

www.qconferences.com
www.qconbeijing.com



伦敦 | 北京 | 东京 | 纽约 | 圣保罗 | 上海 | 旧金山
London • Beijing • Tokyo • New York • Sao Paulo • Shanghai • San Francisco

QCon全球软件开发大会

International Software Development Conference

这一次，
我们重新定义了WAF

吴翰清
@安全宝
2013-4-25

我是谁

- 2005 - Alibaba
- 2012 - 安全宝
- 《白帽子讲Web安全》
- 微信：道哥的黑板报，
ID: taosay

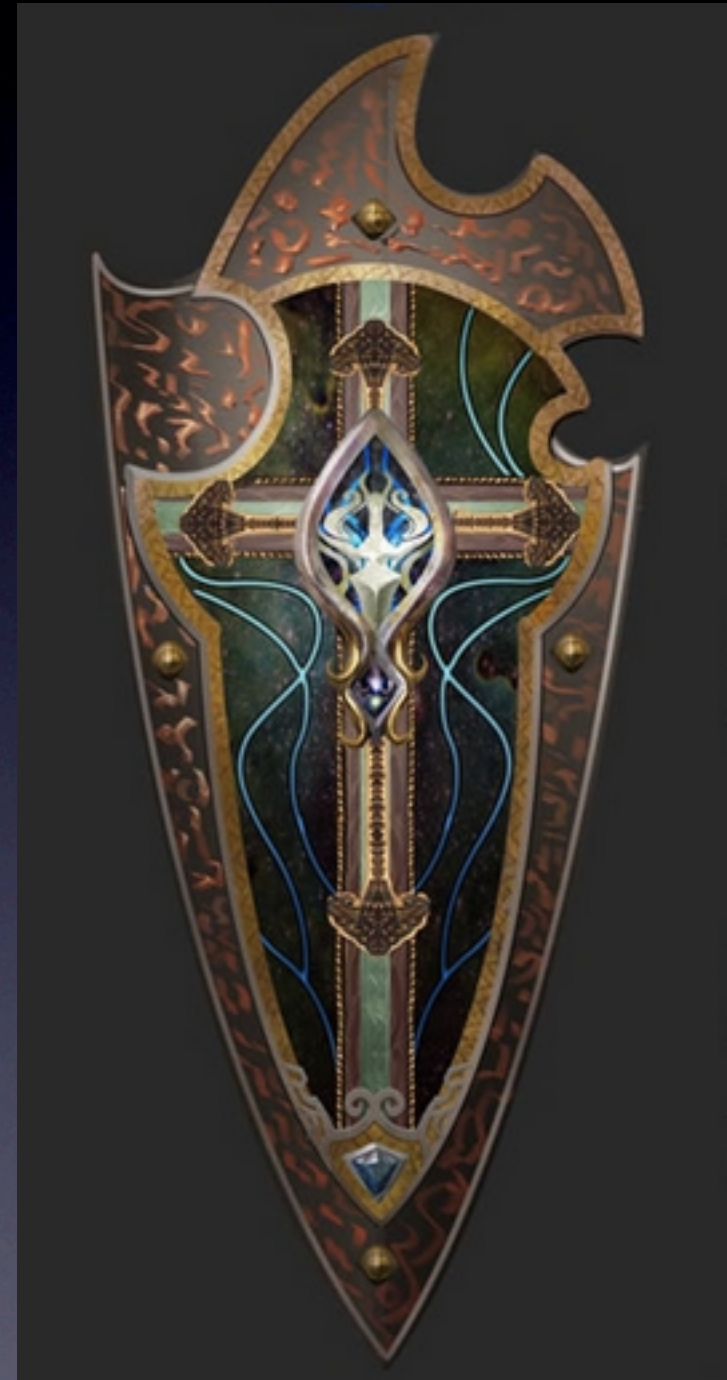


什么是WAF

- Web Application Firewall
