



# 全球运维大会

2016

DevOps 2.0: 重塑运维价值



北京站

会议时间：12月16日 - 12月17日

会议地点：北京国际会议中心

主办单位：





# 去哪儿私有云建设和自动化

叶璐 Qunar DevOPS



# 提纲

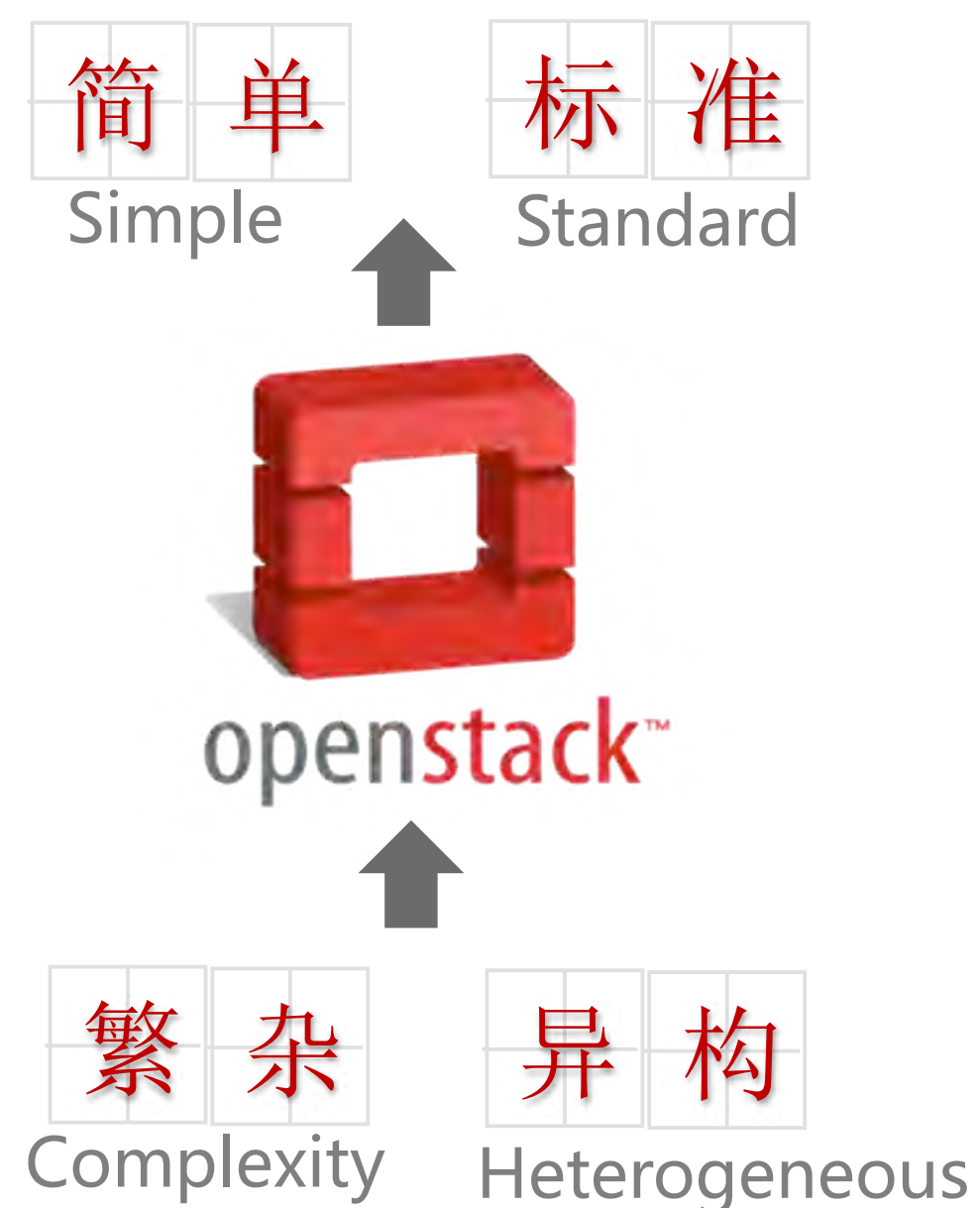
- 架构
- 部署&维护
- 性能调优&排错
- 日志与监控
- 硬件自动化运维
- 周边应用





# OpenStack的意义

Why OpenStack



OpenStack是化“繁”为“简”的框架，也是化“异构”为“标准”的框架  
OpenStack is the framework to transform complexity to simple and heterogeneous resources to standard services



# 现状

- 400+应用
- 3万+虚拟机
- 7+region统一管理
- 支撑容器，大数据，机器学习，支付等各类生产环境应用
- 自动集成测试平台





# Why OpenStack ?

- kvm, Qemu, libvirt
- novnc
- iptables
- guestfs
- dnsmasq





## QUNAR CLOUD

CLOUD MANAGEMENT PLATFORM



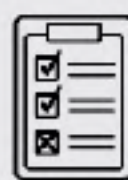
DISCOVERY



CAPACITY  
PLANNING



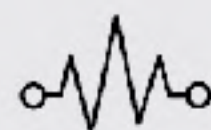
REPORTING



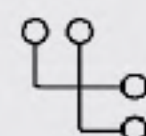
RESOURCE  
APPROVAL



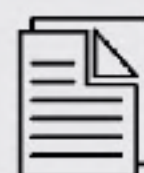
ANALYSIS



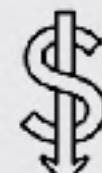
MONITORING



ORCHESTRATION



AUDIT



BILLING

## OPERATIONAL MANAGEMENT

Self-Service Monitoring

Resource Quotas

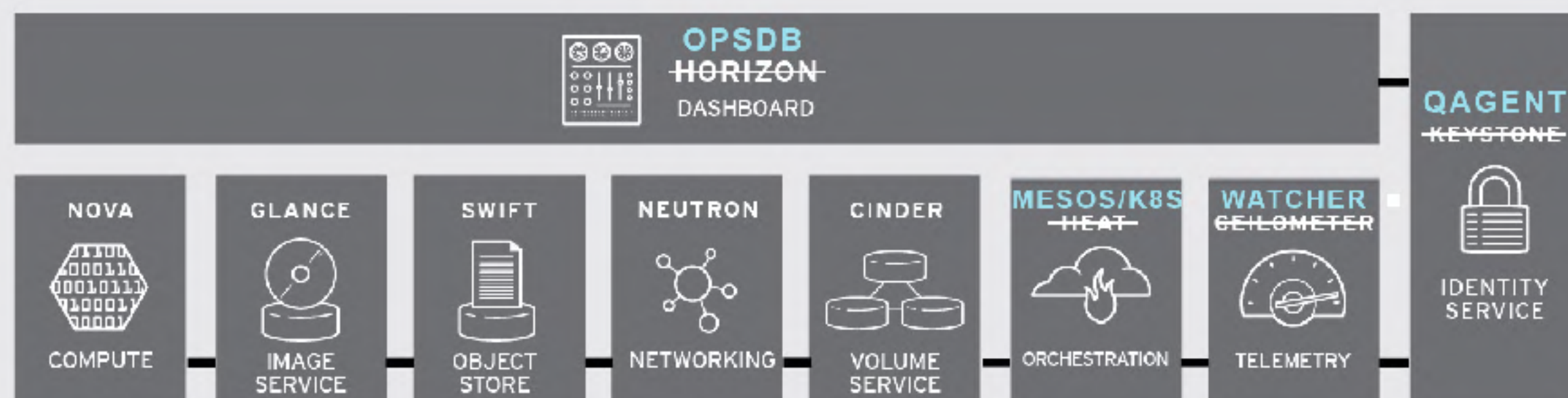
Intelligent workload placement

Unified operations management

## QUNAR

## OPENSTACK PLATFORM

CLOUD INFRASTRUCTURE PLATFORM



## CORE INFRASTRUCTURE

Compute

Storage

Networking

CL0052

04-S10003







# 定制化了什么？

- 全rpm安装，Salt打Patch更新
- OpsDB替换Dashboard
- 应用模型
- 账号体系qagent
- 简单的网络模式:自动DNS, Inject初始化过程
- 监控+日志管理





# OVMS V.S. OpenStack

	OVMS	OpenStack
VM Kernel	<= 2.6.18	ALL
Hypervisor	Xen	KVM、LXC、QEMU、UML、VMWare ESX/ESXi、Xen
迁移方式	手动scp image文件	动态迁移 / 静态迁移
VM Launch Time	5 ~ 30 min	1 ~ 5 min
宿主机选择	基于debian	RedHat、CentOS、Debian、Ubuntu、SUSE, etc.
管理方式	web interface / xm	cli / web ui / api
鉴权	Unix	Role-Based Access Control (RBAC)
镜像/快照管理	N	Y
		



# 部署情况

- 版本：Grizzly/Havana/Mitaka/Newton

- 主控节点：

nova-api

nova-scheduler

nova-conductor

nova-api

keystone

glance

mysql

qpid

- 计算节点

nova-compute

nova-network

nova-metadata-api(for windows hostname)

dnsmasq(for windows dhcp)





# 机器配置

- 主控节点硬件配置

16/24 Core

64GB Ram

600G HDD(4X10k sas Raid1+0)

- 计算节点硬件配置

24/32 Core

128/256GB Ram

2.4TB HDD(8X10k sas Raid1+0)



