

腾讯基于kubernetes的企业级容器 云实践

罗韩梅

腾讯 专家工程师



全球技术领导力峰会

Geekbang | TGO 鲲鹏会
极客邦科技

500+ 高端科技领导者与你一起探讨 技术、管理与商业那些事儿



🕒 2019年6月14-15日 | 📍 上海圣诺亚皇冠假日酒店



扫码了解更多信息

自我介绍

罗韩梅，腾讯 T4 专家工程师，2009 年加入腾讯，现任数据平台部容器云开发组组长。拥有多年分布式系统研发经验，对大数据、云计算、容器等有深刻理解。从事过自研容器云平台，大数据云平台，以及面向公司内部外的通用容器云平台，从无到有，从自研到开源生态，从公司内部平台到同时面向To B市场。目前专注于容器云平台领域，负责腾讯企业级容器云平台。

自研容器云平台



2009年-2013年

腾讯大数据云



2014年-今

通用云平台

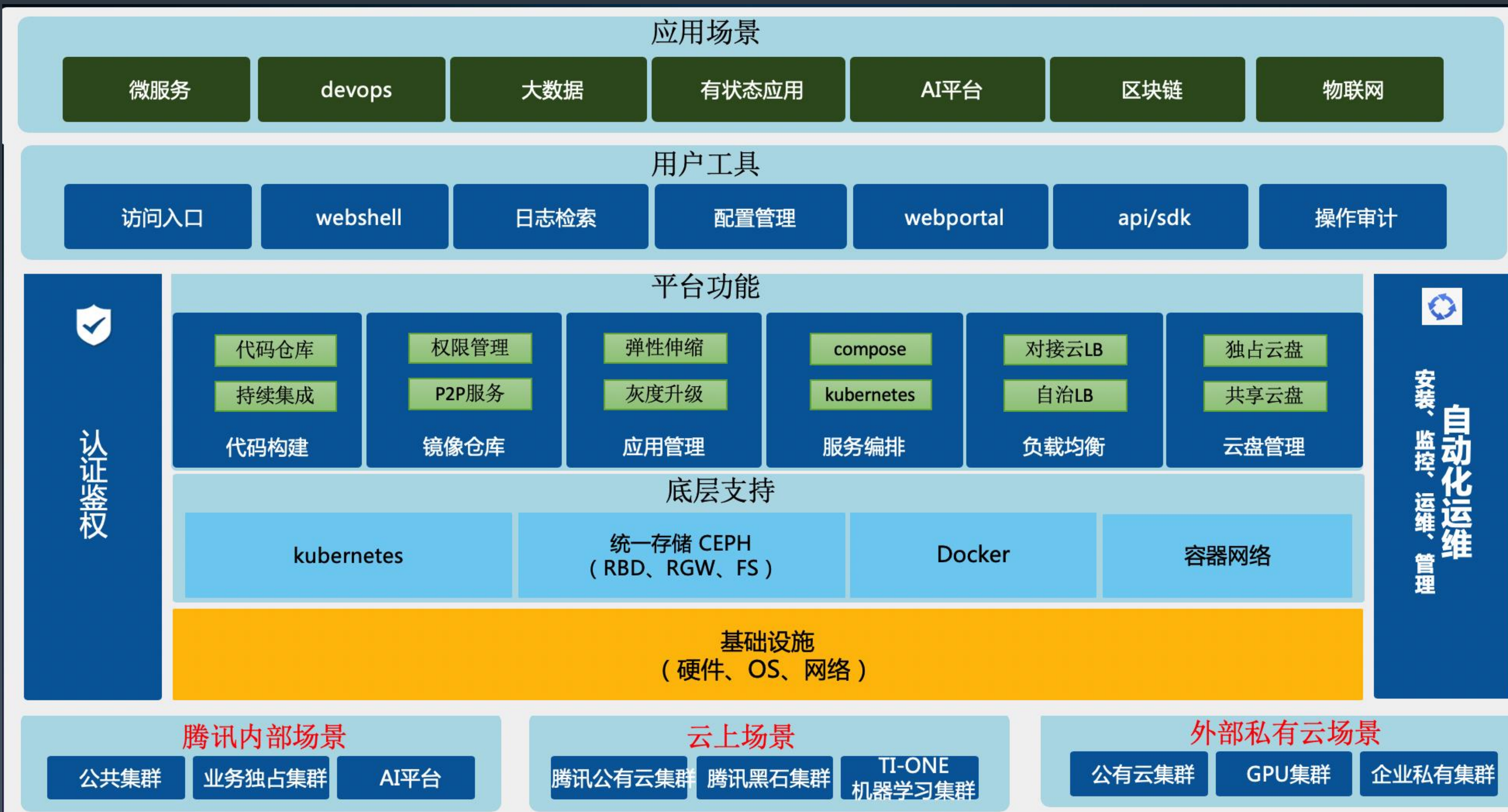


2015年-今

目录

- 架构简介
- 企业级容器云解决方案
- Next

企业级容器云架构



产品功能

集群管理

集群部署

主机管理

扩容

缩容

服务管理

停止

重启

集群监控

资源监控

主机监控

镜像服务监控

存储服务监控

etcd监控

告警管理

告警配置

告警记录

系统日志管理

系统日志

Docker日志

规划管理

用户管理

业务管理

实例配置管理

应用全生命周期管理

CI/CD

关联代码仓库

创建项目

持续集成

构建镜像

构建记录

异常定位

交付中心

镜像仓库

编排模板

个人镜像

业务镜像

公共镜像

Kubernetes
编排

compose
编排

应用管理

Stack管理

应用管理

实例(容器)管理

创建编排

创建应用

删除应用

新增实例

删除编排

停止应用

启动应用

删除实例

配置管理

ConfigMap

Secret

存储管理

云硬盘

快照

创建云硬盘

销毁云硬盘

挂载云硬盘

扩容云硬盘

创建快照

删除快照

存储quota管理

从快照创建云硬盘

应用自动化运维

自动扩缩容

主动扩缩容

灰度升级

操作记录

事件管理

控制台

访问入口

绑定域名

负载均衡

应用监控

应用告警

应用日志

网络配置

Quota准入

存储对接

多集群视图

多业务视图

统计概览

企业级容器云解决方案

易用

可靠

安全

企业级
场景

通用

能力扩展

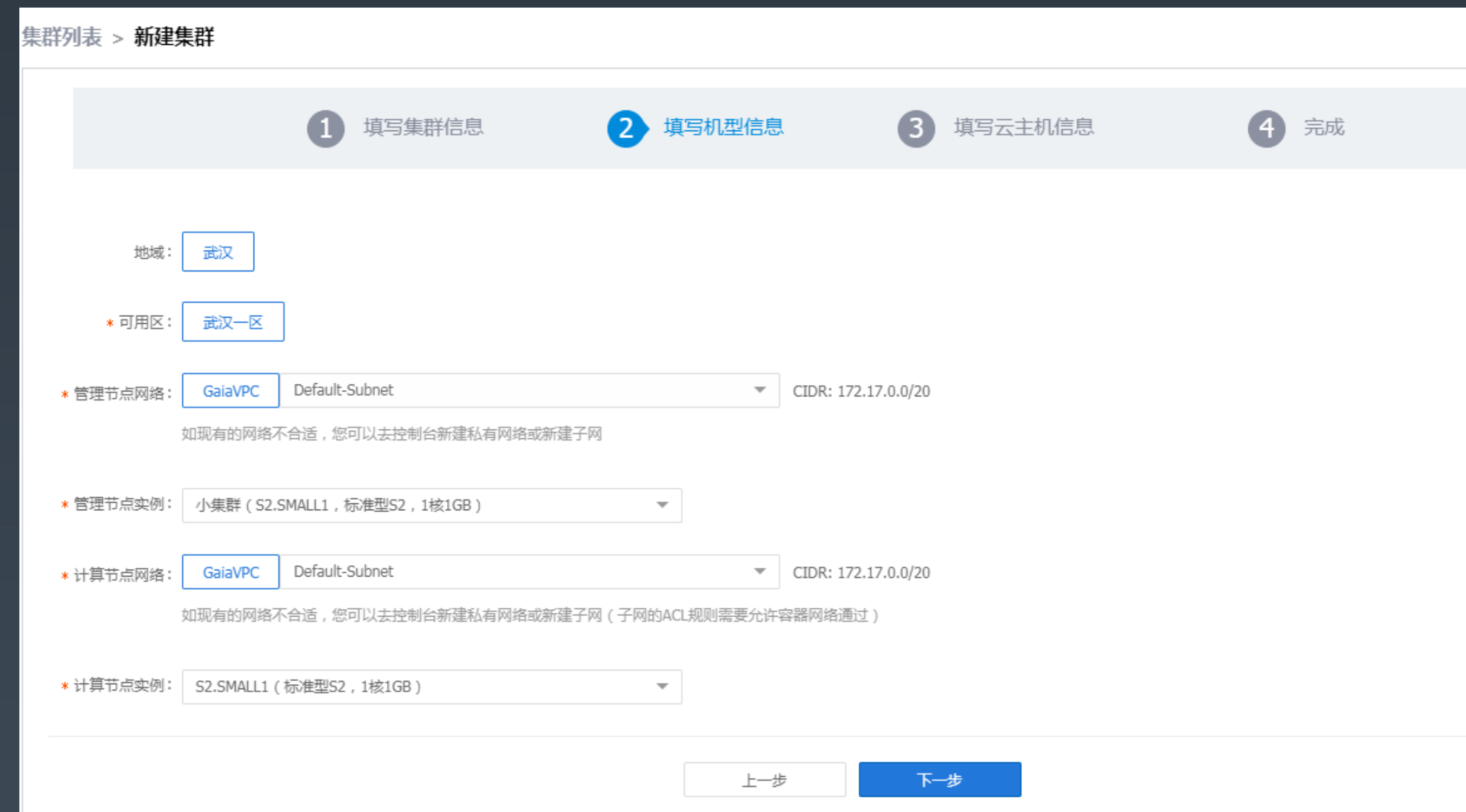
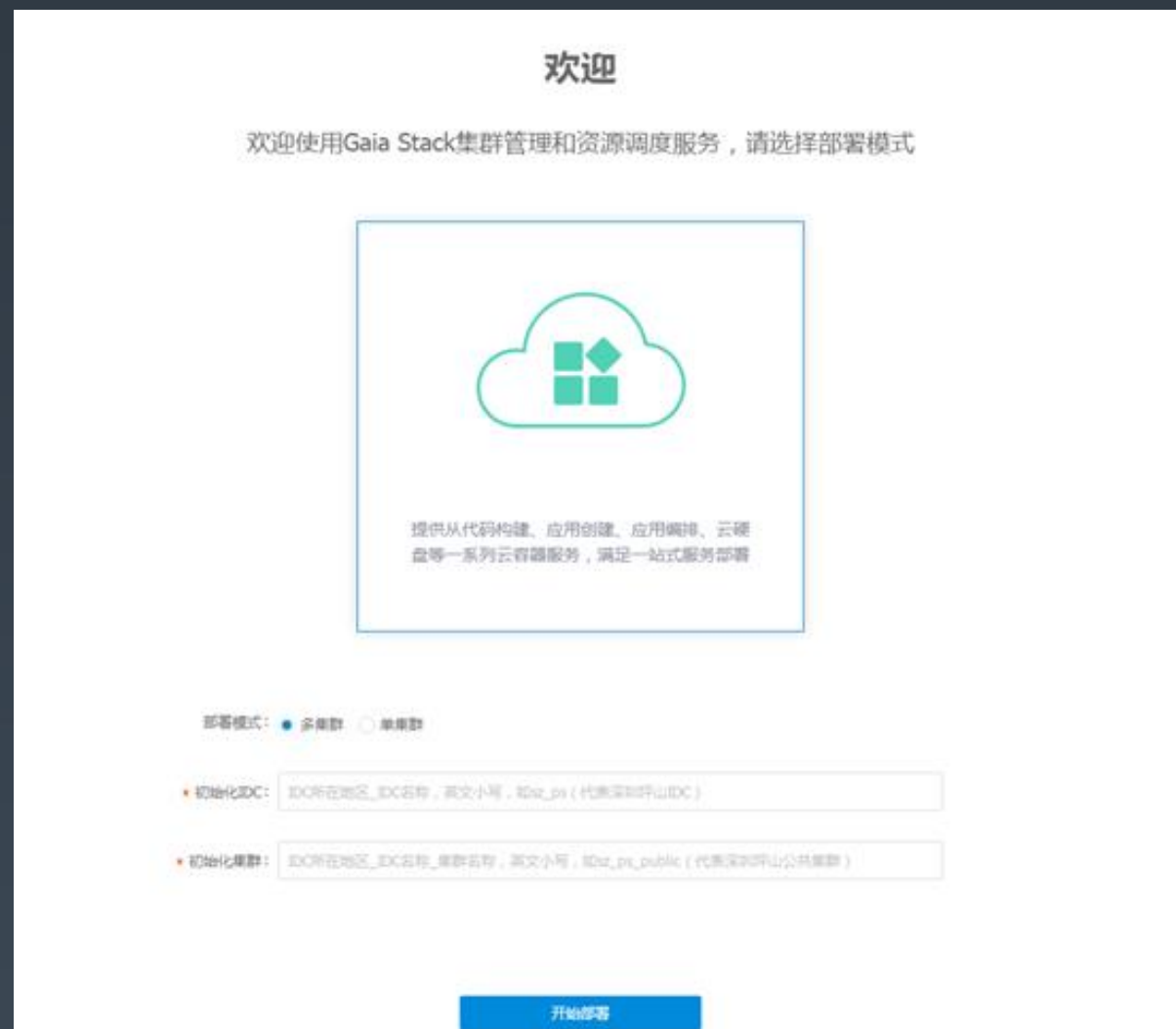
成本

性能

生态

易用

- 全组件自动化部署、统一配置管理、多策略灰度升级
- 提供可视化、自动化的运维能力，降低使用者的人力成本和学习成本



可靠



平台容灾

- 所有组件无单点;
- 平台本身支持**热升级**;
- 组件自身HA机制, 如docker;
- **多地域多可用区**的容灾设计
- 管理机挂掉: 对应用无影响
- 计算节点挂掉: 跨机迁移

- 举例: 1.4升级1.9版本
- Pod Hash发生变化
- Container名称发生变化, 点分隔改为了下划线分隔
- 容器标签发生变化

pause容器的标签io.kubernetes.container.name=POD改为
io.kubernetes.docker.type=podsandbox
io.kubernetes.container.restartCount改为
annotation.io.kubernetes.container.restartCoun

- Cgroup目录结构发生变化, 新增Pod层级



应用容灾

- 健康探针
- ① 存活探针
- ② 就绪探针
- 负载均衡
- 重启机制
- ① **区分异常原因**
- ② **本地重启/跨机重启**
- 黑名单机制



数据容灾

- 集群核心数据的备份和恢复
- ① Etcd
- ② 核心数据库
- 云盘机制保护应用数据

企业内部各个集群灰度运营。

可靠

一次现网事故

群聊的聊天记录

mavisluo: 403b
mavisluo: 有机器挂掉，我们一起看下
mavisluo: @bauerzhou(周俊清) 发下是哪台? ...