- Web应用的中间层，通过定制规则识别XSS、SQL注入等恶意攻击，并拦截

为什么需要WAF

- 拦截Web攻击 - 第一时间
- 拦截CC攻击 - Web Server层
- 虚拟补丁，应用安全集大成者



我对WAF的态度

- 2008年时撰文反思WAF的缺陷
- 2011年以后发现WAF还是有用的
- 现在，投身于此

PCI DSS对WAF的要求

- 2008年6月30号以后，在标准中明确要求安装WAF
- 此前只是一项最佳实践



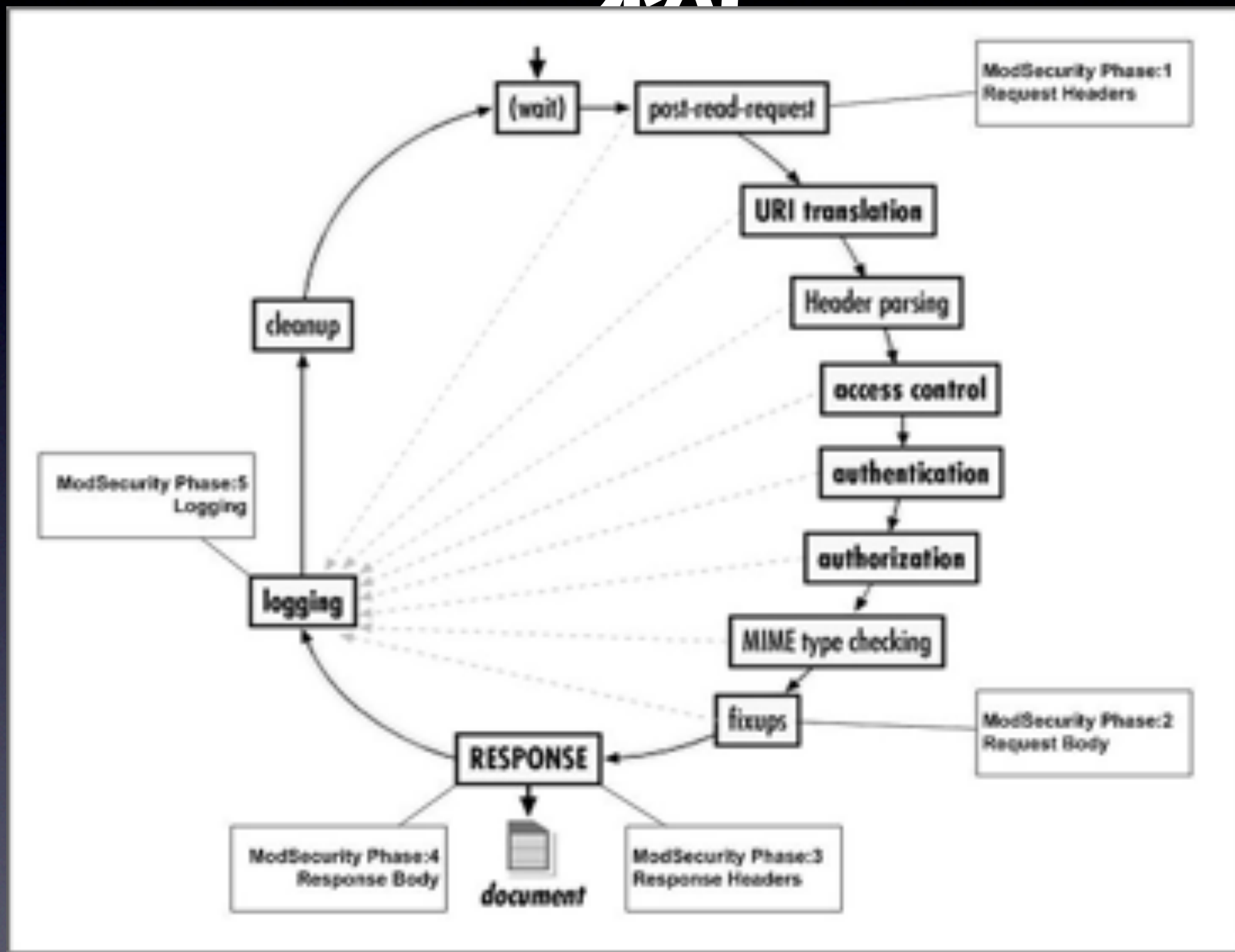
WAF的产值

- 规模在几十亿美元
- 受PCI DSS影响，每年增长18.7%
- 个人观点：远远被低估

WAF的实现方式

- 开源软件实现 (mod-security, phpids)
- 硬件设备实现 (trustwave)
- 云端实现 (Cloudflare, 安全宝)

