网络空间威胁对抗防御技术研讨会
暨 第十届安天网络安全冬训营

沧海横流

执行体全量识别与精细管控

端点执行体采集和全量识别

——安天智甲支撑客户执行体治理的实践

安天 | 端点安全部

目 录

01 端点执行体治理的必要性

执行体是端点内系统与业务运行的载体

网络空间中无论是系统程序、应用软件还是恶意软件、代码模块、漏洞利用工具，本质都是以**执行体作为载体。**

执行体既是一切网空服务和业务活动的支撑，也是**网络空间对抗中的“弹药和装备”。**

全量执行体泛指所有**带有执行机理，可以进行局部或全部执行，以及可能创造执行机会的一切数据体。**

攻击独有	执行体分类	开发形态	执行态格式	I/O存储位置	环境要求	执行入口	样本形态
否	编译可执行程序	编译型语言源码	二进制	磁盘文件/内存	对应的操作系统、硬件架构	交互执行、调用、自动执行	文件/内存镜像
否	伪编译程序	解释型语言源码	二进制	磁盘文件/内存	对应的解释环境、或打包编译了解释器	加载、调用	文件
否	脚本	脚本语言源码	文本	磁盘文件/内存	对应的解释环境	加载、调用	文件
否	Shell	Shell源码	文本	磁盘文件/交互命令/内存	对应的命令解释器	执行、输入	文件
否	宏（脚本分支）	脚本语言源码	二进制/文本	OLE对象/磁盘文件/内存	对应的格式文档/文件解释器	宏加载	载体文件/宏块
否	引导记录	汇编指令	二进制	扇区数据/内存	对应的BIOS/UEFI	BIOS/UEFI自举加载	文件镜像
否	裸机程序	汇编指令/编译型语言源码	二进制	ROM对象/内存	对应的硬件	设备自举加载	文件
否	固件	汇编指令/编译型语言源码	二进制	ROM对象/内存	对应的硬件	BIOS调用设备自举加载	升级包/文件镜像
否	微码	自定义	二进制/文本	ROM对象/内存	对应的硬件	BIOS调用设备自举加载	文件
是	文件格式溢出代码	汇编指令	二进制	磁盘文件/内存	对应的格式文档/文件解释解析器/处理器	文档加载溢出	载体文件
是	网络服务溢出代码	汇编指令	二进制	流量数据/内存对象	对应的网络服务或网络数据处理程序	数据处理溢出	数据包文件

端点环境内执行体是复杂且难以收敛的

- 具备各种业务用途
- 支持复杂交互操作
- 用户具有自主操作权限
- 来源多且难以控制
- 环境难以固化、易改变

按程序类型维度划分

- **类型**：系统软件、工具软件、办公软件、娱乐软件……
- **来源**：系统升级、程序更新、企业分发、用户自行安装……
- **版权**：正版、工具激活版、破解版、第三方汉化……

按数字签名维度划分

- **有无签名**：具有签名、无签名
- **签名有效性性**：合法签名、过期签名、无效签名、国内厂商签名、国外厂商签名……
- **签名安全性**：被盗用签名

按运维管理维度划分

- **活跃情况**：持续活跃、一般活跃、长时间未活跃、从未活跃……
- **合规性**：是否在资产台账内、是否最新版本、是否是禁用程序……
- **生僻对象**：生僻编译器、生僻厂商、生僻数字签名……

按安全性维度划分

- **脆弱性**：是否存在漏洞、有无UAC提示、是否修改系统配置
- **暴露面**：是否有对外服务、是否使用敏感端口
- **权限**：执行权限、配置修改权限、网络访问权限、启动项操作权限……

某办公机执行体采集数据

操作系统：Windows 10 企业版 21H2

系统安装时间：2022/3/1　持续监测时间：5天

综合统计

执行体总数	可执行程序	模块	系统文件	驱动文件	工具软件	办公软件	业务软件	文档	脚本	其他应用
12W	1.5W	9.5W	5.6W	480个	1800个	9000个	8000个	1800	5600个	2.3W

数字签名统计

有签名	无签名	有效签名	无效签名	过期签名	吊销签名	证书格式异常	风险签名	主流厂商签名	生僻签名	配置清单中的签名
7.3W	4.7W	6.2W	2600个	300个	15个	8个	1个	5.8W	120个	1300个

按属性统计

编译器种类	壳种类	供应商	共享文件	隐藏文件	正版软件	未授权软件	持续活跃	偶尔活跃	持续静默	静默
36个	124个	92个	2500个	3200个	92个	4个	6000个	3.4W	3W	5W

安全性统计

有已知漏洞	敏感命令创建的文件	服务执行	启动项执行	生僻技术栈	生僻编译器	生僻壳	生僻厂商	名称伪装	图标伪装	敏感端口监听
7个	600	370个	136个	3个	4个	2个	4个	8个	73个	3个

某家庭主机（兼SOHO办公机）执行体采集数据

操作系统：Windows 11 家庭版 22H2　　系统安装时间：2020/1/15　　持续监测时间：5天

综合统计

执行体总数	可执行程序	模块	系统文件	驱动文件	工具软件	办公软件	业务软件	文档	脚本	其他应用
26W	3W	23W	5.8W	640个	4.7W	2000个	300个	2300	7700个	1.2W

数字签名统计

有签名	无签名	有效签名	无效签名	过期签名	吊销签名	证书格式异常	风险签名	主流厂商签名	生僻签名	配置清单中的签名
20.3W	5.7W	18.2W	6000个	230个	37个	9个	3个	9.8W	2000个	230个

按属性统计

编译器种类	壳种类	供应商	共享文件	隐藏文件	正版授权	未授权软件	持续活跃	偶尔活跃	持续静默	静默
55个	114个	125个	1.3W	1200个	67个	37个	8000个	11.4W	6W	13W

安全性统计

有已知漏洞	敏感命令创建的文件	服务执行	启动项执行	生僻技术栈	生僻编译器	生僻壳	生僻厂商	名称伪装	图标伪装	敏感端口监听
34个	33个	245个	138个	9个	3个	6个	19个	19个	3个	4

某国产化服务器执行体采集数据

操作系统：统信服务器操作系统V20 1050　　系统安装时间：2021/01/15　持续监测时间：5天

综合统计

执行体总数	可执行程序	模块	系统文件	驱动文件	工具软件	办公软件	业务软件	文档	脚本	其他应用
8W	1.1W	6.9W	3.2W	320个	6850个	356个	8650个	1200	1.2W	4000个

