

OpenStack@eBay



Geekbang>

极客邦科技

全球领先的技术人学习和交流平台

扫我，码上开启新世界



Geekbang>

InfoQ | EGO NETWORKS | StuQ

InfoQ

专注中高端技术人员
的社区媒体

EGO NETWORKS

EXTRA GEEKS' ORGANIZATION
高端技术人员
学习型社交网络

StuQ

实践驱动的IT职业
学习和服务平台

InfoQ^{ueue}

促进软件开发领域知识与创新的传播

ArchSummit
全球架构师峰会

实践第一 案例为主

时间：2015年12月18-19日 / 地点：北京·国际会议中心

欢迎您参加ArchSummit北京2015, 技术因你而不同



ArchSummit北京二维码

QCon
全球软件开发大会

[北京站]

2016年04月21日-23日



关注InfoQ官方信息
及时获取QCon演讲视频信息

About me

- 在IT业打混了14年
- 从电信和通讯行业开始职业生涯
- 在2009年进入了互联网行业，从那时起专注于大规模分布式系统的建设
- 是eBay自建云平台Stratus的主要贡献者
- 是eBay新一代基于Openstack云平台的技术lead
- 曾经是一名活跃的开源贡献者，从2005年开始活跃于source forge 社区

Agenda

Why Openstack

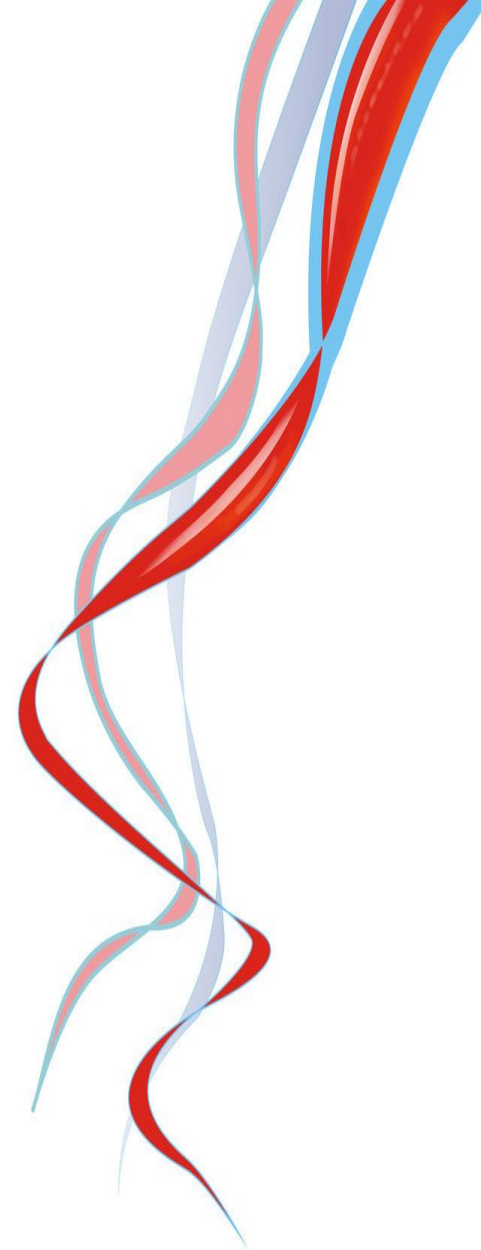
Openstack @ scale

Plan for failure

Monitoring, Alerting, Operation

Build & Deploy

Q&A



Openstack journey

COMPUTE

12+K

Hypervisors

300K Cores

70K VMs