Mod Security的架构



硬件WAF

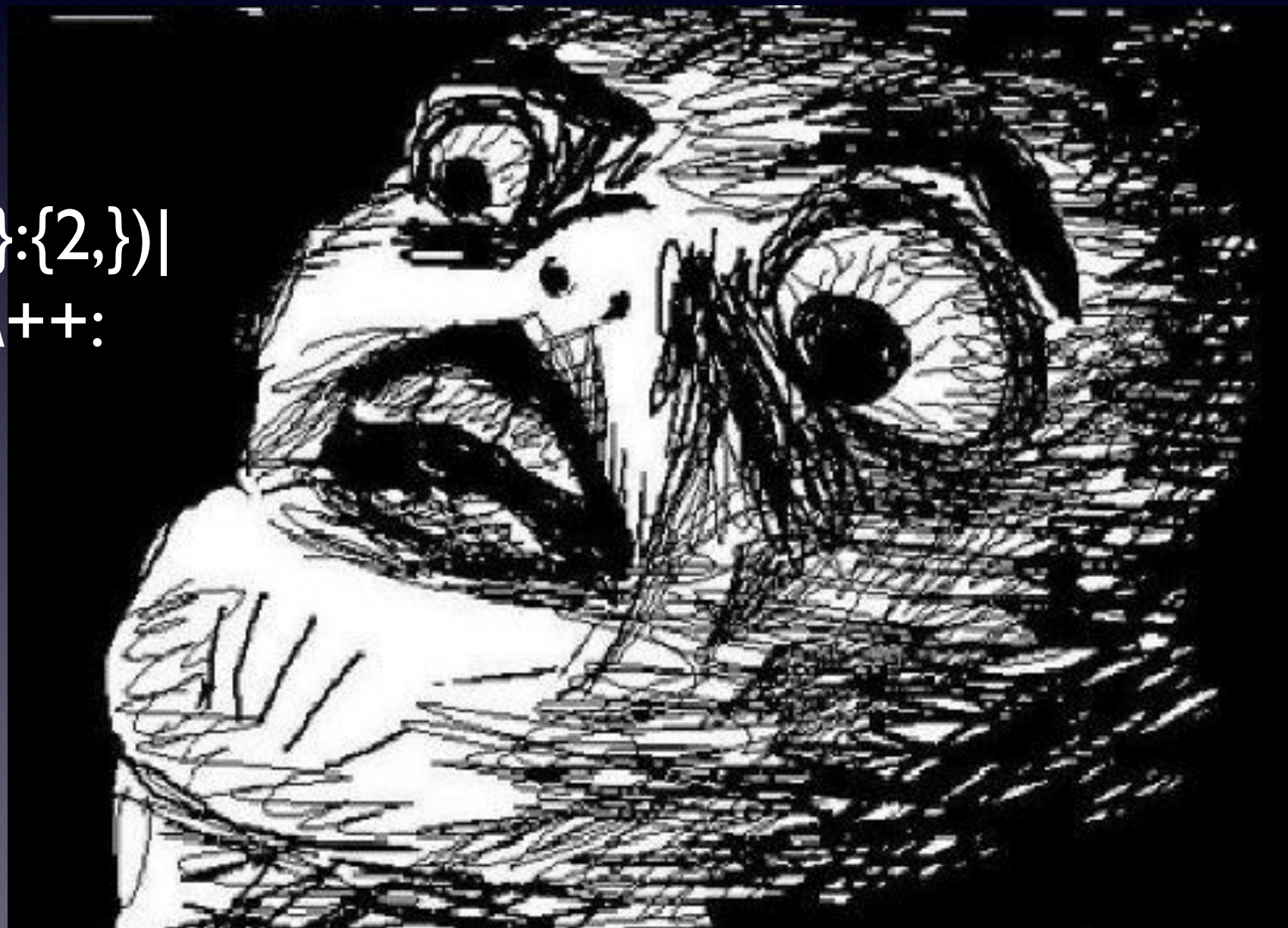
- 面向大客户
- 部署慢 - 以天为单位
- 计算能力有限 - 受制于单机性能
- 需要专家上门调试规则

WAF是给安全专家用的？

What the fuck?

