威胁狩猎 101 文档

在 Kill Chain 攻击框架发布了近 10 年后，ATT&CK 框架做为后继者极大丰富了攻击分析和场景，包含了黑客渗透过程中利用具体的各种技术。在这么多攻击技术和手段的攻击下，传统的安全设备堆叠已经失守。比如各种 Webshell 的混淆、加密流量、社会工程对于终端的渗透，这些技术基本都可以穿透所有的传统安全产品下堆叠出的安全架构和系统。

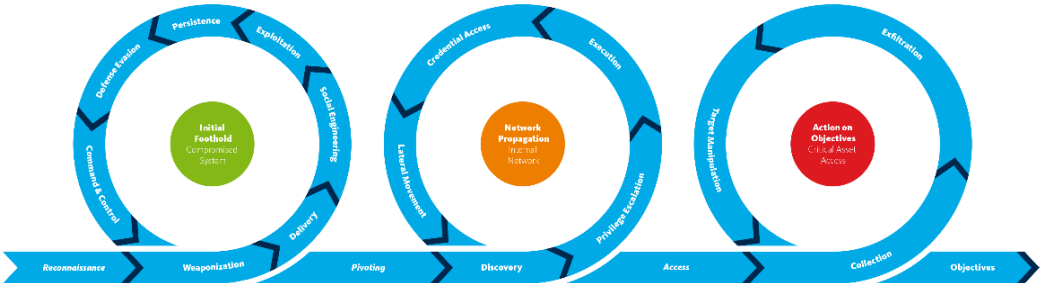

图 1：ATT&CK 和 Kill Chain 的融合图

在 FireEye 的 M-Trends 2020 Reports 中，发现攻击者隐藏或者驻留时间的中位数为 56 天。近几年的威胁检测时间都在不断缩短，主要是由于对于内部威胁的情况发现较早，极大的减少了中位数，但是外部威胁的驻留时间还在 141 天，接近 5 个月之久。

GLOBAL MEDIAN DWELL TIME BY YEAR

Compromise Notifications	2011	2012	2013	2014	2015	2016	2017	2018	2019
All	416	243	229	205	146	99	101	78	56
Internal Detection	–	–	–	–	56	80	57.5	50.5	30
External Notification	–	–	–	–	320	107	186	184	141

图 2：全球驻留时间中位数（按年份划分）

高级威胁的存在背后是拥有高级黑客技术的人和相关的动机。对于攻击行为的动机可以分为随机的、自动化的、报复性的、经济目的、政治目的和军事目的，威胁的等级也不同。攻击者组织方式无论从时间精力和和武器库的丰富度来说，对于防御方来说都是极大的不平衡。缺少高级的攻击防御经验、没有太多时间精力去保证全面的安全，缺乏相关高级的技术手段来应对。

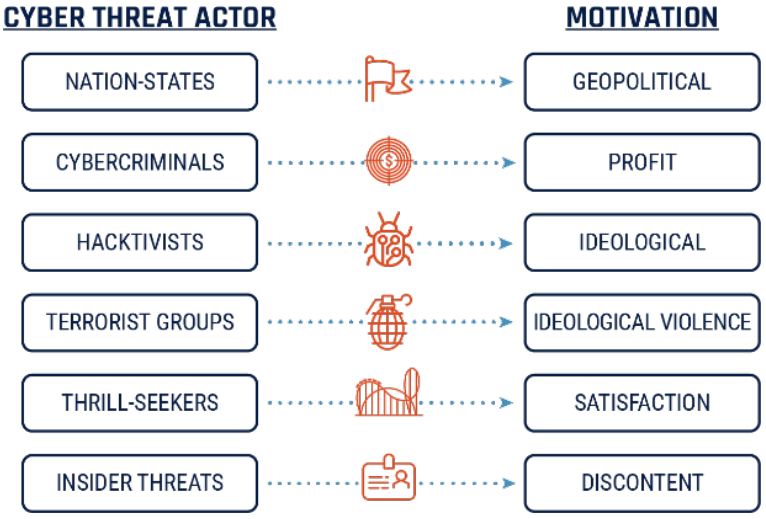

图 3：高级威胁与动机

根据 Sans 网络安全滑动标尺模型有五个阶段，架构安全、被动防御安全、主动防御安全、威胁情报和反制安全。在整个安全建设的过程中，目前整体都是在架构安全和被动防御安全这两个方面努力，而在主动防御方面，投入的技术、人力和产品还严重不足。要提高整个安全态势向更高层面的提升，必须要重视主动防御安全的建设。

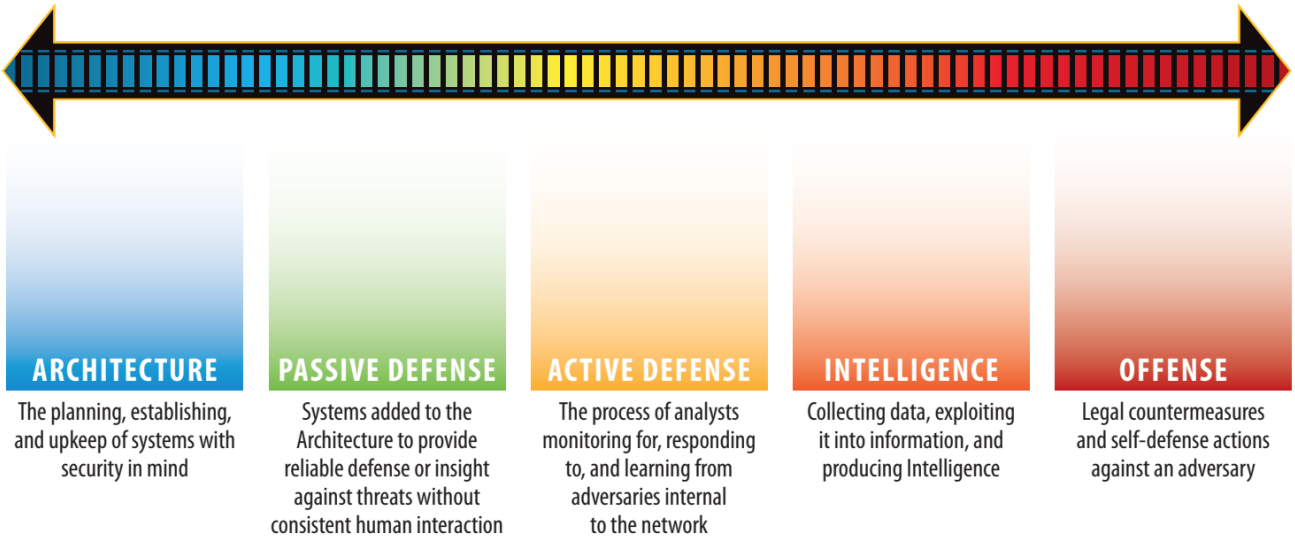

图 4：Sans 网络安全滑动标尺模型

针对以上四种情况：1. 攻击手段多样性，2. 攻击者驻留时间长，3. 高级威胁检测难度高，4. 安全建设的进一步要求，网络威胁狩猎（Cyber Threat Hunting）应运而生。威胁狩猎是主动安全的代表性技术，依赖于相关技术手段和人的知识，利用威胁狩猎可以减少我们目前的威胁。威胁狩猎的定义：威胁狩猎是一个高级安全功能，集成了主动防御、创新技术、技术专家以及深度威胁情报来发现和阻止恶意的并且极难检测的攻击行为。同时，这些攻击行为也是传统自动化的防御无法检测出来的。

