智能检测与响应系统实践

兜哥@百度

百度安全
Baidu Security

个人简介

兜哥

百度WEB防护产品负责人

Freebuf专栏作家

《Web安全之机器学习入门》作者

WEB安全对抗

僵尸网络

机器学习

沙箱

黑产对抗

DDoS

WAF

威胁感知

百度安全 Baidu Security

团队简介：为百度系以及企业客户提供BAT级防护

产品层面	提供给百度系企业安全团队使用	百度云安全产品	2中大B企业级产品 安全宝	2中小B产品 云加速
能力层面	抗D	WEB防护	入侵检测	攻击溯源
技术层面	威胁情报	机器学习	威胁建模	沙箱

百度安全
Baidu Security

我们的挑战：业务复杂

作为中国最大的互联网公司之一，百度是体量庞大、业务线最为复杂的互联网公司。

✓ 安全运维的复杂度相当大。

业务体量庞大

- 服务器数是传统量级的200+倍
- 业务规模是传统行业量级的百倍
- 百度日搜索量50亿次+
- 全球TOP100的APP中，百度系占6个，百度系用户数上亿的APP 14个

✓ 业务线纷繁复杂

- 业务线复杂，覆盖搜索、无人车、金融、地图、APP分发、O2O、社区等众多业务

百度安全
Baidu Security

我们的挑战：树大招风

作为全球最大的中文搜索引擎，百度一直是黑产紧盯的目标，因此安全形势也非常严峻。

1）百度每年的安全漏洞在千级别，高危漏洞占20%；

2）每天受到上亿次黑客攻击；

3）平均每个产品线有4-5个黑产组织紧盯；

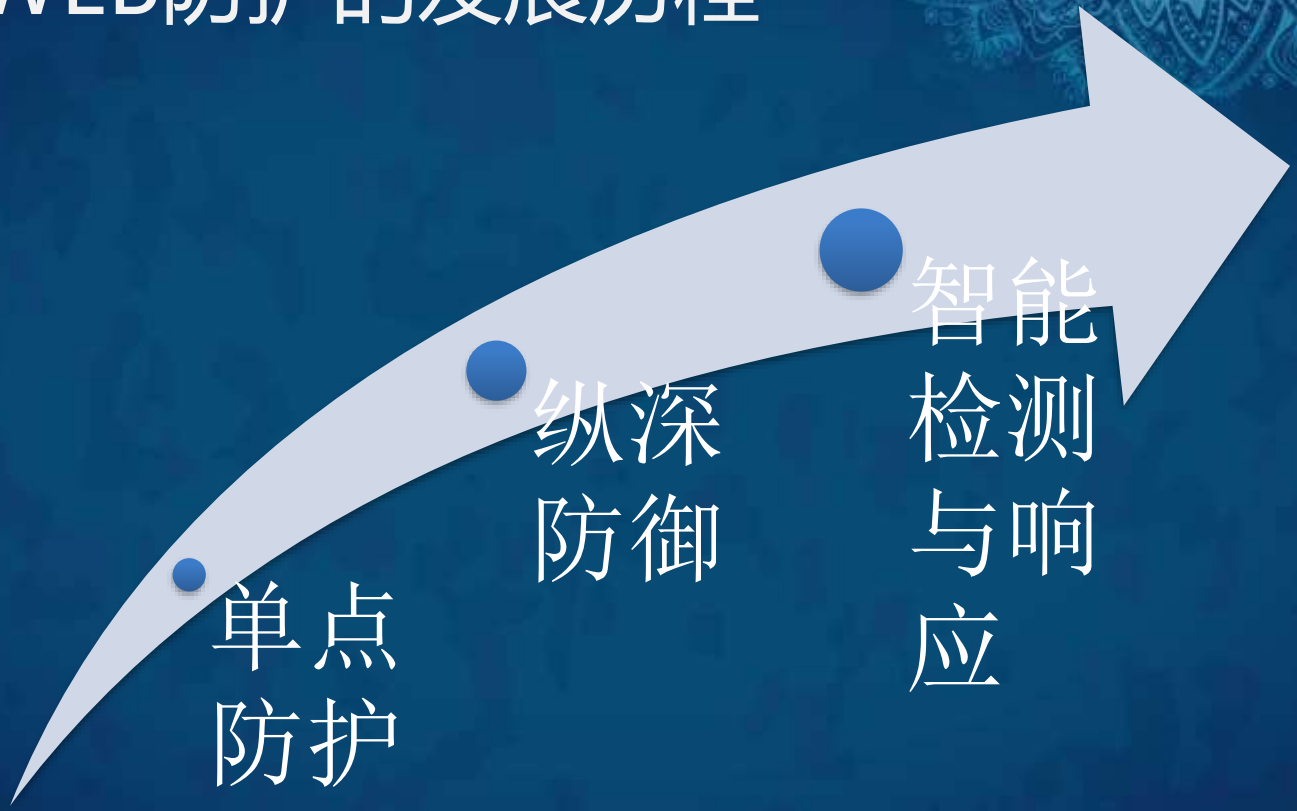
4）有组织的APT攻击，平均每周拦截钓鱼邮件约1.76万封；

百度安全
Baidu Security

百度的安全建设之路：不得不做

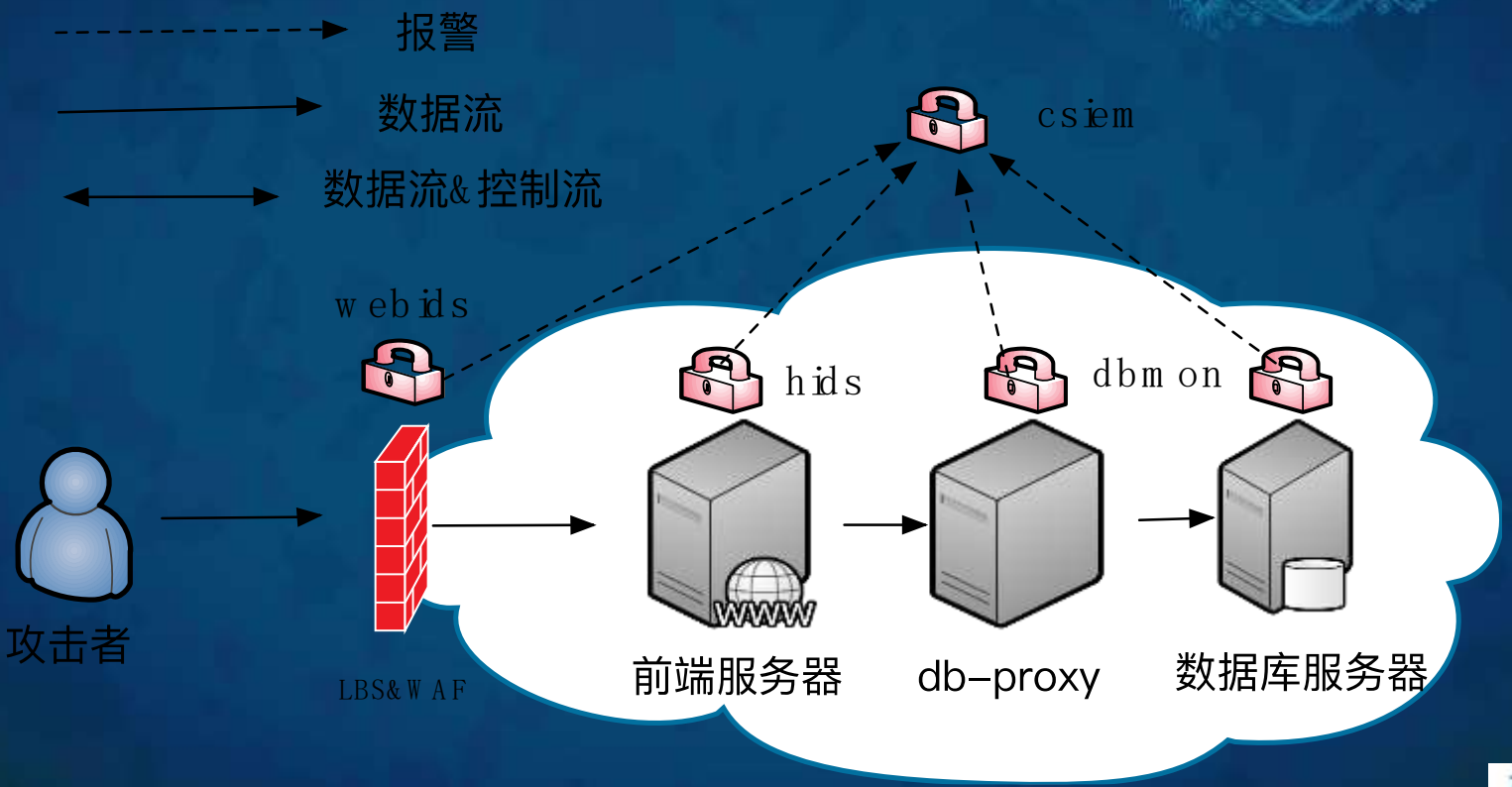
大力投入安全建设，不仅是合规驱动，更是保护自己
大力投入安全研究，不是无痛呻吟，是面对的黑产几乎具备代差的攻击能力