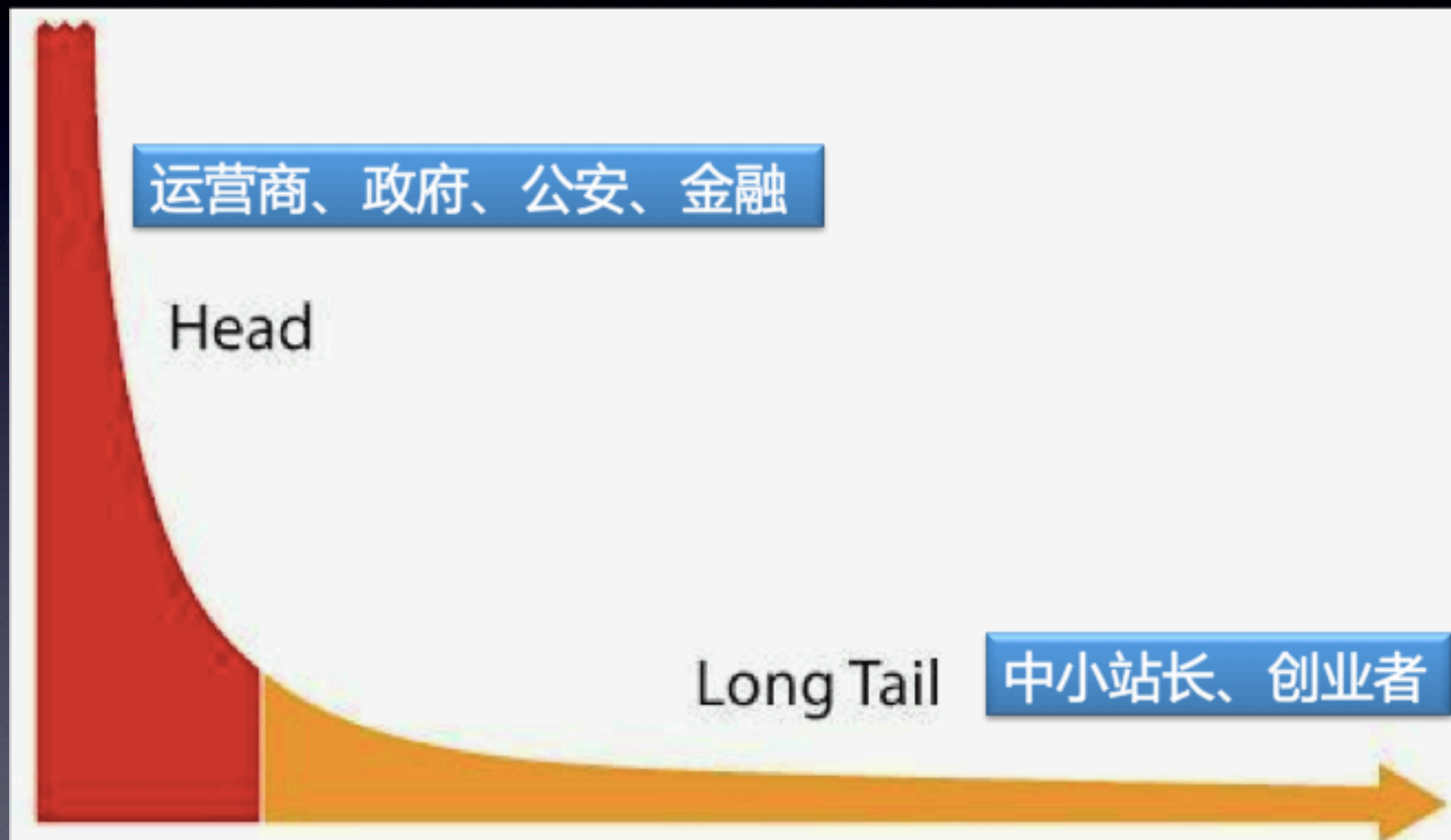
- `<![CDATA[REQUEST_URI^[r](*[\\s\\r\\n\\v\\f]|([^\\-&])%23[\\s\\r\\n\\v\\f]|\\-\\-[\\s\\r\\n\\v\\f]|([^\\-&])#[\\s\\r\\n\\v\\f]))]>`