数字签名统计

有签名	无签名
5.8W	2.2W

由于统信系统签名机制特殊性，难以签名厂商信誉以及其他属性

按属性统计

编译器种类	壳种类	供应商	共享文件	隐藏文件	正版授权	未授权软件	持续活跃	偶尔活跃	持续静默	静默
12个	25个	36个	0个	320个	128个	0个	5000个	1.6W	4W	1.9W

安全性统计

有已知漏洞	敏感命令创建的文件	服务执行	启动项执行	生僻技术栈	生僻编译器	生僻壳	生僻厂商	名称伪装	图标伪装	敏感端口监听
13个	360个	78个	60个	5个	4个	2个	4个	6个	0个	12个

某移动设备执行体采集数据

操作系统：Android 11　　系统安装时间：2020/09/15　　持续监测时间：5天

综合统计

执行体总数	可执行程序	模块	系统文件	系统应用	so文件	工具软件	办公软件	文档	脚本	CLASS文件	其他应用
10W	1200个	7852个	3.2W	273个	6425个	62个	37个	1036个	8652个	4.1W	89个

数字签名统计

系统签名	普通签名
273个	135个

按属性统计

技术栈	加固壳种类	大文件	普通文件	隐藏文件	图片文件	文档文件	后台运行应用	系统预置应用	第三方应用
4个	4个	223个	35254个	3320个	128个	365个	18个	273个	135个

安全性统计

有已知系统漏洞	敏感权限申请使用应用	服务执行	开机启动	申请短信权限	申请摄像头权限	申请位置权限	申请sd卡读写权限	申请网络权限
16个	74个	65个	44个	6个	32个	18个	64个	146个

伴随着执行体而来的脆弱性与安全风险

执行体作为各类执行代码的载体，就会导致所有执行体都是风险的入口，**执行体本身是被利用的对象，是攻击工具，是被篡改目标。**

执行体中所包含漏洞是攻击者可利用入口

典型代表：【CVE-2017-0143】某Windows SMB 远程代码执行漏洞、【CVE-2021-44228】Apache Log4j2远程代码执行漏洞

各类恶意代码本质上都是执行代码，而其载体就是执行体

典型代表：蠕虫，病毒，木马，黑客工具，流氓软件，风险程序，垃圾文件，测试文件

供应链所提供的软件系统或者签名认证机制成为攻击者目标

典型代表：Solarwinds攻击事件、 NVIDIA数字签名被盗事件

需要对执行体进行有效治理，降低其所带来的安全风险

传统防御机制无法及时、准确、全面覆盖的对海量执行体进行细粒度防护

- 病毒库数量限制
- 知识库更新滞后性
- 管控规则未覆盖
- 厂商规则与用户环境的不匹配

有效的执行体治理应是**基于全量执行体采集能力**，提取执行体**共性元数据**（如来源、成分、用途、脆弱性、异常行为等），基于对元数据的统计和分析评估执行体在用户环境内的信誉值，并以此作为依据对执行体行为进行**细粒度约束与控制**，其本质是一个构建端点内**执行体安全运行基线**的过程。

- 在资源有限情况下，将对海量、差异化的执行体的治理变为可能
- 保证了执行体检测与管控的覆盖度与粒度
- 降低了对病毒库、人工管控规则的体量与时效性的依赖
- 让防护模型和知识库与用户环境更加兼容

执行体治理是对已有防护体系的补充

	AV	EPP	EDR
运行机理	主要使用黑名单病毒库对执行体进行检测	依托主防内核，在AV基础上扩展行为管控、网络管控等能力	对端点内多类数据进行采集与分析，并将病毒库升级为大量威胁情报
可解决问题	实现**对已知恶意代码的查杀**，帮助用户发现明确的“黑文件”	**增强动态防御能力**，对可操作行为的约束提升主机环境的安全性	**将单主机检测能力升级到使用全网数据的综合分析**体系，实现对高技术水平威胁的发现
未解决问题	对**大量无法检出的执行体难以准确判定属性**，存在大量“灰文件”	**难以进行精准的场景识别**，管控策略主要面向明显的违规行为，难以精细化管控	采集重点是环境要素和事件，但**并未将执行体本身的属性与行为作为重点**

在传统终端防护产品基础上，叠加执行体治理体系

有效执行体治理体系应具备的几类关键能力

安全管理员可对执行体治理全过程进行管理

执行体治理动作可自动化完成

基于多层次指标体系评估执行体信誉

执行体行为细粒度约束与权限控制

全量端点资产内执行体全要素的数据采集

基于执行体元数据与行为的精细化识别

用户业务系统可稳定运行保证

02

端点执行体全要素采集

建立全面、可落地的采集能力是识别的前提

执行体采集是治理体系中最重要的支撑能力，执行体识别地准确、管控地及时、治理过程与用户环境可良好兼容的前提一定是建立在具有合格的执行体采集能力基础之上的。

采集能力覆盖全量端点资产	执行体识别准确、可管控范围全面的必要条件
执行体的识别分析需要元数据化支持	执行体元数据采集，保障信誉识别的精准度
与业务系统良好兼容	采集过程确保业务运行不受影响
满足不同场景采集需求	满足不同类型资产、不同场景、不同时期下的多种采集要求

全要素执行体采集需要与端点环境相结合

采集8大类60+种数据类型，全面覆盖执行体识别所需

基本对象
- 文件
- 进程
- 目录
- 注册表
- 任务计划
- 内存
- 加载模块
- 服务

系统操作
- 文件操作
- 进程操作
- 打印操作
- FTP操作

网络通讯
- 网络连接
- 数据内容
- 协议
- 端口

资产信息
- 硬件配置
- 软件配置
- 资源性能
- 移动外设
- 账户信息
- 操作系统
- 域信息
- 启动链

身份权限
- 用户
- 账户
- 身份
- 权限

脆弱性
- 漏洞/补丁
- 组策略
- 安全配置
- 防护策略

暴露面
- 对外服务
- 对外开放端口
- 共享目录

日志数据
- 系统日志
- 应用日志

执行体元数据化采集赋能输出区分度统计和决策标签

执行体

执行体元数据

基础数据

文件名	大小
创建时间	数据签名
公司	作者
格式	编译器
壳信息	hash
路径	……

向量数据

字符串信息	算法信息
URL信息	函数
结构	Domain信息
运行环境	依赖库

衍生数据

位置特征	活跃方式
所需权限	资源占用
漏洞	来源
生僻信息	应用分类
应用场景	伪装信息
违规行为	终端覆盖

执行体行为数据

操作文件	操作进程
操作持久化配置	操作注册表
操作系统关键配置	利用系统漏洞
提升自身权限	缓冲区溢出

执行体网络通信数据

IP
端口
协议
原始数据包
请求URL
命令
流量

输出

- 签名清单
- 厂商清单
- 软件清单
- 活跃执行体清单
- 技术栈清单
- 壳信息清单
- 文件格式清单
- 行为标定标签
- 风险标定标签
- 威胁标定标签
- ……