NETWORK

15 VPCs

3 Regions

9 AZs

STORAGE

4.4 PB Local

52 TB Object

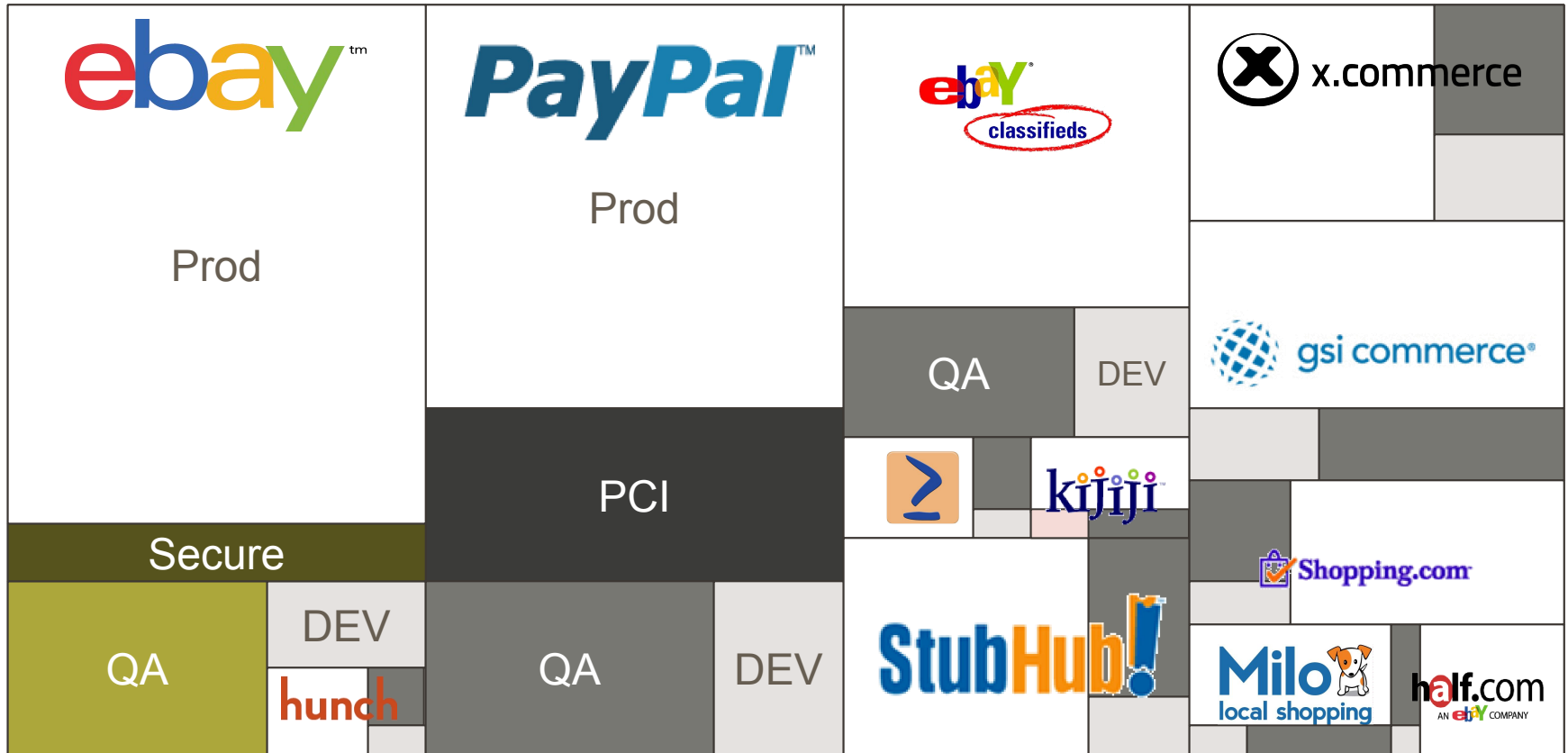
1.6 PB Block

OpenStack 2015 Super User Finalist

Before openstack

- Home grown Cloud : Stratus, includes:
 - OSP (Baremetal, Virtual Machine)
 - LBMS
 - DNS
 - RAS
 - ...
- Scale
 - 99+% eBay site traffic
 - 80,000+ virtual computes
 - 30,000+ physical computes

In 2012



One cloud on shared infrastructure



One
place to
go

- As a service
- Dashboard
- Service catalog
- Service health
- Support



One way
to do
things

- One control plane identity
- One set of public APIs
- Inter-operable



One
logical
control
plane

- Global topology
- Multi-tenant
- Service ecosystem

Principles

- Any application Anywhere
- Soft cabling
- Shared Standardized Infrastructure
- Virtualize everything
- Automate everything