机器没有挂，是有个容器没有把日志挂出来，而且写了400多G日志，导致磁盘满

明白，谢谢，我们马上找vpc的产品去搞

这里应该是网关的东西

其实这个问题本质上是社区的原生k8s的问题，因为没有做磁盘容量管理。

有资源维度没有隔离，就会导致一个容器异常可能影响整个节点

不过gaiatack做了磁盘容量管理了，只是需要用户指定一下作业属性，需要把日志目录属性填一下，整个是k8s没有的属性，所以后面还需要tce这边加上整个属性。否则还有可能导致整个问题。

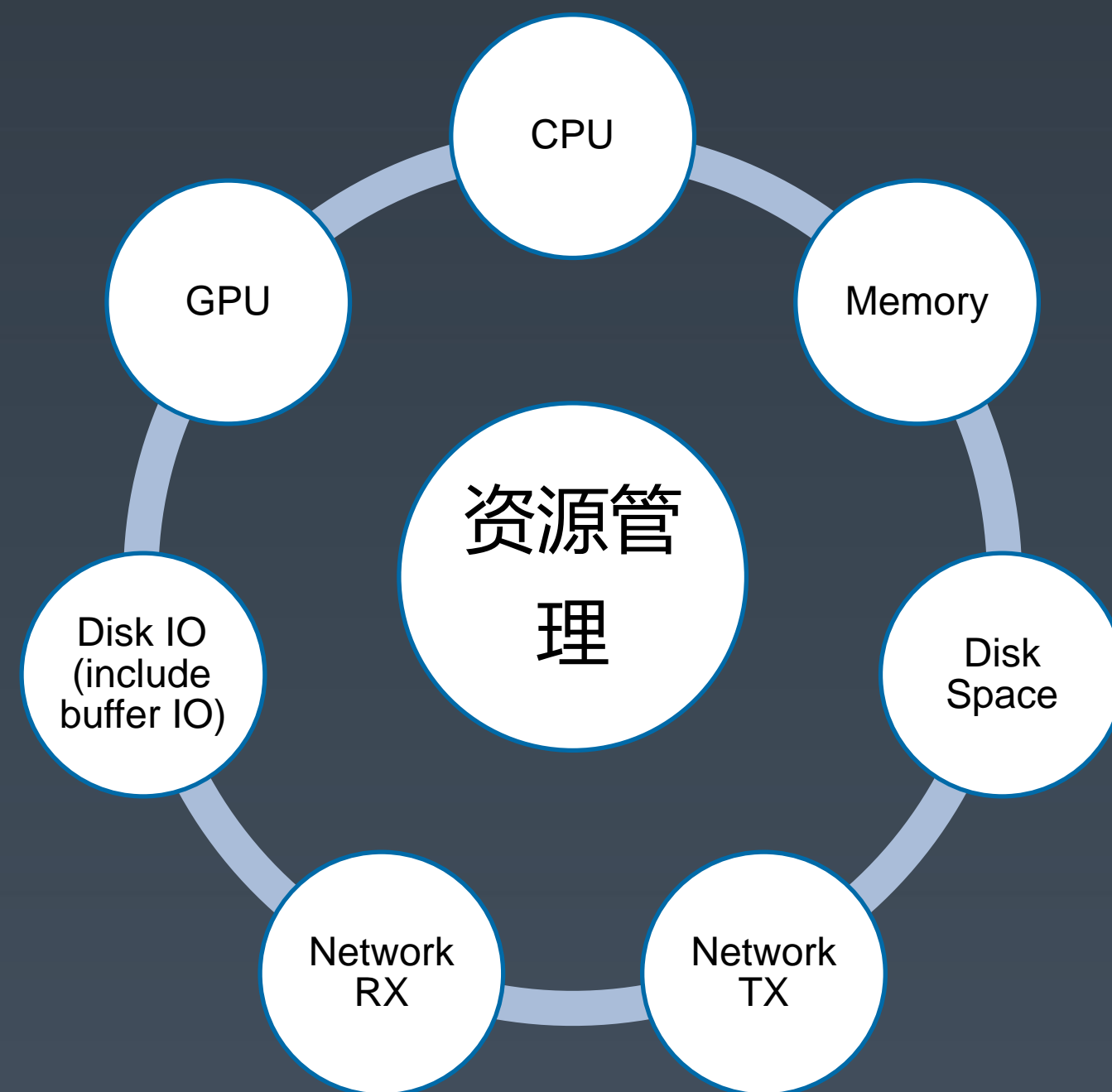
一个用户需求

背景：广告业务，8个集群，4个在线集群，4个离线集群，分布在四个地区：北京、天津、成都、深圳。

需求：减少机器，降低成本。

手段：在线离线集群做合并。

问题：容器只能管理CPU和内存，不能对网络和磁盘IO做管理，导致在线应用受离线业务影响。



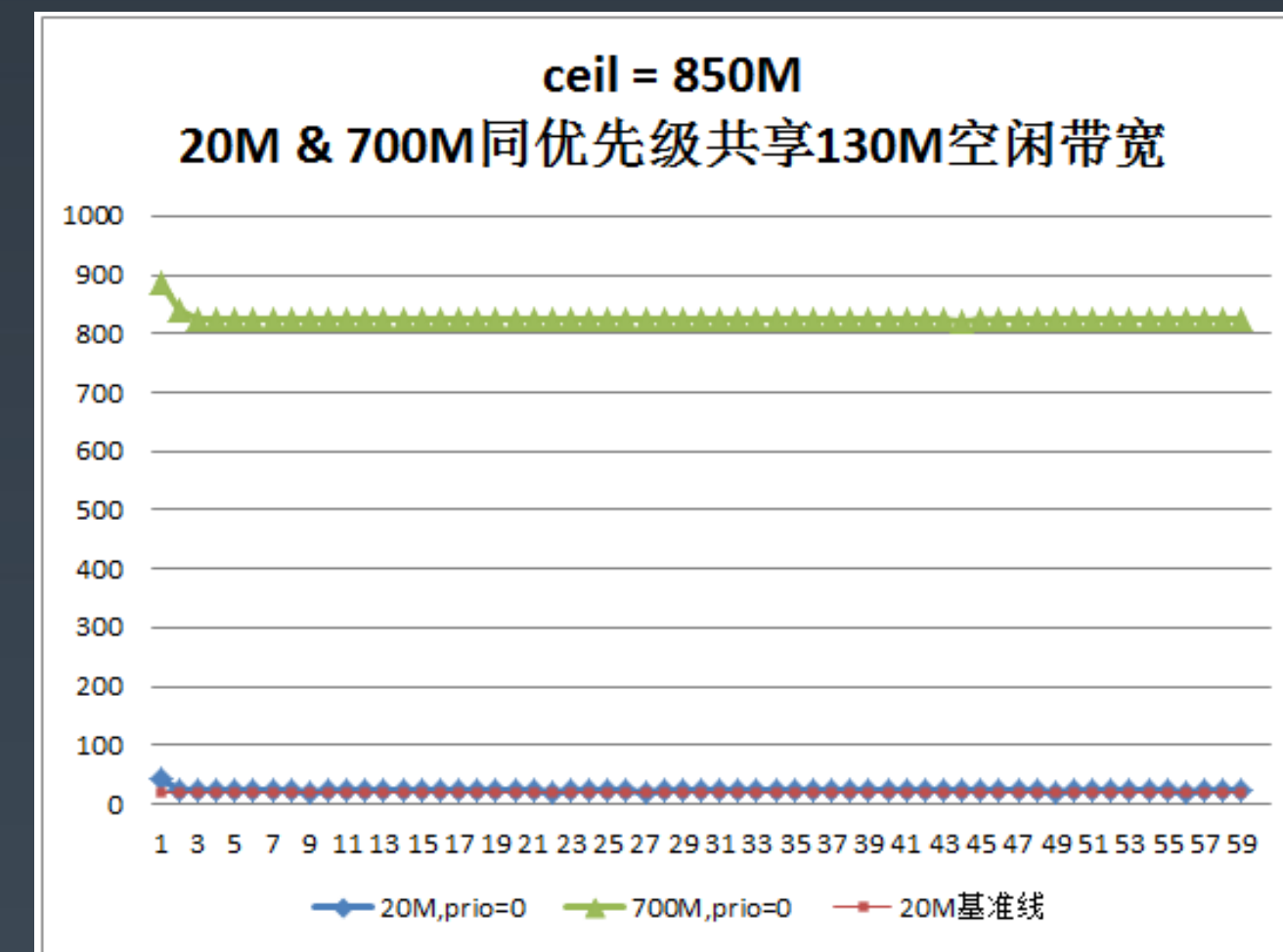
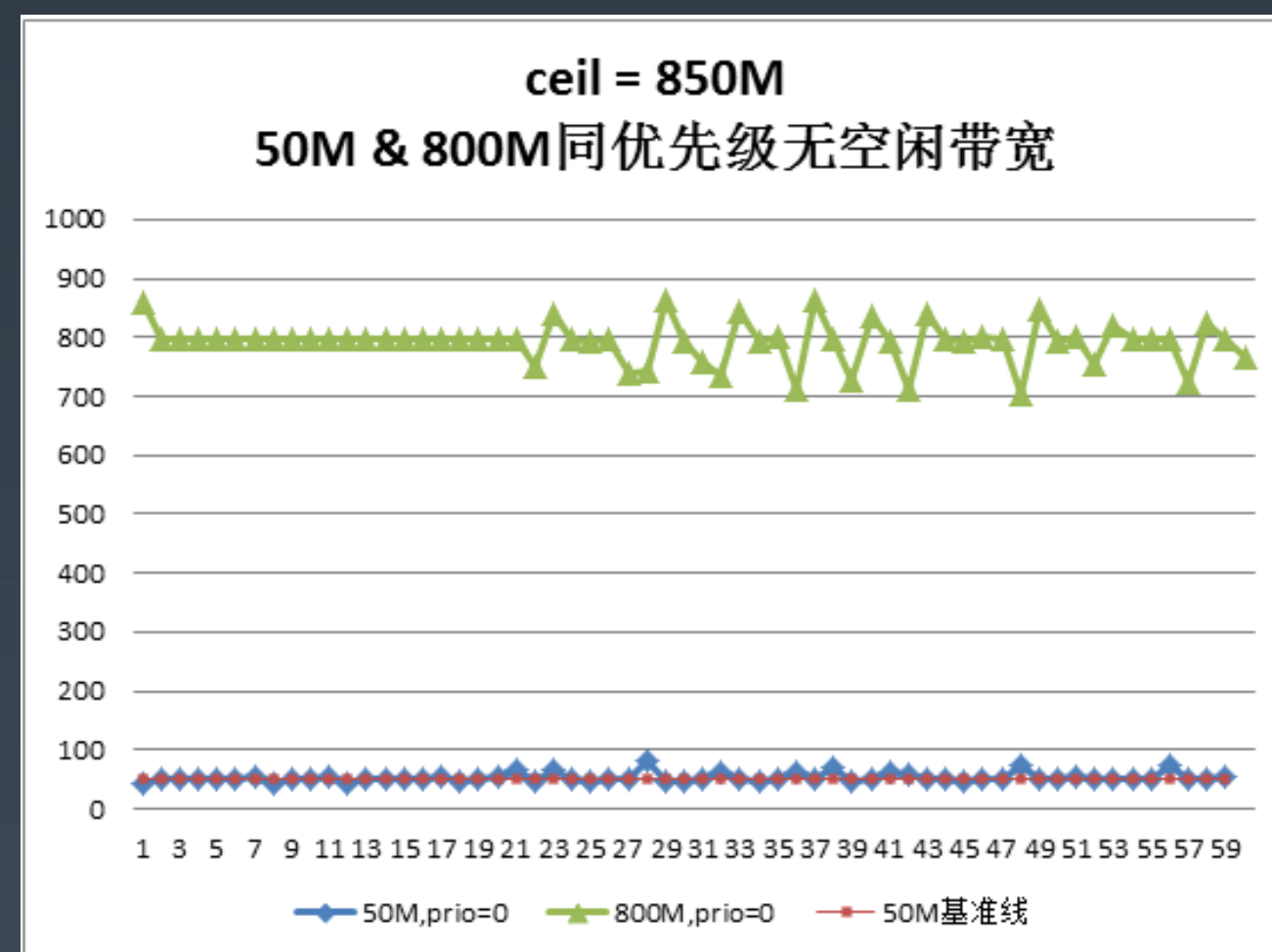
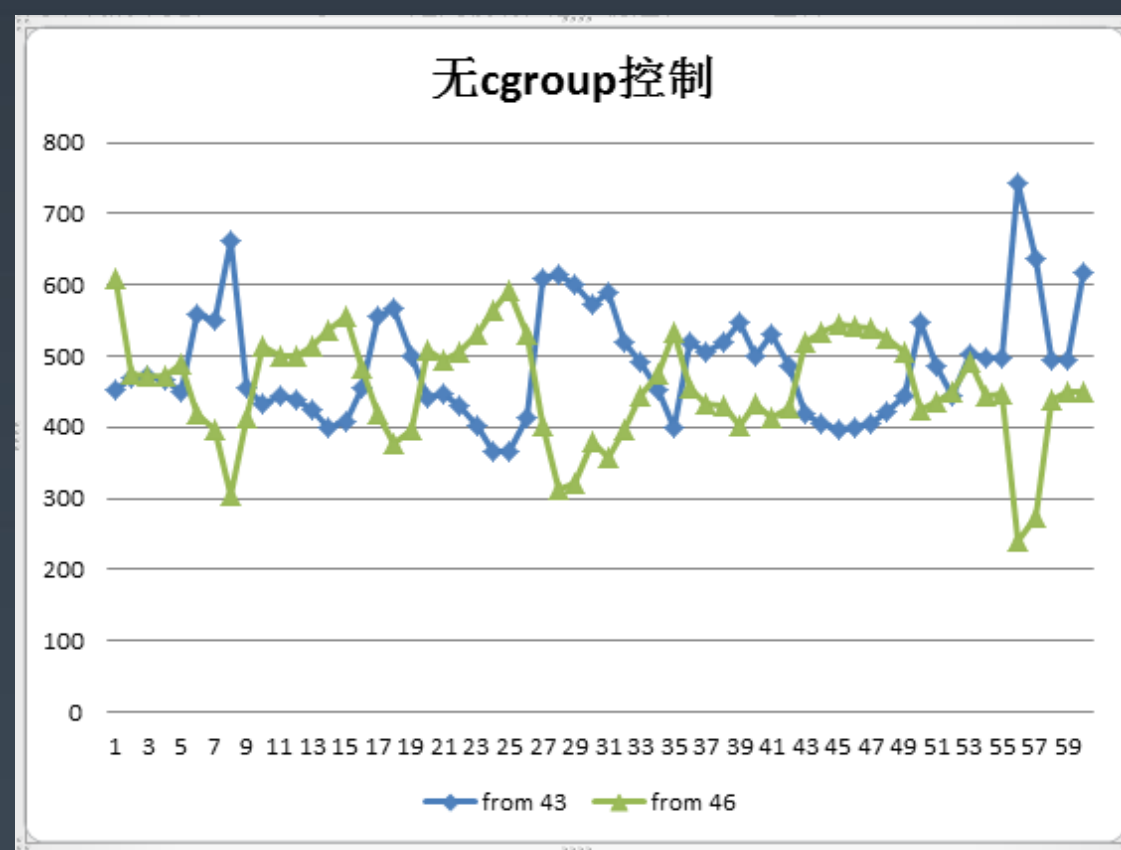
资源使用率

