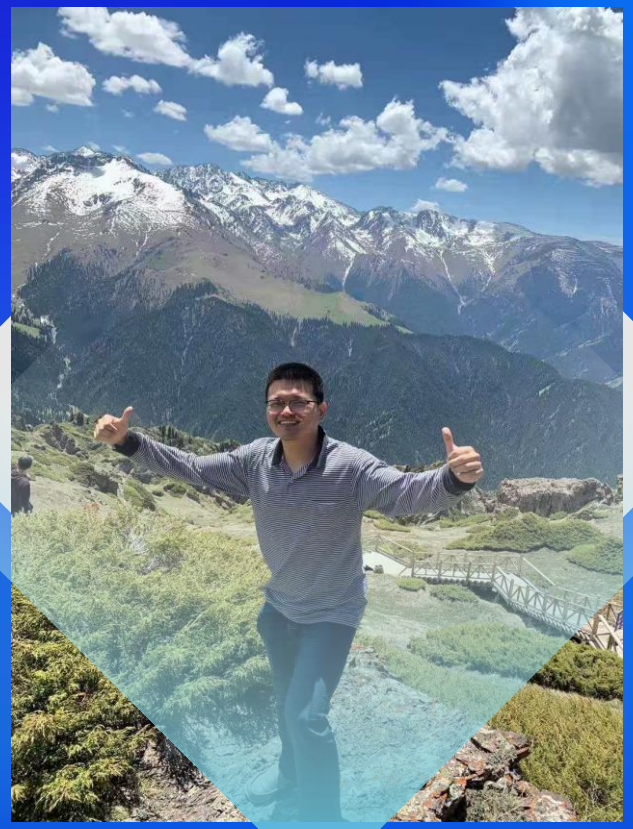
携程云原生基础设施演进之路

周昕毅



周昕毅

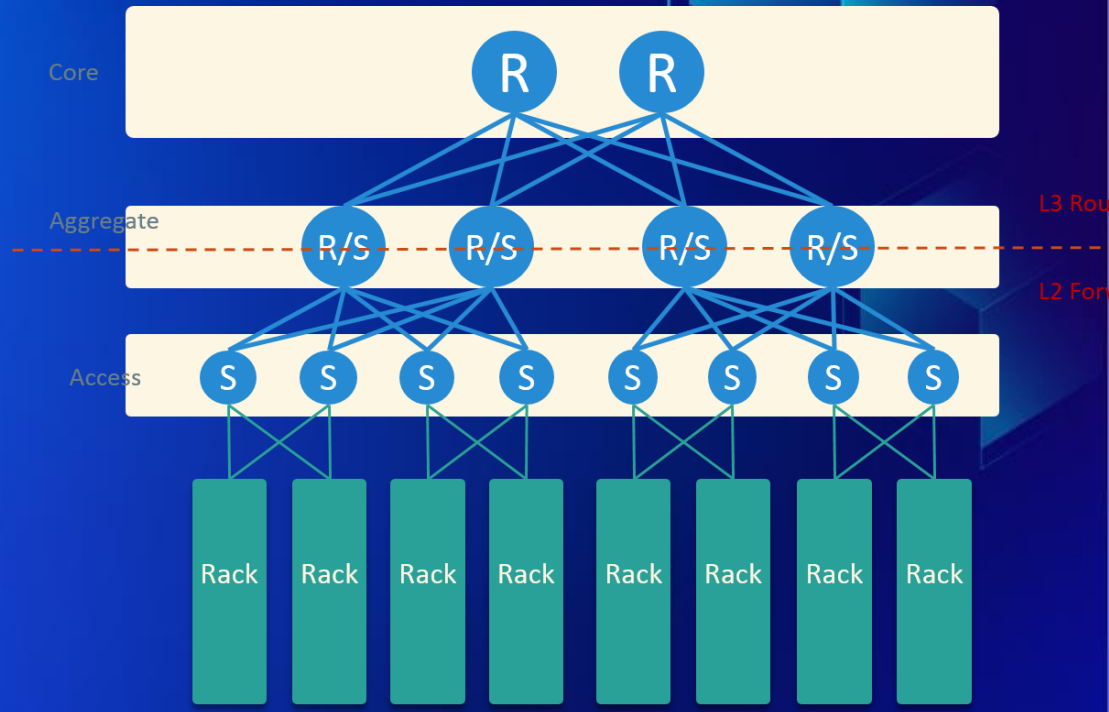
携程云平台高级研发经理

负责携程云网络、分布式存储、K8S集群运维管理

目录

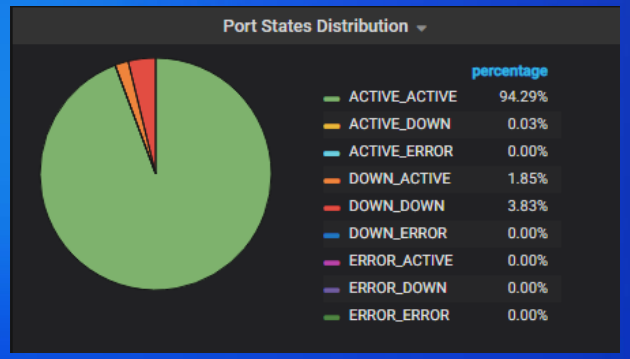
Gen1.0: 2013-2015 OpenStack & IAAS

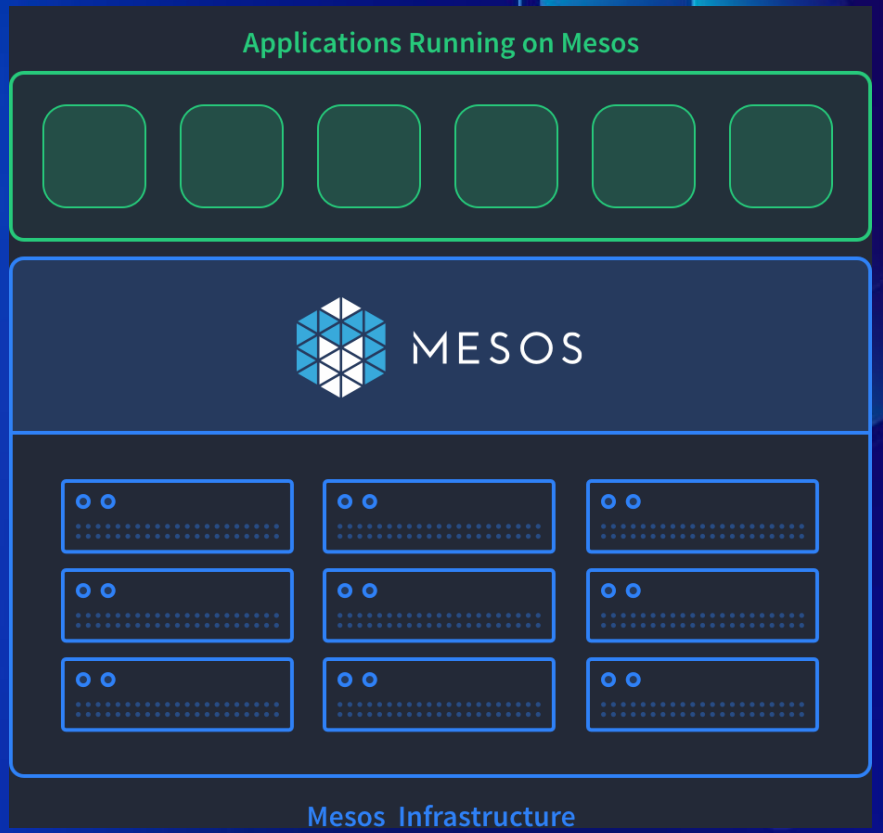
- IAAS based on OpenStack
- VM/BM provision
- Hierarchical network model
- Ctrip internal projects: Tars/CMS/SLB ..
- Focus on efficiency of resource delivery



Gen2.0: 2016-2017 Mesos & SDN & PAAS

- Beginning of Dockerize
- Mesos landing Ctrip
- VM Group migration to Docker Group
- SDN project: H3C / Cisco solution
- VM IP vs Docker IP – remains the same
- Neutron CNC plugin





