AX2400S AX-2.
(4) Set VLAN ID 10 to trunk port 0/2 of AX2400S AX-2.
(5) Add line number 0/8 to tagged-port of VLAN ID 10 of BS1000 embed
ded switch Chassis4/SW&M0/Switch.
(6) Add line number 0/2 to tagged-port of VLAN ID 10 of BS1000 embed
ded switch Chassis4/SW&M0/Switch.
(7) Change VLAN mode of VNIC 0 (segment 1a) of LPAR 1 of BS1000 blad
e Chassis4/Blade2 from Undef to Untag (VLAN ID = 10).
```

- Examinees were instructed to create operating procedures for connecting a specified LPAR to a specified VLAN.

- Examiner provided a network diagram and configuration sheets of each question to the examinees.

Test	BS1000(*) Chassis	BS1000(*) Blade	LPAR	AX2400S(**)
Example	1	2 (2 blades/chassis)	2 (1 LPAR/blade)	1
Question 1	2	4 (2 blades/chassis)	8 (2 LPARs/blade)	1
Question 2	4	14 (3 or 4 blades/chassis)	26 (1 or 2 LPARs/blade)	4

(*) Hitachi's blade server BladeSymphony 1000
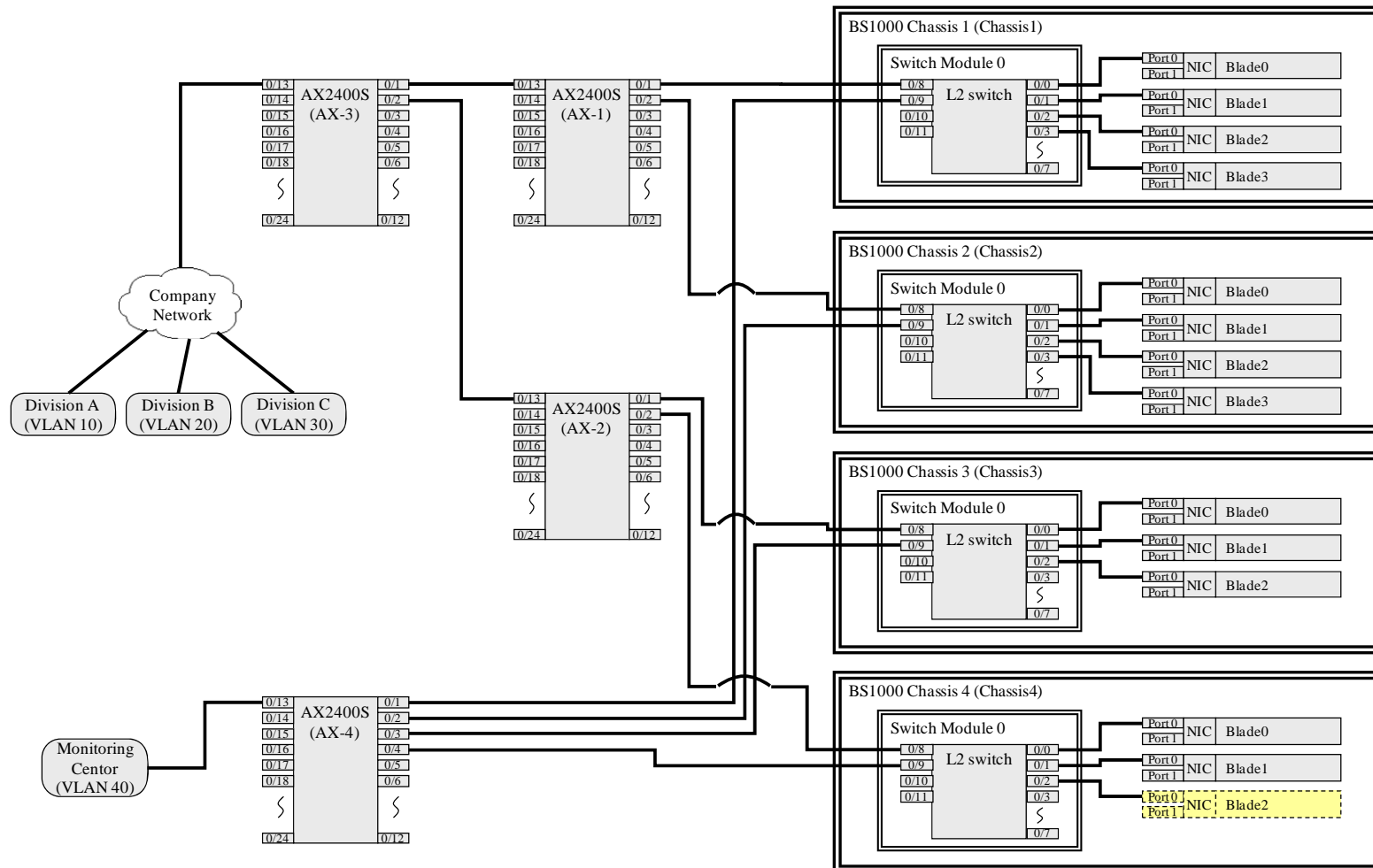(**) Alaxala's gigabit layer 2 switch

- When two LPARs are created on Chassis2/Blade1, connect one LPAR to VLAN 10, and another LPAR to VLAN 20.