百度WEB防护的发展历程

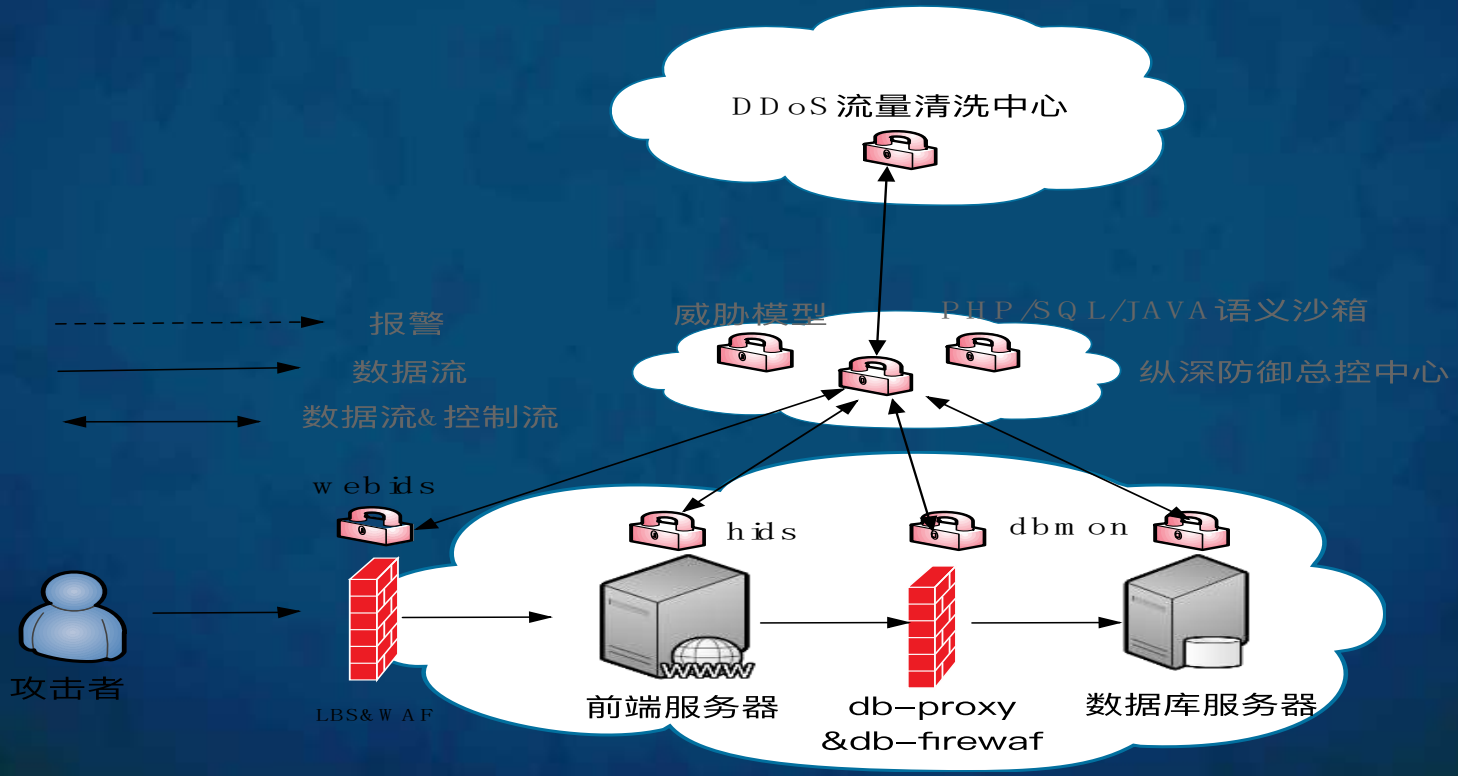


百度安全 Baidu Security

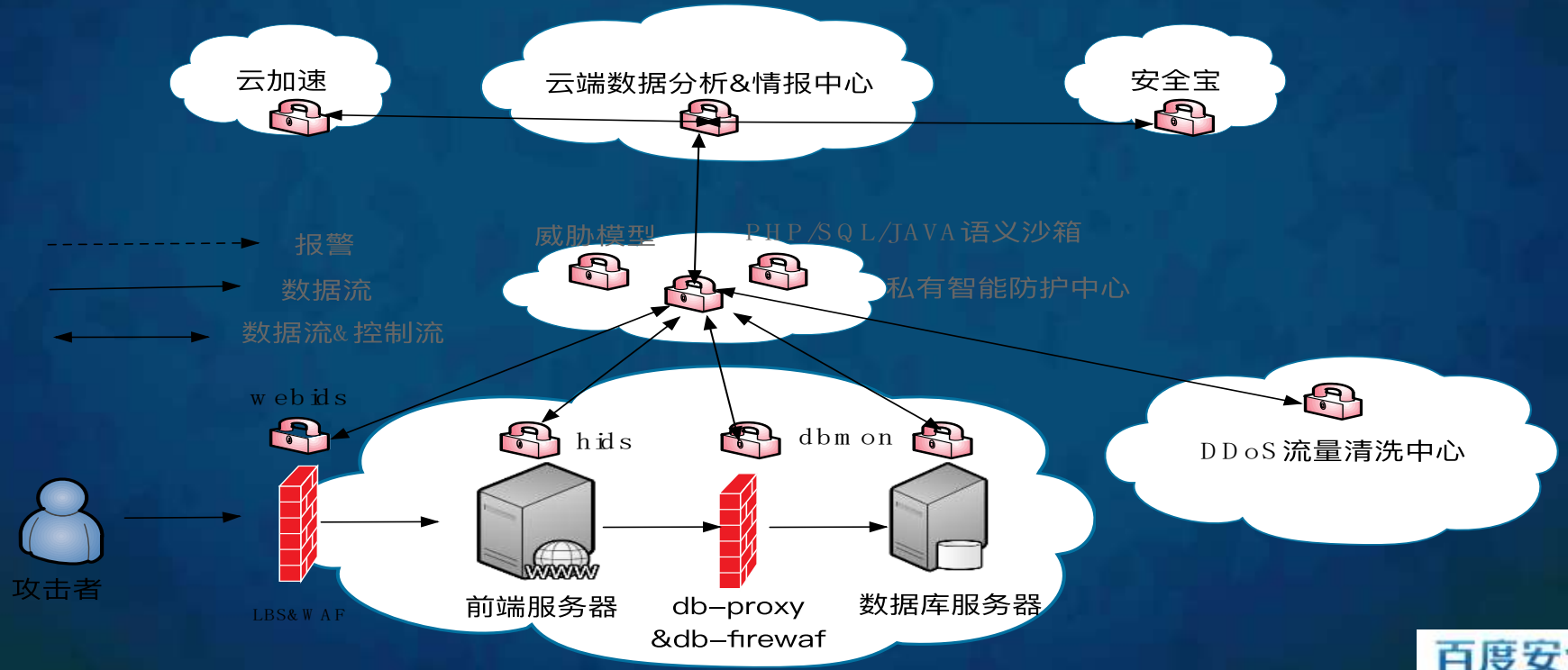
单点防护：各自为战 形式上统一



纵深防御：多点设防 协同联动



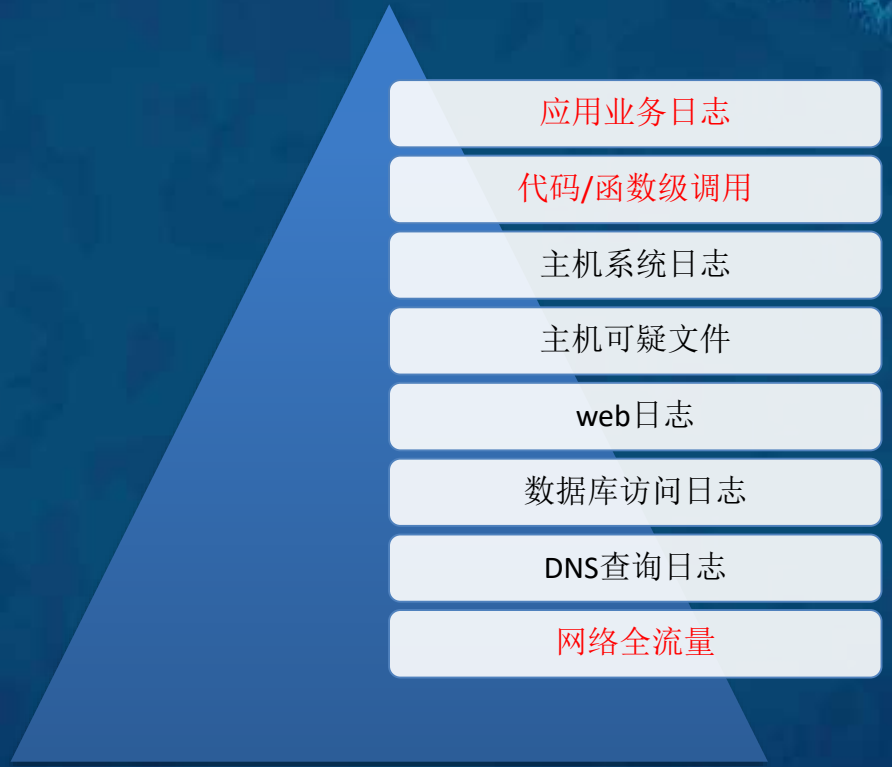
智能防护：数据驱动 检测&响应



智能防护:低调中的演进 润物细无声

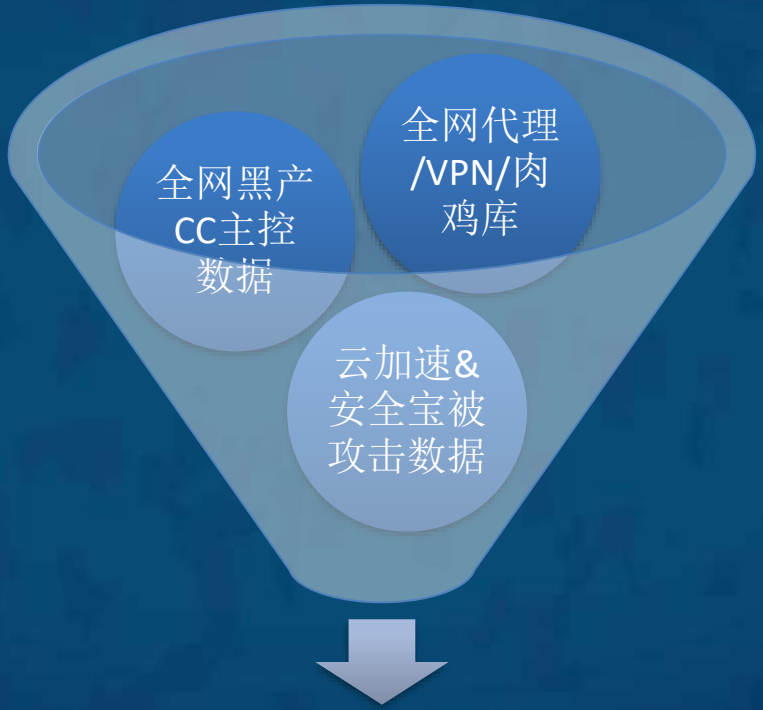
百度安全
Baidu Security