确保业务系统算力需求满足的前提下完成执行体采集

执行体采集需考虑

- 内存占用
- 网络流量占用
- CPU占用
- I/O资源使用

避免无效采集

基于安天引擎自带的HASH和签名库的免采集策略

基于资产环境的免采集策略

减少重复采集

基于本地缓存的单终端数据提取去重

基于全网查询的执行体去重

数据上报性能优化

基于数据差量化数据上报策略

基于优先级的数据采集策略

基于压缩技术的数据上传

面向不同场景优化采集策略

基于细粒度场景的采集配置策略

采集能力需细粒度灵活配置满足多种应用场景（保障场景适配度）

事件&对象

- 文件对象基础属性
- 文件对象向量属性
- 文件对象扩展属性
- 进程执行事件
- 文件操作事件
- 注册表事件
- 计划任务事件
- 网络访问事件
- 系统环境信息

属性

- 基础属性
- 精简属性
- 全量属性

触发器

- 资产类数据：有变更，有差异，触发
- 系统资源数据：仅记录，有需要，有异常，触发
- 运行承载类：通过驱动模块和系统消息实时触发

方式

- 办公模式
- 实时监控
- 闲时扫描
- 定时检测
- 触发器

上报优先级

事件安全等级
事件最小集合
事件补报机制

模板类型

原厂模板
自定义模板
