

云智未来<sup>9</sup><sup>th</sup>

第九届中国系统架构师大会  
SYSTEM ARCHITECT CONFERENCE CHINA 2017

# 小米生态云的应用引擎实践

李波，小米生态云

# 大纲

- 小米生态云简介
- 小米生态云应用引擎演进
- 未来规划

SACC2017

# 小米生态云

- 为小米生态链及合作伙伴提供一站式云服务及解决方案
- 完整的产品和服务
  - 20+
  - 涵盖云计算 大数据 人工智能
  - 统一使用小米账号体系
- 效率 安全 自由 成本 大数据 人工智能
- 国际化布局

# 小米生态云

CLI

## 用户管理控制台

应用管理

认证与授权  
(集成小米账号)

用户权限管理  
(用户/组/角色)

计量计费

事件审计

## 应用引擎

应用1  
(公司A)

应用2  
(公司A)

应用3  
(公司B)

应用n  
(公司N)

文件存储 (FDS)

消息队列(EMQ)

日志采集与分析

结构化存储(SDS)

流式消息队列(Talos)

监控报警

数据处理(EMR)

RPC服务治理

计划任务

数据库服务(RDS)

缓存服务

第三方服务

## 大数据服务

用户画像

推送推广

数据工场

数据通道

.....

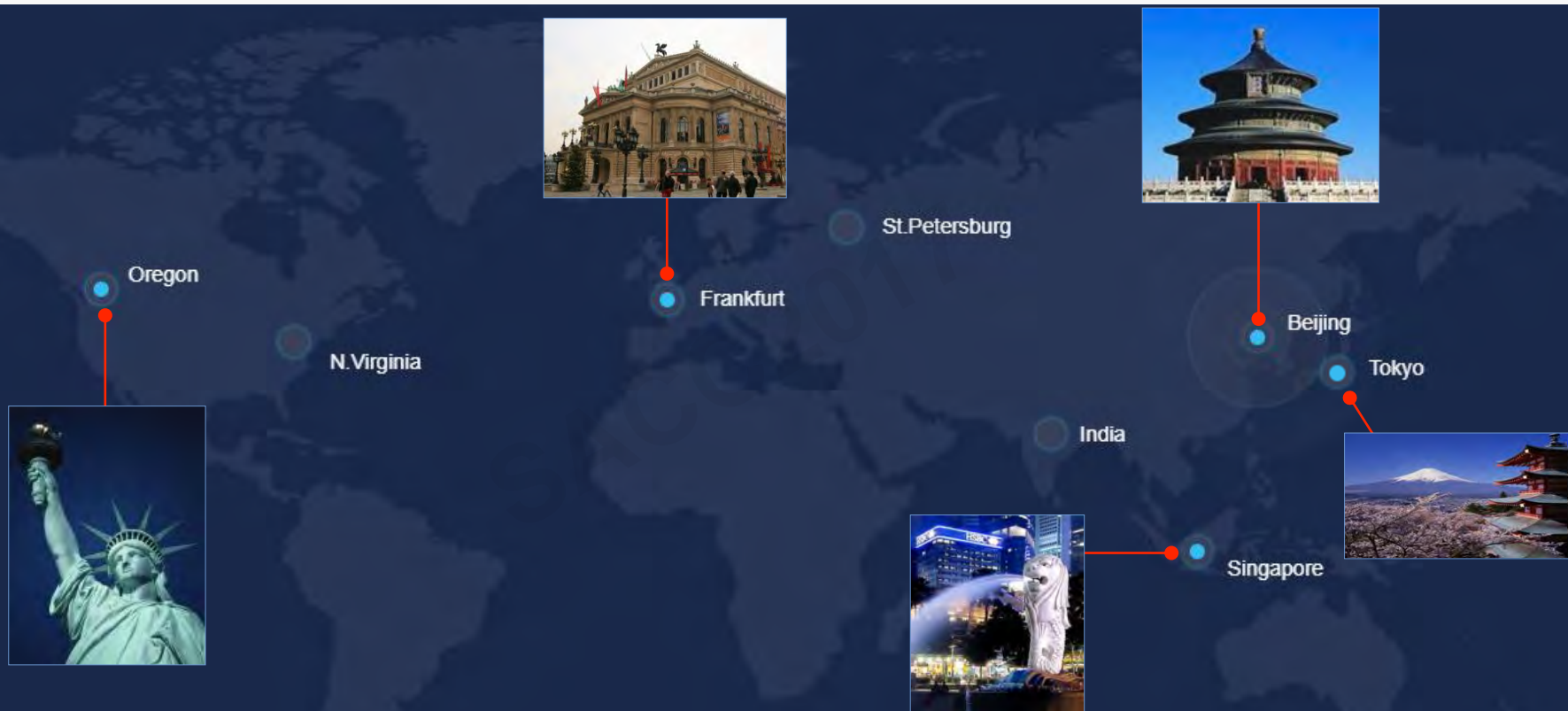
## 人工智能服务

深度学习

智能语音

.....

# 小米生态云区域分布



# 小米生态云应用引擎演进

SACC 2017

# 应用引擎v1

- 基于Cloud Foundry
- 集成小米账号，支持公司及部门隔离，用户和角色管理
- 支持主流开发语言以及静态页面和二进制文件（ Heroku Buildpack ）
- 支持Docker应用
