

构建10亿级商品的电商平台架构（微店）

陈国成

① 微店是谁，技术挑战是什么

- 业务模式
- 第一代架构
- 新的技术挑战

② 为交易构筑安全可靠的防火墙

- 日均经受560万次攻击
- 日均被爬6亿次

③ 低成本的架构建设之道

- 私有云
- 全站分布式
- 性能优化
- SRE, devops

④ 大数据下的基础建设

- 数据层治理
- 中间件

⑤ 大数据面前，业务系统的演进之路

- 搜索引擎



全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

微店

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

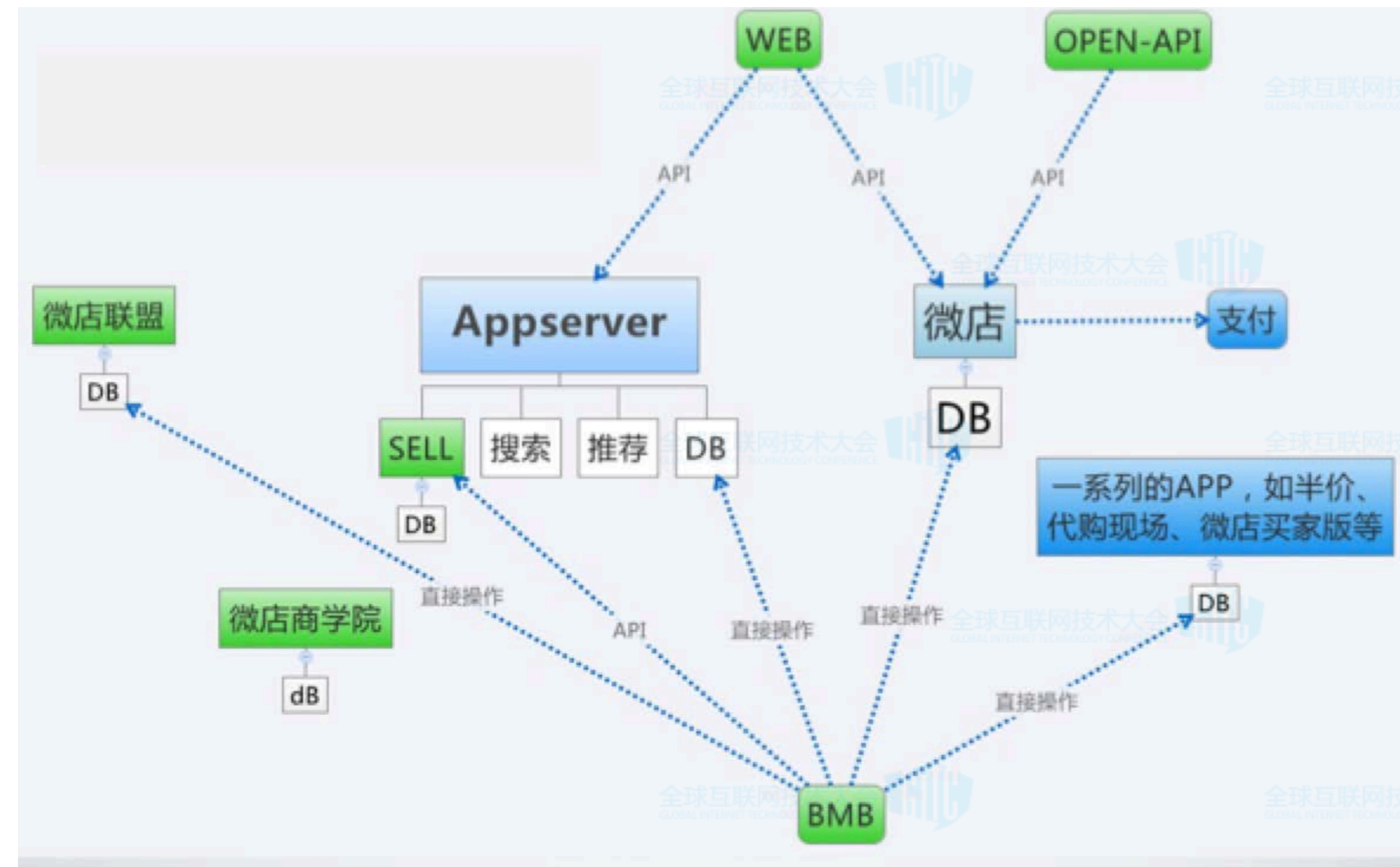
全球互联网技术大会
GLOBAL INTERNET TECHNOLOGY CONFERENCE

手机开店

7000万店铺

月访客过亿

微店



10 - 300人
海豹突击队

LAMP、F5、redis
技术特点

100w - 3000w
注册卖家数



规模

7000W

- 7000万店铺，10亿商品；9P数据
- 支付、搜索、推荐、风控、IM、交易、开放平台、广告、供应链...

安全

6亿/天

- 经受560万次攻击/天；CC、SQLInjection、XSS、CSRF ...
- 每天6亿+爬虫访问...

成本

数千万/年

- IDC方面，2016/2017年整体支出数千万

效率

600+

- 600+ 研发一起协同
- 每天有60次上线，400+次发布（含测试/预发）



① 微店是谁，技术挑战是什么

- 业务模式
- 第一代架构
- 新的技术挑战

② 为交易构筑安全可靠的防火墙

- 日均经受560万次攻击
- 日均被爬6亿次

③ 低成本的架构建设之道

- 私有云
- 全站分布式
- 性能优化
- SRE, devops

④ 大数据下的基础建设

- 数据层治理
- 中间件

⑤ 大数据面前，业务系统的演进之路

- 搜索引擎



②为交易构筑安全可靠的防火墙

A.主要的安全威胁

B.微店安全产品框架

C.WAF（防火墙）

D.代码扫描

E.主机安全防护HIDS

F.实时安全日志分析风险感知

CC
DDOS
设备漏洞

网络安全

SQL Injection
XSS
CSRF
命令执行
移动安全
SSRF
应用逻辑漏洞

应用安全

木马
蠕虫
Rootkit
软件漏洞

主机安全

帐号被盗
恶意数据抓取
内部违规行为
数据安全

其他类型



②为交易构筑安全可靠的防火墙

A.主要的安全威胁

B.微店安全产品框架

C.WAF (防火墙)

