C3安全峰会
CYBER | CLOUD | COMMUNICATION

态势感知——实现自适应安全架构

江玮
普华永道管理咨询（上海）有限公司
信息安全咨询主管高级经理

pwc

议 题

- 自适应安全架构
- 用户及实体行为分析
- 内容感知数据丢失防护
- 积极安全防御整体体系

C3安全峰会
CYBER | CLOUD | COMMUNICATION

自适应安全架构

C3安全峰会
CYBER | CLOUD | COMMUNICATION

普华永道2017全球安全态势报告

52%
入侵检测
52%的受访企业拥有入侵检测工具

51%
主动监测
51%的受访企业主动监测和分析信息安全情报

48%
弱点测评
48%的受访企业实施薄弱环节测评

47%
威胁测评
47%的受访企业实时信息威胁测评

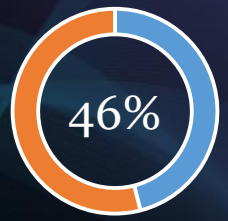
47%
SIEM工具
47%的受访企业拥有信息安全与实践管理(SIEM)工具.

45%
威胁资讯
45%的受访企业订阅信息威胁情报服务.

44%
渗透测试
44%的受访企业实施渗透测试.

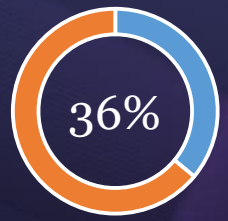
