
OPEN
Compute Project
OCP U.S. SUMMIT 2017
Santa Clara, CA

OCP Trial Results for Telco Infrastructure

Jungsoo Kim/Manager/SKTelecom

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

OPEN Compute Project

Agenda

SK telecom

SKT Infrastructure

- Infrastructure Evolution
- COSMOS
- Open Hardware & Software in SK Telecom

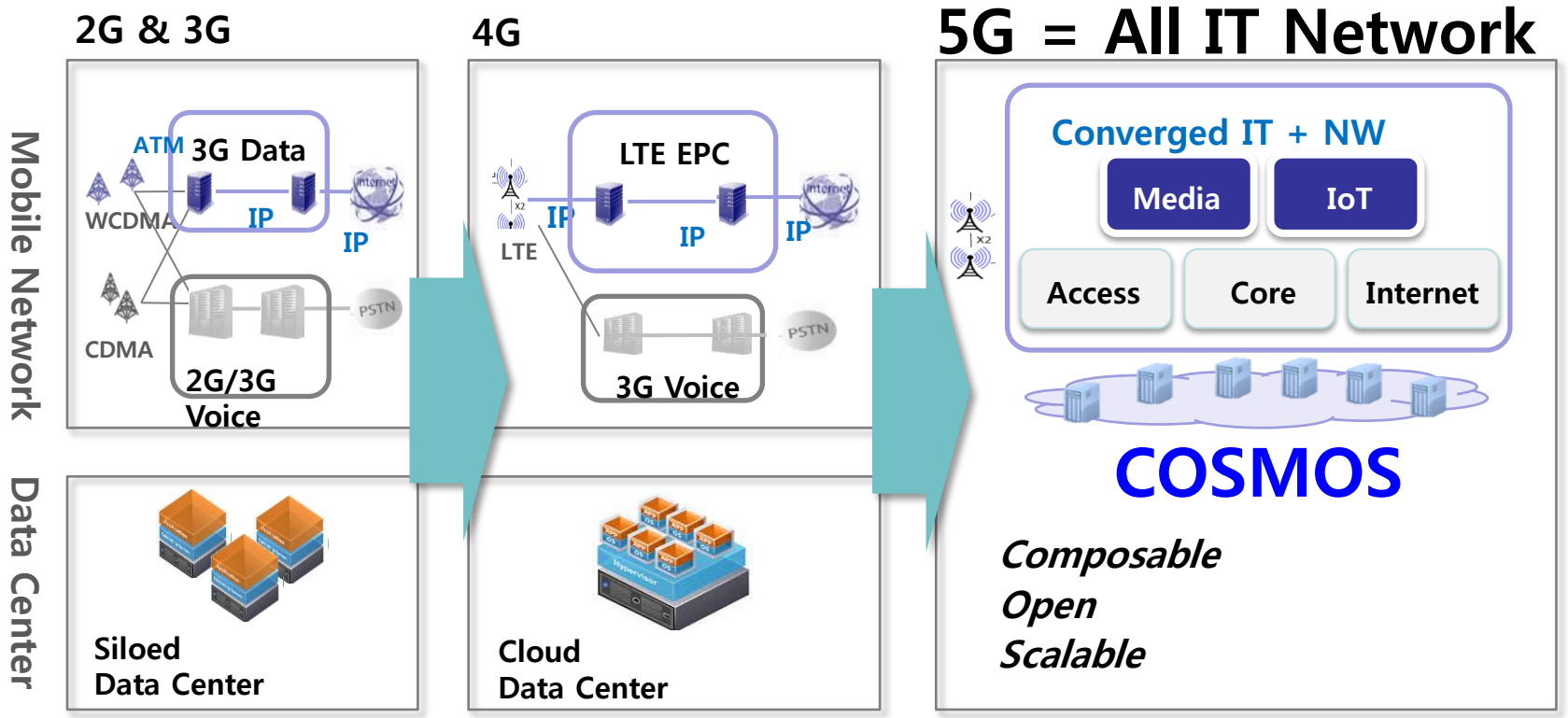
OCP Trial Results

- The Purpose of OCP Trial
- Characteristic Differences
- Trial Scope and Details
- Testbed Design
- Power Consumption Comparison
- Deployment Issues

Our Approach

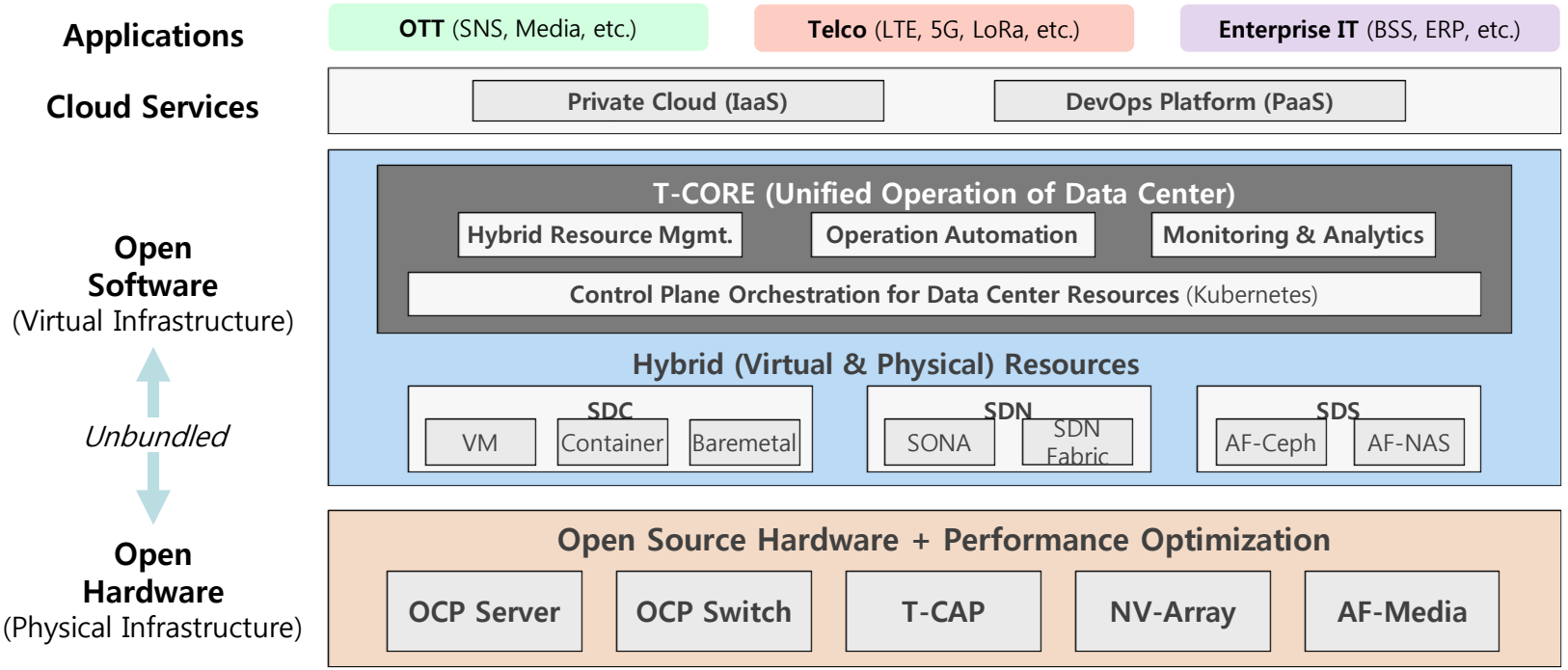
Next Step

Infrastructure Evolution



COSMOS

- **Composable, Open, Scalable, Mission-critical Optimized System**



Open Hardware/Software in SK Telecom

OPEN Compute Project

Gold Member
OCP Telco Project

Server | Switch | Storage

TELECOM INFRA PROJECT

Founding Member
Board Chair

RAN | Backhaul | Core

Community Lab TEAC

All-IT Network

Data Center

Telco Network

*TEAC: TIP Ecosystem Acceleration Center

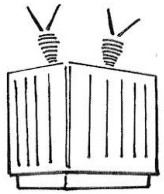
The Purpose of OCP Trial

Benefits of Open Compute at Facebook

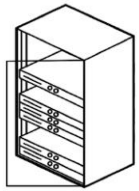
Compared to traditional servers.....