Gen2.0: Challenges

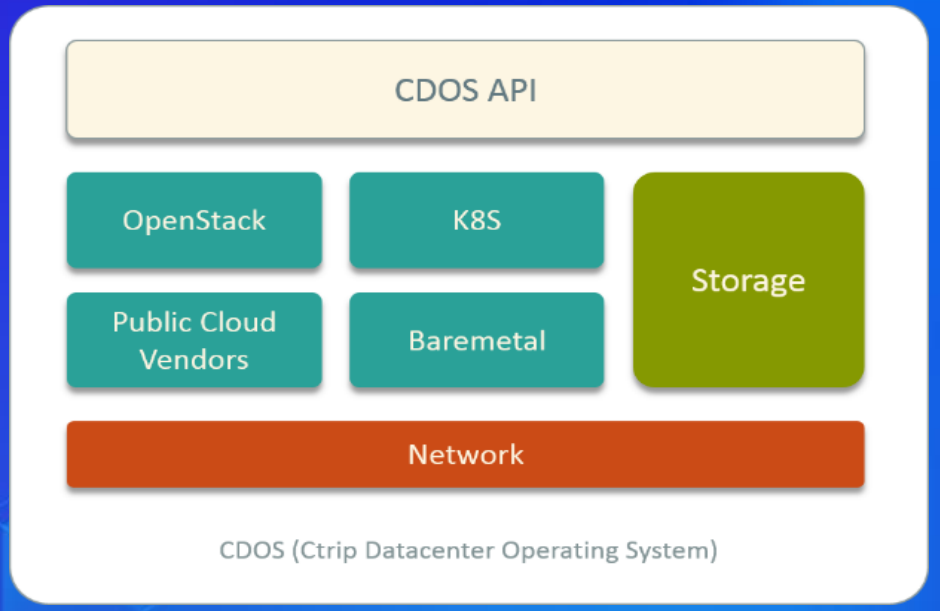
IAAS -> PAAS

Focus on Stability-> Efficiency & Cost

Provision Tools -> Immutable Infra

CPU Utilization -> Capacity Plan&HPA

Gen2.5: 2018 CDOS + Ctrip Paas



Gen2.5: New Challenges

Resource	Application
Multi-Regions	~10k apps
~nK Hypervisor	10K+ prod release per week
IAAS/PAAS	java/nodejs/C#/golang
Private/Public Cloud	Stateful & Stateless

Ctrip Cloud - Challenges Never Ends

Hyper’s CPU/Memory/Disk -> lxcfs

Defunct Process in side Docker -> Reboot?

