

The background features a light blue grid pattern with various circular icons: an hourglass, a power button, a recycling symbol, a gear, a laptop, and a starburst. On the left, there are clusters of blue dots of varying sizes and colors (light blue, medium blue, dark blue).

SDCC 2016
中国软件开发者大会
SOFTWARE DEVELOPER CONFERENCE CHINA

华为全球IT系统基于PaaS的实践分享

嘉宾：吴建林

目录

- 华为IT需求背景
- PaaS架构设计
- 需求实践
- 问题回顾
- Q&A

引言



技术



业务



生产力

华为IT介绍



华为IT现状

全球8大数据中心
17万+固定用户



业务支撑

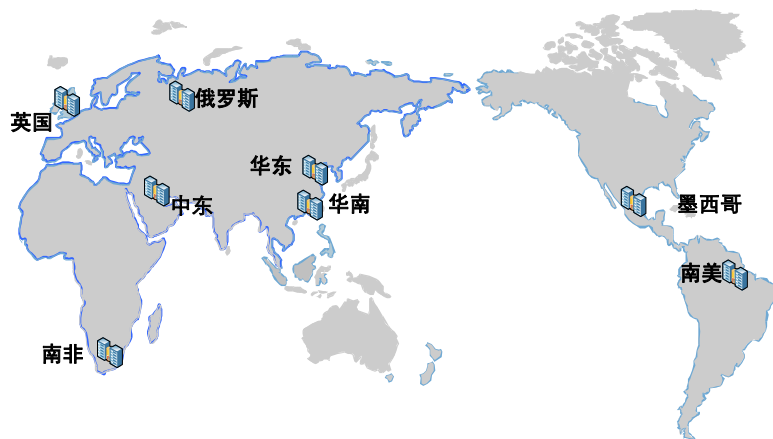
采购、供应链、交付、销售、财经
办公IT、流程IT



管理规模

数十万台虚拟机
800+应用、2000+子应用 (2016)

华为IT需求背景



01 传统IT模式业务不够敏捷
人工审批流程繁多，无法快速Devops迭代

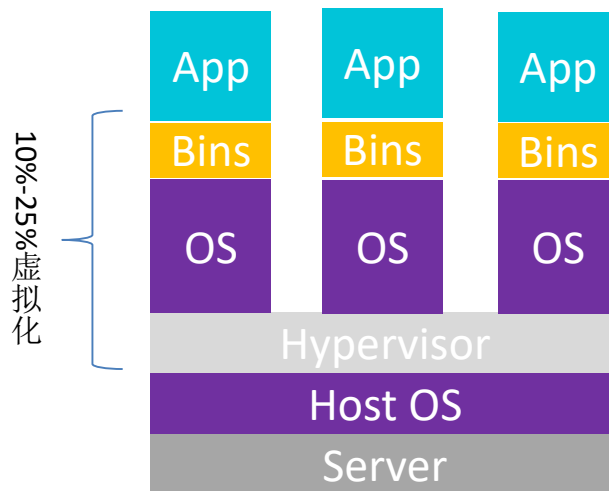
02 管理规模庞大，运维难度高
全球8大数据中心，管理虚拟机数量极速增长，运维一体化成本高