培训

46%的公司加入了隐私训练和意识培训



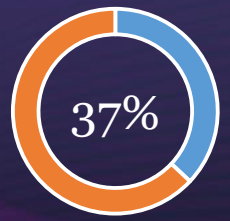
政策

36%的公司更新隐私政策与程序



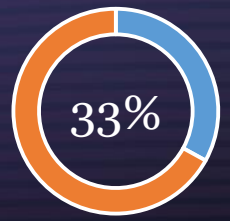
隐私

37%的公司加强了隐私事件应对机制



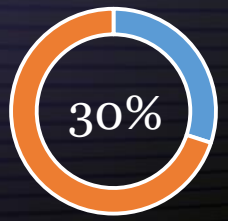
评估

33%的公司对隐私进行评估



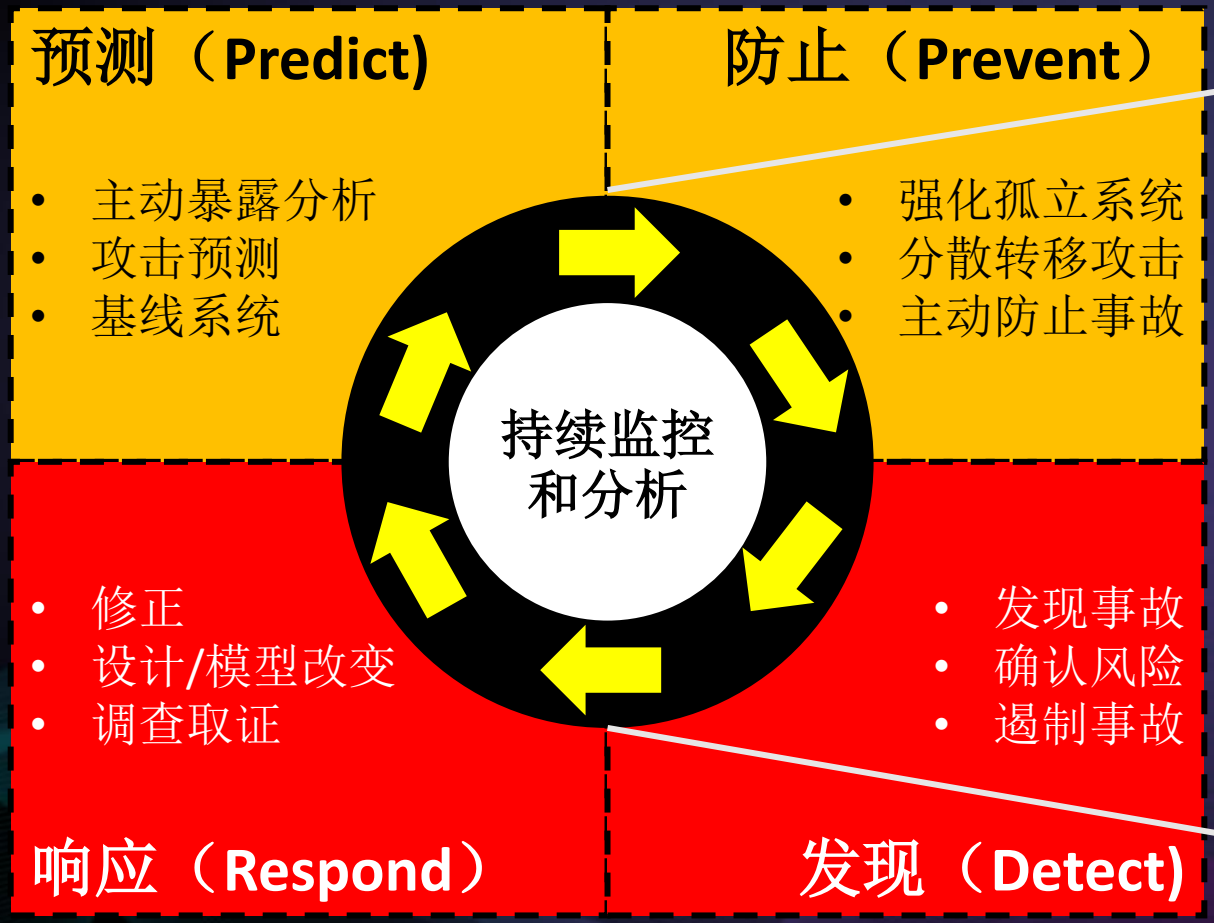
大数据

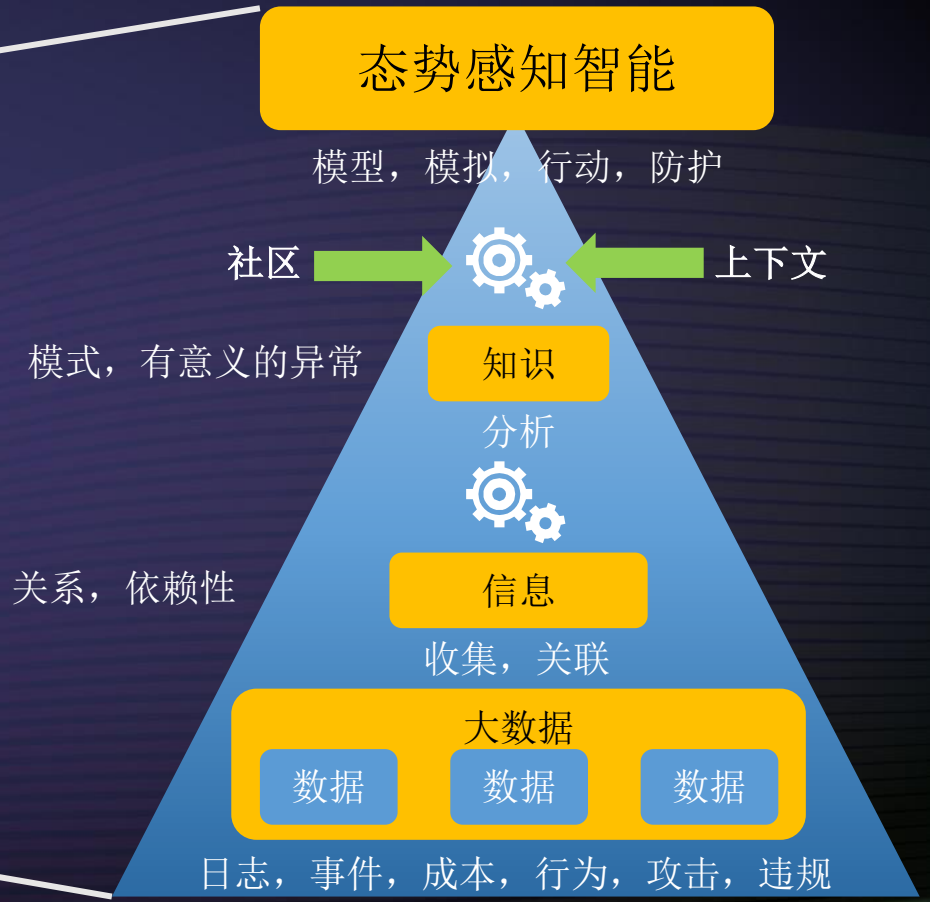
30%的公司使用大数据，数据分析或者数据去身份化技术



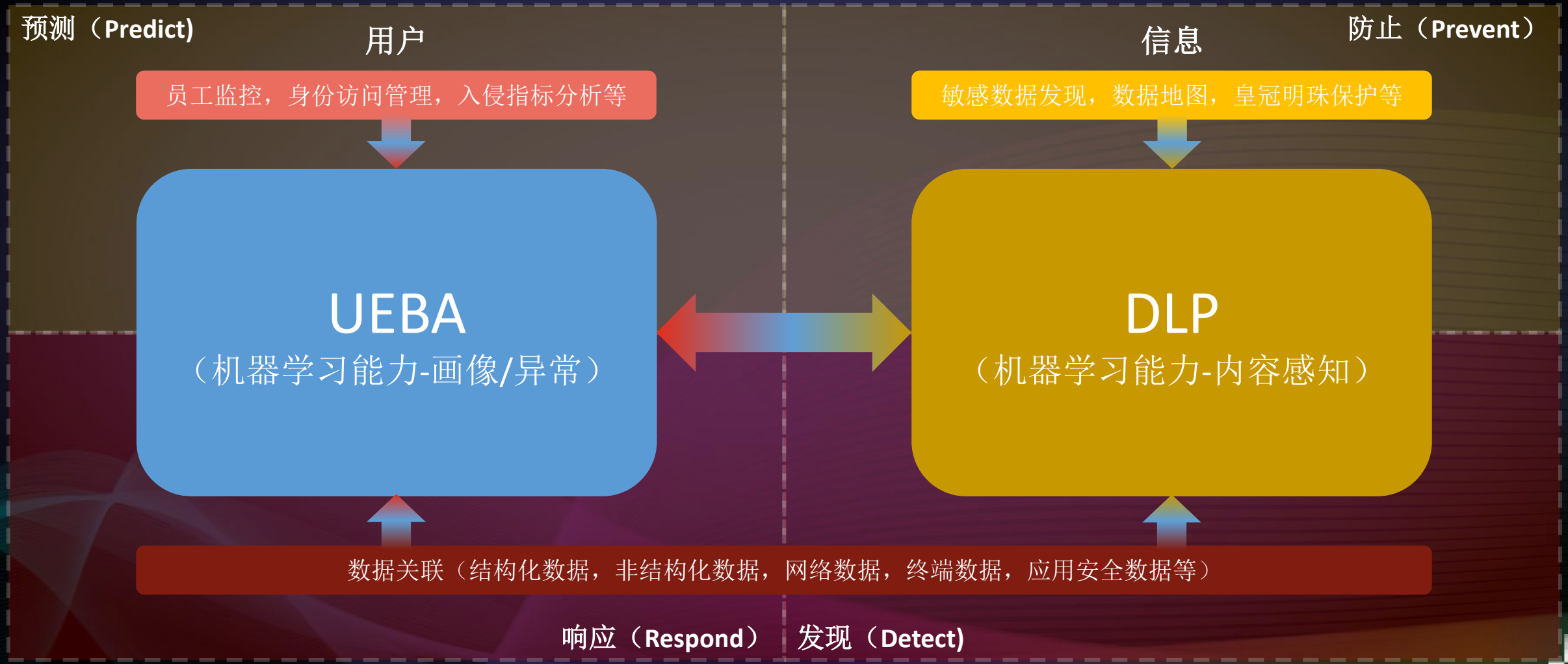
pwc

自适应安全架构


预测（Predict)
主动暴露分析
攻击预测
基线系统
防止（Prevent）
强化孤立系统
分散转移攻击
主动防止事故
持续监控
和分析
修正
设计/模型改变
调查取证
响应（Respond）
发现事故
确认风险
遏制事故
发现（Detect)