CPU Throttle Time -> CPU set

100x System calls -> Kernel dead lock

Docker live restore & Docker Daemon hang

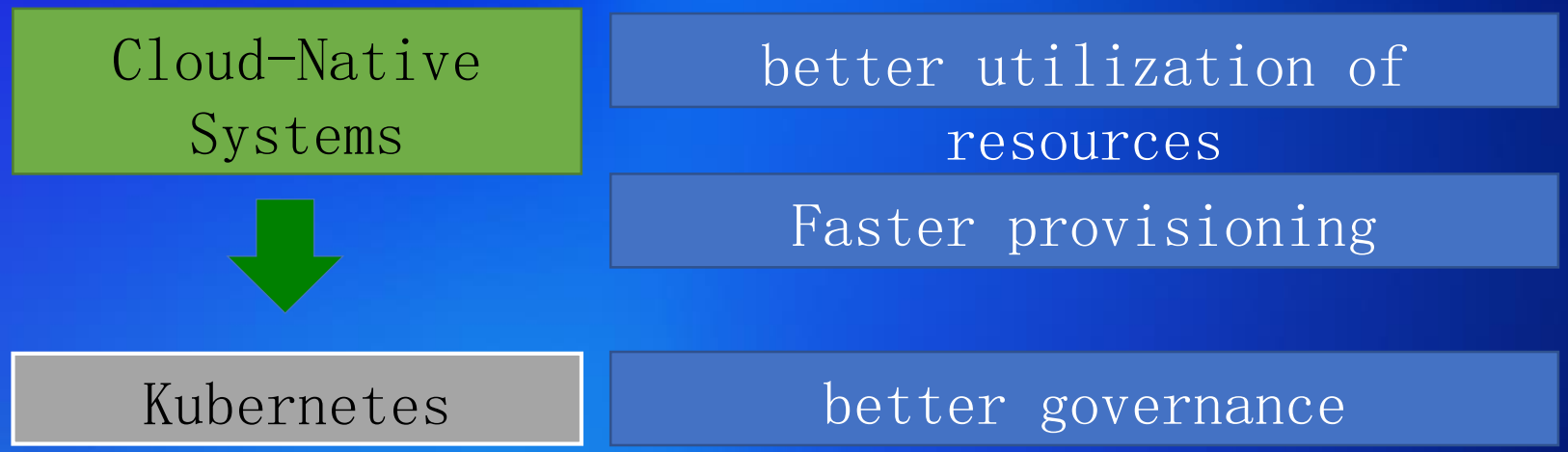
Central IPAM(neutron) become bottleneck

Ctrip Cloud – Infrastructure Stability

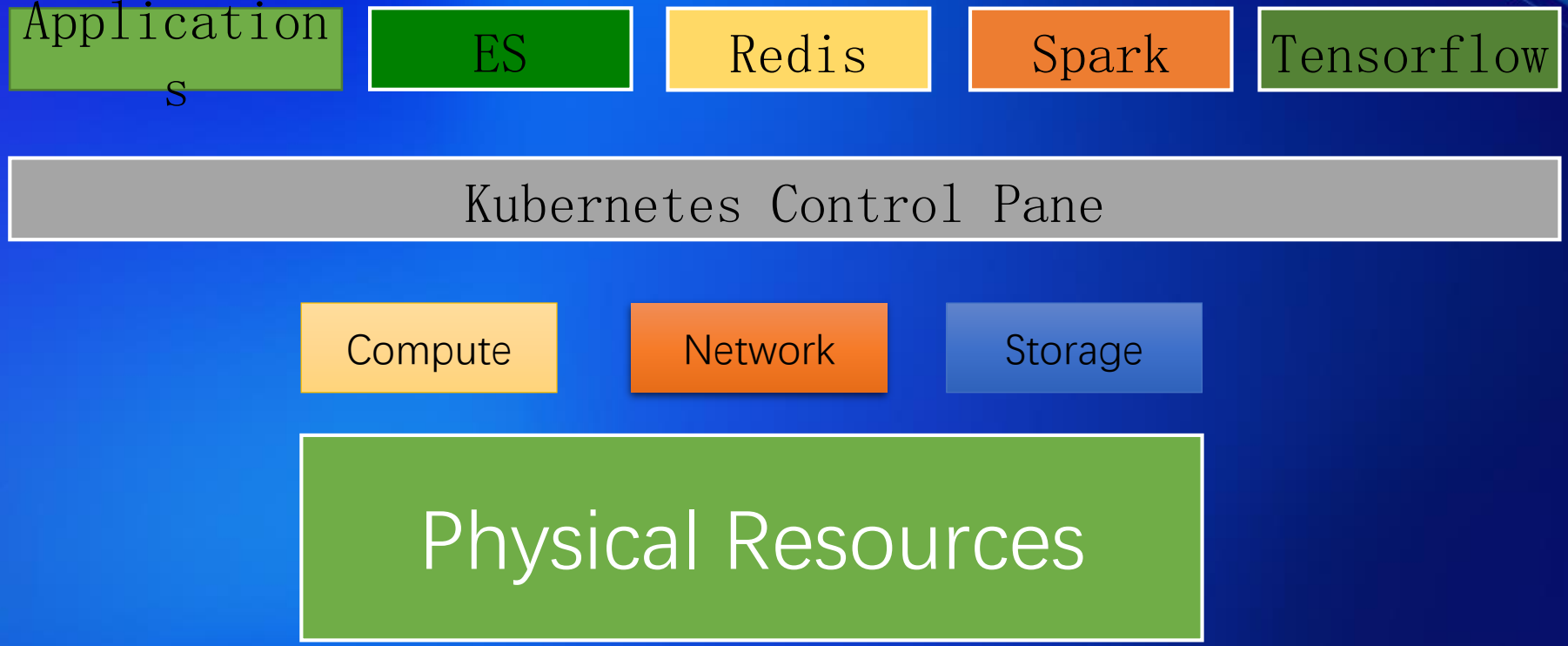
Hypervisor	Centos7.4 / Docker18.09 / Kernel4.14
Docker Image	Harbor / BuildPortal / Ceph
Resource Limit	CPU Quota / CPU Set / Network Qos
Scheduler	OpenStack -> Mesos -> Kubernetes