# 用法

- Single Tenant
- Multiple Vlan
- 配额 = 实际物理资源(memory&disk)
- 资源调度标准

cores=1.5

ram\_allocation\_ratio=1

disk\_allocation\_ratio=0.95

max\_instances\_per\_host=17



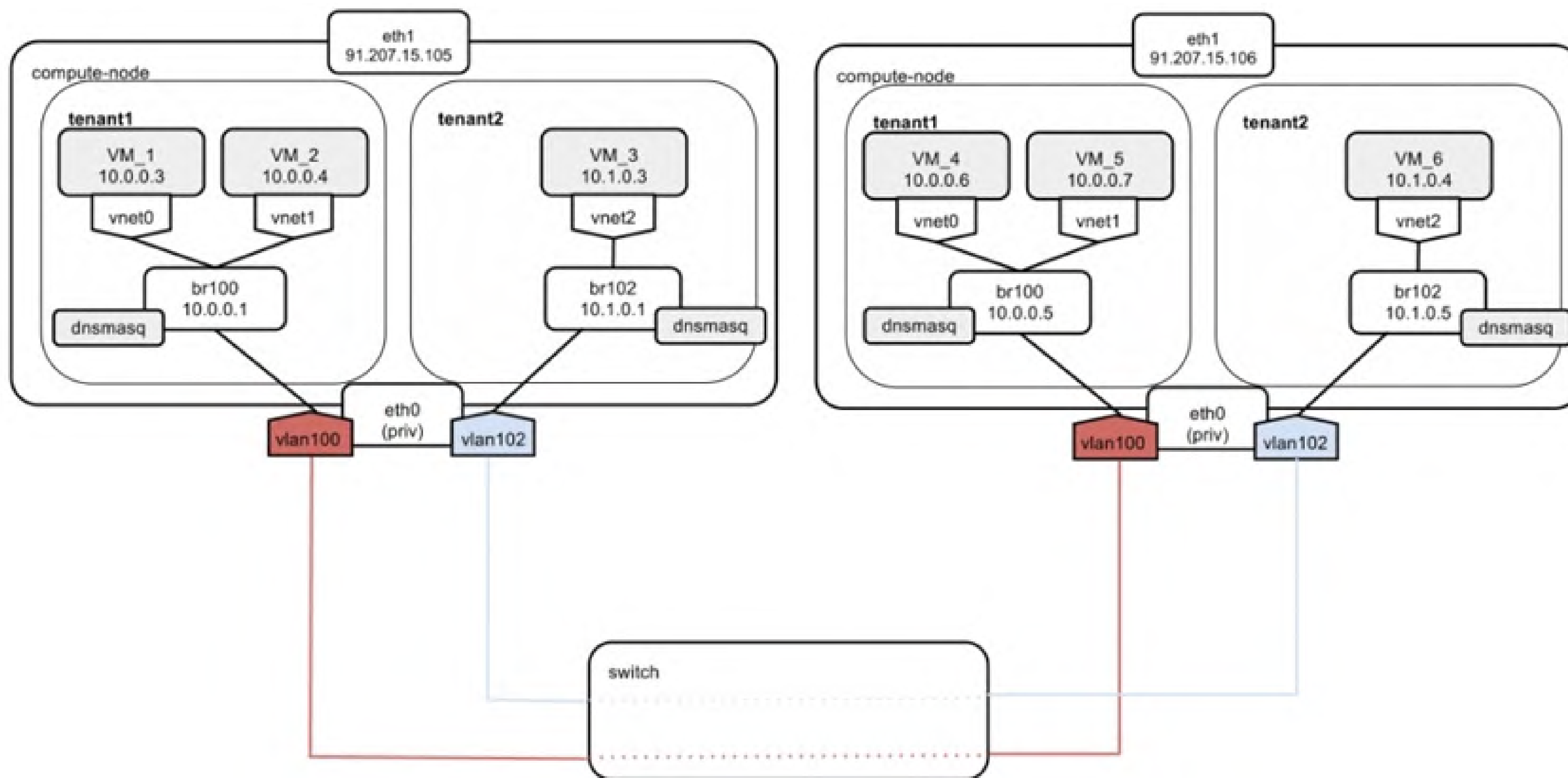


# 网络

- nova.network.manager.VlanManager
- Linux Bridge + Vlan tag
- Hardware Gateway
- 无floating-ip



# Vlan架构





# 网络规则

- Migration —block-migrate
  - 宿主1X1Gbps
  - live\_migration\_bandwidth=50
- share\_dhcp\_address=True
  - dnsmasq的监听地址
  - 数据库nova.networks.gateway字段
- HW Gateway
  - vnet—bridge—vlan—HW GW
  - file injection(for linux)
  - dhcp-option=tag:vlan203,3,10.xx.xx.4(for windows)



# 网络规则

- vm images

libguestfs-tools-c

cloud-utils(resizefs)

- Disable ebtables

虚拟机多IP需求

nova/virt/libvirt/firewall.py:

```
-filter_set = []
```





# 存储

- 镜像存储(glance)

控制节点本地存储

filesystem\_store\_datadir = /var/lib/glance/images

镜像格式：qcow2

- 虚拟机磁盘镜像

计算节点本地存储

镜像格式：qcow2

- 公司内部孵化器：Ceph做共享存储

- 桌面云：Ceph实现个人盘随主机漂移



# 调优

- multiple nova-conductor  
Count.=Cores
- rpc\_call tuning
  - rpc\_thread\_pool\_size = 4096
  - rpc\_conn\_pool\_size = 512
- qpidd tuning
  - qpidd\_tcp\_nodelay = false
  - qpidd\_heartbeat = 240
- image\_cache\_manager\_interval = 0