总计已使用

CPU核	0.3/1 个	
内存	384/1000 M	
本地磁盘	3/10 G	
普通云盘	0/0/100 G	
共享云盘	0/0/100 G	
网络出带宽	3/100000 Mbit/s	
GPU	0/100 个	
GPU内存	0/1024 MB	

可靠

下图是两个进程都拼命争抢网络带宽时的效果。两个进程的带宽和时延都得不到任何程度的保证。



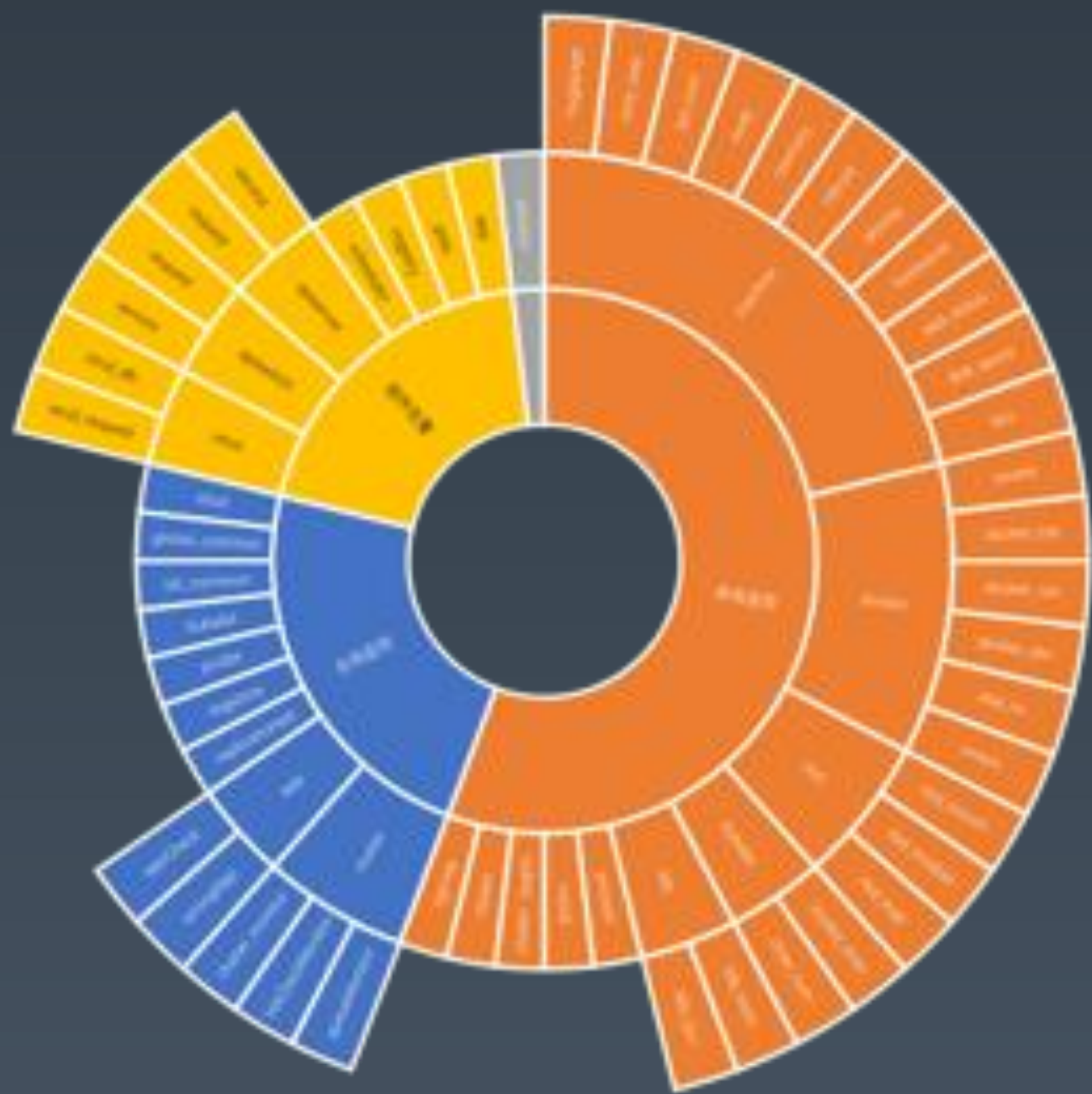
设计目标

- 在某个cgroup网络繁忙时，能保证其设定配额不会被其他cgroup挤占
- 在某个cgroup没有用满其配额时，其他cgroup可以自动使用其空闲的部分带宽
- 在多个cgroup分享其他cgroup的空闲带宽时，优先级高的优先；优先级相同时，配额大的占用多，配额小的占用少
- 尽量减少为了流控而主动丢包

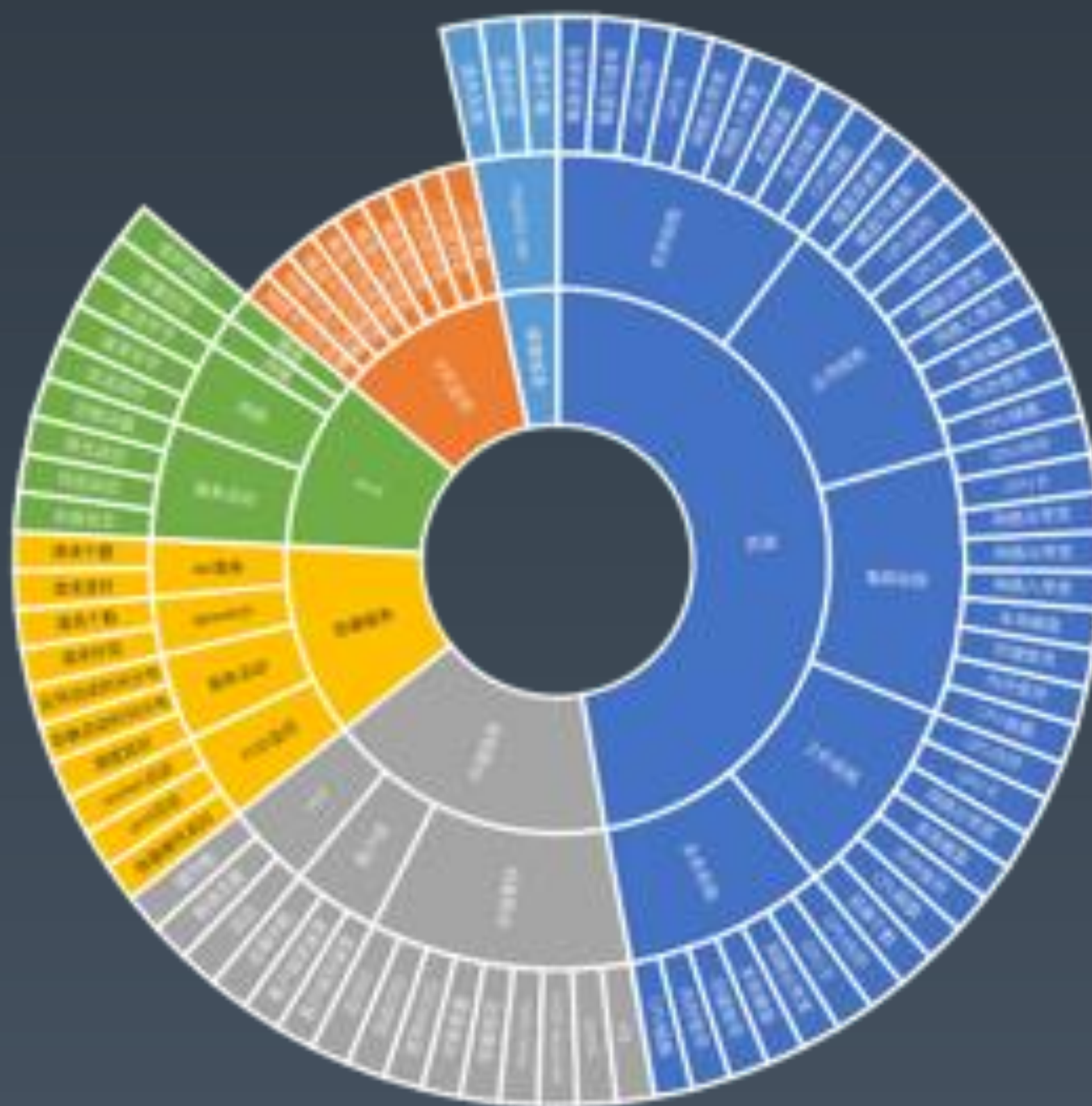
- 队列：不增加队列，对每个报文直接在正常代码路径上进行决策
- Cgroup区分(标记)：在正常处理流程中，报文查找到目标socket结构之后，根据socket的owner process来确定cgroup
- 报文决策：令牌桶 + 共享令牌池 + 显式借令牌
- 限速方式：ECN标记 + TCP滑窗 + 丢包

可靠

158项告警



87项指标采集



4种告警方式——可随时修改

短信

Email

微信

自定义渠道

通用

自研容器网络解决方案Galaxy（CNI网络插件+调度器插件+控制器），面向所有场景：高性能互联网业务、离线业务、在线离线混合场景、传统有状态服务、公有云...

