360 态势感知与安全运营平台

产品技术白皮书

■文档编号	■密级
■版本编号	■日期

1 引言 2
2 产品设计目标 4
2.1 产品价值 5
3 产品关键技术 6
3.1 万兆网络及 IPv4/IPv6 网络环境下数据还原技术 6
3.2 数据处理与计算分析的自动化关联技术 8
3.3 规模化沙箱动态检测技术 9
3.4 基于大数据挖掘的恶意代码智能检测技术 10
3.5 基于搜索引擎技术的大流量行为检索与存储 13
3.6 云端基于大数据的 APT 发现与跟踪技术 13
3.7 可视化技术 16
4 产品组成与架构 17
4.1 产品组成 17
4.2 产品功能架构 19
5 产品功能 20
6 产品部署 22
7 产品优势与特点 23

1 引言

近年来，关于 APT（Advanced Persistent Threats，高级持续性威胁）攻击的报道日益增多，例如：2010 年攻击伊朗核电站的“震网病毒”、针对 Google 邮件服务器的“极光攻击”、2013 年韩国金融和电视媒体网络被大面积入侵而瘫痪、卡巴斯基在 2014 年揭露的 Darkhotel 组织和 2015 曝光的方程式组织（Equation Group）等等。

2016 年初，360 天眼实验室发布了《2015 年中国 APT 研究报告》。报告中指出，截至 2015 年 11 月底，360 威胁情报中心监测到的针对中国境内科研教育、政府机构等组织单位发动 APT 攻击的境内外黑客组织累计 29 个，其中 15 个 APT 组织曾经被国外安全厂商披露过，另外 14 个为 360 威胁情报中心首先发现并监测到的 APT 组织。

中国是 APT 攻击的受害国，国内多个省、市受到不同程度的影响，其中北京、广东是重灾区，行业上教育科研、政府机构是 APT 攻击的重点关注领域。

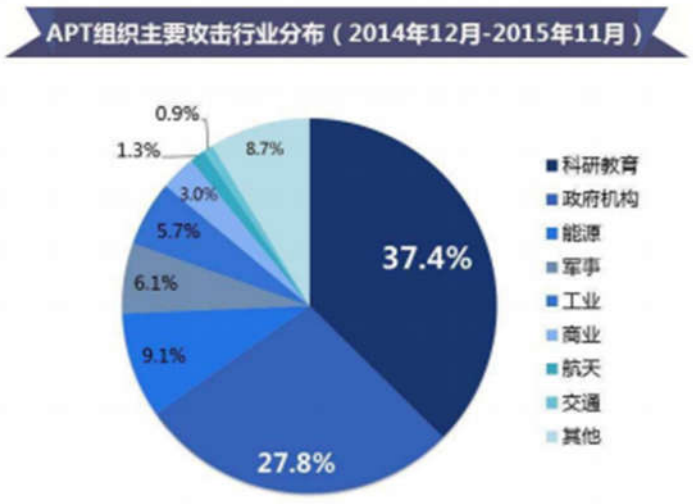

图 1 APT 组织主要攻击行业分布

根据调研，我们发现这些 APT 攻击的受害者中几乎都是具备一定规模的企事业单位，而且都已经部署了大量的安全设备或系统，也有明确的安全管理规范和制度。既然已经有了防御措施，为什么仍然会有部分威胁能绕过所有防护直达企业内部，对重要数据资产造成泄漏、损坏或篡改等严重损失？

我们认为原因主要有以下几个方面：

1. 传统安全防御手段基于已知威胁，对未知威胁没有防御作用

传统安全防御体系的设备和产品遍布网络 2 ~ 7 层的数据分析。其中，与 APT 攻击相关的 7 层设备主要是 IDS、IPS、审计，而负责 7 层检测 IDS、IPS 采用经典的 CIDF 检测模型，该模型最核心的思想就是依靠攻击特征库的模式匹配完成对攻击行为的检测。反观 APT 攻击，其采用的攻击手法和技术都是未知漏洞（0day）、未知恶意代码等未知行为，在这种情况下，依靠已知特征、已知行为模式进行检测的 IDS、IPS 在无法预知攻击特征、攻击行为模式的情况下，理论上就已无法检测 APT 攻击。

2. 攻击者与防御者在信息上不对称

攻击者在发起攻击前通常都会精心策划每一个攻击环节，包括：攻击工具的开发、控制网络的构建、木马程序的投递、本地的突防利用、通信通道的构建等等。这些精心策划的过程在正式攻击发起前就早已经开始了，例如：攻击者在制造木马程序时会将木马在常见的防病毒软件中进行免杀测试，甚至会在互联网上小范围内进行投放测试，以验证木马效果。而真正在受害者网络中进行的攻击操作则非常精准、隐蔽，使得安全人员能从本地发现的攻击线索非常少，即便发现了异常，也无法定位攻击的来源和过程。因此，对于当前的攻击检测和防御，要求安全分析员不仅仅关注本地数据，还要了解互联网上的威胁情报信息，结合来自互联网的威胁情报来对本地线索进行关联分析。

3. 缺少本地原始数据，难以溯源分析

攻击者通常都会在内网的各个角落留下蛛丝马迹，真相往往隐藏在网络的流量和系统的日志中。传统的安全事件分析思路是遍历各个安全设备的告警日志，尝试找出其中的关联关系。但依靠这种分析方式，传统安全设备通常都无法对高级攻击的各个阶段进行有效的检测，也就无法产生相应的告警，安全人员花费大量精力进行告警日志分析往往都是徒劳无功。如果采用全流量采集的思路，一方面是存储不方便，每天产生的全流量数据会占用过多的存储空间，企业或组织通常没有足够的资源来支撑长时间的存储；另一方面是全量数据来源于网络流量、主机行为日志、网络设备日志、应用系统日志等多种结构化和非结构化数据，无法直接进行格式化检索，安全人员也就无法从海量的数据中找到有

价值的信息。

4. 缺少能在海量数据中快速分析的工具

对高级攻击进行检测需要从内网全量数据中进行快速分析，这要求本地具备海量的数据存储能力、检索能力和多维度关联能力，而传统的数据存储和检索技术很难达到这样的要求。例如：在一个中型规模的企业中记录全年的网络出口流量，大约有 2000 亿条日志，需要约 300 多 TB 的存储空间，如果使用传统的检索技术进行一次条件检索，大概需要几个小时的时间。这种效率明显不能满足攻击行为分析的需求。

2 产品设计目标

360 态势感知与安全运营平台是基于 360 威胁情报和本地大数据技术的对用户本地的安全数据进行快速、自动化的关联分析，及时发现本地的威胁和异常，同时通过图形化、可视化的技术将这些威胁和异常的总体安全态势展现给用户的系统。360 态势感知与安全运营平台一方面可基于 360 自有的多维度海量互联网安全数据，进行情报挖掘与云端关联分析，提前洞悉各种安全威胁，并将威胁情报以可机读格式推送到本地系统，供本地威胁检测和分析时使用，另一方面，360 态势感知与安全运营平台可对本地全量数据进行采集和存储，利用大数据技术在本地进行安全数据分析和威胁溯源。整个设计将遵循发现、阻断、取证、溯源、研判、拓展的安全业务闭环设计，使得用户能通过产品各个功能模块完成威胁处置的全过程。