Gen3.0: Kubernetes & Cloud Native Infrastructure

- Jim Zemlin: Kubernetes is becoming the Linux of the cloud

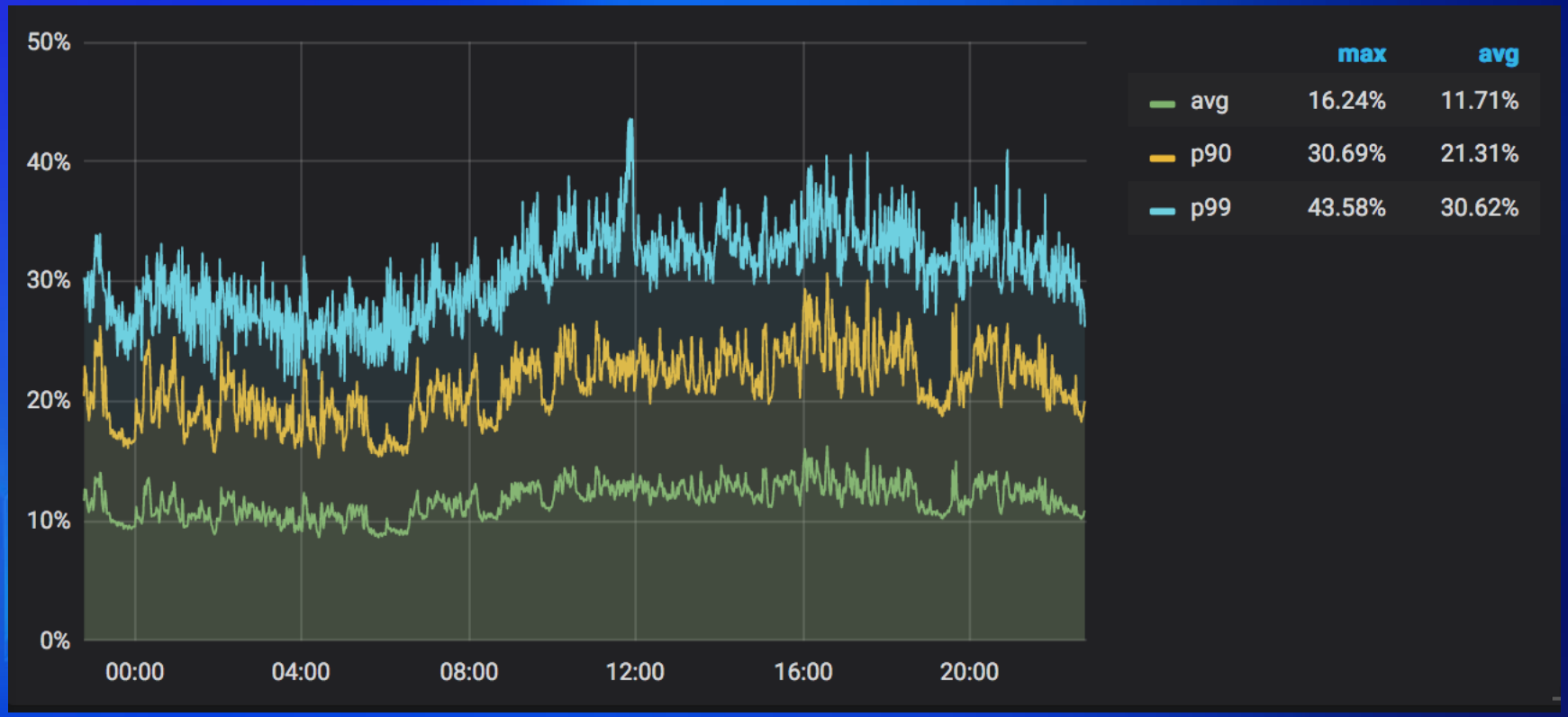


Better utilization of Resources (1)