D.代码扫描

E.主机安全防护HIDS

F.实时安全日志分析风险感知

安全运营

安全流程

安全规范

渗透测试

安全评审

安全培训

应急响应

WAF

漏洞扫描

代码缺陷扫描

设备
指纹

入侵检测

DDOS防护

主机安全agent

数据库防火墙

SOC

漏洞管理

实时安全
日志分析
风险感知
系统



②为交易构筑安全可靠的防火墙

A. 主要的安全威胁

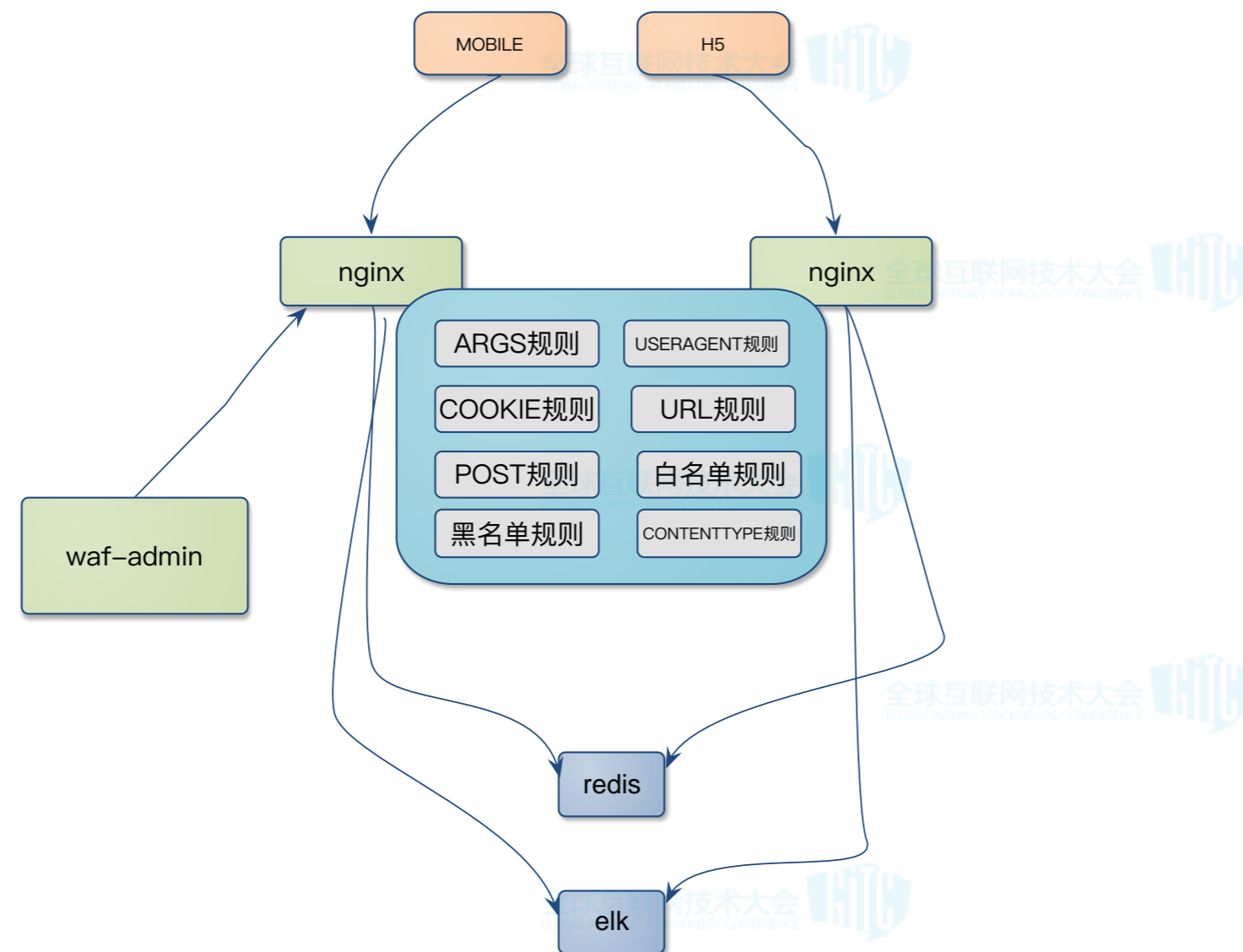
B. 微店安全产品框架

C. WAF (防火墙)

D. 代码扫描

E. 主机安全防护HIDS

F. 实时安全日志分析风险感知



WAF

- 统一接入层 (lua on nginx) 运行。对业务无侵入，且100%全覆盖。
- 实时规则防护，0day防护。
- 虚拟补丁，规则动态实时更新

微店

②为交易构筑安全可靠的防火墙

A. 主要的安全威胁

B. 微店安全产品框架

C. WAF (防火墙)

D. 代码扫描

E. 主机安全防护HIDS

F. 实时安全日志分析风险感知

微店代码

任务

漏洞管理 任务管理 项目详情

Vulnerabilities

- AjaxPage.class.php
- vc-web/vcweb_third
- auth_util.php
- vc-web/vcweb_third
- examples.php
- vc-web/vcweb_third
- Process.php
- vc-web/vcweb_third
- functions.php
- vc-web/vcweb_third
- functions.php
- vc-web/vcweb_third
- phprpc_server.php
- vc-web/vcweb_third

```
256         else {
257             $arguments = $args;
258         }
259         return call_user_func_array($function, $arguments);
260     }
261     function getRequest($name) {
262         $result = $_REQUEST[$name];
263         if (get_magic_quotes_gpc()) {
264             $result = stripslashes($result);
265         }
266         return $result;
267     }
268     function getBooleanRequest($name) {
269         $var = true;
270         if (isset($_REQUEST[$name])) {
271             $var = strtolower($this->getRequest($name));
272             if ($var == "false") {
273                 $var = false;
274             }
275         }
276         return $var;
```

状态:

描述: Reflection Injection

类型: rips

加白 确认 向下 向上

代码扫描

- DSL
 - ✓ 支持java、php、node等语言
- 扫描模式
 - ✓ 动态（语法分析）、静态（正则表达）以及集成第三方扫描引擎（rips、findsecbugs等）相结合的扫描模式
- 发布系统强结合，强制安全扫描



②为交易构筑安全可靠的防火墙

A. 主要的安全威胁

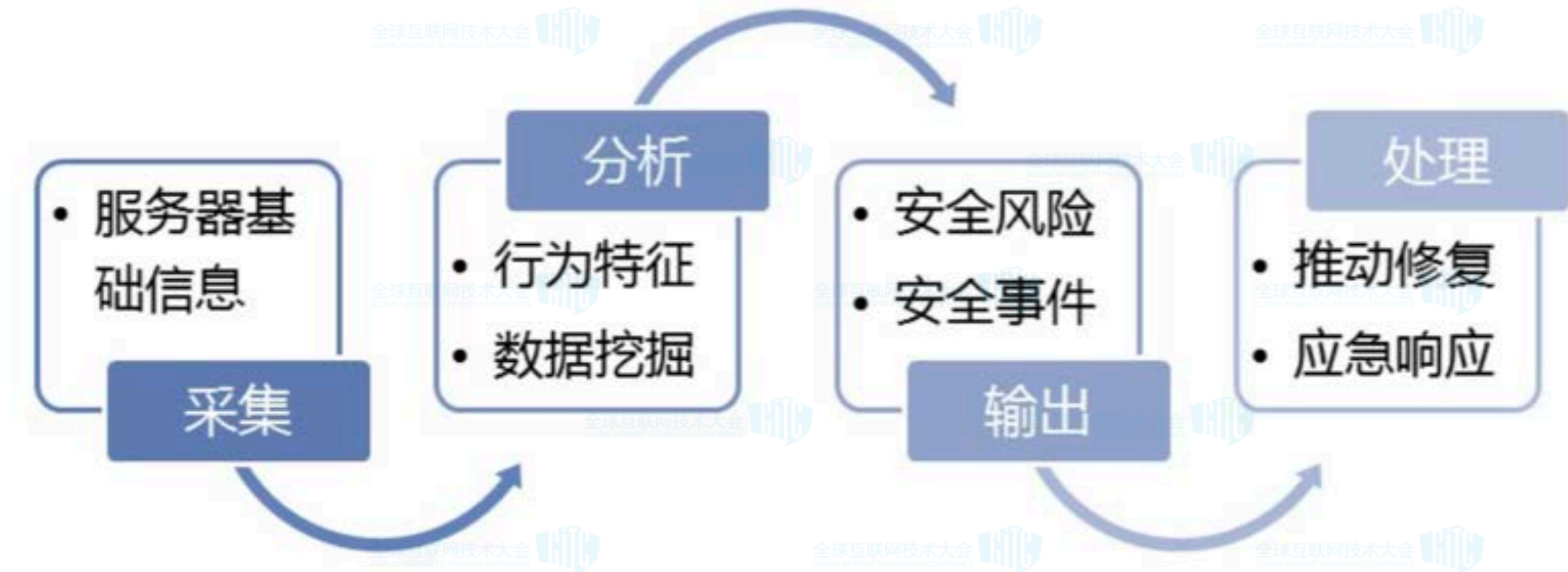
B. 微店安全产品框架

C. WAF (防火墙)

D. 代码扫描

E. 主机安全防护HIDS

F. 实时安全日志分析风险感知



主机安全防护HIDS

- 文件监控
- 命令监控
- 网络监控
- 基线巡检

微店

②为交易构筑安全可靠的防火墙

A.主要的安全威胁

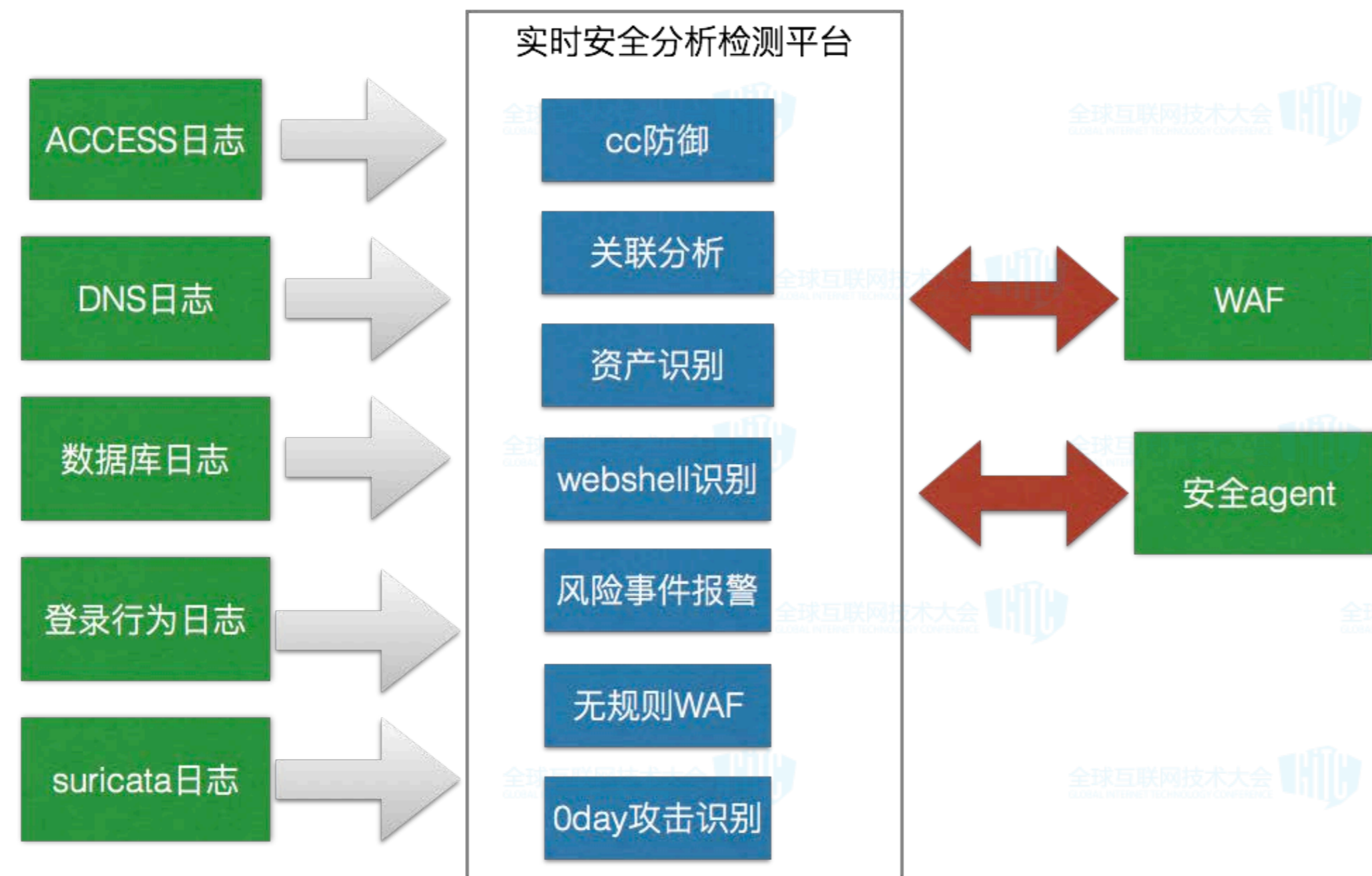
B.微店安全产品框架

C.WAF（防火墙）

D.代码扫描

E.主机安全防护HIDS

F.实时安全日志分析风险感知



实时安全日志分析风险感知系统

- 建模：
 - 关键字、频率、webshell、httpheader等
 - 去噪
 - 菜刀请求特征
 - 基于统计的模型
- 机器学习攻击检测（进行中）



① 微店是谁，技术挑战是什么

- 业务模式
- 第一代架构
- 新的技术挑战

② 为交易构筑安全可靠的防火墙

- 日均经受560万次攻击
- 日均被爬6亿次

③ 低成本的架构建设之道

- 私有云
- 全站分布式
- 性能优化
- SRE, devops

④ 大数据下的基础建设

- 数据层治理
- 中间件

⑤ 大数据面前，业务系统的演进之路

- 搜索引擎



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G. 以应用为中心的运维体系

每年净增400+服务器

IDC成本

70%的服务器利用率在15%以下

服务器利用率

全站大部分页面
90%的访问首屏渲染完成时间在3s以上

性能

测试开发比1:2
OP高峰时达到18人

人效及成本



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

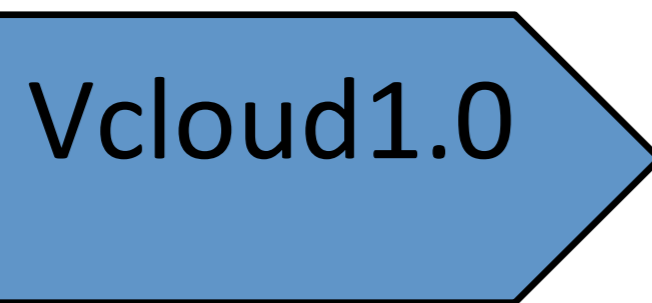
C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