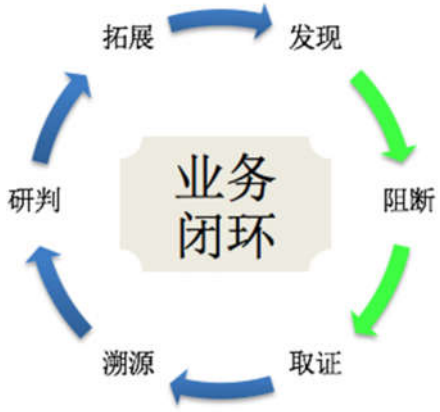

图 2 威胁处理闭环

2.1 产品价值

360 态势感知与安全运营平台可为用户带来以下几个方面的价值：

1. 将大数据技术应用到本地，建立本地大数据分析平台，使用户能将全量的网络行为日志、主机行为日志和设备日志保存下来。
2. 内置基于大数据量的流式计算的高效关联分析引擎，实现对所采集数据的实时关联分析，发现疑似的网络攻击和违规访问行为。
3. 支持加载指定的分析规则，对满足指定条件的历史数据进行关联分析，从历史数据中发现可疑访问行为；
4. 能在本地海量日志中通过优化的快速搜索技术进行日志检索和关联分析，发现攻击者留下的蛛丝马迹，绘制出攻击者的攻击路径。
5. 将 360 互联网威胁情报中心的 APT 跟踪和研究成果做成可机读的高级威胁情报，并推送到用户本地，使用户能及时发现隐藏在客户网络中的攻击事件。
6. 提供高可视化的事件溯源分析工具，使安全分析人员可以方便、快速地对攻击事件和可疑网络访问行为进行溯源分析。
7. 提供丰富的云端安全基础数据，通过可视化的展现方式为用户提供其外网网站的安全态势和遭受 DDOS 攻击的信息，提升用户对外网网站安全风险的把控能力。

8. 通过可视化技术为用户提供基于其网络内威胁和异常的整体安全态势大屏，使得用户的业务管理和决策者能通过大屏总览其网络的全网威胁态势，有利于帮助业务管理和决策者迅速判断做出决策。
9. 与 360 天擎终端安全管理系统和 360 天堤下一代防火墙进行联动，能通过天擎和天堤收集主机行为日志和网络行为日志，同时还能将经过 360 态势感知与安全运营平台关联分析后产生的告警和防御建议推送到天擎和天堤管控中心。
10. 产品功能基于发现、阻断、取证、溯源、研判、拓展的安全业务闭环设计，使得用户能通过产品各个功能模块完成威胁处置的全过程。

3 产品关键技术

采用大数据技术，实现事件的分布式采集、分析、存储和检索，对海量的日志数据、流量数据、数据包数据等做到实时关联分析、快速检索、高效统计，并以高度可视化的方式进行数据展现；同时，可利用 360 海量的云端安全数据提供安全态势分析与展示的数据服务；利用 360 提供的可机读的高级威胁情报在本地数据中进行比对，发现 APT 攻击和本地网络中的 Botnet 主机；提供与网关设备和终端控制系统的联动（NDR/EDR），及时阻断攻击和违规访问，实现安全运营的闭环操作。

通过部署 360 态势感知与安全运营平台和采用 360 的数据数据服务，能够真正提高用户处理海量安全事件的处理性能，提升发现未知威胁攻击的能力，提高用户对安全的“看见”的能力。

3.1 万兆网络及 IPv4/IPv6 网络环境下数据还原技术

本技术主要是采取分光器镜像网络入口上下行数据，输入到几台实体机器做相关分析。流量搜集和分析模块使用自定义的高性能内核和驱动程序，使用独立部署模式，可以支持最高 1Gbps，以及使用集群部署模式，可以支持高达 20Gbps 的线速流量搜集。

流量分析和数据还原使用自主知识产权的协议分析模块，可以在 IPv4/IPv6 网络环境下，支持 HTTP （网页）、SMTP/POP3（邮件）等主流协议的高性能分析，并拥有自主研发的碎片文件侦测和 P2SP 重组模块，可以还原通过迅雷等国内主流 P2SP 软件下

载的文件。

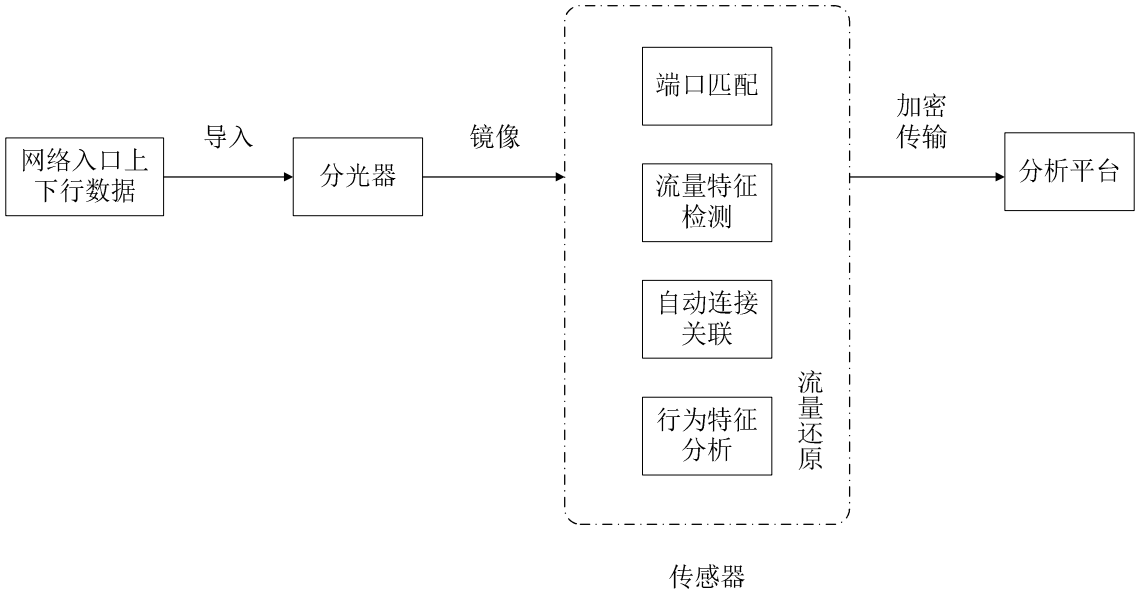

图 3 流量还原流程图

流量的还原当中使用了多种技术，包括端口匹配、流量特征检测、自动连接关联和行为特征分析。

1）端口匹配：在网络协议发展的过程当中，已经形成了一系列的标准协议规范，其中规定了不同协议使用的端口，而很多广泛使用的应用程序虽然没有别标准化，但已经形成了事实上的标准端口。端口匹配就是根据这些标准或非标准的对应关系，根据 TCP/UDP 的端口来识别应用。这种方式具有检测效率高的优点，弱点是容易被伪造，因此在端口检测的基础上，还需要增加一些特征检测的判断和分析，来进一步分析这部分数据。

2）流量特征检测：相对于端口，不同的应用程序使用的协议也存在大量的共性。这些共性就是所谓的流量特征。对于流量特征的识别，大致分为两种：一种是有标准协议的识别，标准协议规定了特有的消息、命令和状态迁移机制，通过分析应用层内的这些专有字段和状态，就可以精确可靠地识别这些协议；另一种是未公开协议的识别，一般需要通过逆向工程分析协议机制，直接或解密后通过报文流的特征字段来识别该通信流量。

3）自动连接关联：随着互联网应用的发展，在互联网上传输的数据越来越多，单个连接完成所有任务的模式也逐渐开始出现瓶颈，因此很多协议开始采用动态协商端口

的方式进行传输，这种模式最早出现在标准的 FTP 协议上，后来逐渐在语音、视频和文件的传输上面被广泛使用。为了识别这种数据，需要根据控制链接上面的报文信息，自动关联到数据传输的链接并对其进行还原，这种技术成为自动连接关联。

4）行为特征分析：针对一些不便于还原的数据流量，可以采用行为特征的方法进行分析。这种方法不试图分析出链接上面的数据，而是使用链接的统计特征，如连接数、单个 IP 的连接模式、上下行流量的比例、数据包发送频率等指标来区分应用类型。如网络电话应用通常语音数据报文长度较为稳定，发送频率较为恒定，P2P 网络应用单 IP 的连接数多、每个连接的端口号都不同、文件共享数据包包长大而稳定等等，都是可以利用来进行应用特征检测的特征指标。

对互联网数据的识别是上述多种技术综合运用。流量分析和文件还原模块会使用这些技术，能够支持 Web、文件、邮件等主流的协议，并能够支持具有中国本土特征的应用程序协议。