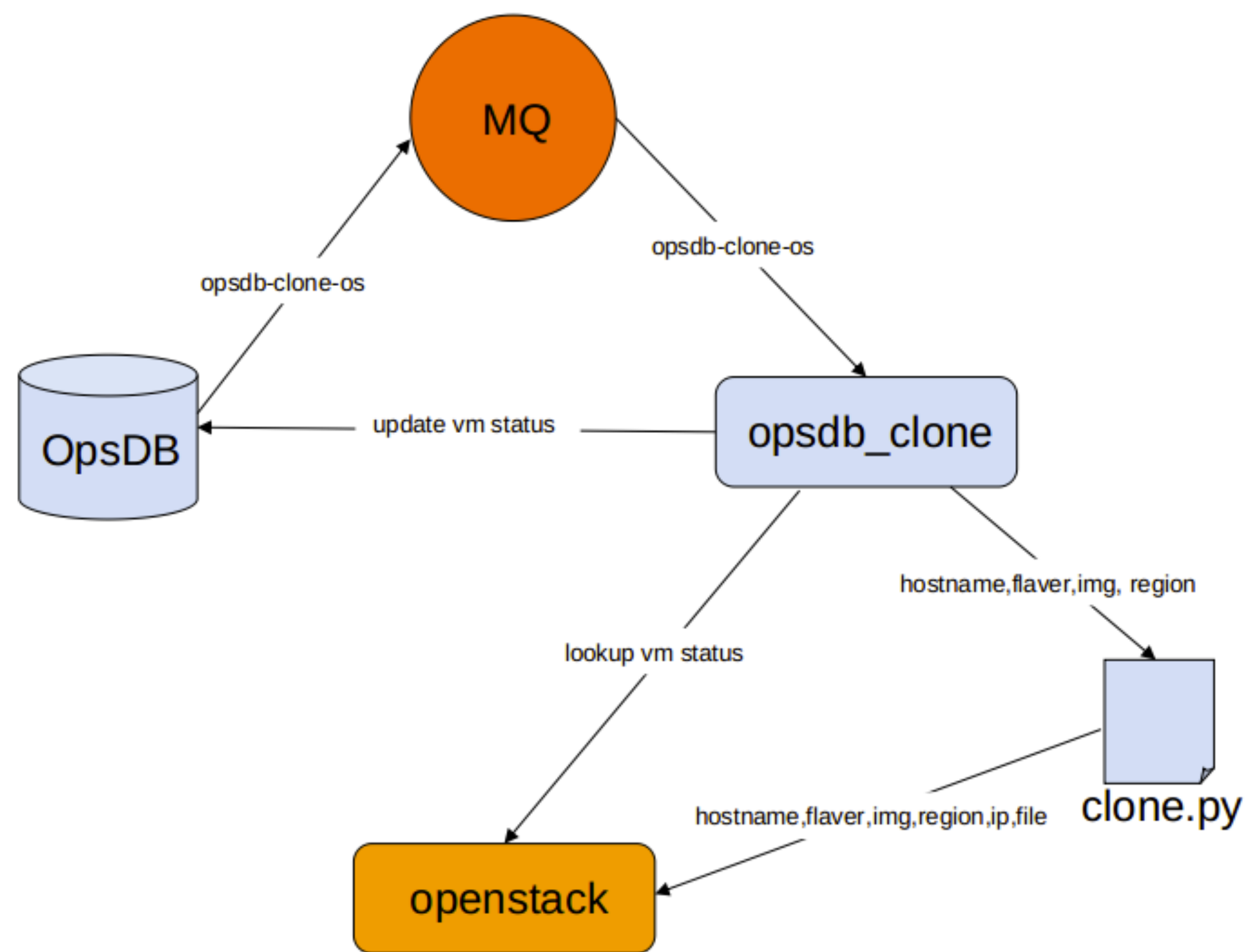


# 集群部署&管理

- Saltstack批量部署计算节点
- 自动化建立新集群（主控节点 + 计算节点）
- pillar变量替换
- 各组件使用RPM安装

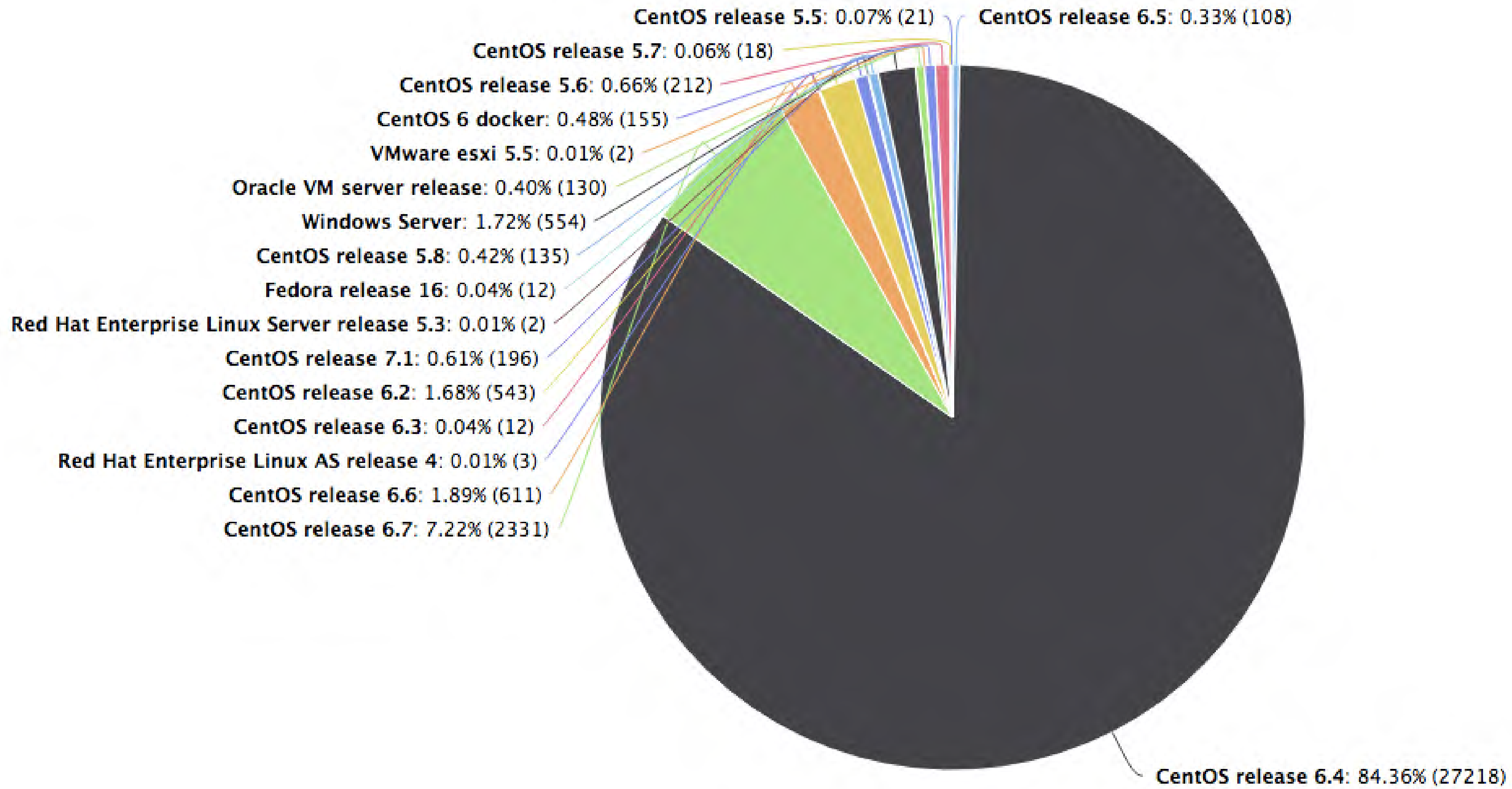


# opsdb clone

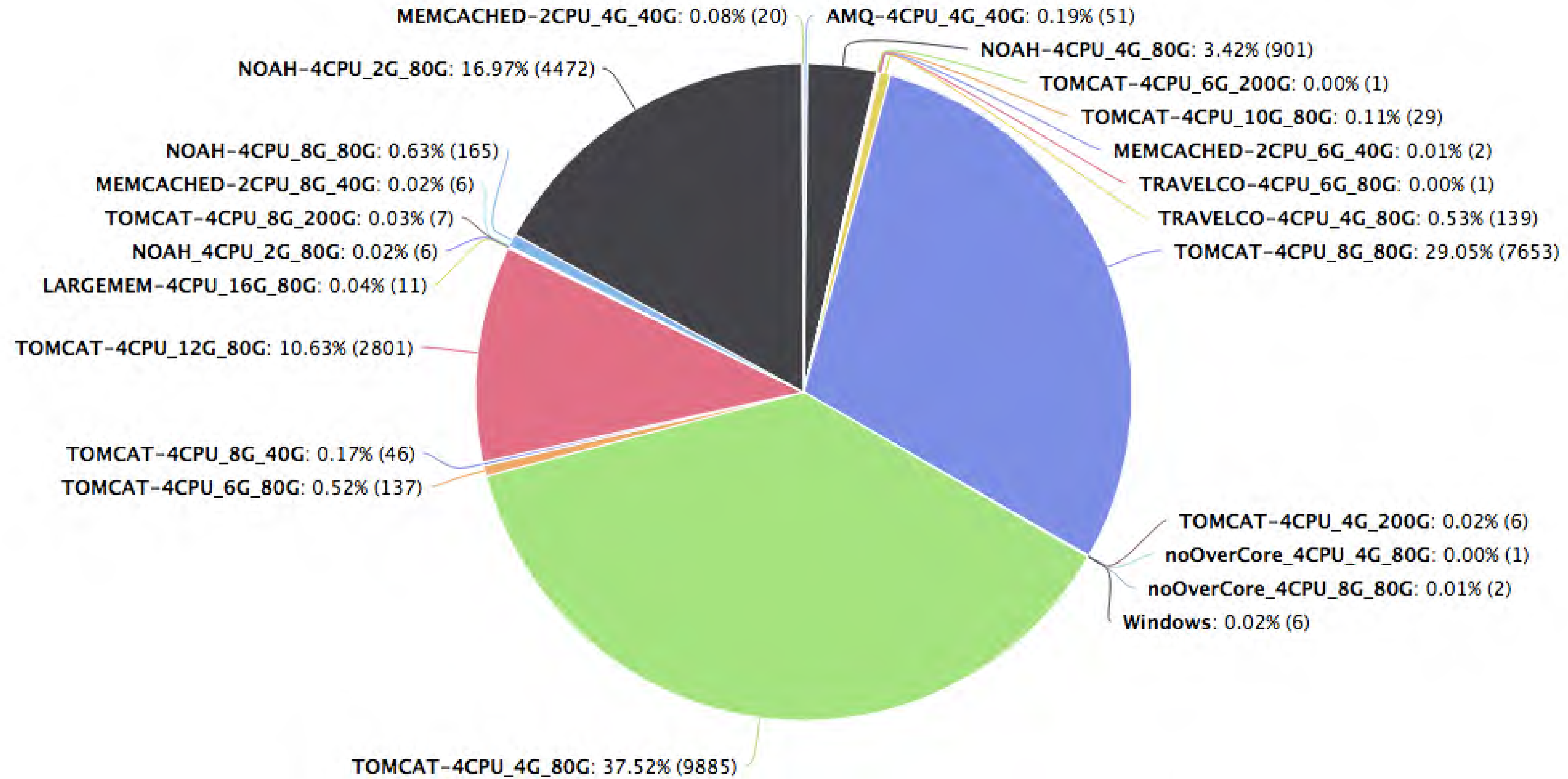




### os使用分布情况



### flavor使用分布情况





# 机架管理

CN1机房	CN2机房4层	CN2机房6层	CN5机房5层	CN5机房6层	CN6机房	CN8机房	CN9机房	ZH机房	HK1机房			
a01_1	a02_1	a03_1	a04_1	a05_1	a06_1	a07_1	a08_1	a09_1	a10_1	a11_1	a12_1	a13_1
a01_2	a02_2	a03_2	a04_2	a05_2	a06_2	a07_2	a08_2	a09_2	a10_2	a11_2	a12_2	a13_2
a01_3	a02_3	a03_3	a04_3	a05_3	a06_3	a07_3	a08_3	a09_3	a10_3	a11_3	a12_3	a13_3
a01_4	a02_4	承重柱	a04_4	a05_4	承重柱	a07_4	a08_4	承重柱	a10_4	a11_4	a12_4	承重柱
a01_5	a02_5		a04_5	a05_5		a07_5	a08_5		a10_5	a11_5	a12_5	
a01_6	a02_6	a03_4	a04_6	a05_6	a06_4	a07_6	a08_6	a09_4	a10_6	a11_6	a12_6	a13_4
a01_7	a02_7	a03_5	a04_7	a05_7	a06_5	a07_7	a08_7	a09_5	a10_7	a11_7	a12_7	a13_5
a01_8	a02_8	a03_6	a04_8	a05_8	a06_6	a07_8	a08_8	a09_6	a10_8	a11_8	a12_8	a13_6
a01_9	a02_9	a03_7	a04_9	a05_9	a06_7	a07_9	a08_9	a09_7	a10_9	a11_9	a12_9	a13_7
a01_10	a02_10	a03_8	a04_10	a05_10	a06_8	a07_10	a08_10	a09_8	a10_10	a11_10	a12_10	a13_8
a01_11	a02_11	a03_9	a04_11	a05_11	a06_9	a07_11	a08_11	a09_9	a10_11	a11_11	a12_11	a13_9
a01_12	a02_12	a03_10	a04_12	a05_12	a06_10	a07_12	a08_12	a09_10	a10_12	a11_12	a12_12	a13_10
a01_13	a02_13	a03_11	a04_13	a05_13	a06_11	a07_13	a08_13	a09_11	a10_13	a11_13	a12_13	a13_11
a01_14	a02_14	a03_12	a04_14	a05_14	a06_12	a07_14	a08_14	a09_12	a10_14	a11_14	a12_14	a13_12
a01_15	a02_15	a03_13	a04_15	a05_15	a06_13	a07_15	a08_15	a09_13	a10_15	a11_15	a12_15	a13_13
a01_16	a02_16	a03_14	a04_16	a05_16	a06_14	a07_16	a08_16	a09_14	a10_16	a11_16	a12_16	a13_14