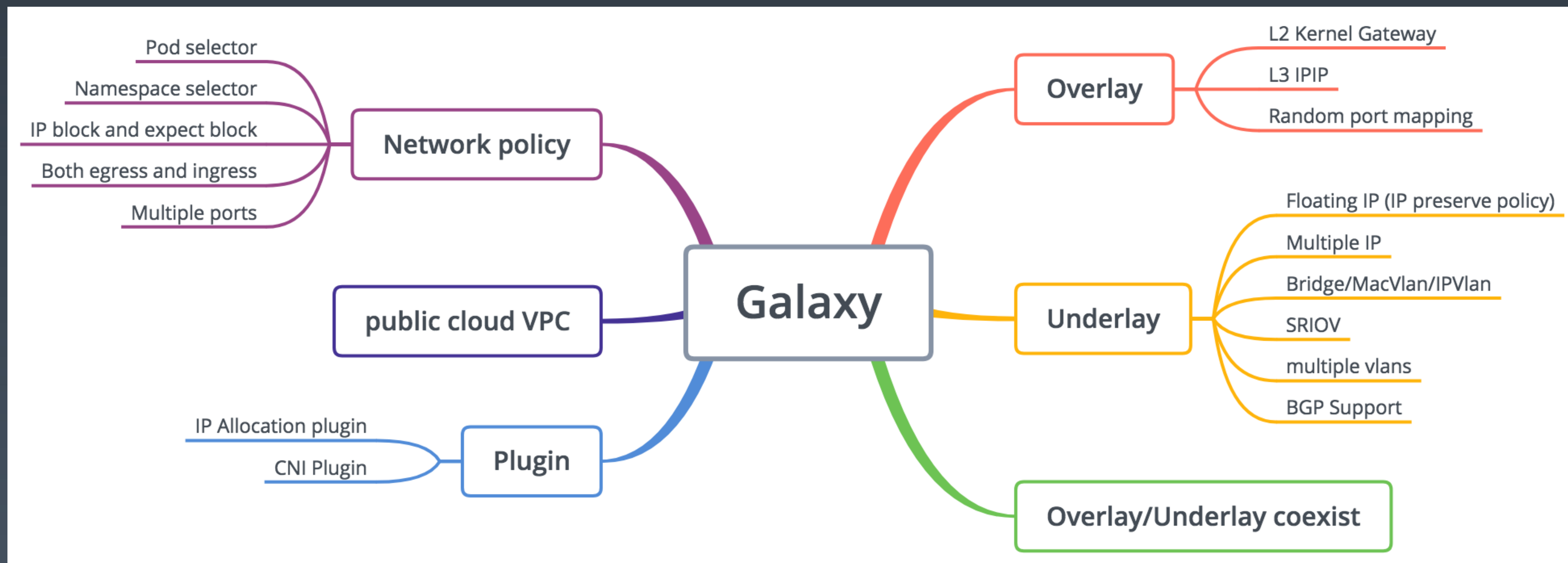
网络模式： Floating IP（浮动IP）

IP漂移： Floating IP（浮动IP）

NAT（端口映射）

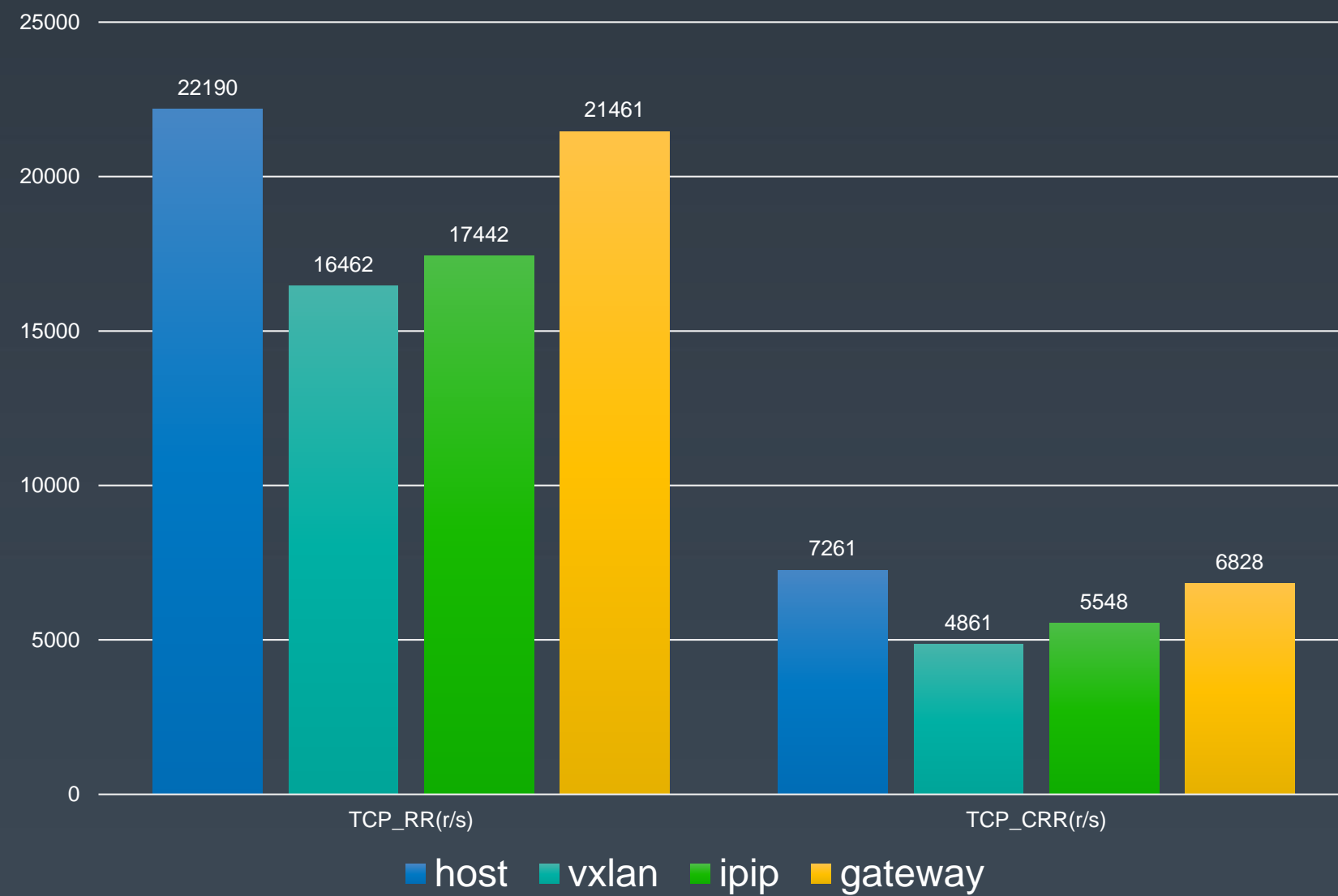
Host（宿主机网络）

- 不同的应用可以选择不同的网络模式
- 同一主机的不同容器可以选择不同的网络模式

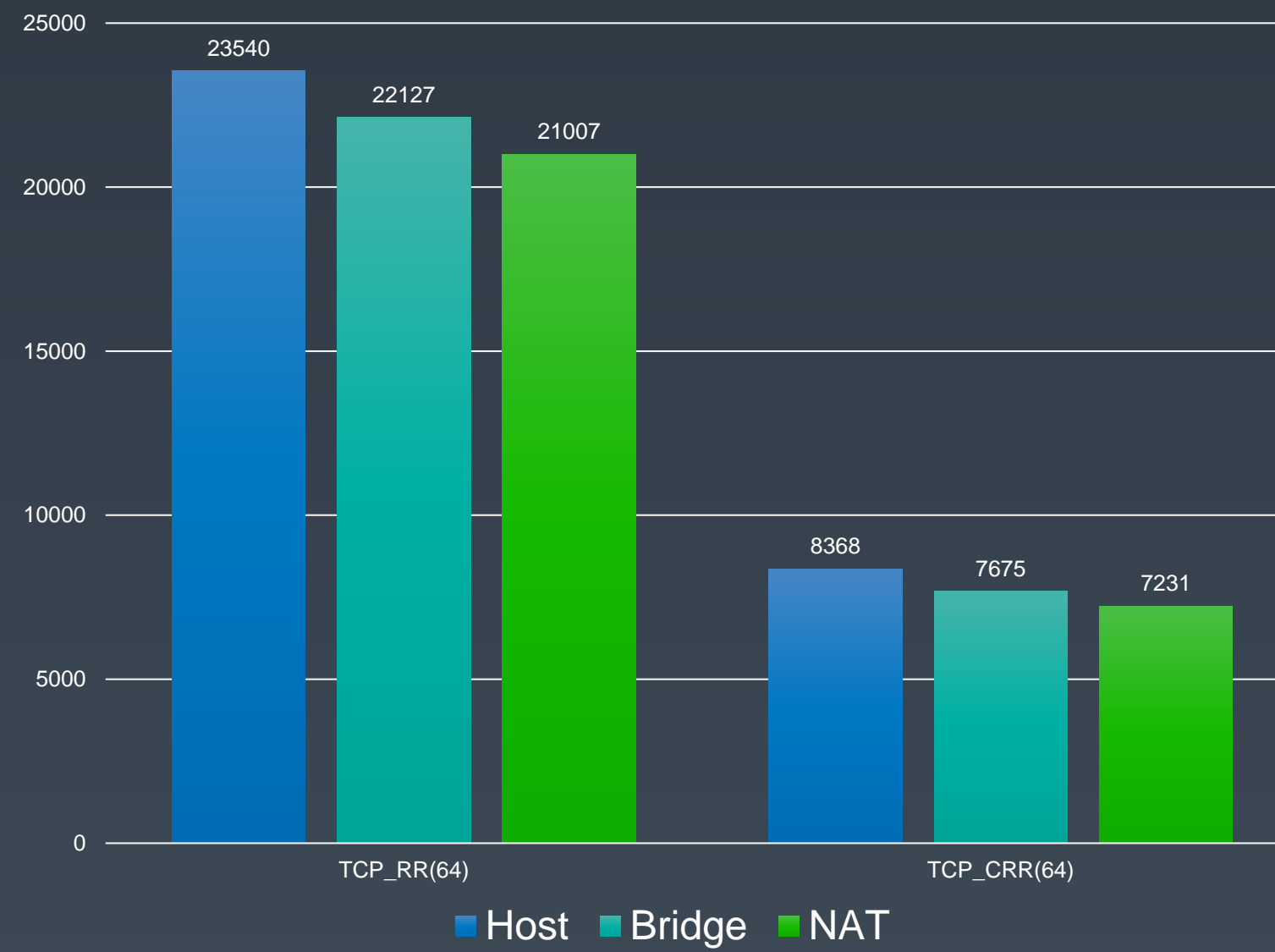


性能

Overlay方案性能



Underlay方案性能

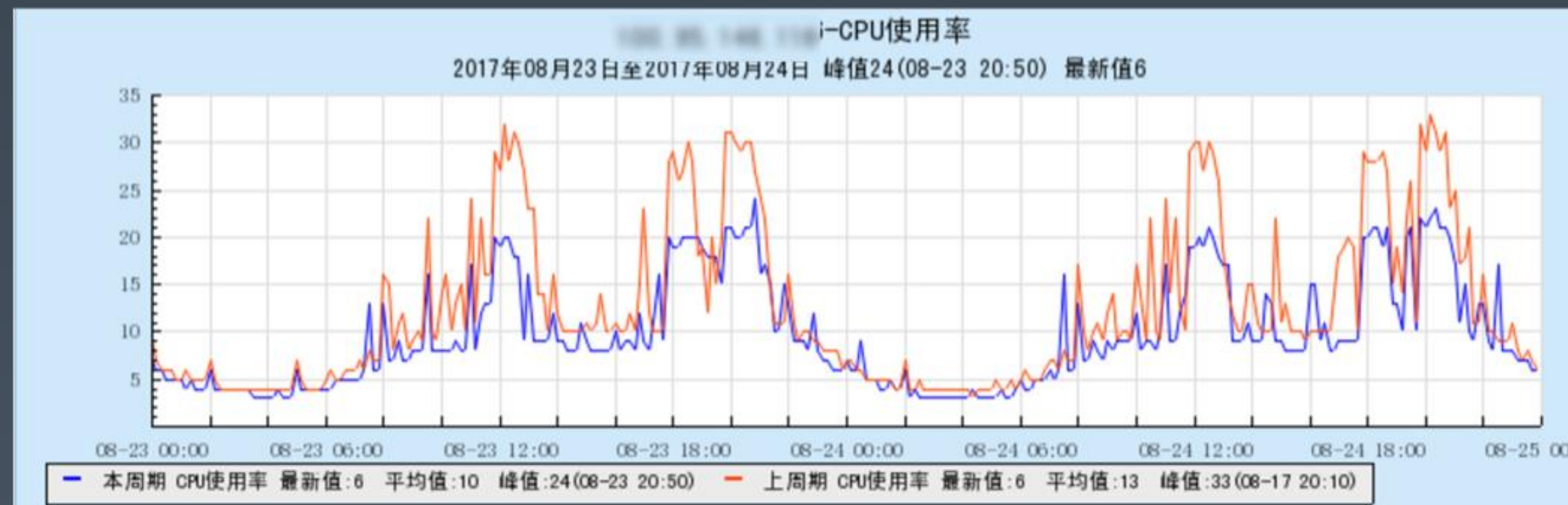


IPIP+Gateway混合Overlay方案

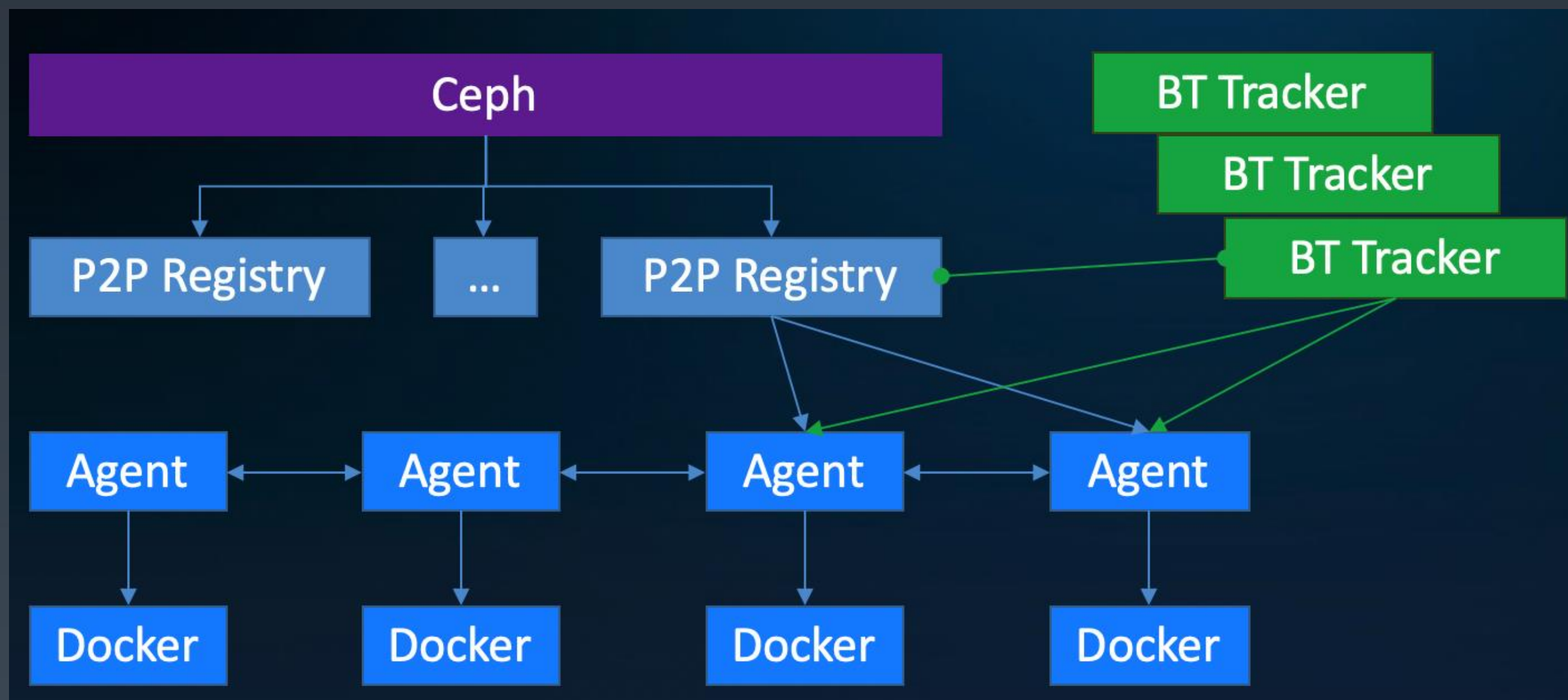
- 短链接IPIP包量比Vxlan多**14.1%**
- Gateway比Vxlan多**40.5%**
- 方案被Flannel社区合并

Underlay方案

- Bridge方式仅比Host差**6%**，一般overlay比Host差**20~40%**
- SRIOV方式比Bridge CPU下降**38.3%**，包量+6%