一、威胁狩猎详解

威胁狩猎、SOC 和事件响应的关系

威胁狩猎与 SOC 运营中心以及 IR 事件响应的关系如下图所示：

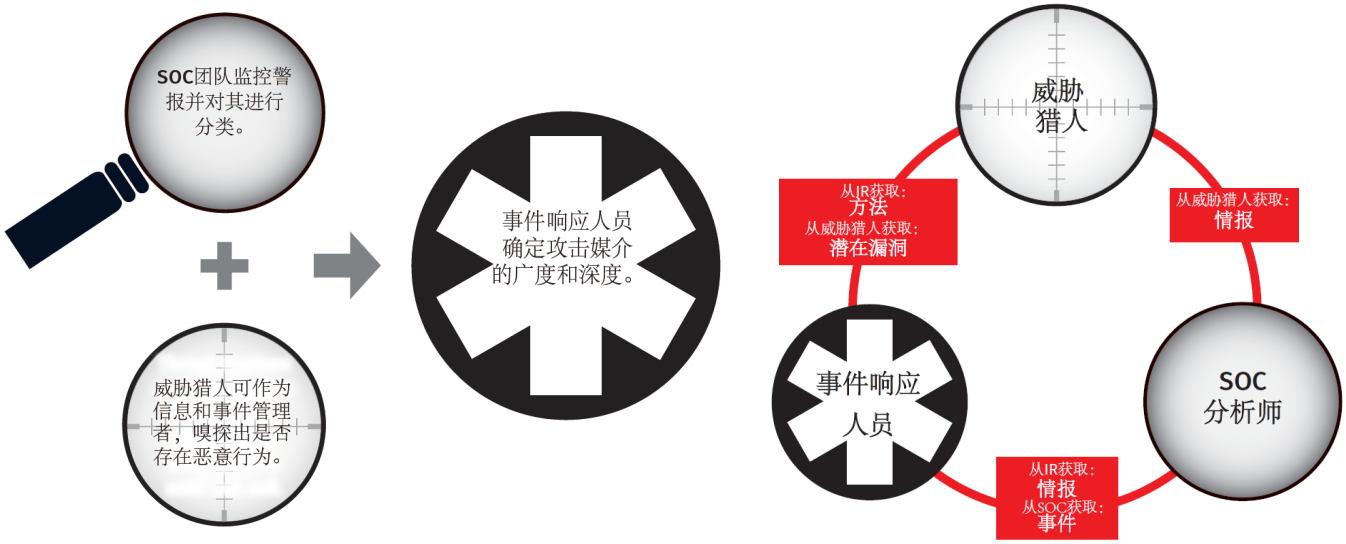

图 5：威胁狩猎、SOC 和事件响应的关系

SOC 团队主要是运营维护日常的安全设备和 SIEM 报警以及如何分类处置这些事件；威胁狩猎团队主要是基于一些数据和征兆进行分析安全事件，而不是直接的安全报警；事件响应团队根据这两个团队提供的信息进行相关的动作以及后面如何处理、取证、恢复等。

三者的联系在于，威胁狩猎将安全情报输送给 SOC 团队，并且从事件响应团队获取狩猎方法论。SOC 团队将安全事件发送给事件响应团队，并从该团队获取情报。事件响应团队接受来自威胁狩猎和 SOC 团队的事件或者潜在的渗透行为。

在 2019 年的 Threat Hunting 报告中，受访者关于 SOC 有一些问题是目前很难解决的，比如一些高级威胁、处理误报时间、以及缺乏专家型人才、响应时间太慢等等。

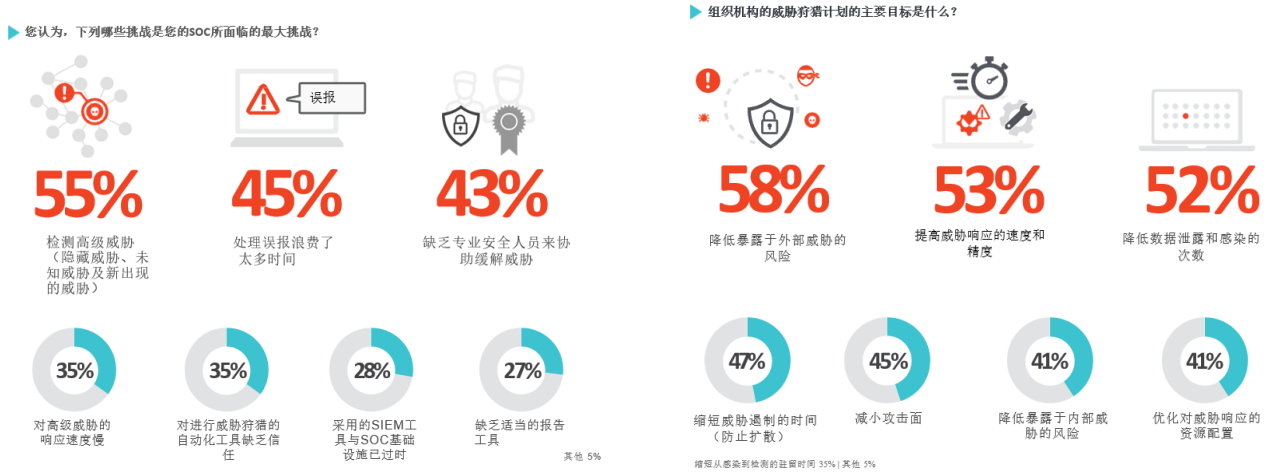

图 6：威胁狩猎中面临的挑战与主要目标

针对这种情况，威胁狩猎的主要的目的是减少外部威胁暴露面、提高威胁响应的速度和准确性，减少入侵的数量以及响应时间。

威胁狩猎的过程

威胁狩猎是一个持续的过程，也是一个闭环。基本都是基于假设作为狩猎的起点，发现 IT 资产中的一些异常情况，就一些可能事件提前做一些安全假设。然后借助工具和相关技术展开调查，调查结束后可能发现新的攻击方式和手段（TTP），然后增加到分析平台或者以情报的形式输入到 SIEM 中，可能触发后续的事件响应，从而完成一次闭环。