数据源：纵向全IT协议栈数据



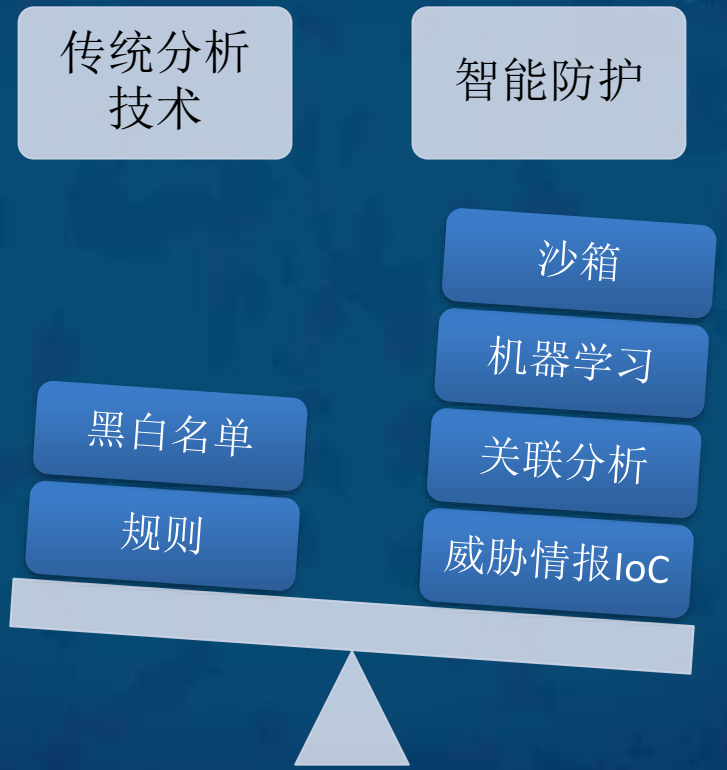
百度安全
Baidu Security

数据源：横向关联云端大数据



百度安全
Baidu Security

检测方式

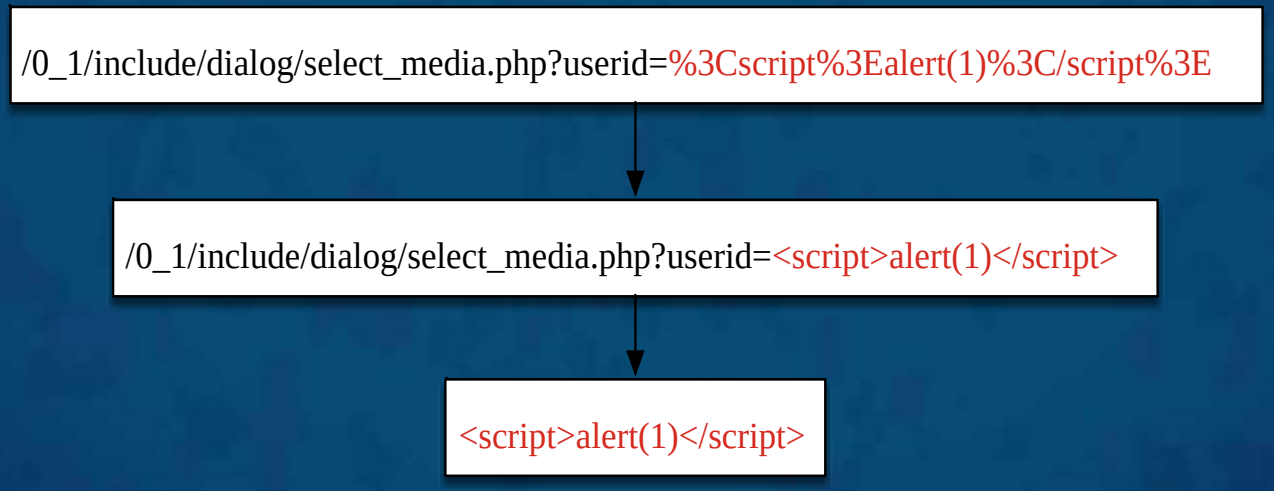


百度安全
Baidu Security

检测方式：语义沙箱

文本层面用正则表达式检测SQL注入、XSS犹如用人的语言理解鸟语，误报漏报难以避免

检测方式：语义沙箱



单向语义沙箱本质上是识别符合SQL/PHP/JAVA/JS语法的代码片段

百度安全
