IT@Intel Brief
Intel IT
Risk Management
July 2013

Fast Threat Detection with Big Data Security Business Intelligence

- To reduce risk, IT must rapidly correlate event log data
- We are able to selectively monitor 1.5 billion directory service events per day
- Our new solution enables large-scale log management and custom analytics

Intel IT's new Security Business Intelligence (BI) platform incorporates a large-scale common logging service (CLS), real-time correlation engine, and various custom analytics platforms to deliver faster detection and response to security threats. The ability to implement custom analytics solutions enables our security team to filter and distill specific event logs from over 6 billion events recorded daily. The benefits include improved compliance, better protection of high-risk assets, and faster, more intelligent response to advanced persistent threats.

After operating a near-real-time correlation engine on smaller data sets for several years, we saw the need for a comprehensive log management solution capable of recording a full year of Intel's server event log activity. The ability to analyze current logs and historical data helps investigators and threat management analysts better track and identify actionable events.

To build our Security BI platform, we first tested a smaller design using commercial-off-the-shelf (COTS) relational database technology. Based on its success, we scaled out our Security BI platform. This solution enabled us to reduce data collection analysis throughput from two weeks to 20 minutes. Using the Security BI platform, we can selectively monitor 1.5 billion directory service events per day, and then, using a custom analytics platform, focus on the approximately 1,500 account lockout events per day that could reveal brute force attacks (see Figure 1).

Example: Directory Service Attacks

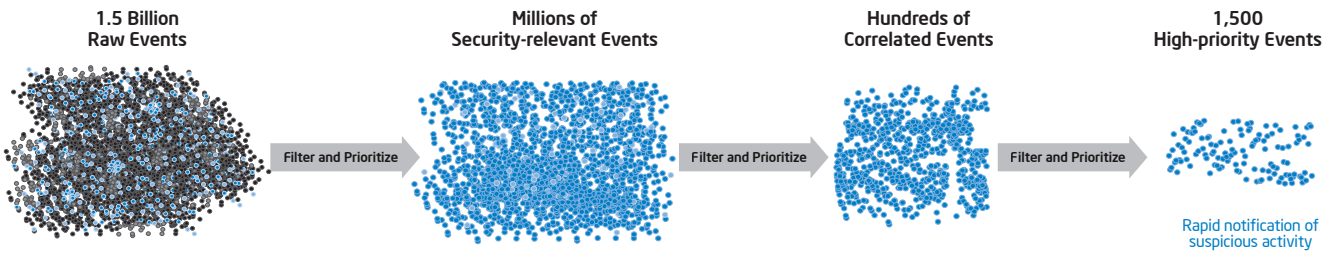

Figure 1. Intel IT's Security Business Intelligence solution collects and correlates billions of new server log events daily. Our custom analytics platform distills these raw events down to a manageable set of high-priority events to derive actionable items.

Background

The importance of Security Business Intelligence (BI) solutions is continuing to grow; as enterprises generate ever-increasing volumes of data, the amount of threats also escalates in number, type, and sophistication. Factors driving these increases include rapid growth in cloud computing, mobile form factors, and new consumer and enterprise applications.

Factors in particular that contribute to Intel's growing potential for information risk exposure include the following:

- IT consumerization, including a 200-percent increase in handheld device use over the past two years
- An annual 35-percent growth in Intel IT's data storage
- The continued blurring of the distinction between infrastructure in the data center and in the cloud

To reduce risk exposure without impacting business velocity, Intel IT formulated our "Protect to Enable" security strategy. This strategy balances enterprise-wide protection with Intel's need to maintain fast information flow for a productive work environment.

Through the judicious use of preventive controls balanced with detection and corrective measures, we provide reasonable protection for data, privacy, and intellectual property, while meeting systems availability goals and regulatory compliance objectives. At the same time, we recognize that in today's fast-paced, competitive business environment, absolute protection is impossible. It is vital that risk management controls do not impede business velocity or adoption of new usage models such as bring-your-own device (BYOD), cloud services, and collaboration solutions.

Security BI Platform Solution

A key component of our Protect to Enable strategy is Intel IT's Security BI platform. This platform uses big data technologies to collect, aggregate, and analyze enterprise activity for incidents such as the unauthorized transfer of data.

