The 8th China Cloud Computing Conference
第八届中国云计算大会
技术融合 应用创新

云WAF与大数据实时分析实践

携程信息安全部—张亮

关于我

个人

- 张亮
- 携程信息安全部

方向

- WEB安全、网络安全、安全产品开发

大纲

背景	闭环设计	大数据分析	构建实践	以后的路
• 痛点 • 难点 • 方案	• 规则源 • 部署 • 日志	• 前置分析 • 接口提供 • 后置学习	• 架构 • 特色 • 现状 • 可视化	• 自动化 • 计划

背景—痛点

- 业务安全需求
  - 恶意IP封禁
  - 恶意用户封禁
  - ...
- 应用安全防护需求
  - SQL注入
  - XSS跨站
  - LFI、RCE
  - ..
- 应急响应需求
  - 不能精确拦截
  - 发布更新周期长
  - 规则生效慢

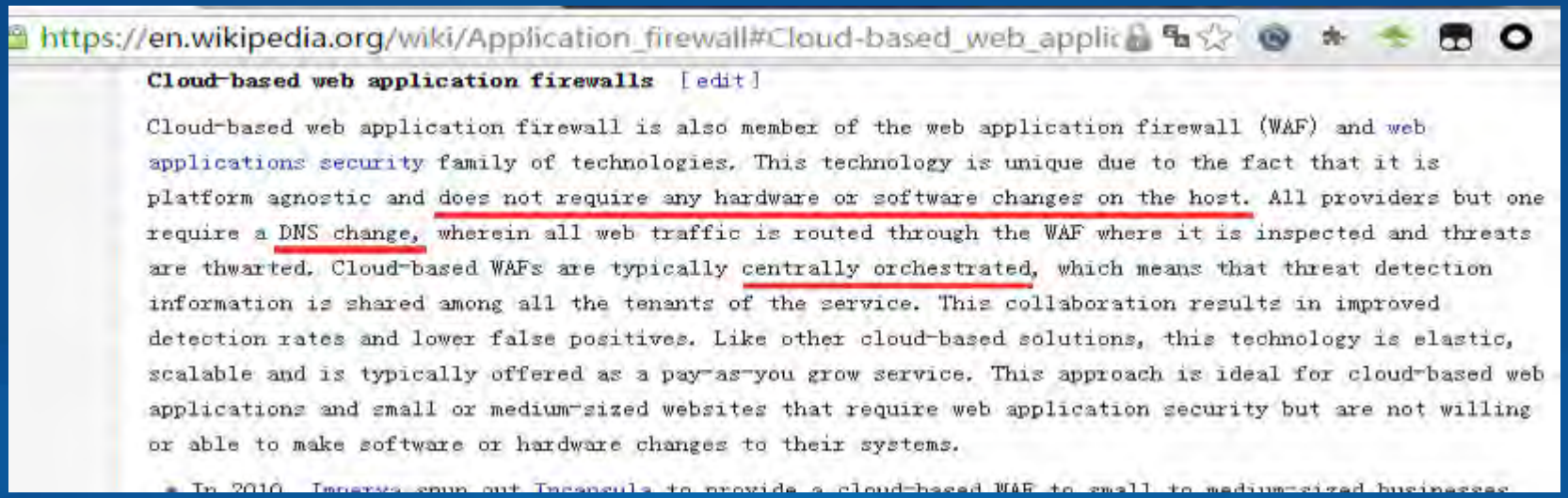
- 硬件WAF
  - 带宽、效率瓶颈
  - 成本高昂
  - 不适合分布式多机房
  - Struts 2 s2-032之类的0day或者1day很多商业WAF都不支持，并且响应更新慢
- 商业云WAF
  - 规则定制难
  - 业务线长难以全部迁移