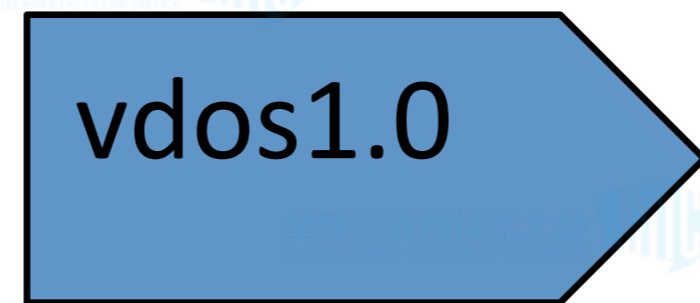
G. 以应用为中心的运维体系



- KVM
- IAAS化



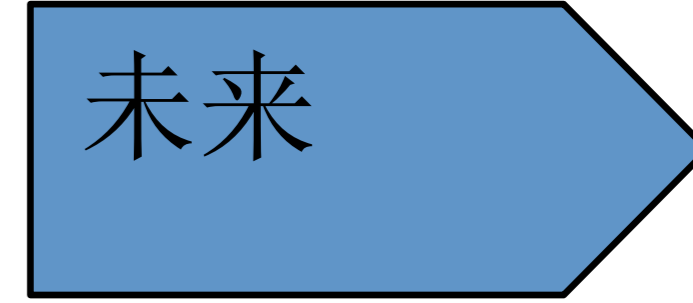
- 完成docker的IAAS化平台管理
- 非线上环境全面推广docker
- 资源快速交付；秒级



- PAAS化，容器or虚拟机创建即部署
- 应用平滑上下线
- docker（通过镜像）
- KVM（通过开机脚本）
- 性能监控
- 容器服务编排功能；降低TCO



- 自动化、平台化
- IDC级别的弹性
- 服务自动上线。
- 日志分析收集
- KVM，容器，物理机混合编排



- 在线迁移
- 在线离线混部
- 超融合云
- 混合云
- 智能调度



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G. 以应用为中心的运维体系

- OS版本
 - Centos7.2
- 虚拟化技术
 - KVM
 - Docker
- 开发语言
 - 前端
 - Golang
 - Bootstrap, angularjs
 - 后端
 - Golang
 - shell

- 网络模式
 - 网桥
 - Pipework
 - libvirt
 - 配置唯一IP，且全网互通。
- 存储模式
 - 本地LVM（OS+数据盘/数据卷）
 - Ceph统一镜像存储
- 资源控制
 - 网络—TC
 - CPU, MEM, DISK---cgroup



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

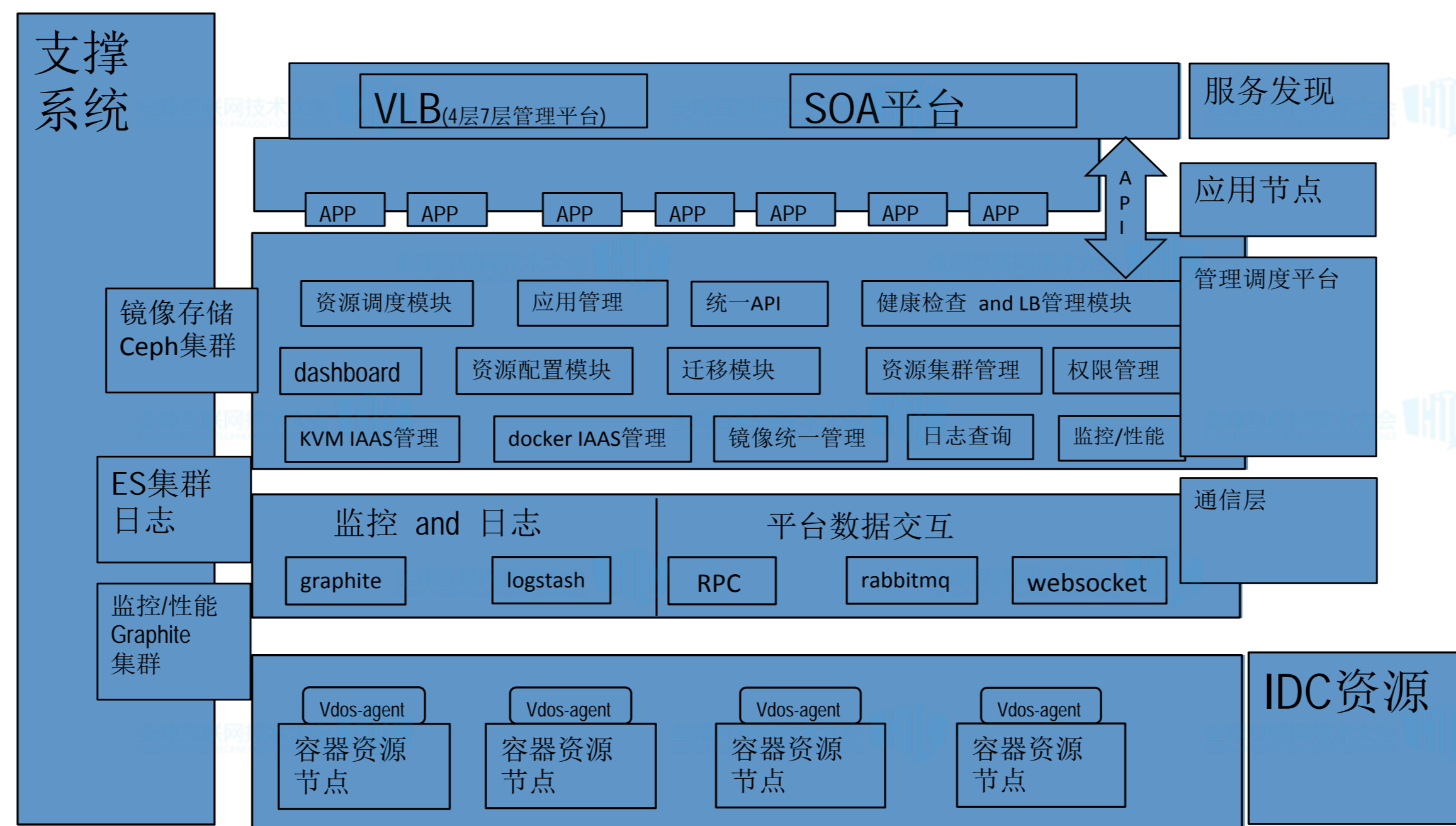
C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G. 以应用为中心的运维体系



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

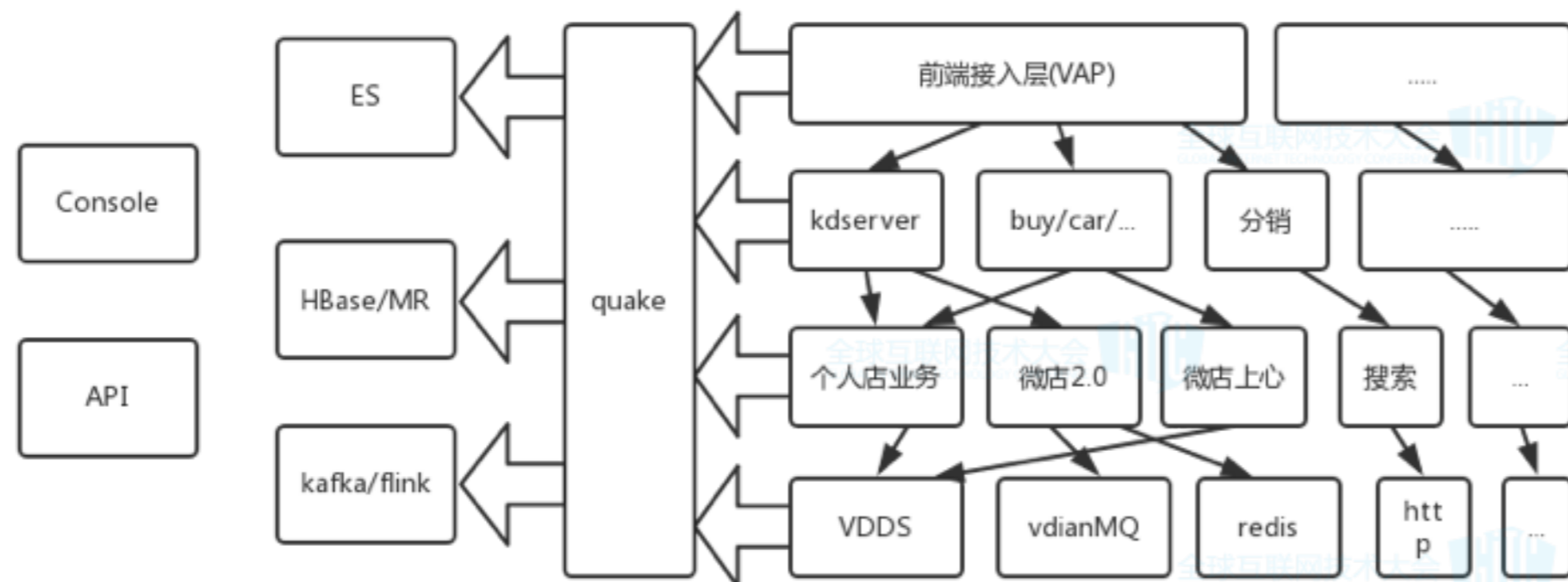
C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G.以应用为中心的运维体系