38% Increase in Power Efficiency

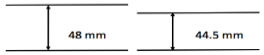
24% Reduction in Costs



Innovative
Power Distribution



21 inch wide



Open U

- ✓ **Verify OCP Technology**
  - **Efficiency, Performance, Serviceability**
- ✓ **Analysis Deployment Environment**
  - **Procurement, IDC Environment, Specification …**
- ✓ **Check OCP Ecosystem in Local**
  - **Vendor, Tech Support, Delivery …**

Characteristic Differences

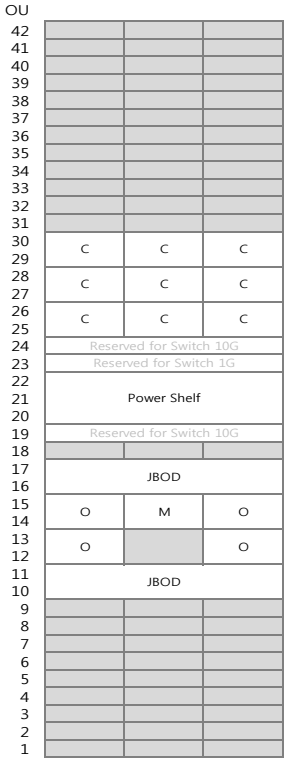
- Understanding of environmental difference is important key to success

	Hyper-Scale	SK Telecom
Volume	> Hundreds of Thousands	< Tens of Thousands
Age of Datacenter	Young	Old
Datacenter	A Few Big IDCs	Small Distributed IDCs
Infra Management	Centralized	Separated
Deployment	Rack Scale	System Scale
Major Supplier	ODM Vendor	Brand Vendor

Trial System Configuration

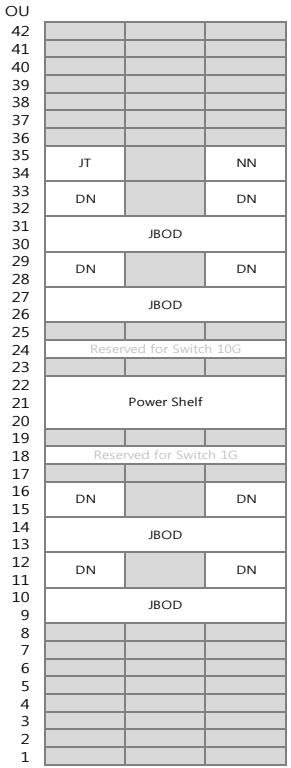
- Selected three rack configurations to test various kinds of application

OpenStack Rack



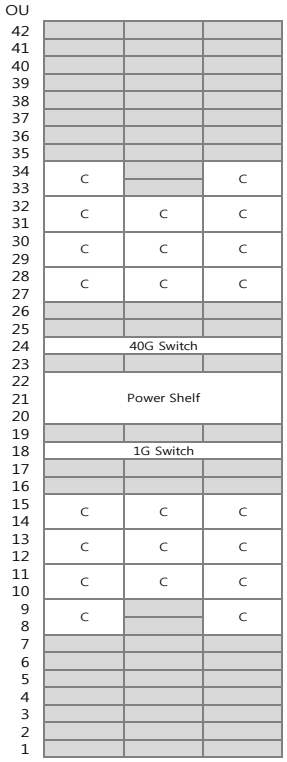
- Private Cloud
- Open Rack (21”)
- 9 Compute Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  *SSD 300GB*2ea*
- 5 Ceph Nodes
  *E5-2690 V4*2ea*
  Memory 256GB
  M.2 300GB
  *SSD 1T*6ea*

Hadoop Rack



- Big Data Analytics
- Open Rack (21”)
- 10 Hadoop Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  *SSD 450GB*2ea*
- 4 JBODs
  *HDD 4TB*12ea*

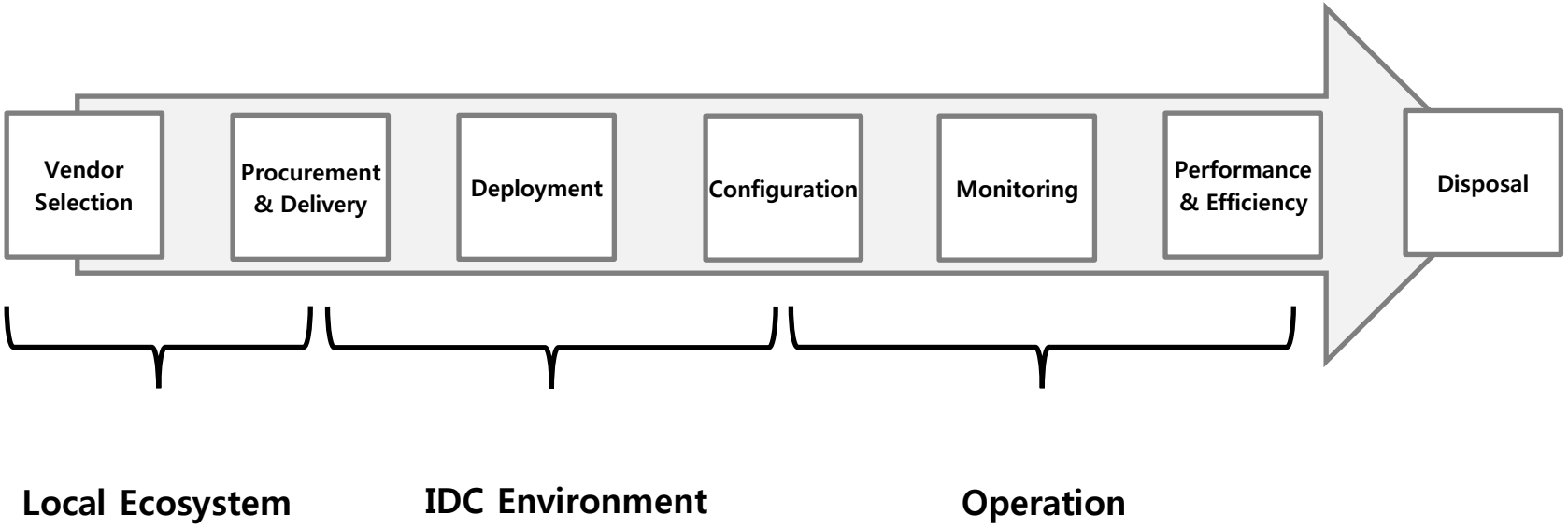
R&D Rack



- OpenStack/SDN, NFV Container
- Open Rack (21”)
  22 Compute Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  M.2 300GB

Trial Scope

- From the selection of vendor to operation environment, we wanted to check practical issues of whole lifecycle of OCP elements in local environment