03 应用微服务化改造，个数急剧增长
通过微服务改造应用数量增长3倍以上，传统IaaS无法支持快速弹性部署

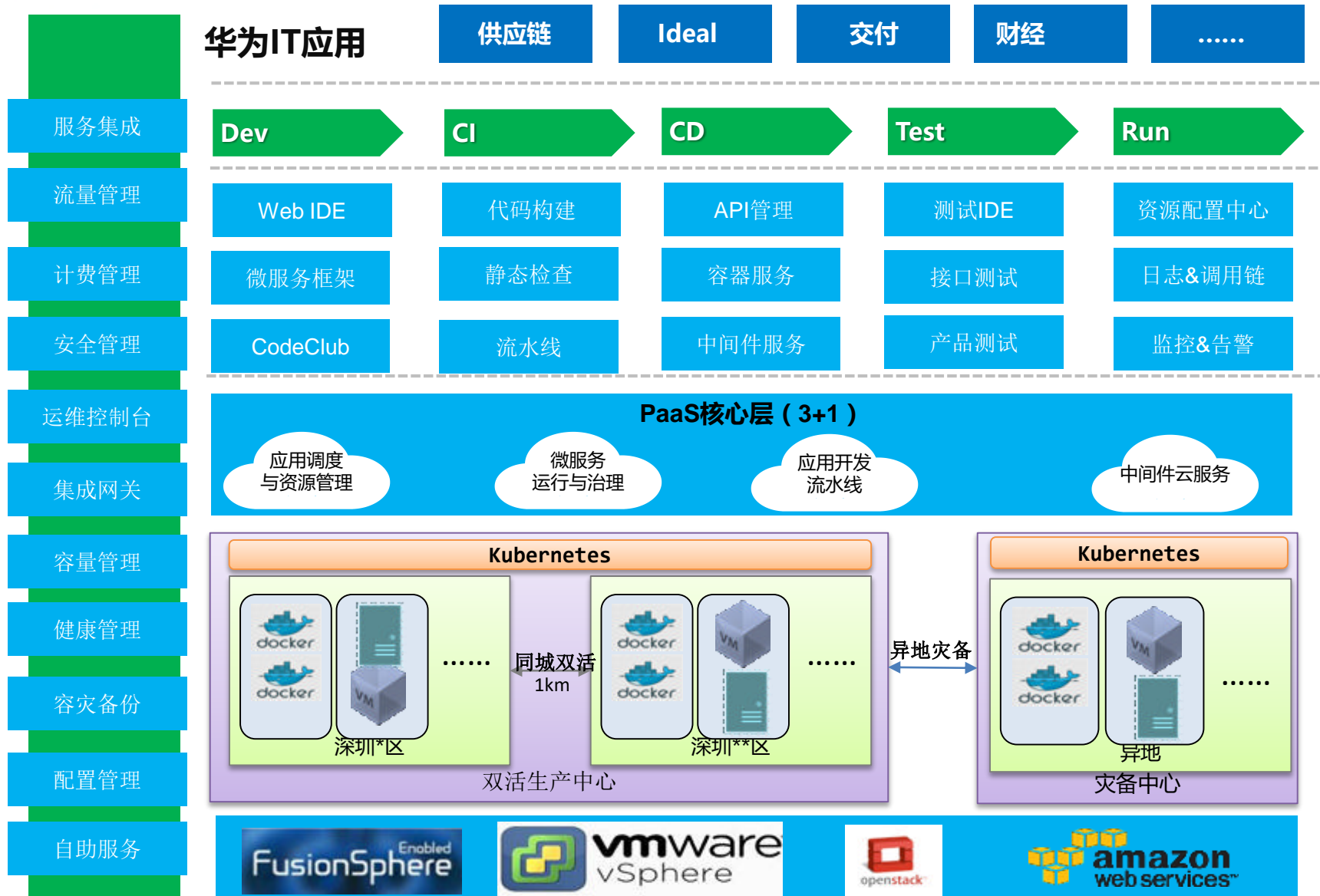
04 虚拟化成本高，资源利用率低
IaaS虚拟化成本高，应用独占资源，利用率低

05 全球业务访问体验差
缺乏全球分布式动态路由，用户访问速度无法保障

06 平台异构，支撑难度大
多种开发语言、多种IaaS、多种中间件增加平台支撑难度



PaaS架构设计



- 服务集成
- 流量管理
- 计费管理
- 安全管理
- 运维控制台
- 集成网关
- 容量管理
- 健康管理
- 容灾备份
- 配置管理
- 自助服务

FusionStage在华为IT的进展

1

2015年2月份开始容器试点

- 局部应用容器试点，基于Docker
- 应用无状态改造

2

2015年10月份开始上线生产环境

- 基于Kubernetes、Docker方案，管理规模：**1000+VM**，**5000+容器**
- 解决WS-AT网络协议问题，支持企业级应用的容器化方案，完成传统SOA向微服务的改造。