理论基础：分布式跟踪系统鼻祖 dapper

traceID：应用生成，放到tomcat

每个RPC调用生成一个span，同时记录父span

在RPC框架上完成，对业务零侵入

Hex string输出

traceID：seq+时间戳+机器IP+进程号

与其他中间件的全方位整合：tomcat/dubbo/http/vdianMQ/vdds/redisClient

Rpc ID	AppName	Remoteip	Trace Type	Status	Trace Name	Size	Time Line	Message
0	vap-server(10.2.131.202)	172.19.35.116	HTTP	OK	/h5/ans/item.getRecommendItems/1.0	1	200ms	200
0.1	v aree(10.2.101.81)	10.2.101.81	DUBBO	OK	com.vdian.vap.common.aresService.item.getRecommen15982		292ms	OK
0.1.1	pluto(10.2.114.91)	10.2.114.91	DUBBO	OK	com.vdian.pluto.client.service.RecEngineService.execu13895		230ms	OK
0.1.1.1	mercury(10.2.129.88)	10.2.129.88	DUBBO	OK	com.vdian.mercury.service.MercuryService.service:1.0	846	5ms	OK
0.1.1.2	mercury(10.2.129.88)	10.2.129.88	DUBBO	OK	com.vdian.mercury.service.MercuryService.service:1.0	3000	14ms	OK
0.1.1.3	mercury(10.2.129.88)	10.2.129.88	DUBBO	OK	com.vdian.mercury.service.MercuryService.service:1.0	861	14ms	OK
0.1.1.4	mercury(10.2.129.88)	10.2.129.88	DUBBO	OK	com.vdian.mercury.service.MercuryService.service:1.0	861	13ms	OK
0.1.1.5	mercury(10.2.129.88)	10.2.129.88	DUBBO	OK	com.vdian.mercury.service.MercuryService.service:1.0	26594	20ms	OK
0.1.1.6	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	DUBBO	OK	com.koudai.ferxiao.client.service.FxItemService.queryI763		144ms	OK
0.1.1.6.1	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	SINGLE	0	2ms	SELECT,fx_item_info,10.2.117.
0.1.1.6.2	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	1ms	ok
0.1.1.6.3	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.4	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	0ms	ok
0.1.1.6.5	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.6	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	0ms	ok
0.1.1.6.7	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.8	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	0ms	ok
0.1.1.6.9	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.10	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	0ms	ok
0.1.1.6.11	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.12	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	1ms	ok
0.1.1.6.13	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.14	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	supplyItemSetting_mGet	0	0ms	ok
0.1.1.6.15	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	REDIS	OK	ferxiao-core1_get	0	0ms	ok
0.1.1.6.16	ferxiao-core(10.2.1210.2.129.168)	10.2.1210.2.129.168	DUBBO	OK	com.vdian.vmp.client.service.detail.DetailPromotionQue1016		56ms	OK
0.1.1.6.16.2	vmpcoupon(10.2.1210.2.129.168)	10.2.1210.2.129.168	DUBBO	OK	com.koudai.vmp.service.SellerShopCouponReadService443		0ms	OK
0.1.1.6.16.2.1	vmpcoupon(10.2.1210.2.129.168)	10.2.1210.2.129.168	VDDS	OK	SINGLE	0	1ms	SELECT,vmp_shop_coupon,10.



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

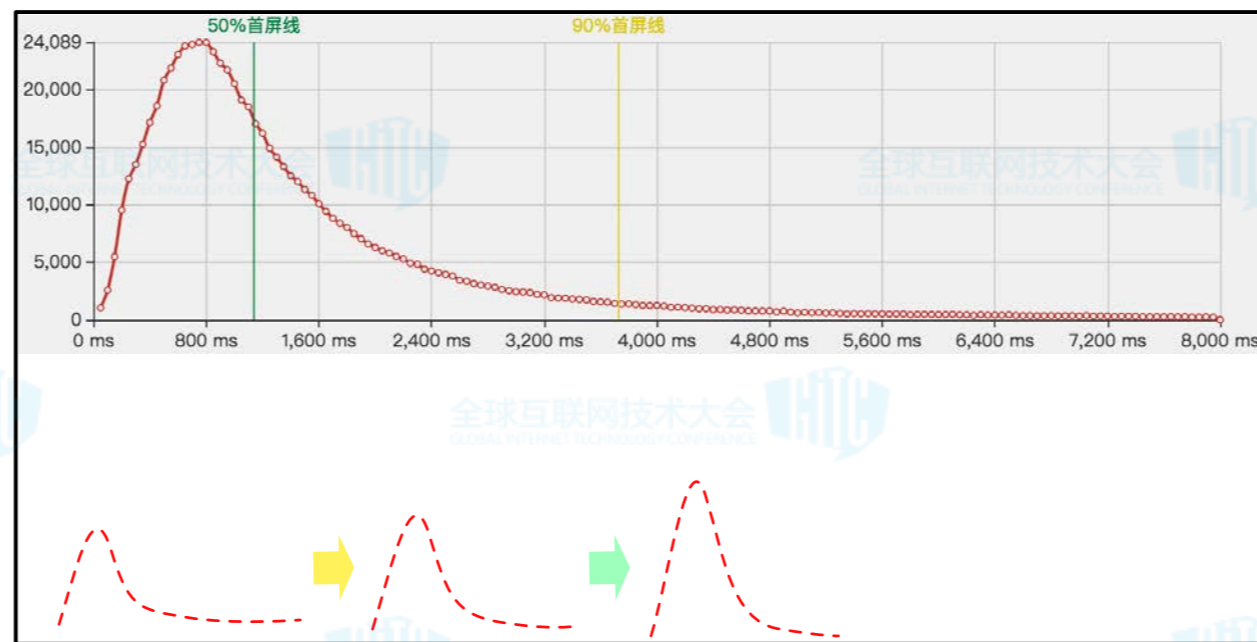
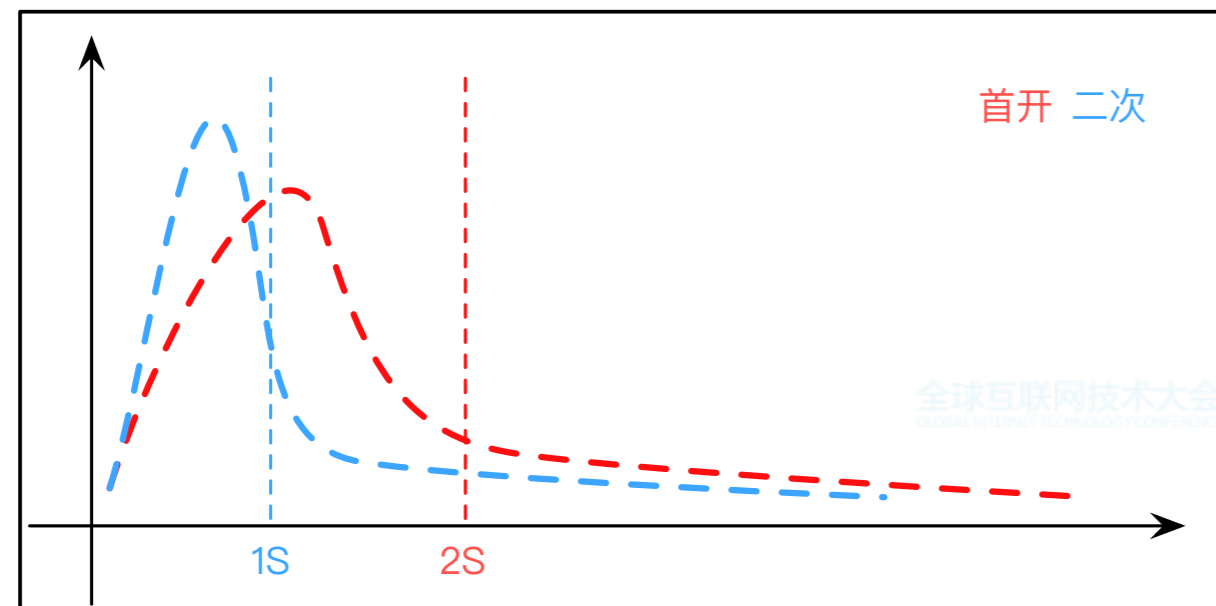
C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G. 以应用为中心的运维体系