`<![CDATA[(?:\\({2,}\\+{2,}:{2,})|
(?:\\({2,}\\+{2,}:+)|(?:\\({3,}\\++:
{2,})|(?:\\$\\[!!!\\))]]>`



为什么需要云WAF?

令人难以置信



云WAF闪亮登场

- DNS托管接入 - 10分钟完成部署
- 云端计算能力可以线性扩展
- 云端优化规则，用户勿须关心
- 及时的事件响应（误报、漏报、稳定性）

一不小心发现玩大了!

- 加速(70%的流量不需要访问源站, 60%的网站感觉明显变快)
- 抗D (110G抗D中心)
- Web安全



taosay.net的加速效果

- linode 日本的vps，速度提升2-4倍

[首页](#) [Get Ping](#) [TraceRoute](#) [Dns](#) [Cdn](#)

http://taosay.net/webscan_360_cn.html

http://www.taosay.net/webscan_360_cn.html

检测一下

[取消对比](#) [高级](#) [帮助](#)

<http://www.17ce.com/site/ping/f07a042ff24f3994428f75a97> [分享](#) [发给好友](#) [复制地址](#)

检测目标: [taosay.net](#) [www.taosay.net](#)

检测时间: 2013-03-22 11:47:44



云WAF的挑战

风险集中化

CloudFlare 宕机导致 78 万网站下线，服务中断超过 1 小时

xinzhi 发表于 2013/03/04-00:04 [DNS错误](#) / [云计算](#) / [宕机](#) / [CloudFlare](#)



四川雅安芦山县7.0级地震 相关专题：[百度](#)[新浪](#)[360](#)[腾讯](#)[Google](#)寻人



分享到QQ



分享

赵武360 ，知道创宇  等216人分享过

→ 分享到

几个小时前，CloudFlare 由于 DNS 路由配置错误导致使用其 CDN 和安全服务的 785000 多个网站遭受影响，其中不乏 4chan、Wikileaks, Metallica 等大型网站，故障在 30 分钟左右被排除，影响持续了一个多小时。



CLOUDFLARE®

DDOS引起的不稳定

- 一个站点被攻击，可能引起整个节点的不稳定
- 对用户类型进行划分

备案问题

- 接入后IP变化，导致网站备案被取消



SEO问题

- 百度降权
- 搜索引擎IP被拦截



用户源站安全拦截

- IDC防火墙拦截
- “安全狗” 拦截
- D盾拦截

源站不能访问

- 自动探测源站信息
- 更人性化的出错页面



您访问的页面正在维护，我们正在尽全力恢复访问，请您稍候...

[站长点击查看详情](#)

ServerName: /

误报 漏报

- 评价一个WAF好坏的关键指标
- 把关键指标做到极致，就是创新



由于您访问的URL有可能对网站造成安全威胁，您的访问被阻断。

[误报反馈](#)

[站长点击查看详情](#)

从一个**idea**到成功的**产品**，
有漫长的路要走

超越传统WAF

为什么以前的WAF很傻？

- 针对人，
- 而非针对单次请求



基于行为



恶意爬虫

- Alibaba B2B数据曾被恶意爬取

规则系统的设计

- 安全专家最精通的是正则表达式
- 出于工作效率考虑，使用正则写规则是最佳方案
- 高性能的正则表达式规则解决方案

性能测试

- 正则表达式的DOS
- 故障：3台服务器CPU跑满
- $(.[\backslash V].+)^*$

海量数据的回归测试

- “每条”规则发布前都需要回放数亿的真实请求数据
- 灰度发布，确保无误后再开启

自动识别误报

- 每天几十万到上百万的请求被拦截
- 人肉?
- 抽样?