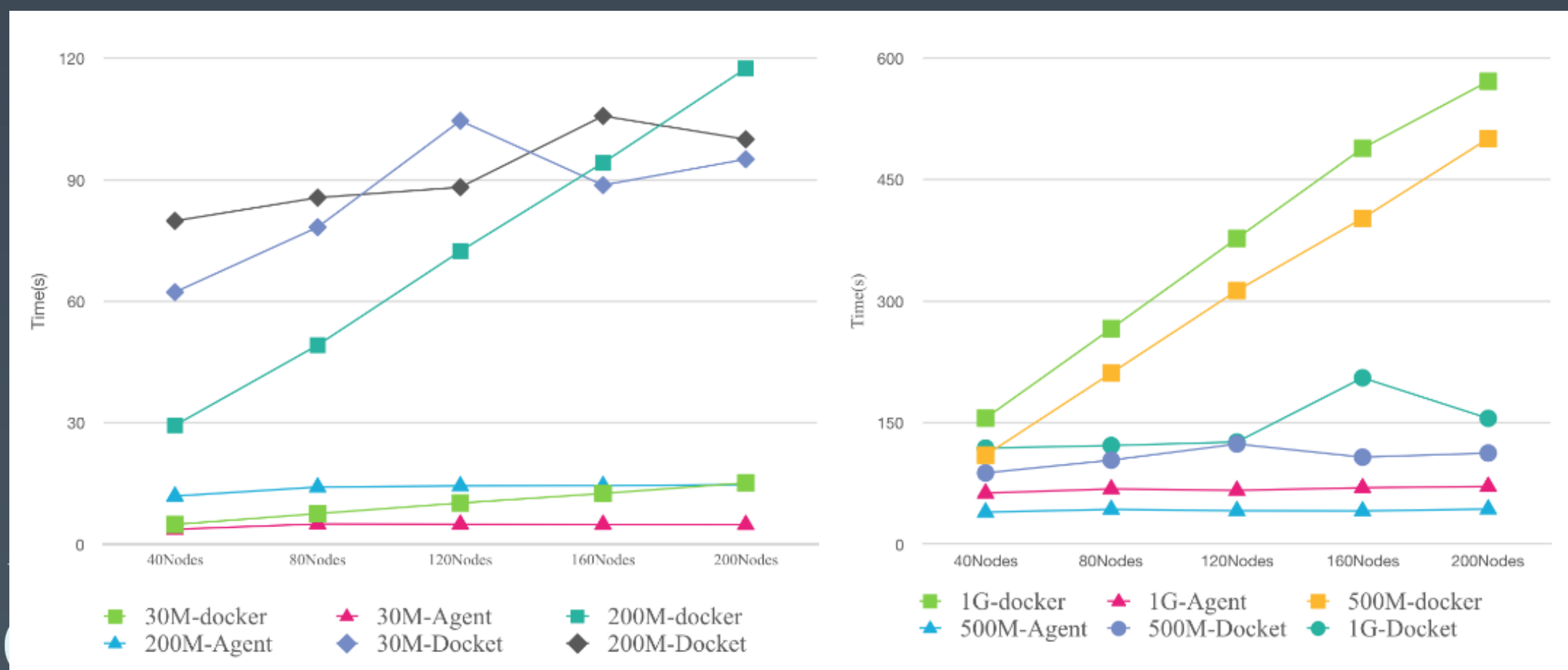
性能



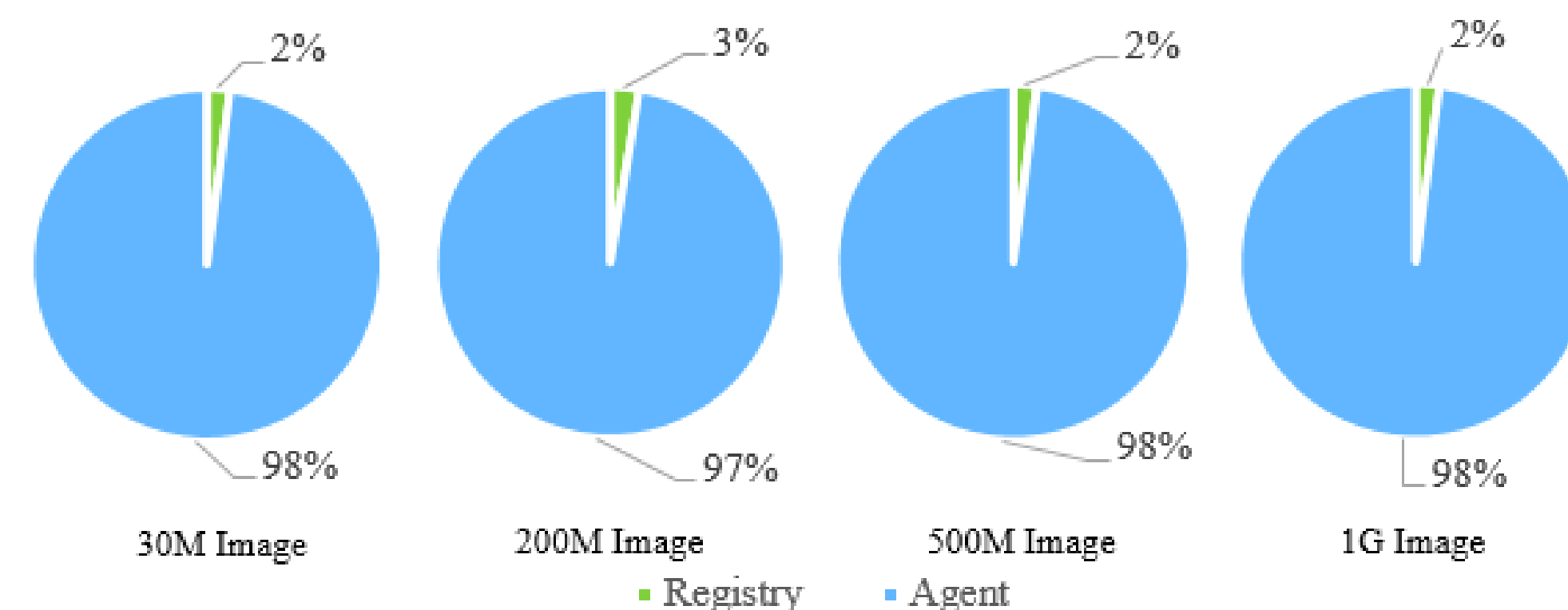
- 镜像下载引入BT协议
- 对Docker Daemon零入侵
- 每层分别做种
- 优化blob下载策略

发表论文: 《FID: A Faster Image Distribution System for Docker Platform》
2017 IEEE 2nd International Workshops on FASW

Docker、Docket、Gaiastack P2P Agent下载镜像对比



Registry与P2P Agent流量占比对比



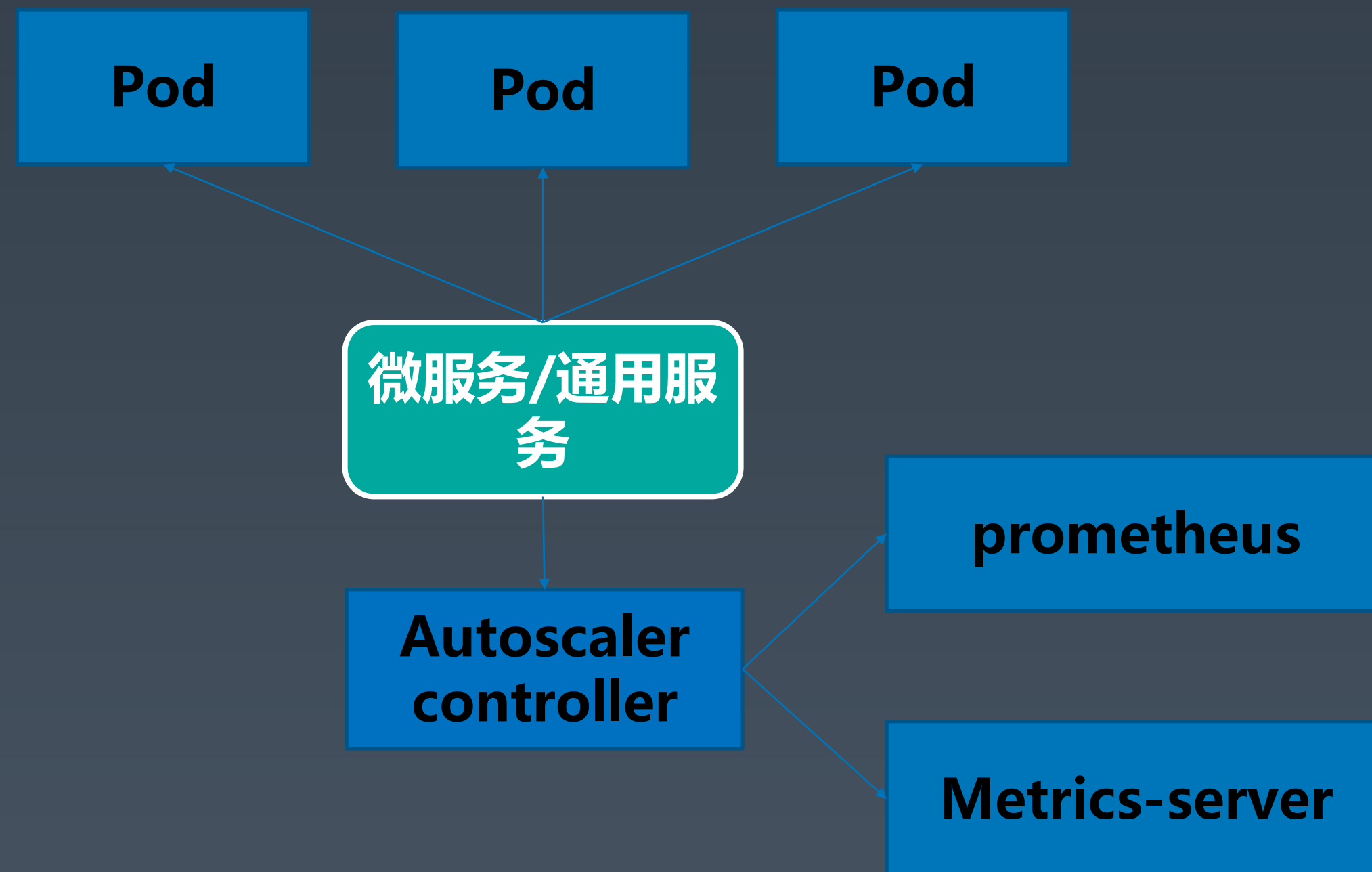
安全



能力扩展：弹性伸缩

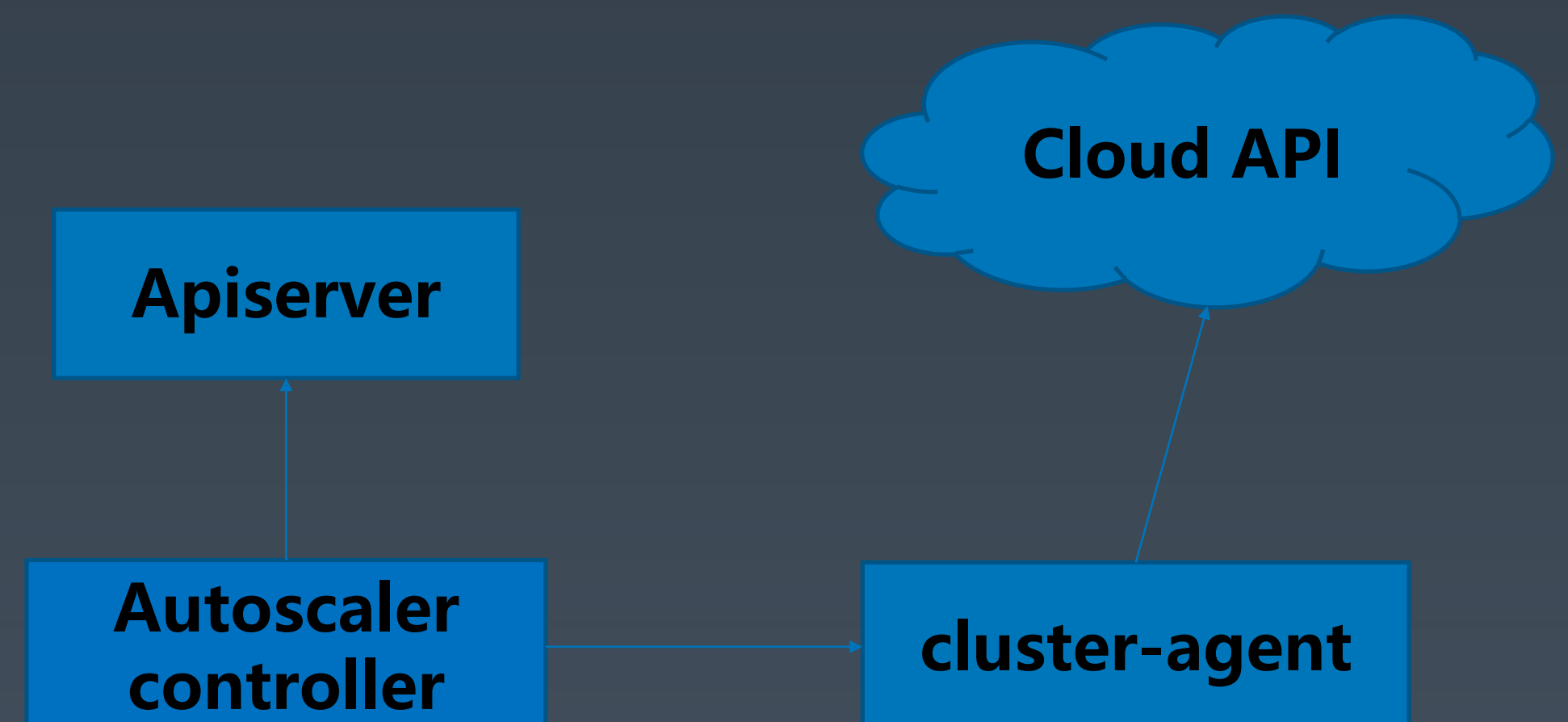
APP弹性伸缩：

- 主动扩缩容
 - 扩容可以指定新版本
 - 缩容可以定点裁撤
- 自动扩缩容
 - 资源阈值
 - 自定义指标阈值
 - 实例个数范围
 - 周期性自动伸缩




集群弹性伸缩：

- 监控节点资源使用率
- 自动迁移低负载Node上的Pod，完成缩容
- 一定数量Pod因资源不足pending时，自动扩容



能力扩展：灰度升级



prd-cloud-str-003-p40-cluster1运行中

启动

停止

更多操作

实例数及状态数：25 / 运行中：25

应用类型：通用服务App

访问地址：<http://prd-cloud-str-003-p40-cluster1.teg-qboss-ocr-admin.szjx.sp>

所属：业务

部署镜像：cloud.str.003-p40 v8

创建信息：创建于 2018-02-06 11:11:14

概况

实例

监控

日志

事件

操作记录

Compose 类型应用实例仅有一个容器，因此实例信息与容器信息一致

<input type="checkbox"/>	运行状态	实例名称 ↑	Pod IP 复制	Host IP 复制	部署镜像
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-0			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-1			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-2			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-3			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-4			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-5			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-6			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-7			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-8			cloud.str.003-p40 v8
<input type="checkbox"/>	运行中	prd-cloud-str-003-p40-cluster1-9			cloud.str.003-p40 v8

所属：业务

部署镜像：cloud.str.003-p40 v8

实例配置：4个 10GB 4GB 10Mbit/s 2个 GPU

pod IP

- 在GPU集群中有一个长时间服务应用prd-cloud-str-003-p40-cluster1。该应用有25个实例，每个实例需要2个GPU卡。用来提供图片识别的OCR服务。
- 当该服务要升级新的版本时，如果对所有实例停止，则会造成服务中断；如果采用滚动升级，无法保证升级过程是否有异常，以及无法充分验证新版本的可用性（即使经过了测试阶段的测试）。
- 通常采用灰度升级的方式：即选择某一个或N个实例先升级到新版本，在充分稳定验证后，再考虑升级其他实例，而该灰度的过程可以分为任意批次。有时为了验证多个版本，一个应用内也可以同时又多个版本并行存在。充分保证现网的服务质量以及版本的可控性。


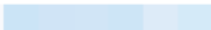



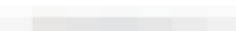

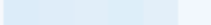
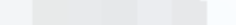

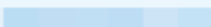
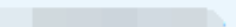



能力升级：灰度升级

事件	出现次数	详细说明	类型	用户名	首次出现时间
已调度	1	Successfully assigned prd-cloud-str-003-p40-cluster1-4 to tbd	Normal		2018-02-06 11:24:45
正在拉取镜像	1	pulling image "docker.oa.com:8080/g_bergxu/los-2.1.2-online-2and1-cloud.str.003-p40:v7"	Normal		2018-02-06 11:24:46
已创建	1	Created container with docker id 4fd68f3bba7d	Normal		2018-02-06 11:46:38
已拉取镜像	1	Successfully pulled image "docker.oa.com:8080/g_bergxu/los-2.1.2-online-2and1-cloud.str.003-p40:v7"	Normal		2018-02-06 11:46:38
已启动	1	Started container with docker id 4fd68f3bba7d	Normal		2018-02-06 11:46:38
终止	1	Killing container with docker id 4fd68f3bba7d: pod "prd-cloud-str-003-p40-cluster1-4_teg-qboss-ocr-admin(4766ae25-0aed-11e8-a5e8-6c0b84fff863)" container "prd-cloud-str-003-p40-cluster1" hash changed (3426724710 vs 3590433639), it will be killed and re-created.	Normal		2018-02-09 09:33:02
正在拉取镜像	1	pulling image "docker.oa.com:8080/g_bergxu/los-2.1.2-online-2and1-cloud.str.003-p40:v8"	Normal		2018-02-09 09:33:02
已启动	1	Started container with docker id d320b150b7c0	Normal		2018-02-09 09:35:16
已拉取镜像	1	Successfully pulled image "docker.oa.com:8080/g_bergxu/los-2.1.2-online-2and1-cloud.str.003-p40:v8"	Normal		2018-02-09 09:35:16
已创建	1	Created container with docker id d320b150b7c0	Normal		2018-02-09 09:35:16