调试模板

模板配置

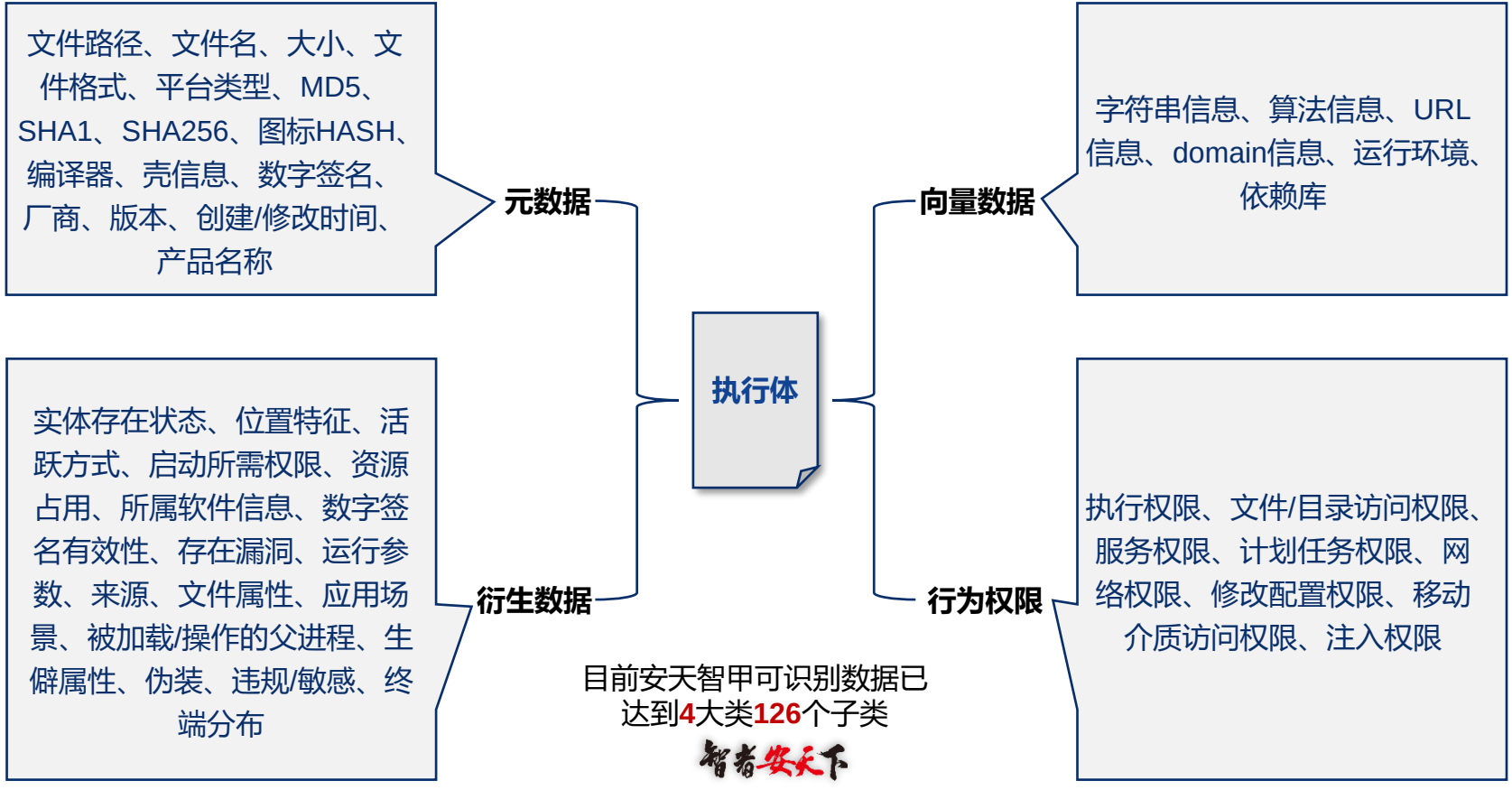
全局模板
群组模板
单机模板

03

端点执行体精细化识别

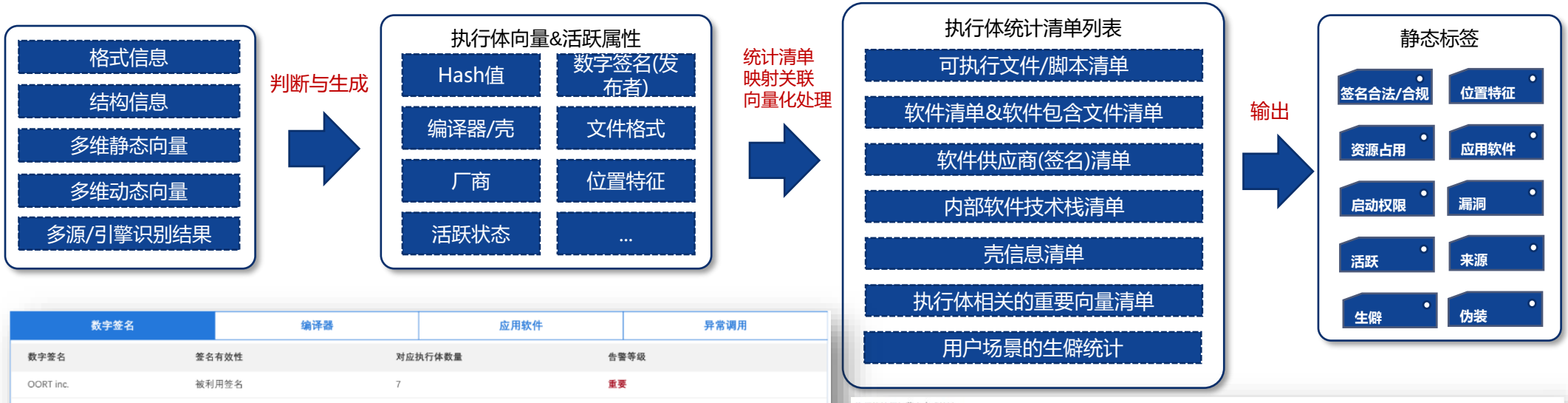
执行体的治理需构建细粒度的执行体识别

执行体识别的目标是提取**执行体中的共性资源**从执行体来源、环境、成分、用途、脆弱性、异常行为和信誉等多个维度提取共性信息。

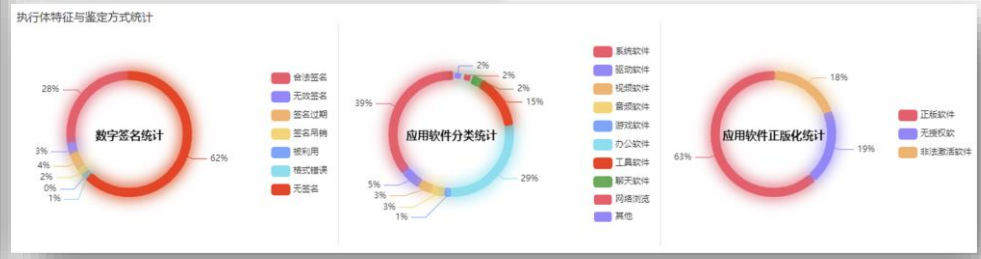


基于执行体共性信息识别多维度清单和标签

通过对执行体格式、结构、向量、活跃状态等信息采集，对属性数据计算统计归纳，对全网终端执行体在供应商、内部技术手段等情况全面摸底，**实现对执行体静态信息清单化，并输出执行体的属性识别标签。**



数字签名的生僻统计

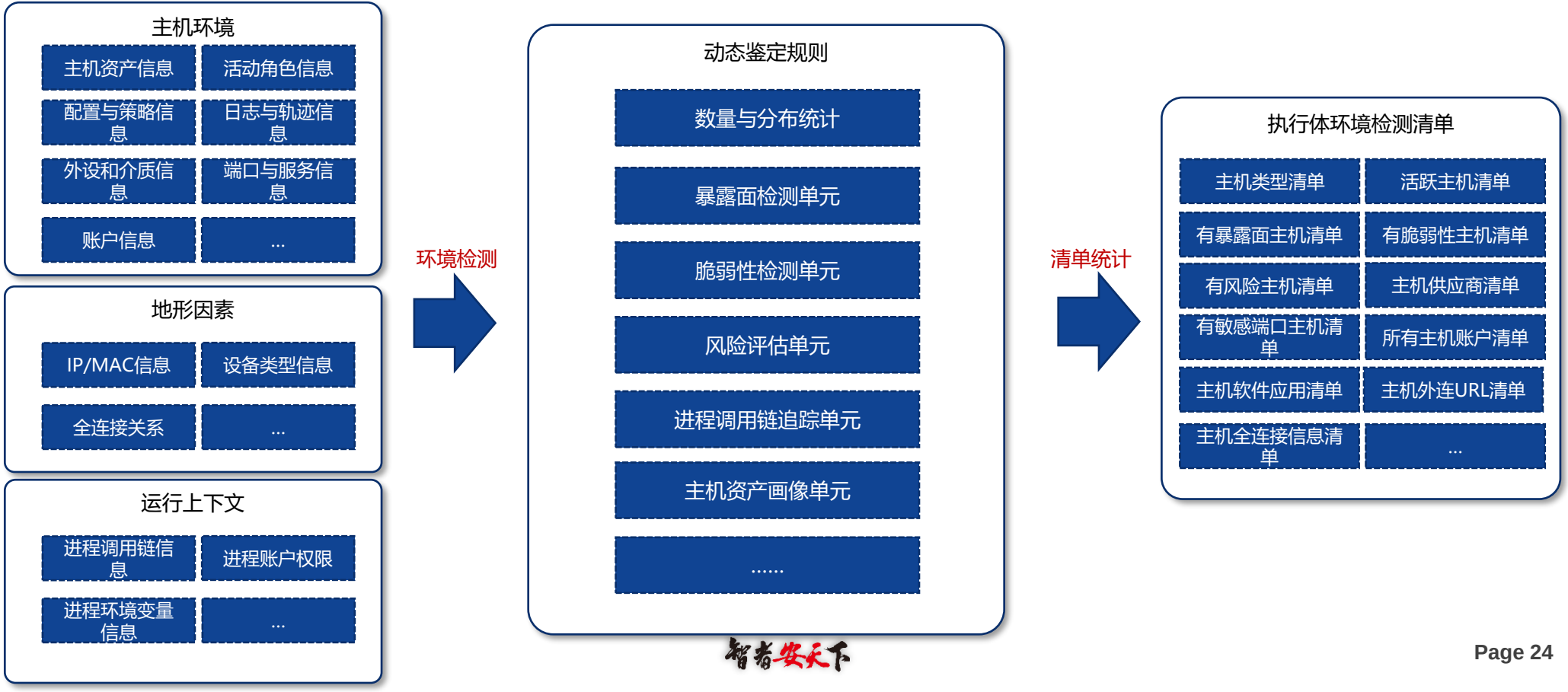


基于规则标定执行体动态行为标签

执行体动态检测并非传统的使用黑名单库对恶意代码进行判定的思路，而是对识别出的动态行为序列，**使用多种行为鉴定单元并结合人工规则设置对执行体动态行为进行评估，并以动态分析标签结果直观呈现。**