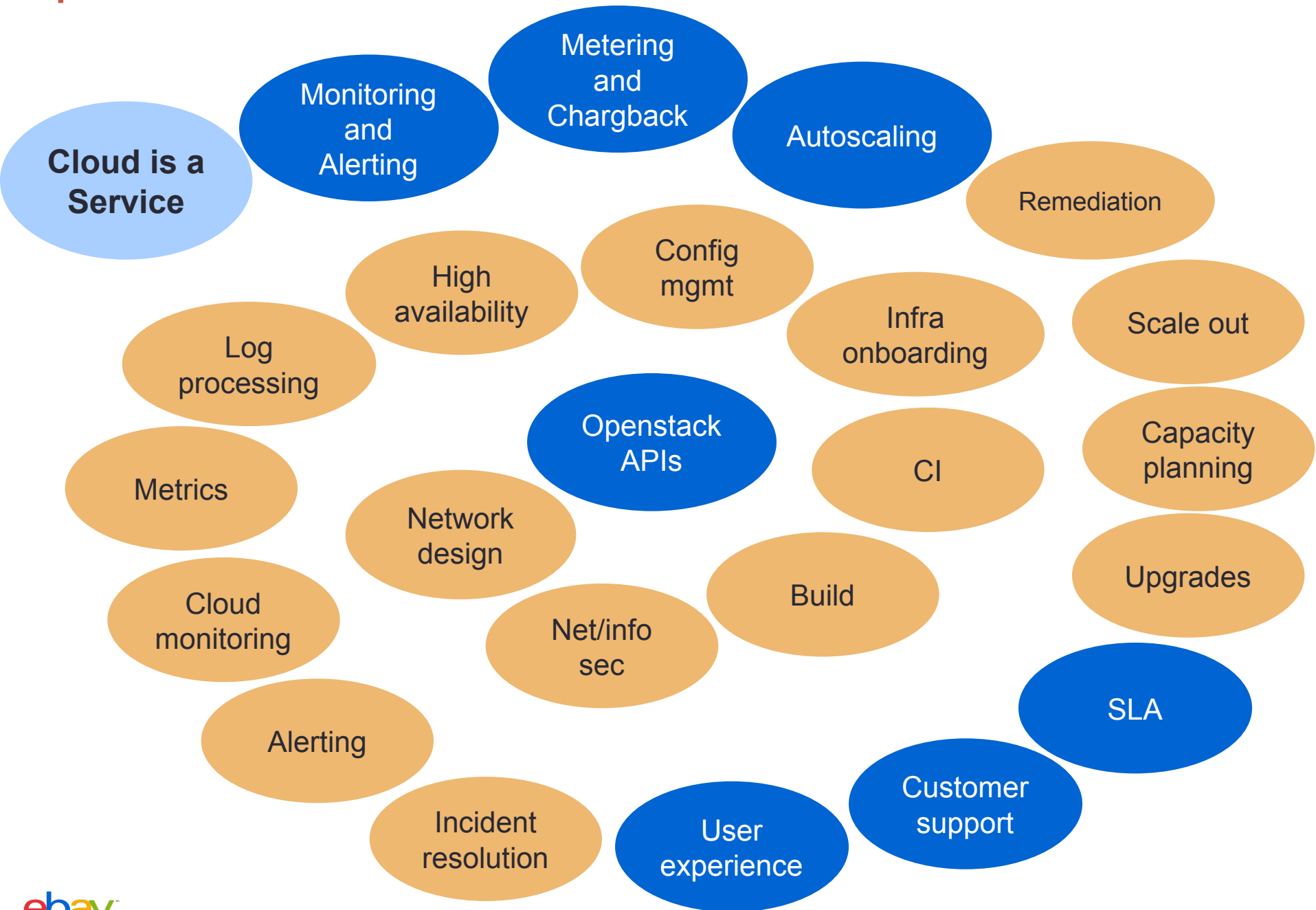
Openstack

- Community
- Vendor ecosystem
- Technology
- Adoption

Start from Dev Cloud in 2012

- A logical environment defined as a VPC on top of shared infrastructure
 - Self service VM for developers
 - Access is same as to their desktops
- Implemented as a set of L2 networks (/24) within a given L3(/20)
 - No private networks: all developers on same shared networks
 - No private IP space: traffic is routed within core, no need for floating IPs