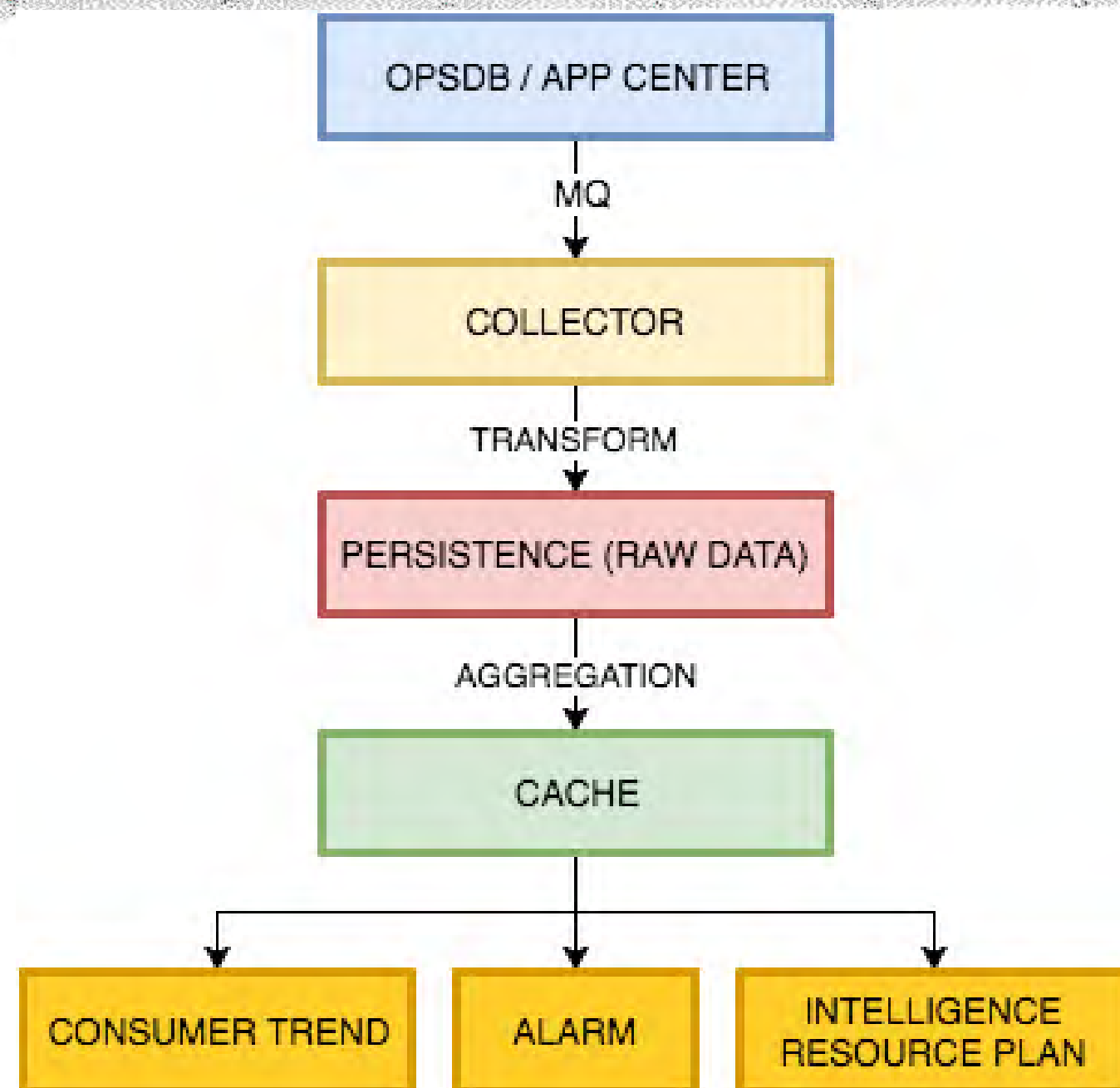


# 机架管理





# 计费



4CPU_6G_80G	5	¥386.53
4CPU_8G_200G	5	¥516.75
4CPU_8G_40G	26	¥2,480.40
4CPU_8G_80G	776	¥52,293.33
Dell Inc. - PowerEdge R720	85	¥29,372.88
Dell Inc. - PowerEdge R720xd	366	¥126,476.16
Dell Inc. - PowerEdge R730xd	177	¥842.68
HJ210-BDOCXO	1	¥345.56
HJ210-BDRB	2	¥691.13
HP - ProLiant DL160 G6	22	¥7,602.39

```

{
  "openstack-vm": {
    "count": 449,
    "price": 449,
    "detail": {
      "4C8G80G": {
        "count": 168,
        "price": 168
      },
      "4C12G80G": {
        "count": 35,
        "price": 35
      },
      "4C4G80G": {
        "count": 246,
        "price": 246
      }
    }
  },
  "server": {
    "count": 3,
    "price": 3,
    "detail": {
      "HP - ProLiant DL380p Gen8": {
        "count": 2,
        "price": 2
      },
      "Dell Inc. - PowerEdge R620": {
        "count": 1,
        "price": 1
      }
    }
  }
}
  
```



# 集群Quota

应用树 我的应用 我的主机

qunar.corp.ops.dev.mesos.esaas

新增 重命名 ApiKey 删除

- hadoop 6
- hardware\_dev 21
- hr 45
- hyper\_stats
- ia\_system 3
- iphone 3
- isapi 3
- itdb 2
- it\_workstation 1
- ivrproxy 17
- mesos 308
  - beta 1
  - bufslave 82
  - dns
  - esaas 110
    - mesos\_master 9
    - proxy
    - slaves 93
    - zookeeper 8
- mfs\_test
- miaotest 1
- mobileapproval
- nghcapi 2
- ngxdb 4
- ngx\_qps\_bps\_opentsdb 1
- notification 21
- oa 2
- oa\_workflow