图 7：威胁狩猎的过程

威胁狩猎的方式

威胁狩猎过程的起点是假设，但是这种假设有三种假设来源，也是狩猎的方式：

◆ 基于分析的方式：分析分为两种，基本数据分析以及机器学习的 UEBA 的高级分析方式。

◆ 基于重点的方式：皇冠珍珠分析法，基于 IT 资产中比较重要的资产进行重点关注。

◆ 基于情报的方式：根据威胁情报提供的内容，进行威胁狩猎。

威胁狩猎的成熟度

威胁狩猎也有成熟度评价，可以根据自身的安全建设情况进行相关的成熟度规划。主要有两个维度进行评价：分析水平和数据收集水平。从低到高主要分为：

◆ Level 0：基本的自动化报警但没有数据收集；

◆ Level 1：有一定的威胁情报处理能力和一定的数据收集能力；

◆ Level 2：遵循数据分析的流程和较高级别的日常数据采集；

◆ Level 3：有一些新的分析流程和高级别的日常数据收集；

◆ Level 4：自动化大部分的分析过程和高级别的日常数据收集。

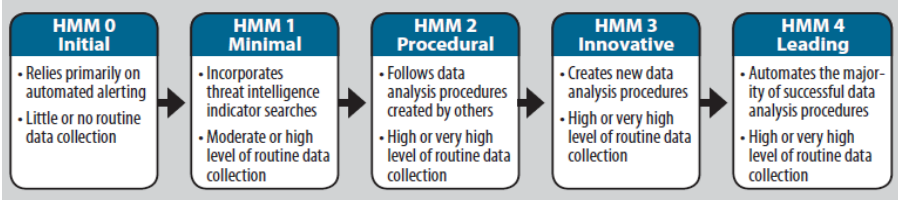

图 8：威胁狩猎的成熟度

更全面的成熟度加入了假设的来源、使用的工具以及对于威胁情报的使用水平。

威胁狩猎成熟度水平

威胁狩猎循环步骤	HM0 起步	HM1 基础	HM2 流程化	HM3 创新	HM4 领先
数据收集	很少收集或不收集数据	从IT环境中的一些关键点，收集一定量的、某些类型的数据	大量收集整个IT环境中特定类型的数据	大量收集整个IT环境中特定类型的数据	大量收集整个IT环境中多种类型的数据
制定假设	对来自SIEM、IDS、防火墙等来源的现有自动化警报做出响应	审查情报信息，制定新的假设	审查威胁情报和其他有用情报，制定新的假设	审查威胁情报、其他有用情报及手动进行的网络风险评分（例如，“皇冠珍珠分析”），制定新的假设	审查威胁情报、其他有用情报和自动执行的网络风险评分，制定新的假设
确定用于验证假设的工具和技术	警报中台、SIEM搜索；但不进行主动调查	利用SIEM或日志分析工具，通过全文查询或类SQL查询，进行基本的搜索	根据现有的狩猎流程，利用简单的工具和统计图来检索和分析数据	充分利用可视化和图表搜索；制定新的狩猎流程	高级可视化和图表搜索。发布并自动执行新的狩猎流程
类型和TTP检测	无；只有SIEM/IDS警报	识别PoP底层的IoC，例如域名、URL和哈希值	识别PoP底层和中间的IOC，以及这些IOC随时间变化发生的映射趋势	能够检测出PoP顶层的对抗TTP和其他IOC	自动执行发现的复杂TTP和进行追踪；与信息共享组织机构共享发现的IOC
分析自动化	无	集成威胁情报信息，进行自动化报警，实现基本匹配	构建一个威胁狩猎有效流程库，并按定期执行这些流程	构建一个威胁狩猎有效流程库，并经常执行这些流程；基础的数据科学（标准差、异常值检测）	持续执行有效的自动化狩猎流程，提高警报能力；高级数据科学（机器学习）

图 9：威胁狩猎成熟度的全面分析

二、威胁狩猎的开展

SANS 2019 年的威胁狩猎调查报告显示，对于威胁狩猎的预算情况排序，大部分的预算会放在技术和产品的采购上，其次是在员工的招聘上，需要有新员工补充，排在后面的是培训和服务。