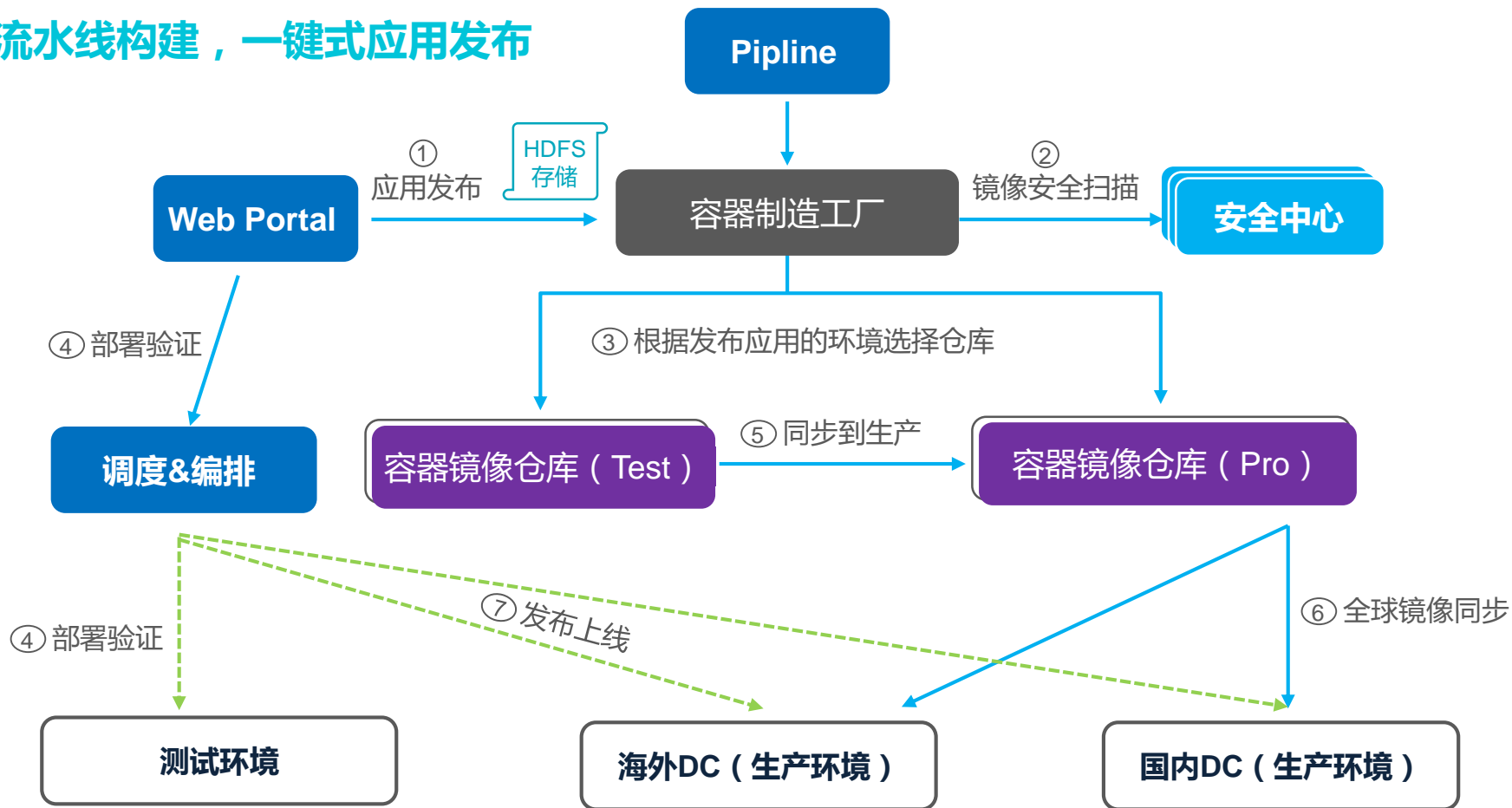
3

2016年10月份开始在内外部IT大规模推广

- 基于Kubernetes、Docker方案，管理规模：**4000+VM**，**20000+容器**
- 提高Devops效率**10倍**以上，资源利用率提升**1-3倍**，全自动化流程，替换繁琐的人工审批。
- 支持弹性伸缩，亲和反亲和调度，ELB自动对接等

需求实践

流水线构建，一键式应用发布



基于流水线，完成自动化版本构建，实现一键式应用发布；完成容器安全扫描，全球镜像同步与自动化调度编排，打通部门墙，实现Devops。部署效率由原来的**1周**提升到**5分钟**，开发迭代速度**提高10倍**。