从客服统计误报？

- 可能只有不到1/100的网站站长会上报客服
- 有时候站长自己都不知道

我们分析总量！

- 大数据的思路是分析总量，而非抽样！



CMS后台

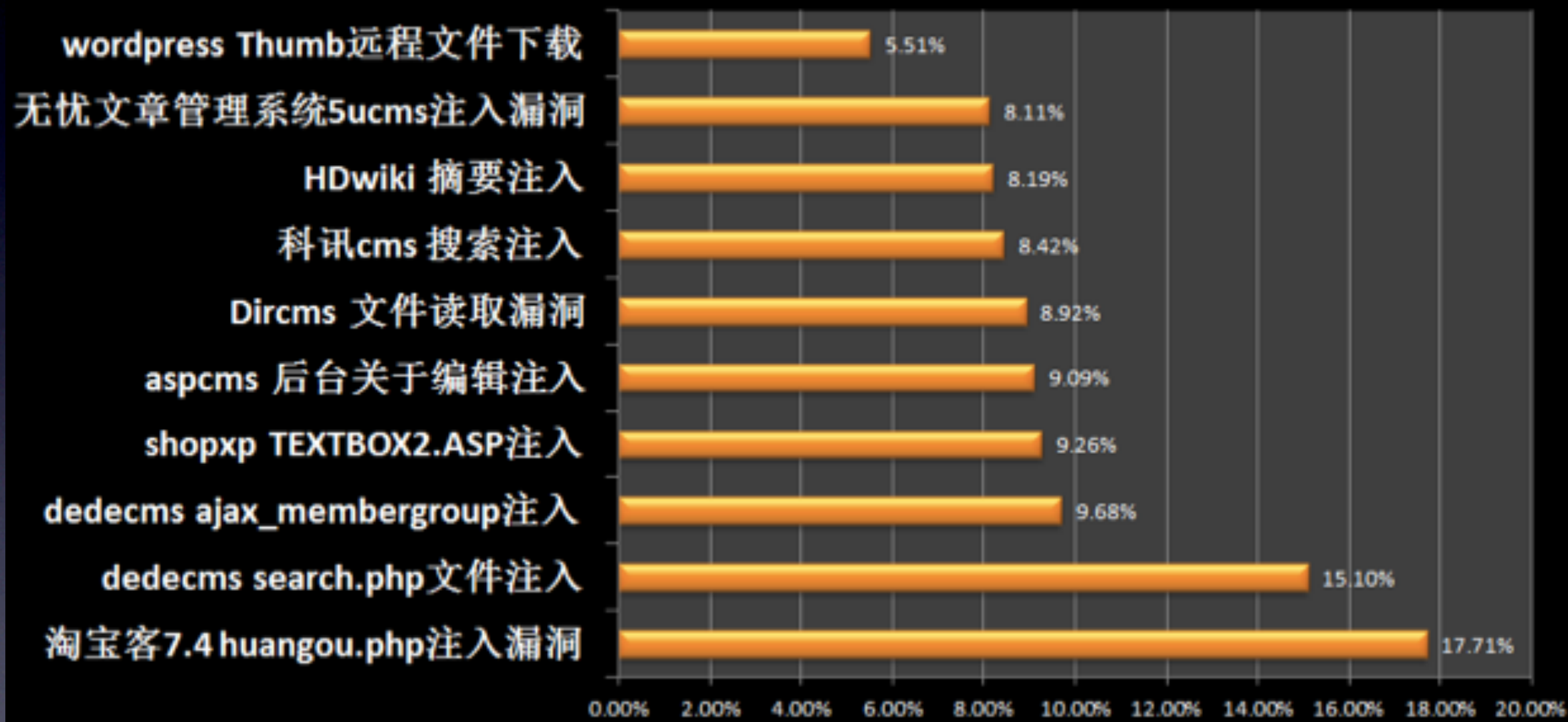
- 误报的最大来源
- 如何区分富文本?
- 区分POST类型

如何统计漏报？

- 灰度规则，观察模式