- 域名及证书

# 优点和缺点

- 优点
  - 开箱即用的PaaS平台
  - 完整的权限和授权体系
  - 成熟稳定，非常适用于无状态Web应用
- 缺点
  - Buildpack机制和基础文件系统不灵活，定制难度大
  - 生态链公司开发测试逐步迁移到Docker平台，与生产环境不一致
  - 无法限制应用的CPU绝对用量
  - 不支持Cluster应用、UDP应用
  - Docker支持不完整，非原生体验
  - 自有体系，组件繁多，部署运维复杂
  - 社区参与度和活跃度下降



# 应用引擎v2

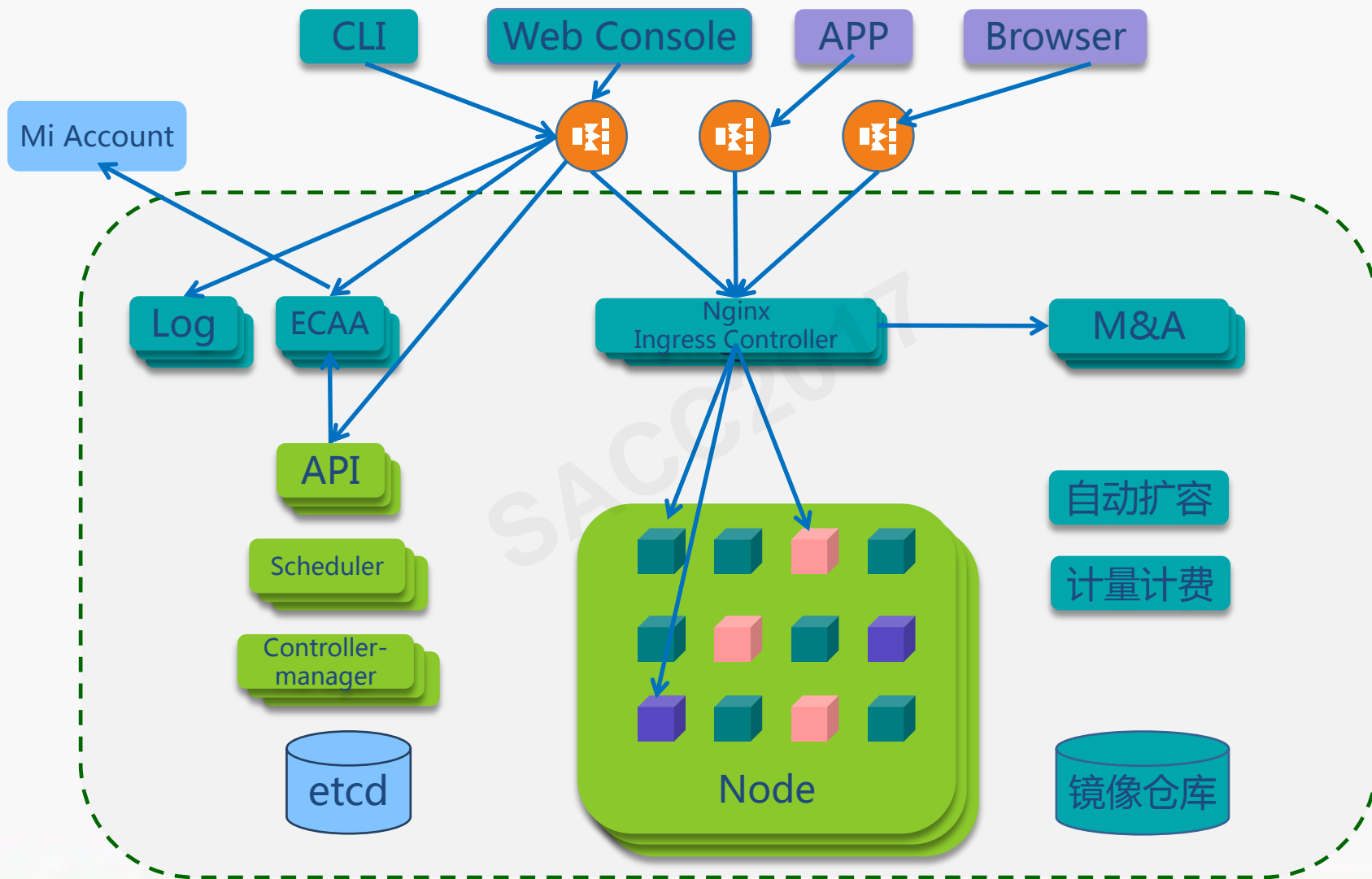
- 基于Kubernetes
- 原生Docker体验
- 支持TCP/UDP应用
- 原生计划任务支持
- 配置和敏感信息管理
- CPU的绝对用量限制
- 支持Cluster应用