背景—难点

- 10种左右的系统版本
- 5种左右的web容器
- N种不同的编码语言

背景—方案

- Cloud-based web application firewalls
- 集中式平台，客户端无需任何改变
- 仅需要DNS解析指向
- 检测信息共享


https://en.wikipedia.org/wiki/Application_firewall#Cloud-based_web_applic

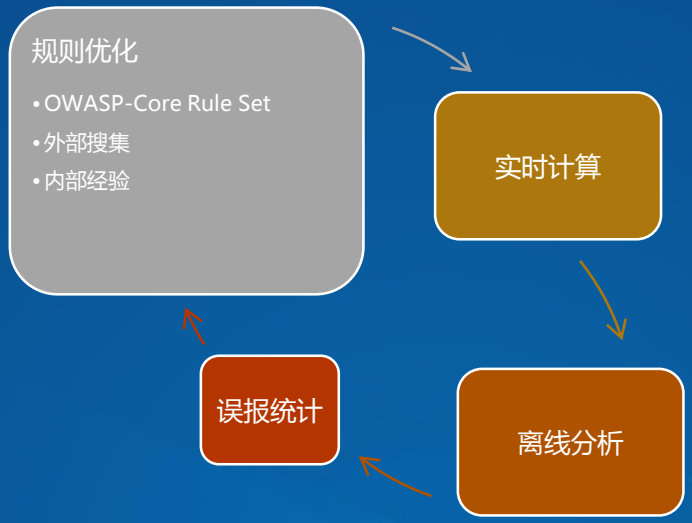
Cloud-based web application firewalls [edit]

Cloud-based web application firewall is also member of the web application firewall (WAF) and web applications security family of technologies. This technology is unique due to the fact that it is platform agnostic and does not require any hardware or software changes on the host. All providers but one require a DNS change, wherein all web traffic is routed through the WAF where it is inspected and threats are thwarted. Cloud-based WAFs are typically centrally orchestrated, which means that threat detection information is shared among all the tenants of the service. This collaboration results in improved detection rates and lower false positives. Like other cloud-based solutions, this technology is elastic, scalable and is typically offered as a pay-as-you grow service. This approach is ideal for cloud-based web applications and small or medium-sized websites that require web application security but are not willing or able to make software or hardware changes to their systems.

In 2010, Imperva spun out Incapsula to provide a cloud-based WAF to small to medium-sized businesses

闭环设计—规则源

- ✓外部搜集，内部整理
- ✓通过日志、流量运用规则做实时计算
- ✓将实时计算的结果进行分析
- ✓将误报率低、漏报率低的规则策略发至线上



大数据分析—前置分析

- 基于STORM对流量进行实时计算
- 更宽泛的规则策略
- 实时分数计算
- 可自动或手动与WAF联动

时间	攻击源	攻击目标	全部访问的域名	归属地	分数	攻击类型	攻击次数/访问总次数	深度分析	操作
开始时间：2016-04-22 07:43:01 结束时间：2016-04-22 07:48:01	[redacted]	[redacted](2028)	[redacted](2309)	河南 郑州	6451	SQLI-2008(454) SQLI-2005(394) SQLI-2009(182) XSS-3001(116) RFI-1004(109)	2309/2028	攻击详情 访问详情	确认 误报 拉取IP 加入黑名单库 加入白名单库
开始时间：2016-04-22 07:58:01 结束时间：2016-04-22 08:03:01	[redacted]	[redacted](140)	[redacted](101)	北京 北京	615	RFI-1004(40) SQLI-2005(23) SQLI-2010(22) SQLI-2008(21) SQLI-2007(20)	101/140	攻击详情 访问详情	确认 误报 拉取IP 加入黑名单库 加入白名单库
开始时间：2016-04-22 07:58:01 结束时间：2016-04-22 08:03:01	[redacted]	[redacted](10)	[redacted](2)	河南 郑州	39	SQLI-2008(2) SQLI-2009(2) SQLI-2010(2) RCE-1001(1) SQLI-2003(1)	2/10	攻击详情 访问详情	确认 误报 拉取IP 加入黑名单库 加入白名单库
开始时间：2016-04-22 07:53:01	[redacted]	[redacted](298)	[redacted](366)	北京 北京	1155	SQLI-2008(46) SQLI-2005(39)	366/298	攻击详情 访问详情	确认 误报