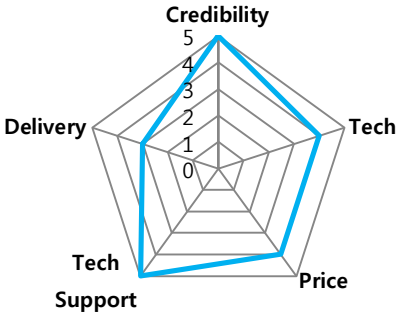


Vendor Selection

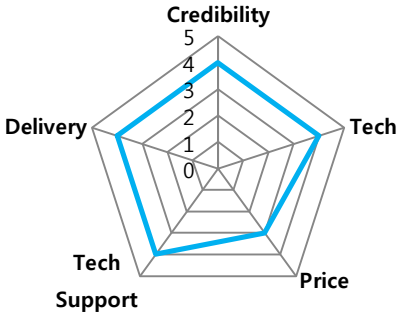
- Four Vendors responded to RFP
  - ✓ One vendor had Korea office, two vendors had a local partner, one vendor hadn't entered local market yet
- Assessed each vendor based on Credibility, Technology, Price, Tech Support and Delivery Time
- Based on the checklist, two vendors were selected as a OCP trial supplier

➢ Issues: Long delivery time (BTO based process), Wide price variations, Unskilled Technical Support

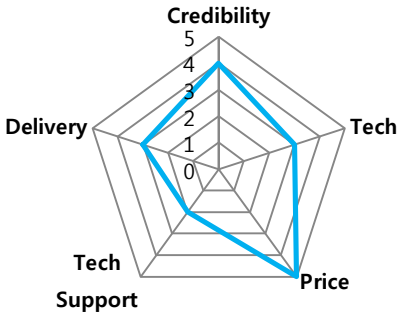
Vendor A



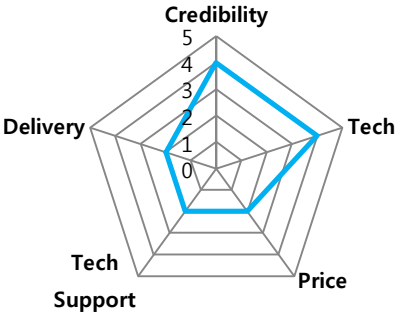
Vendor B



Vendor C

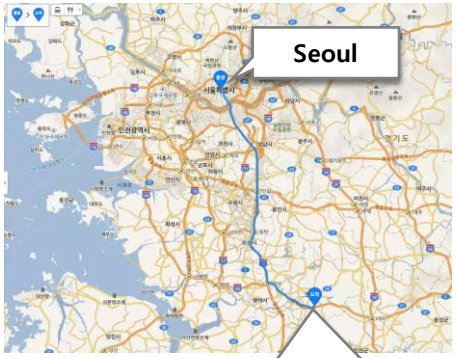


Vendor D

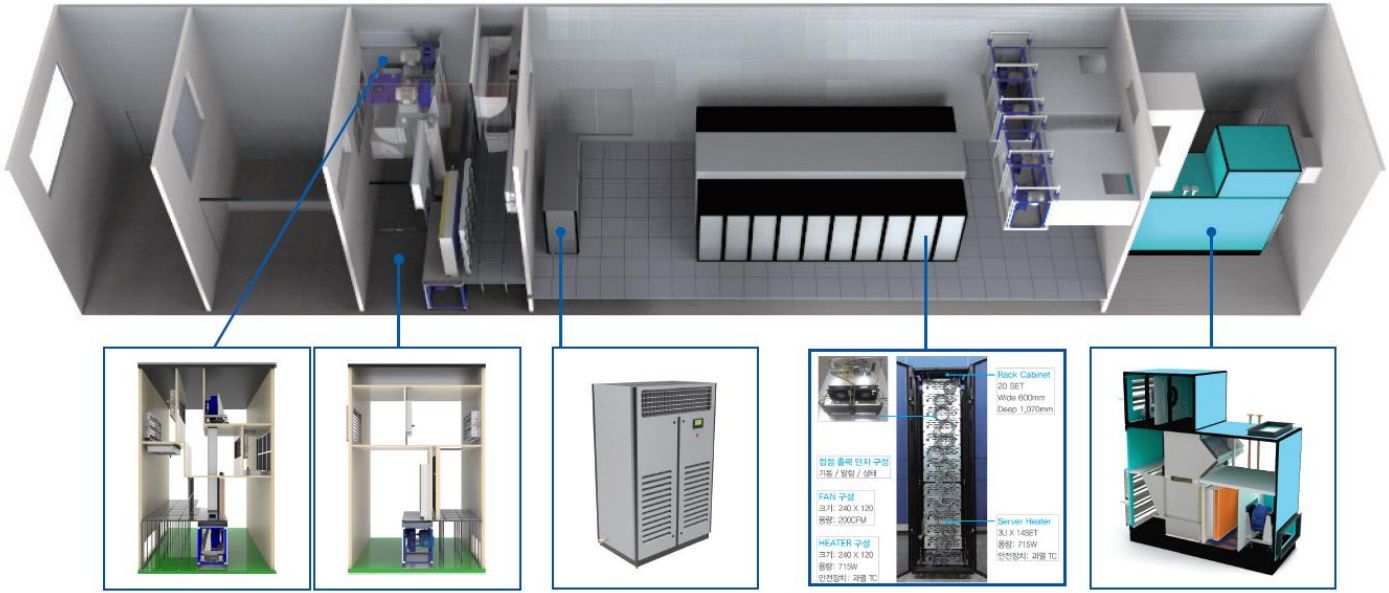


CRAC Testbed

- Rented CRAC Testbed for environmental & performance test
- Temp/Humidity Control, 20 Racks with Cold-Aisle Containment, Server Simulator, Sensors and loggers etc.