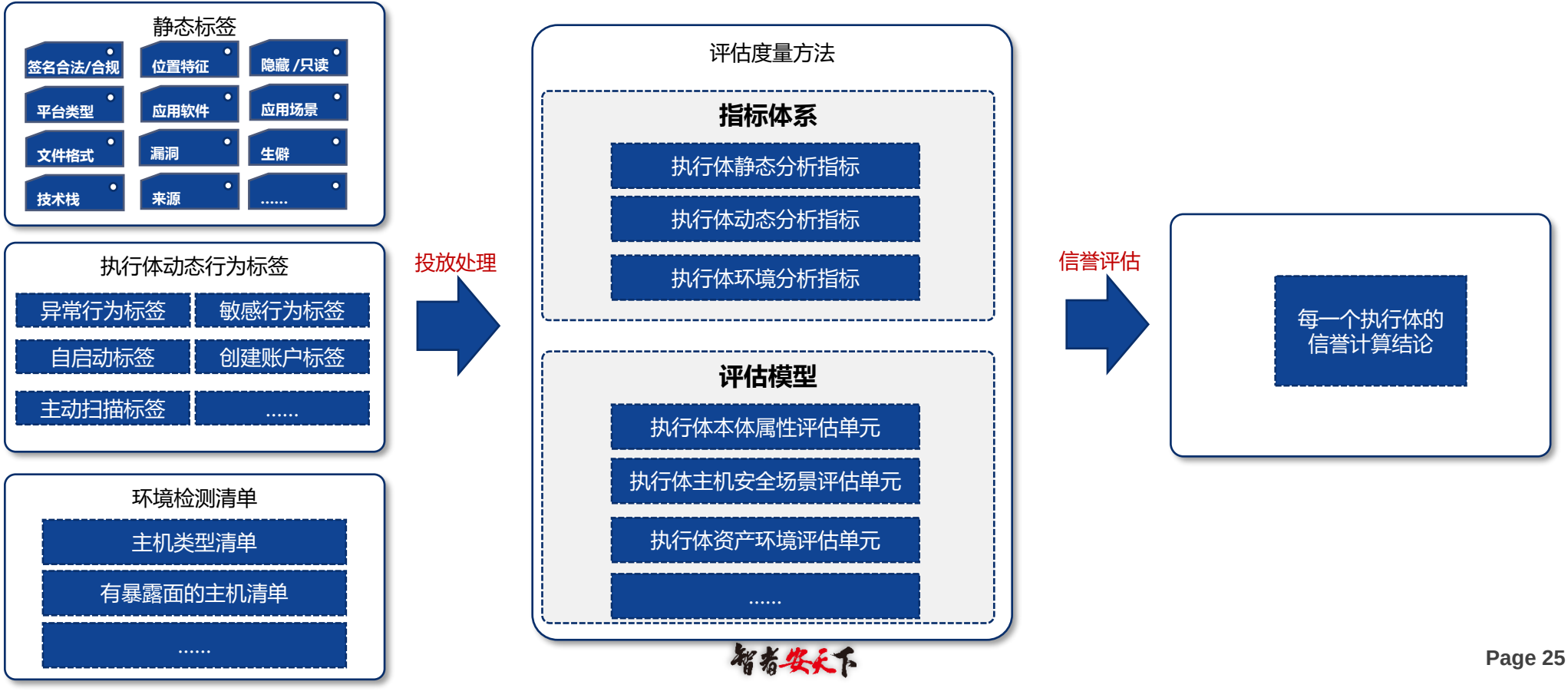
基于运行上下文构建执行体环境检测清单

执行体环境检测是基于执行体主机环境、地形因素及运行上下文，**使用数量与分布统计、脆弱性及风险评估、进程调用链追踪等多种手段，实现对执行体环境信息清单化。**



基于多层次指标体系评估执行体信誉

执行体信誉评估是基于执行体共性统计清单、运行行为标签、环境检测清单的多维度指标体系，**使用执行体本体属性评估、主机安全场景评估、资产环境评估等多种评估模型，实现对主机每一个执行体的信誉评分。**