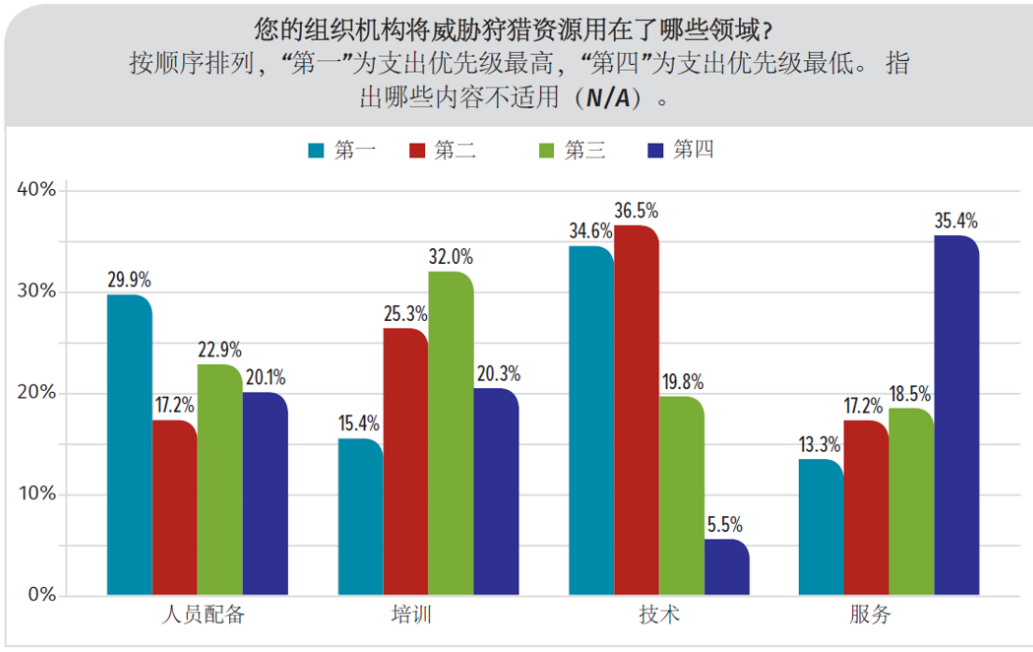

图 10：威胁狩猎预算分配情况

对于威胁狩猎人员技能方面，其中 75% 的受访者认为威胁狩猎团队需要具备网络知识、事件响应、威胁情报分析以及终端相关知识等等。

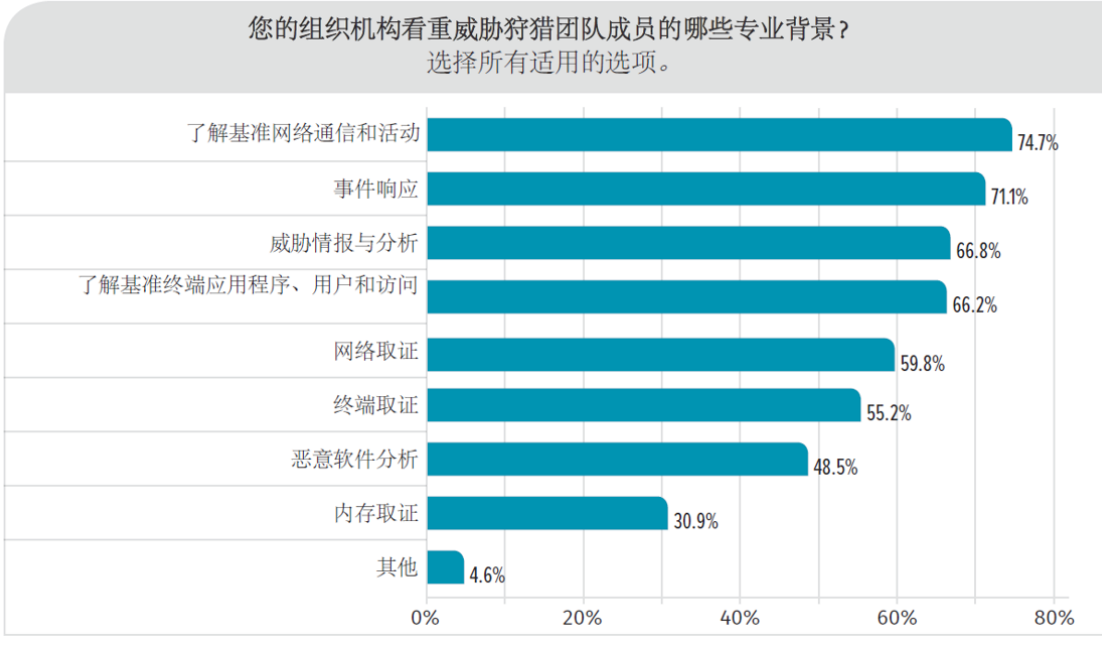

图 11：威胁狩猎团队成员应具备的技能

对于工具和数据收集的方面，SIEM 报警、终端事件数据、IPDS 数据、威胁情报、终端日志数据是排名靠前的几类数据来源。

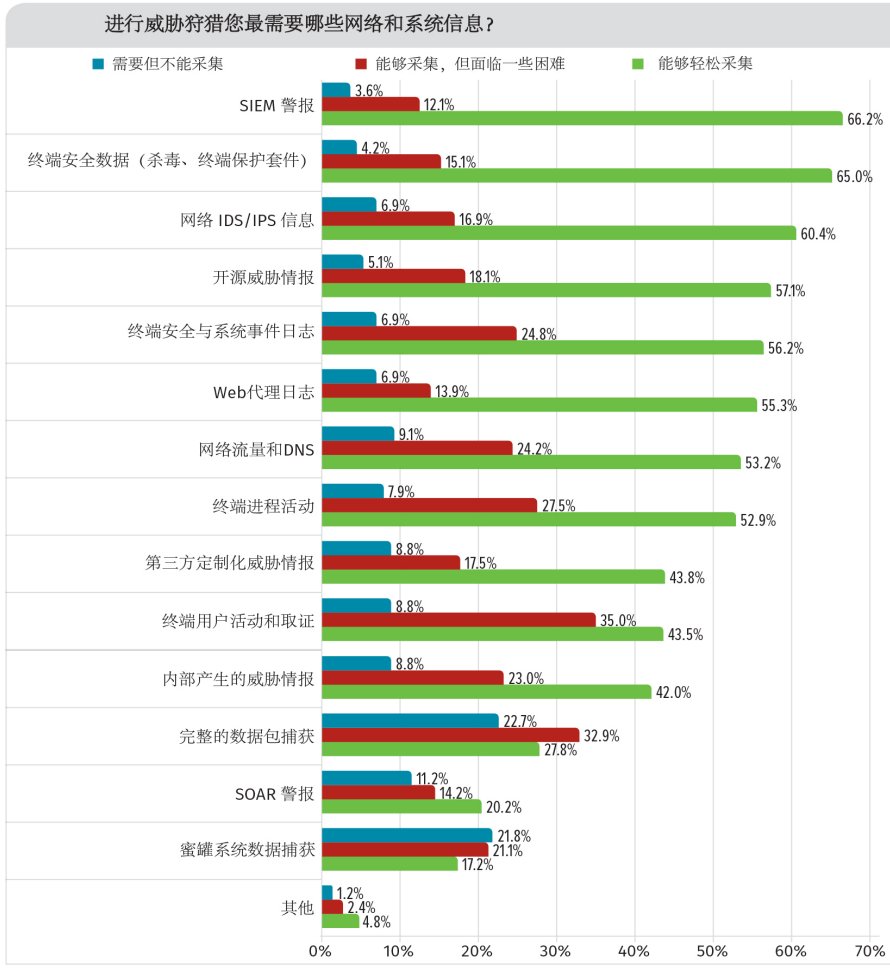

图 12：威胁狩猎的数据来源

开展威胁狩猎活动需要考虑三点：人员、流程和技术。

其中，对于人员的规划，需要考虑招聘、培训以及服务外包三种方式。对于未来预期有长期需求但内部员工短时间内很难习得的技能，采用招聘的方式来引入能力；对于现有员工的能力提升，需要加强培训增强员工对某些知识的理解；对于一些高级技能，若招聘成本太高，同时培训周期太长时，可以采用外包服务的形式来解决临时需求。威胁狩猎的团队组织架构如下：

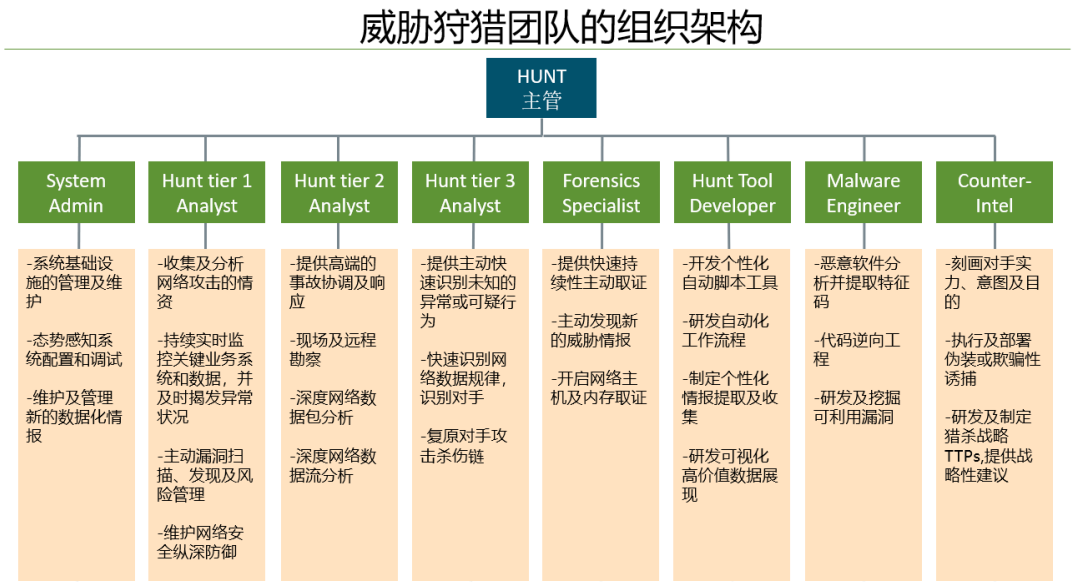

图 13：威胁狩猎团队的组织架构

威胁狩猎团队的人员组织，需要 7 种角色，有些角色可以合并为一个人，不一定是 7 个角色 7 个人。第一个角色是系统管理员，主要针对 SIEM 系统的维护以及威胁狩猎平台的管理；狩猎初级分析师可以使用 SIEM 系统和威胁狩猎平台，处理报警和一些基本平台使用。狩猎中级分析师具有对威胁情报、日志的分析能力，同时也具有渗透测试和网络协议的知识。狩猎高级分析师具有风险等级评估、漏洞管理、网络包和日志的深度分析能力、以及恶意软件分析能力。取证专家对于内存、硬盘要有专业的取证知识，可以做时间链分析。狩猎工具开发人员要具备开发经验，可以自动化一些狩猎场景。恶意软件分析工程师，主要负责恶意软件的逆向，熟悉汇编语言等内容。安全情报人员，具有情报资深经验，能够筛选、使用、开发威胁情报。

上文从流程方面介绍了威胁狩猎的具体过程，下面从管理角度描述一下整个威胁狩猎的过程，一共分为六个步骤：目的确认、范围确认、技术准备、计划评审、执行、反馈。

图 14：威胁狩猎的六个步骤

对于威胁狩猎的目的确认，必须要描述清楚相关的目的和预期达到的结果，跟前文中解决的那四个问题相关。范围确认阶段主要是为了确认要达到预期结果，需要开发的威胁狩猎的假设用例。技术准备阶段要确认，在基于假设用例的情况下，需要采集哪些数据和哪些技术和产品。计划评审是对范围确认和技术准备的内容进行评审，确认其是否能真正能满足目的。接下来就是执行阶段，主要是看实际效果。最后进行复盘来检查每项活动中的一些不足，进行持续改进。

比较重要的是两个阶段：范围确认和技术准备。范围确认首先要进行测试系统的选择。对于测试系统，要确认需要哪些数据和技术手段来进行威胁狩猎。其次，假设用例的开发尤为关键。假设用例作为威胁狩猎的核心，是威胁狩猎分析的起点，来源于对数据的一些基本分析和高级分析，威胁情报的使用和收集以及对 TTP 的理解，甚至是一些核心能力的使用，比如使用搜索的分析能力。