Baidu Security

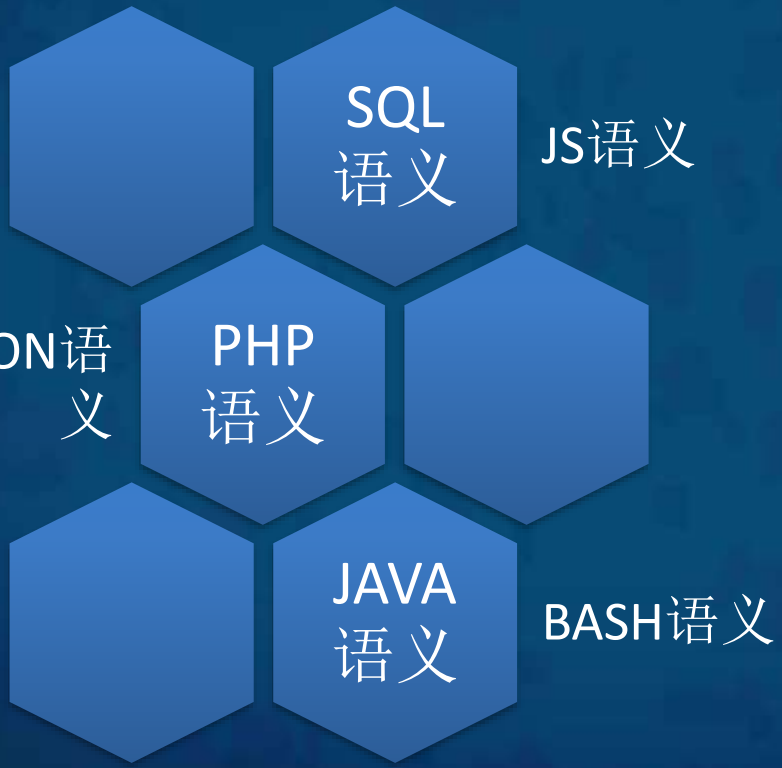
检测方式：语义沙箱

GET /0_1/include/dialog/select_media.php?
userid=cat+%2fetc%2fpasswd

200 OK
root:*:0:0:System Administrator:/var/root:/bin/sh
daemon:*:1:1:System Services:/var/root:/usr/bin/
false

双向语义沙箱检测的本质是同时观测请求内容是否包含语法片段以及应答内容是否匹配

检测方式：语义沙箱