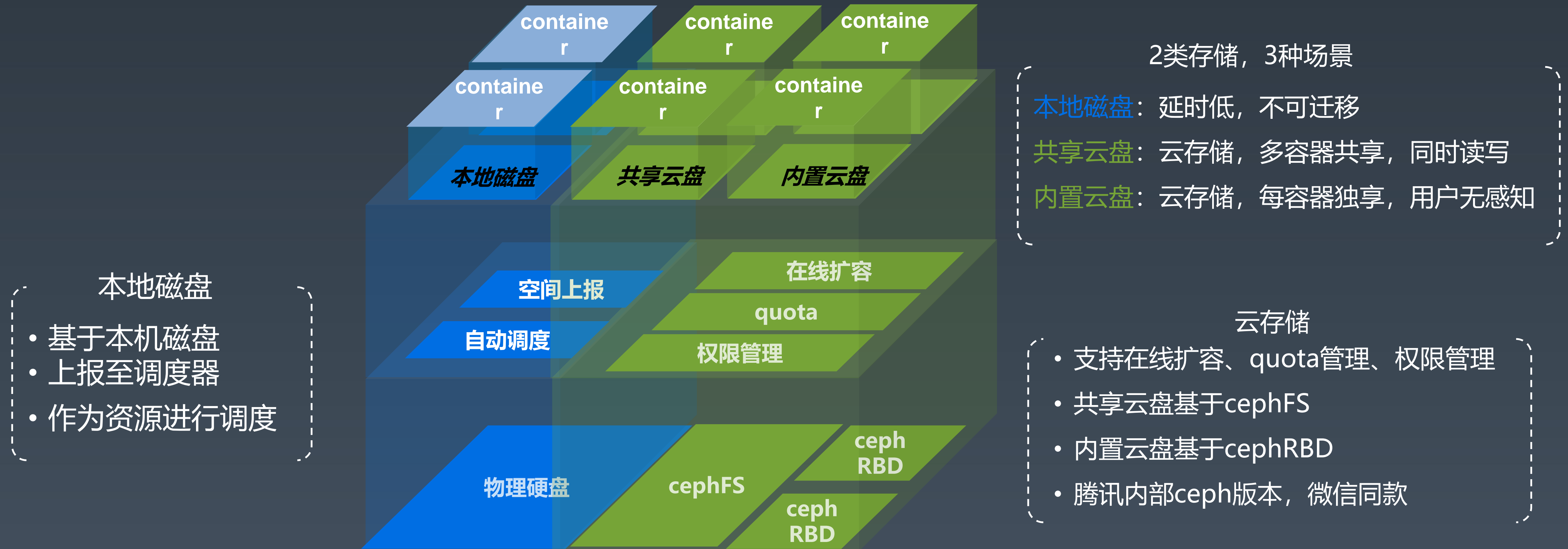
如左图所示，对某一个实例从v7升级到v8版本。

- 2018-02-06 11:46:38 V7版本开时候运行
- 2018-02-09 09:33:02 对该实例做灰度升级，从V7版本升级到V8版本
- 2018-02-09 09:33:02 开始pull V8版本的image

PS：灰度升级属于原地升级，因此不需要重新过调度，升级的效率也会提升。
每次升级可以选择要升级的实例个数以及具体哪些(个)实例。

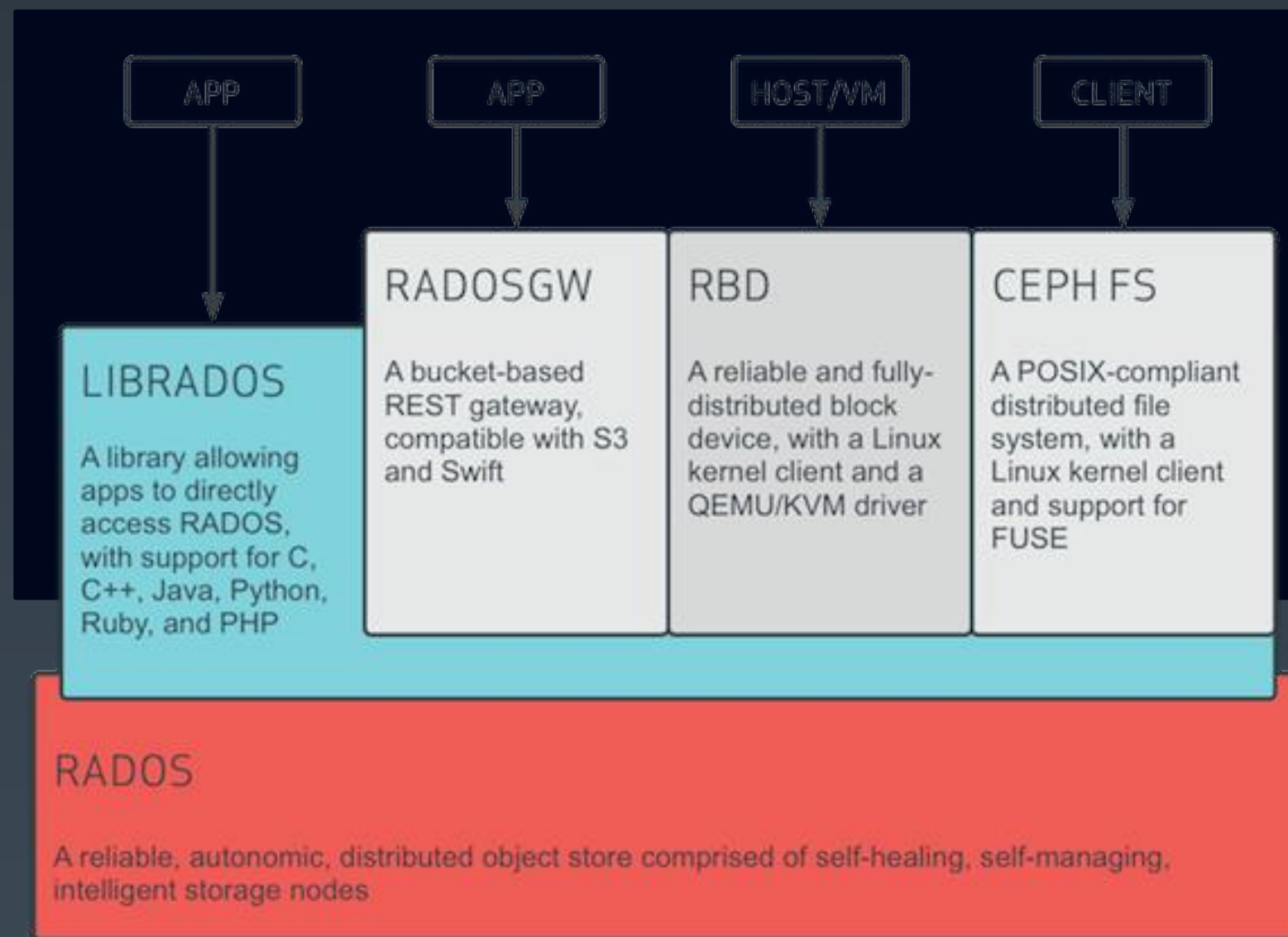
<div>启动</div> <div>停止</div> <div>灰度升级</div>				
<input type="checkbox"/>	运行状态	实例名称 ↑	Pod IP 复制	Host IP 复制
<input checked="" type="checkbox"/>	 运行中	prd-cloud-str-003-p40-cluster1-0		
<input type="checkbox"/>	 运行中	prd-cloud-str-003-p40-cluster1-1		
<input checked="" type="checkbox"/>	 运行中	prd-cloud-str-003-p40-cluster1-2		
<input type="checkbox"/>	 运行中	prd-cloud-str-003-p40-cluster1-3		
<input checked="" type="checkbox"/>	 运行中	prd-cloud-str-003-p40-cluster1-4		

能力扩展：存储场景



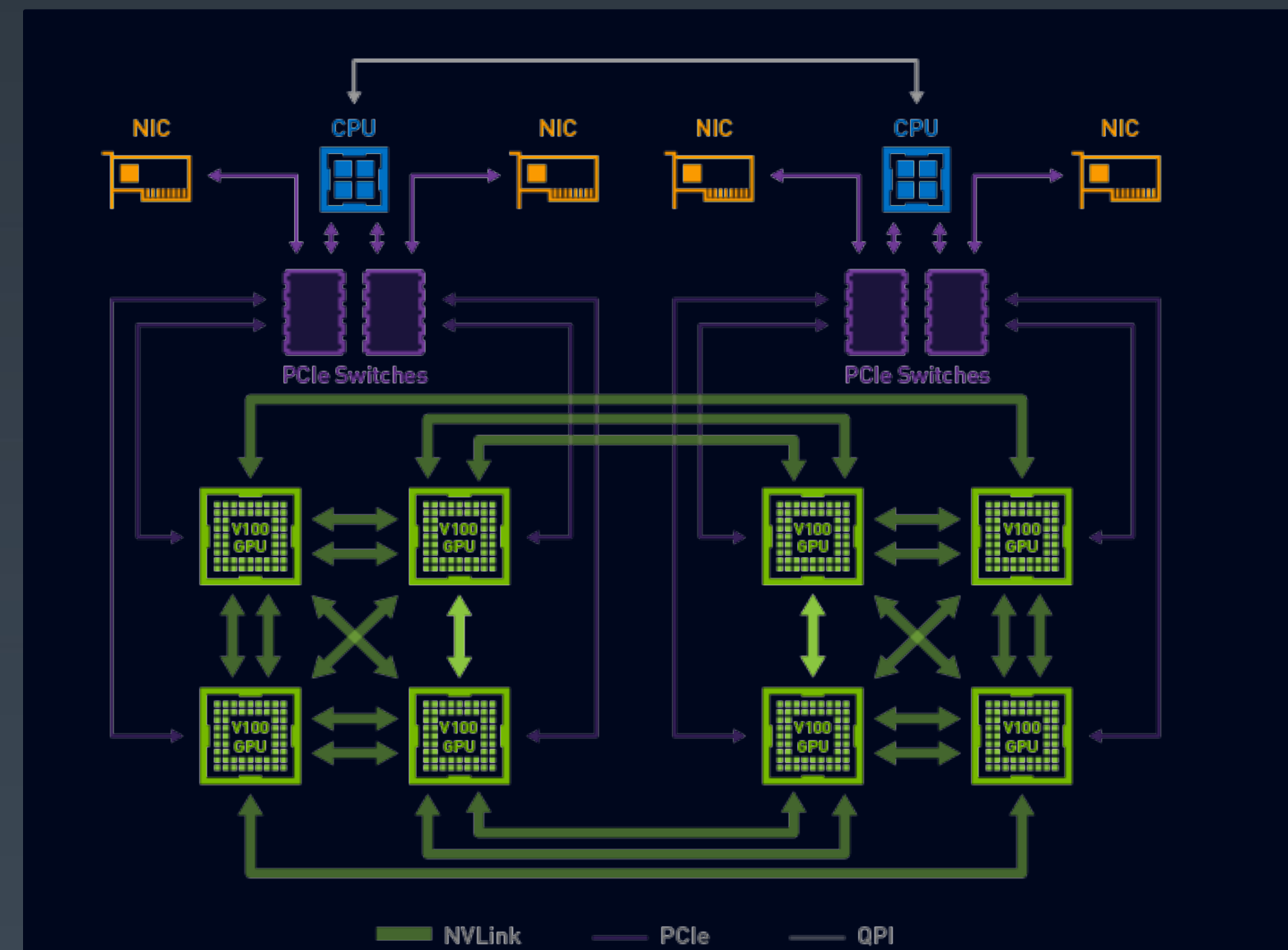
能力扩展：GPU支持

发表论文：《Gaia Scheduler: A Kubernetes-based Scheduler Framework》
The IEEE ISPA 2018 (16th IEEE International Symposium on Parallel and Distributed Processing with Applications)