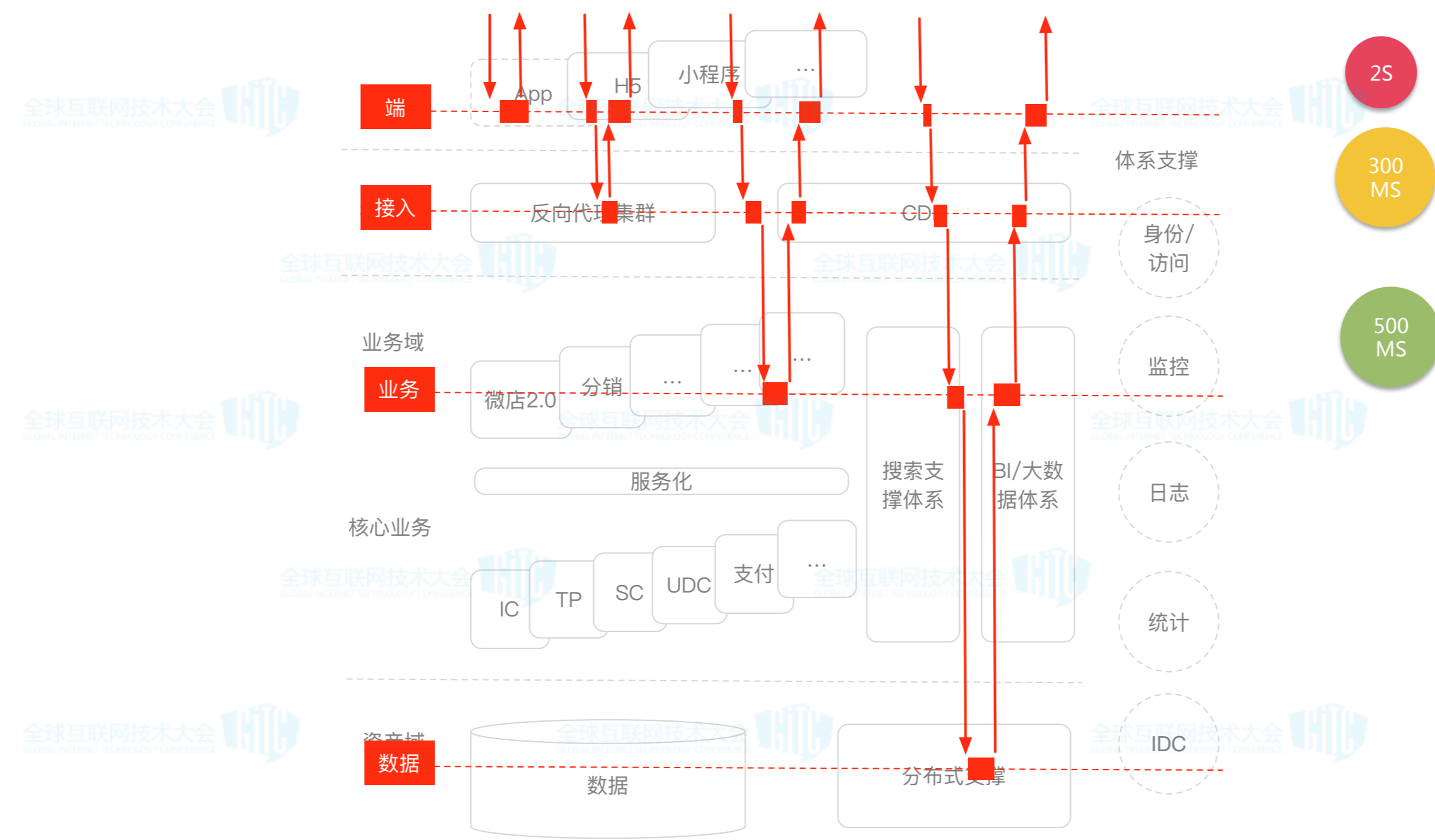
数据评估参考：

- 页面加载速度每提高1秒，转化率增加2%...反之，如果超过4秒，25%的用户会选择离开...
- 用户最满意的打开网页时间是2-5秒，如果等待超过10秒，99%的用户会关闭这个网页
- Google：网站访问速度每慢400ms就导致用户搜索请求下降0.59%
- Amazon：每增加100ms网站延迟将导致收入下降1%
- 雅虎：如果有400ms延迟会导致流量下降5-9%



③低成本的架构建设之道

- A. 成本和性能的挑战
- B. 微店私有云发展历程
- C. 私有云技术选型
- D. 容器管理平台架构
- E. 分布式治理：链路追踪
- F. 性能优化：秒开
- G. 以应用为中心的运维体系