SACC2017

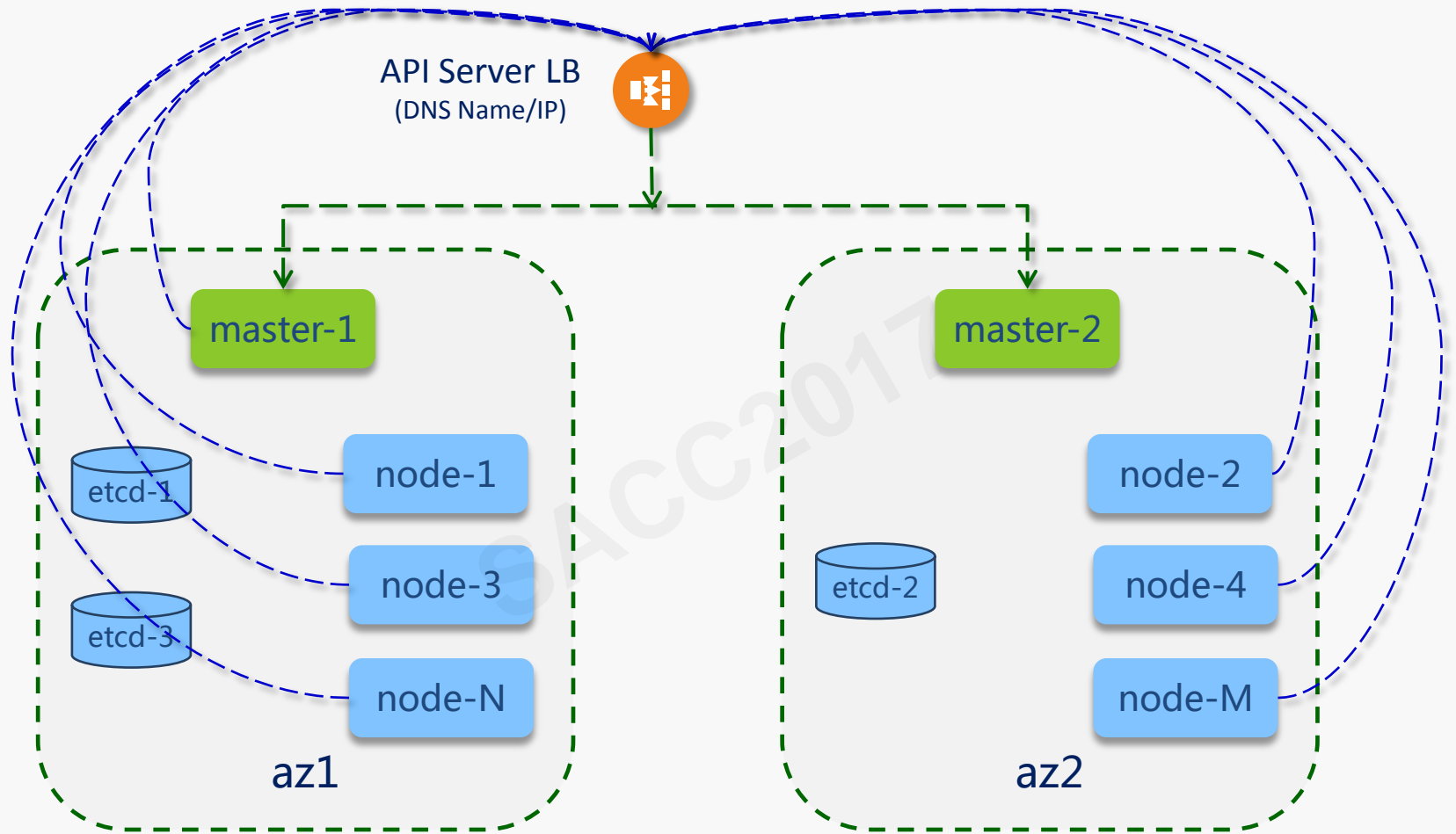
# 扩展 定制

- 高可用部署
- 认证授权
- 安全隔离
- 应用抽象和封装
- 应用自动扩容
- 容器直连
- 外部HTTP/HTTPS以及TCP/UDP服务
- 命令行工具
- 日志
- 监控报警
- 镜像安全扫描

# 整体架构



# 高可用部署



kubectl label nodes node-1 failure-domain.beta.kubernetes.io/region=cn-bj-6 failure-domain.beta.kubernetes.io/zone=cn-bj-6a

