

Use Case of OpenStack 50000 VM in Yahoo! JAPAN

2015/8/17

Yahoo! JAPAN
Sr. Manager, Infrastructure Engineering Dept.
Takuya Ito

takitou@yahoo-corp.jp



- 1. About Yahoo! JAPAN
- 2. Operation statistics of OpenStack
- 3. Practical uses and operations of OpenStack
- 4. Why we use OpenStack?

Introduction of the statistic data



Introduction of the statistic data

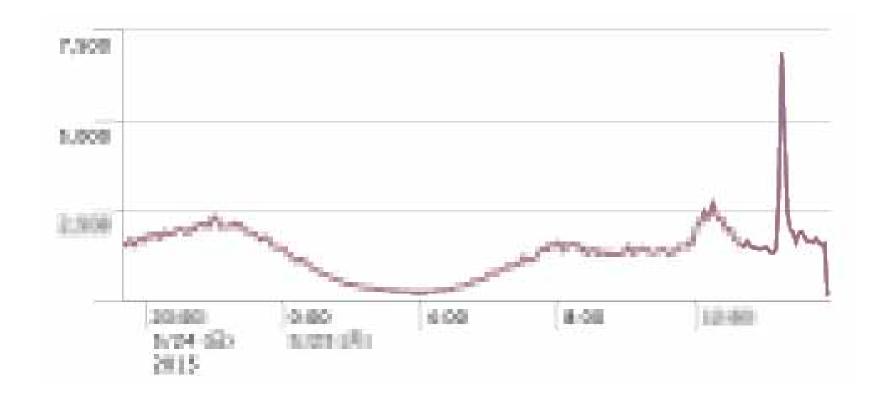


Introduction of the statistic data



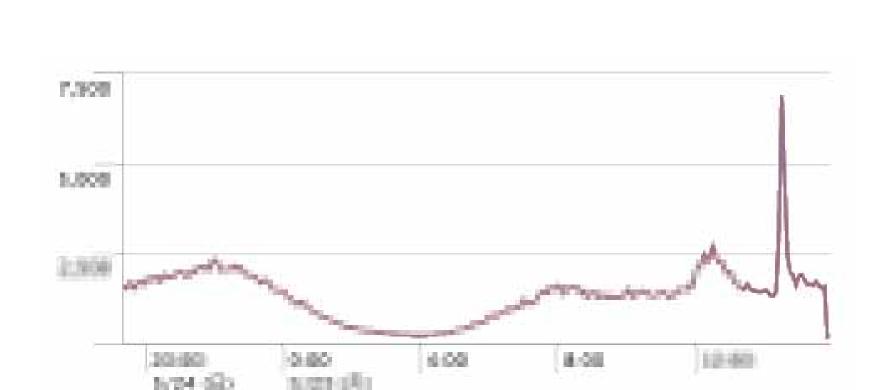






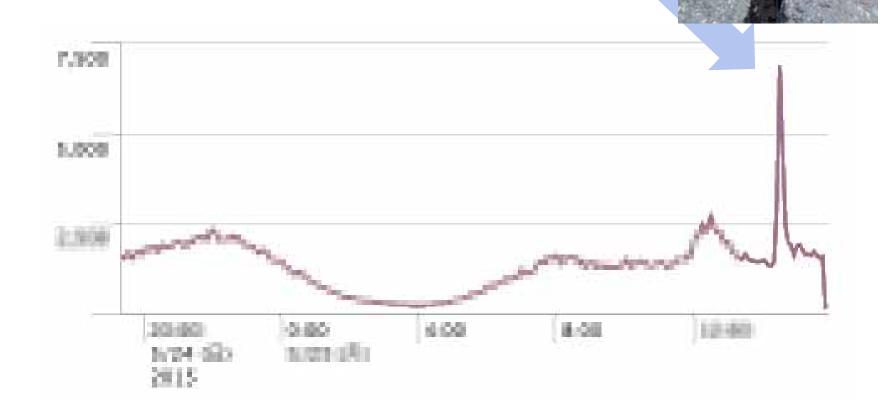
2015

Answer: Data Center Utilization

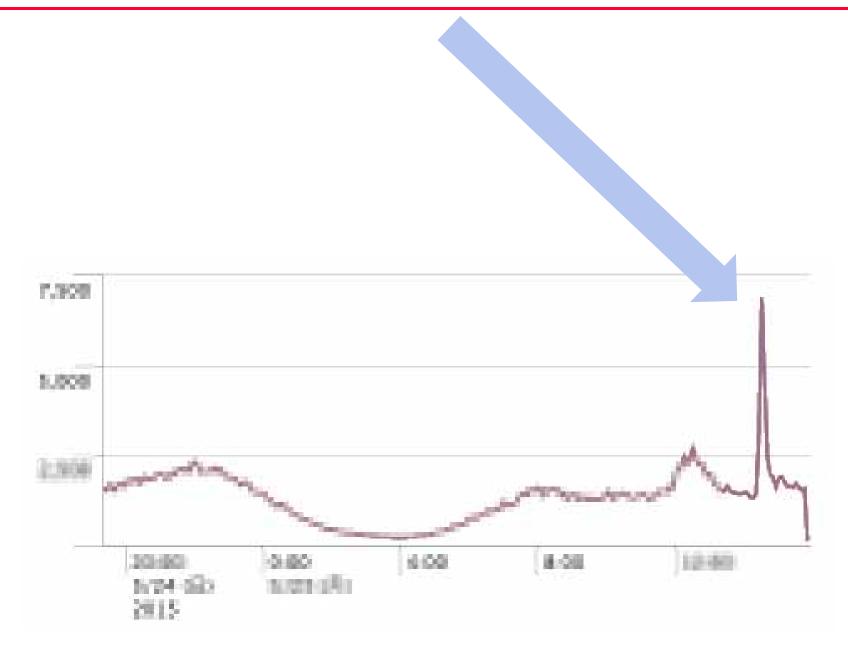




We have given only few ten seconds till the traffic comes to the peak.







- Urgent notification for disaster prevention
- Weather report

Examples of emergency notification

Earthquakes, Tsunami, Volcano, Heatstroke, Rain storms



We use the applications that cannot be stopped on OpenStack.