SAMHWA CRAC Testbed



*CRAC: Computer Room Air Conditioner

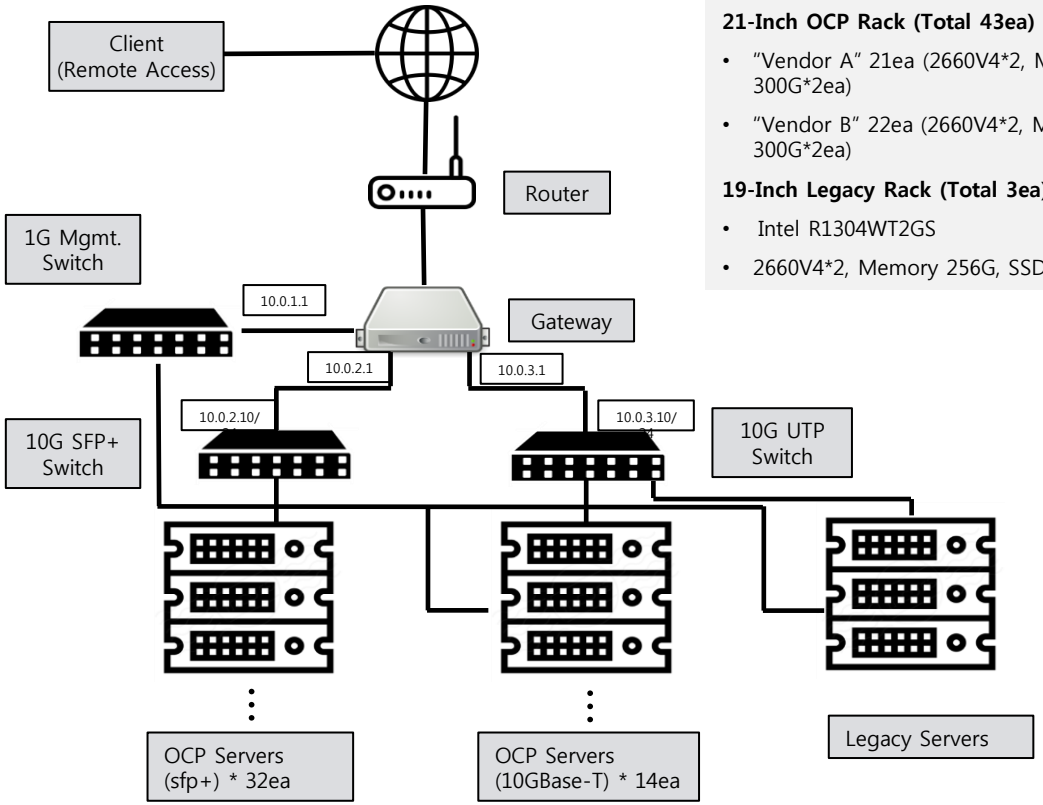
Testbed Design – Architecture

Open Rack

OU			
42	Void		
41	Void		
40	Void		
39	B	B	B
38			
37	B	B	B
36			
35	B	B	B
34			
33	B	B	B
32			
31	B	B	B
30			
29	Void		
28	Void		
27	B	B	B
26			
25	B	B	B
24			
23	Void		
22	Void		
21	Power Shelf		
20			
19			
18			
17	Void		
16	A	Void	Void
15			
14	A	A	A
13			
12	A	A	A
11			
10	A	A	A
9			
8	A	A	A
7			
6	A	A	A
5			
4	A	A	A
3			
2	A	A	A
1			

19-Inch Rack

U	
42	Void
41	Void
40	Void
39	Void
38	Void
37	Void
36	Void
35	Void
34	Void
33	Void
32	Void
31	Void
30	Void
29	Void
28	Void
27	Void
26	Void
25	L3
24	
23	L2
22	L1
21	Void
20	Void
19	Void
18	Void
17	Void
16	Void
15	Void
14	Void
13	Void
12	Void
11	Void
10	Void
9	Void
8	Void
7	Void
6	Void
5	Void
4	Void
3	Void
2	Void
1	Void



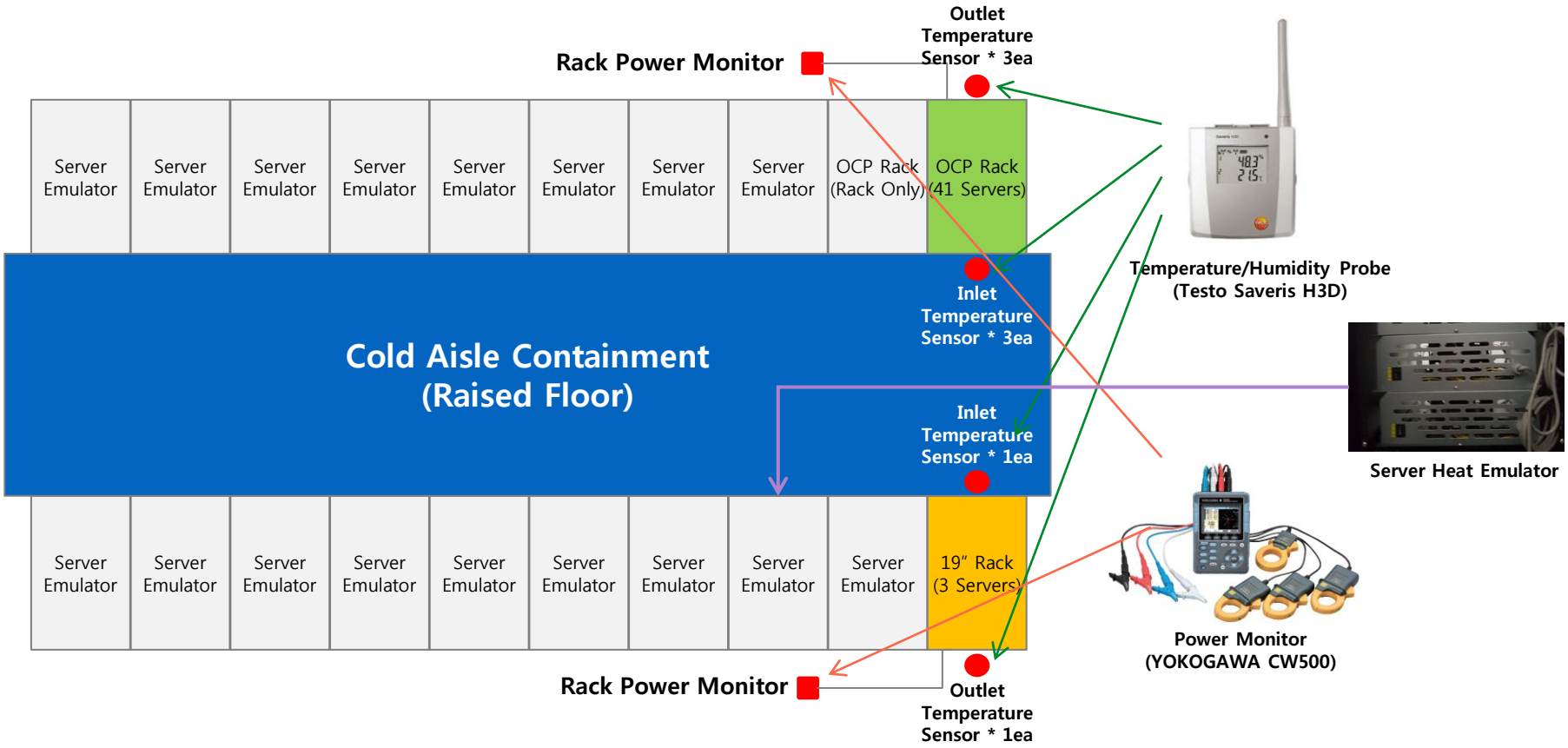
21-Inch OCP Rack (Total 43ea)

- "Vendor A" 21ea (2660V4*2, Memory 256G, SSD 300G*2ea)
- "Vendor B" 22ea (2660V4*2, Memory 256G, SSD 300G*2ea)

19-Inch Legacy Rack (Total 3ea)

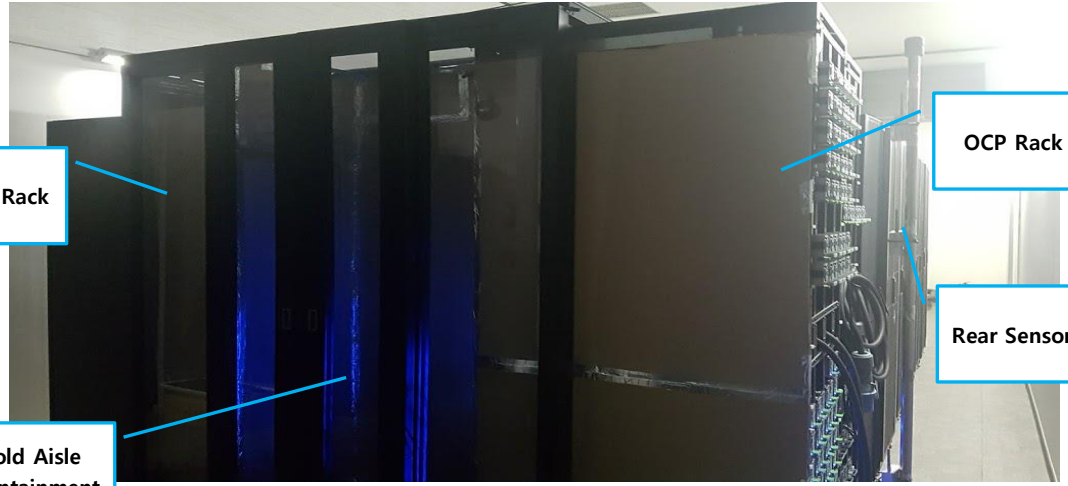
- Intel R1304WT2GS
- 2660V4*2, Memory 256G, SSD 480G*2ea