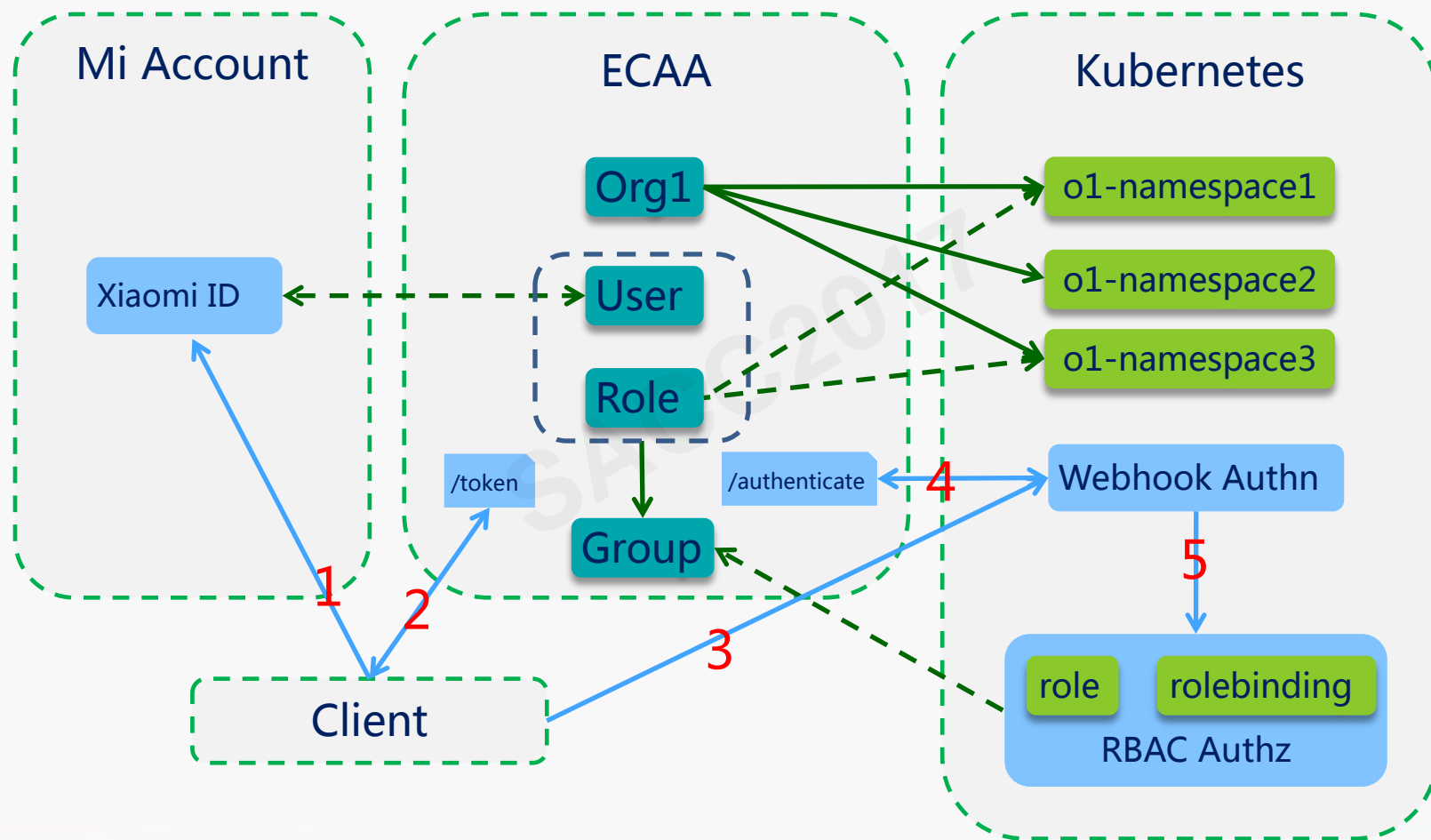
kubectl label nodes node-2 failure-domain.beta.kubernetes.io/region=cn-bj-6 failure-domain.beta.kubernetes.io/zone=cn-bj-6b