- Rapid provisioning of the resources are required.
- Works properly in case of an emergency.
- It is important to have the same API in any operating environments whether it s KVM, VMware, or Container.

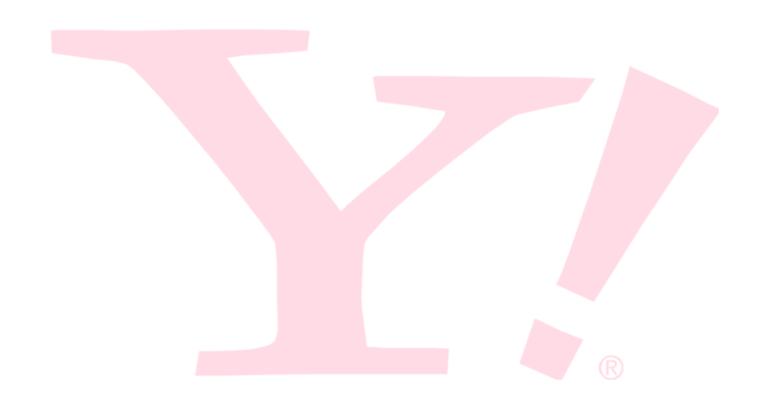
Our mission is abstraction of the data center and OpenStack is in the core.

Things that we want to in the beginning

In Yahoo! JAPAN, OpenStack is operated as an infrastructure to support the people.

- One of the good thing of the OSS is that everyone s effort creates various opportunities.
- Activities in the OpenStack community supports important applications.

Thank you so much to everyone of OpenStacker.





Use Case of OpenStack 50000 VM in Yahoo! JAPAN

2015/8/17

Yahoo! JAPAN
Sr. Manager, Infrastructure Engineering Dept.
Takuya Ito

takitou@yahoo-corp.jp





- About Yahoo! JAPAN
- 2. Operation statistics of OpenStack
- 3. Practical uses and operations of OpenStack
- 4. Why we use OpenStack?