需求实践

支持应用混合编排，实现虚拟机、物理机、容器的统一调度

Web Designer

图形化编排部署

混合编排

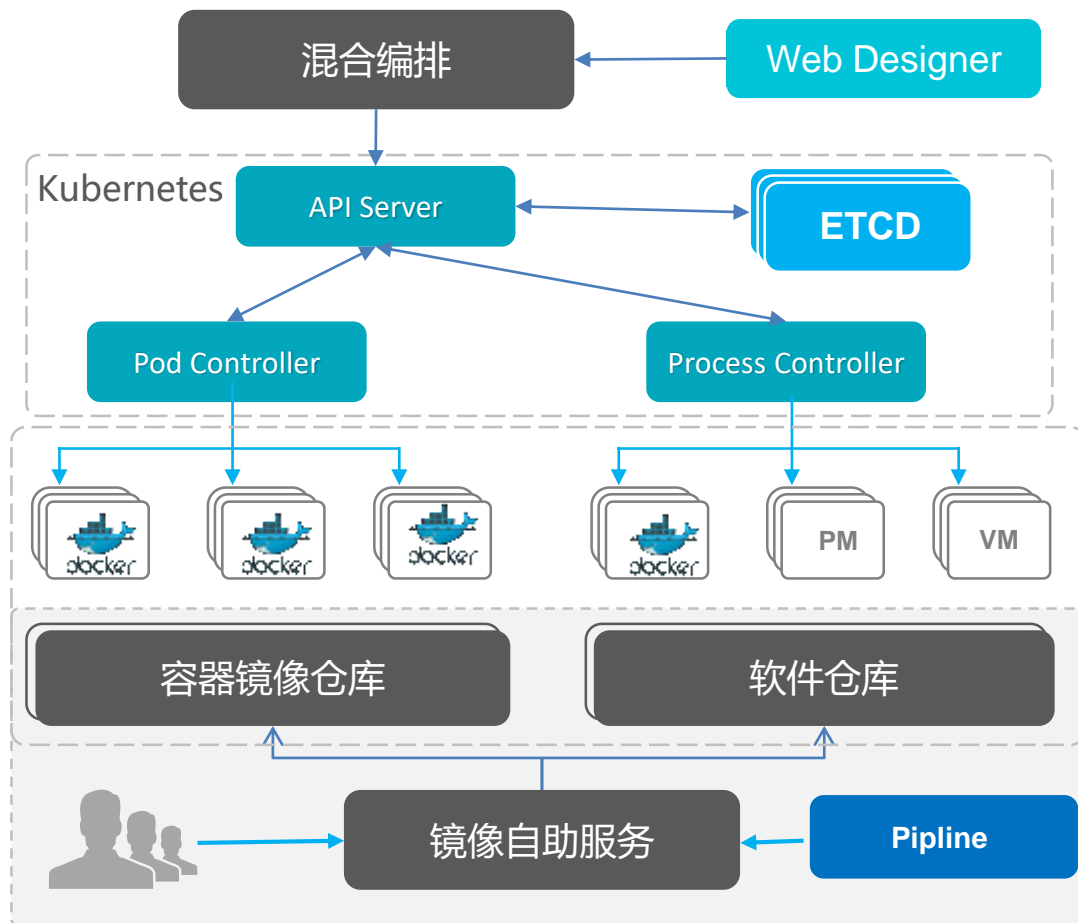
支持物理机、虚拟机、容器的统一资源调度，网络互通

安全仓库

Docker镜像仓库、软件仓库，支持权限控制、安全扫描、分布式存储

自助服务

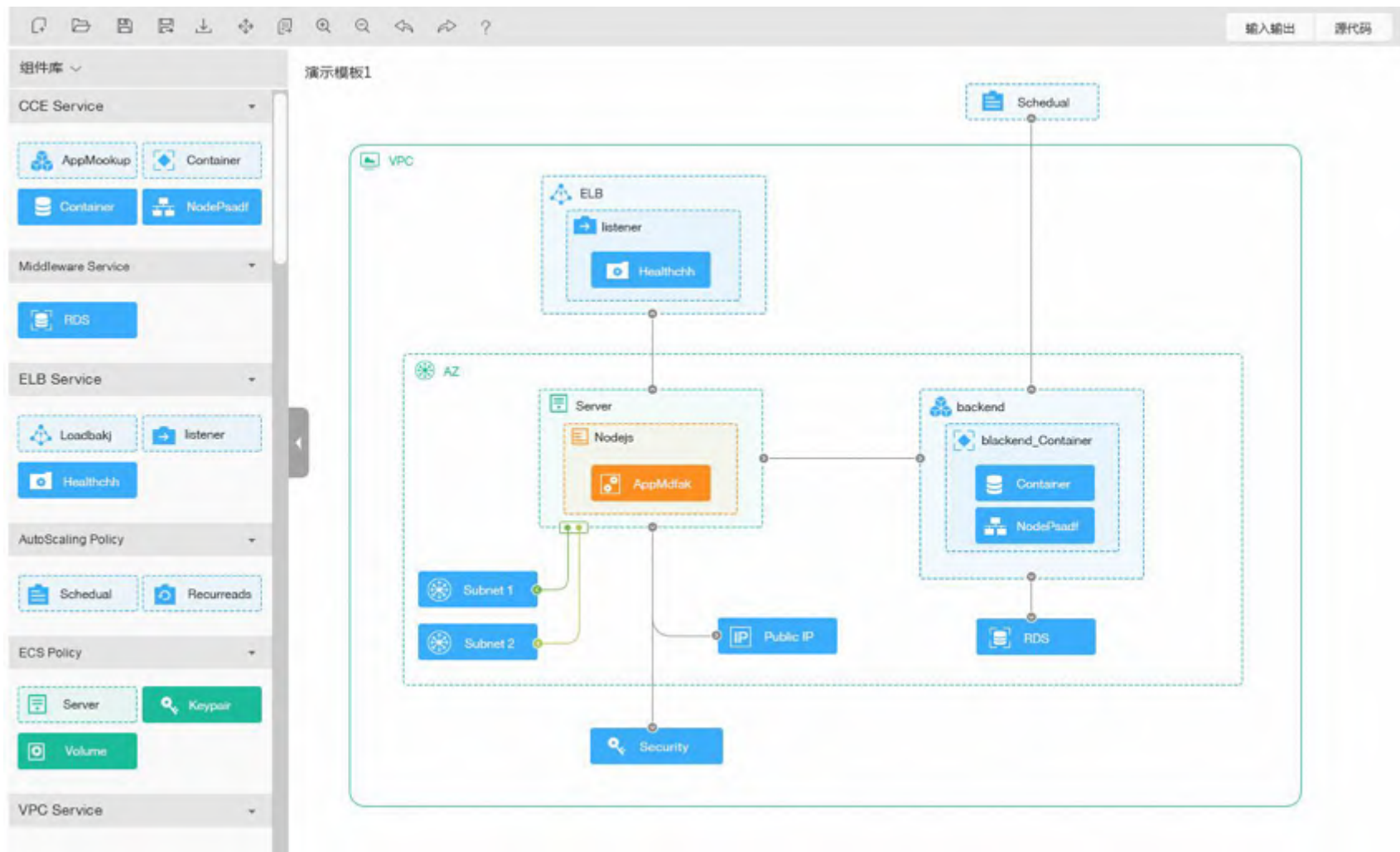
自动化镜像构建，一键式自助服务



拖拽式应用部署，**资源一体化**，屏蔽异构环境，**节省95%**的人工操作时间，操作性强

需求实践

支持应用混合编排，实现虚拟机、物理机、容器的统一调度



万级容器运维管理系统

采集端

统一Agent管理框架

存储

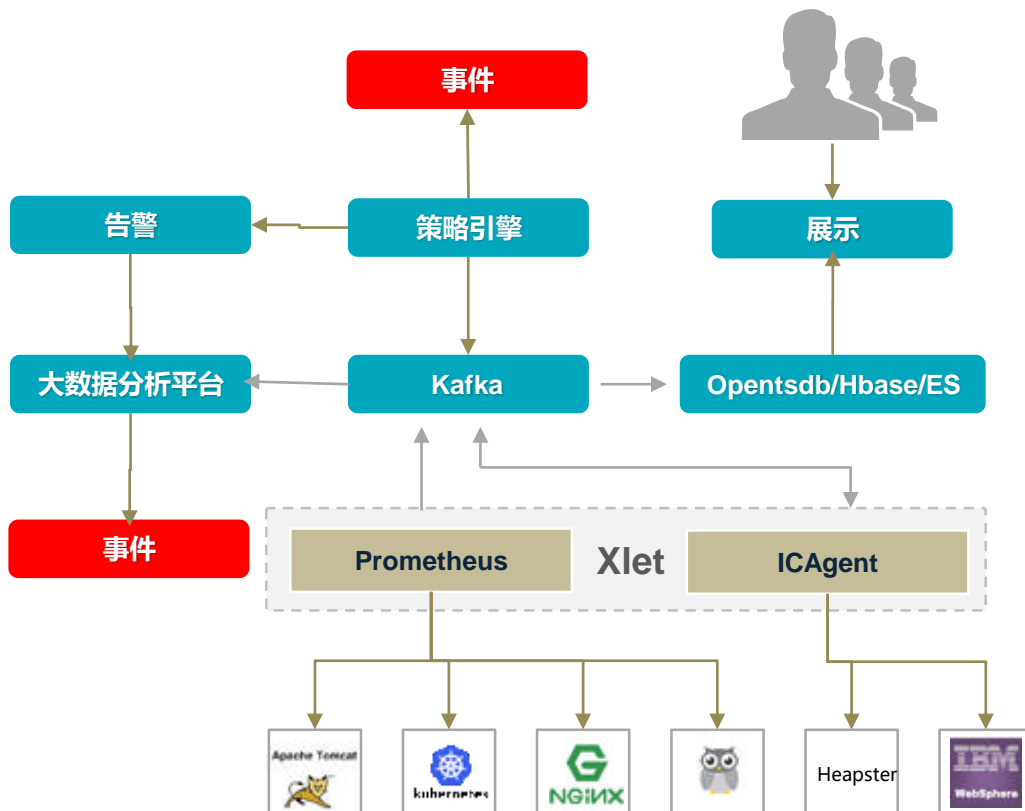
支持Opentsdb/Hbase/ES, 支持跨数据中心

智能分析