威胁狩猎技术方面包含三项准备工作：数据收集、产品技术选型和威胁情报的使用。关于数据收集，要利用数据收集管理架构 CMF（Collection Management Framework），来评估收集的数据。可以根据以下几个维度进行考虑：位置、数据类型、KillChain 阶段、收集方法和存储时间。当然也可以参考更细、更有针对性的 ATT&CK 的 TTP 收集粒度，DeTT&CT 项目就是可以看出数据收集的范围、质量和丰富度。总体来说，数据收集的内容主要有三种——终端类型数据、包数据和日志数据。在每类数据中，要按照要求的格式和接口提供相关数据。收集形式主要有拉和推两种方式，即主动拉取数据和推送数据。

	Source	Source	Source	Source	Source	Source	
Location							
Data Type							
Kill Chain Step							
Collection Method							
Storage Duration							

图 15：数据评估的考虑因素

在产品技术选型中，核心产品要考虑两种平台型产品：一种是围绕 SIEM 产品展开的威胁狩猎内容，要有威胁狩猎的模块，另外一种威胁狩猎的平台型产品来实现这个功能，当然也可以考虑这两个产品进行联动。以 SIEM 为核心进行威胁狩猎平台的对接，其他类型的安全产品的数据接口要能开放，并能对接 SIEM 产品。作为威胁捕猎的核心产品或者模块要有以下几种能力：安全大数据的分析能力、查询搜索能力和威胁情报处理能力。分析能力作为核心能力，不仅仅在于基本的筛选、分类以及排序，还需要高级的分析能力，比如 UEBA 的能力。根据机器学习算法来进行建模分析，来定位一些异常行为，极大地降低了分析的难度。查询搜索能力是维持威胁狩猎日常运营的能力，一些疑似的攻击行为可以通过查询进行定位并可以进行深度定点分析。查询搜索也是实现 ATT&CK 场景的基础，很多场景的检测可以通过查询搜索能力来完成对于 ATT&CK 场景的覆盖。最后一个方面就是威胁情报的使用，威胁情报的识别和使用是威胁狩猎平台比较常见的功能，可以根据痛苦金字塔的威胁情报的使用，可以先从简单的文件 hash 和恶意 IP 开始使用，然后逐步加强对威胁情报的使用能力，到最后的 TTP，乃至自己生产威胁情报。

三、威胁狩猎解决方案

关于威胁狩猎解决方案，从产品到运营落地，重点介绍三类产品和服务：SIEM 类产品、终端类产品和 MDR 服务。

SIEM 类产品

SIEM 类产品是安全品类中的集大成者，也是威胁狩猎的核心。参照 2020 年 SIEM 魔力象限的领导者象限，这些厂商都以 Threat Hunting 作为 SIEM 的主要功能之一，作为下一代 SIEM 的一项主要功能。

图 16：SIEM 魔力象限的领导者象限

IBM 的 QRadar 作为 SIEM 的主流产品，已经广为大家所知，i2 是威胁狩猎的产品，X-Force 是将威胁情报产品作为整体解决方案。IBM i2 核心提供了多种可视化分析方法，主要包括可视化查询 (Visual Search)、链接分析（Find Linked）、路径分析（Find Path）、群集分析（Find Clusters）、社会网络分析（SNA）等分析算法与分析工具。自动布局一直是可视化分析能力的难点和重点，在这块十分出色。通过这些可视化的分析工具来实现最终的威胁狩猎。

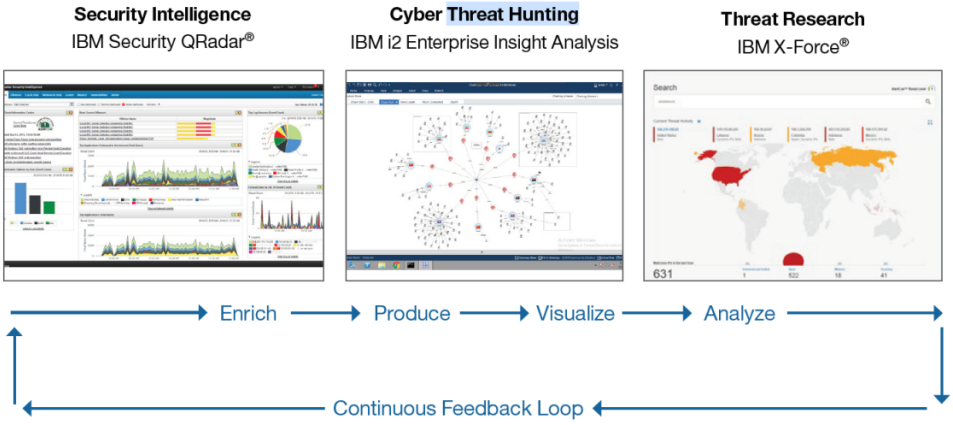

图 17：IBM 威胁狩猎解决方案

Splunk 的 Threat Hunting 能力主要是通过其强大的 SPL 语言实现的。作为大数据平台的领导者，安全只是其中一块业务。在其 splunkbase 里面有相关的 app，基本思路就是将 sysmon 采集的数据导入 Splunk 然后进行 ATT&CK 映射，相对比较局限于 windows 系统。

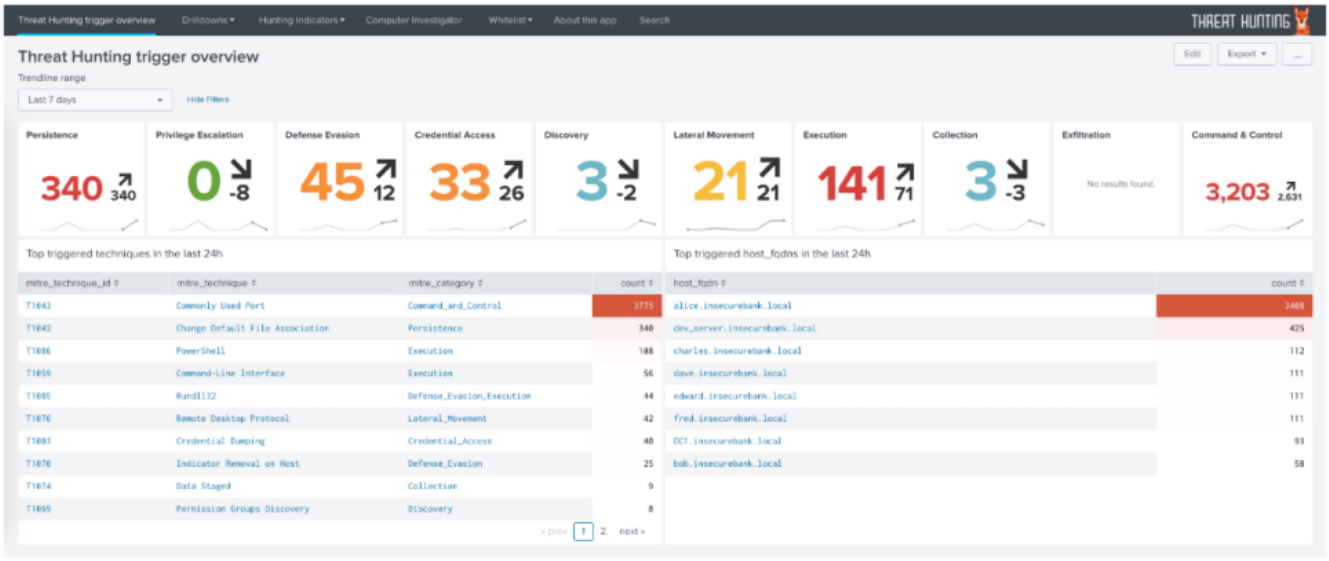

图 18：Splunk 威胁狩猎解决方案

Logrhythm 的威胁狩猎是通过 Threat Hunting Automation app 实现的，可以看出来主要依托于威胁情报进行自动化分析。这种能力相对来说比较简单，可以应付一些场景。