概览 主机 负责人

有监控主机 **110**  
低使用率 **26**

CPU **5.39 %**  
内存 **51.9 %**  
流量 **19.57 Mib/s**

主机账单

7天账单(¥) **9.41 K**

30天账单(¥) **40.4 K**

Compute nodes

190 servers

Running VMs

3078 instances



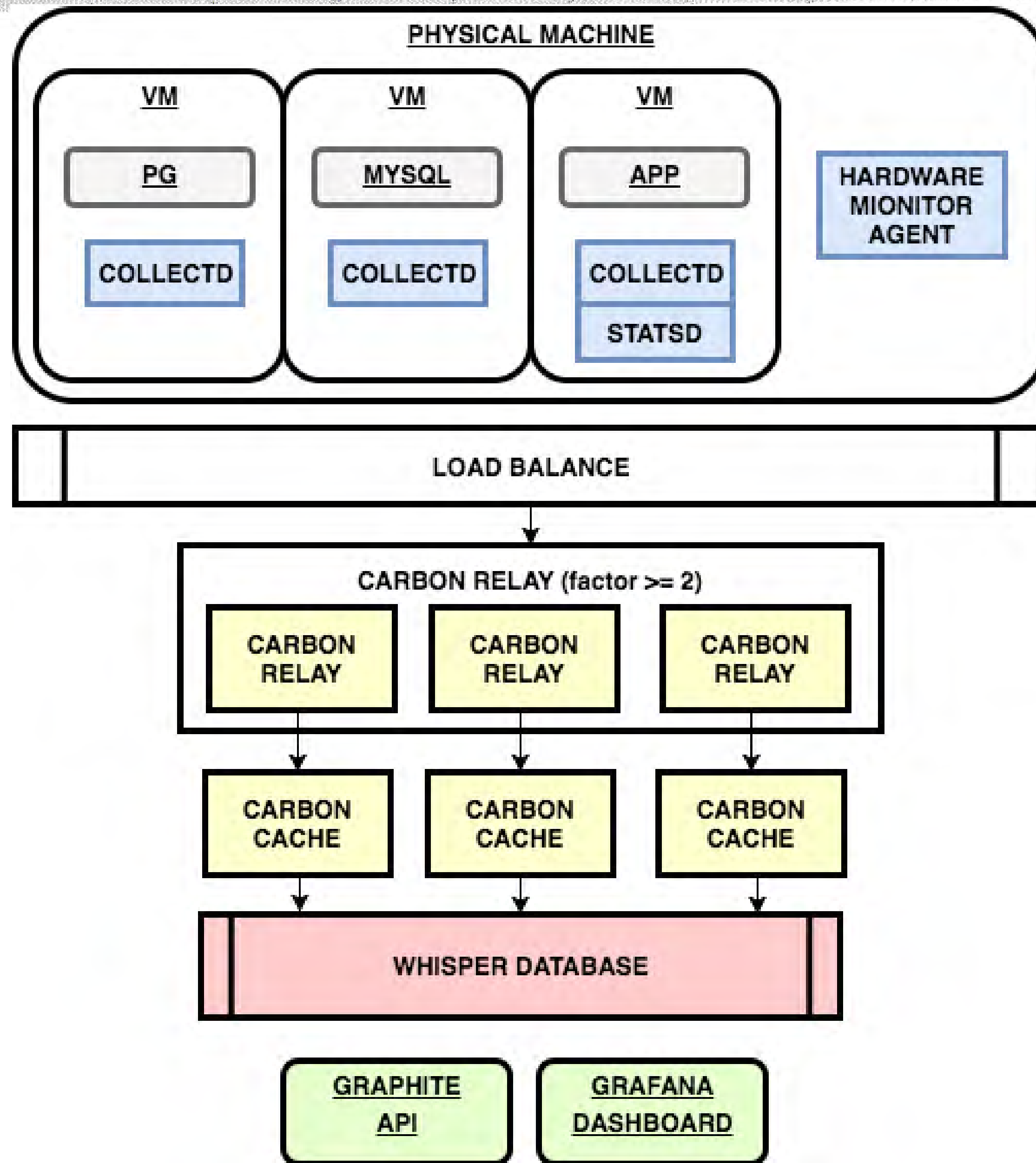


# 集群Quota

Region: cn8_qstack_prod		
Memory	Disk	Physical cores
<b>22.75 TiB</b>	<b>332 TiB</b>	<b>4832 cores</b>
Compute nodes	Running VMs	
<b>190 servers</b>	<b>3078 instances</b>	



# 监控架构

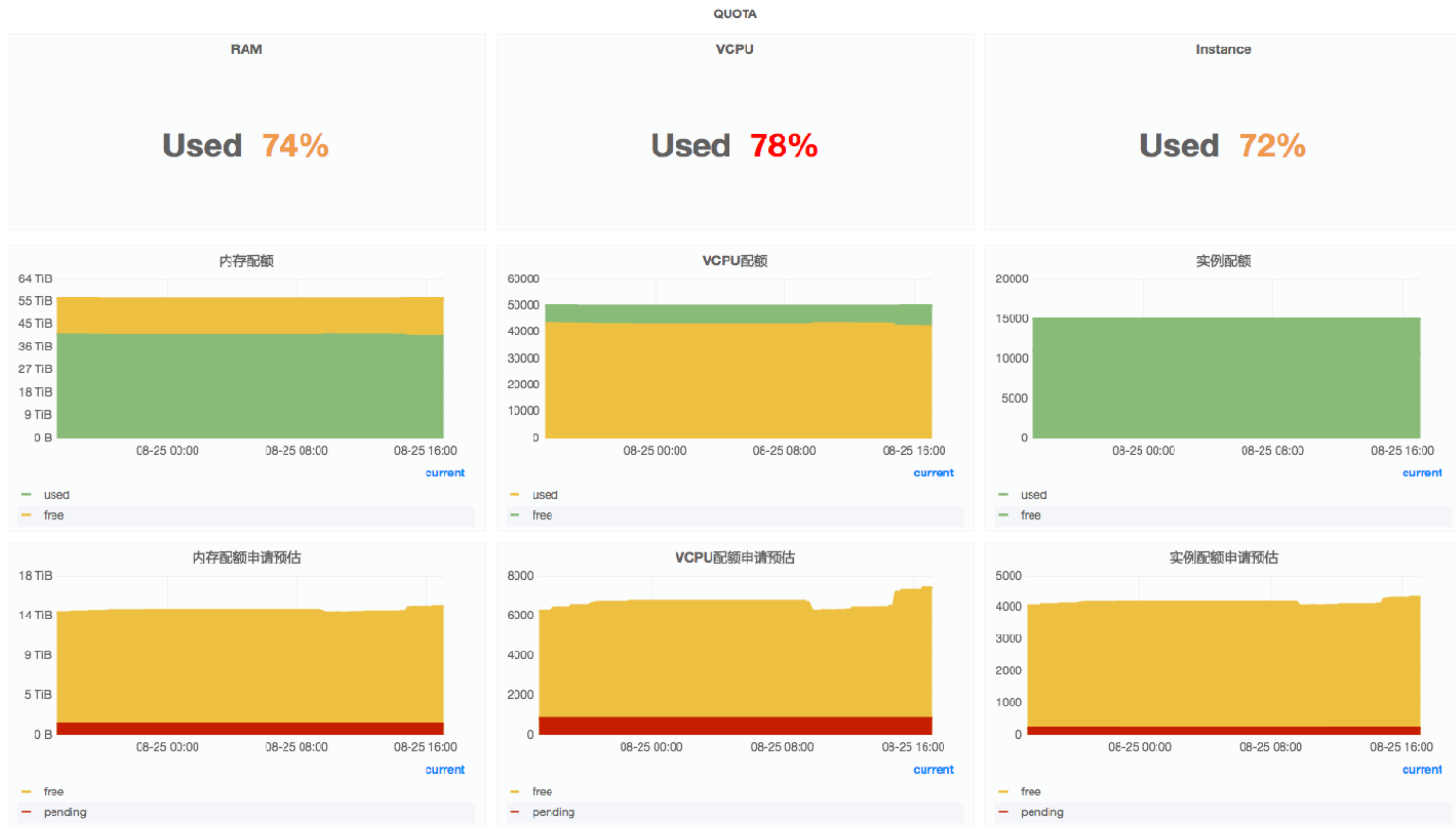




# Clone监控

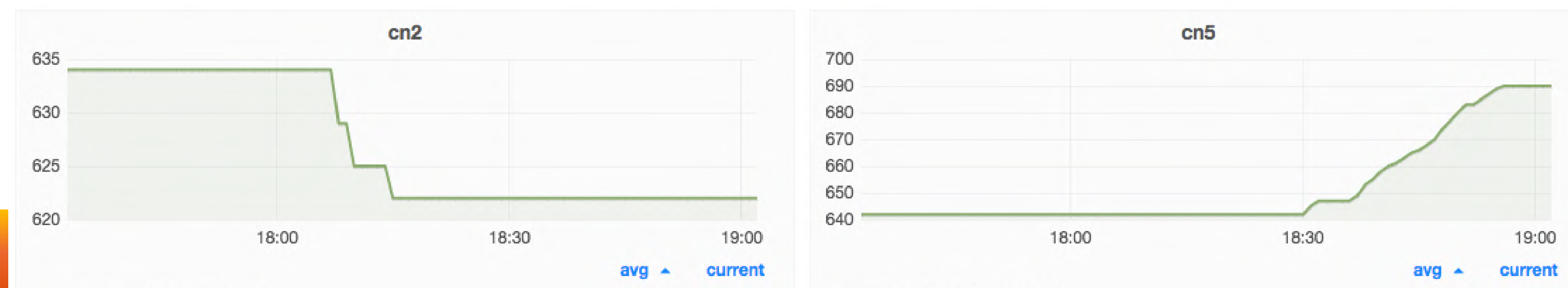
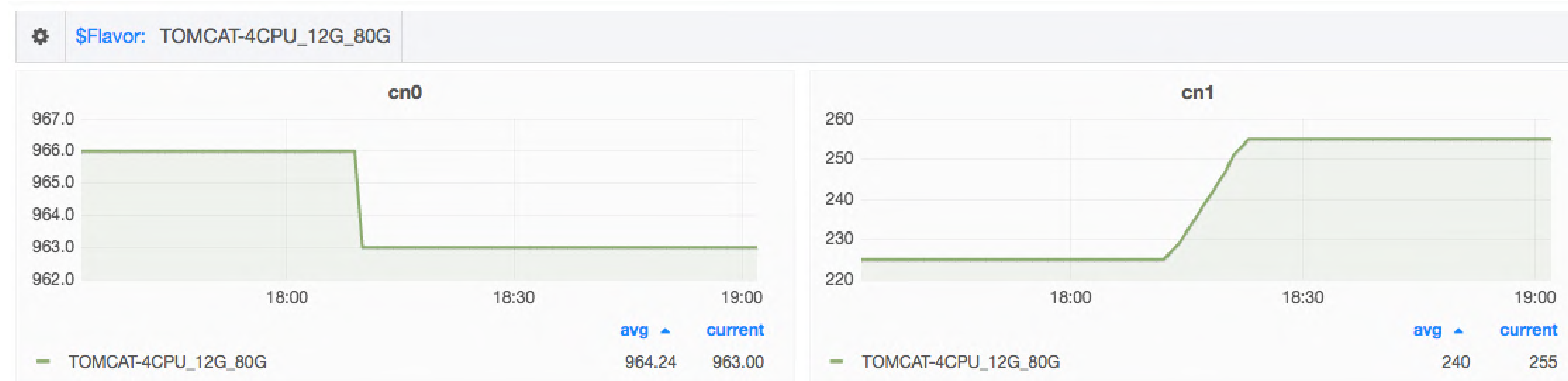
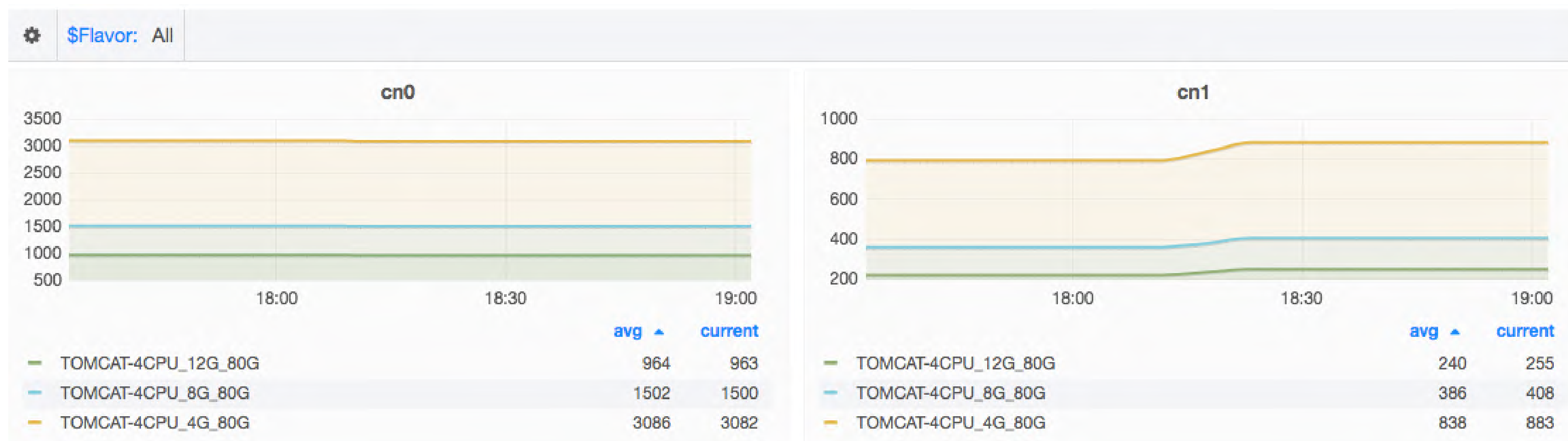