分布式存储Ceph

海量小数据读写优化
不同用户配额管理
任务带盘迁移



智能拓扑感知

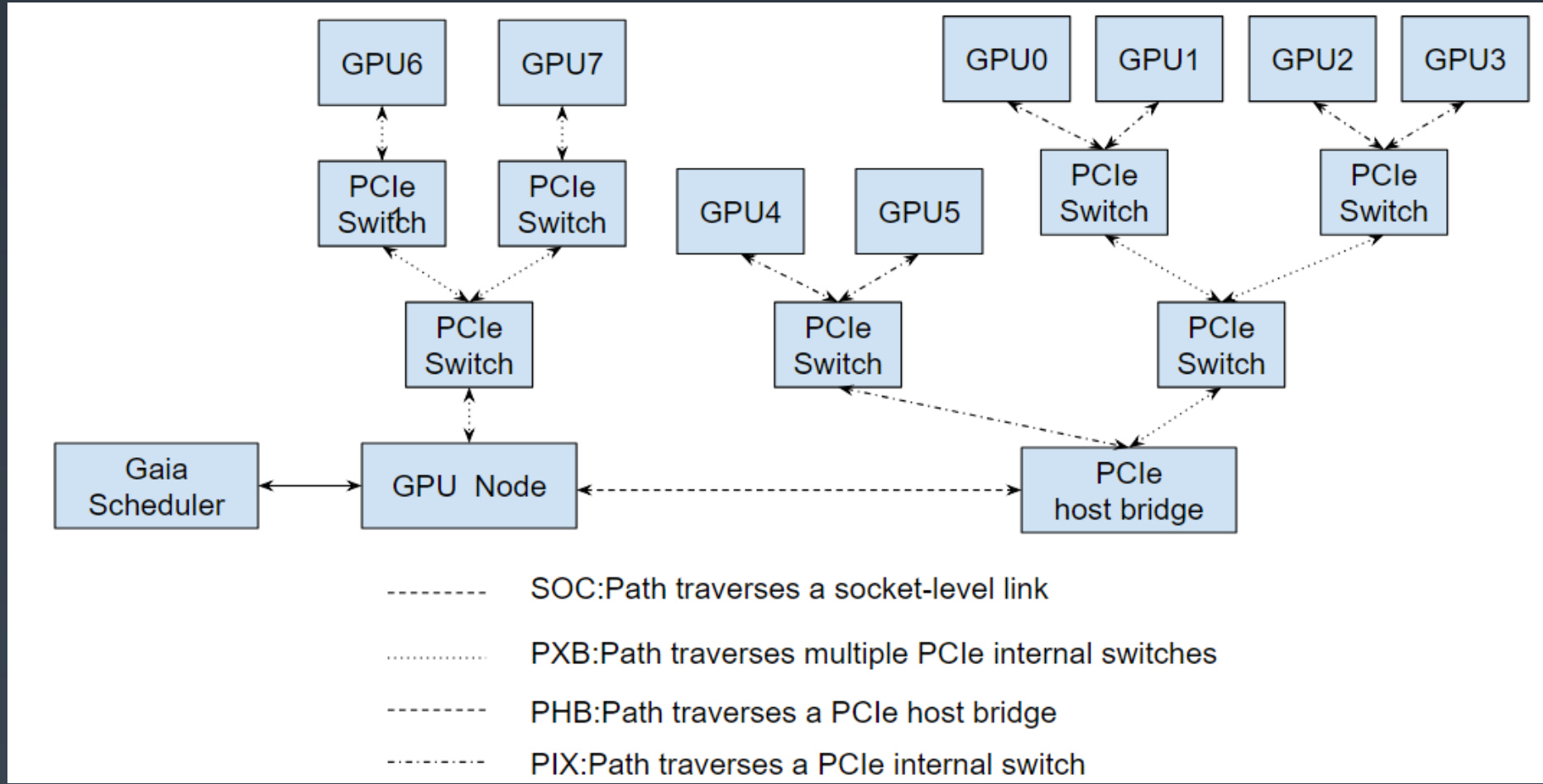
GPU卡拓扑感知
资源访问代价树决策
资源调度算法解决碎片化



异构GPU统一管理

多种调度策略，多租户管理GPU卡
与CPU核自动绑定
支持单机多卡和多机多卡

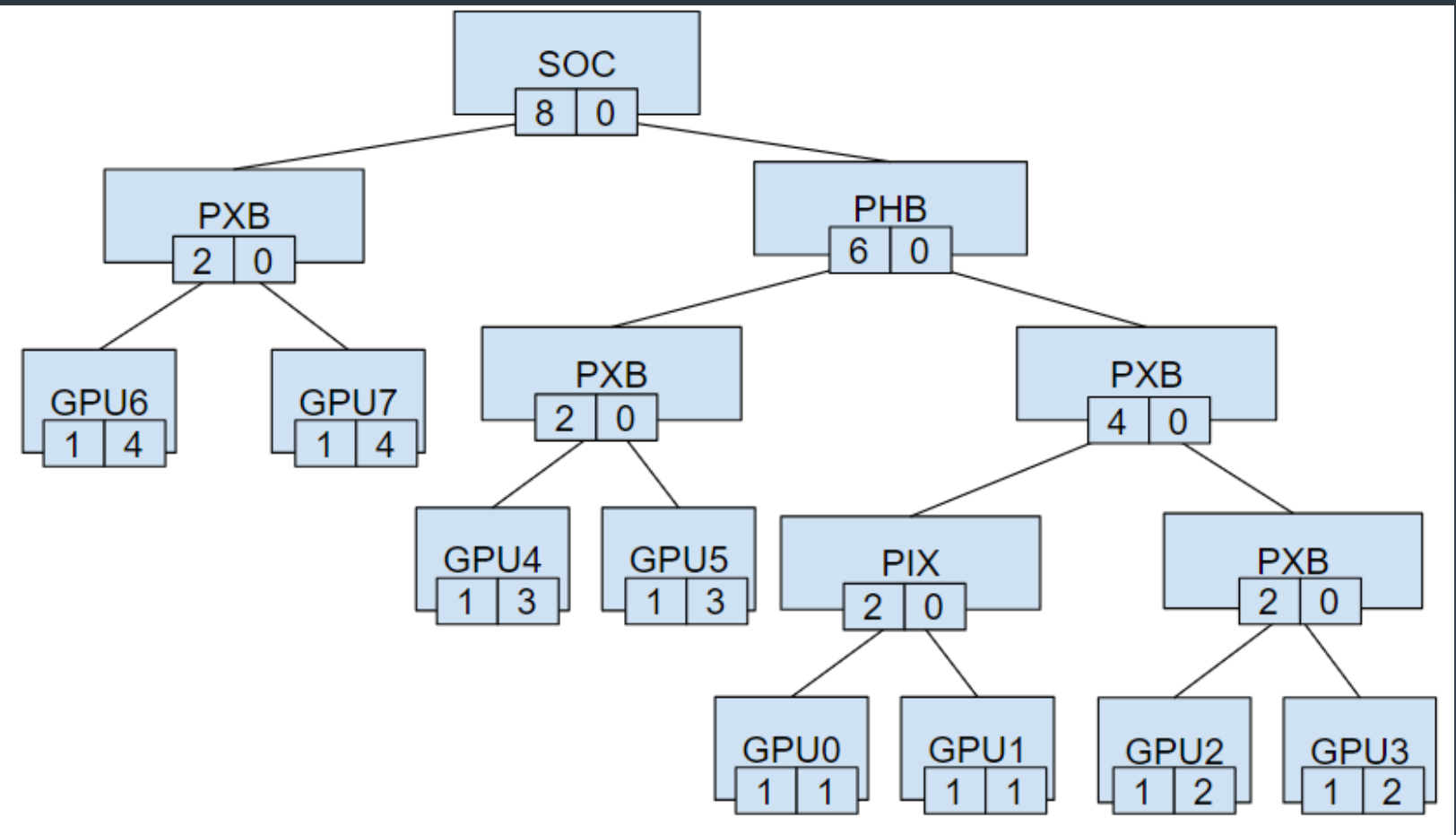
能力扩展：GPU支持



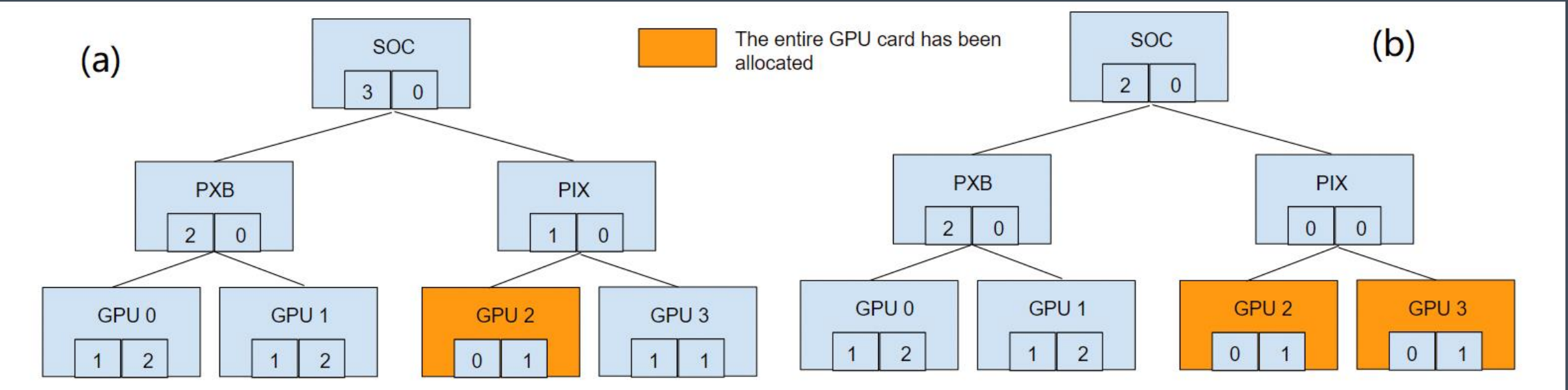
资源-访问代价树

- 拓扑节点中存储3个信息：
- 子节点的GPU通信方式(SOC、PXB、PHB或PIX)
 - 可用的GPU资源数(如果下属n张GPU卡则为n)
 - 节点通信开销(非GPU节点为0)

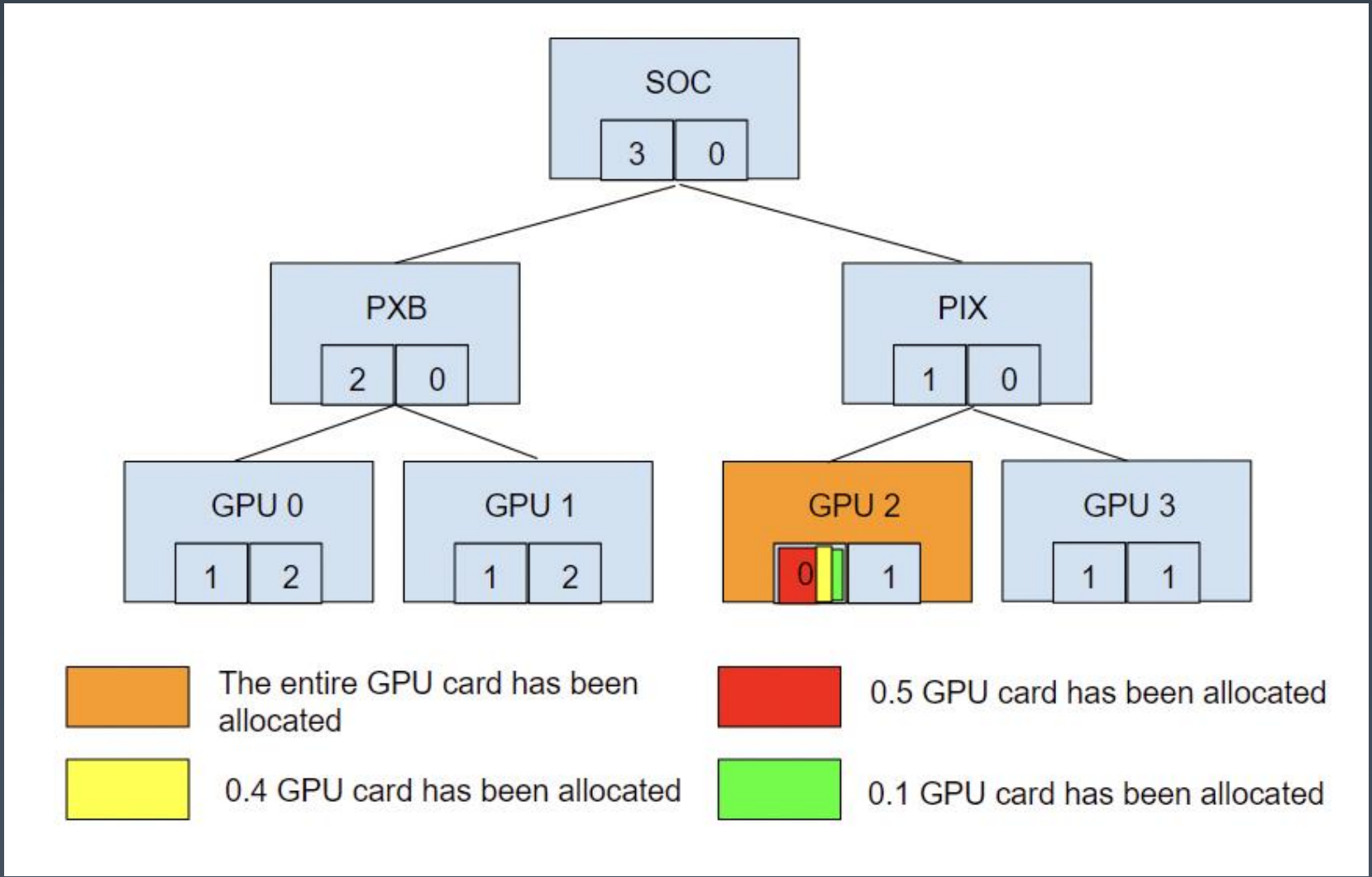
- GPU节点存储3个信息：
- GPU id
 - 可用的GPU资源数(GPU节点为1)
 - 节点通信开销(数字越小，访问代价越低)