# 多租户环境

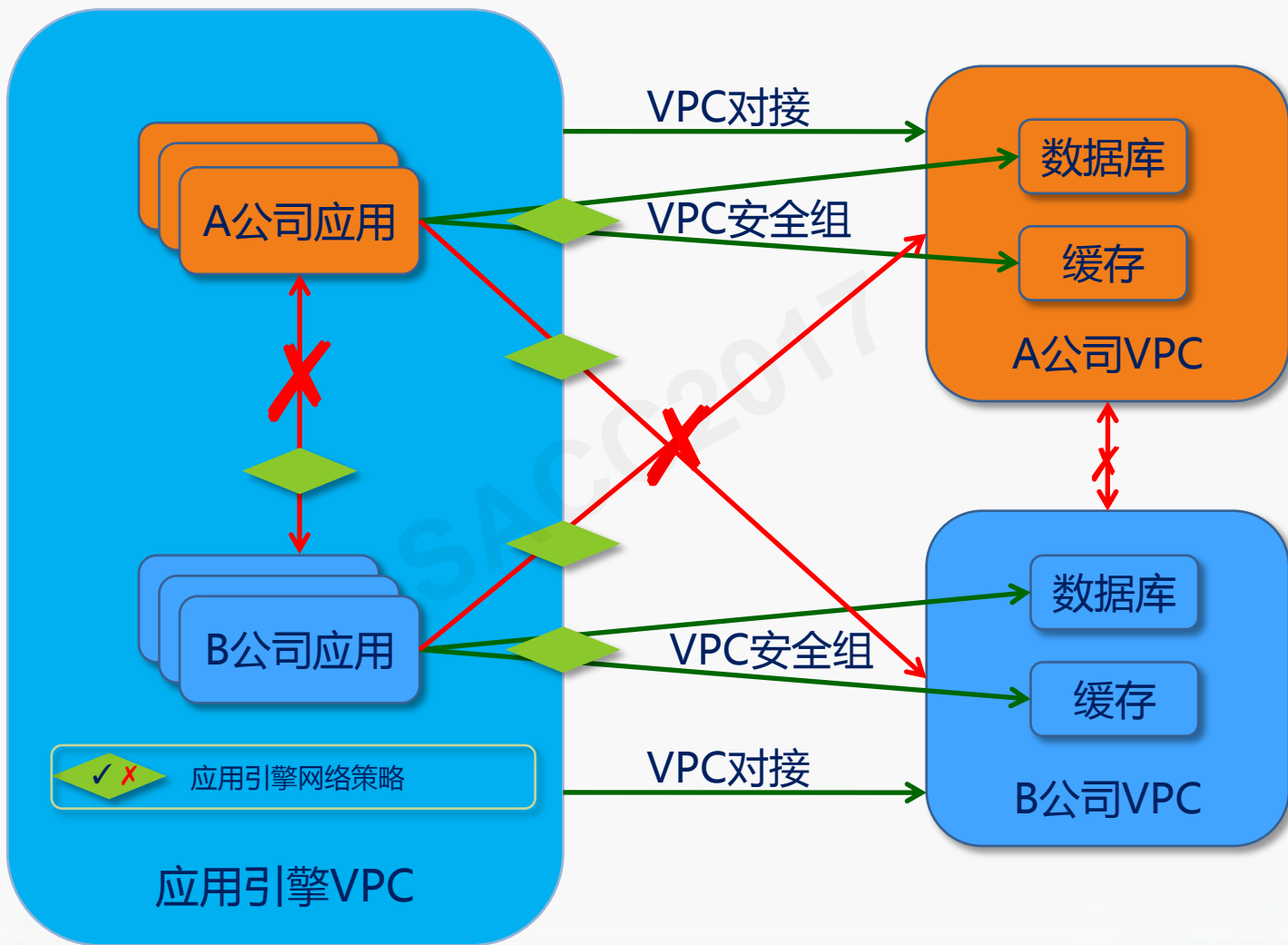
- 安全
  - 权限控制
  - 网络隔离
- 资源公平分配
  - CPU
  - 流量

SACC2017

# 认证授权



# 网络隔离



# 网络策略实现

- 容器网络: Calico
- Calico Policy
  - 优先级从高到低
    - 允许某公司的应用访问该公司的VPC网段 ( egress )
    - 允许同一公司的应用之间相互访问 ( egress+ingress )
    - 允许所有应用访问kube-dns地址 ( egress )
    - 允许Nginx Ingress Controller访问所有应用 ( ingress )
    - 禁止所有应用访问所有私有地址 ( egress )
    - 允许所有应用
      - 访问所有地址 ( egress )
      - 被所有地址/应用访问 ( ingress )



# Calico Policy

- 允许同一公司的应用之间相互访问

```
- apiVersion: v1
kind: policy
metadata:
  name: ${ORG}.allow-access-among-same-org
spec:
  egress:
    - action: allow
      destination:
        selector: k8s_ns/label/org == '${ORG}'
      source:
        selector: k8s_ns/label/org == '${ORG}'
  ingress:
    - action: allow
      destination: {}
      source:
        selector: k8s_ns/label/org == '${ORG}'
  order: 800
  selector: k8s_ns/label/org == '${ORG}'
```

# Calico Policy

- 禁止所有应用访问所有私有地址

```
- apiVersion: v1
kind: policy
metadata:
  name: cluster-policy.deny-private-egress
spec:
  egress:
    - action: deny
      destination:
        nets:
          - 192.168.0.0/16
          - 172.16.0.0/12
          - 10.0.0.0/8
        source: {}
      order: 900
      selector: has(calico/k8s_ns) && calico/k8s_ns != 'kube-system'
```