# Quota监控

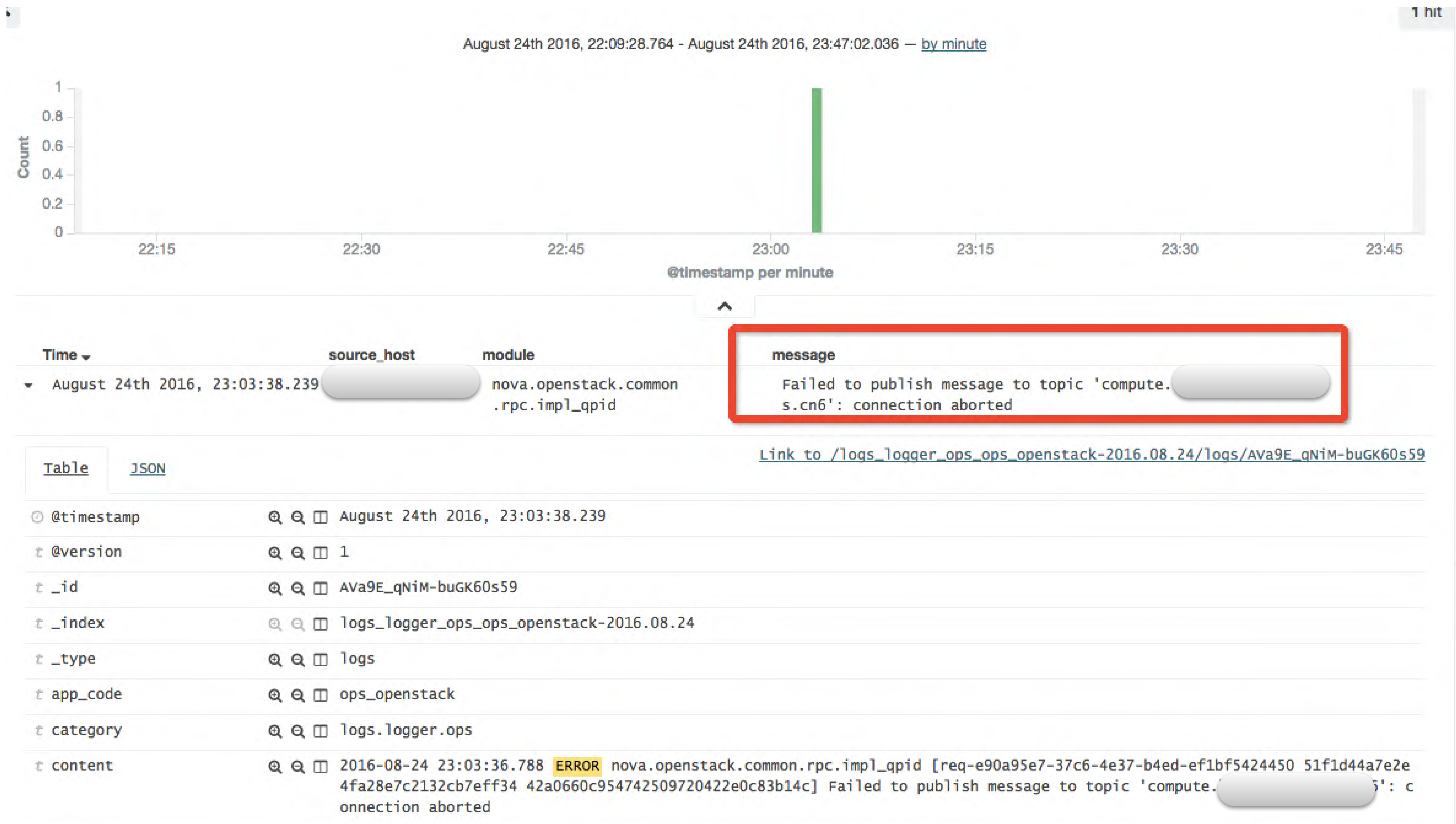




# Available Resource by Flavor



# 日志与监控





# 硬件自动化运维

- 新服务器选型测试
- 服务器到货自动化检测
- 服务器故障自动化运维
- 机房数据可视化
- 网络设备配置运维平台



# 新服务器选型测试

到货检测记录   到货检测性能数据   到货检测异常

到货检测性能数据

共 10 记录

搜索:

序列号	CPU整型评分	内存带宽评分	硬盘顺序256k读	硬盘顺序256k写	硬盘随机4k读	硬盘随机4k写	测试时间
70C39X1	21414	35864.96	1240.70		6.72		2016-12-13 20:19:47
1FSFS72	23987	39213.26	380.75	373.19	5.07	5.18	2016-12-12 16:57:42
6CU4350AYF	26658	37368.4	425.62	327.04	4.77	6.30	2016-12-12 13:09:20
2L28K82	25660	49549.36	396.87	386.77	4.82	5.64	2016-12-12 11:23:05





# 新服务器选型测试

- 收集线上全部设备配件配置
- 建成配件库
- 配件自动盘点
- 配件故障率分析
- 批次问题预警



# 服务器故障自动化运维FaultRepair

- 完成线上全部服务器故障监控工具(hwmon)部署，准确报警；
- 故障自动发单系统(FRS)结合hwmon报警，实现除开死机之外的全部硬件故障自动报警发单
- NOC值班服务器故障基本不用手动处理硬件故障；平均处理周期缩减至3.5天