- 四叶草&i春秋 CTF攻防赛
- 开放合作，近期将以SaaS形式逐步对外开放能力

百度安全
Baidu Security

检测方式：关联分析

- 单一数据源分析 容易盲人摸象 过于片面
- 多数据源关联分析 可以提升准确度
- 关联分析本身不是新技术 过去数据处理能力有限 不敢关联的现在都可以尝试去做

百度安全
Baidu Security

检测方式：关联分析

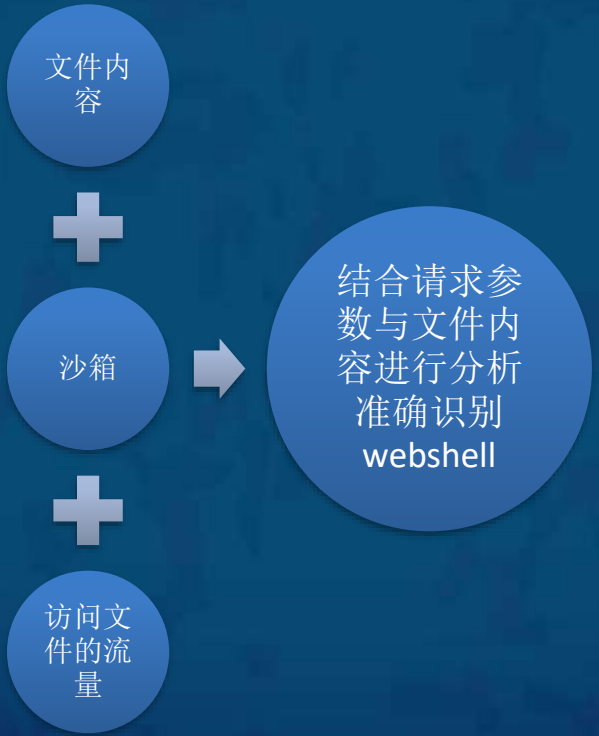
以webshell检测为例

单纯分析文件，如果重要攻击载荷在请求中将难以识准确别

单纯分析流量，加密以及编码的流量难以准确识别