策略告警汇聚, 经过大数据平台分析产生事件
实时流分析&线下分析
产生事件以后的自动弹性扩缩容

展示

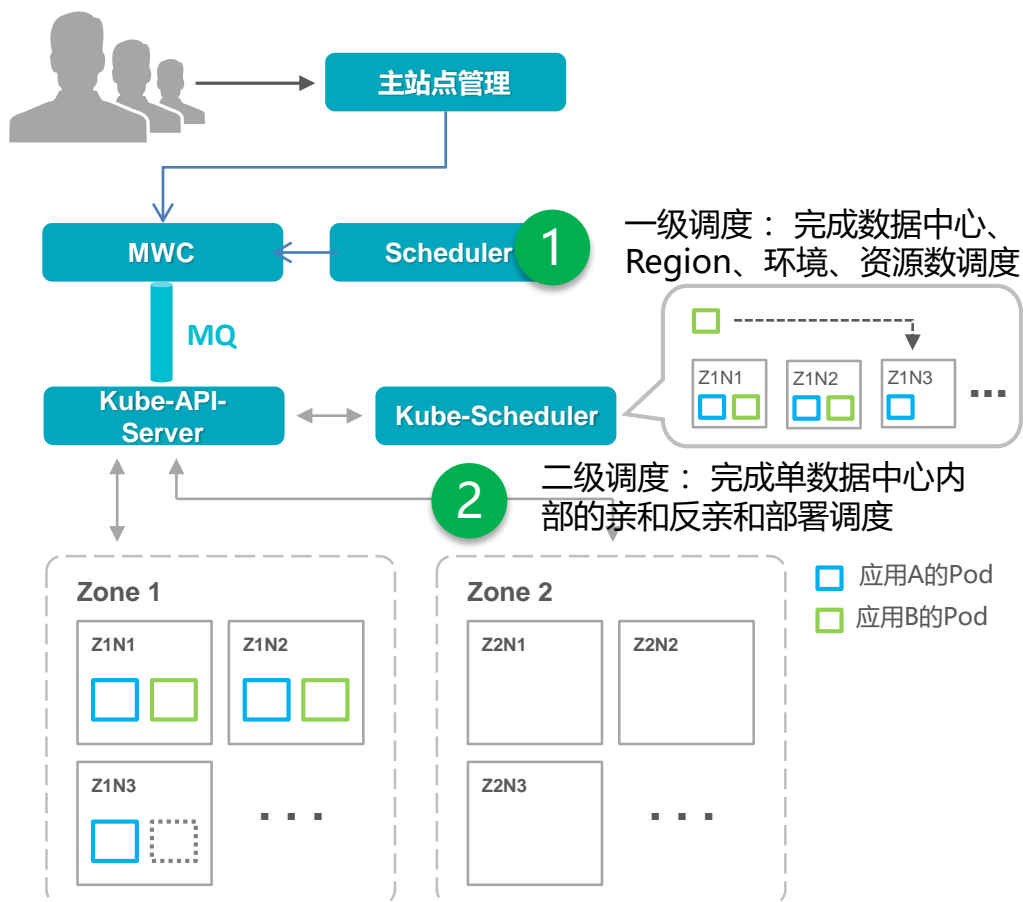
支持管理员视角、应用视角多种展示
自定义Dashboard制作



支持**万级容器**的实时监控数据采集、存储、聚合分析与展示

需求实践

跨数据中心的两级调度：亲和反亲和，提高部署可靠性&部署效率



自研调度算法：

1. 本地镜像：亲和（提速）
2. 创建失败次数：反亲和（Failover）

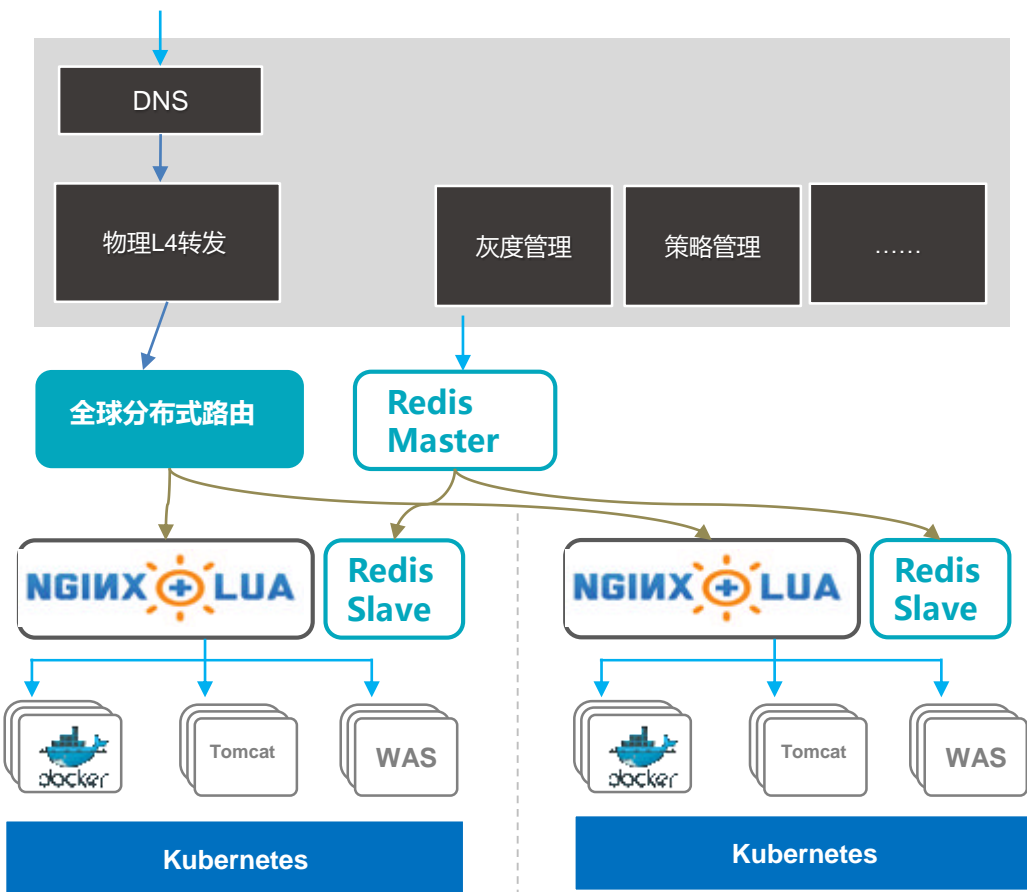
开源调度算法：

1. CPU余量：亲和
2. Mem余量：亲和
3. 容器个数：反亲和
4. 同个应用部署：反亲和
5. 同个app实例尽量分散在不同的虚拟机

解决跨数据中心调度问题，应用端到端部署调度时间从原来的5分钟降低到20s

需求实践