态势感知智能
模型，模拟，行动，防护
社区
上下文
模式，有意义的异常
知识
分析
关系，依赖性
信息
收集，关联
大数据
数据
数据
数据
日志，事件，成本，行为，攻击，违规

自适应安全架构下的防信息泄漏



C3安全峰会
CYBER | CLOUD | COMMUNICATION

用户与实体行为分析

UEBA定义

Gartner对UEBA的定义：

UEBA提供**画像**及基于各种分析方法的**异常检测**，通常是**基本分析方法**（利用签名的规则，模式匹配，简单统计等）和**高级分析方法**（监督和无监督的机器学习等），用打包分析来评估用户和其他实体（主机，应用程序，网络，数据库等），来发现与用户或实体标准画像或行为相异常的活动所相关的潜在事件。这些活动包括受信内部或第三方人员对系统的异常访问（**用户异常**），或者外部攻击者绕过防御性安全控制的入侵（**异常用户**）。

UEBA的主要特征

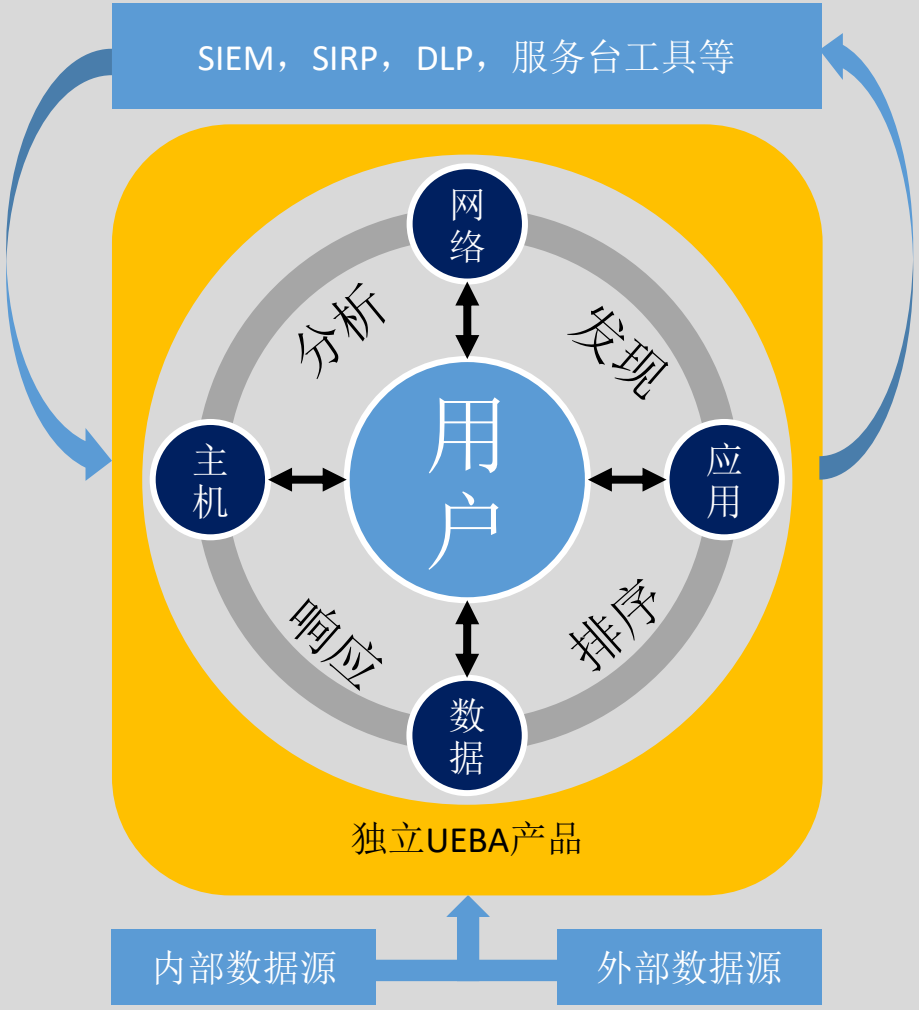
- **画像** – 采用基线使用户，用户组或其他实体的行为可视而对其画像
- **异常检测** – 通过一系列预先定义打包的分析（统计模式，机器学习，规则和签名等）和用户或实体的画像进行对比以发现异常
- **关联** – 将用户及其他实体的活动和行为关联起来，汇总个别风险行为，来突出异常活动。
- **上下文** – 利用IT目录（如AD）获得的信息作为主要数据源，并提供用户相关的**上下文**信息
- **“可信”用户** –主要用于解决安全和风险管理相关的案例，围绕组织内部“可信”的用户，无论是用户展现违规违法的行为还是其帐户或主机遭到了外部黑客的入侵
- **实时** – 提供几乎实时的监控和预警

UEBA主要功能及服务领域

C3安全峰会
CYBER | CLOUD | COMMUNICATION

融合UEBA功能的产品/平台

- 安全信息及事件管理（SIEM）
- 网络流量分析
- 数据为中心审计与保护（DCAP）
- 员工监控工具
- 身份访问管理（IAM），特权访问管理
- 终端检测和响应（EDR）