百度安全
Baidu Security

检测方式：关联分析

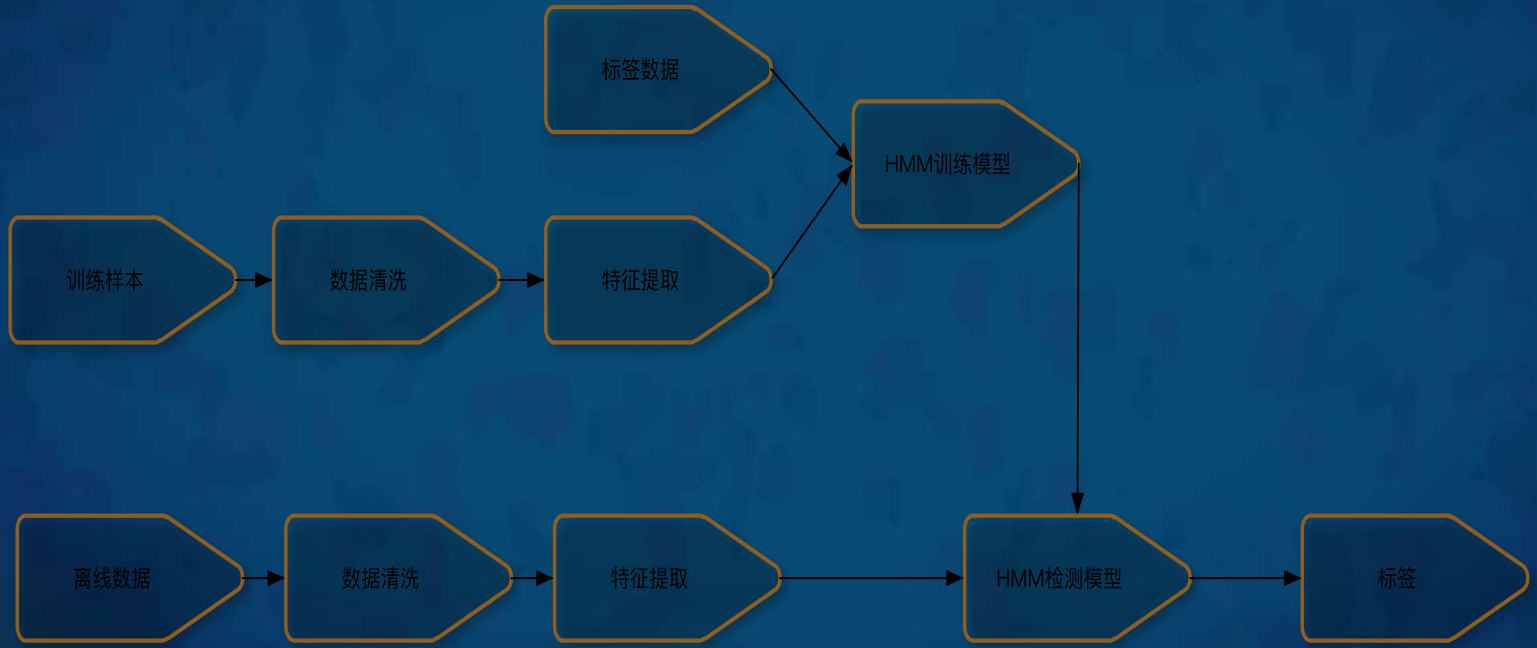


检测方式：机器学习&威胁情报落地

以白找黑 vs 以黑找黑

百度安全
Baidu Security

检测方式：机器学习&威胁情报落地



以白找黑：异常识别

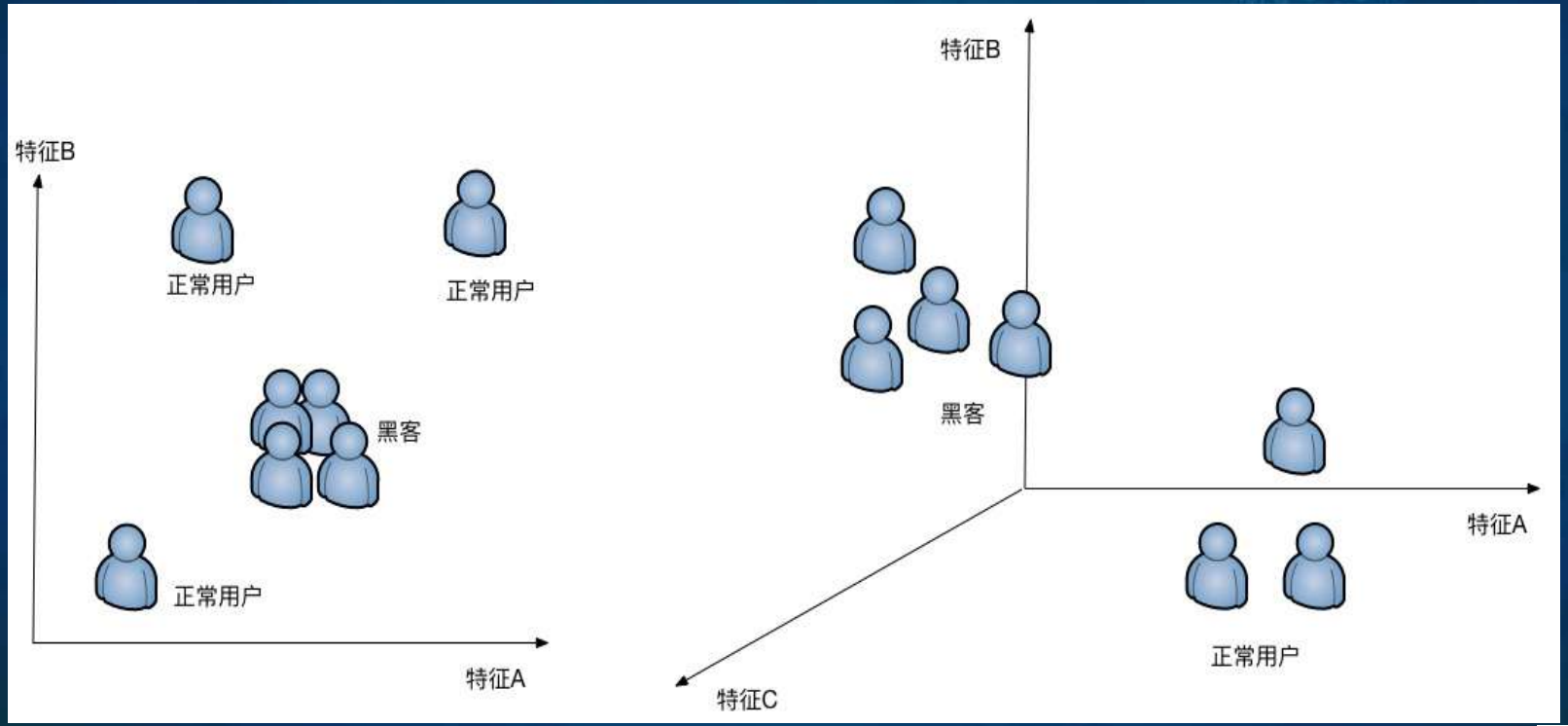
百度安全
Baidu Security

检测方式：机器学习&威胁情报落地

以白找黑：异常识别

百度安全
Baidu Security

检测方式：机器学习&威胁情报落地