Openstack is not cloud



PLAN FOR FAILURE

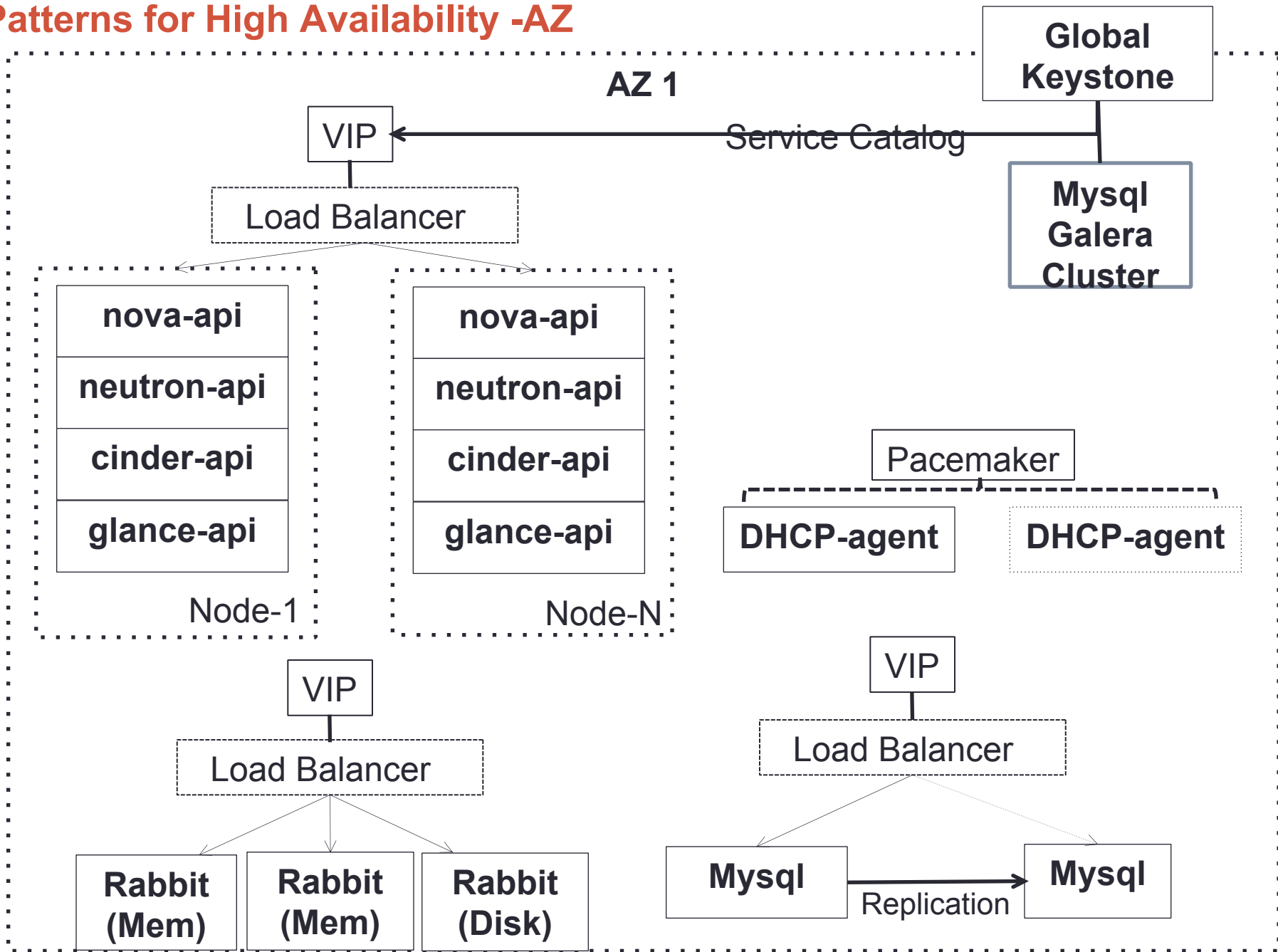
Patterns for High Availability -global

Multiple Geographic Regions for DR

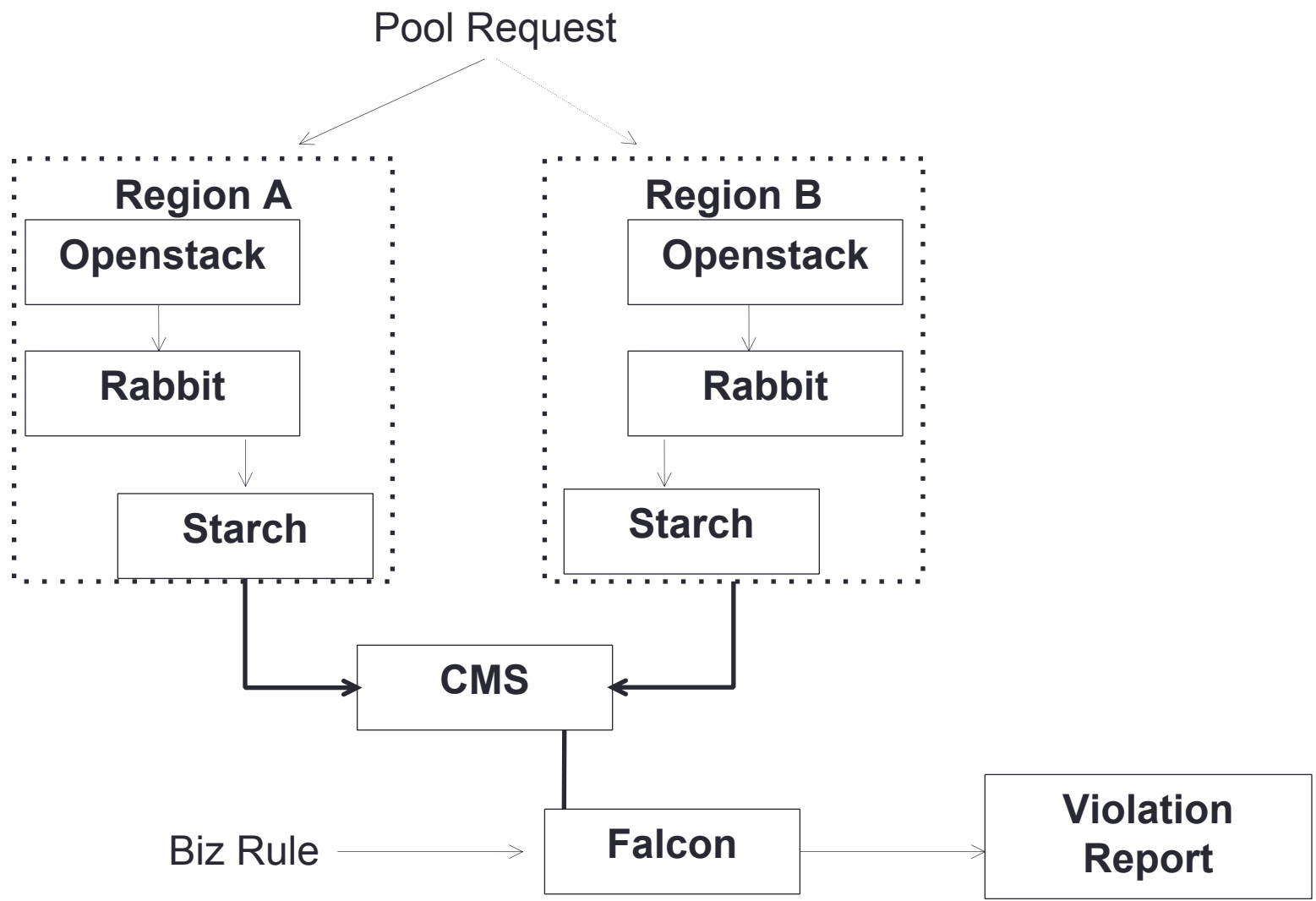


✓ 2 or More Availability Zones per Region - No sharing, low latency

Patterns for High Availability -AZ