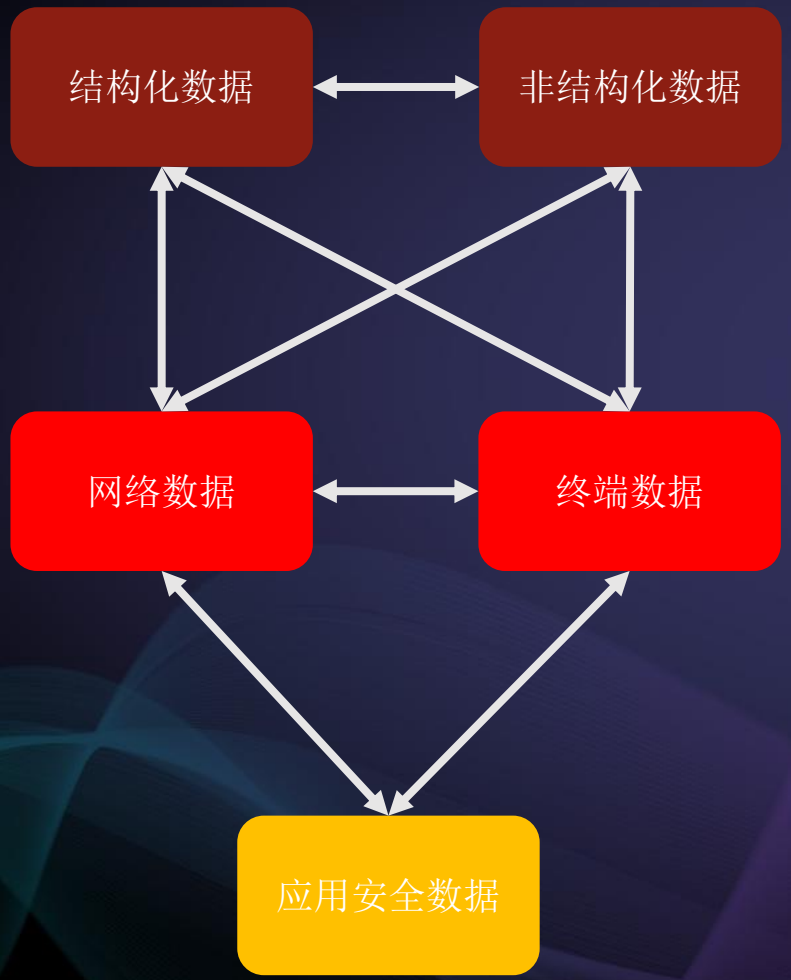


服务领域

- 外部黑客入侵（与SIEM集成）
- 内部用户威胁（聚焦用户及相关非结构化数据）
- 数据泄漏（与DLP联动）
- 员工监控（终端部署）
- 身份访问管理（强化IAM）
- 云安全（CASB服务）

pwc

UEBA数据源，集成及分析



预先定义分析的质量是关键：

- 知道哪些**数据和变量**必须被分析
- 确保从可以提供**完整信息**的”正确 “数据源读取数据
- 知道关键变量的**权重**

基于规则的分析

- 必须对数据非常了解
- 会产生过多缺乏恰当优先级的警报
- 维护更新规则的成本非常高

基于机器学习的统计分析

- 建立学习模型
- 指定各种参数的权重
- 监督或非监督学习

深度学习

画像及异常检测注意事项

- 当涉及到检测**特权用户、开发人员和有丰富IT知识的内部人员**的可疑行为时，用户和实体画像以及机器学习仍然没有得到充分证明。在这些情况下，组织仍然必须部分依赖**自己的规则**，而不是仅仅依靠基于机器学习的统计分析。这些规则可以很好地与成熟的产品模型一起使用，但用户必须负责编写它们。

- **UEBA的使用者需要注意**
- 特权用户的行为，IT开发人员和其他相关用户因为他们的工作职能，其行为可能非常不“寻常”，使得基线行为分析的画像和异常检测变得更不可靠
- 一些特定用户或者用户组的行为**从画像的一开始就可能是有问题的**，使得所建立的“正常”画像是错误的，从而异常行为无法根据画像基线检测到。这对于特权用户和普通用户都可能发生
- **重新画像**和重新建立基线是需要的，应确定其周期以及其功能是否内建在产品中或需特别安排

防数据泄露案例举例

对数据泄露相关的可能威胁进行识别并划分风险等级

案例

监控哪些用户，为什么?

- 离职用户 (离职风险)
- 特权用户 (系统及财务数据)
- 不满的用户 (情绪分析)

监控什么?

- 重复数据泄露尝试
- 大量数据传输
- 传给公司竞争对手的数据 (邮件等)
- 发给新闻网站，维基解密或其他网站的数据-第三方情报
- 与同组人相比可疑的数据访问

可视化

- 分析的关联和风险排序
- 风险最高的用户，领域，访问，组织等的图表

交叉关联

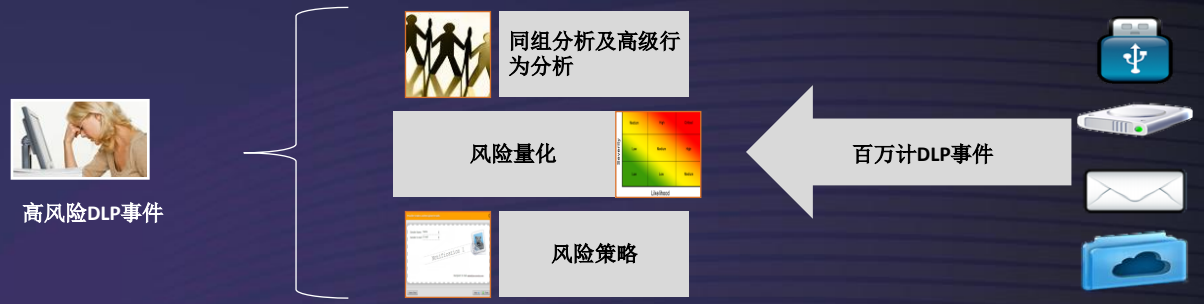
- 关联不同数据源的数据 (人事, 访问, 日志, DLP)
- 使用更多数据来开发“固有风险画像”