闭环设计—部署

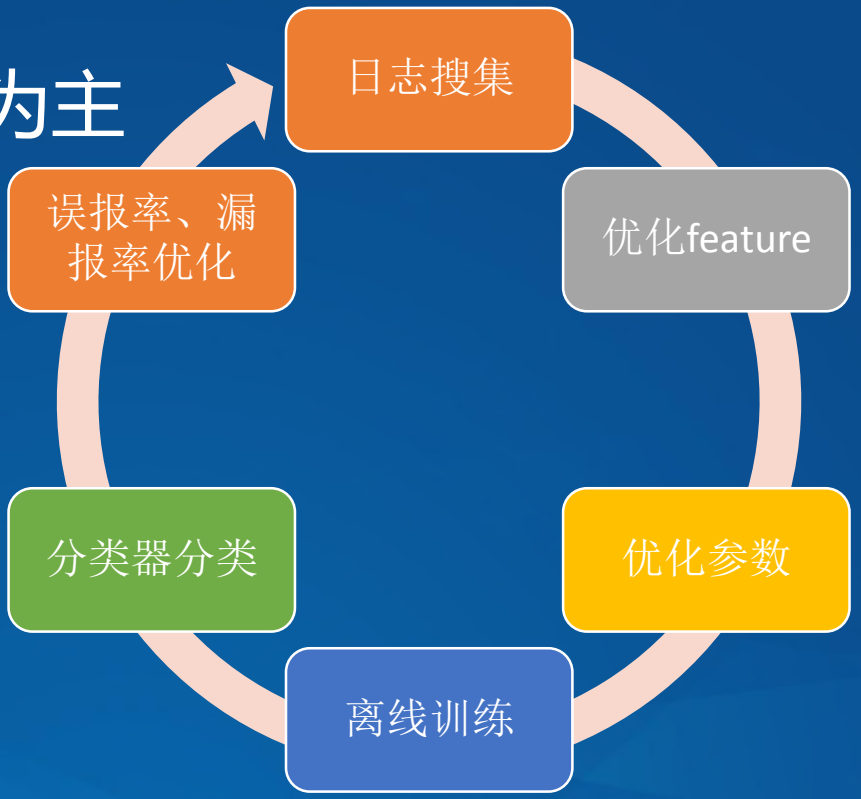
- 结合nginx，一行配置，无缝开启waf
- 结合Load Balance，低成本部署产生高性能
- REST API用于管理

大数据分析—接口提供

- ✓RESTful API接口
- ✓支持动态规则，拦截基于IP、UA、UID等策略组合
- ✓支持静态规则，拦截基于参数、内容的web攻击拦截
- ✓支持一键开启关闭WAF、一键切换拦截、检测模式、一键bypass
- ✓支持状态查询、规则下发、规则更新等功能