Better utilization of Resources (2)

- Online Applications running together with offline Job
- CPU utilization under full control



Better utilization of Resources (3) – Capacity

Better utilization of Resources (4)

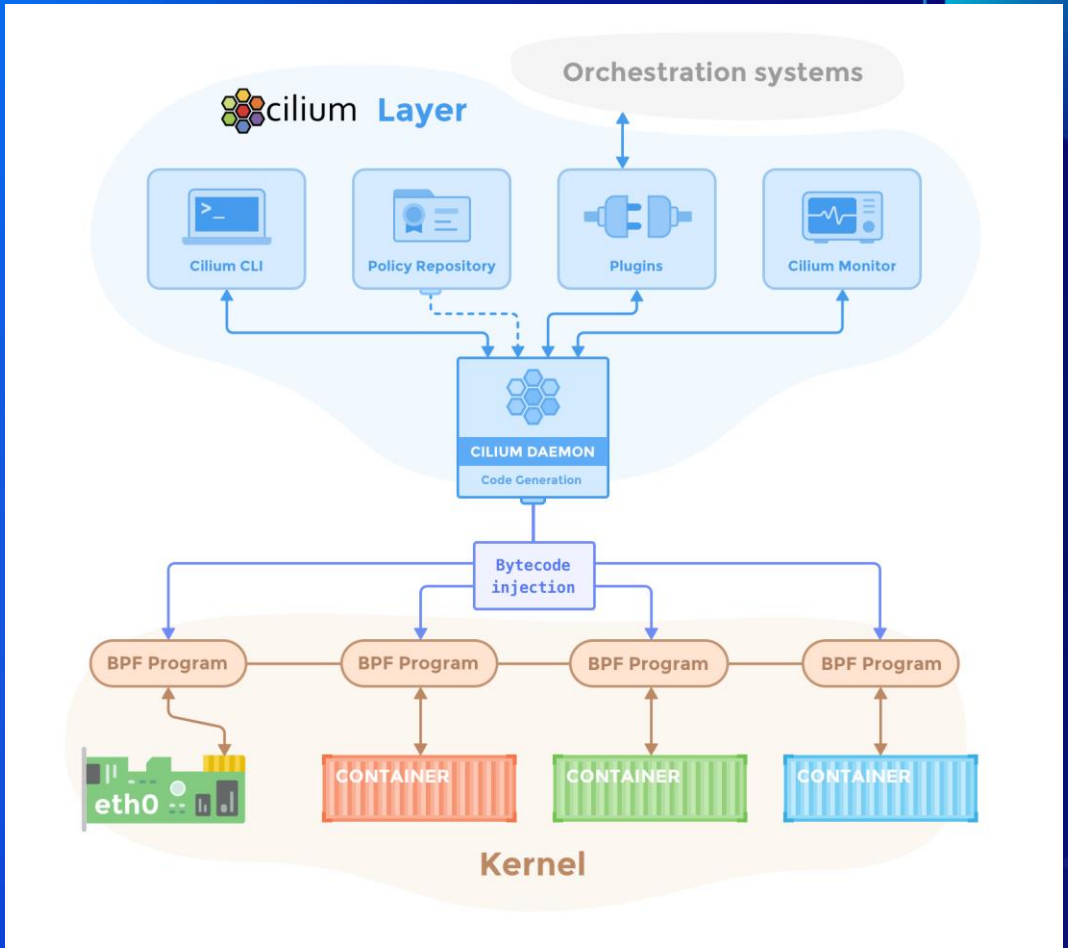
Cloud Health Checker

Faster Provisioning (1)