案例管理

- 自动案例管理和工作流
- 三振出局

组织风险评分

- 启用连续风险监控
- 业务单元风险态势10日平均



收益:

- 为DLP事件增加上下文和相关性以避免信息过载
- 过滤噪音并关注于高风险DLP报警
- 减少误报，提供智能识别，根据风险高低确定事件优先级
- 涵盖打印，上传，邮件或者USB的违规行为
- 利用高级行为分析和同组行为分析来区分意外数据丢失和恶意内部行为
- 提供离职用户情报并阻止数据泄露
- 对风险最高的组织/部门提供有针对性的DLP意识培训
- 当用户向未经授权域发送信息时进行收件人分析
- 提供报告以调整DLP策略
- 进行关联分析调查的平台

从UEBA规划实施中学到的经验

C3安全峰会
CYBER | CLOUD | COMMUNICATION

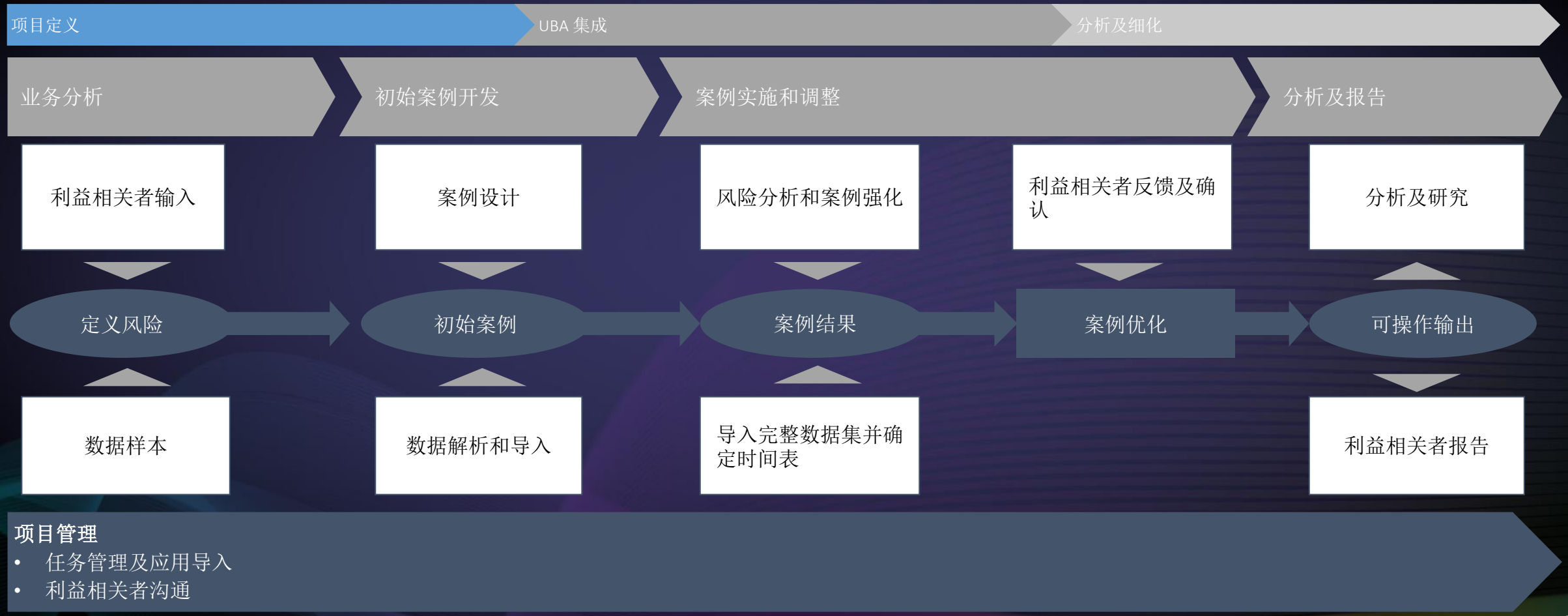
	关键经验
合适的项目团队	• 有一个专注于协调和部署的项目组将为成功奠定基础 • 专门的案例开发和初始分析专家是积极成果的必需

关注特定案例和结果	• 重点关注特权用户监控，异常行为和基于基线行为的模式检测
	• 多重分析能力
	  ○ 支持业务风险聚焦的案例优先级列表
	  ○ 确定分析所需数据源
	  ○ 数据源整合及新分析能力部署

潜在的约束	• 从多种数据源接收数据的困难 • 数据源可能没有设置正确的细节级别以获取需要的信息或必须进行显著的修改以实现各种关联 • 必须尽早评估不同司法管辖区的法律限制

pwc

总结的实施路径


项目定义
UBA 集成
分析及细化
业务分析
初始案例开发
案例实施和调整
分析及报告
利益相关者输入
案例设计
风险分析和案例强化
利益相关者反馈及确认
分析及研究
定义风险
初始案例
案例结果
案例优化
可操作输出
数据样本
数据解析和导入
导入完整数据集并确定时间表
利益相关者报告
项目管理
• 任务管理及应用导入
• 利益相关者沟通

C3安全峰会
CYBER | CLOUD | COMMUNICATION

内容感知数据丢失防护

DLP定义

DLP的定义：

DLP是一系列的**可集中管理**的技术工具或流程通过**深度内容分析**来保证**存储，使用或传输中的敏感数据**不被盗取或丢失，通常的保护对象包括个人信息（保护/合规），知识产权，重要商业信息等，通常的威胁来自内部员工，第三方以及外部黑客或网络犯罪行为等