Operation statistics of OpenStack



- Running 50000+ instances
- Availability 99.996%
- 6 times more traffic density
 (Compared to the physical environment)
- 20PB in the data storage
- 20 clusters operating
- 6 developers, 4 operators

1 year ago

Instance

25k+



50k+

Today



Machine

2k+

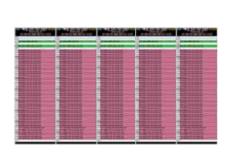


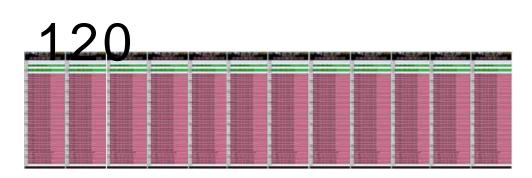
4k+



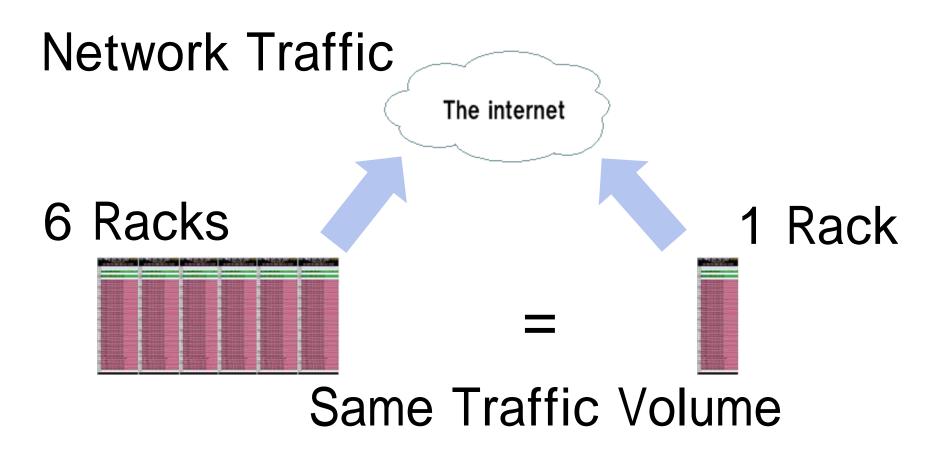
Rack

50





Different Environment OpenStack Environment



2 years ago

Today

Instance Density 400/Rack



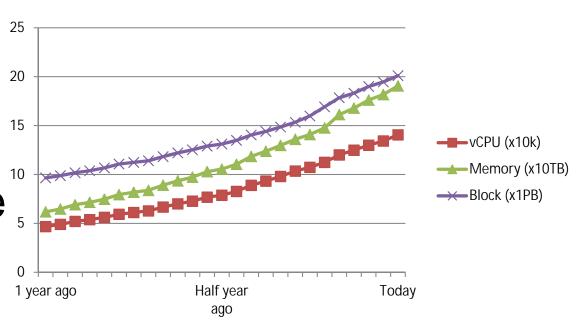
2000 / Rack

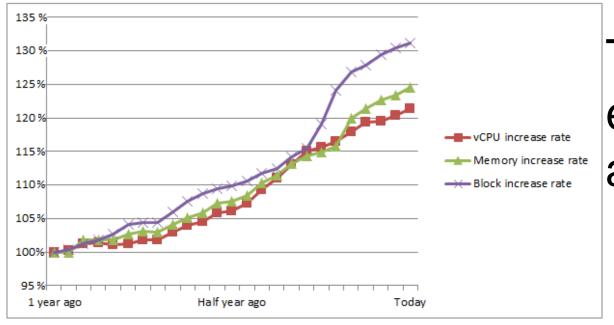






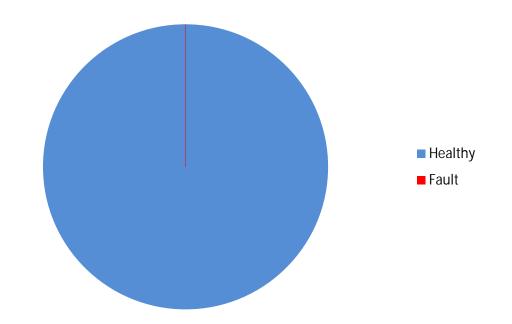
150k Cores200TB Memory20PB Data Storage





The increase rate of each resources for an instance.