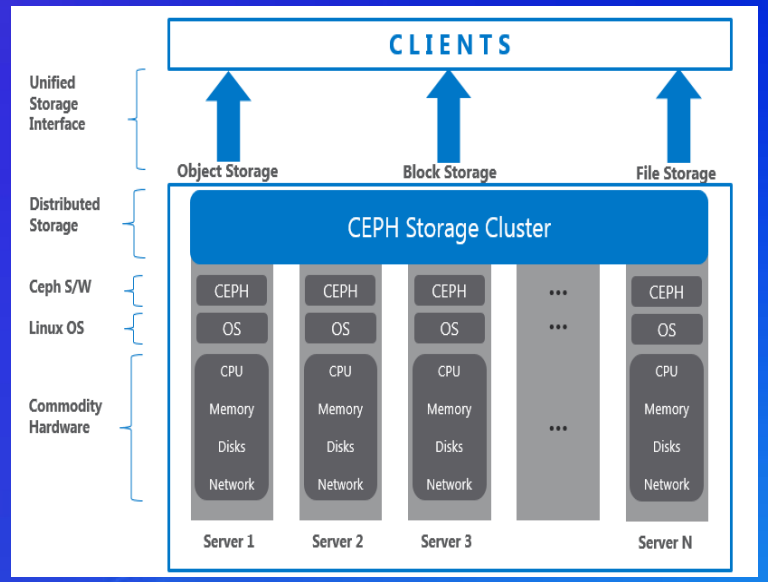
- Harbor Federation
- Image prefetch/dispatch
- Jenkins on K8S

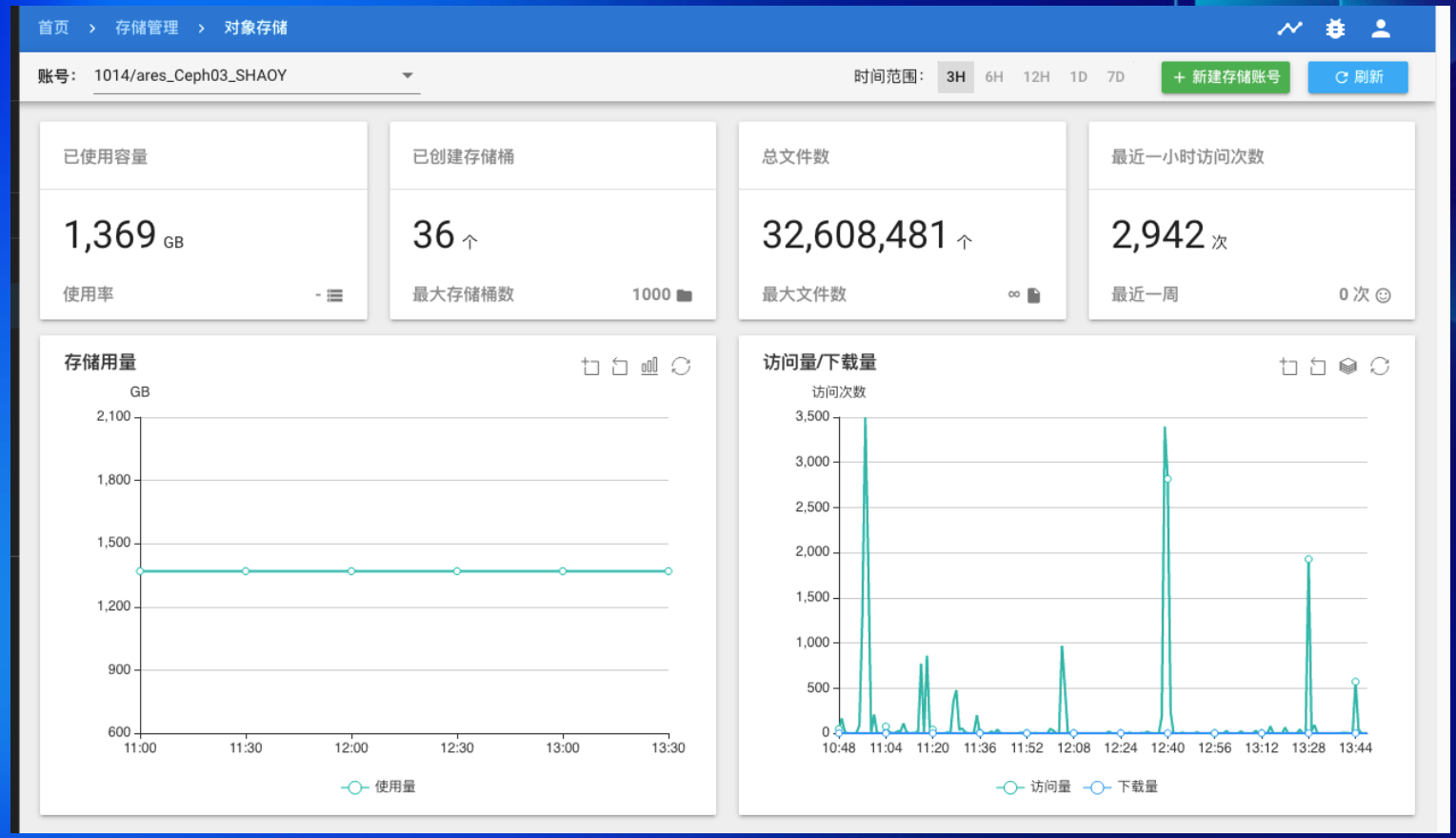
Faster Provisioning (2) – Network Bottleneck