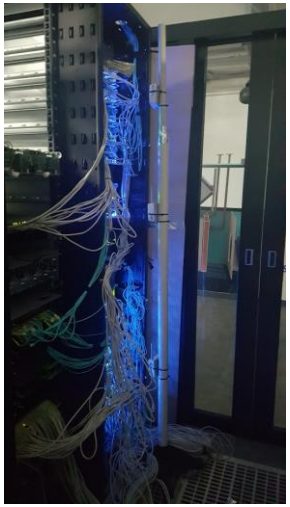
Testbed Design – Top View



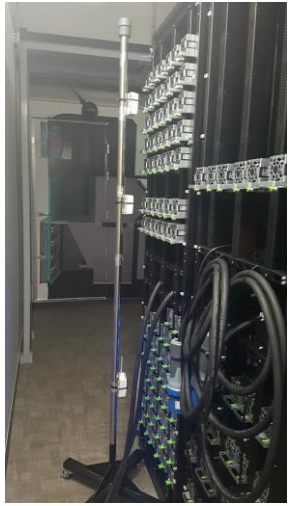
Testbed Design – Installation

CRAC Server Room





Front



Rear

Test Details

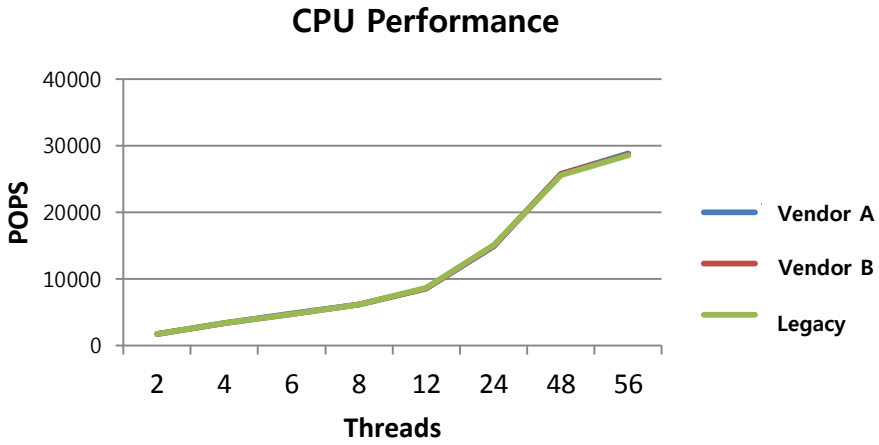
Category	Item	Details			
Performance (Testbed)	CPU	Purpose	To compare CPU performance of each compute node		
		Summary	Measurement	Unit	Tool
			POPS (Prime Operations Per Second)	POPS	Sysbench
	Memory	Purpose	To compare memory performance of each compute node		
		Summary	Measurement	Unit	Tool
			Throughput	MB/s	STREAM
Efficiency (Testbed)	Power Consumption	Purpose	Power consumption comparison between legacy and OCP server under different room temperature(25C~35C) and workloads (Idle~100%)		
		Summary	Measurement	Unit	Tool
			Inlet/Outlet Temp, Power Consumption	Watt, degree C	Stress
Application (Datacenter)	OpenStack / Hadoop / NFV Container	Purpose	To compare application environment between OCP and legacy system		
		Summary	Measurement	Unit	Tool
			Feedbacks from operation and development	N/A	Various

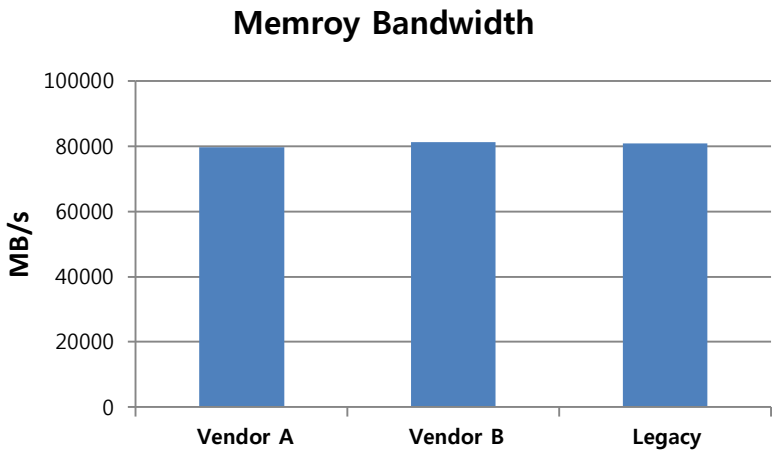
Performance

- No significant difference was found between OCP and legacy system
- Can check configuration differences or problems among systems before measuring power usage

➢ Test Tools

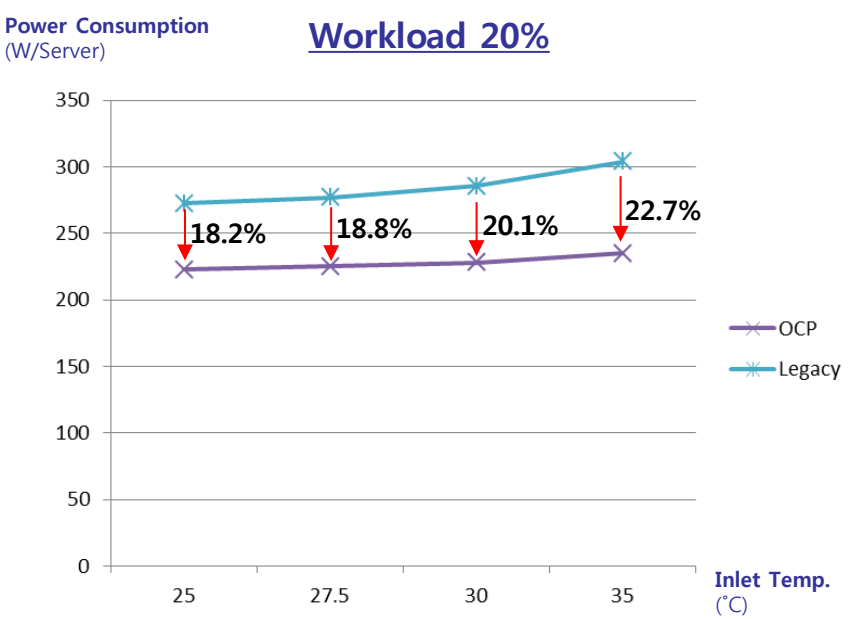
Category	Unit	Tool
CPU	▪ POPS (Prime Operations Per Second)	▪ sysbench
Memory	▪ MB/s	▪ STREAM