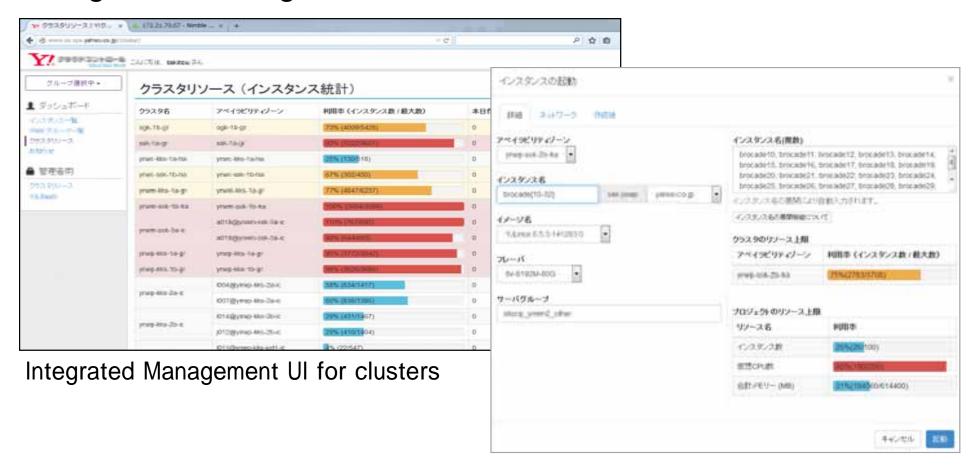
 We achieved 99.996% availability for the total operating time of all hypervisors in Yahoo! JAPAN.



Made a pie chart, but "failure rate" is not visible...:)

- About Yahoo! JAPAN
- 2. Operation statistics of OpenStack
- 3. Practical uses and operations of OpenStack
- 4. Why we use OpenStack?

Since there are more than 20 clusters, we provide integrated management UI.



- Most of the data center resources can be self served by OpenStack.
 - The developers can make their services by their wish.

- The devenlopers can use it any time.
 - No waiting time like in physical environments.

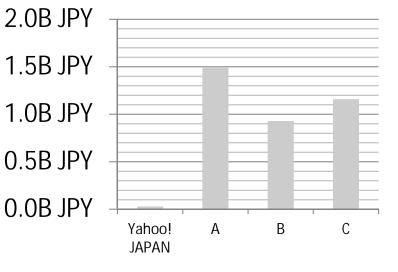


- We provide the resources to 2000 people equally.
 - OpenStack support team responses to the incidents immediately.
 - Recently, non-engineers started to use them.
- 500 new instances created each day.
 - Half of them are removed in a month.
 - Created automatically without engineer s operation.

Low cost

 Reduced by 97% if we use the same resources from the public cloud.

Monthly fee





- Bill when used
 - Billed based on the time that used.

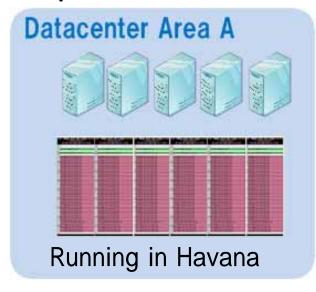
- Salvage the unused instances
 - Show the instances that uses less CPU, I/O, and network to the user and salvage them if it is OK.

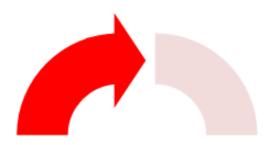
We must avoid wasting resource.

Internal users should be cost conscious.

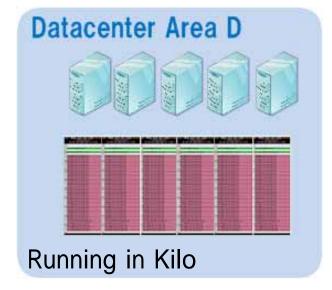
Explain the managements of the data center using OpenStack