According to the Verizon 2012 Data Breach Investigations Report, server-log evidence is believed to be "more effective than nearly all other methods" for discovering breaches.[1] Averaging over 6 billion new logged events every day, Intel's enterprise generates a tremendous amount of data that is searched for breaches and anomalies. The challenge is collecting and analyzing this data fast enough to contain threats and perform fast remediation.

Commercial-off-the-shelf (COTS) solutions, such as Security Information and Event Management (SIEM) packages, tend to offer limited functionality for advanced analytics. We wanted a solution stack that included capabilities for an extreme data warehouse (XDW) that was large enough to run custom applications for billions of events per day. We also wanted to use smaller, relational database platforms and analytics libraries where they are practical and cost effective.

This led us to create our own big data Security BI platform. Our design sought to deliver specific advantages in the collection and analysis of server event log data:

- The capacity for a year's worth of enterprise log events to enable historical views
- The ability to deploy advanced analytics platforms for the rapid detection, correlation, and analysis of specific data sets indicating potential security and privacy threats
- Greater accuracy in identifying threats to significantly reduce false positives
- Reduction of data collection analysis throughput from two weeks to 20 minutes

STARTING SMALL

To develop such a powerful platform, we tested our concept first with a smaller-capacity common logging service (CLS) solution designed to collect and analyze millions of events per day for breach investigation. For the test phase, we selected high-value assets we wanted to protect and a few core infrastructure services that had proved instrumental in identifying past breaches.

[1] See: www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2012_en_xg.pdf

TEST PHASE GOALS

In the initial test phase, we laid the groundwork for our security and privacy measures, designing the data warehouse as a secure enclave with strict access controls. Role-based controls and entitlement rules ensured use for only security-event analysis. We set up secure networks to move data from the data warehouse to the correlation engines.

The initial test phase also included a solution for reducing false positives that used an advanced set of criteria for determining likely threats. This solution helped ensure IT staff would trust and act on results.

BUILDING A TEAM

We assembled a diverse team of experienced security professionals: an overarching program manager, architects, investigators, engineers, project managers, analysts, and operations personnel. These people worked closely with our privacy experts to design and document the tools, policies, processes, and privacy guidelines.

ENSURING PRIVACY

We recognize that protecting our employees' privacy is paramount. Prior to design and deployment, the team is educated on privacy principles to ensure that all parties understand, respect, and abide by Intel's privacy policies. We then implement these policies and processes to ensure appropriate management of personal information throughout the data lifecycle.

TEST PHASE RESULTS

The platform successfully proved that centralization of events and standardized queries dramatically reduce the time required for investigators and analysts to determine if an actionable event occurred. The platform also scaled up to billions of events per day, though this quickly revealed the need for more compute power and data storage.

BUILDING OUT THE PLATFORM

After our successful test phase, we decided to extend our Security BI platform enterprise-wide, adding new capabilities to handle larger data volumes and new use cases.

Our objectives included the following:

- Build a cost-effective solution that maintains regulatory compliance while keeping data available and protected
- Apply advanced big data analytics to improve the ability to predict, detect, prevent, and respond to cyberthreats and incidents, zeroing in only on what is relevant
- Use the results to identify less-effective security controls so we can either improve or eliminate them

Three-Stage Process

The Security BI infrastructure is a three-stage process (see Figure 2).

EXTRACTION AND LOAD

This first stage collects and analyzes log files from proxy servers, domain name servers (DNS), Dynamic Host Configuration Protocol (DHCP), Active Directory* databases, and other systems. It extracts data from event sources such as server logs, security sensors, intrusion systems, and management platforms. The platform also stores and loads contextual information such as asset owners and individual IP addresses. It notes the approximate location associated with a device to help find anomalies such as an account or device logged into a system from two geographically separated places at the same time.

DATA STORAGE AND ANALYSIS

In the second stage, Intel IT's CLS collects and parses the extracted event and contextual data at more than a million events per second. The platform holds a rolling year's worth of data comprising trillions of events in 1.5 petabytes of compressed storage. Depending on task size, investigators can draw specific data from the CLS into a custom data warehouse, such as our XDW, and use a massively parallel processing Structured Query Language engine to perform advanced analytics. Investigators and other authorized users employ standard queries or perform custom searches on recent or historical data sets using various tools. To enable automation, the CLS has an API that enables pushing large data amounts to our real-time correlation engine or custom solution platforms in our Security BI solution. The CLS also manages access control.