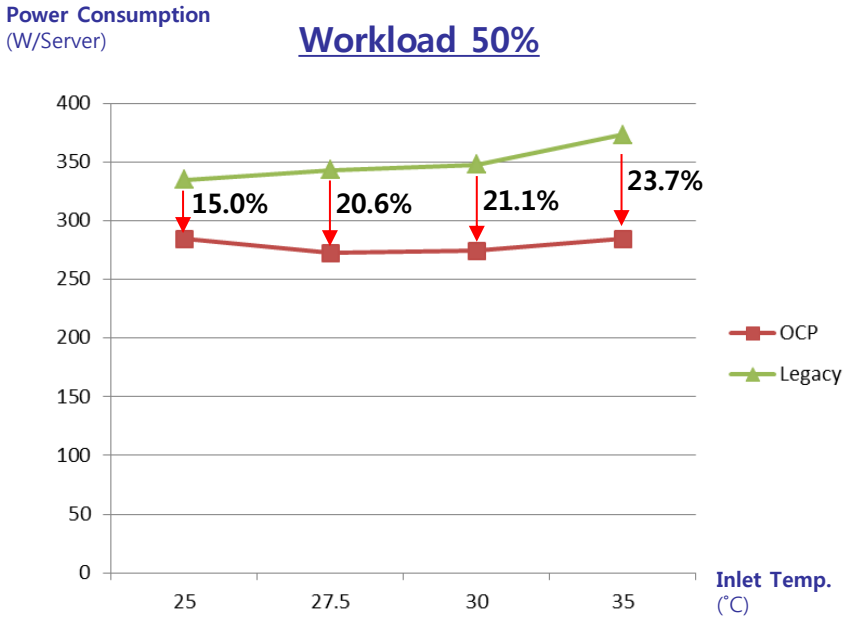




Power Consumption Comparison

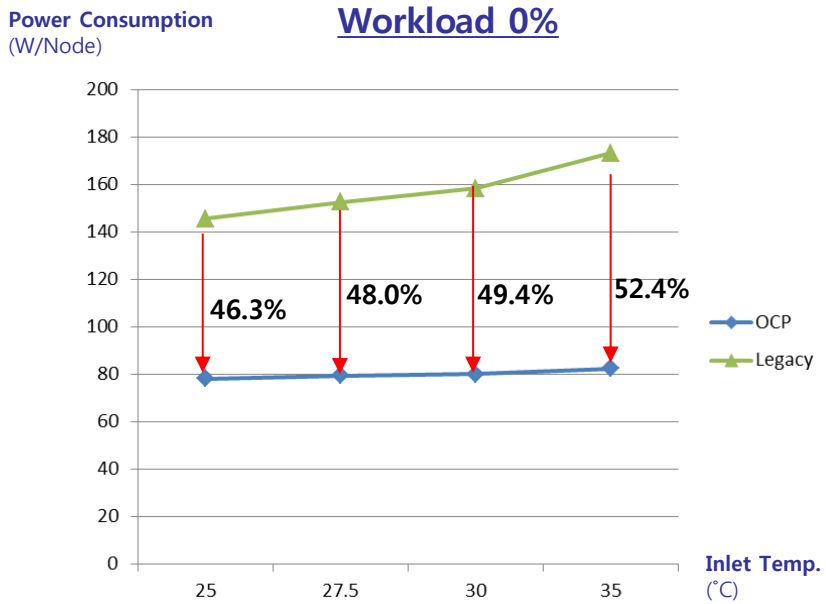
- Compared power consumption of OCP and Legacy system under different room temperature and workload
- As the inlet temperature rises, the efficiency also rises

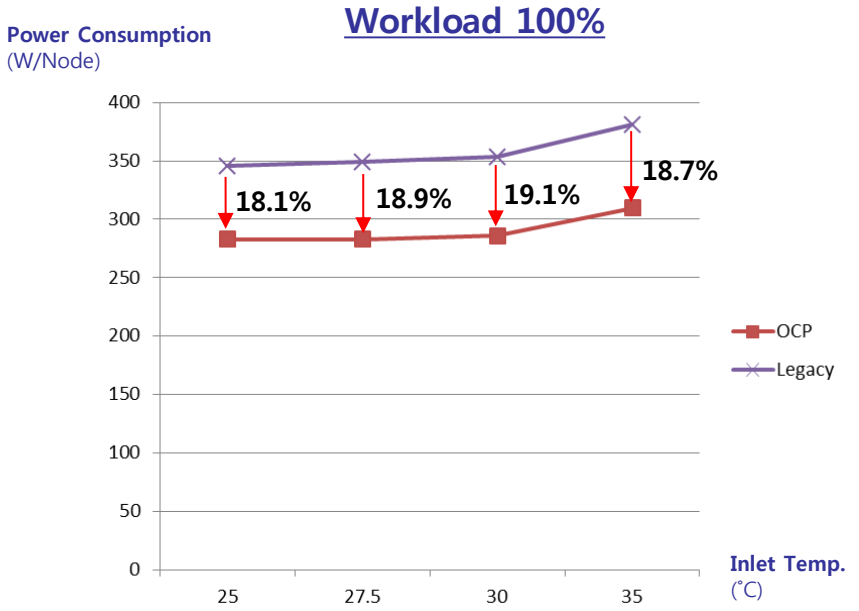




Power Consumption Comparison

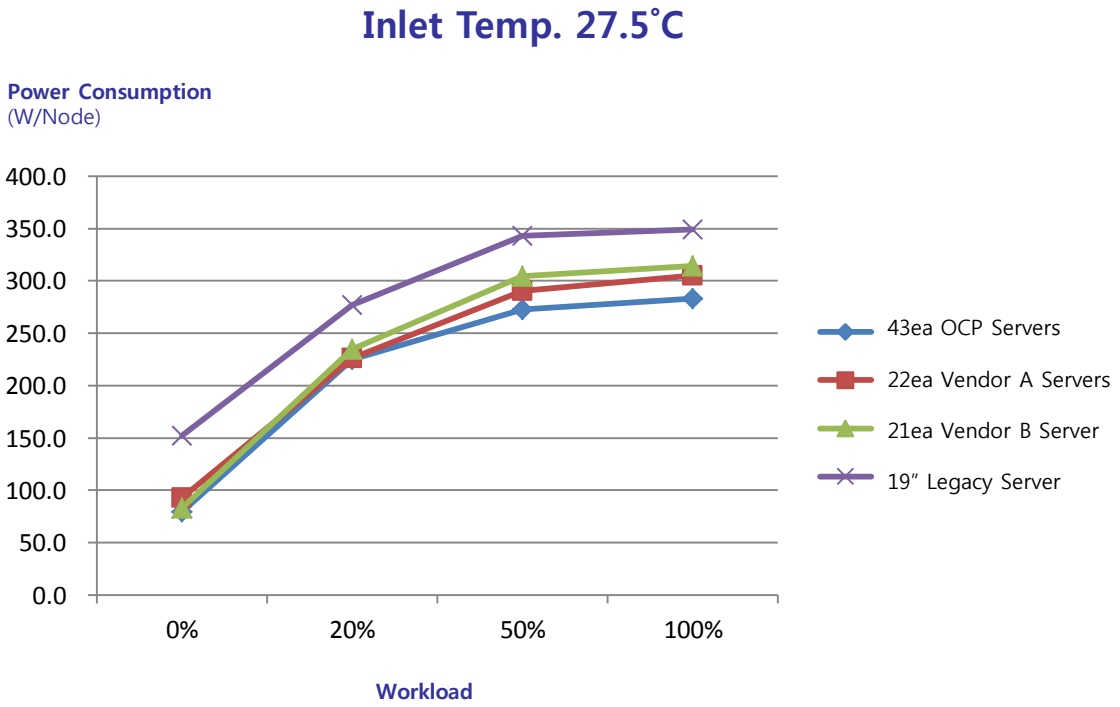
- In idle state, OCP server bettered legacy server under all the room temperatures
- In extreme workload, power consumption difference drops slightly





Power Consumption Comparison

- From idle to full workload, OCP server bettered 19-inch legacy server
- Most power efficiency can be gained through full rack configuration (43 servers)
- Vendor differences in power consumption exists (may need firmware optimization)



Power Efficiency Summary

- On general workload, OCP system is about 20% more power efficient than 19-inch legacy system
- As the room temperature rises, the efficiency of OCP system also rises
- Vendor difference exist in power consumption (about 3%~5%)
  - ✓ Fan & Power Curve tuning may need to optimize
- Should deploy full-rack configuration, to gain most energy efficiency
  - ✓ With full-rack configuration (43 Servers), OCP server saved about 20% compare to 19-inch server
  - ✓ With half-rack configuration (21 Server), the power efficiency decreases by 10%~40%