1. OpenStack clusters operating

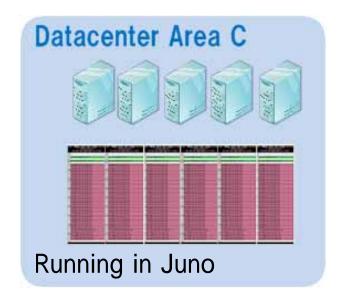




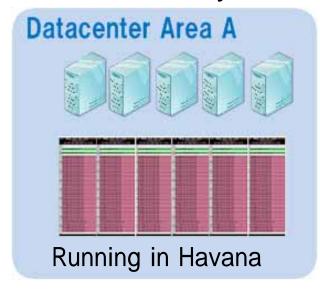


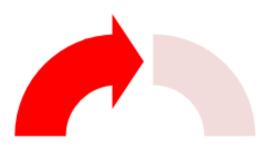




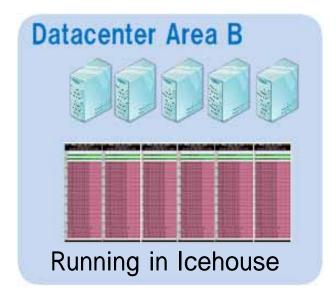


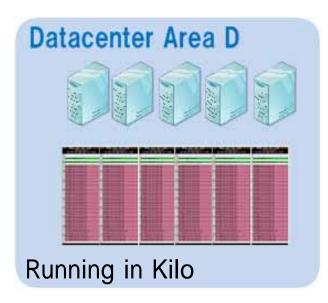
2. When Liberty is released

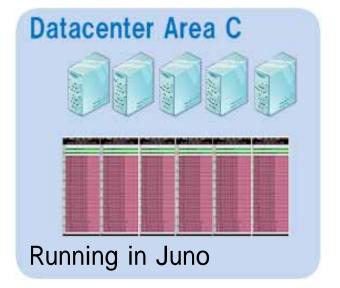








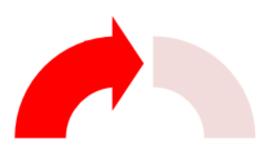


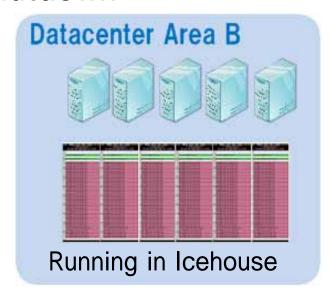


Data Center Life Cycle

3. Instances in the old cluster will be shutdown.

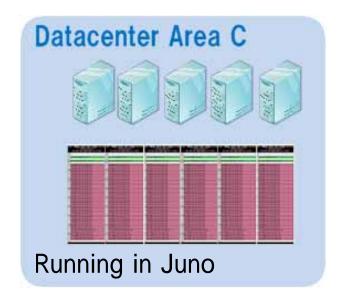




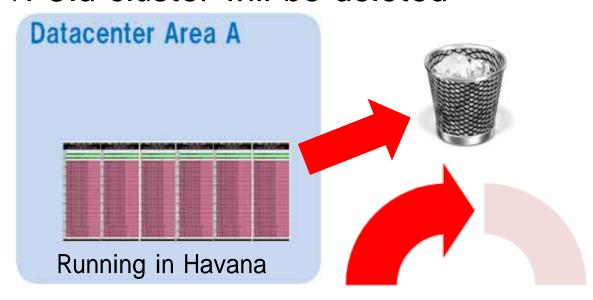


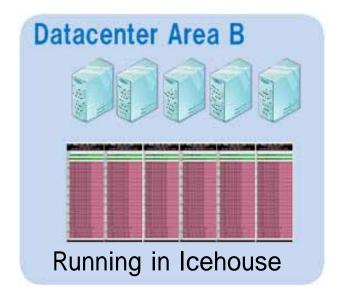






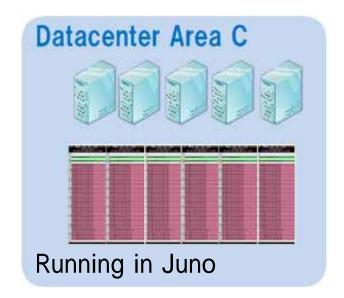
4. Old cluster will be deleted