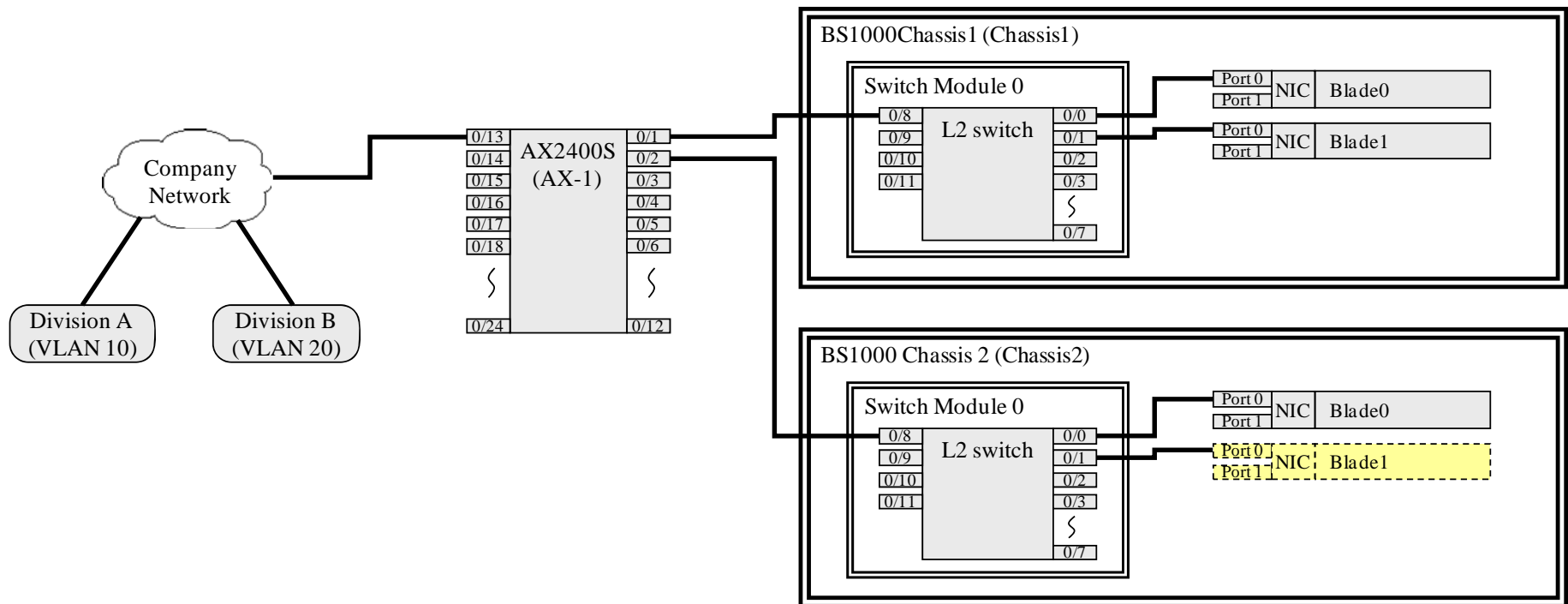
- When two LPARs are created on Chassis4/Blade2, connect one LPAR to VLAN 10, and another LPAR to VLAN 20.

5-5. Configuration Sheets

- Data structures of VLAN setting are device-specific.