3.2 数据处理与计算分析的自动化关联技术

360 态势感知与安全运营平台采用数据处理与计算分析的自动化关联技术自主研发的 SecStream 作为整体的事件处理流程框架，对各类数据按照预定的流程进行流式处理，以保证各种数据处理的准确性。SecStream 自身具有的特性可以满足大数据平台在数据处理阶段的实时、高效、并发、可靠的要求。

SecStream 框架自身是一个分布式的结构，支持水平扩展，通过增加集群节点即可提高集群的并发处理能力。SecStream 还具有自动容错机制，可自动处理进程、机器以及网络异常，保证事件处理流程的稳定运行。在处理数据时，由于数据不写入磁盘，均是缓存在各个节点的内存中，因此 SecStream 具有延迟低，实时性强的特点，通过预先设定的 SecStream 事件处理拓扑，可以快速的对事件处理流程进行搭建，可根据不同的处理要求构建相应的事件处理拓扑模型，满足业务要求。

- 数据处理：在 SecStream 流程中对数据通过正则表达式、协议还原等方式进行归一化处理，将各种厂商及设备类型的日志信息以及流量数据归一化为平台通用的数据对象，作为整个平台分析、存储的数据元。
- 情报知识库关联：数据元经过情报库与知识库关联，补充企业支持信息及

情报信息，为后续的检测分析提供数据基础。情报库与知识库均存储在平台的分布式搜索引擎 SecSearch 中，可为平台数据关联提供快速的检索功能。

- 计算分析：SecCEP 引擎可以满足对事件进行复杂分析并作出相应反应等这类要求吞吐量高，响应时间低以及复杂处理逻辑的需求，SecCEP 可以提供 EPL 语句去分析指定基于 EPL 表达式的事件匹配模式，重点是分析以时间为基础的各个事件之间的关系。SecCEP 通过事件过滤、时间滑动窗口聚集、事件分组窗口输出率限制、事件的内外连接等逻辑对事件进行分析。平台使用 SecCEP 作为实时关联规则引擎，作为事件处理流程中的一个 Bolt 节点，所有事件处理完成后将汇总进入 SecCEP 规则引擎入口，规则引擎内置多种分析规则，结合日志数据、流量数据等数据元分析数据流中的异常，从而触发告警。规则引擎的配置文件可以通过大数据展示平台进行统一配置管理，并支持在线更新。

3.3 规模化沙箱动态检测技术

大数据安全分析及态势感知系统采用规模化动态沙箱的技术对 APT 攻击的核心环节“恶意代码植入”进行检测，与传统的采用基于恶意代码特征匹配的检测方法不同，大数据安全分析及态势感知系统所采用的规模化动态沙箱的方法可以对未知的恶意代码进行有效检测，这种利用对恶意代码的行为进行动态分析的方法，可以避免因为无法提前获得未知恶意代码特征而漏检的问题，亦即在无需提前预知恶意代码样本的情况下仍然可以对恶意代码样本进行有效的检测，因为未知恶意代码是 APT 攻击的核心步骤，因此对未知恶意代码样本的有效检测，可以有效解决 APT 攻击过程的检测问题。

大数据安全分析及态势感知系统相对于其他同类产品的最大特点在于：将会提供了非常丰富的沙箱环境，这种规模化的沙箱环境可以有效保障每种待检测的文件样本都有其适合打开、运行的沙箱环境，同时大数据安全分析及态势感知系统的沙箱采用了高级优化技术，可以有效降低样本文件在沙箱之中打开、运行过程中的内存资源消耗、CPU 资源消耗，与其他同类型产品相比，可以以最小的资源消耗、最快的速度得出准确的检测结果。

目前大数据安全分析及态势感知系统需要模拟沙箱环境包括：PDF 沙箱、Word 沙箱、浏览器沙箱、邮件沙箱、图片沙箱等。同时，借助于大数据安全分析及态势感知系统的多核平台，大数据安全分析及态势感知系统中的各种规模化沙箱可以绑定在处理器的物理核心上进行快速运行，这种进程与处理器绑定的方式可以有效降低进程在处理器的不同处理核心上切换所带来的资源开销，降低并发检测线程之间的资源竞争，有效提高资源利用率。

大数据安全分析及态势感知系统与国内外其他同类型产品相比，最大的优势将在于所模拟的沙箱类型众多，可以提供更多、更精确的文件类型的沙箱检测，与国内同类型产品星云系统相比，大数据安全分析及态势感知系统模拟的沙箱类型将会超过星云系统的一倍，而与国外的同类型产品 Fireeye 相比，大数据安全分析及态势感知系统将会支持更多类型的中国国产软件及系统的沙箱模拟，比如：大数据安全分析及态势感知系统可以模拟国产软件 WPS 的沙箱环境，但是美国的同类型产品 Fireeye 则无法模拟 WPS 文件的打开与运行环境，这样，在 WPS 文件中含有恶意代码的情况下 Fireeye 产品将会产生漏报。

3.4 基于大数据挖掘的恶意代码智能检测技术

该技术是一个人工智能引擎，依靠海量数据挖掘、引入机器智能学习算法，能够有效准确识别未知恶意软件，能够根据已知的正常软件和恶意软件的大量样本，通过数据挖掘找出两类软件最具有区分度的特征，建立机器学习模型，使用机器学习算法，得到恶意软件的识别模型。通过获得的模型对未知程序进行分析判断，即可获得软件的恶意概率，从而在可控的误报率之下尽可能多的发现恶意程序。

该引擎的学习流程如下图所示：

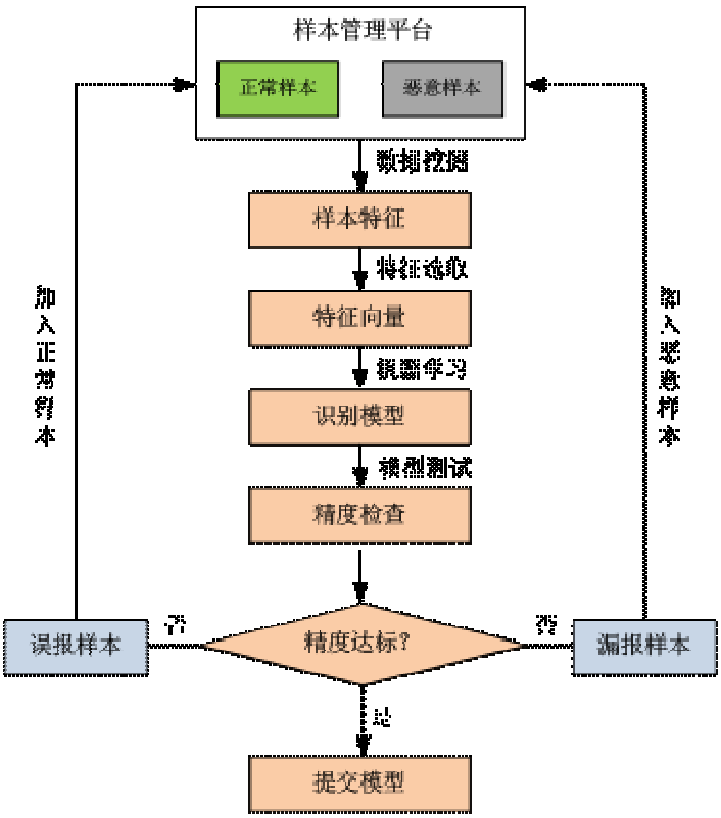

图 4 QVM 引擎机器学习流程示意图

样本管理平台负责管理训练样本，并且对可疑样本可进行人工分析，保证训练样本的纯度，并给下面的阶段提供数据。

通过对训练样本的数据挖掘，例如导入 API 函数、PE 头部信息、代码反汇编信息等等进行海量数据挖掘，找到海量 PE 文件特征。应用特征选取算法，选取最有效的特征，建立特征模型。