闭环设计—日志处理

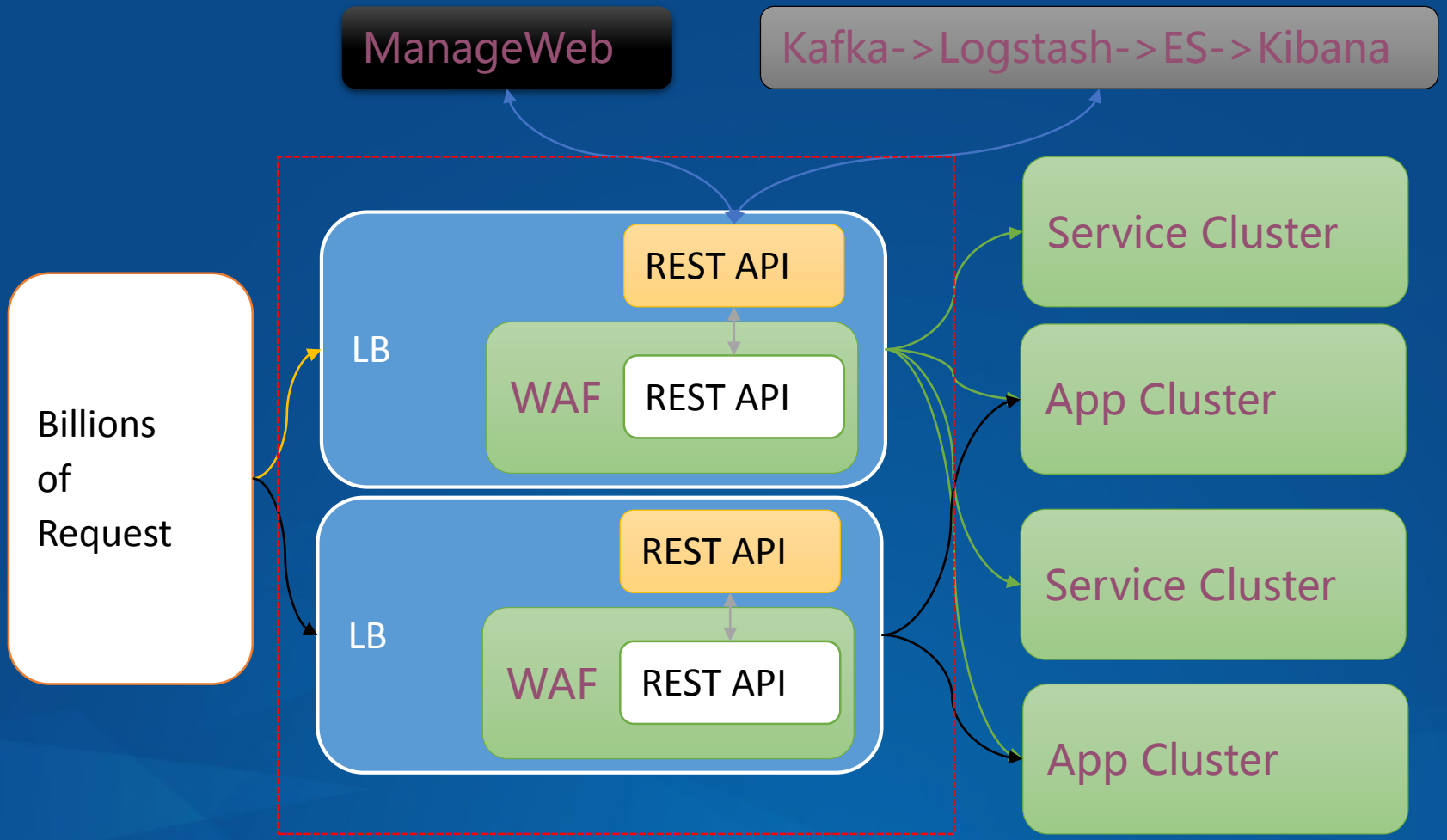
- Supervised Learning
- 误报率、漏报率
- 机器学习为辅、人工为主



大数据分析—后置学习

- 基于SPARK stream、SPARK mllib
- 针对WAF日志进行解析、分析、统计、学习
- 自动分析每日上百万日志中误报条目
- 人工确认结果作为训练数据输入