- 扫描器自动化运营
- 来自第三方报告

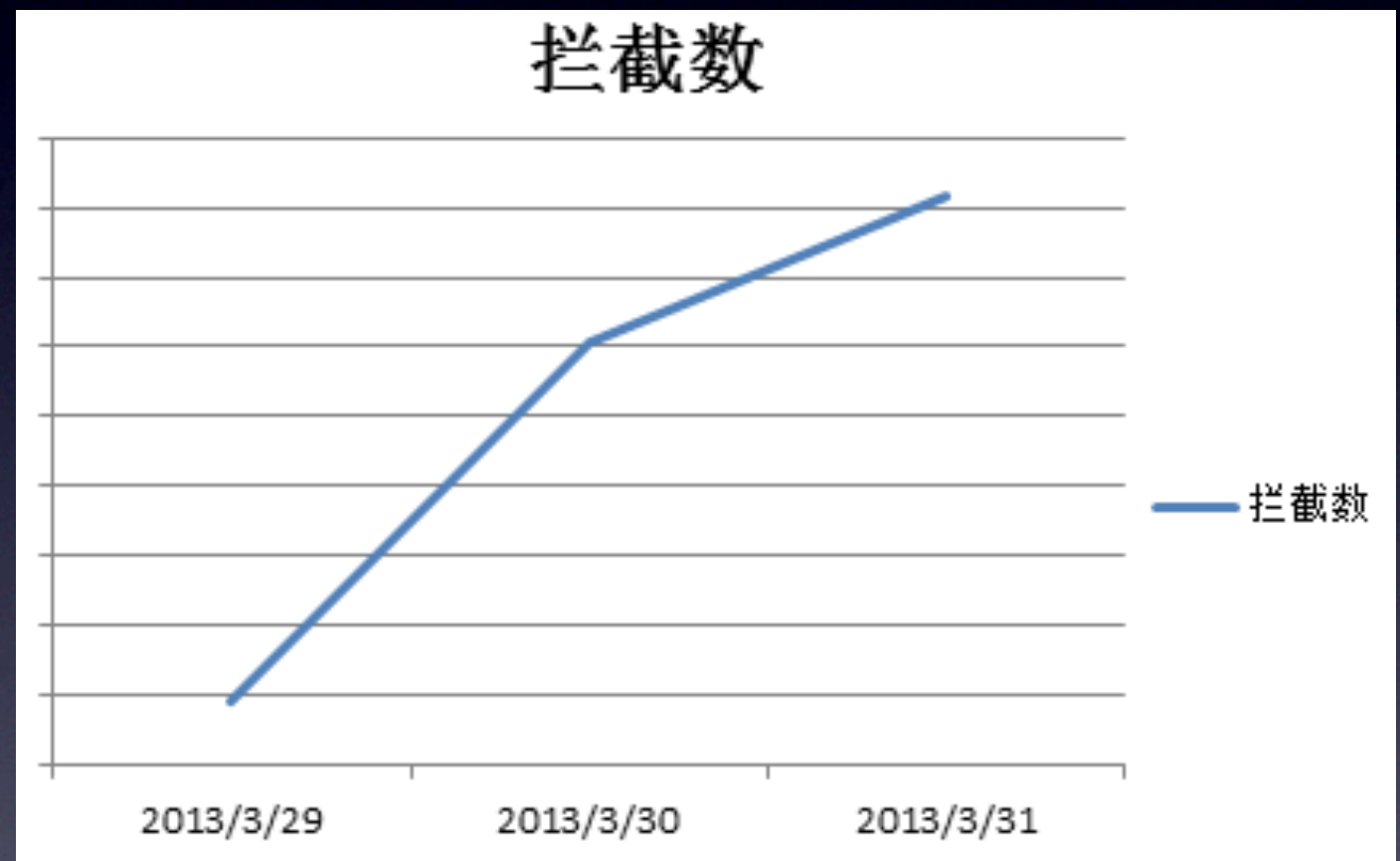
金矿： 0day捕获

2012年Web攻击统计

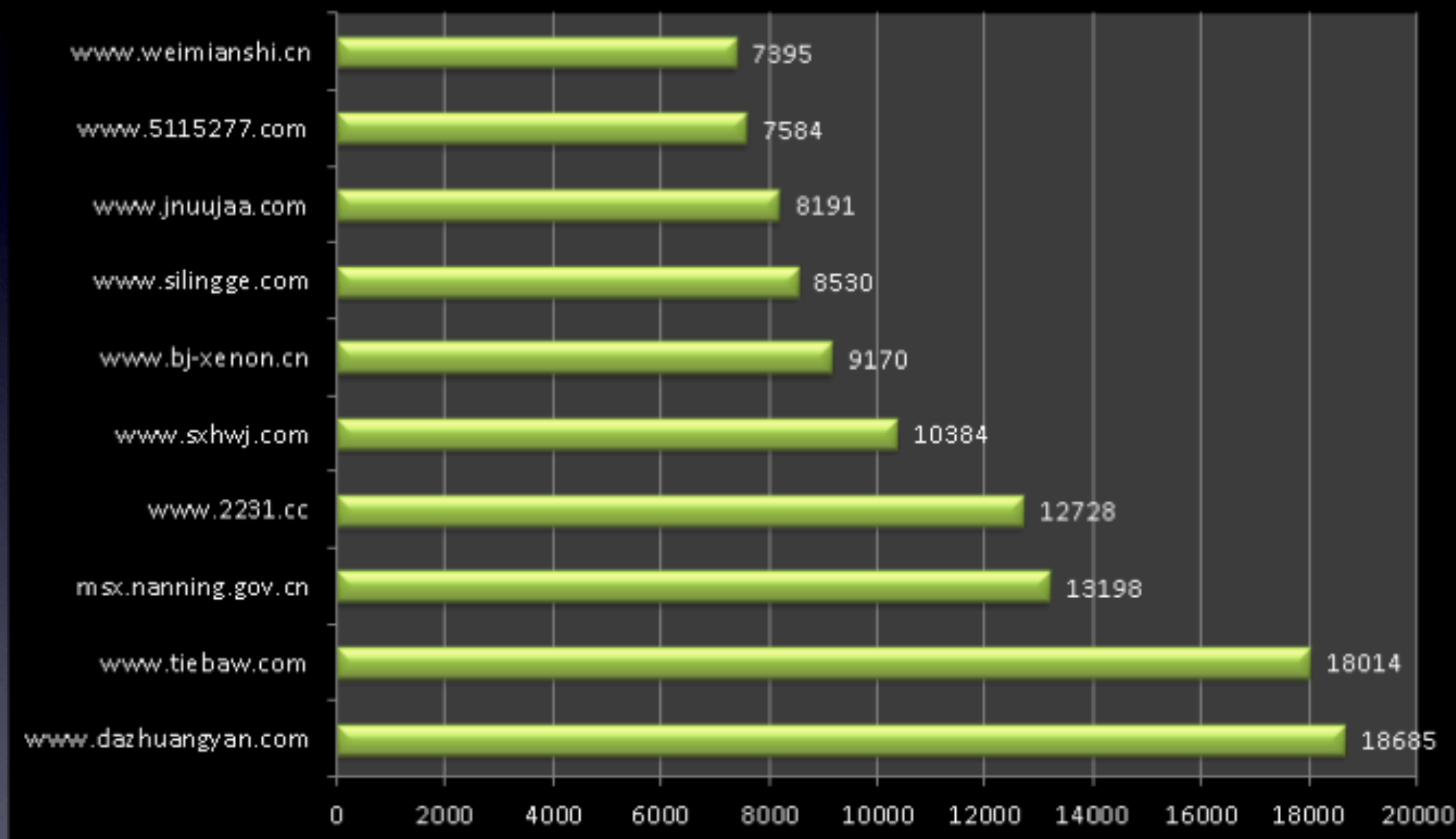


Dedecms 0day

- 0day在曝光前地下流传时间
- 漏洞细节公布后, 拦截数上升了**10倍**



基于网站的攻击预警



定制规则

- 白规则的修复思想

基于大数据计算引擎

- 流式计算引擎
- 离线数据分析引擎

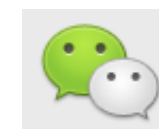
Thanks!

www.infoq.com/cn

InfoQ^{ueue}



@InfoQ



infoqchina

软件
正在改变世界!