利用特征模型对训练样本数据进行数据特征化变换，生成对应的特征向量，利用成熟的机器学习算法（例如 SVM)，对样本进行训练，得到恶意程序识别问题的识别模型。

对生成的模型进行测试，如果精度达到要求，则终止。否则对误判样本进行分析（在样本不确定的情况下，需要人工分析确认),调整样本的分类属性，再次迭代。

该引擎的运行环境如下图所示：

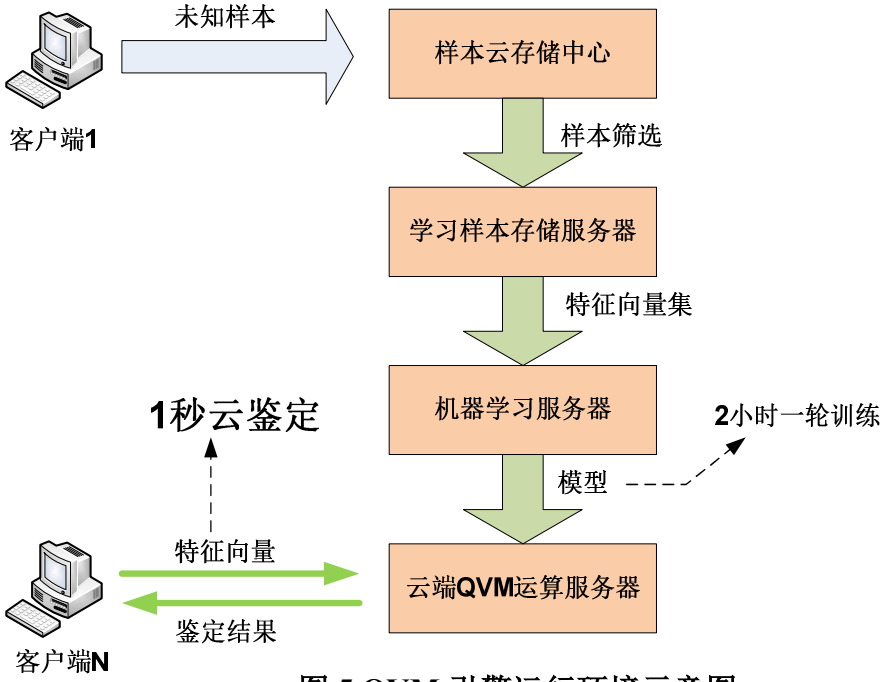

图 5 QVM 引擎运行环境示意图

由于安全领域用户对误报的敏感性和特殊性，导致长期以来机器学习算法在本领域一直作为不大，多数研究者尝试后，无法达到预期的精度而放弃。所以对本技术而言，首要处理的就是降低误报率。根据前期研究的结果，一个合适的机器学习算法的选择对误报控制是相当重要的，目前初选 SVM 作为基本学习算法，并设计了快速的参数选择，和快速训练方法。

机器学习算法在人工少量干预样本（添加、删除、修改黑白属性)指导的情况下，系统能够实现自我学习，自我进化。目前该引擎学习一轮的时间仅为 2 小时。

机器学习有效的解决了大部分未知恶意程序的发现问题。由于传统杀毒技术严重依赖于样本获得能力和病毒分析师的能力，基本只能处理已知问题，不能对可能发生的问题进行防范，具有严重的滞后性和局限性。本技术对海量样本进行挖掘，能够找到恶意软件的内在规律，能对未来相当长时期的恶意软件技术做出前瞻性预测，实现不更新即可识别大量新型恶意软件。

传统杀毒软件技术基本基于简单的特征或者规则进行查杀，很容易被病毒作者免杀。本算法单特征贡献相当微弱，所以简单免杀很难奏效。

机器学习使得对样本分析人员的要求相对较低，仅仅需要分析员能够区分文件是否恶意，而不需要人工分析恶意软件实现方法和识别方法，降低了人员参与门槛，大大节约了人力成本。

3.5 基于搜索引擎技术的大流量行为检索与存储

在本地数据的存储和检索方面，360 使用 ElasticSearch 检索平台做为平台基础，并进行了定制化修改，并配套了大量的检索和分析软件以对数据做到高效分析。

ElasticSearch（简称 ES）是一个基于 Lucene 构建的开源，分布式，RESTful 搜索引擎。其实时性能优越；安装配置简单；RESTful API 和 JSON 格式的文档型数据，降低开发调试的难度；具备面向文档的全文检索能力；而且具有分布式部署，高可靠，稳定等优点。非常适宜作为企业本地的数据存储和检索平台。

ES 可将索引以多个分片和多个副本的形式存储于分布式系统当中，即可提高检索性能，又能保证数据的可靠性。而且其默认使用的内存索引方式可以保证系统对近期录入的数据做到近乎实时的查询，而对于存储于硬盘的 TB 级数据也可做到秒级查询。同时为了便于将来自采集设备或终端上的数据录入 ES 平台，天眼将使用 JSON 格式定义所有的采集后数据。