构建实践—架构

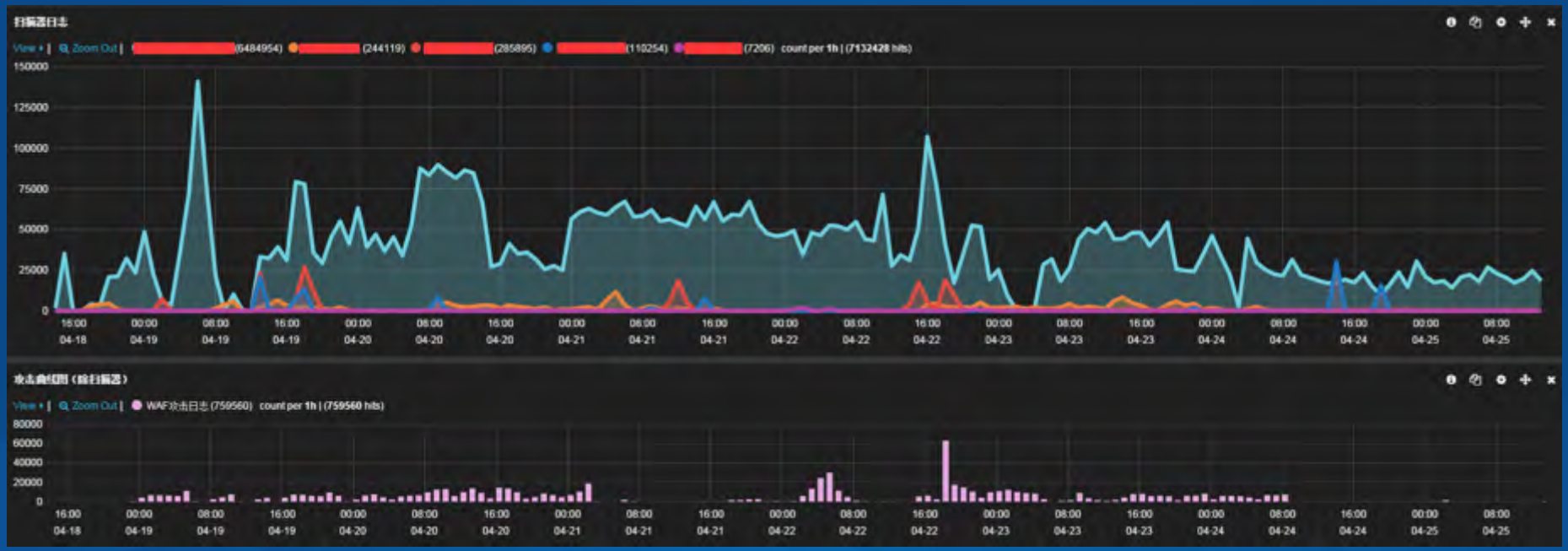

ManageWeb
Kafka->Logstash->ES->Kibana
Billions of Request
LB
REST API
WAF
REST API
LB
REST API
WAF
REST API
Service Cluster
App Cluster
Service Cluster
App Cluster