REPORTING AND WORKFLOW AUTOMATION

The third stage involves report generation and workflow automation using both COTS and custom algorithms developed by Intel security analysts to identify unusual events. This platform automates many reports to save investigation time. One early custom use case we devised is our server anomalies solution that uses proxy server data analysis to look for evidence of anomalous connections to the outside. Using advanced analytics, we drill down into selected anomalies to make inferences that enable investigators to identify potentially abnormal behavior from our server population out to the Internet.

We use workflow automation to help us accelerate the evaluation of new indicators of compromise. For example, we run daily processes that distill raw logs down to a set of searchable metadata. Using this data set, we can rapidly determine if there is any historic evidence of suspicious activity when analyzing new intelligence data. Other Security BI functions include various decision-support tools, such as analytical and operational reports, real-time trending, and historical reporting. Certain types of event data go directly into the real-time correlation engine where an automated process quickly analyzes and identifies anomalies that might require immediate follow-up.

To enable workflow automation, the Security BI platform includes filters that provide analytics to flag items that require action. This collection of customizable event processors requires periodic tuning and updates of its sensors and advanced analytics solutions to help accurately distill billions of events down to a manageable level of actionable items.

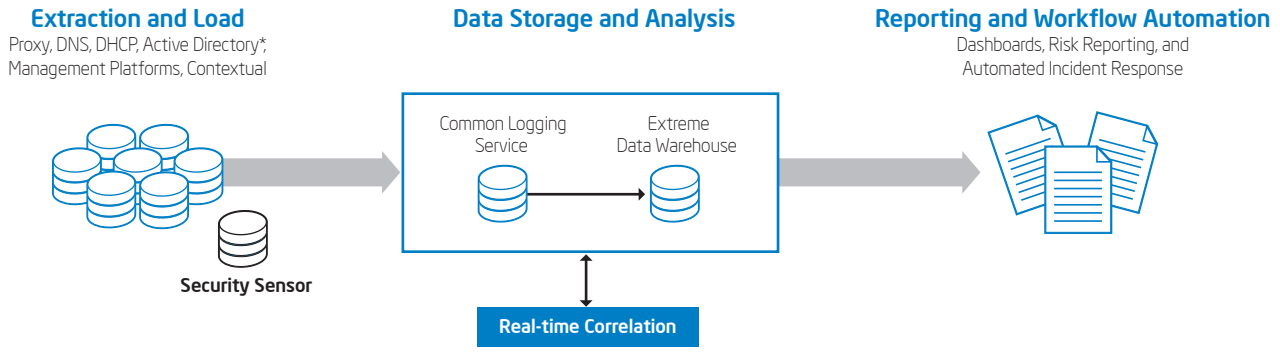

Figure 2. The Intel IT Security Business Intelligence platform is designed to handle three process stages: extraction and load, data storage and analysis, and reporting and workflow automation.

Future Plans

We are currently developing a "My Security Alerts" tool, which will provide employees with the ability to view activity connected with their accounts. Events will be filtered down and analyzed to flag potentially suspicious activity and the summarized view will be available to each individual employee. This information, combined with the employee's knowledge of his or her use of company resources, will allow them to help us identify suspicious activities. Employees will also be able to advise Intel IT if further investigation is warranted.

Intel IT is continuing to scale its Security BI platform to increase its ability to proactively find advanced threats, react quickly, and develop preventive and corrective controls for the future. We are also looking at ways to use trusted sensor and event information from our platforms to improve the quality and reliability of our Security BI system.

Conclusion

We have built a robust Security BI platform with new tools, capabilities, and experienced professionals as a cornerstone of our Protect to Enable enterprise security strategy. This Security BI platform stores a year's worth of server event log data and performs big data correlation to detect potentially inappropriate data-handling abnormalities and issue alerts to security responders. The security responders can ask questions and the Security BI platform generates fast, actionable answers. With this data, the team can make informed decisions and deploy intelligent responses.

Related Topics

Visit www.intel.com/it to find white papers on related topics:

- "Granular Trust Model Improves Enterprise Security"
- "Rethinking Information Security to Improve Business Agility"
- "Using a Multiple Data Warehouse Strategy to Improve BI Analytics"

AUTHORS

Stacy Purcell
Senior Security Architect, Intel IT

Catherine Franke
Threat Database Analyst, Intel IT

Paul Dockter
Privacy Engineer, Intel Privacy Office

For more straight talk on current topics from Intel's IT leaders, visit www.intel.com/it.

Printed in USA ♻ Please Recycle 0713/ERAI/KC/PDF 328757-001US