3.6 云端基于大数据的 APT 发现与跟踪技术

为支持大量数据的采集和处理，首先在云端需要有一套可处理 PB 级数据的大数据平台。该平台基础架构如下：

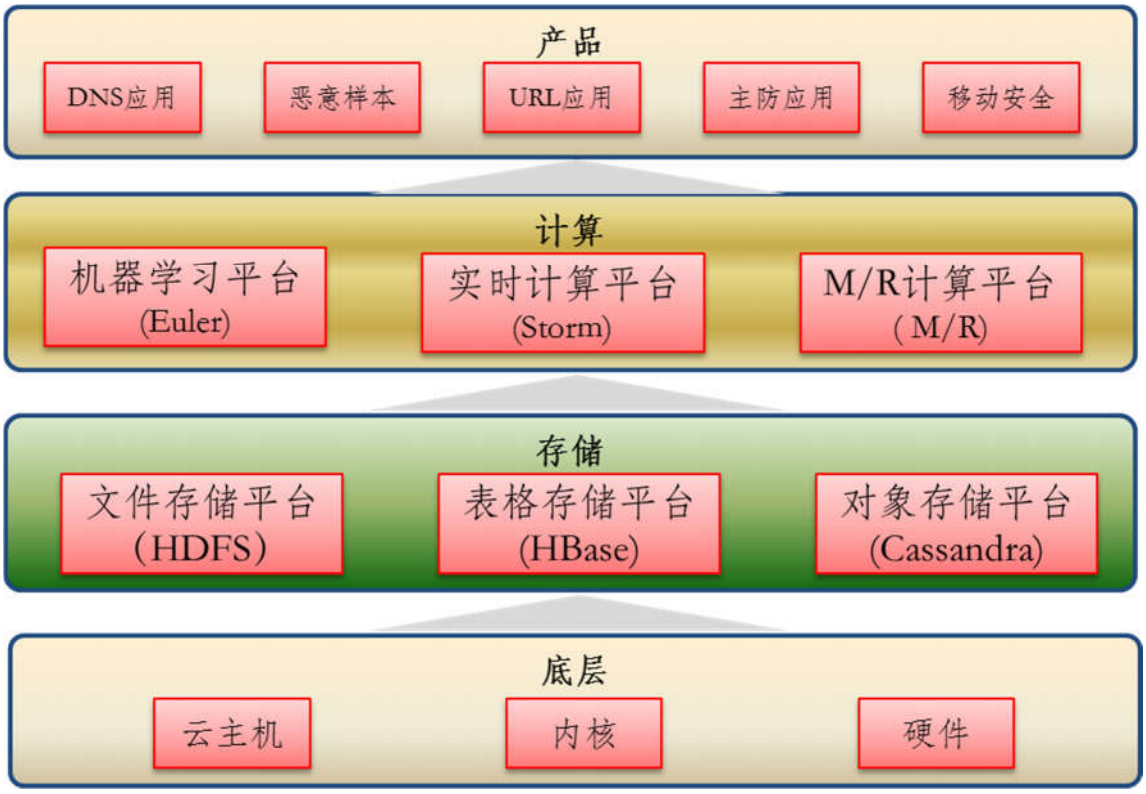

图 6 大数据平台架构图

该平台几个核心技术如下：

1、Hadoop Distributed File System，简称 HDFS，是一个分布式文件系统。

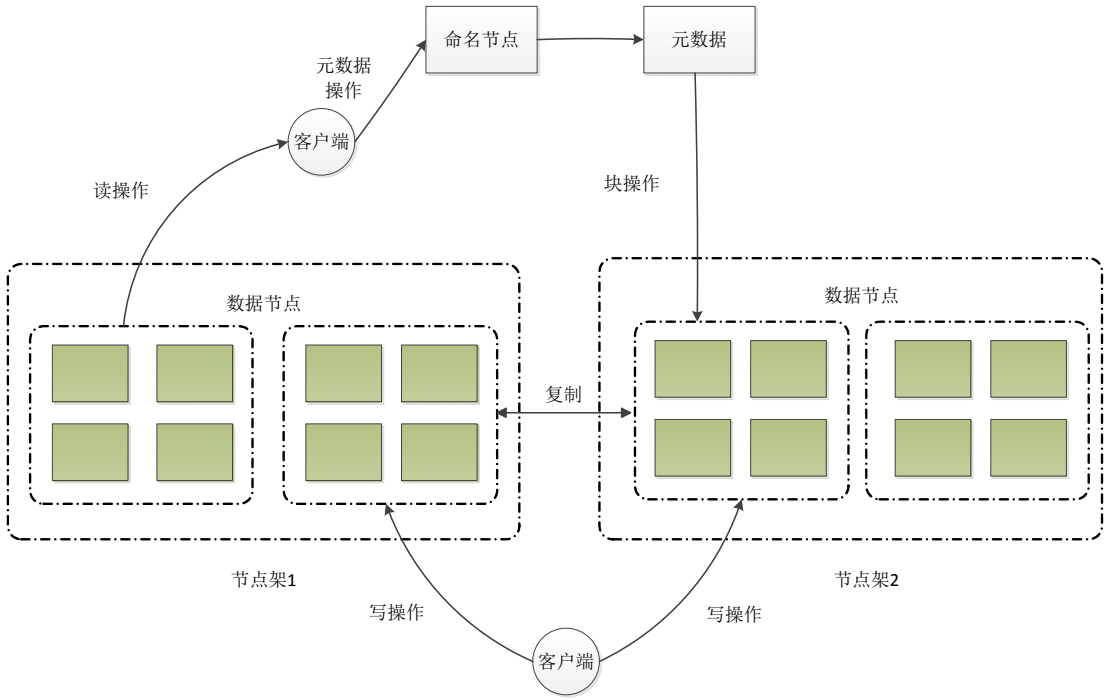

图 7 分布式文件系

HDFS 有着高容错性（fault-tolerant）的特点，并且设计用来部署在低廉的（low-cost）硬件上。而且它提供高吞吐量（high throughput）来访问应用程序的数据，适合那些有着超大数据集（large data set）的应用程序。HDFS 放宽了（relax）POSIX 的要求（requirements）这样可以实现流的形式访问（streaming access）文件系统中的数据。HDFS 开始是为开源的 apache 项目 nutch 的基础结构而创建，HDFS 是 hadoop 项目的一部分，而 hadoop 又是 lucene 的一部分。Hadoop 分布式文件系统(HDFS)被设计成适合运行在通用硬件(commodity hardware)上的分布式文件系统。它和现有的分布式文件系统有很多共同点。但同时，它和其他的分布式文件系统的区别也是很明显的。HDFS 是一个高度容错性的系统，适合部署在廉价的机器上。HDFS 能提供高吞吐量的数据访问，非常适合大规模数据集上的应用。HDFS 放宽了一部分 POSIX 约束，来实现流式读取文件系统数据的目的。

2、MapReduce 分布式计算平台，系统采用 MapReduce 计算模型，MapReduce 是一种编程模型，用于大规模数据集的并行运算。

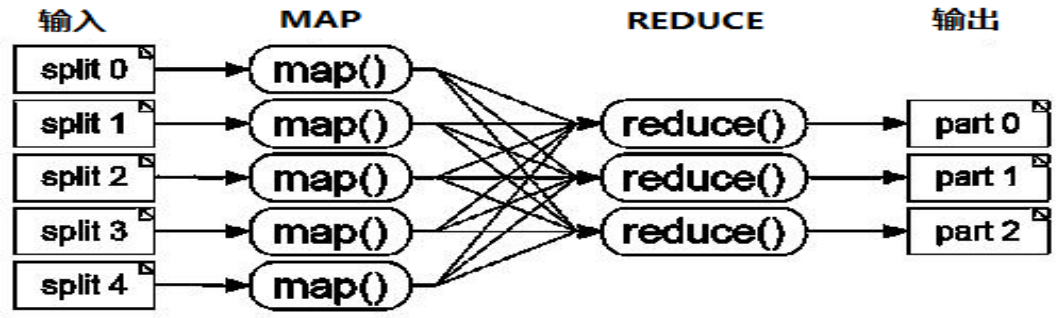

图 8　MapReduce 计算模型

概念"Map（映射）"和"Reduce（化简）"，和他们的主要思想，都是从函数式编程语言里借来的，还有从矢量编程语言里借来的特性。他极大地方便了编程人员在不会分布式并行编程的情况下，将自己的程序运行在分布式系统上。当前的软件实现是指定一个 Map（映射）函数，用来把一组键值对映射成一组新的键值对，指定并发的 Reduce（化简）函数，用来保证所有映射的键值对中的每一个共享相同的键组。

3、HBase - Hadoop Database，是一个运行于 HDFS 顶层的 NoSQL(=Not Only SQL，泛指非关系型的数据库)数据库系统。其具有高可靠性、高性能、面向列、可伸缩的特点。HBase 以表的形式存储数据，表由行和列组成，列划分为若干个列簇(row family)。例如：一个消息列簇包含了发送者、接受者、发送日期、消息标题以及消息内容。每一

对键值在 HBase 会被定义为一个 Cell，其中，键由 row-key(行键)，列簇，列，时间戳构成。而在 HBase 中每一行代表由行键标识的键值映射组合。Hbase 目标主要依靠横向扩展，通过不断增加廉价的商用服务器，来增加计算和存储能力。

在拥有了数据的采集和处理能力后，还需要依赖机器学习、重沙箱和关联分析等能力对大量数据进行筛选，提取重要信息输送到人工运营团队产生结果。其中机器学习部分将使用到 Euler 平台，基于此平台所提供的各种算法，云端可对 DNS、文件等信息进行聚类和相似度分析，以从海量的低价值信息中筛选出可能和已知的攻击行为或攻击特征相关的新生威胁。同时，可视化关联分析平台可为攻击行为和攻击背景的进一步发现提供帮助，可视化的技术手段可以将任意两个互联网信息间的关联性直观的展现在面前，比如两个毫不相干的域名可能拥有相同的注册人，这种相关性分析也为威胁发现提供了进一步帮助。而重沙箱则单纯从文件角度将大量恶意文件或未知文件的网络行为和进程行为输出到机器学习平台和人工运营团队，即为分析提供了数据支撑，也直接输出了部分攻击线索。

3.7 可视化技术

对于安全大数据的 3D 可视化展现，我们使用了 WebGL / Canvas + SVG（Scalable Vector Graphics 即可缩放矢量图形）技术实现。

WebGL 即 OpenGL（开放图形库） 的 Web 版本——OpenGL ES 2.0，现在移动端设备的渲染接口也是 OpenGL ES 2.0，所以理论上 WebGL 程序可以直接在手机、平板电脑等设备上运行。

CPU 和 GPU 都具有运算能力，不使用 CPU 进行图形运算的原因是 CPU 的架构天生是用来做核心运算控制电脑的，而且 CPU 核心数都很少，并发能力差。GPU 则不同，一个 1080p 的显示器有 200 多万的像素点，想要快速渲染这么多像素点，并发能力必须要强，所以 GPU 一般都会有数百个处理单元。而且 GPU 的渲染逻辑是流水线式的，不需要关心其他设备的状态，你给我数据我给你结果，所以用 GPU 做图形运算无疑是最好的方法。