AX2400S

ID	Port #	Port Type	VLAN Type	VLAN ID
AX-1	0/1	trunk	Port VLAN	10
	0/2	trunk	Port VLAN	20
	0/3	access	Port VLAN	1
	0/4	access	Port VLAN	1
	0/5	access	Port VLAN	1
	0/6	access	Port VLAN	1
	0/7	access	Port VLAN	1
	0/8	access	Port VLAN	1
	0/9	access	Port VLAN	1
	0/10	access	Port VLAN	1
	0/11	access	Port VLAN	1
	0/12	access	Port VLAN	1
	...	...	...	...

BS1000 Blade with Virtage

Chassis #	Blade #	LPAR #	VNIC Settings			
			VNIC #	VNIC Assignment	VLAN Mode	VLAN ID
1	0	1	0	1a	Untag	10
			1	1b	Undef	
			2	*		
			3	*		
			4	*		
			5	*		
			6	*		
			7	*		
		2	0	1a	Untag	10
			1	1b	Undef	
			2	*		
			...	...	...	...

5-6. Experimental Procedures

1) Explanation on VLAN configuration (about 30 min.)

   – VLAN configurations of BS1000 (blade) and AX2400S (switch), and how to change them.

2) Explanation of Prototype NMS (about 30 min.)

   – How to create the operating procedures with the prototype NMS.

   – The examinee uses the prototype and asks an example question.

3) First Experiment: Question One and Question Two

   – With prototype (Group A) / Without prototype (Group B)

4) Second Experiment: Question One and Question Two

   – Without prototype (Group A) / With prototype (Group B)

1) The prototype reduced the examinees' mistakes. The result implies that the prototype is effective to reduce the cost of managing data center networks.

2) The prototype reduced the time required to create the operating procedures, although the amount of the reduced time was small.

6-2. Analysis of Examinees' Mistakes

- Examinees' mistakes (i.e., wrong answers and correct answers that included redundant procedures) were reduced.

Without the NMS

Examinee	Question 1	Question 2
1	[1st] correct but redundant	[1st] correct but redundant
2	[1st] wrong [2nd] correct	[1st] wrong [2nd] correct
3	[1st] correct but redundant	[1st] correct
4	[1st] correct but redundant	[1st] correct
5	[1st] correct but redundant	[1st] correct
6	[1st] wrong [2nd] correct	[1st] correct

With the NMS

Examinee	Question 1	Question 2
1	[1st] correct	[1st] correct
2	[1st] correct	[1st] correct
3	[1st] correct	[1st] correct
4	[1st] correct	[1st] correct
5	[1st] correct	[1st] correct
6	[1st] correct	[1st] wrong [2nd] correct

- The time to answer the questions was reduced in many cases.
- Examinees 2 and 4 took longer to complete Question 2.

Examinee	Q.	Without NMS	With NMS	Reduced Time	Reduced Rate
1	1	13:03	12:25	0:38	4.9%
	2	21:18	21:10	0:08	0.6%
2	1	17:36	11:16	6:20	36.0%
	2	14:34	21:19	−6:45	−46.3%
3	1	7:56	5:17	2:39	33.4%
	2	10:00	6:51	3:09	31.5%
4	1	15:24	13:49	1:35	10.3%
	2	11:09	13:28	−2:19	−20.8%
5	1	20:44	18:11	2:33	12.3%
	2	18:44	15:07	3:37	19.3%
6	1	14:01	9:57	4:04	29.0%
	2	12:01	11:30	0:31	4.3%

Reason: Difficulty in using the prototype's CLI

⇩

The usability should be improved

7-1. Conclusions

- The prototype NMS based on the proposed information model reduced operator errors in creating VLAN configurations.

- The NMS is expected to be effective in reducing the cost of network management at each phase.

7-2. Future Work

- Apply the NMS to other phases of network management. (construction phase and monitoring phase)

- Implement the plugins that support other hypervisors (e.g., VMware) to verify the feasibility of the information model.

- Standardize the information model, and confirm whether part of the information model already meets some existing standards (e.g., DMTF).

Implementation and Evaluation of Network Management System to Reduce Management Cost Caused by Server Virtualization

September 6, 2010

Hitachi, Ltd., Central Research Laboratory
M. Yoshizawa, T. Tarui, and H. Okita