OCP Trial Results for Telco Infrastructure

Jungsoo Kim/Manager/SKTelecom
Agenda

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

2G & 3G
- 3G Data
- 3G Voice
- 2G/3G Voice
- WCDMA
- CDMA
- ATM
- PSTN

Mobile Network

Data Center
- Siloed Data Center

4G
- LTE EPC
- LTE
- IP
- 3G Voice
- IP

Converged IT + NW
- Media
- IoT
- Access
- Core
- Internet

5G = All IT Network
- COSMOS
- Composable
- Open
- Scalable
COSMOS

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

Applications

Cloud Services

Open Software
(Virtual Infrastructure)

Unbundled

Open Hardware
(Physical Infrastructure)

T-CORE (Unified Operation of Data Center)

Control Plane Orchestration for Data Center Resources (Kubernetes)

Hybrid (Virtual & Physical) Resources

OH (SNS, Media, etc.)

Telco (LTE, 5G, LoRa, etc.)

Enterprise IT (BSS, ERP, etc.)

Private Cloud (IaaS)

DevOps Platform (PaaS)

Hybrid Resource Mgmt.

Operation Automation

Monitoring & Analytics

Open Source Hardware + Performance Optimization

OCP Server

OCP Switch

T-CAP

NV-Array

AF-Media
Open Hardware/Software in SK Telecom

Gold Member
OCP Telco Project

Server  Switch  Storage

Founding Member
Board Chair

RAN  Backhaul  Core

Community Lab  TEAC

*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

- Verify OCP Technology
  - Efficiency, Performance, Serviceability

- Analysis Deployment Environment
  - Procurement, IDC Environment, Specification ...

- Check OCP Ecosystem in Local
  - Vendor, Tech Support, Delivery ...
Understanding of environmental difference is important key to success

<table>
<thead>
<tr>
<th>Characteristic Differences</th>
<th>Hyper-Scale</th>
<th>SK Telecom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>&gt; Hundreds of Thousands</td>
<td>&lt; Tens of Thousands</td>
</tr>
<tr>
<td>Age of Datacenter</td>
<td>Young</td>
<td>Old</td>
</tr>
<tr>
<td>Datacenter</td>
<td>A Few Big IDCs</td>
<td>Small Distributed IDCs</td>
</tr>
<tr>
<td>Infra Management</td>
<td>Centralized</td>
<td>Separated</td>
</tr>
<tr>
<td>Deployment</td>
<td>Rack Scale</td>
<td>System Scale</td>
</tr>
<tr>
<td>Major Supplier</td>
<td>ODM Vendor</td>
<td>Brand Vendor</td>
</tr>
</tbody>
</table>
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
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
CRAC Testbed

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

*CRAC: Computer Room Air Conditioner
Testbed Design – Architecture

Open Rack

19-Inch Rack

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

Client (Remote Access)

1G Mgmt. Switch

10G SFP+ Switch

10G UTP Switch

10.0.1.1

10.0.2.10/24

10.0.3.10/24

10.0.1.1

10.0.2.10/24

10.0.3.10/24

Legacy Servers
Testbed Design – Top View

Cold Aisle Containment (Raised Floor)

- Rack Power Monitor
- Temperature/Humidity Probe
- Inlet Temperature Sensor
- Outlet Temperature Sensor
- Server Heat Emulator
- Power Monitor (YOKOGAWA CW500)
Testbed Design – Installation

CRAC Server Room

- Legacy Rack
- Cold Aisle Containment
- OCP Rack
- Rear Sensors

Front

Rear
## Test Details

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

<table>
<thead>
<tr>
<th>Category</th>
<th>Unit</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>POPS (Prime Operations Per Second)</td>
<td>sysbench</td>
</tr>
<tr>
<td>Memory</td>
<td>MB/s</td>
<td>STREAM</td>
</tr>
</tbody>
</table>

**CPU Performance**

**Memroy Bandwidth**
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**

<table>
<thead>
<tr>
<th>Power Consumption (W/Node)</th>
<th>Inlet Temp. (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Workload 0%</strong></td>
<td></td>
</tr>
<tr>
<td>OCP</td>
<td>25</td>
</tr>
<tr>
<td>Legacy</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
</tbody>
</table>

- 46.3%                      
- 48.0%                      
- 49.4%                      
- 52.4%                      

**Workload 100%**

<table>
<thead>
<tr>
<th>Power Consumption (W/Node)</th>
<th>Inlet Temp. (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Workload 100%</strong></td>
<td></td>
</tr>
<tr>
<td>OCP</td>
<td>25</td>
</tr>
<tr>
<td>Legacy</td>
<td>27.5</td>
</tr>
<tr>
<td></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
</tbody>
</table>

- 18.1%                      
- 18.9%                      
- 19.1%                      
- 18.7%                      

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%

<table>
<thead>
<tr>
<th>Utilization</th>
<th>Legacy</th>
<th>OCP</th>
<th>Power Saving (OCP/Legacy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>152.5W</td>
<td>79.3W</td>
<td>48%</td>
</tr>
<tr>
<td>20%</td>
<td>276.9W</td>
<td>225.0W</td>
<td>19%</td>
</tr>
<tr>
<td>50%</td>
<td>343.1W</td>
<td>272.5W</td>
<td>21%</td>
</tr>
<tr>
<td>100%</td>
<td>348.9W</td>
<td>282.9W</td>
<td>19%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>OCP</th>
<th>Power Saving (OCP/Legacy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fully Packed</td>
<td>88.0W</td>
<td>42%</td>
</tr>
<tr>
<td>Half Packed</td>
<td>230.6W</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>297.4W</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>309.7W</td>
<td>11%</td>
</tr>
</tbody>
</table>

※ Inlet Temp: 27.5℃
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

<table>
<thead>
<tr>
<th>Description</th>
<th>Datacenter A (Excluded)</th>
<th>Datacenter B (OpenStack/R&amp;D)</th>
<th>Datacenter C (Hadoop)</th>
</tr>
</thead>
</table>
| 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
### 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

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

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

Deployment
- Relatively long delivery
- High deviation of purchasing price

Configuration
- No serious problems are reported
- VGA port may need

Monitoring
- Vendor monitoring system

Performance & Efficiency
- No performance difference
- Confirmed high power efficiency

Disposal
Trial Summary

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

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

- Deployment
  - High deviation of purchasing price

- Configuration
  - VGA port may need

- Monitoring
  - No serious problems are reported
  - VGA port may need

- Performance & Efficiency
  - Unified monitoring system needed

- Disposal

Open Source Hardware Ecosystem Growth

Organization Consensus

Standardization

- No performance difference
- Confirmed high power efficiency
Our Approach

Graph of total number of open source software projects

We are somewhere around here

Participate

Collaborate

Contribute

**Next Step - Hardware Management System**

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

<table>
<thead>
<tr>
<th><strong>Collection Data</strong></th>
<th><strong>AS-IS</strong></th>
<th><strong>HW Mgmt. System</strong></th>
</tr>
</thead>
</table>
| **Path**            | - OS Level Data Collect  
  - hostname  
  - serial number  
  - OS version  
  - CPU info  
  - VM info | - HW Level Data Collect  
  - Chassis/Board info  
  - BMC info  
  - Temperature  
  - Power  
  - Fan |
| **Method**          | - In-Band | - In-Band & Out-Of-Band |
| **Change Mgmt.**    | - Agent  | - IPMI/RedFish/SNMP |

Extend Mgmt. Range and Level
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