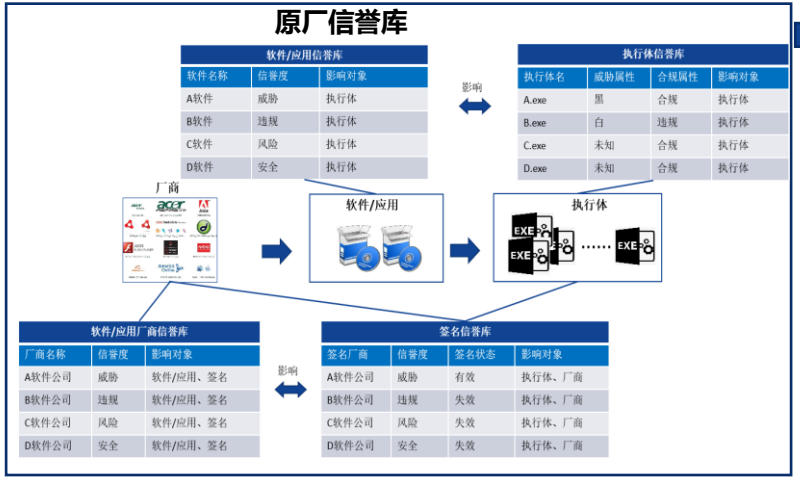


基于执行体信誉构建细粒度安全基线

按照原厂赋能与客户分析能力结合，**以客户侧基线为优先，以原厂基线为辅助**的思路构建差异化的基线。

以执行体库为基础，构建执行体信誉评价机制，**建立符合管理要求的安全基线。**

构建终端场景下差异化基线。



客户侧私有信誉库

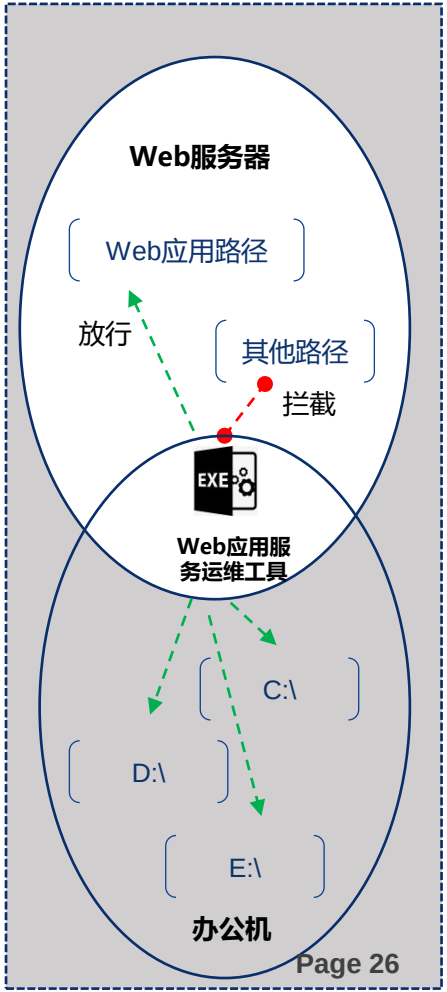
评价因素	威胁属性	合规属性
厂商	白	合规
软件	白	合规
签名	安全	合规
威胁	黑	合规
评价给结果	黑	合规
执行体基线	不符合基线要求	

非安全基线

安全基线

评价因素	威胁属性	合规属性
厂商	白	合规
软件	白	合规
签名	安全	合规
威胁	白	合规
评价给结果	白	合规
执行体基线	符合基线要求	

评价因素	威胁属性	合规属性
厂商	白	合规
软件	白	合规
签名	未知	合规
威胁	未知	合规
评价给结果	白	合规
执行体基线	符合基线要求	



基于规则灵活配置管理执行体识别规则

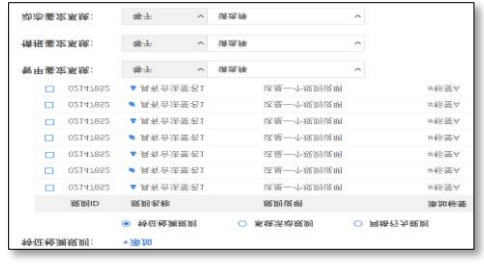
为实现对全量执行体的精准信誉判定就需使用执行体多种数据对执行体进行综合分析评估其信誉，而为了**满足针对不同类型、不同资产内执行体的识别准确性，系统支持可管理执行体识别规则能力。**

识别规则类型	可配置对象
执行体元数据	文件类型、文件名称、路径特征、数字签名、签名有效性、厂商/版权、编译器、壳、读写属性、创建/修改时间、PDB符号、hash
执行体行为	执行对象元数据特征、执行动作（进程操作、文件操作、加载操作、注册表操作、系统服务操作、外设操作、启动项操作、外设操作）、操作对象元数据特征
执行体网络访问	访问动作（使用敏感端口、访问指定网络）、执行对象元数据特征、访问地址
信誉综合分析	执行体元数据、执行体标签、情报鉴定系统结论、动态鉴定系统 结论、威胁引擎检测结论

执行体元数据识别规则配置

执行体行为识别规则配置

执行体网络访问识别规则配置



信誉综合分析规则配置

以执行体识别对全量执行体进行细粒度管控

执行体管控的核心目标是**限制全量执行体的运行权限、收敛执行体可执行动作、控制执行体可访问资源范围**，使攻击者无法借助执行体获取到攻击资源或者完成攻击动作，以此保护用户重要资产不被风险执行体篡改、破坏和窃取。

执行体管控依据是**基于执行体识别数据、信誉评估结论、执行环境、资产类型、厂商规则、用户自定义规则综合完成**，执行体管控要保证业务的稳定，不能采用“一刀切”方式，而是细粒度控制，**将执行体可执行动作收敛到最小集合**，以此即保证业务可运行也降低执行体运行风险。