传统的 Web 程序都是在浏览器绘制 dom 节点，大多数情况下使用 CPU 进行运算，当 dom 节点数量超出了 CPU 的处理能力时就会非常卡顿。WebGL 开发的程序是直接

GPU 加速的，直接通过浏览器连接通用图形编程接口的 API，所以使用 WebGL 做可视化程序可以绘制数量非常庞大的数据量，这是天生的优势。

Canvas 是 HTML5 中新增的图形绘制 API，可以实现复杂的图形以及动画，在正常显示器分辨率下效率较高，性能随着分辨率的提升而下降，与 WebGL 结合时可以用作 WebGL 的输出画布。国内知名的可视化框架“百度 Echarts”就是使用 Canvas 开发，但是在处理 3D 图形或高分辨率画面时该技术的优势荡然无存。

在传统的地理信息可视化中很难将同时间、同经纬的信息同时展现出来，即使通过大小、颜色以及形状来区分也都会叠加在一起难以辨识，无法提供详细的信息，更无法体现某段时间的整体态势，在传统 2D 可视化（缺少深度信息）中更是捉襟见肘。所以为了解决上述的问题，必须将可视维度从二维扩展至三维甚至四维空间中，于是就诞生了 DDoS 攻击的树状展示——四维空间地理位置可视化组件，即在三维空间中增加树形结构的深度信息，将本来零散的数据结构化，形成高维度的可视化方案，以便于用户理解。树形结构的深度可以根据数据种类、定义方法以及理解方式的不同定制化，所以这也是一个通用的地理位置可视化组件。

4 产品组成与架构

4.1 产品组成

360 态势感知与安全运营平台主要包括流量传感器、日志采集器、关联分析引擎和分析平台四个硬件模块，如下图所示：

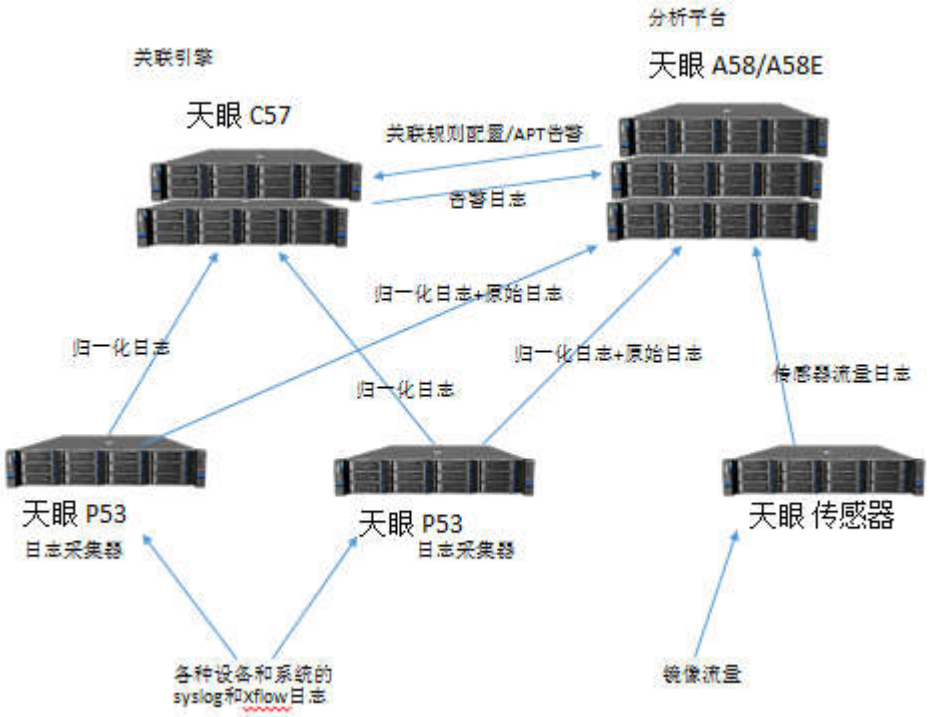

图 9 各模块数据流程图

- 流量传感器

  流量传感器通常部署在网络出口交换机旁，或者其他需要监听流量的网络节点旁。流量传感器主要负责对网络流量的镜像文件进行采集并还原，还原后的流量日志会加密传输给分析平台。

- 日志采集器

  日志采集器主要负责对网络内各业务应用系统、设备、服务器、终端等设备通过主动采集或被动接收等方式对日志进行采集。同时日志采集器还负责对内网资产进行扫描，收集资产数据。在确保网络可达的前提下，日志采集器可与分析平台和关联分析引擎部署在同一位置，如：安全管理区。

- 关联分析引擎

  关联分析引擎主要负责对来自日志采集器的大量日志信息进行实时流解析，并匹配关联规则，对异常行为产生关联告警。通常关联分析引擎与分析平台和日志采集器部署在同一位置。

- 分析平台

  分析平台用于存储流量传感器和日志采集器提交的流量日志、设备日志和系统

日志，并同时提供应用交互界面。分析平台底层的数据检索模块采用了分布式计算和搜索引擎技术对所有数据进行处理，可通过多台设备建立集群以保证存储空间和计算能力的供应。

4.2 产品功能架构

360 态势感知与安全运营平台的产品功能架构分为：数据采集层、存储与工具层、应用交互层、云端数据层，如下图所示：

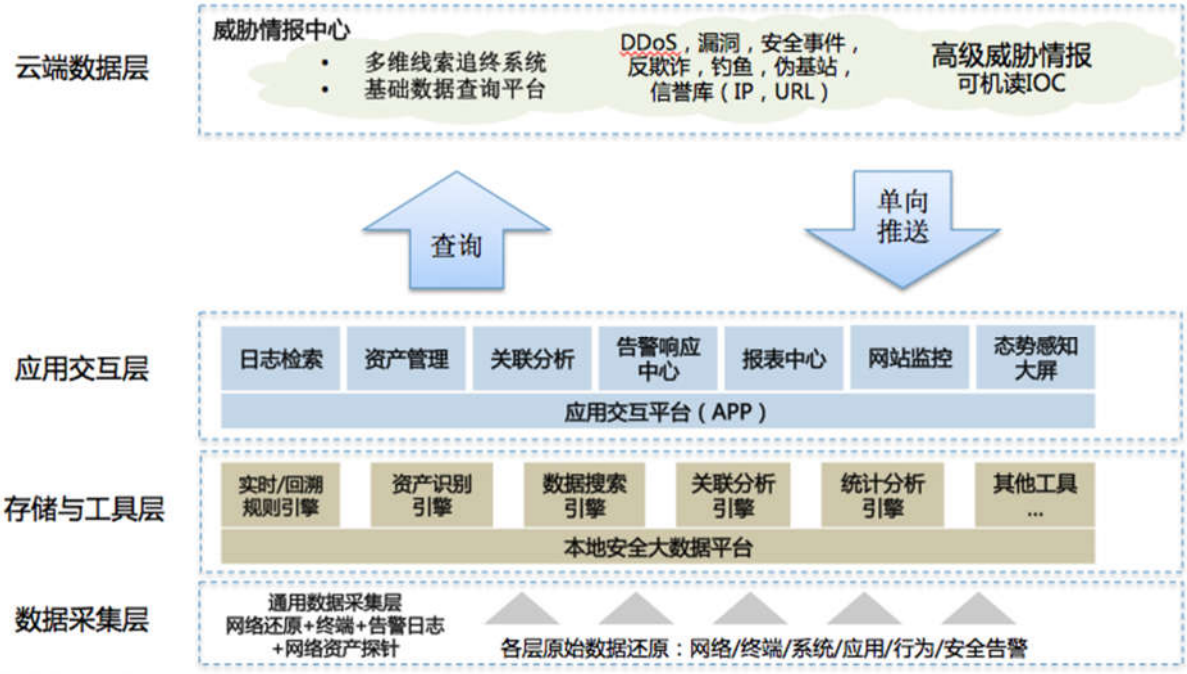

图 10 功能架构图

- 数据采集层

数据采集层由流量传感器和日志采集器两个硬件组成。其中流量传感器的主要功能是接收来自内网的镜像流量，将网络原始数据进行解析、还原，并形成统一的流量日志格式上传到分析平台进行保存。日志采集器的主要功能是对内网各业务应用系统、设备、服务器、终端等设备通过主动采集或被动接收等方式对日志进行采集，对网络设备的 NetFlow 数据的采集。另外日志采集器还负责对内网资产进行主动扫描，以获得内网资产数据。