四类通信方式分类中，通信开销最大的是SOC，其次是PXB，再次是PHB，PIX通信方式的GPU之间的通信开销最小。



Singular and link



Fragment

成本

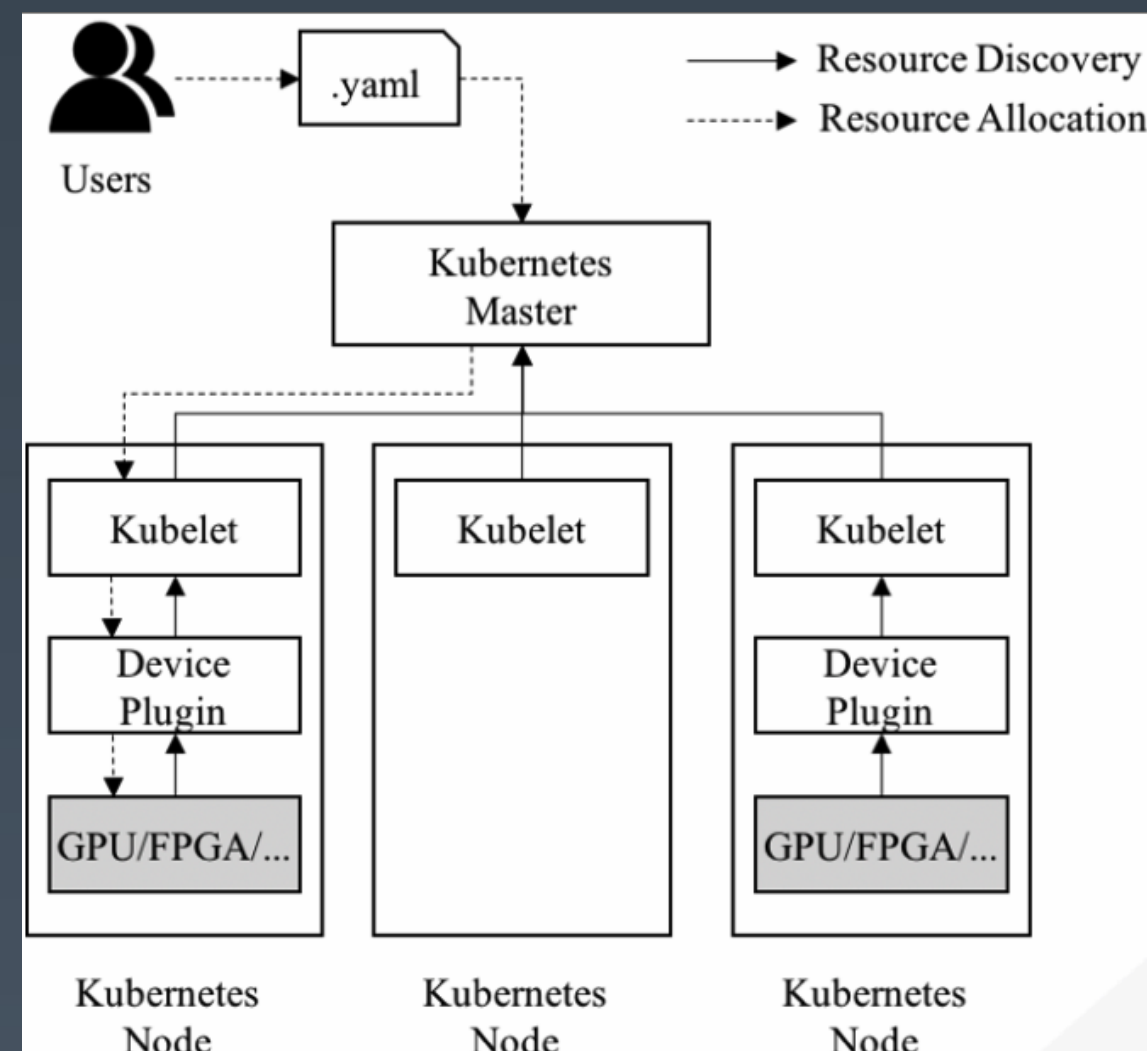
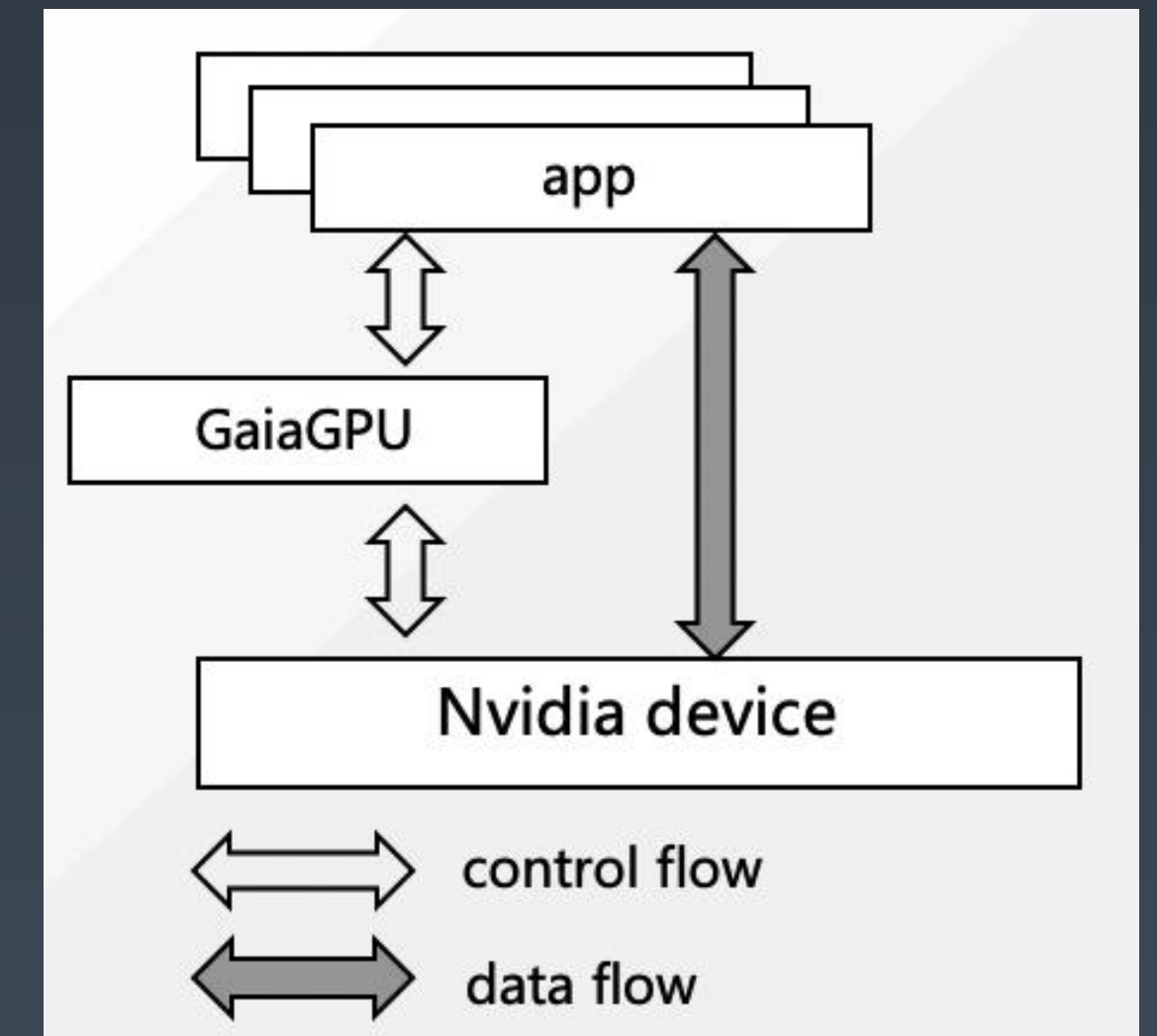
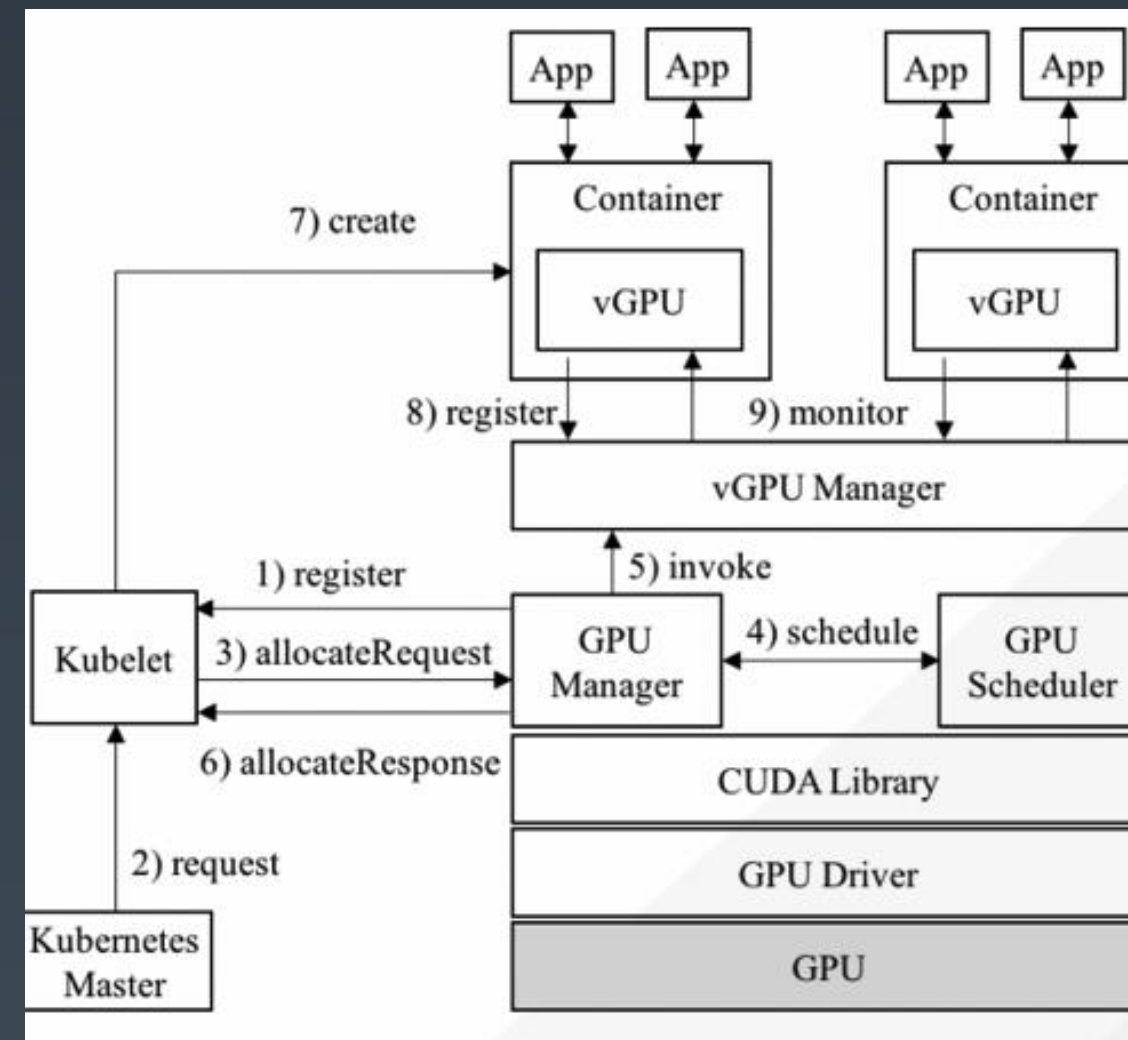
发表论文：《GaiaGPU: Sharing GPUs in Container Clouds》
The IEEE ISPA 2018 (16th IEEE International Symposium on Parallel and Distributed Processing with Applications)

transparent. GaiaGPU不应修改Kubernetes代码或容器镜像以共享GPU。使用共享GPU执行应用程序应该就像在物理GPU上执行一样。

Performance. GaiaGPU应当保证vGPU的性能与原生GPU性能相近。

Isolation. GaiaGPU可以有效的分配和回收每个容器使用的GPU资源并实现不同容器间的资源隔离。

GPU使用方式	实现
VCUDA	在vm中构建wrapper library以拦截GPU调用并将这些调用重定向到宿主机执行
Amazon	将设备直接挂在到vm中
GPUvm	在Zen的hypervisor层实现了全虚拟化。为了隔离运行在物理GPU上的多个VM，GPUvm将物理GPU分成几个部分，并将每个部分分配给单个VM。
NVIDIA GRID	在硬件层面实现GPU虚拟化，每个容器可以绑定一个虚拟GPU
NVIDIA Docker	通过将GPU设备及运行时的库转为volume挂载到容器中实现了容器与驱动的解耦。但是一个GPU设备仅能挂载到一个容器中，不支持容器间共享GPU设备
ConvGPU	仅支持内存资源的共享且仅处理单个GPU



容器使用GPU的问题：

- 需要特定的硬件设备
- 不支持容器共享
- 仅支持内存资源虚拟化
- 仅支持单个GPU卡

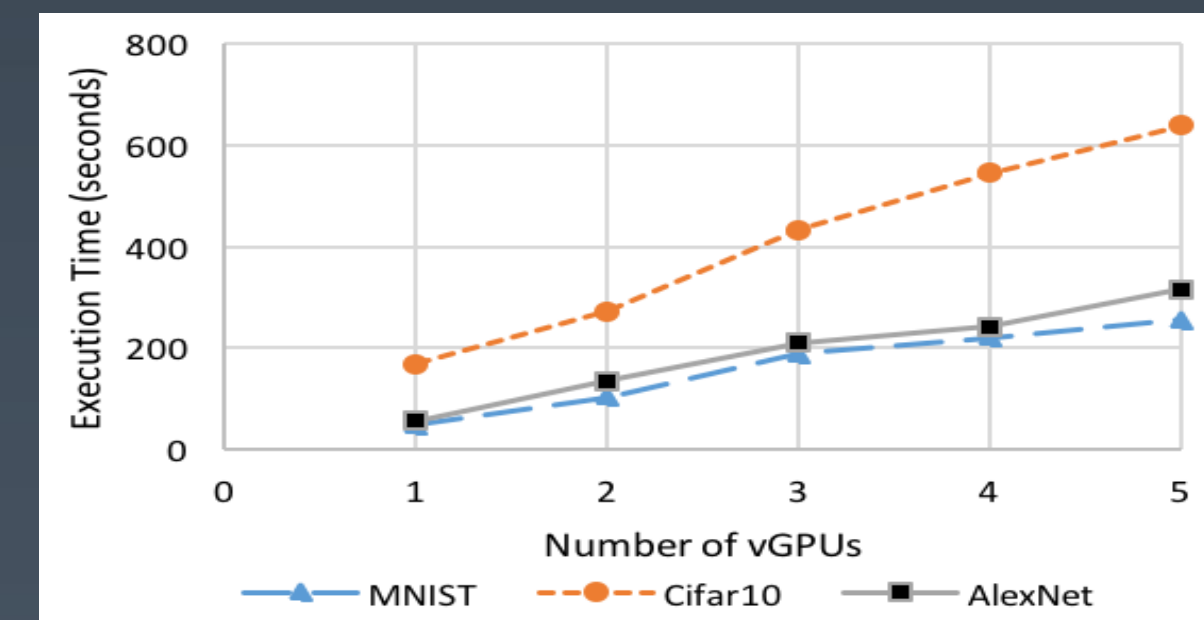
采用Device Plugin：

- GPU资源的发现
- 为任务分配相应的硬件资源及配置容器运行时环境

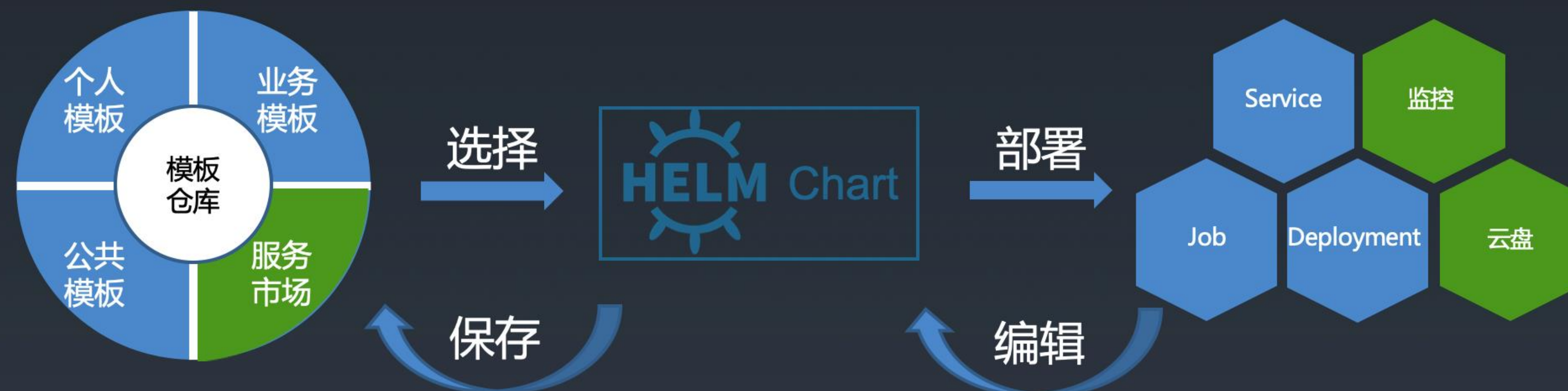
$$\text{Difference} = \frac{\text{GaiaGPU_time} - \text{Native_time}}{\text{Native_time}} \times 100\%$$

	Native_time (seconds)	GaiaGPU_time (seconds)	Difference (%)
Tensorflow	47.82	47.88	0.13
Caffe	22.47	22.50	0.15
PyTorch	69.33	69.64	0.44
CNTK	7.39	7.41	0.27

Number of vGPUs	1	2	3	4	5
resource per vGPU	1	0.5	0.3	0.25	0.2



生态



模板仓库

- 兼容HELM模板仓库
- 模板在线编辑、上传、下载
- 多级管理
- RBAC

服务市场

- 由 **GAIASTACK** 提供的最佳实践模板
- 极简配置，一键部署
- 高可用保证
- 多种常用服务及开源软件的腾讯自研版本



- 支持K8S原生接口
- 支持 **GAIASTACK** 的高级功能，如全维度资源管理、监控、云盘、日志管理等

Next



Tencent Kubernetes Engine

TGO 鲲鹏会

汇聚全球科技领导者的高端社群

🏢 全球12大城市

👤 850+ 高端科技领导者

使命

Mission

为社会输送更多优秀的
科技领导者

愿景

Vision

构建全球领先的有技术背景
优秀人才的学习成长平台



扫描二维码，了解更多内容

想做团队的领跑者 需要迈过这些“槛”

成长型企业，易忽视人才体系化培养
企业转型加快，团队能力又跟不上

VS

从基础到进阶，超100+一线实战
技术专家带你系统化学习成长

团队成员技能水平不一，
难以一“敌”百人需求

VS

解决从小白到资深技术人所遇到
80%的问题

寻求外部培训，奈何价更高且
集中式学习

VS

多样、灵活的学习方式，包括
音频、图文 和视频

学习效果难以统计，产生不良循环

VS

获取员工学习报告，查看学习
进度，形成闭环



课程顾问「橘子」

回复「QCon」
免费获取
学习解决方案

极客时间企业账号 # 解决技术人成长路上的学习问题

THANKS!

QCon  th