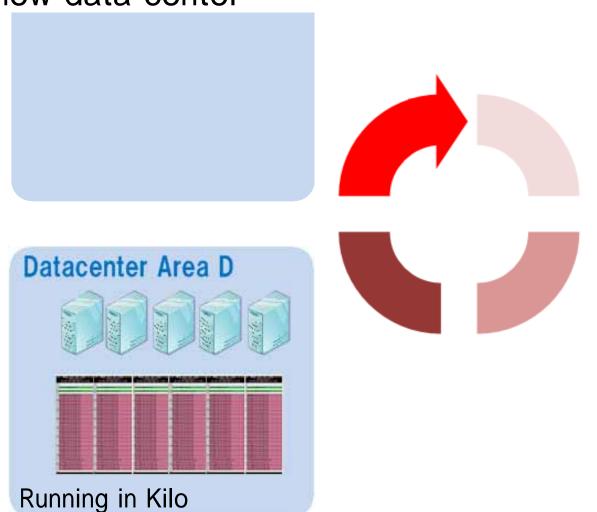


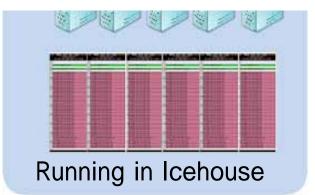


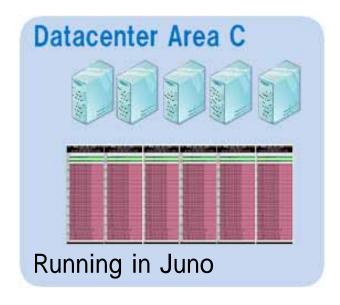


5. The area had been vacant.

If the data center is old, scrap the data center and build the new data center



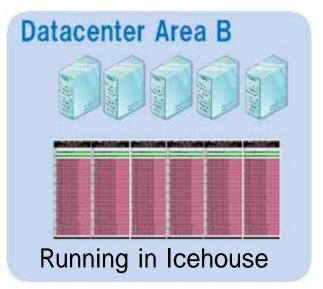


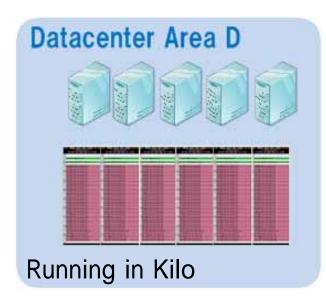


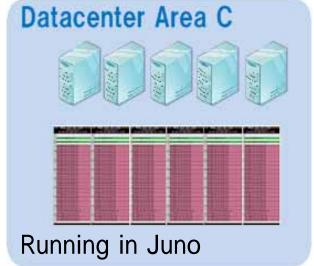
6. Liberty will be created



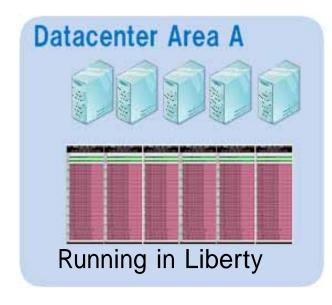




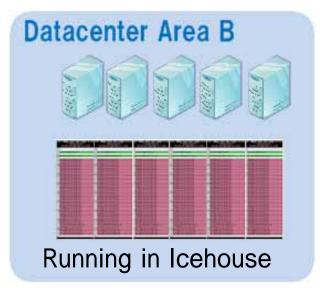




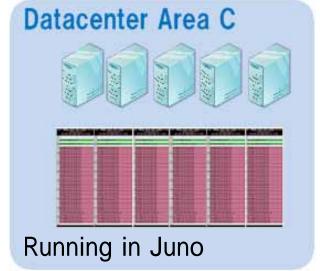
7. Instances will be activated





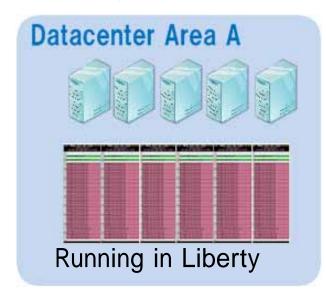




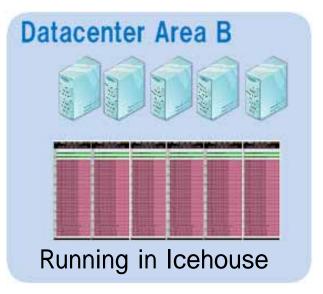


Data Center Life Cycle

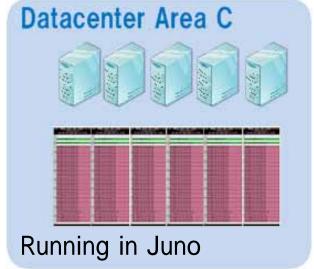