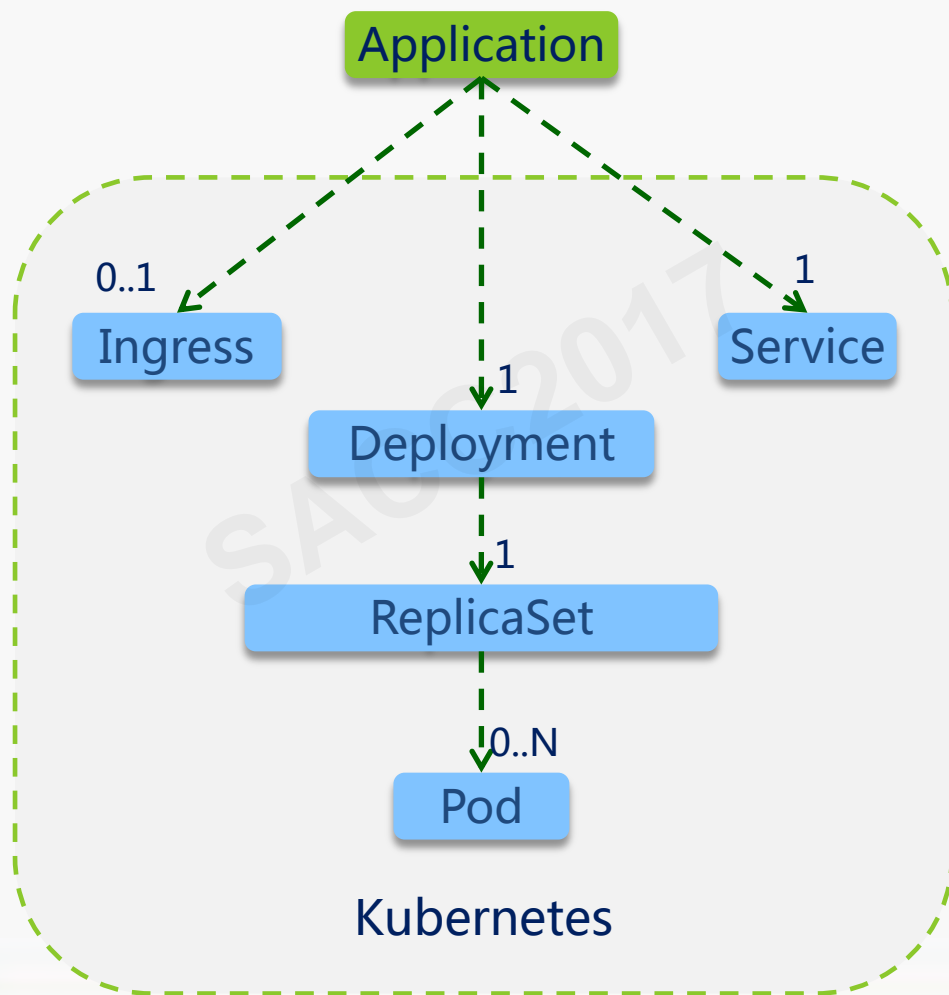
SACC2017

# 流量隔离

- 配置不同的LB
  - 独立IP
  - 独立带宽
  - 独立的Nginx ( 如果需要 )

SACC2017

# Web应用抽象和封装



# 应用自动扩容

## 我的扩容策略

最小实例数：2 最大实例数：50

修改

## 按规则扩容

间隔周期：3分钟

规则 1



平均 CPU 使用率高于 80%，达到 2 个周期，按 个数 增加实例，每次增加 1个

平均 CPU 使用率低于 20%，达到 2 个周期，按 个数 减少实例，每次减少 1个

## 按计划扩容

### 重复计划

计划 1



开始时间：09:55 结束时间：10:30 重复频率：周日，周一，周二，周三，周四，周五，周六

计划实例数：20 结束后实例数：2

### 特定日期

计划 1



开始时间：2017-11-10 20:00 结束时间：2017-11-12 08:00

计划实例数：50 结束后实例数：2

# 容器直连



- `wss://<apiserver>/api/v1/namespaces/<ns>/pods/<pod>/exec?stdout=1&stdin=1&stderr=1&tty=1&container=<container>&command=%2Fbin%2Fsh&command=-i`