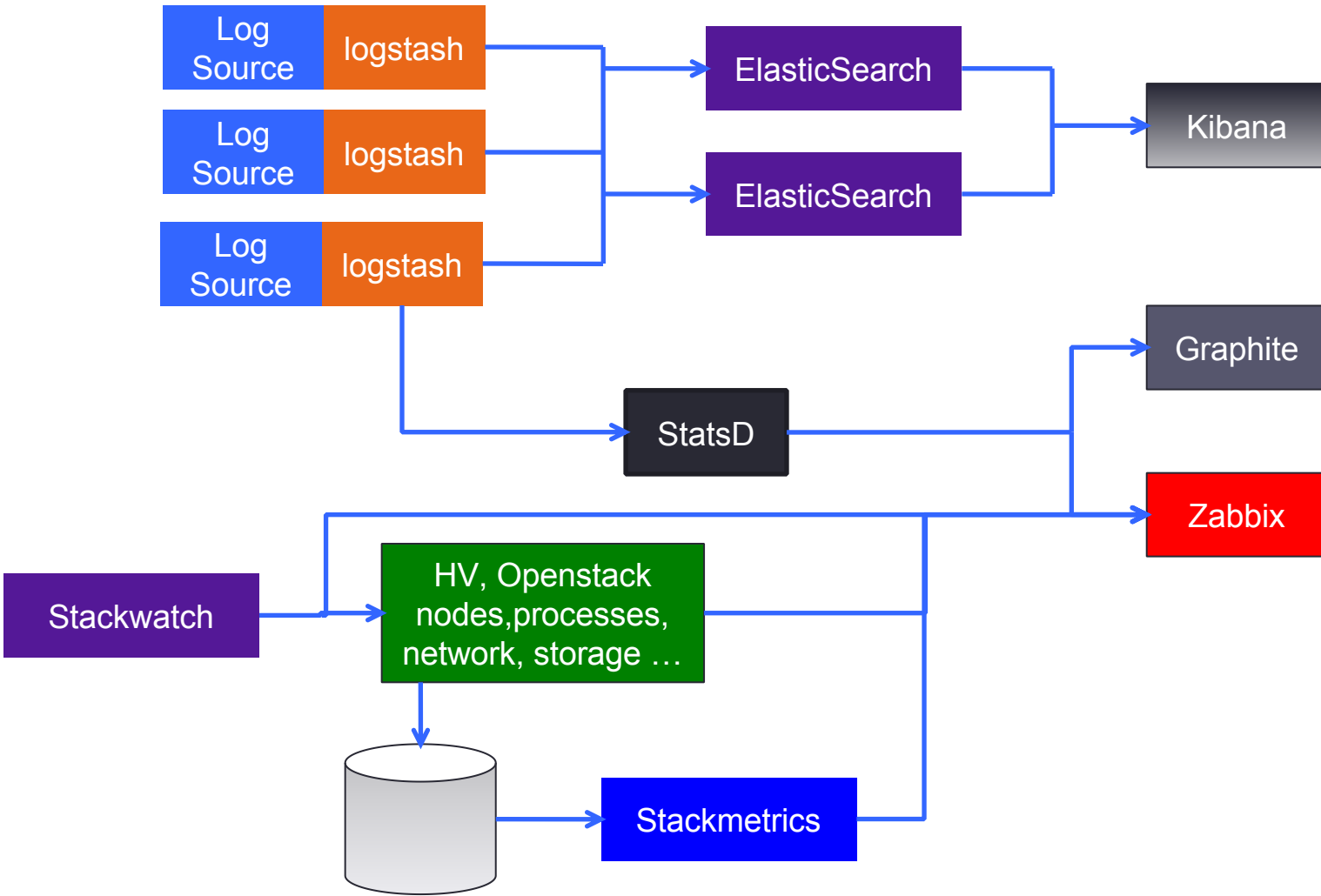


High Availability for your customer



Monitoring, Alerting, Operation

Monitoring evolution



Monitoring control plane

- Zabbix
- Stackmetrics
 - Including compute capacity monitoring
 - Including network capacity monitoring
- StackWatch
 - Customer simulator

Monitoring data plane

- Ping test