③低成本的架构建设之道

A.成本和性能的挑战

B.微店私有云发展历程

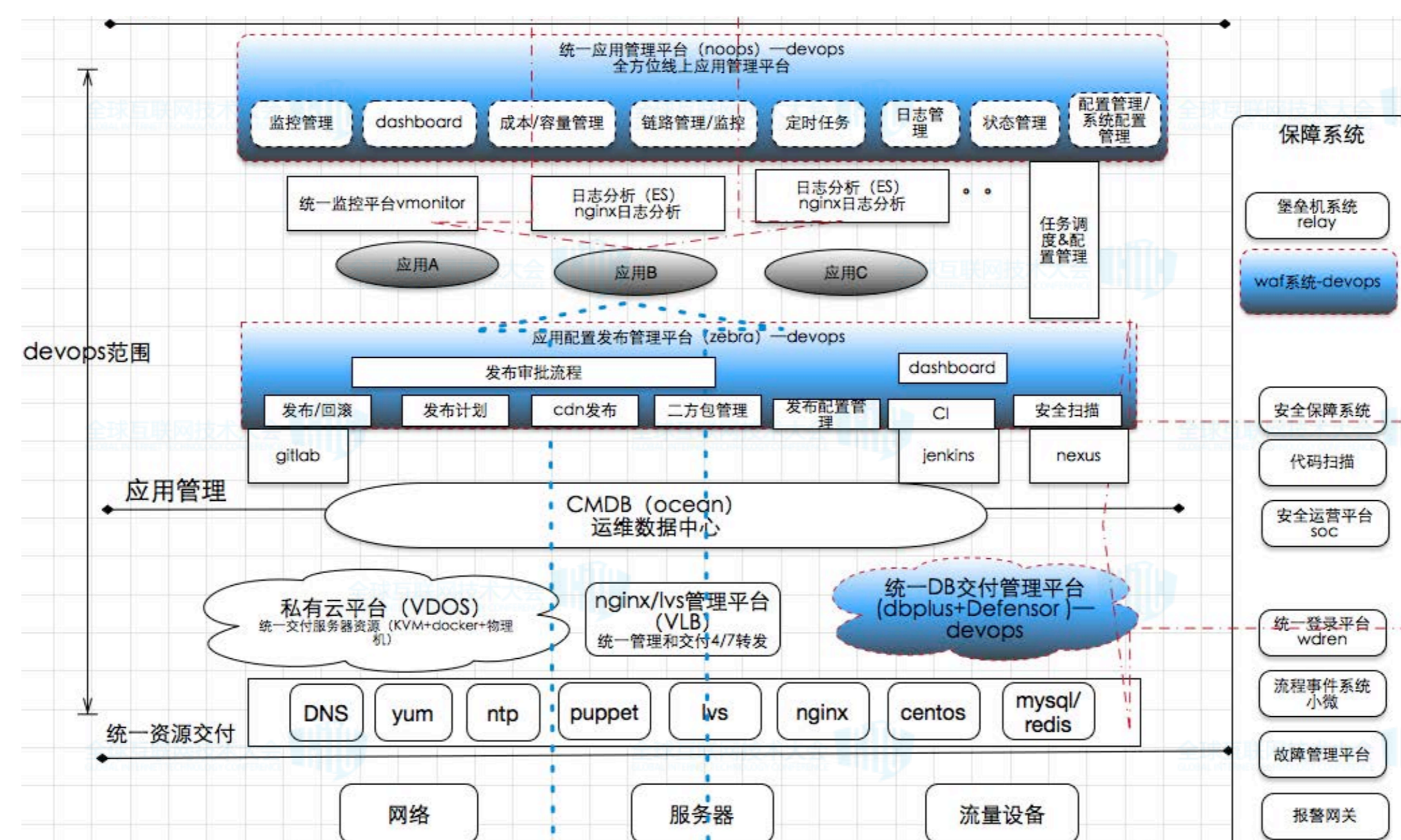
C.私有云技术选型

D.容器管理平台架构

E.分布式治理：链路追踪

F.性能优化：秒开

G. 以应用为中心的运维体系



① 微店是谁，技术挑战是什么

- 业务模式
- 第一代架构
- 新的技术挑战

② 为交易构筑安全可靠的防火墙

- 日均经受560万次攻击
- 日均被爬6亿次

③ 低成本的架构建设之道

- 私有云
- 全站分布式
- 性能优化
- SRE, devops

④ 大数据下的基础建设

- 数据层治理
- 中间件

⑤ 大数据面前，业务系统的演进之路

- 搜索引擎

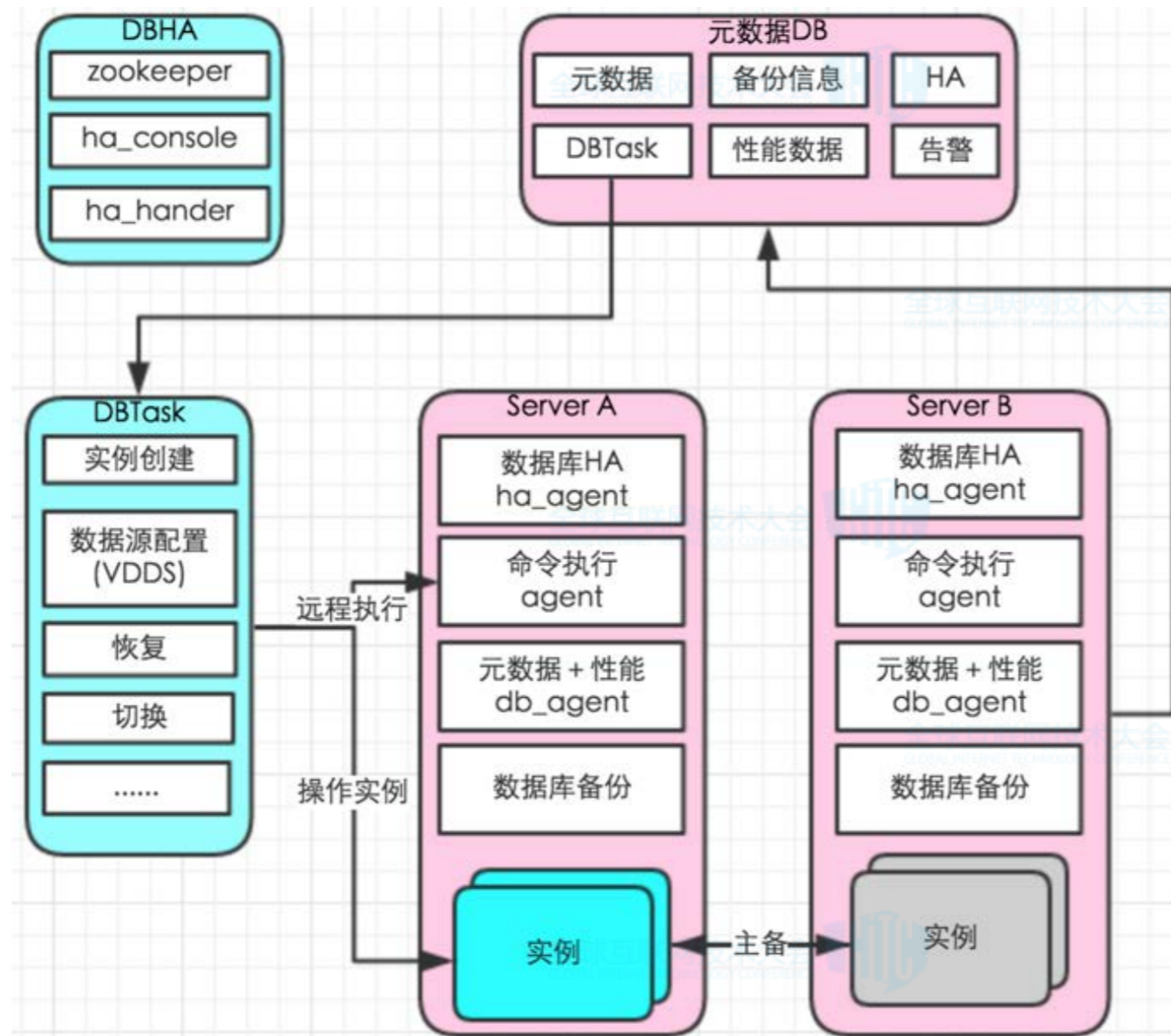


④大数据下的基础建设

A.数据库治理

B.数据层相关中间件

C.分布式事务框架



●建立数据库标准:

- 单机多实例

- 应用独享实例

- MS结构, 主备分工

- HA

- 分布式监控, DBAgent

 - 自动发现实例, 自动采集实例数据, 主机性能数据, 磁盘数据, 自动添加监控

 - 实时慢SQL

 - 每秒实时更新心跳 (show slave status不可靠)

- DBTask

 - 数据库创建, 配置VDDS, 数据库迁移, 拆库扩容, 恢复



④大数据下的基础建设

A.数据库治理

B.数据层相关中间件

C.分布式事务框架

- 构建中间件，提供数据库治理相关的框架支持
 - VDDS，分库分表中间件
 - VDDS-Proxy，分库分表中间件（for PHP）
 - VSS，全量数据同步
 - VTS，增量数据同步
 - vdianMQ，消息中间件
 - Tcc，分布式事务框架
- 建立数据库自动化运维体系

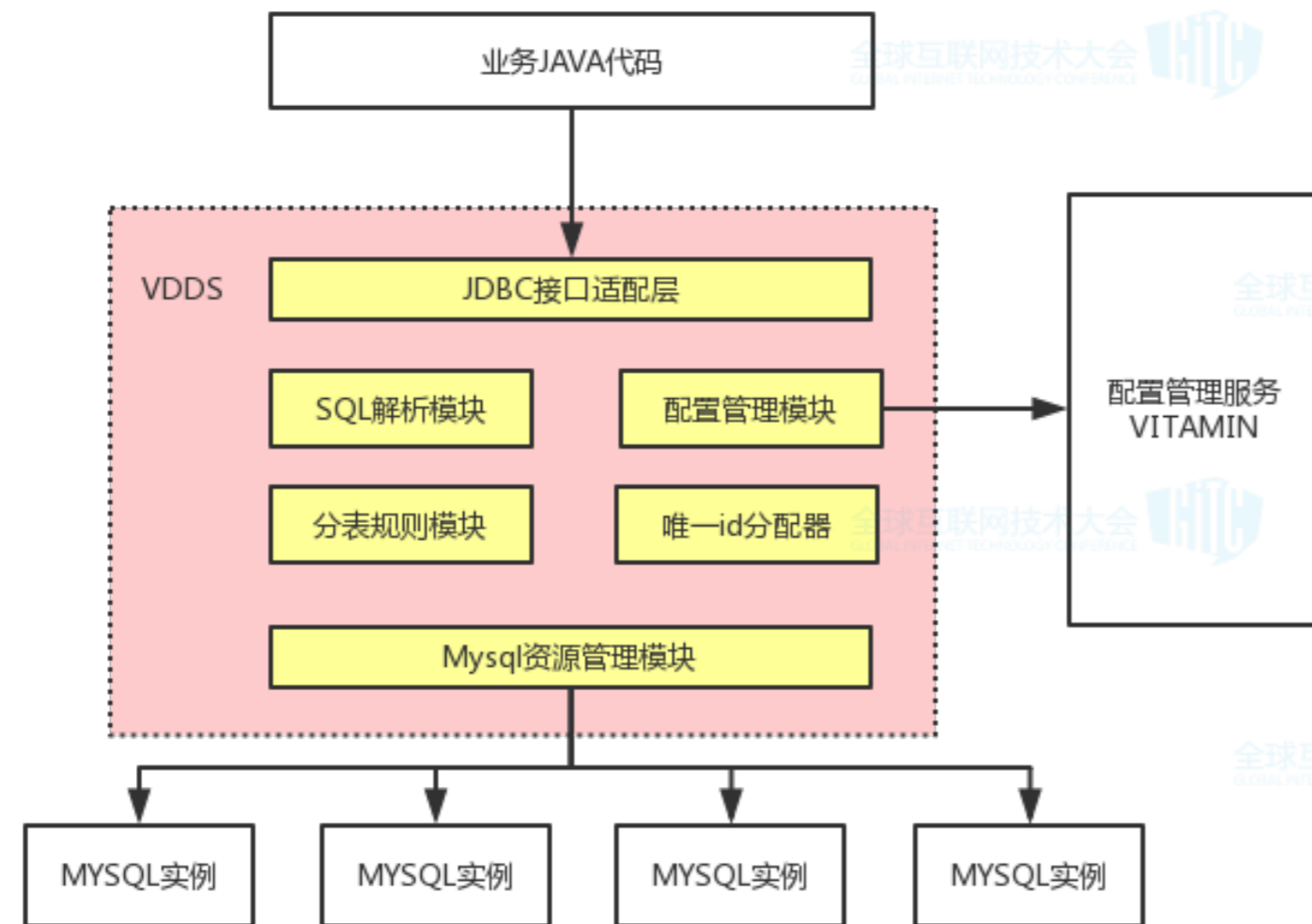


④大数据下的基础建设

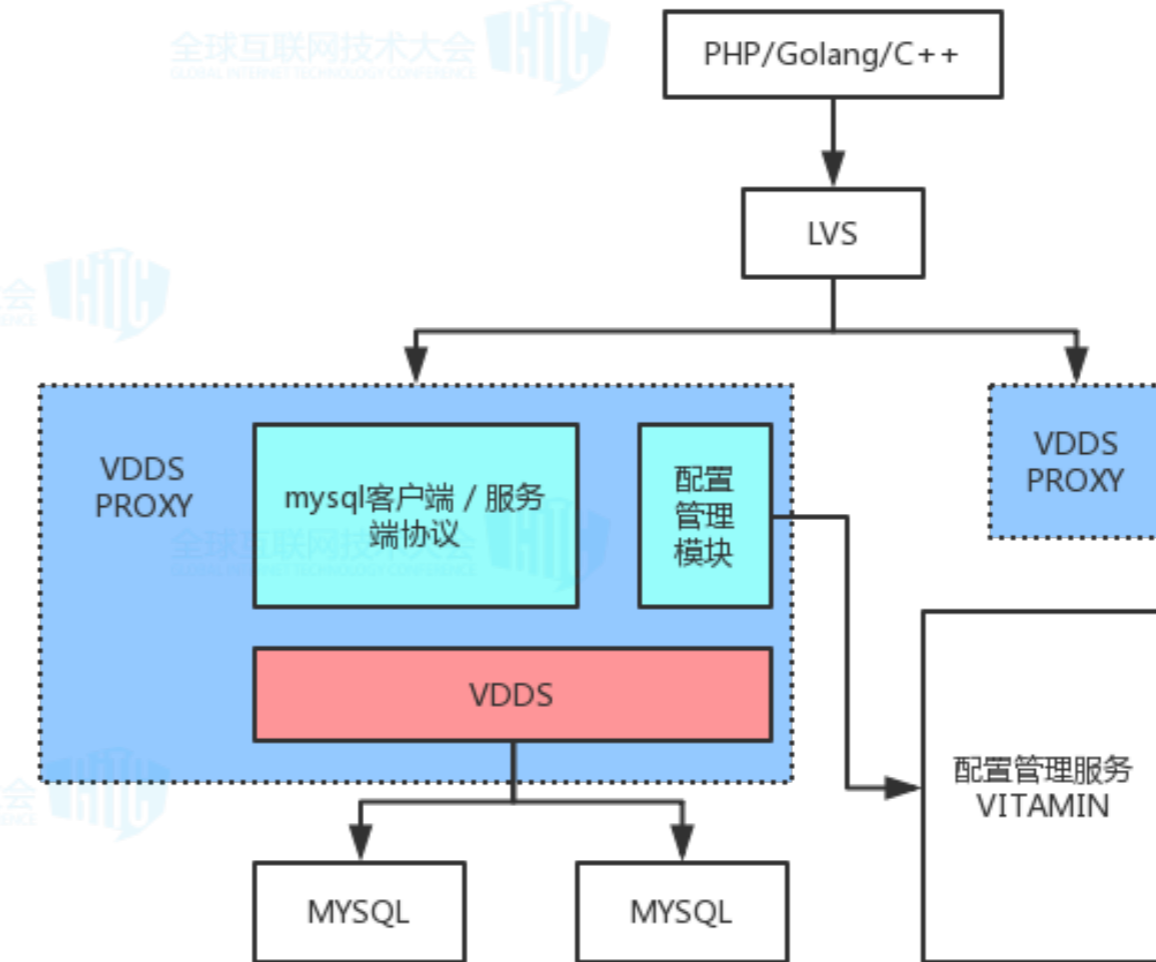
A.数据库治理

B.数据层相关中间件

C.分布式事务框架



VDDS
分库分表
读写分离
独立的账号体系
配置自动变更
灵活的hint机制



VDDS PROXY (for php)