8. Then, M is released



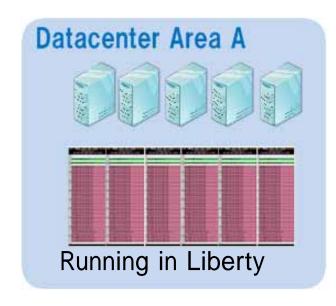


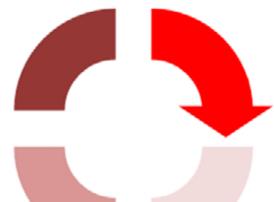






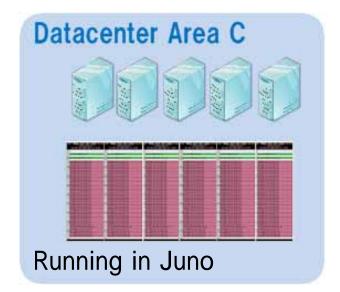
9. Instances are activated in M



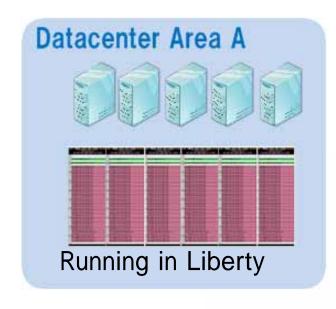




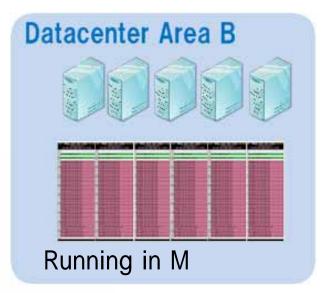




10. N has been released and the instances are activated in N.

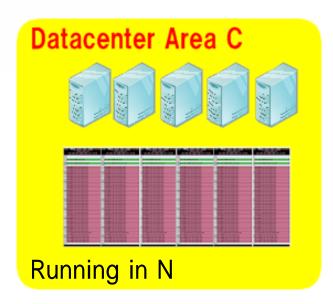




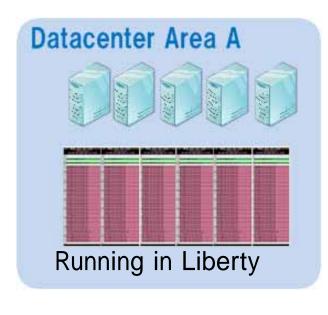




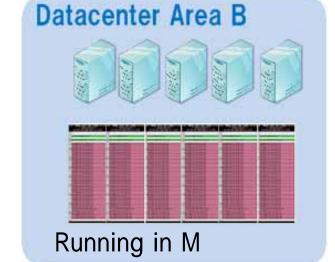


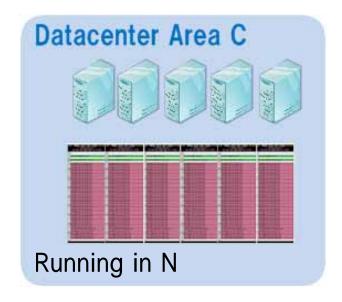


11. 0...

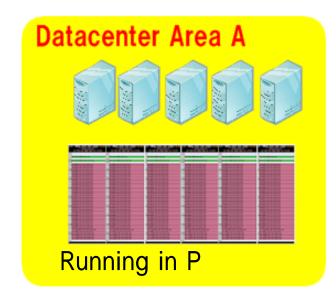


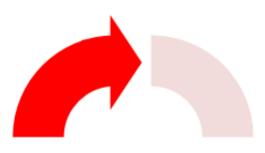


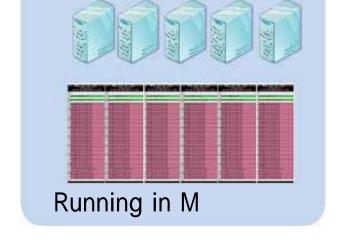




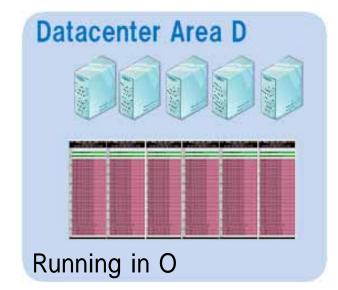
11. P...