- Quark
 - VM Diagnostic
- HV hardware healthy check

Build and Deployment

Cattle cloud

Ubuntu,apt, fabric

Ubuntu,apt, puppet and foreman

Pet test clouds

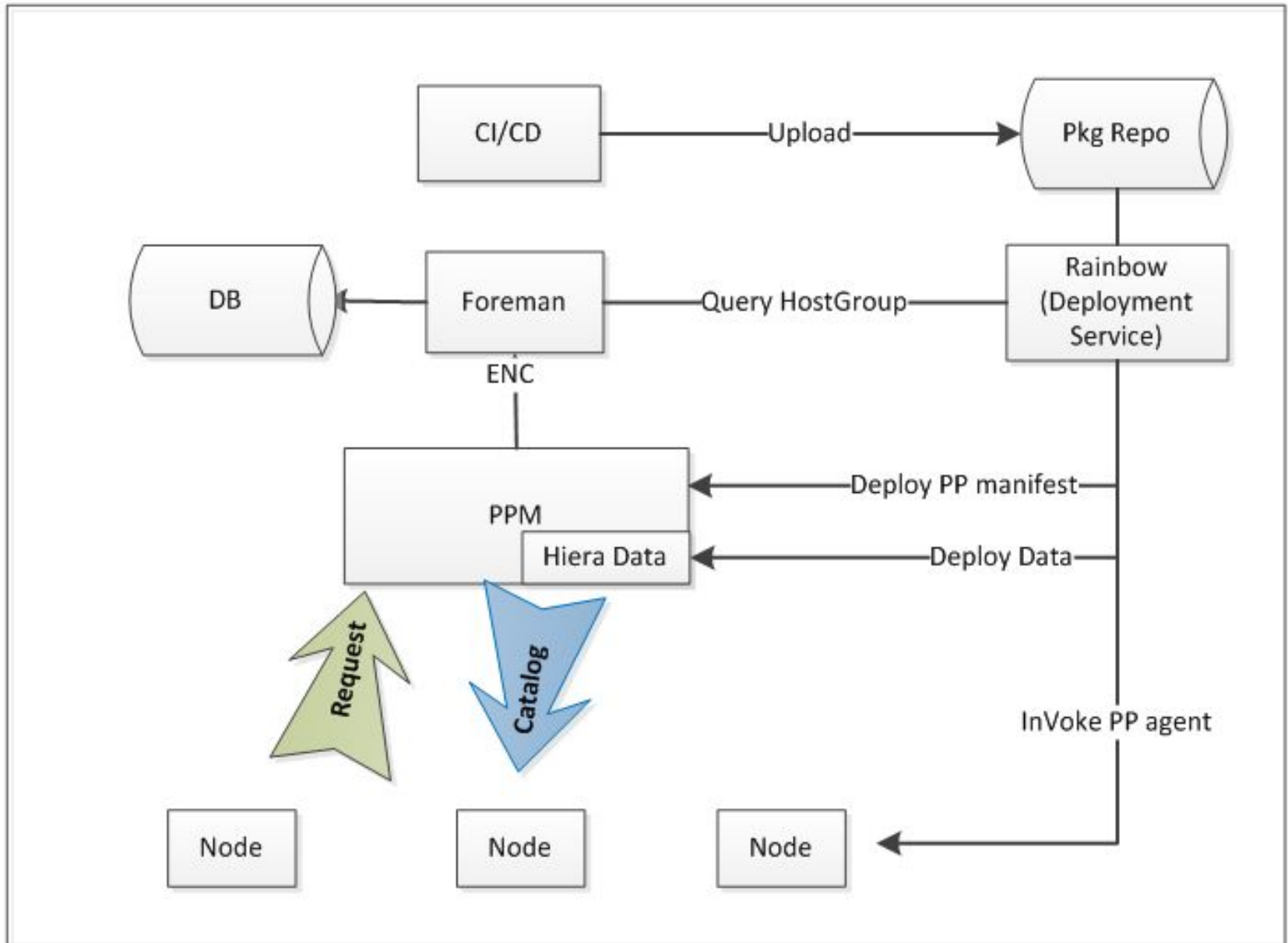
1. Bare-metal provisioning for HV, controllers, NVP,...
2. Virt env tarballs, puppet and foreman
3. A full working deployment