- 存储与工具层

存储与工具层主要是将采集到的数据进行分部式存储和索引，以便上层应用根据需求随时调用。同时本层还包括多种数据处理引擎，包括：规则引擎、资产识别引擎、数

据搜索引擎、关联分析引擎、统计分析引擎等。这些引擎分别提供各种功能接口，上层应用通过接口调用这些工具引擎对采集到的数据进行分析和处理。

- 应用交互层

行为发现、核查处治、证据固化、攻击回溯、深度分析、事件溯源、背景研判以及拓展深化是处理不同威胁的过程中必须遵循的流程闭环。应用交互层根据威胁处置的闭环流程设计，将产品功能按照威胁处置的不同应用场景分成了日志检索、资产管理、关联分析、威胁情报利用、告警响应中心、报表中心、网站监控、安全仪表板、可视化的事件溯源分析和态势感知大屏等多个应用模块。

- 云端数据层

互联网端的威胁情报数据来源于 360 互联网威胁情报中心，目前包括超过 100 种未公开（首次发现）APT 攻击行为，并且每个月以不低于 10 个的速度在增长。依托互联网端对超过 5 亿 PC 终端和 6 亿移动终端实时保护产生的海量大数据，以及全球最大的 IP、DNS、URL、文件黑白名单四大信誉数据库，同时可以对互联网上活跃的未被发现的攻击进行记录。目前 DNS 库拥有 90 亿 DNS 解析记录，每天新增 300 万； 样本库总样本 95 亿，每天新增 900 万;360 URL 库每天处理 100 亿条，覆盖国内 60%客户端；主防库,覆盖超过 5 亿客户端,总日志数 18 万亿条,每天新增 100 亿条。

5 产品功能

360 态势感知与安全运营平台主要实现以下功能：

- 日志检索

日志检索 APP 的主要功能是对采集到的全量原始日志进行快速检索，可实现千亿条日志秒级检索的性能。

- 资产管理

资产是攻击者的目标，也是安全防御的核心。资产管理 APP 是用来识别和管理本地资产的程序。通过资产管理 APP，安全管理员可对资产赋予不同的安全属性，如安全权重、业务组、资产管理员信息等，并通过可视化技术将资产用不同的拓扑类型进行展示。在进行威胁分析的时候安全分析人员可以通过资产分组来进行安全数据统计和分析，大大提高了安全分析人员的工作效率。

- 关联分析

关联分析 APP 是方便安全分析人员对多维度数据进行关联并分析攻击路径、

取得攻击证据链的工具。在此 APP 上安全分析员可以将原始网络流量日志、原始主机日志、安全设备告警、威胁情报、互联网基础数据等多维度数据进行关联，寻找攻击者的在内网留下的痕迹，对攻击进行溯源和研判，并按照时间维度形成攻击证据链。

■ 威胁情报利用

通过从 360 云端获取（在线查询、云端推送或离线拷贝）可机读威胁情报，本地系统可自动创建分析规则，对本地网络中采集的数据进行实时比对比对，发现可疑的连接行为；同时，可利用威胁情报对历史数据进行比对，以发现曾经发生过的 APT 攻击行为或本地网络中的 Botnet 主机，并可利用情报对安全事件进行溯源分析。

■ 告警响应中心

360 态势感知与安全运营平台采集的数据维度较多，太多的日志和告警反而让安全管理员无从下手。通过告警响应中心，安全管理员可将多个不同维度的数据进行关联后再做研判，这样可大大减少有效告警数量，提升安全管理效率。在告警响应中心，安全管理员可将潜在的威胁的判定逻辑做成关联规则，实时的发现符合威胁判定逻辑的内网行为，并产生告警。

在发现关联告警后，安全管理员可将告警内容和响应建议通过邮件、短信等方式发送给指定的安全事件处置人员，或者将其推送到下级处置中心，如：天擎终端管控中心。在下级处置中心完成事件处置后，告警响应中心会将告警事件标记为“已处置”。

■ 报表中心

提供丰富的报表管理功能；根据时间、数据类型等定期自动生成报表，提供打印、导出以及邮件送达等服务；直观地为管理员提供决策和分析的数据基础，帮助管理员掌握网络及业务系统的状况。报表可以保存为 HTML、EXCEL、文本、PDF、WORD、PNG 等多种格式，提供报表模版的导入、导出功能，用户可根据需求自定义相关报表模版进行数据的导入、导出。

■ 网站监控

网站监控 APP 是用于监测网站安全性与可用性的系统，与常见监控技术不同的是网站监控 APP 采用了云扫描、互联网漏洞众测平台及云多点探测等新技术，

解决了以往网站监测仅依靠本地漏洞扫描及人工值守存在的时效性和准确性等问题。

网站监控系统能通过云扫描技术对大量网站同时进行漏洞扫描，并具备篡改、挂马及暗链的发现能力，通过互联网漏洞众测平台保证漏洞（包括 WEB 0Day）发现的准确性及时效性，通过云多点监测全天候对大量网站进行可用性监测。

- 安全仪表板

利用系统采集的海量数据，并根据用户不同的安全分析应用场景，精心定义了四类不同的安全仪表板，分别为资产威胁视图、互联网威胁视图、安全告警视图、资产安全监控视图。资产威胁视图中定义了内网资产拓扑、资产安全事件告警及漏洞统计、资产风险统计、组件状态等仪表板；互联网威胁视图中定义了外部威胁视图、外联安全事件告警统计、外联安全事件告警详情等仪表板；安全告警视图中主要定义了安全事件告警详情、安全事件告警处置、安全事件处置状态等仪表板;资产安全监控视图中主要定义了安全域资产信息统计、安全域各维度告警及风险统计、安全域所包含资产的漏洞详情等仪表板。通过这些仪表板的使用，极大提高了安全事件的可视化展现能力，同时帮助分析人员从视图中快速发现异常，提供深入分析的线索。

- 可视化的事件溯源分析

通过利用威胁情报发现 APT 攻击或 Botnet 主机后，可进一步通过系统提供的可视化分析工具和海量的云端安全数据，对事件进行溯源分析，在互联网范围内发现类似的攻击事件，找出攻击事件背后的团伙和实际的攻击者，提高分析人员对安全事件的溯源分析能力。