	Fully Packed (43 Severs)			Half Packed (21 Server)	
Utilization	Legacy	OCP	Power Saving (OCP/Legacy)	OCP	Power Saving (OCP/Legacy)
0%	152.5W	79.3W	48%	88.0W	42%
20%	276.9W	225.0W	19%	230.6W	17%
50%	343.1W	272.5W	21%	297.4W	13%
100%	348.9W	282.9W	19%	309.7W	11%

※ Inlet Temp: 27.5℃

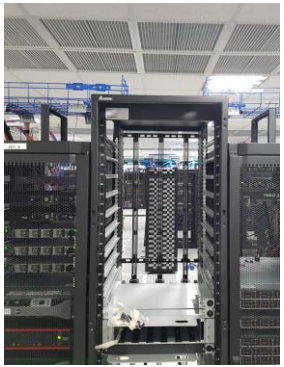
Datacenter Deployment

- Moved OCP gear from Testbed to IDC to test service applications
- Among the three candidates, the oldest datacenter were excluded because of rack size problem

	Datacenter A (Excluded)	Datacenter B (OpenStack/R&D)	Datacenter C (Hadoop)
Description	• The oldest building built as a central office in early 90's • Reconstructed office building to a server room • Small freight elevator • Height of server room entrance is lower than Open Rack	• Colocation center • The newest building • Height of server room entrance is lower than rack	• Constructed as a central office in mid 90's • Height of server room entrance is lower than Open Rack
Rack Power	• 220V Single Phase (6.6kw)	• Provides various rack power options (220V/380V, 6.6/13.2kw etc.)	• 220V Single Phase (6.6kw)
Operation Temp.	~25℃	~25℃	~25℃

Deployment Issues

- Removed network tray



- Disassemble containment

- Negative pressure

- Server room entrance

- Small freight elevator

- Incompatible rack shelf

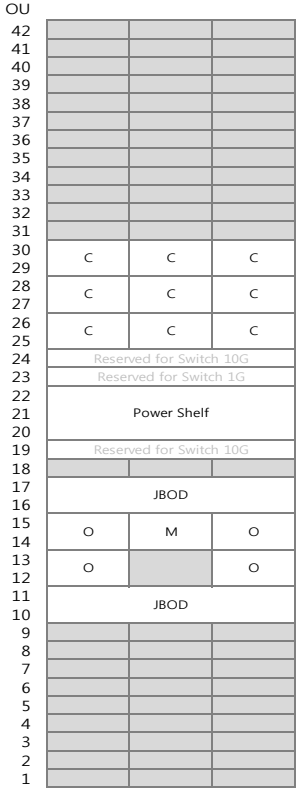
Vendor A

Vendor B

Application Test (ongoing)

OpenStack Rack

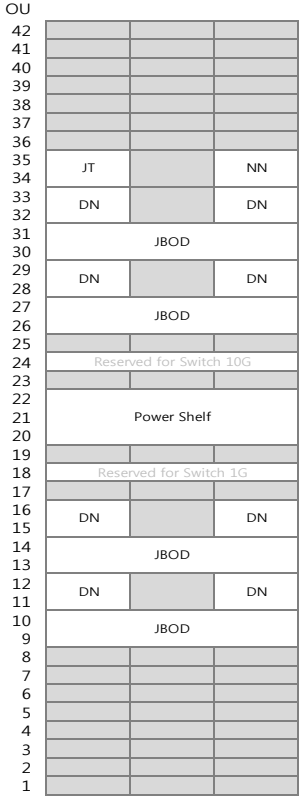
Datacenter B



- Open Rack (21")
  Power 12kW
- Private Cloud
- 9 Compute Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  *SSD 300GB*2ea*
- 5 Ceph Nodes
  *E5-2690 V4*2ea*
  Memory 256GB
  M.2 300GB
  *SSD 1T*6ea*

Hadoop Rack

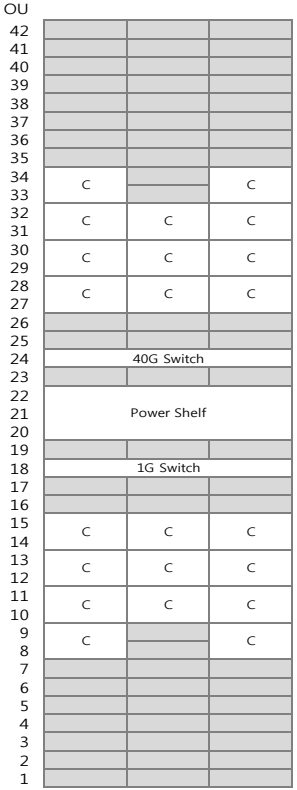
Datacenter C



- Big Data Analytics
- Open Rack (21")
  Power 12kW
- 10 Hadoop Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  *SSD 450GB*2ea*
- 4 JBODs
  *HDD 4TB*12ea*