全球分布式动态路由，快速访问体验，应用灰度，达5个9可靠性



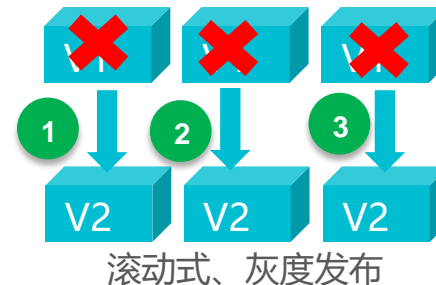
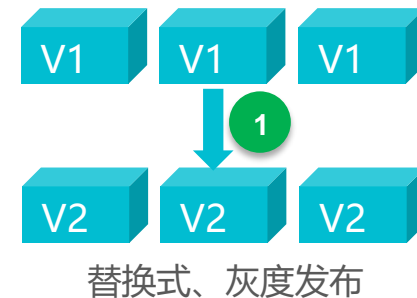
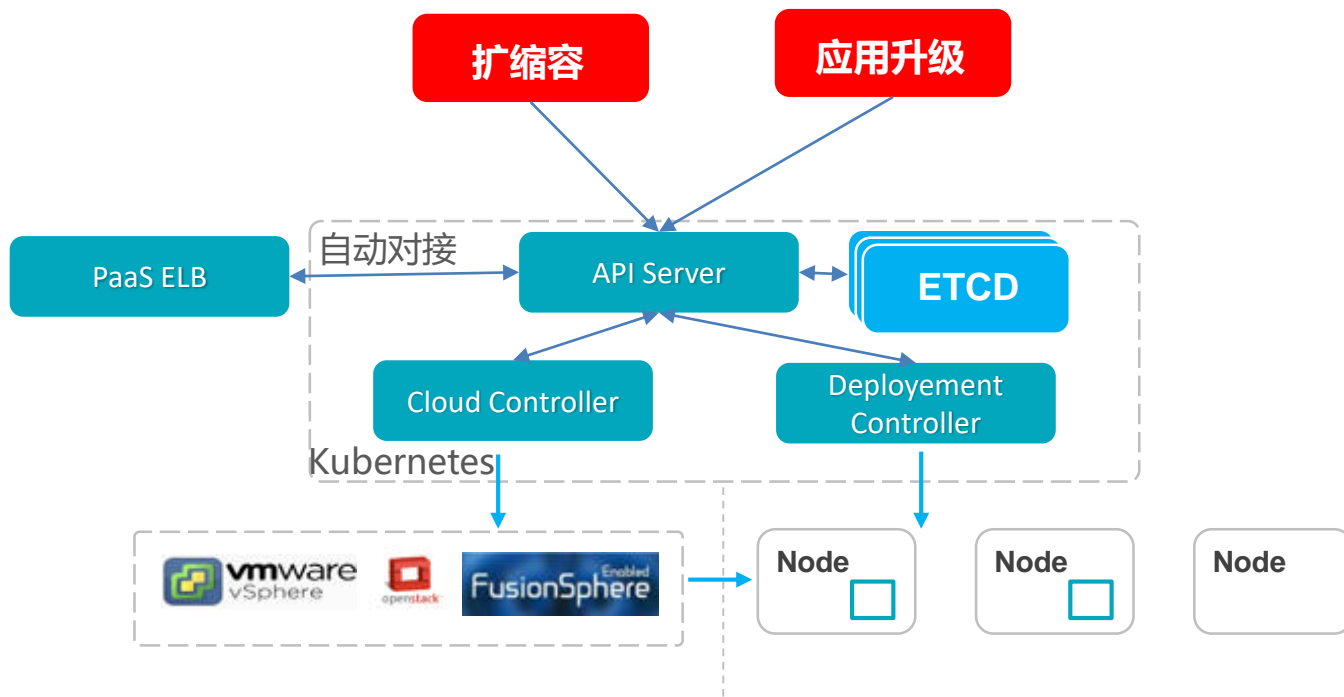
中间件ELB功能：

- 支持企业级WebSphere应用WS-AT事务协议
- 动态路由配置，无需重启
- 提供Nginx Plus商用版本级功能，会话保持、健康检查
- 支持跨EDC的失效转移
- 基于策略的灰度发布

支撑华为IT内部系统**域名260多个**，平均访问流量达**2千万次/天**，访问时延微秒级别，达5个9可靠性。

需求实践

应用灰度发布、自动弹性扩缩



扩缩容
应用自发现，ELB自动对接
虚拟机、容器混合扩缩容支持

应用升级
滚动灰度升级应用无中断
替换方式升级，资源需求量较大

支持应用灰度发布，异构资源弹性扩缩容，自动ELB对接，**实现无中断升级与扩缩容**

需求实践

传统IT

-  100个虚拟机
-  20%虚拟化消耗
-  部署&弹性时间：一周
-  运维部署靠人工，量小
-  中断升级，路由切换
-  CI CD未打通
-  不支持跨站点迁移
-  大网直通消耗大量网络资源

Fusionstage

-  20个虚拟机
-  8%虚拟化消耗
-  部署&弹性时间：5分钟
-  百万级全自动化运维平台
-  灰度发布应用0中断升级
-  Devops全流程打通
-  跨站点应用快速迁移
-  网络共享、端口分配

价值分析

问题回顾

K8s Service引起的访问性能问题

服务自发现，网络4层直通，端口分配，性能相比Overlay提升2-5倍

问题描述：

测试发现1000台节点，600个Service，4000个Pod，每个Pod一个容器，最终Kube_proxy会在所有的节点上产生近18000条Iptable数据，且每次更新路由表都会耗时5-10s。