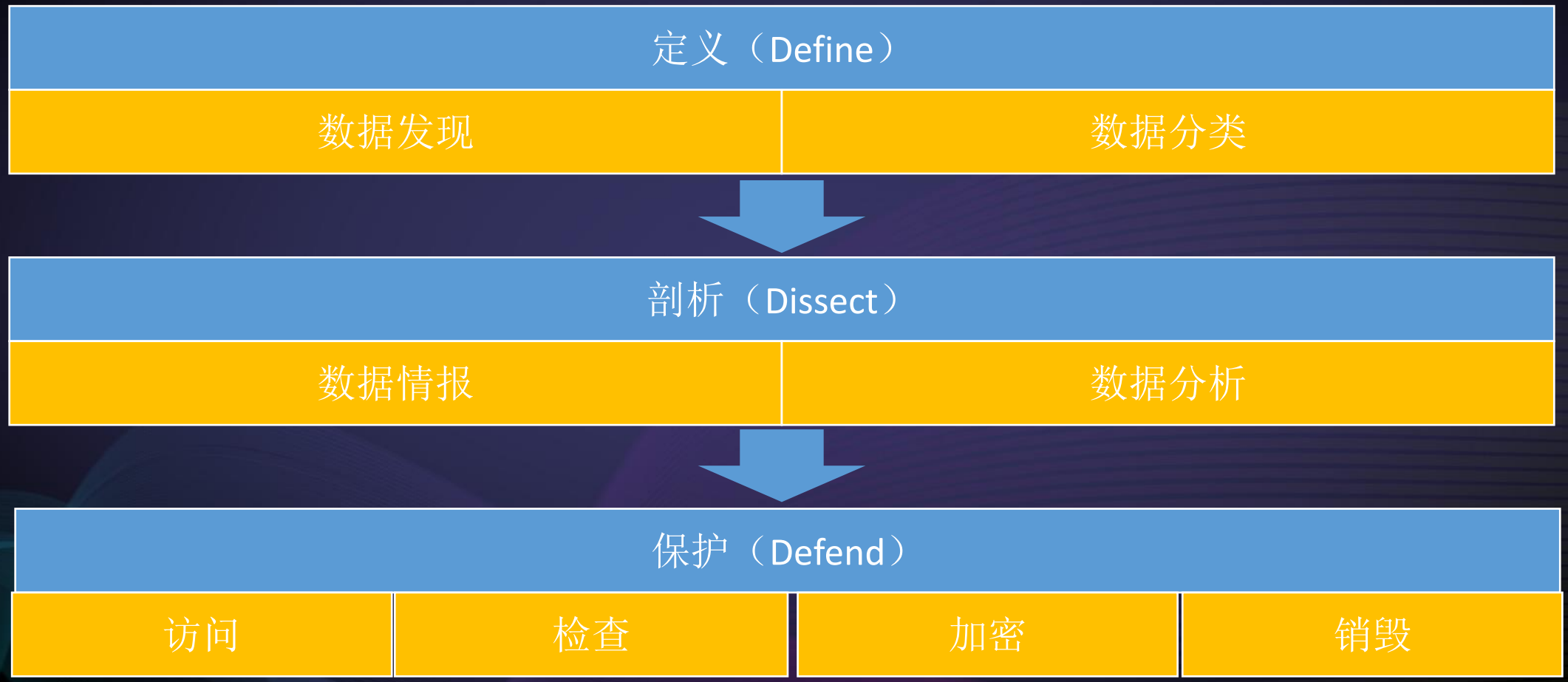
DLP主要手段

- **扫描** – 对存储中或使用中的数据进行扫描
- **定义** – 对敏感数据进行定义
- **纠正** – 提示，报警，隔离，阻止，加密
- **报告** – 合规，审计，鉴证，事故响应等

DLP主要误区

- DLP需要大量内部专家来管理和维护 – **智能DLP以及专业DLP外包服务**极大减低了这方面的需求
- DLP需要起码18个月才能见效 – **模块化的**DLP解决方案允许阶段化的实施来实现持续渐进迅速收益的数据保护
- DLP必须先建立详细的业务逻辑才能运行 – **内容感知**DLP先收集数据的使用和传输信息再与业务部门根据

以数据安全为中心的DLP



敏感数据发现

自上而下的方法聚焦于评估已知环境

通过定制化的访谈和一系列专题讨论并对相关文档的审查来更好地**理解已知环境**并收集**数据生命周期活动**的信息和所支持的控制

针对结构化和非结构化数据自下而上的方法

通过部署各种**自动化工具和扫描程序**来识别以前未知的风险区域，扫描类型可包括正则表达式，关键字匹配和精确数据匹配等，存储或传输中的数据都可被扫描以便进行“**内容感知**”检测