用户自定义规则>系统预设规则

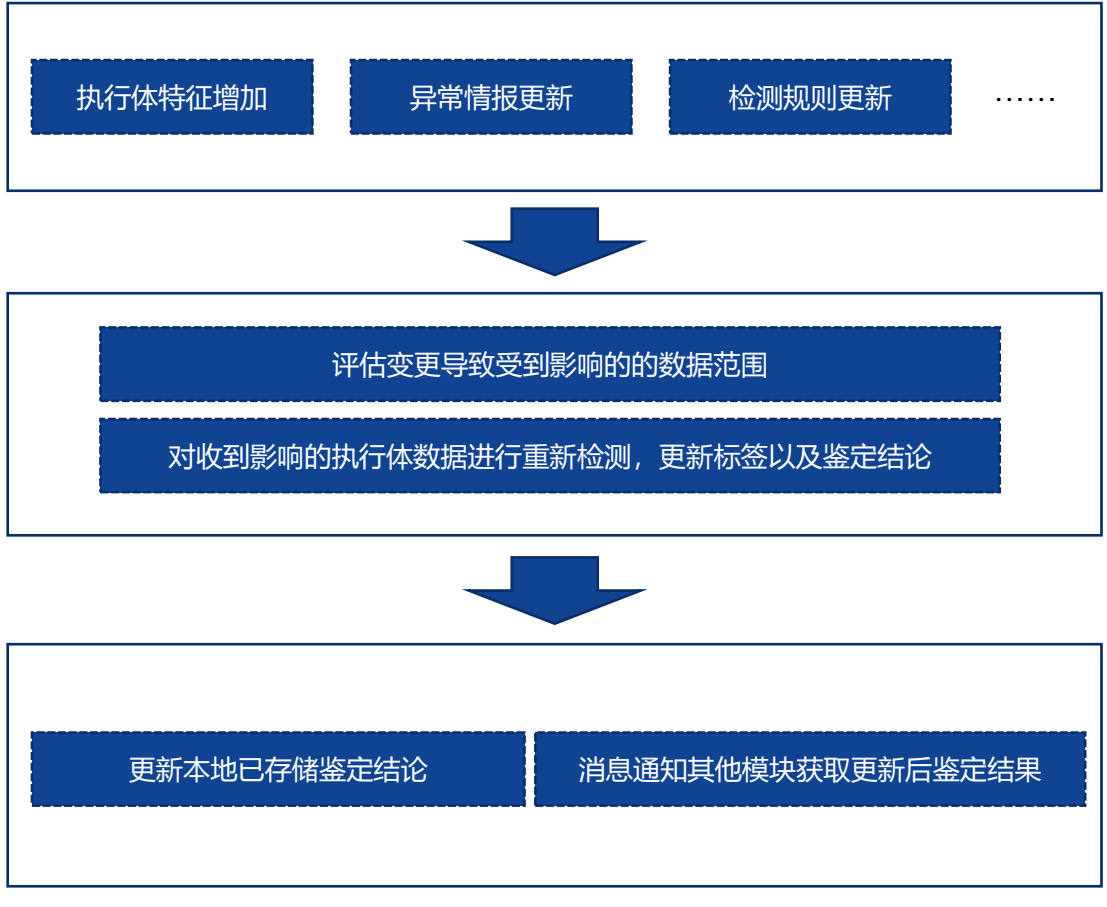
类型	管控能力
执行	1.路径/本地/远程执行权限控制 2.可执行体时间区间控制 3.父进程/执行方式/执行参数控制
数据访问	1.可访问数据范围控制 2.禁止访问数据范围控制
环境变更	1.注册表操作权限控制 2.启动项操作权限控制 3.系统服务操作权限控制 4.计划任务操作权限控制
网络访问	1.网络访问权限控制 2.可访问/禁止访问网络控制 3.端口使用权限控制
存储	存储权限控制
操作账户	可操作账户控制

基于全生命周期监控与动态化信誉评估确保执行体识别准确

执行体信誉评估并不一定能在第一时间就准确结果，**随着获取到执行体信息逐渐丰富，执行体画像逐渐完整才可以获得更加准确的评估结果。**

异常情报的扩展、检测规则变更、用户执行环境的调整都有**可能对已有鉴定结论产生影响**。

发生的安全事件、曝光的安全漏洞等也可能**导致原本安全性相对高的执行体存在降级情况。**



04

实际应用和持续化运营

某科研研发机构全网二进制执行体治理情况

终端总数：**704**　　- Windows终端：**514**　　– Linux终端：**132**　　- 国产终端：**58**

统计周期：2022.07.01-2022.12.01

——执行体静态属性分类统计——

基于HASH已去重

执行体总数	实体大小	系统文件	应用软件（归类）	有效数字签名	无效数字签名	编译器/壳种类	执行态格式	名称伪装
479万	3.7TB	86万	1950个	134W	14W	1239种	49种	31个

生僻技术栈	生僻编译器	生僻壳	生僻厂商	生僻目录	隐藏文件	高CPU占用	高内存占用	高网络占用
121个	24个	14个	19个	23个	4322个	58个	28个	26个

——执行体活跃统计——

已去重

活跃总量	日活（平均）	日增（平均）	工作时间	非工作时间	人工执行	自动执行	被动执行	被动加载
13万	0.3万	0.1万	12万	1万	10万	0.7万	2万	0

——执行体信誉评估统计——

已去重

属性信誉规则条数	行为信誉规则条数	产出信誉标签总数	规则命中执行体总数	自动化信誉评估占比	运营支撑信誉占比
546条	60条	260个	427万	93%	7%

某科研研发机构全网二进制执行体治理情况-2

终端总数：**704**　　- Windows终端：**514**　　– Linux终端：**132**　　- 国产终端：**58**

统计周期：2022.07.01-2022.12.01

——执行体鉴定统计——

已去重

安全文件	1级受限文件	2级受限文件	3级受限文件	异常文件
184万	20万	41万	40万	142万

——执行体治理管控统计——

已去重

拦截受限文件操作	禁止非必要端口监听	禁止外连访问请求	清理恶意病毒文件	修复风险主机
12万次	52个	3860条	653个	143台
禁止非必要启动项	**禁止非必要服务**	**发现异常日志主机**		
602个	211个	210台		

- 说明 -

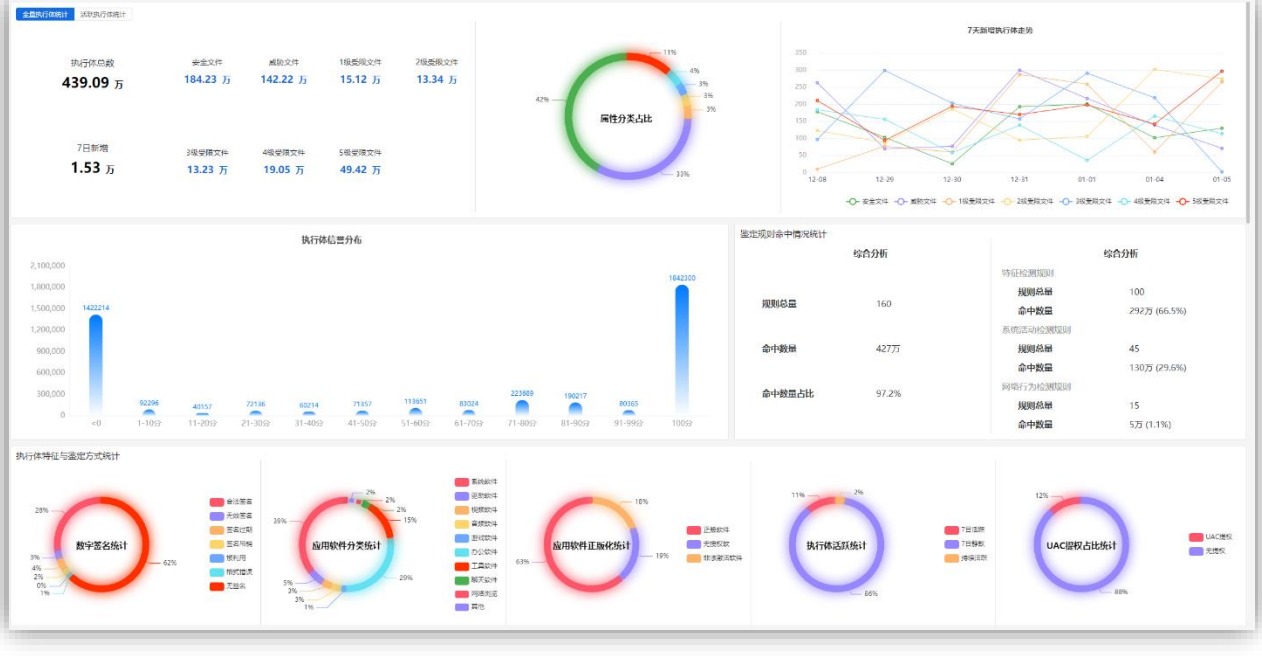
执行体安全属性排序：安全文件>1级受限文件>2级受限文件>3级受限文件>异常文件