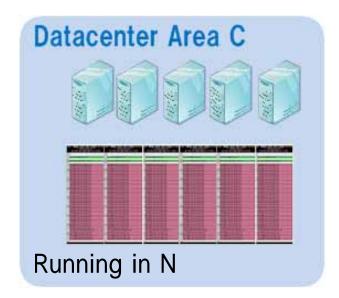




Datacenter Area B

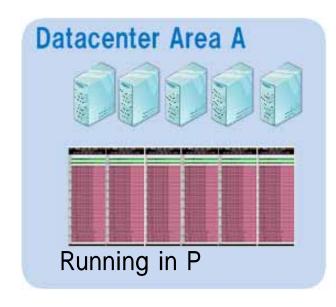


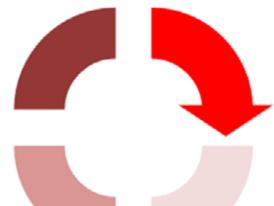




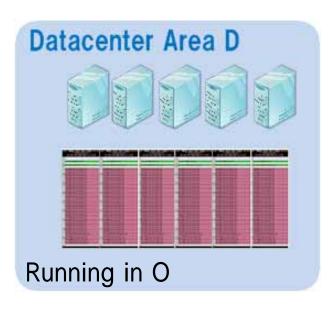
Data Center Life Cycle

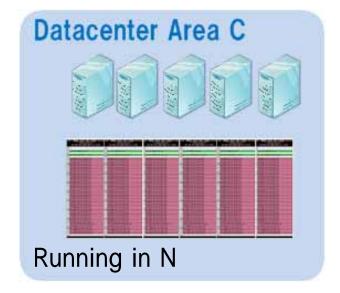
12. Q...





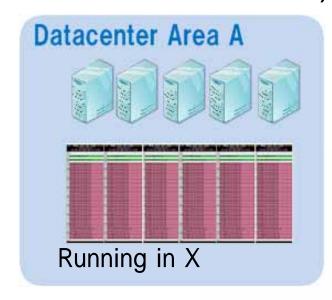






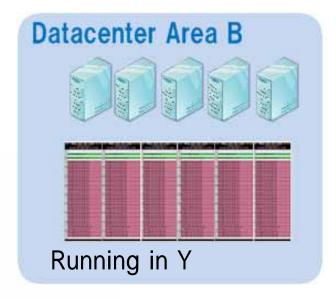
Few years later..

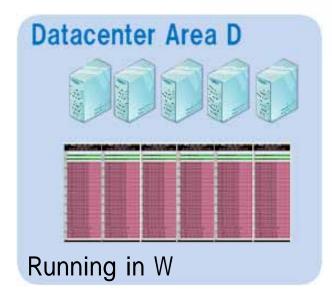
What s next of Z...:)

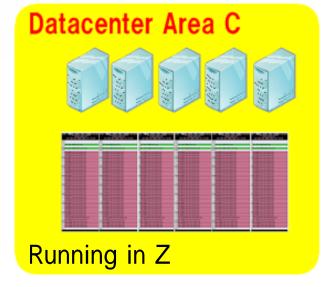












- About Yahoo! JAPAN
- 2. Operation statistics of OpenStack
- 3. Practical uses and operations of OpenStack
- 4. Why we use OpenStack?

- In-house development
 - provide function of laaS
 - Running 10k+ instances
 - unique APIs
 - OSS does not cooperation

Basic Functions of laaS

Minor Functions of vendors

Unique Functions of in-house

In-house development

Use OpenStack (FY2013 or later)



- Common APIs are available
- Basic functions are those community will provide
- Appliance functions are those we will develop with vendors
- In-house functions are those in-house will develop.

Basic Functions of laaS

Minor Functions of vendors

Unique Functions of in-house

Growth with communities

Development with vendors

In-house development



- Data Center Life Cycle Management
 - We want to use the evolving data center
- Data Center Abstraction (Hardware Abstraction)
 - We can use the same APIs in all environment
 - Users are not aware of the physical differences
 - Good things can be using in good timing
- Reduce the cost
 - It is necessary that user should be aware of the cost

We believe Co-Creation is important

 We have jointly developed with the vendors, return it to the community.