内容分析技术

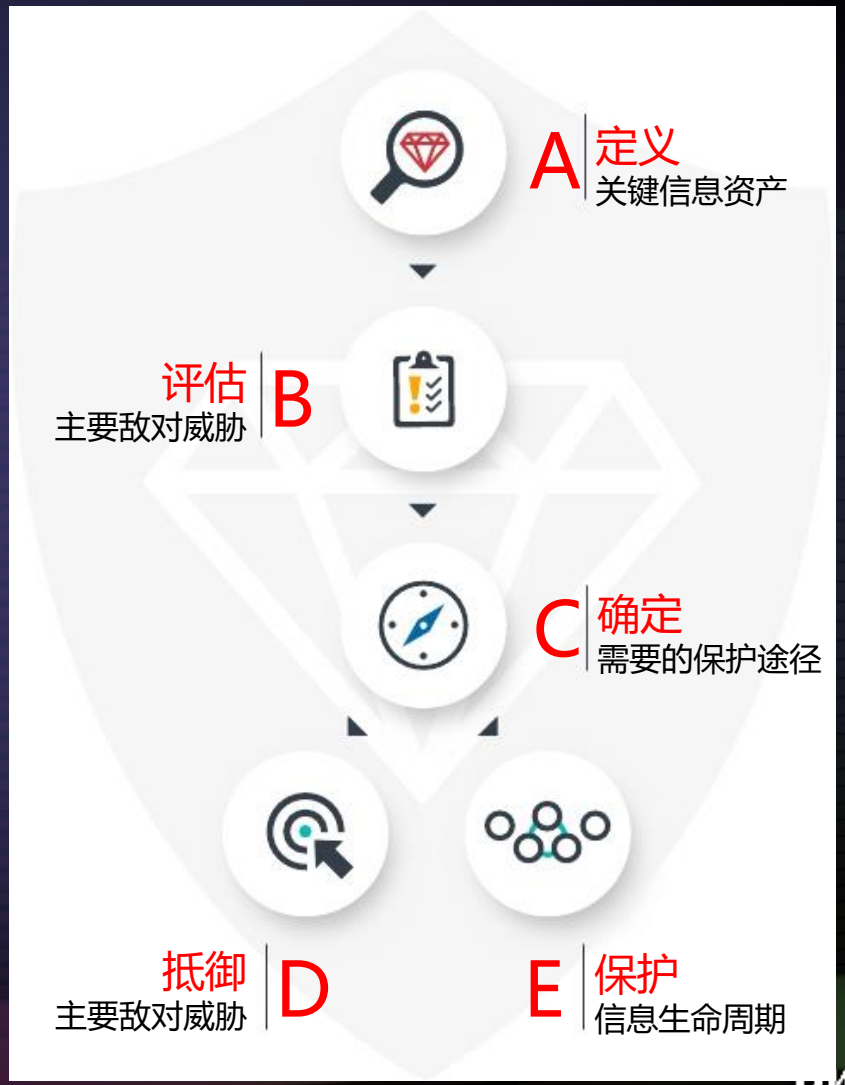
技术	说明	最适于	优势	劣势
正则表达式	通过清晰的特定规则分析内容	第一道过滤，非常结构化的数据如信用卡号等	规则处理极快，非常容易部署	误报率偏高，对敏感的非结构化数据如知识产权等几乎无用
数据库指纹	也叫精确数据匹配，通过数据库转储或者实时连接来寻找精确匹配	来自数据库的结构化数据	误报率低，保护客户/敏感数据而忽略其他类似数据	夜间数据转储不会包括上次提取以后的交易数据，实时连接将影响性能，尤其是大型数据库
精确文档匹配	根据一个文件的Hash值对所有具有同样指纹的文件进行监控	文本分析可能不可行的媒质文件或其他二进制文件	可针对任何文件，如Hash值足够大则误报率很低	对内容编辑过的文件没有意义，如通常的Office文档
部分文档匹配	通过如循环Hash之类的方法对受保护文档进行完全或部分匹配	保护具有文字信息的敏感文件，已知的非结构化内容	可以保护非结构化数据，相对低的误报率，不用依赖对大文档的完全匹配	对可保护内容总量的性能限制；受保护文档里面的一些常用词可能引起误报；必须要知道你要保护的文档是哪个
统计分析	使用机器学习，贝叶斯分析等统计分析工具来分析内容集以发现类似于被保护内容的内容中的违规行为	确定性的技术如部分文档匹配低效的非结构化内容	可分析更模糊的内容，可以强制如“对任何类似于此目录下文档的文件的外传进行报警”这种策略的执行	容易误报，需要大量的源内容，越多越好
概念/词汇	通过字典，规则以及其他分析手段的组合来保护类似于“想法”的模糊内容，如内部交易，性骚扰，找工作等	无法根据已知的文档，数据库或其他已注册数据源进行匹配的完全非结构化的内容	可以发现其他技术连监视都不可能的松散规则的违反	很难由用户自定义，规则集需要供应商大量经验的积累。误报率较高
预置的类别	为通用的敏感数据种类预先设立的具有规则和字典的类别	任何符合规定类别的东西，如隐私合规，行业指引等	配置及其简单，通常来说，预置的类别可以满足许多企业大部分的数据保护需求	只适用于容易分类的规则和内容

皇冠明珠数据保护

数据生命周期

网络攻击链

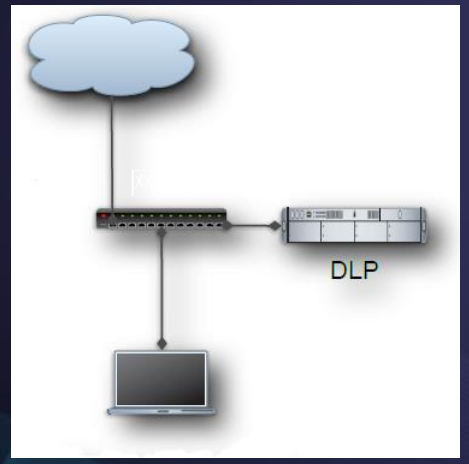
执行侦测 → 获取访问 → 维持控制 → 损害信息 → 利用信息



DLP种类 – 网络DLP（传输中的数据）

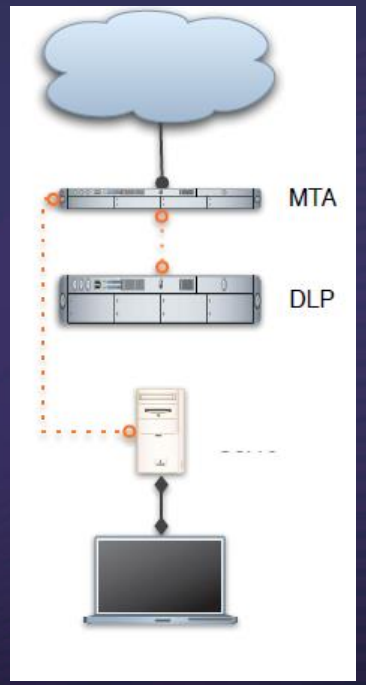
通常是无代理的DLP，提供对网络传输的数据的可视化和控制。部署物理的或虚拟的服务器对邮件，Web或即时消息等通讯进行监控