Cattle test clouds

Pinocchio

- Cloud on Cloud
- Dog fooding
- CI

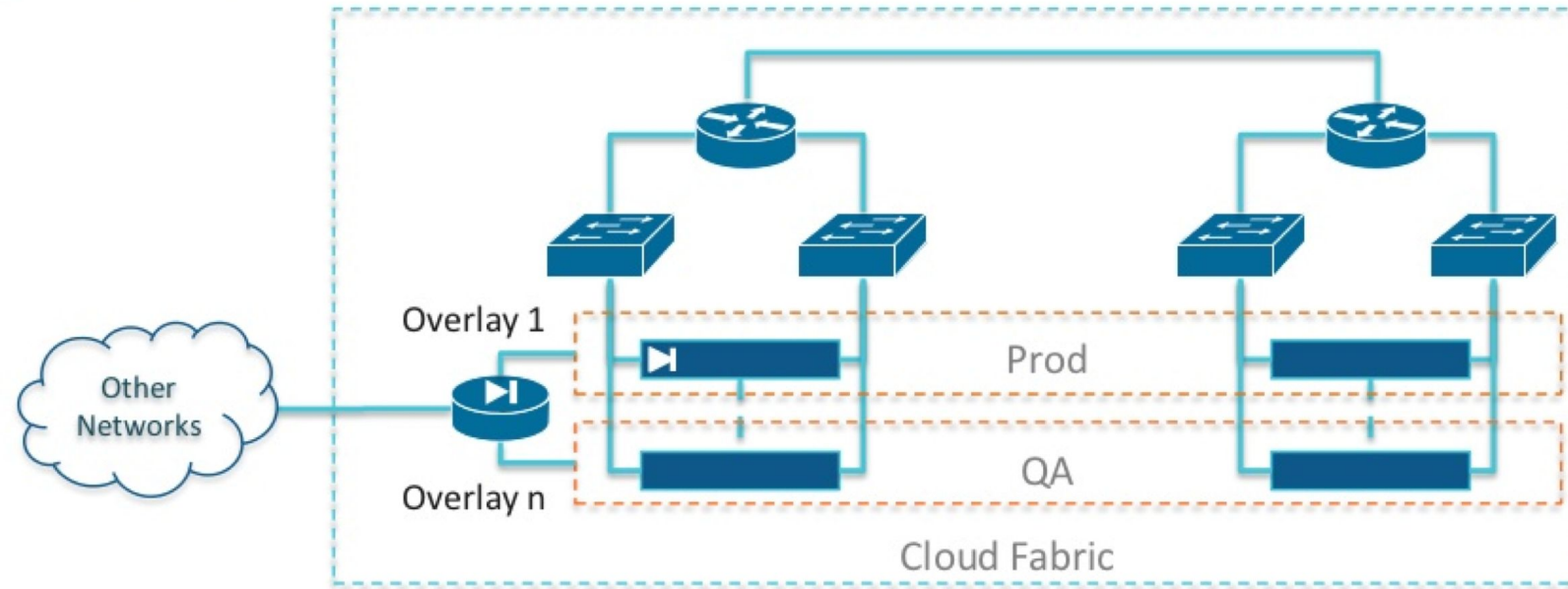
A full picture of Control plane deployment





Appendix Isolation :Virtual L2 Networks

Virtual Networks using Software Defined Networks



•Pros

- L2 isolation
- Compatible with large scale
- can be fully automated
- Firewall can be interposed between virtual networks
- Large number of networks

•Cons

- Tunnel overhead

Appendix Grafana overall dashboard

C3 Dashboards

Availability And Reliability

C3 Overall Reliability (One Dashboard)	Unreachable VM, Ping Failure Rate, HV Drifts, etc.
Individual C3 Services Reliability and Performance	Request errors, API response time, etc.
PHX01 Reliability	Unreachable VM, Ping Failure Rate, HV Drifts, etc.
LVS01 Reliability	Unreachable VM, Ping Failure Rate, HV Drifts, etc.
SLC01 Reliability	Unreachable VM, Ping Failure Rate, HV Drifts, etc.
Control Plane and Data Plane Availability SLA	VIP availability
LBMS Availability and Reliability	From LBMS Stackwatch routine
LBaaS Availability and Reliability	From LBaaS Stackwatch routine
Ping Failure Rate and Latency	From StackWatch ping test
Compute Provisioning Reliability	Reliability of VM creation and deletion
Cinder Reliability	From Cinder StackWatch test
Horizon Availability and Reliability	From StackWatch tests and logs.
Keystone Availability and Reliability	From StackWatch tests and logs.