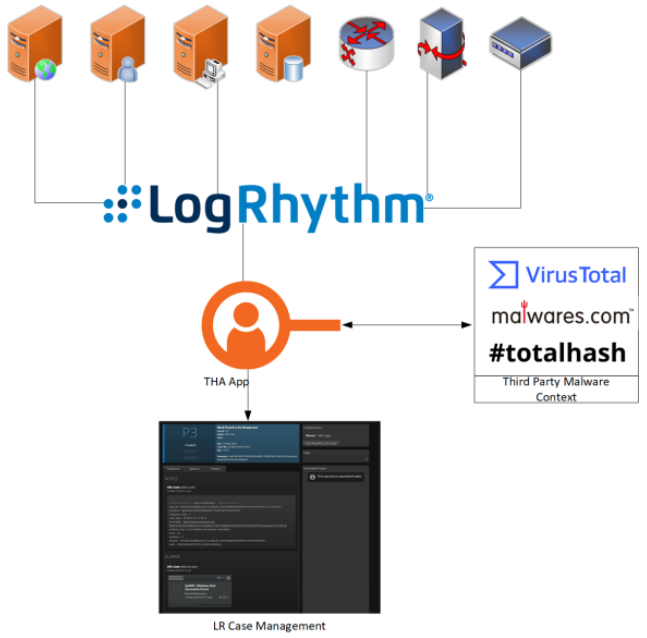

图 19：Logrhythm 威胁狩猎解决方案

Securonix 将搜索和威胁狩猎作为核心能力。威胁狩猎方面可以进行自然语言搜索，可以很快捷地进行搜索来进行威胁假设验证，也可以进行威胁情报的 ioc 的验证，同时可将数据导出并可视化。

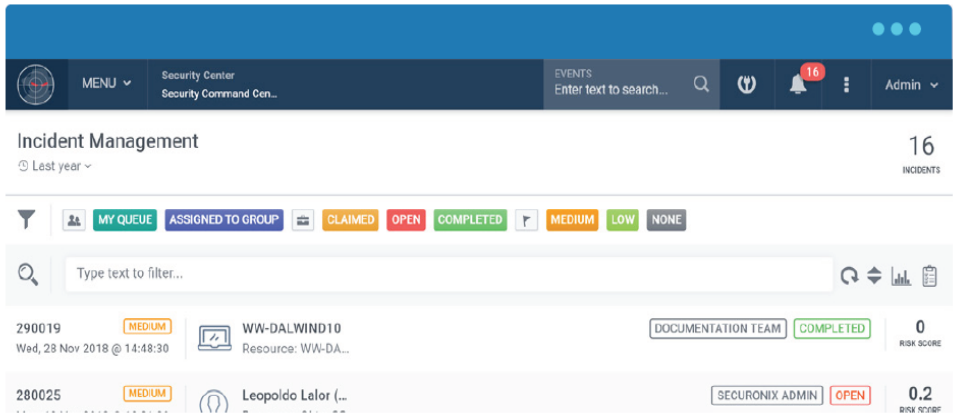

图 20：Securonix 威胁狩猎解决方案

Exabeam 威胁狩猎产品 Threat Hunter 依托的能力主要是搜索、查询、旋转、钻取能力，同时也有威胁情报的使用能力。

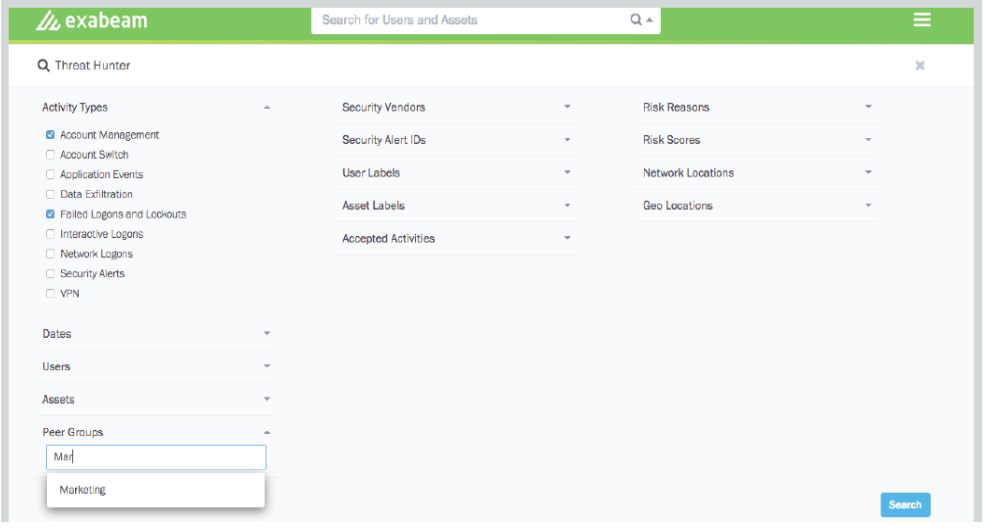

图 21：Exabeam 威胁狩猎解决方案

终端类产品

鉴于终端类产品位置的重要性，且端点有时候也是威胁高发地和最终的落脚点，所以，通过终端安全产品实现威胁狩猎更是事半功倍。

Crowdstrike 的 Falcon OverWatch 是其威胁狩猎模块，核心能力是实时的威胁可见性，也是很重要的一个 feature。其次也有威胁情报的处理能力。