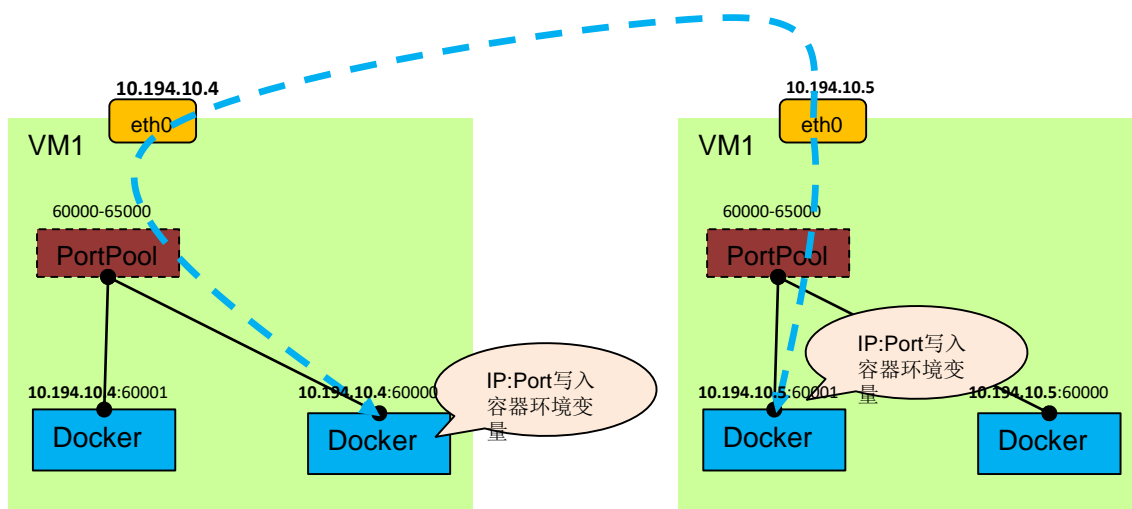
网络优化：

1. 端口管理：Kubelet端口池统一管理。
2. 应用自发现：容器所在节点IP:Port发布到环境变量。
3. 动态对接ELB：动态更新路由。

 容器应用自发现

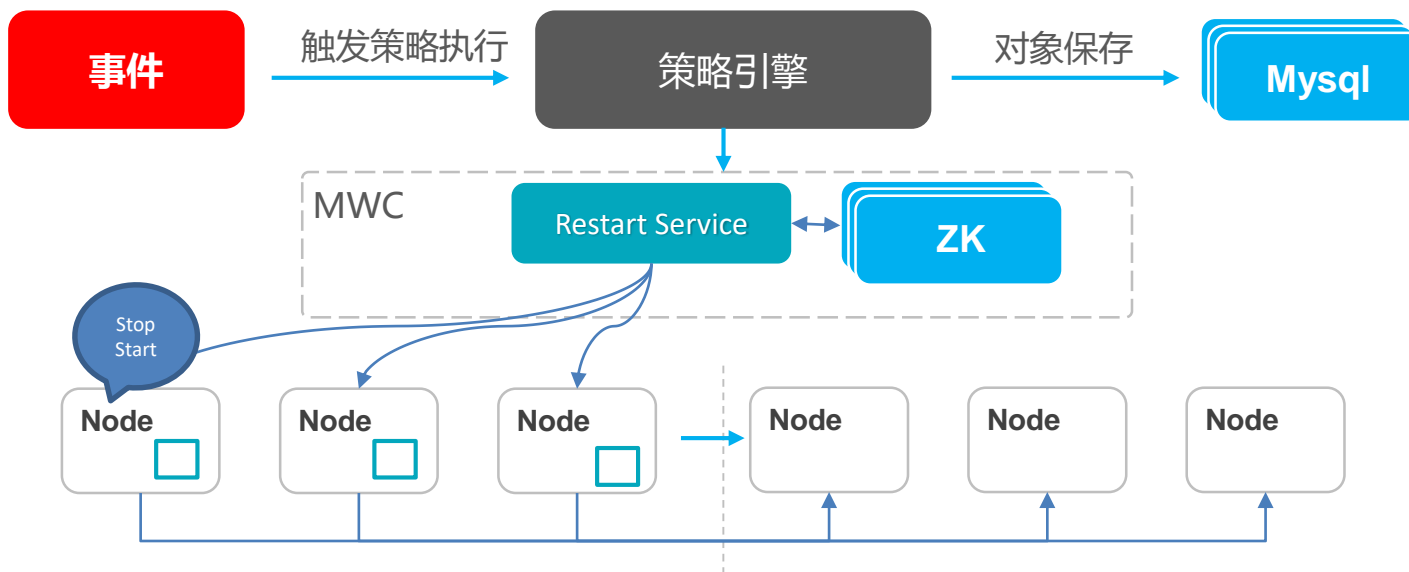
 全网四层打通

 端口管理、防止冲突



机器部署Agent繁多，莫名Hang住！

定期机器重启，减少遗留垃圾，作为应用平滑升级的试炼场



Why ?

How ?

Q&A



平台共建 · 市场合作 · 产品咨询

时间：2016年11月20日



SDCC 2016

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

谢谢！

吴建林

wujianlin@huawei.com