平滑下线

配置自动生效

支持mysql preparedStatement协议

负载均衡

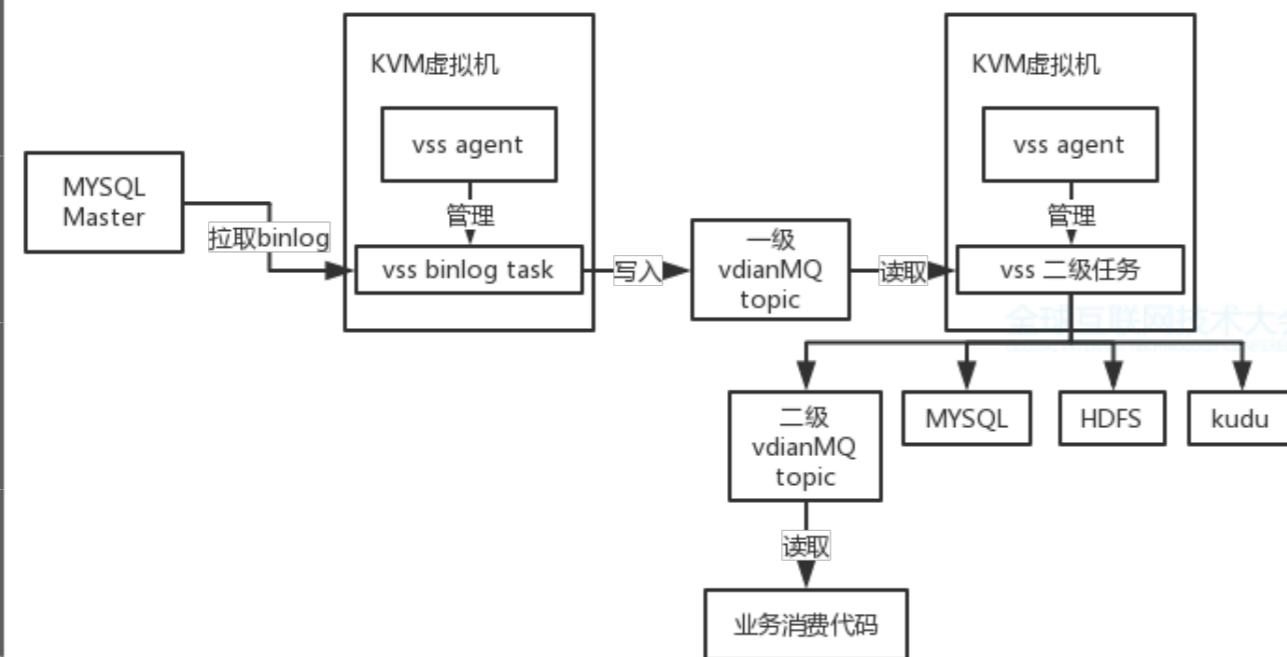


④大数据下的基础建设

A.数据库治理

B.数据层相关中间件

C.分布式事务框架



VSS (增量数据同步)

mysql, 消息, hdfs, kudu的支持

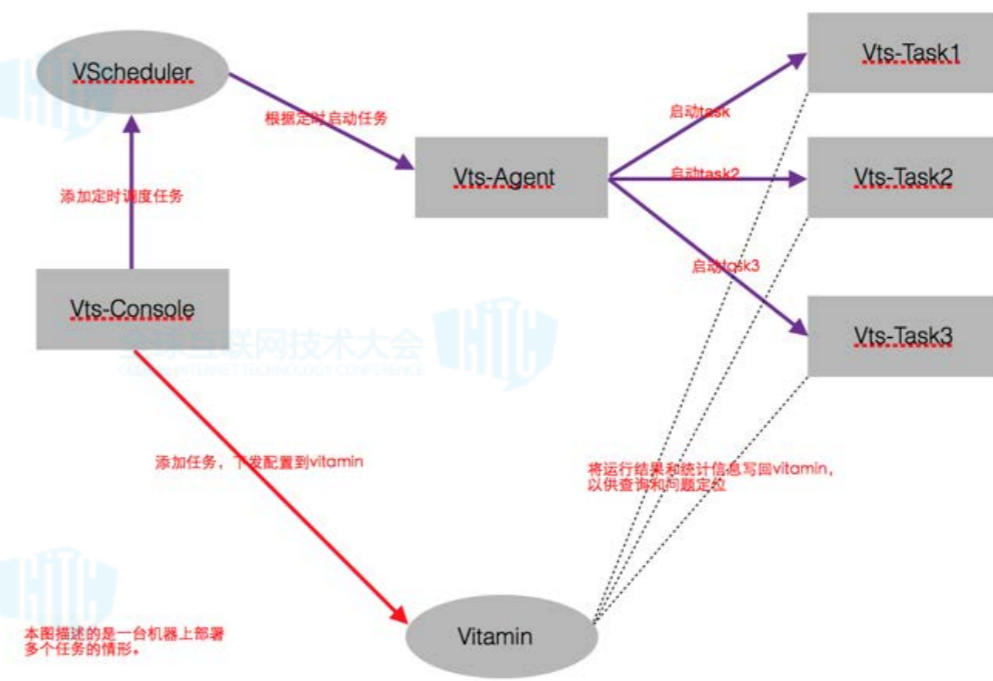
高可用

灵活的过滤规则

动态加载目标写入代码

任务配置自动生成

负载均衡



vts (全量数据同步)

存储, mysql, 消息, hdfs, redis, 本地文件支持

高可用

动态加载过滤规则

数据拉取和写入的速度可控

负载均衡



④大数据下的基础建设

A.数据库治理

B.数据层相关中间件

C.分布式事务框架

	方式	代码侵入性	数据库侵入性
两阶段提交	同步；阻塞协议	弱	弱
Ebay 基于消息	异步	较强	
Alipay XTS	Try同步；confirm/cancel异步	强	主事务分支事务记录
Taobao TXC	同步	弱	Log表业务同库

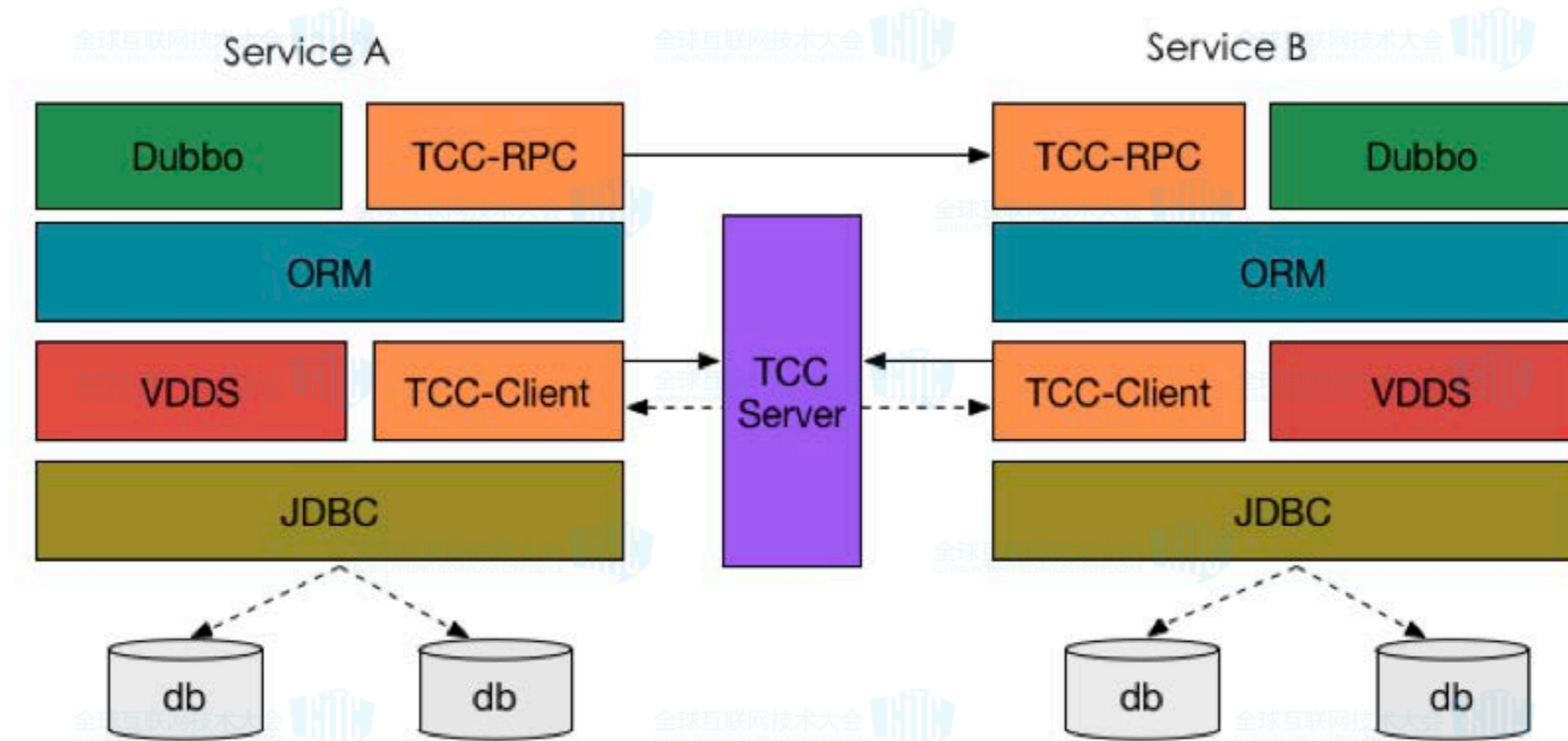


④大数据下的基础建设

A.数据库治理

B.数据层相关中间件

C.分布式事务框架



① 微店是谁，技术挑战是什么

- 业务模式
- 第一代架构
- 新的技术挑战

② 为交易构筑安全可靠的防火墙

- 日均经受560万次攻击
- 日均被爬6亿次

③ 低成本的架构建设之道

- 私有云
- 全站分布式
- 性能优化
- SRE, devops

④ 大数据下的基础建设

- 数据层治理
- 中间件

⑤ 大数据面前，业务系统的演进之路

- 搜索引擎



⑤大数据面前，业务系统的演进之路

A.大数据产品框架

B.搜索引擎的挑战

C.搜索引擎2.0

D.搜索引擎2.0

E.搜索引擎3.0



●搜索的业务场景

- 商品搜索
- 店铺搜索
- 实时搜索
- 10亿商品
- 7000万店铺

技术挑战

数据行数大
实时性、一致性要求高
TPS/QPS在1-2k左右

业务挑战

排序逻辑复杂



⑤大数据面前，业务系统的演进之路

A.大数据产品框架

B.搜索引擎的挑战

C.搜索引擎2.0

D.搜索引擎2.0

E.搜索引擎3.0



1.0架构

- 引擎基于solr，主从复制框架，一主，N个从
- mapreduce join 数据，批量直接更新主 solr 索引中。分钟级增量
- 排序模型简单使用少量几个销量数据加权计算。solr 函数排序