构建实践—特色

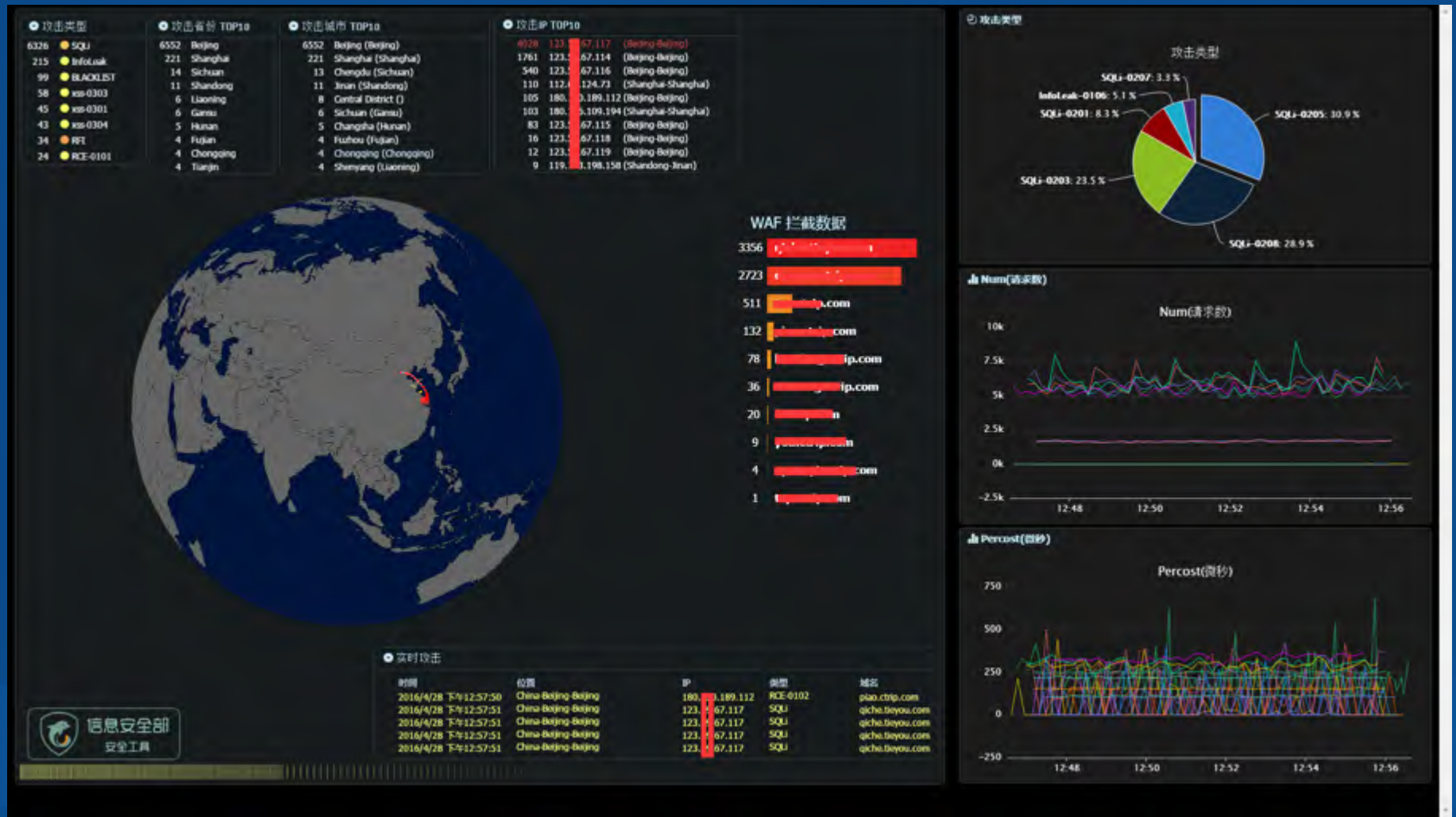
- 结合LB，对HTTPS友好
- 快速部署、规则策略秒级生效
- 引擎一键进入拦截、检测、关闭模式
- 预留REST API接口，配合风控、反爬
- 规则来源于Storm实时计算优化后

构建实践—现状

- 每天处理十亿级请求，每次处理耗时0.2ms左右
- 每天拦截百万次攻击，误报率低于千万分之一
- 保护成千上万的应用



构建实践—可视化


攻击类型
攻击省份 TOP10
攻击城市 TOP10
攻击IP TOP10
WAF 拦截数据
实时攻击
信息安全部
安全工具
攻击类型
SQLi-0207: 3.3 %
InfoLeak-0106: 5.1 %
SQLi-0201: 8.3 %
SQLi-0205: 30.9 %
SQLi-0203: 23.5 %
SQLi-0208: 28.9 %
Num(请求数)
Percost(微秒)

以后的路—自动化

- 无人值守？
- 机器学习判定误报、漏报？
- 自动加白？
- 自动识别可疑访问？

以后的路—计划

- High performance
- More feature
- More api

携程云安全-security.ctrip.com

✓免费

✓200+企业用户

The 8th China Cloud Computing Conference

Thank you