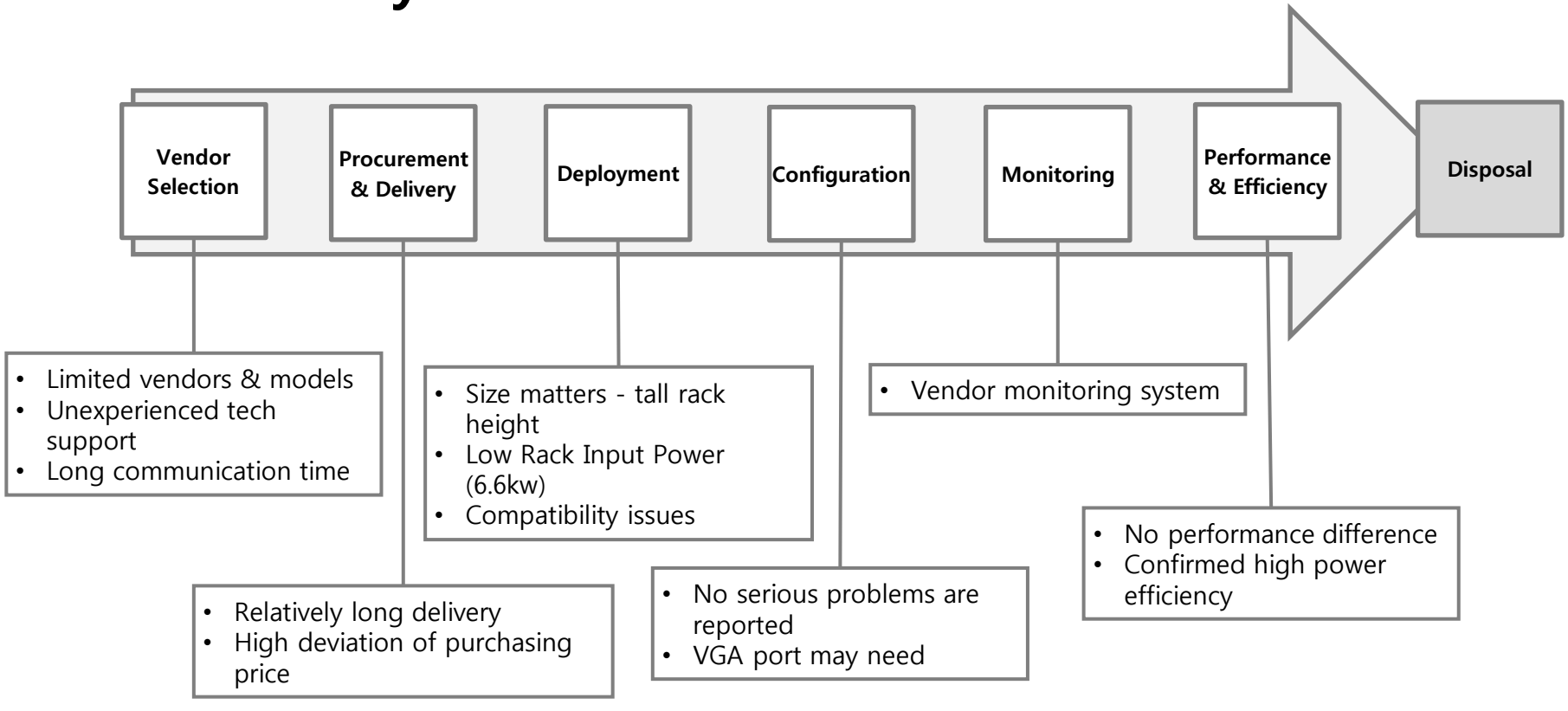
R&D Rack

Datacenter B

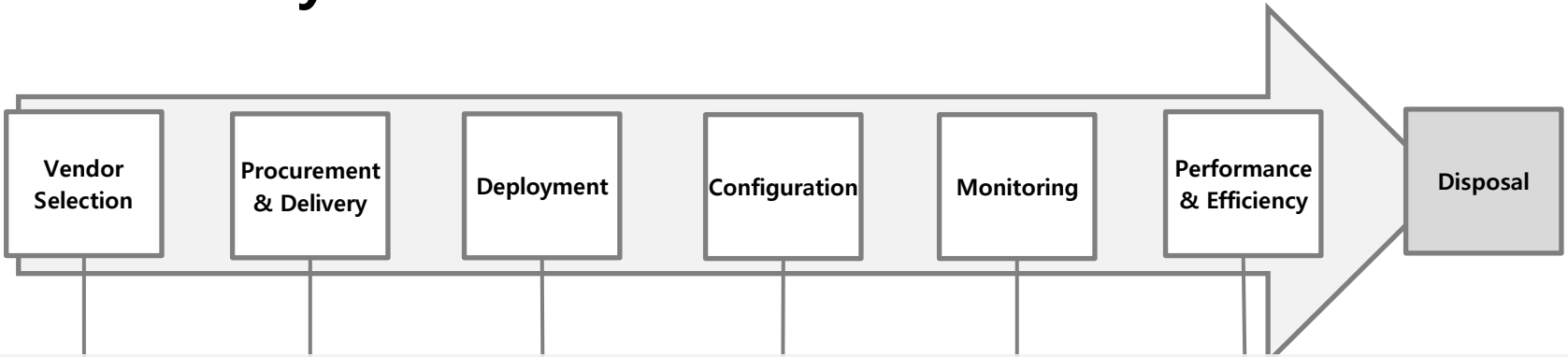


- OpenStack/SDN, NFV Container
- Open Rack (21")
  Power 12kW
- 22 Compute Nodes
  *E5-2660 V4*2ea*
  Memory 256GB
  M.2 300GB

Trial Summary



Trial Summary



- Limited vendors & models
- Unexperienced tech support
- Long communication time

- Size matters - tall rack height
- Low Rack Input Power (6.6kw)
- Compatibility issues

- Unified monitoring system needed

- No performance difference
- Cont… efficiency

- …delivery
- High deviation of purchasing price

- No serious problems are reported
- VGA port may need

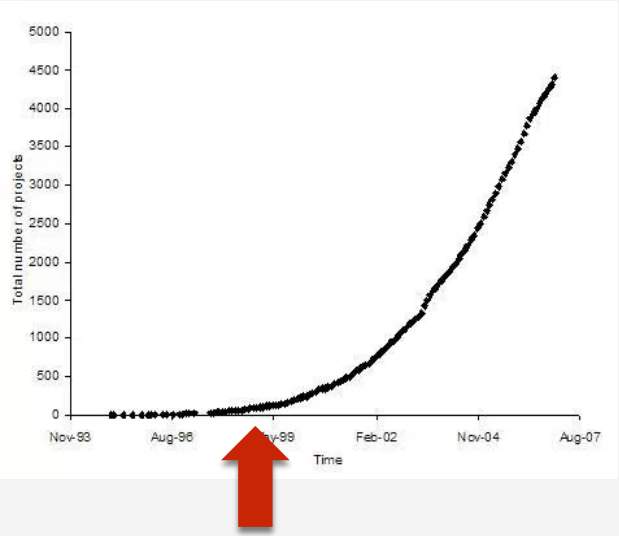
Open Source Hardware Ecosystem Growth

Organization Consensus

Standardization

Our Approach

Graph of total number of open source software projects



We are somewhere around here

Participate

OPEN GOLD

Collaborate

OPEN Compute Project

OCP Telco Project

Contribute

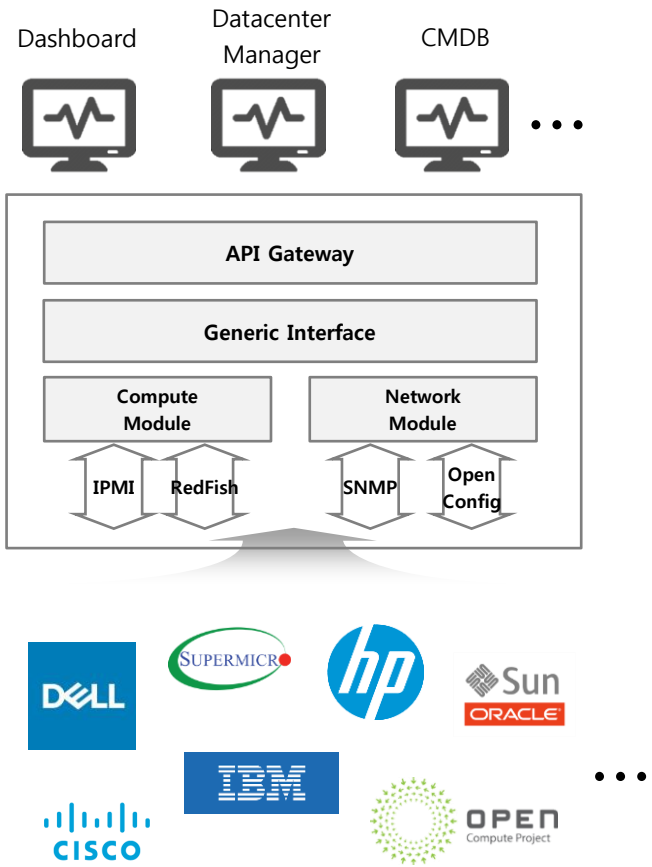
Source: http://dirkriehle.com/publications/2008-2/the-total-growth-of-open-source/

Next Step - Hardware Management System

Due to the proliferation of new platforms, like AI and Container clusters, the management of hardware information becomes important

	【 AS-IS 】		【 HW Mgmt. System 】
Collection Data	• **OS Level Data Collect** - hostname - serial number - OS version - CPU info - VM info	**Extend Mgmt. Range and Level**	• **HW Level Data Collect** - Chassis/Board info - BMC info - Temperature - Power - Fan
Path	• **In-Band**		• **In-Band & Out-Of-Band**
Method	• **Agent**		• **IPMI/RedFish/SNMP**
Change Mgmt.	• **Manual**		• **Manual + Automatic**

Next Step - Hardware Management System



- Provides standard API
- Supports various protocols (IPMI, RedFish, SNMP …) and provides HW abstraction layer

Call for Collaboration

- ✓ To define standard API
- ✓ To develop hardware abstraction layer
- ✓ To reflect telco requirements

Interested Parties

OCP HW Management Project,

DMTF (RedFish), HW/Chip Vendors

Thank you

email: jungsoo.kim@sk.com

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

OPEN Compute Project


OPEN
Compute Project