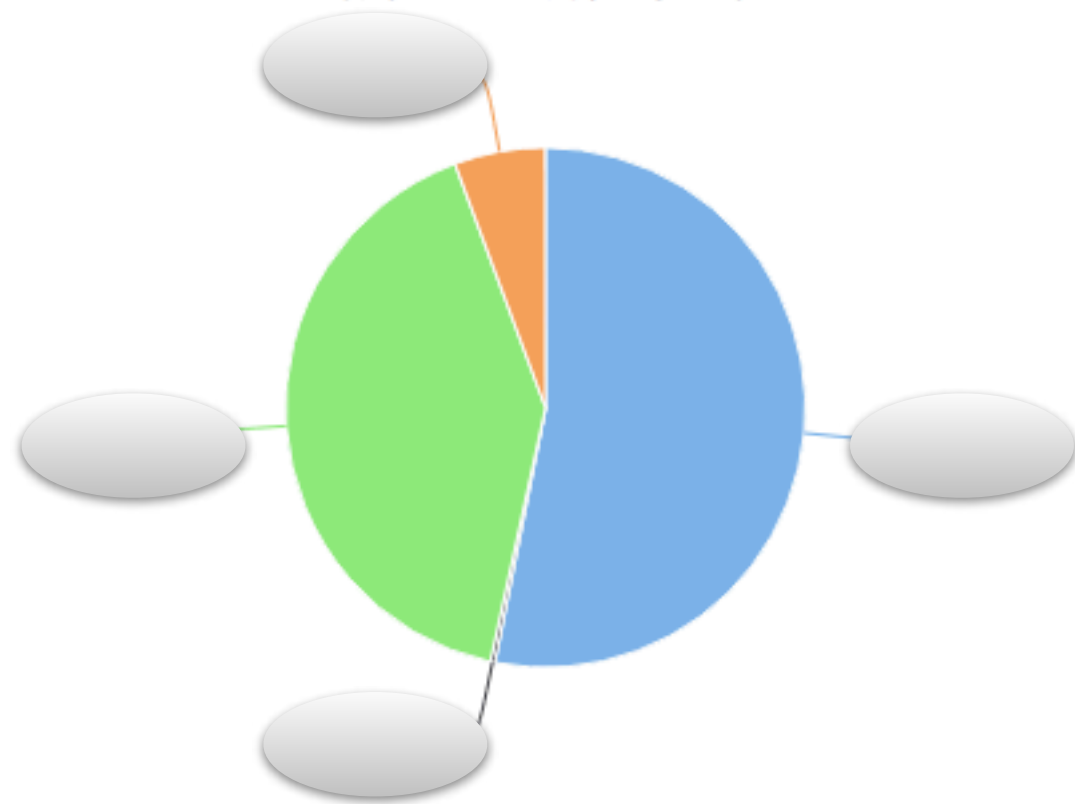




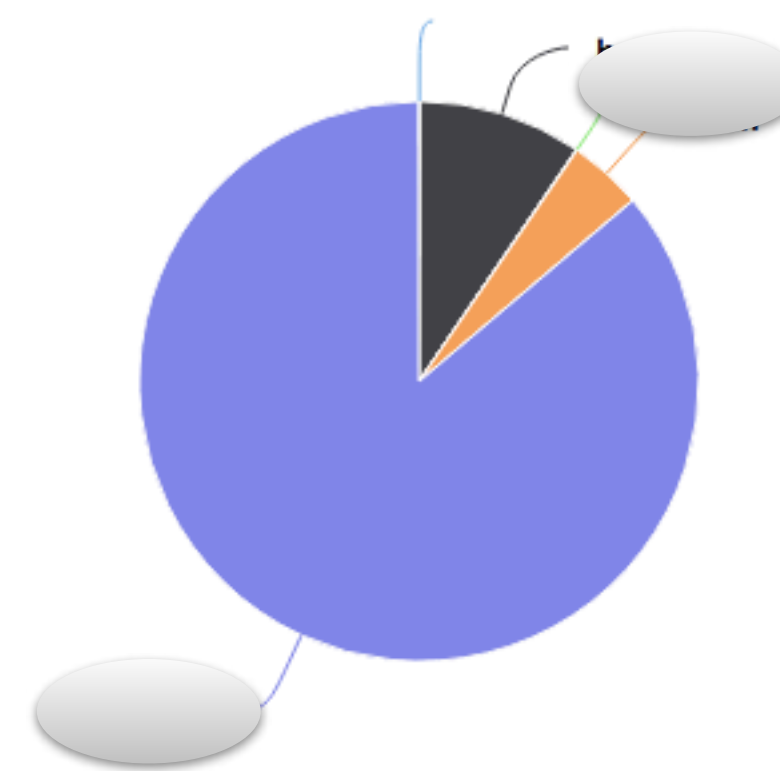
# 服务器故障自动化运维

## 硬件异常统计

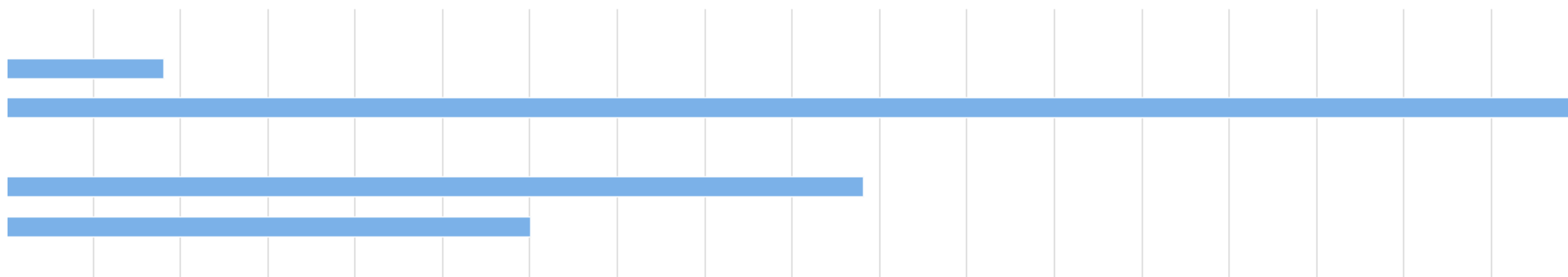
各个OEM总故障记录分布



各个OEM总异常记录分布



各个OEM故障率对比(%)