- **网络监控**
  - 通常是被动式监控，在网关的调试端口部署

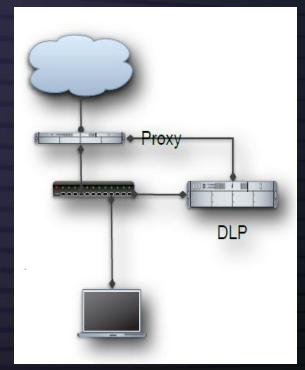


  - 需要50M/s~500M/s的网络开销

- **邮件**
  - 可以具备隔离，加密集成，过滤等功能。
  - 通常与MTA（邮件传输代理）集成，但对内部邮件基本无效。



- **过滤/阻断及代理**
  - **桥接**
  - 将DLP直接串联在链路中对所有通讯进行检查（为不影响通讯效率可能因为系统本身性能而漏过应该过滤或阻断的内容）
  - **代理**
  - 通过与代理服务器集成对特定的协议进行监控如果与反向SSL代理服务器集成甚至可以监控加密的SSL报文



  - **TCP毒药**
  - 监控到违规行为时通过注入TCP重置字段强行断开TCP连接，适用于所有TCP协议

DLP种类 – 内容发现DLP（存储中的数据）

通常是通过各种扫描工具对存储在各种地方的数据进行内容发现，可以通过DLP管理服务器对某种无论如何存储，共享或使用的数据实施单一规则

- **内容发现组件**
  - **终端发现**：对**工作站或笔记本**进行内容扫描
  - **存储发现**：对**文件服务器，SAN或NAS**进行扫描
  - **服务器发现**：具体**应用程序**扫描，如邮件服务器，文档管理服务器，数据库服务器等。

- **内容发现技术**
  - **远程扫描**：通过**文件共享或者应用协议**与服务器连接进行远程扫描并将结果发送到中央规则服务器
  - **代理扫描**：在**本地安装代理程序**使用本地资源进行扫描，通常比远程扫描高效，在终端处应适用与终端DLP同样的代理程序
  - **内存驻留代理扫描**：不在本地安装完整的代理程序，仅在需要扫描时**将代理驻留在内存中**执行，扫描完后即完全清除。

- **措施**
  - **警报/报告**：在管理服务器创建类似网络违规的事件
  - **警告**：通过邮件等手段通知用户违规行为
  - **隔离/通知**：将**文件移到管理服务器**并留下如何恢复的文本通知
  - **隔离/加密**：将**文件本地加密**并留下如何申请解密的文本通知
  - **隔离/访问控制**：改变**访问权限**
  - **移除/删除**：不通知地将文件**移至管理服务器或直接删除**

DLP种类 – 终端DLP（使用中的数据）

- **关键能力**
  - **在网络堆栈中监控和执行**：执行与**受限网络**中一样的规则
  - **在系统内核中监控和执行**：通过**植入操作系统内核**，可以监控诸如复制粘贴等用户行为，并且可以发现并阻止诸如通过加密或修改原文件以躲避监控的尝试
  - **在文件系统中监控和执行**：在文件**存储的地方进行监控**，可防止敏感信息通过USB或其他端口进行传输。
- **重要使用案例**
  - 强制执行网络规则或针对更不安全的网络更改规则
  - 限制**可移动存储中**的敏感内容，如USB，CD/DVD，家用存储，智能手机或平板电脑等
  - 限制**复制粘贴**敏感数据
  - 限制可使用敏感内容的**应用程序**
  - 与企业级DRM结合自动**基于文档内容部署访问控制**
  - 对敏感内容的适用进行**审计**形成合规报告
- **部署注意事项**
  - 终端代理和规则应该由控制传输中的数据和存储中的数据的DLP管理服务器**统一管理**
  - 规则制定和管理应和其他DLP规则**完全集成在一个界面里**
  - 事故应该由中央管理服务器**统一管理**
  - 应该使用与网络/服务器**相同的内容分析技术和规则**
  - 规则应可以根据终端是否在线**自动调整**
  - 应通过企业级软件**部署工具**进行代理部署
  - 规则更新应通过DLP管理服务器或企业级**软件更新工具**

DLP部署框架


关键数据保护目标
知识产权保护
商业伙伴合规要求
合规要求
洞察驱动框架
洞察
建造
教育
行动
评估
获得洞察：
记录行为
构建方向
正常行为
定义异常
确定关注点并与业务领导分享
构建基线规则：
将已知敏感度结合上下文来构建基线
开发用户提示规则
开发管理员报警规则
对用户进行实时安全规则培训：
各种提示
正当理由
游戏化
行动：
显示阻断
强制规则
通知管理员
出口加密
评估并迭代安全规则：
定期审核风险及规则执行
重构方向
优化分类和规则
分析
教育
评估
合规框架
对合规性数据进行分析和控制：
对外邮件
USB/CD/DVD
未加密FTP
上传云
对用户进行实时安全规则培训：
各种提示
正当理由
游戏化
评估并迭代安全规则：
定期审核风险及规则执行
重构方向
优化分类和规则

C3安全峰会
CYBER | CLOUD | COMMUNICATION

积极安全防御整体体系

C3安全峰会
CYBER | CLOUD | COMMUNICATION

积极安全防御体系:
管控、侦测、分析、监控、处理、响应

降低系统风险: **安全可见性、洞察力和攻击防范**

pwc

C3安全峰会
CYBER | CLOUD | COMMUNICATION

Thank You

pwc