- 态势感知大屏

态势感知大屏 APP 提供 4 种面向不同安全场景的态势监控界面： APT 攻击态势、DDoS 攻击态势、僵木蠕毒安全态势和网站安全态势。

6 产品部署

360 态势感知与安全运营平台在企业内的部署模式如下图所示：

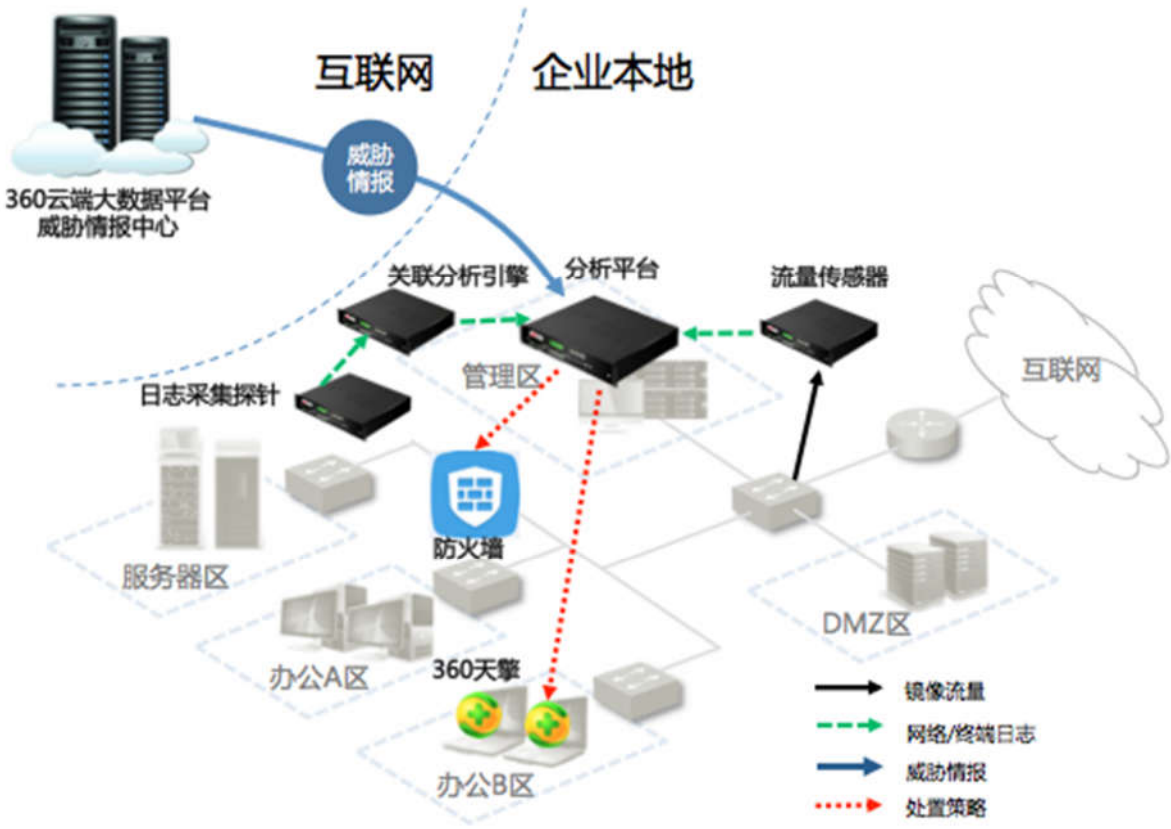

图 11 产品部署图

- 360 态势感知与安全运营平台的各个组件平台均采取旁路部署的模式，组成一个独立的网络，不会和用户本身的网络产生交集。
- 关联分析引擎通过对本地的设备日志、流量日志、本地的安全规则和云端威胁情报进行自动化关联分析，可有效发现本地威胁和异常
- 威胁情报采取单向推送的方式传给部署在用户本地的分析平台，不会造成用户本地数据的泄漏
- 分析平台可采用在线或离线模式获取威胁情报升级包
- 分析平台可轻松进行水平扩展

7 产品优势与特点

1、 首创使用互联网数据发掘 APT 攻击线索，提升企业对威胁看见的能力

传统的 APT 防护技术专注于从企业客户自身流量和数据中通过沙箱或关联分析等手段发现威胁。而由于企业网络防护系统缺少相关 APT 学习经验，而且攻击者的逃逸

水平也在不断的进步发展，本地设备会经常性的出现误报和漏报现象，经常需要人工的二次分析进行筛选。而且由于 APT 攻击的复杂性和背景的特殊性，仅依赖于单一企业的数据经常无法有效的发现 APT 攻击背景，难以做到真正的追踪溯源。而天眼则创新性的从互联网数据进行发掘和分析，由于任何攻击线索都会有相关联的其他信息被互联网数据捕捉到，所以从互联网进行挖掘可极大提升未知威胁和 APT 攻击的检出效率，而且由于数据的覆盖面更大，可以做到攻击的更精准溯源。

2、 以威胁情报形式打通攻击定位、溯源与阻断多个工作环节，帮助企业从源头上解决安全问题

传统的防护体系在多台设备间进行联动往往需要通过特别开发的接口对一种或几种特殊类别的告警或信息进行分发和通知，这种设计往往会制约多种不同设备或系统之间的信息传递。同时由于对于消息接口缺乏一个系统化的规范化的描述，很难对复杂的攻击行为进行准确定义。而天眼的一大创新点在于用威胁情报的形式对各种攻击中常出现的特点和背景信息进行记录和传输，而威胁情报将通过统一的规范化格式将攻击中出现的多种攻击特征进行标准化，可满足未来扩展攻击特征以及后续扩展联动设备的需要。

3、 高效的快速搜索技术帮助企业提升数据查找的能力

传统的安全方案中，对于企业本地数据的处理往往采用 mysql 等关系型数据库。这种设计早已不能满足当前数据量的处理性能需要。天眼创新性的采用搜索引擎技术作为本地数据存储和检索核心技术，采用 json 格式作为引擎的输入输出格式，这样可极大提高检索性能，可以为企业提供 TB 级的数据快速搜索能力，同时相比传统架构也能够降低大量接口上的开发量。天眼可为企业本地的大规模数据保存、攻击证据留存和查询、实时关联分析提供坚实的技术保障。

4、 基于数据处理与计算分析的自动化关联技术，提升了客户发现本地异常行为的能力

360 态势感知与安全运营平台通过基于数据处理与计算分析的自动化关联技术自主研发的 SecStream 作为整体的事件处理流程框架，对各类数据按照预定的流程进行流式处理，以保证各种数据处理的准确性。依靠其延迟低，实时性强的特点，通过预先设定的 SecStream 事件处理拓扑，可以快速的对事件处理流程进行搭建，可根据不同的处理要求构建相应的事件处理拓扑模型，满足业务要求。360 态势感知与安全运营平台使用 SecCEP 作为实时关联规则引擎，作为事件处理流程中的一个 Bolt 节点，所有事件处理

完成后将汇总进入 SecCEP 规则引擎入口，规则引擎内置多种分析规则，结合日志数据、流量数据等数据元分析数据流中的异常，从而触发告警。通过这种自动化关联的方式，360 态势感知与安全运营平台可及时发现用户网内的异常行为并进行告警，保证用户可及时监控并处理这些异常行为，减少用户可能遭受的损失。

5、 基于大数据挖掘分析的恶意代码智能检测技术，提升了客户检测恶意代码的能力

360 态势感知与安全运营平台在云端采用了机器学习等人工智能算法，针对海量程序样本进行自动化分析，有效解决了大部分未知恶意程序的发现问题。由于传统杀毒技术严重依赖于样本获得能力和病毒分析师的能力，基本只能处理已知问题，不能对可能发生的问题进行防范，具有严重的滞后性和局限性。本技术对海量样本进行挖掘，能够找到恶意软件的内在规律，能对未来相当长时期的恶意软件技术做出前瞻性预测，实现不更新即可识别大量新型恶意软件，在全球处于领先水平。

6、 基于可视化技术，使得用户网内的威胁和异常清晰可见

通过可视化技术的利用，将原本碎片化的威胁告警、异常行为告警、资产管理等数据结构化，形成高维度的可视化方案，以便于用户理解。大数据的存储与实时运算能力保证了 360 能够实现数据的实时推送，配以可以实时交互的 3D 可视化界面，与其美观的 3D 展示效果相得益彰。3D 图形可视化有时会将大量的数据抽象地显示出来，我们则在地图旁添加 2D 的统计信息，更便于阅读，区域安全态势一目了然，省去了读繁复报告的过程。可视化技术的利用使得用户可以更直观地感受到网内的安全态势，使得安全由不可见变为可见，不但带来了更好的用户体验，同时还有效地提高了安全监控的效率。

7、 专业的专家运营团队，全天候为企业保驾护航

为了推进自动化分析技术的发展，并对未知威胁做最终定性和跟踪，360 长时间维持了一个近百人的庞大安全分析团队，该团队技术能力覆盖了操作系统、逆向、漏洞挖掘、渗透等安全的各个技术领域，该团队成员的经验为云端分析系统的运行提供了宝贵输入，并支持了国内多次重大 APT 事件的深度挖掘和定位。360 的安全专家团队可为企业提供及时有效的安全服务，帮助企业保护自身网络的安全，减少企业遭受攻击时收到的损失。