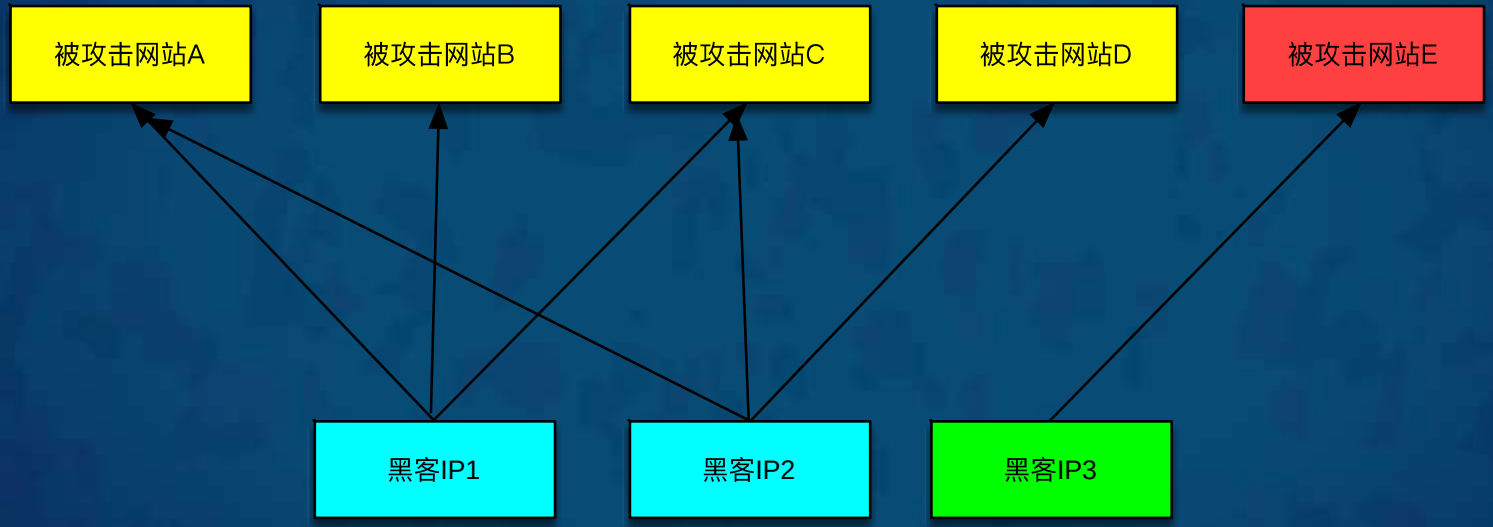
百度安全
Baidu Security

检测方式：机器学习&威胁情报落地

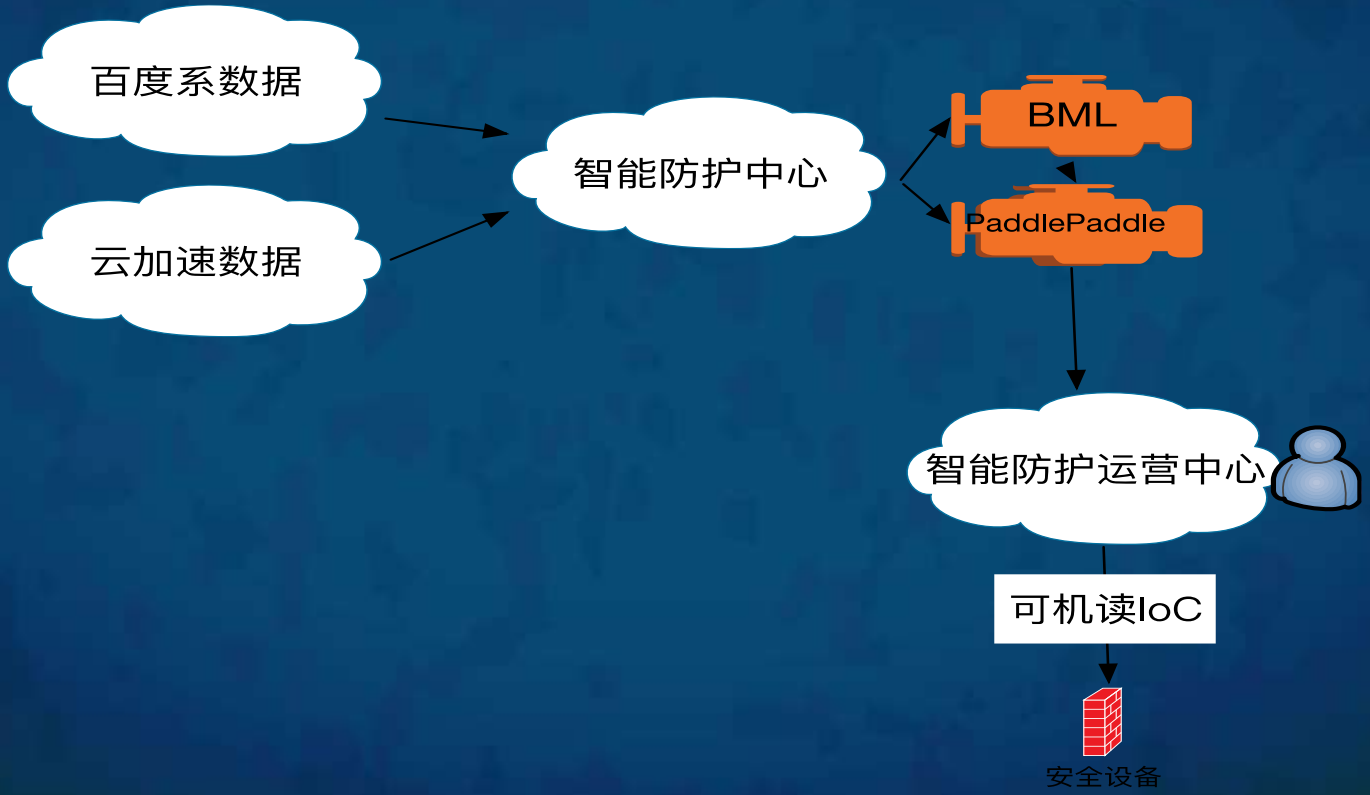
通过SVM、RNN等算法，结合云端黑样本以及海量数据，以黑找黑

百度安全
Baidu Security

检测方式：机器学习&威胁情报落地



检测方式：机器学习&威胁情报落地



百度安全
Baidu Security

线下问题讨论：欢迎关注我的公众号

百度安全
Baidu Security

THANK YOU

百度安全
Baidu Security