⑤大数据面前，业务系统的演进之路

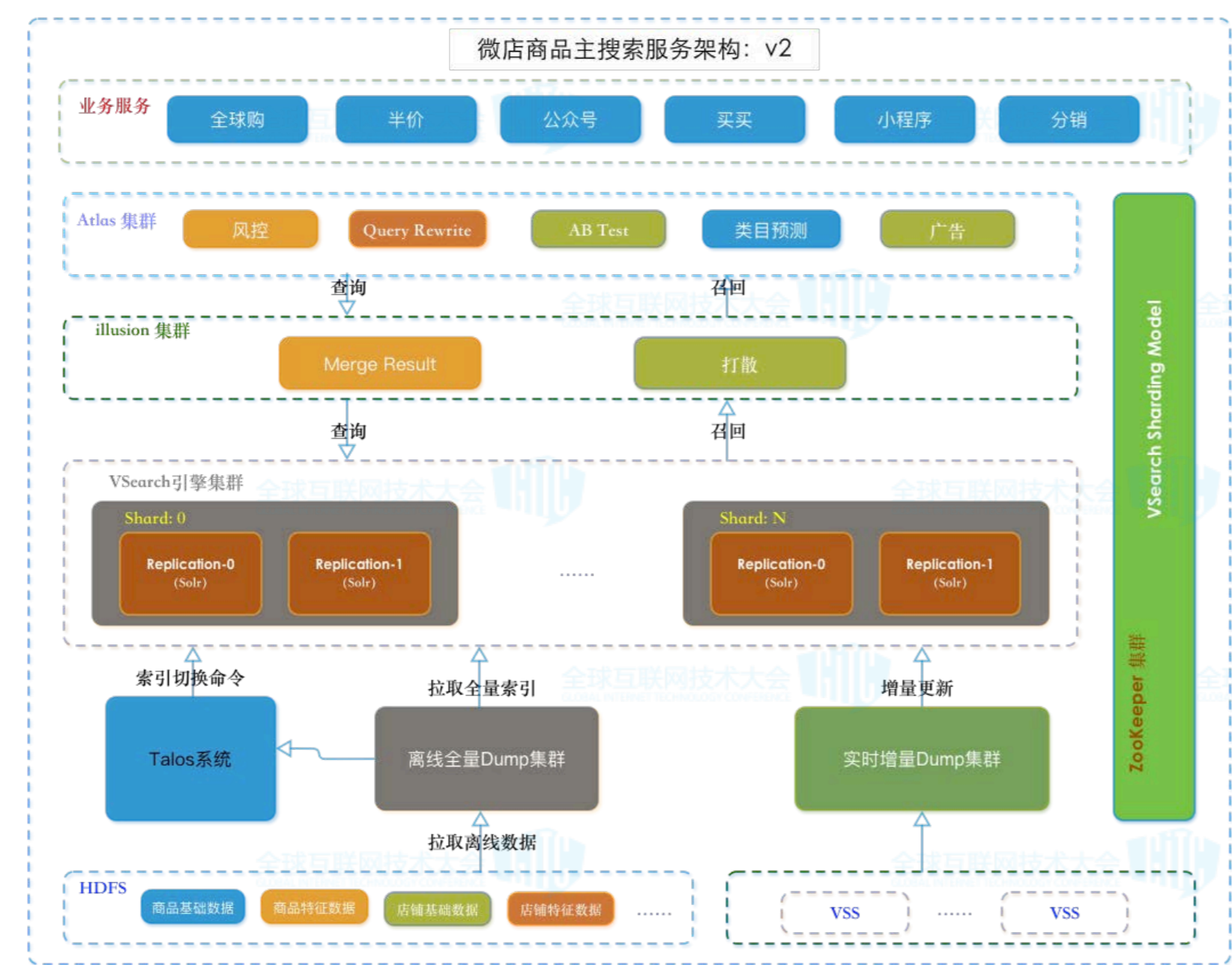
A.大数据产品框架

B.搜索引擎的挑战

C.搜索引擎2.0

D.搜索引擎2.0

E.搜索引擎3.0



- 分布式架构
- 引入merge层（叫 illusion），采用海选+精排架构
- 引入统一接又层，实现QP
- 构建独立的dump集群，全量+增量
- 粗糙的后台
- 算法：海选+精排，LTR，特征丰富



⑤大数据面前，业务系统的演进之路

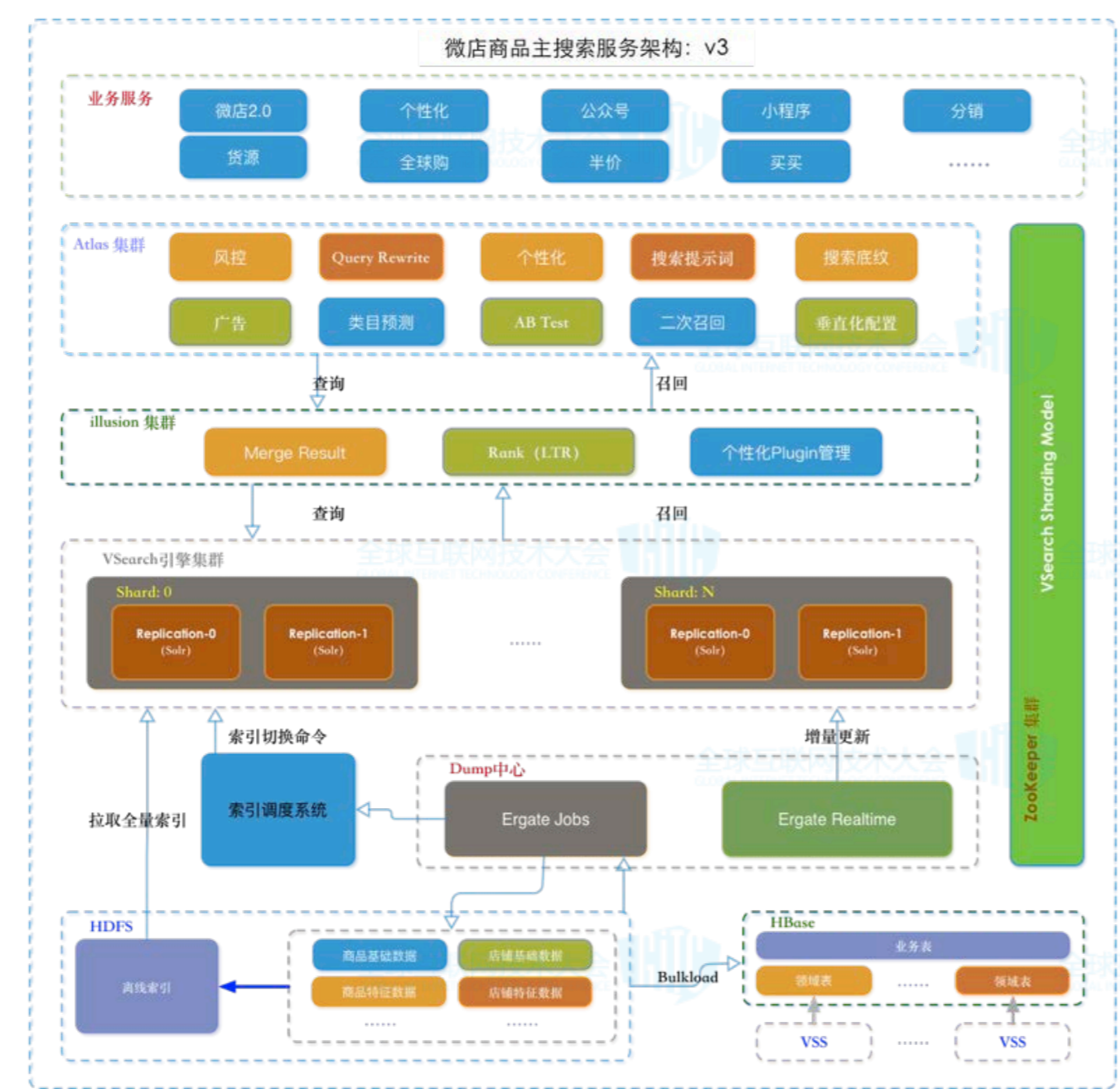
A.大数据产品框架

B.搜索引擎的挑战

C.搜索引擎2.0

D.搜索引擎2.0

E.搜索引擎3.0



- 引擎支持个性化
- 实时搜索上线，全链路时效性在500ms之内，引擎内在100ms之内
- 索引管理引入调度器
- dump 升级，hbase 成为 dump 的基础
- 算法在线学习，模型更新更加实时





谢谢大家!