# 外部服务

- HTTP/HTTPS
  - LB+ Nginx Ingress Controller
- TCP/UDP
  - EIP + Nginx (Stream Proxy)
  - 端口规划和分配

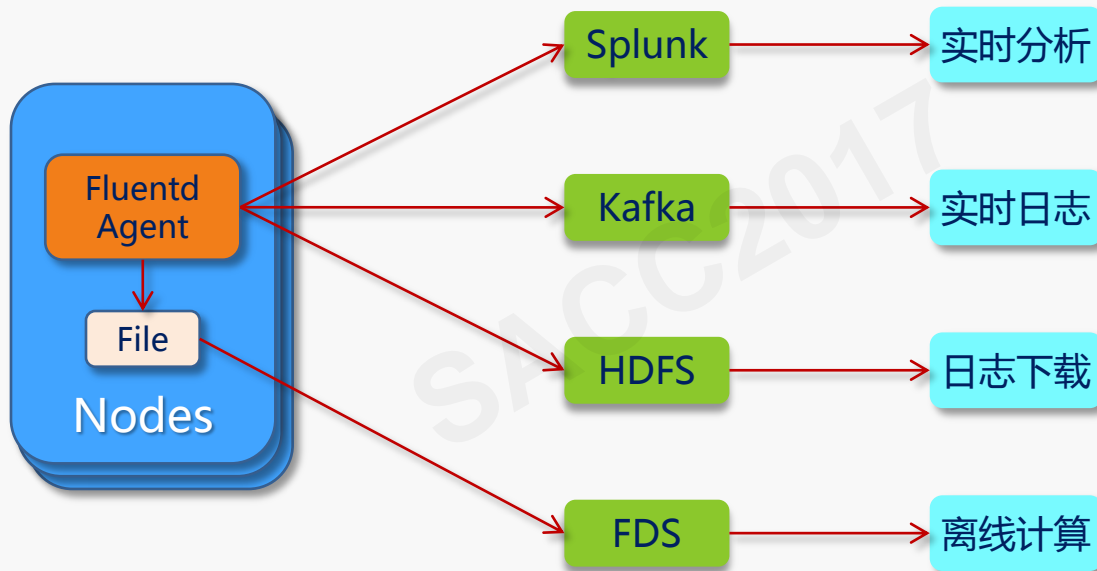
# 命令行工具

- 不使用kubectl
  - 暴露过多细节，功能过于强大
  - 认证不便
- 自研工具
  - ak/sk认证
  - 有限功能
    - 自动化运维
      - 应用部署，更新，删除
      - 应用水平扩容
      - 资源指标(CPU,内存)
    - 实时日志

SACC2017



# 日志



# 监控 报警

- 监控
  - Nginx Ingress Controller + Lua + Open Falcon
  - HTTP相关指标
- 报警
  - Open Falcon
  - 短信
  - 邮件

SACC2017

# 镜像安全

- 镜像安全扫描
  - Harbor + Clair

SACC2017

# 经验分享

- 资源限制
  - ResourceQuota
  - LimitRange
- CronJob
  - successfulJobsHistoryLimit, failedJobsHistoryLimit
  - concurrencyPolicy: Forbid/Replace
  - activeDeadlineSeconds
- Events
  - 导出到外部存储
- Docker
  - log rotation: --log-driver json-file --log-opt max-size=100m --log-opt max-file=10
- Calico
  - AWS多AZ: CALICO\_IPV4POOL\_IPIP=always
  - 重命名策略: k8s-policy-no-match

# 应用引擎未来规划

- 应用版本管理
- 多应用映射同一域名
- 限制容器磁盘大小
- 应用资源监控
- 报警服务
- 精细化的资源调度
- 集群自动扩容
- Windows支持

THANKS

The background features a dark, almost black, space filled with numerous small, bright blue particles. These particles are arranged in several distinct, curved paths that sweep across the frame from the bottom left towards the top right. A bright, white-to-blue gradient light source is positioned near the center, casting a soft glow and illuminating the particles it passes. The overall effect is one of dynamic movement and digital energy.