C3 Dashboards

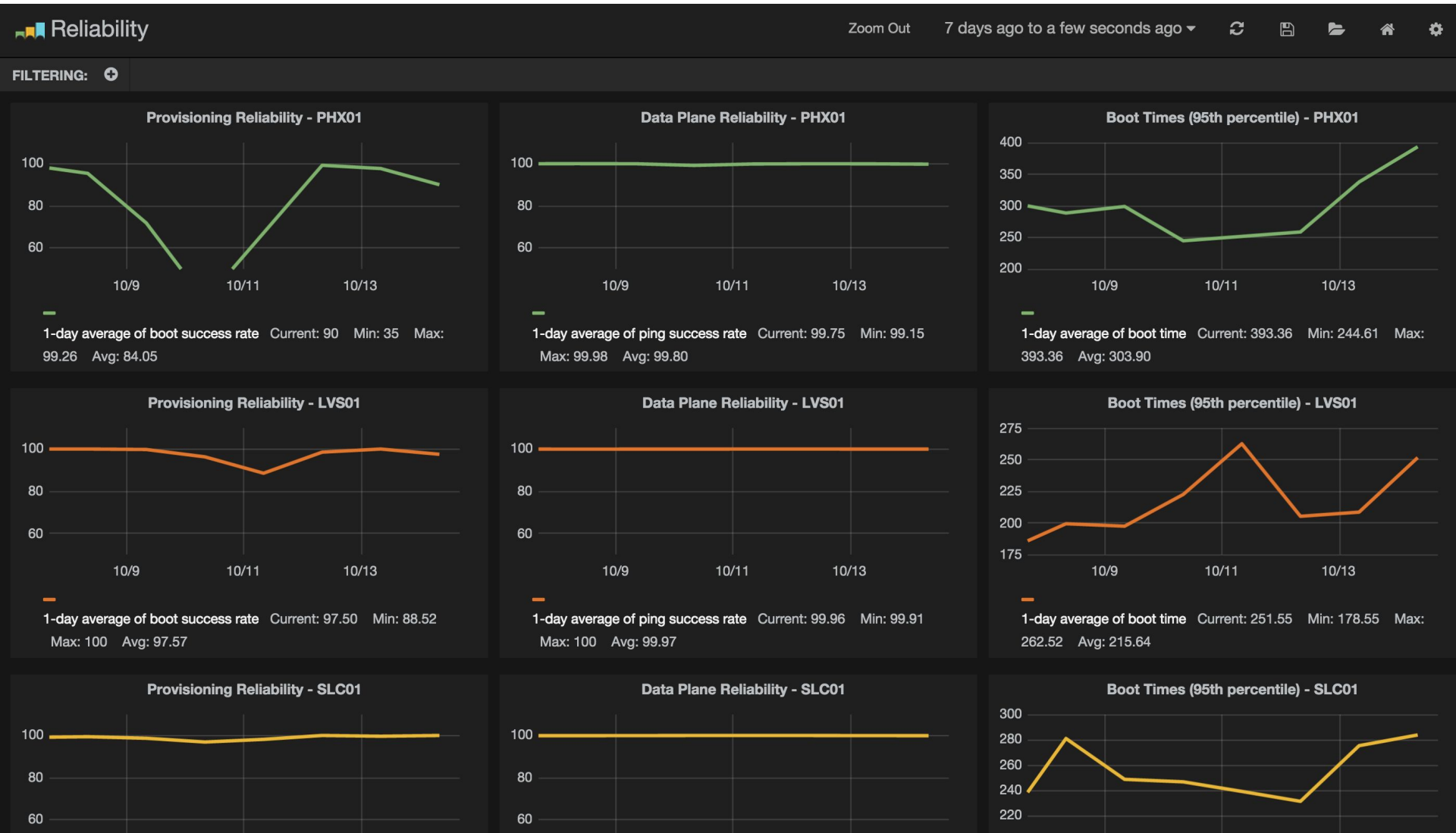
Capacity and Utilization

Compute Capacity - Available VM	
Network Capacity - Available IP	Data computed by StackMetrics
Block Storage Capacity and Utilization	
C3 BM Capacity	
Hypervisor Utilization	Generated by StackMetrics based on Zabbix data
Number of Active Instances	Data computed by StackMetrics
SolidFire Metrics	
Swift Usage	Number of accounts, containers, objects, etc.
Tenants and Users	Active tenants and users in different AZ
Database Utilization	Various metrics on Trove

Global Assets

- [BM Global Cold/Warm Cache](#)
- [BM Global Planning](#)
- [BM Global Utilization](#)

Appendix Grafana reliability dashboard -example



Appendix-Kibana reliability dashboard -example