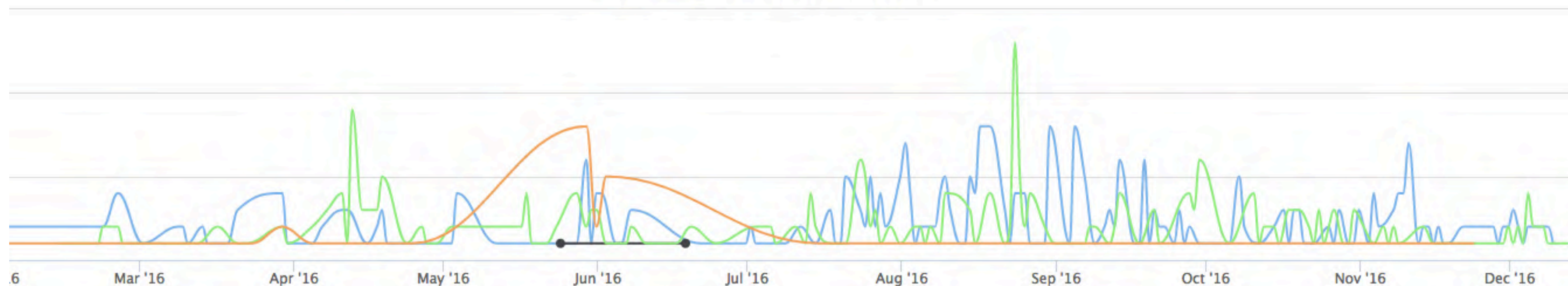


# 服务器故障自动化运维

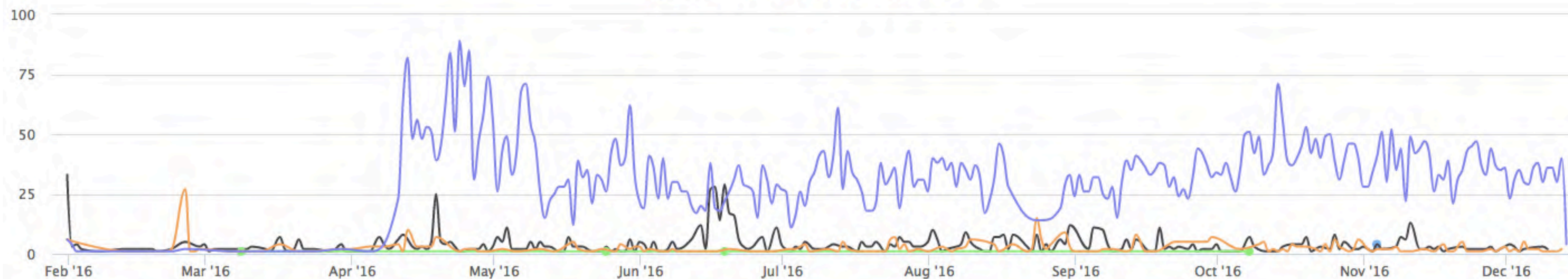
FRS 故障报修 故障统计分析 配置管理

- 点击报修故障
- 自助故障报修
- 全部报警信息

各个OEM月故障记录对比



各个OEM月异常记录对比



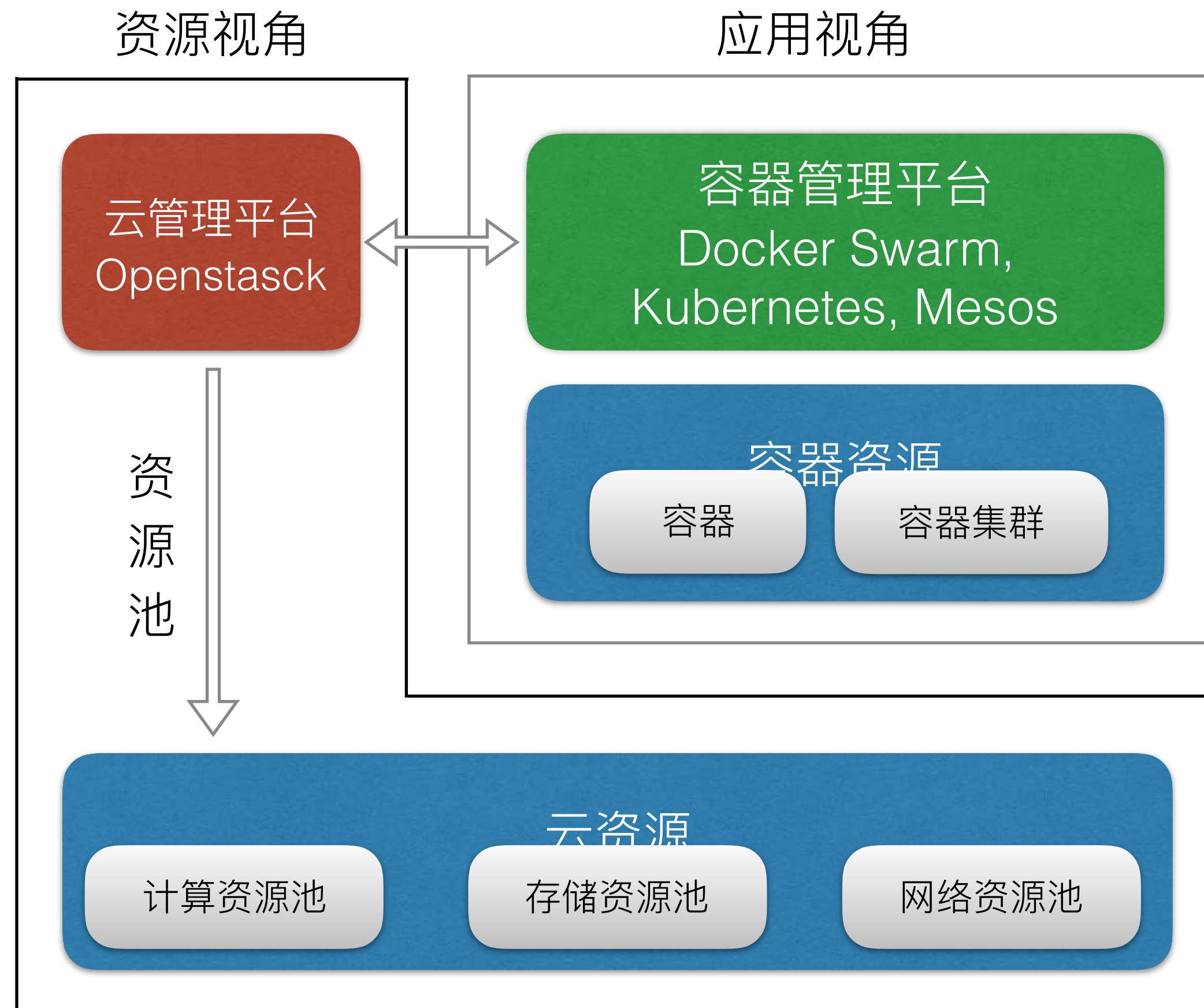


# Docker vs Openstack

Docker和OpenStack生态在多个领域有互补, 存在互相调用集成的关系

**资源视角: Openstack对容器资源的纳管**  
如何支持更加灵活的弹性调度  
如何支持快速的副本复制

**应用视角: Openstack对主流容器管理平台的支持**  
如何支持容器的永久化存储  
如何支持容器集群多层次的网络互通方案





# OpenStack应用: Noah on OpenStack

## 项目环境 查看及配置用于项目测试的环境

首页 > 环境实例 > 项目环境

测试工具

全部应用

环境模板

项目环境

联调环境

测试工具/项目环境/testconsul1 ( 环境ID:10381 )

←返回列表

锁定

SALT

环境变量预览

HOSTS

编译顺序

部署顺序

展开列表

全选

测试任务

更多

关联环境

Empty

git@gitlab.corp.qunar.com:qa/ntp.git

	类型	应用名称	端口	服务器列表	域名
<input type="checkbox"/>	WAR	qa/ntp/ntp-timer [betaa][a]	8082	<input type="text" value="...cn0.quna..."/> >_ H 文件 刷新 编辑 分享	Empty



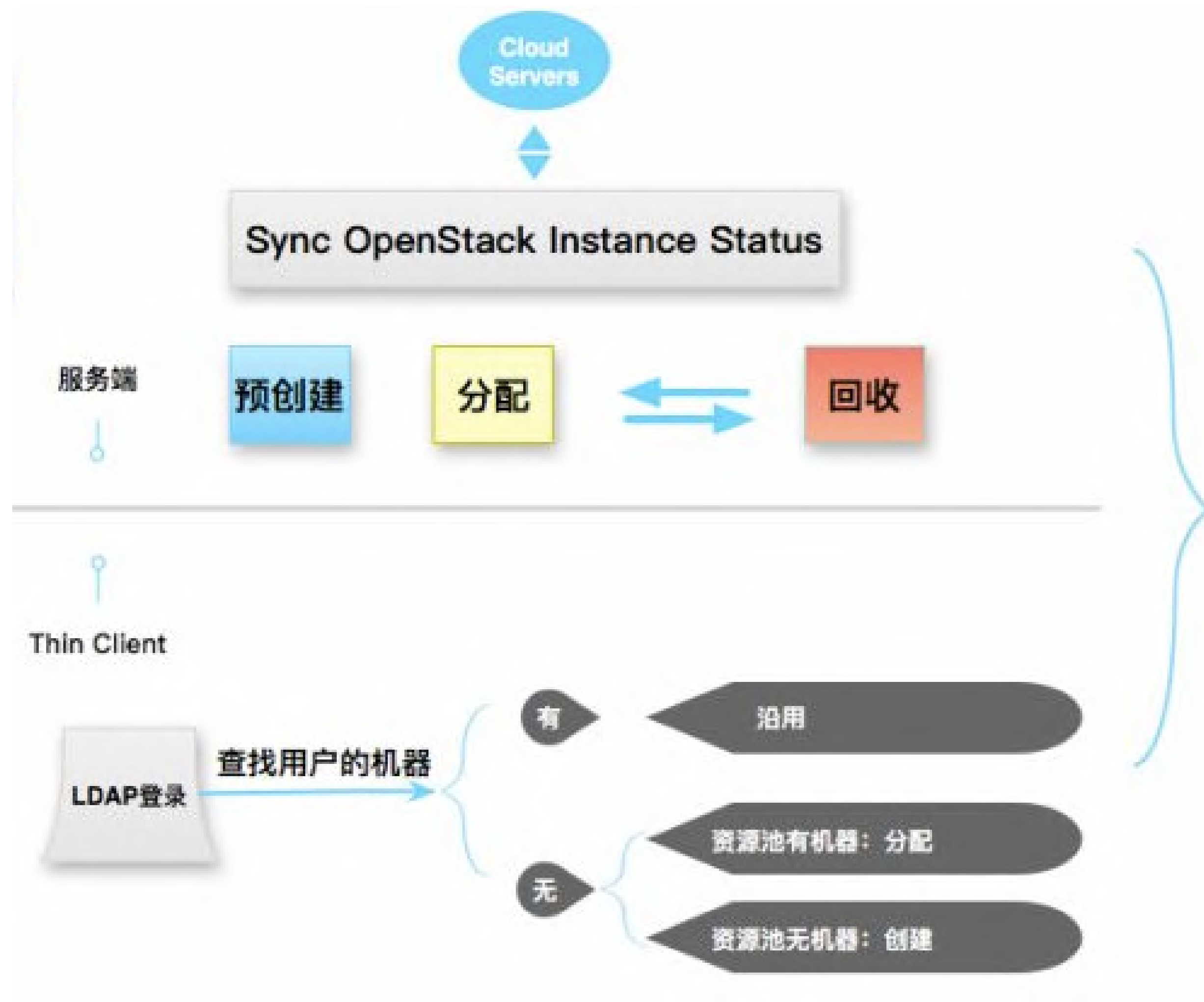


# OpenStack Desktop SaaS

- 可添加多个OpenStack集群
- 原生接口与OS集群兼容，无需修改OS代码/SDK
- Thin client性能结合硬件调优+低成本开源方案
- 实时监控客户端状态




# OpenStack Desktop SaaS





# OpenStack Desktop SaaS

 Qunar Desktop SAAS

- CLOUD
- Cloud servers
- Running Machines
- USERS
- Users

## Cloud Servers

You can manage cloud server in this page

Add Cloud Server

Name	Description	Username	Url	Actions
openstack	corp	admin	<input type="text" value="m:5000/v2.0/"/>	<a href="#">Images</a> <a href="#">Edit</a> <a href="#">Remove</a>
cn0	cn0-for-test	admin	<input type="text" value="5000/v2.0/"/>	<a href="#">Images</a> <a href="#">Edit</a> <a href="#">Remove</a>

 Windows 7	centos	centos	CentOS-7-x86_64	m1.medium	1	<a href="#">Edit</a> <a href="#">Destroy</a>
 Windows 7	windows	windows	Win2k8-r2-thin	m1.large	2	<a href="#">Edit</a> <a href="#">Destroy</a>

 Qunar Desktop SAAS 11.18.16

- cloud
- Cloud servers
- Running Machines
- users
- Users

## Running machines

[同步](#) [新增创建](#) [回分配创建机器](#) [回收](#)

System	Image	Instance's User Name	UUID	机器状态	Spice Info	Actions
windows	2	zhaodong.sun	55ee09d7-8657-4809-a878-52befb5eee1a	active	<input type="text"/>	<a href="#">Start</a> <a href="#">Reboot</a> <a href="#">Destroy</a>
windows	2	shaoping.li	f8c3a8ce-5f23-452e-8ab8-04a4a7a52408	active	<input type="text"/>	<a href="#">Start</a> <a href="#">Reboot</a> <a href="#">Destroy</a>
windows	2	未分配	5d88698b-94d2-43fc-908b-c0276df657bc	shutdown	<input type="text"/>	<a href="#">Start</a> <a href="#">Reboot</a> <a href="#">Destroy</a>



# DevOpsDays 即将首次登陆中国



DevOps 之父 Patrick Debois 与您相约

DevOpsDays 北京站 2017年3月18日



门票早鸟价仅限前100名，请从速哟

<http://2017-beijing.devopsdayschina.org/>







想第一时间看到  
高效运维社区公众号  
的好文章吗？

请打开高效运维社区公众号，点击右上角小人，如右侧所示设置就好





# Thanks

高效运维社区  
开放运维联盟

荣誉出品