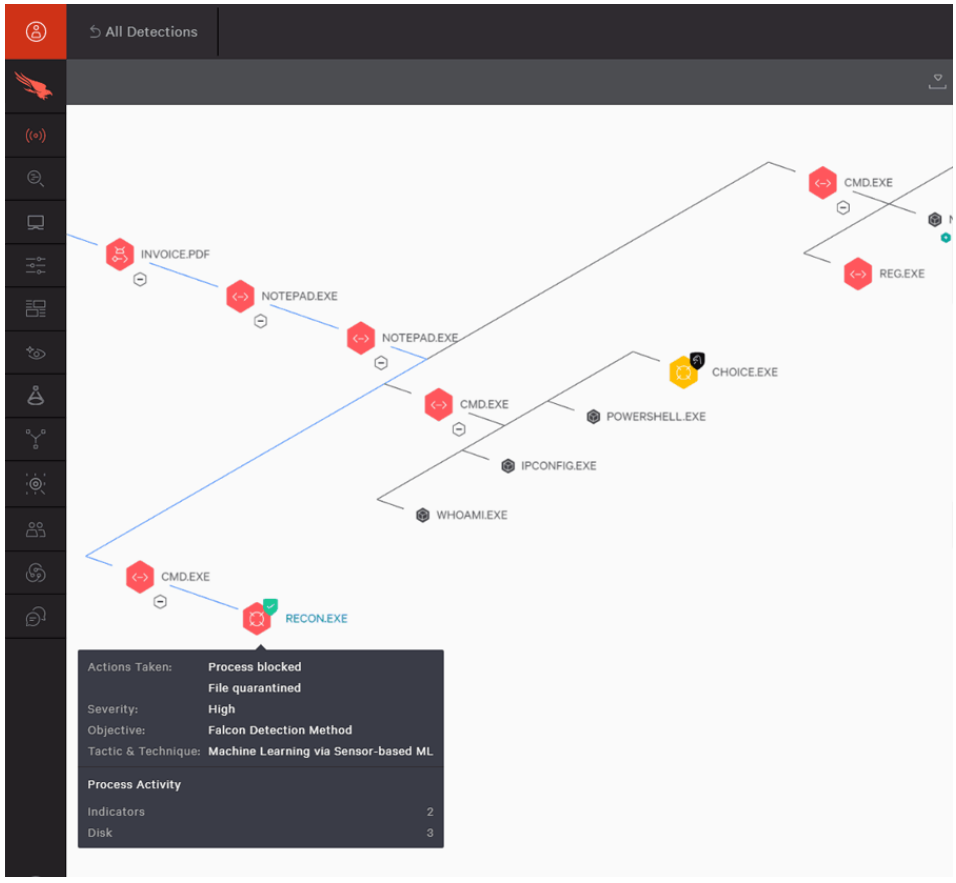

图 22：Crowdstrike 威胁狩猎解决方案

Cylance 的威胁狩猎主要突出了其查询引擎 CylanceOPTICS InstaQuery，可以查询文件、注册表、进程、网络连接等安全信息。

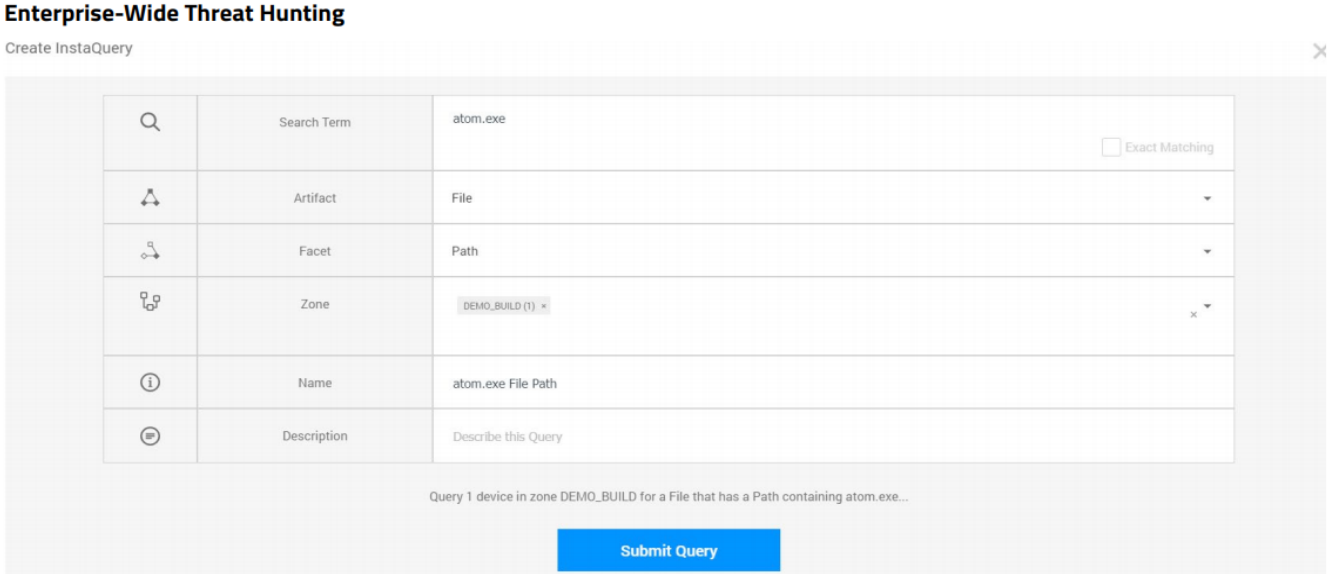

图 23：Cylance 威胁狩猎解决方案

Cybereason 威胁狩猎的能力较为全面，有事件关联分析，有时间轴展示，同时也有搜索能力。对于一个威胁，Cybereason 能把其来龙去脉解释得很清楚，而不是只是简单的报警。

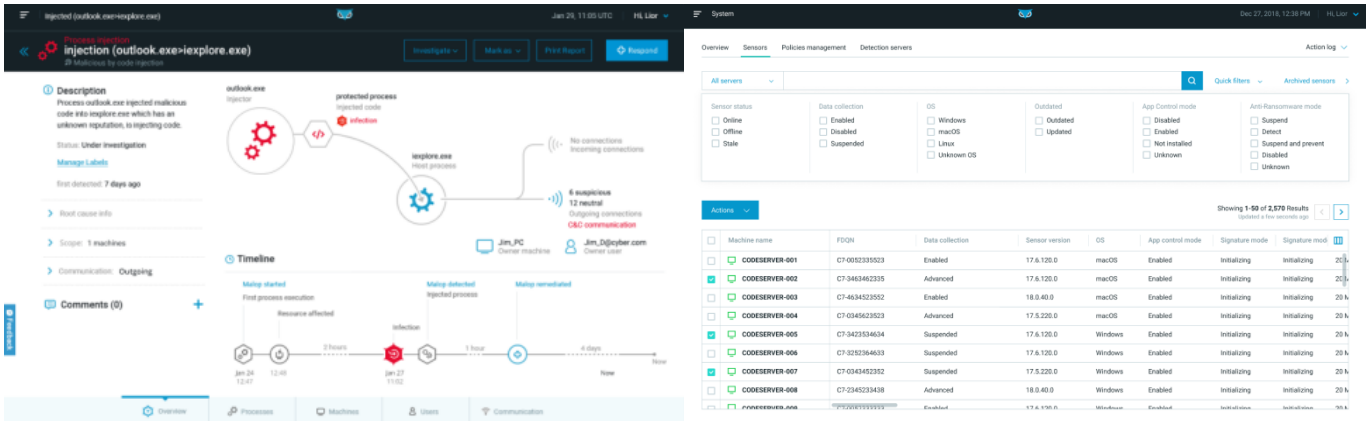

图 24：Cybereason 威胁狩猎解决方案

Carbon Black 也有可视化能力和查询能力，同时也有事件关联分析和威胁情报处理能力。

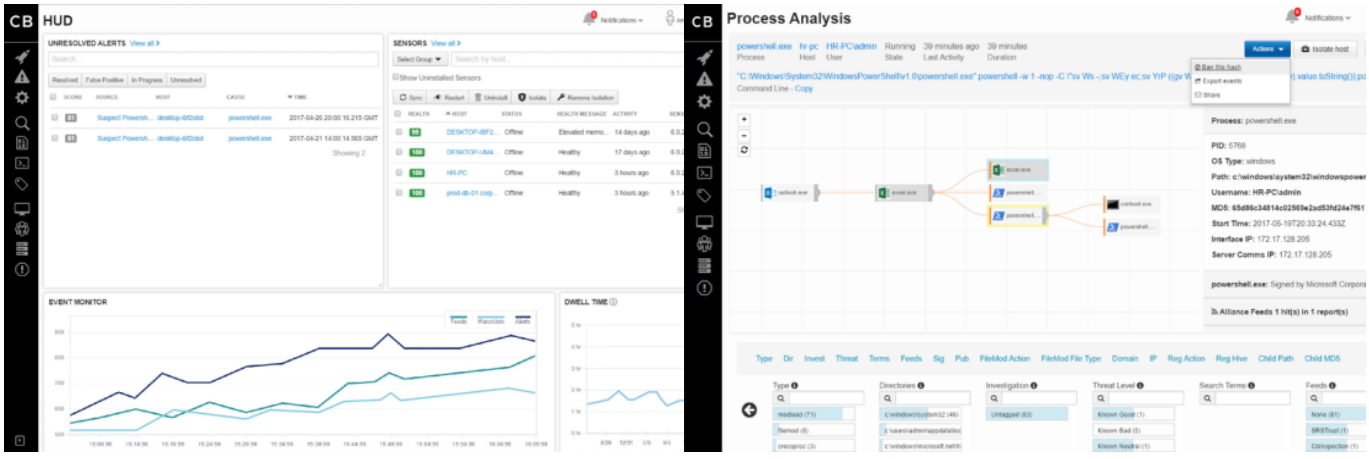

图 25：Carbon Black 威胁狩猎解决方案

Endgame 已经被 Elastic 收购，其开源的 EQL 是一个很好实现威胁狩猎的查询语言，可以进行 ATT&CK 检测场景的实现。同时结合 Elastic 的产品堆栈，能够实现更多的分析场景。