- Cloud Native Network Design
  - High Performance
  - Local IPAM
  - eliminate performance bottle neck
  - L4-L7 Network Policy
  - Routable instance IP
  - Ease of Develop & Operation



Faster Provisioning (3) – Cloud Storage





Better governance(1) – Infrastructure As Code

Infra Tools Owner

Code

DockerFile

HelmCharts

Cloud Platform

CI/CD

Helm Manager

Service Mesh

Prometheus

CoreDNS

Auto Scaling

A/B Testing

Multi Region

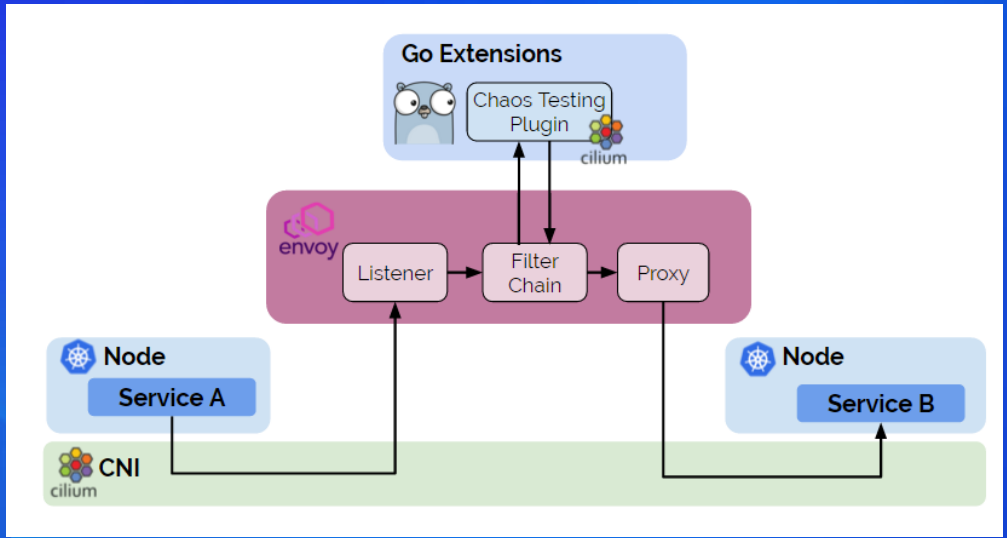
Node 1

Node ..

Node N

Better governance(2) – Cilium on K8S

- Network Control as Code (Chaos Engineering)
- Kernel & Network layer Monitoring
- Enhanced Security Capability – (Application level Access Control)




```yaml
apiVersion: "cilium.io/v2"
kind: CiliumNetworkPolicy
[...]
specs:
- endpointSelector:
  matchLabels:
    app: myService
  ingress:
  - toPorts:
    - ports:
      - port: "8000"
        protocol: TCP
        l7proto: chaos
        l7:
        - probability: "0.8"
          delay-response: 50ms
        - probability: "0.2"
          delay-response: 1s
```

Thanks For Watching

本PPT来自2019携程技术峰会

更多信息请关注“携程技术中心”微信公众号~