可基于执行体信誉评分以及人工设置等方式动态设置/调整执行体安全属性

行为约束方式：基于执行体安全属性约束程序的启动、配置修改、文件操作、进程操作、网络行为……

安天持续赋能运营支撑客户侧执行体采集和自动化识别能力

IT安全治理的同时，为安全管理员最大化减少无效告警和展示数据，聚焦高价值信息，减少运维操作。



存在活跃风险事件终端TOP10

终端名称	数量
DESKTOP-3MHFK2F	12098
DESKTOP-6Q0QLES	9821
DESKTOP-3MHFK2F	7892
DESKTOP-8UFUNV9	6569
zsy	4900
qgj	4021
DELL7080-WHITE	3481
DESKTOP-NCJRLQH	1983

违规端口占比终端TOP10

终端名称	IP	数量
Admin-PC	[redacted]	9
SuperZ	[redacted]	9
TOPDESK-WIN10	[redacted]	6
L	[redacted]	5
DESKTOP-8LOA58O	[redacted]	4
LAPTO	[redacted]	4
MIRACLE	[redacted]	4
Ago	[redacted]	3

违规内网越界访问终端TOP10

终端名称	数量
DESKTOP-3MHFK2F	34
WIN-CLPU8LMQNIK	31
LAPTOP-0LHIEJ5O	29
DESKTOP-704A4S0	24
LAPTOP-KUT97GJ3	20
YHY	16
WIN-H63O7TMCTF3	14
MIRACLE	9

存在隐藏文件终端TOP10

终端名称	IP	数量
SWORDBREAKER	[redacted]	20973
VENUS	[redacted]	19827
TOPDESK-WIN10	[redacted]	18736
DESKTOP-KS5O5FM	[redacted]	15739
DESKTOP-8LOA58O	[redacted]	12034
LAPTOP-QAKPJJKD	[redacted]	10983
MIRACLE	[redacted]	9823
DESKTOP-P6PPDG8	[redacted]	7016

安天持续赋能运营支撑客户侧执行体采集和自动化识别能力

违规端口占比终端TOP10

终端名称	IP	数量
Admin-PC	[illegible]	9
SuperZ	[illegible]	9
TOPDESK-WIN10	[illegible]	6
L	[illegible]	5
DESKTOP-8LOA58O	[illegible]	4
LAPTO	[illegible]	4
MIRACLE	[illegible]	4
Ago	[illegible]	3

- 应用软件
- 文件
- 占用端口
- 当前端口链接状态
- 协议
- 链接溯源
  - 端口开启时间周期
  - 远端主机链接时间
    远端主机IP
  - 远端主机链接时间
    远端主机IP
- 违规原因

用户操作

合规性操作

管控策略

细粒度访问域控制

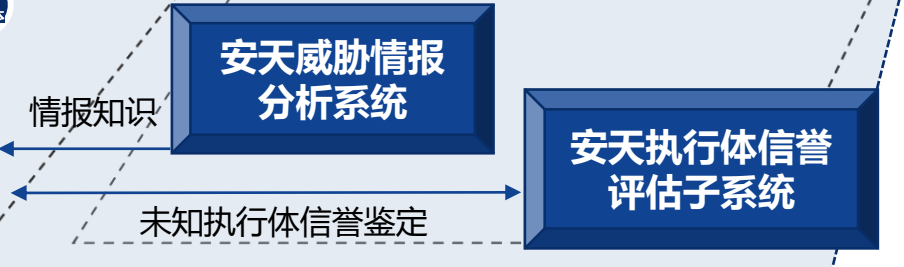
断开

安天持续赋能运营支撑客户侧执行体治理能力

用户侧

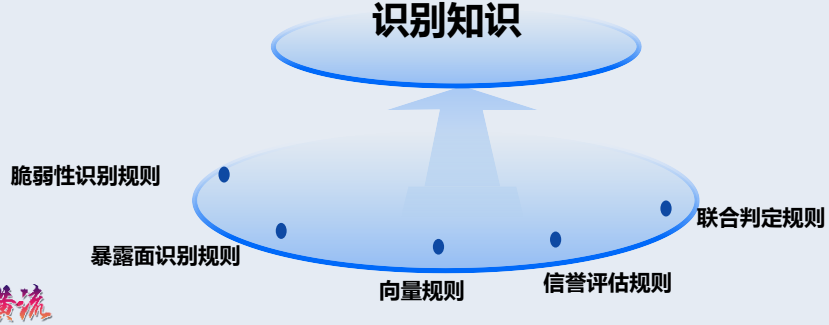
产品自主运行

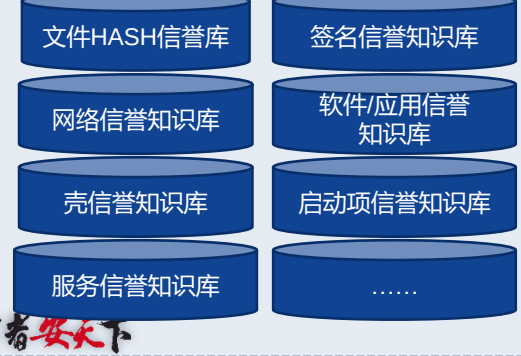
情报动态联动赋能运营



安天

安天持续为客户提供算力、知识和服务方面的运营支撑





感谢大家的关注

安天冬训营 wtc.antiy.cn