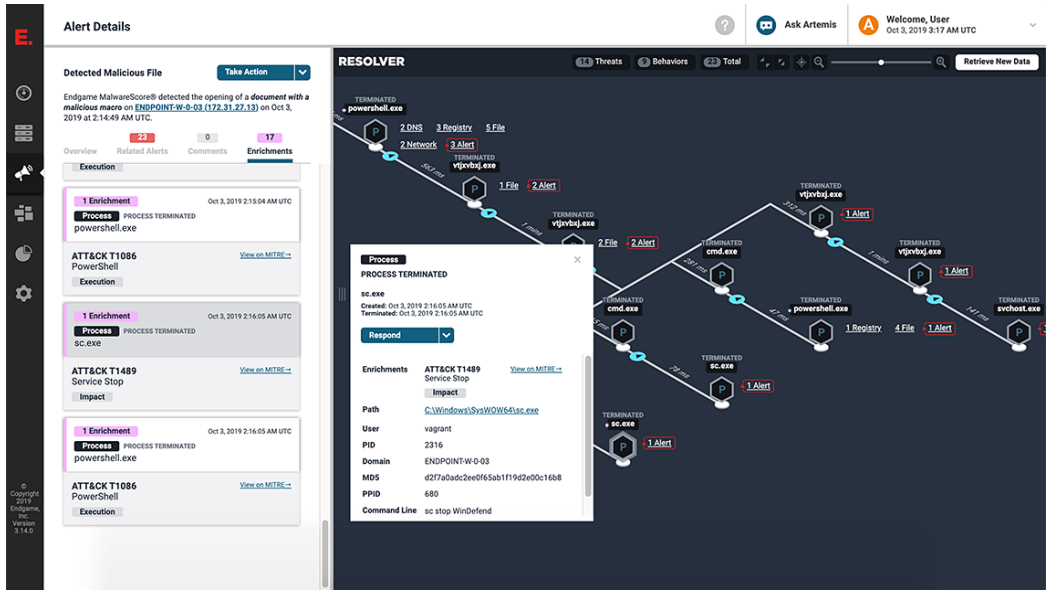

图 26：Endgame 威胁狩猎解决方案

SentinelOne 拥有 True Context ID 专利技术，可以对每个终端进行数据建模。如果某个终端发生异常事件，通过这个技术就可以迅速查询当时现场一些细节信息，包括进程、文件。

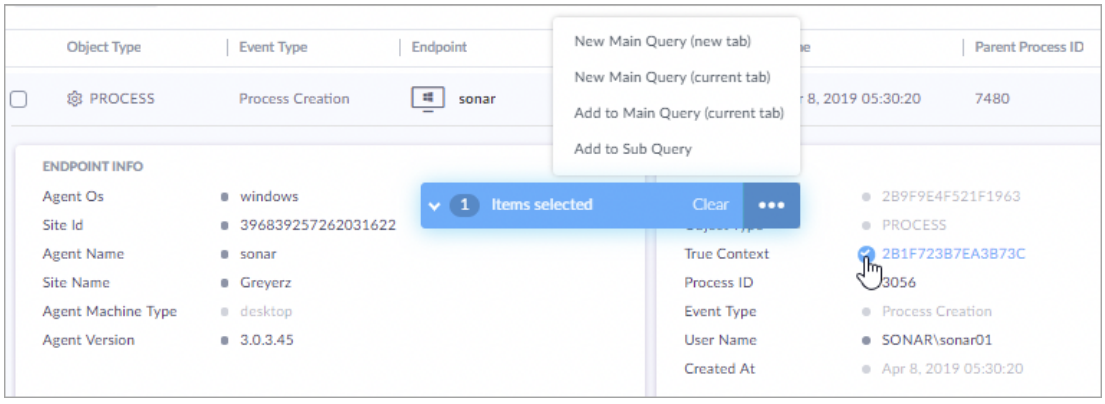

图 27：SentinelOne 威胁狩猎解决方案

MDR 类服务

对于 MDR 类服务，由于甲方可能存在的人员不足或者是技能不足，还需要依赖相关高级的威胁狩猎能力外包。由于能够提供威胁狩猎的上述产品的公司都会推出相关的 MDR 服务，这里就只选取一家专注于 MDR 威胁狩猎服务的公司。

Red Canary 一般会使用 Carbon Black 进行服务，威胁狩猎的内容包括收集终端数据、建立建设、进行狩猎验证、进行用例开发、进行威胁检测、验证威胁。

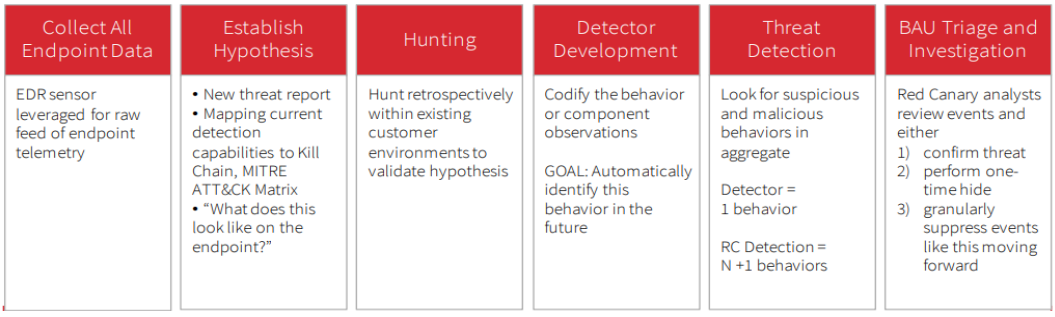

图 28：Red Canary 威胁狩猎解决方案

综上所述，威胁狩猎作为可以减少攻击驻留时间的重要能力，已经得到了业内大部分人的认可和重视。威胁狩猎的方式和成熟度已经进行了定义，可以参照定义进行相关能力建设。开展威胁狩猎需要从人员、流程和技术三个方面进行充分考虑。最后，从能力角度而言，威胁狩猎解决方案要具备三个基本能力：强大的查询能力、分析能力和威胁情报处理能力。也应该考虑 MDR 服务的形态，让威胁狩猎更好地落地。关于威胁狩猎的下个阶段——事件响应（Incident Response），也是 RSAC 2020 的热点议题之一，将在后续的文章中讲解。

写在最后

青藤威胁捕获平台基于 ATT&CK 框架，旨在帮助用户解决安全数据收集、数据挖掘、事件回溯、安全能力整合等各类问题；该平台提供了 100 余类 ATT&CK 攻击场景，用户可以直接对数据进行深度挖掘，发现潜在威胁；该平台还与青藤万相深度集成，提供资产、风险、入侵检测、日志、任务等功能，提供 50 余类原始数据供使用；内嵌 QSL（Qingteng Search Language）数据检索与威胁捕获语法，能对异构数据进行统一查询、统计、分析；同时该语法还集成了多个机器学习与异常行为分析能力，同时允许调用各类外部检测能力，进行补充分析；其将是企业回溯与分析能力的极大补充，充分发挥企业的数据优势，真正捕获常规手段难以发现的威胁，帮助企业用户主动发现威胁，减少威胁驻留时间，全方位提升安全能力！