Monitoring HPC Security at LLNL

4th NIST HPC Security Workshop

Ian Lee
HPC Security Architect

2024-05-20
User Centric View of LC

- Each bubble represents a cluster
- Size reflects theoretical peak performance, color indicates computing zone
- Only large user-facing compute clusters are represented
HPC Zones – User View
More Complete View of LC

- Each bubble represents a managed cluster or system
- Size reflects **number of nodes**, color indicates **type** of system
- All production systems are shown, including non-user-facing systems
HPC Zones – Wider View

- Secure Compute Facility
- FIS
- Open Compute Facility
- Infrastructure Zone
- Restricted Zone
- Collaboration Zone
- Internet
- LLNL Enterprise
- GDO / ATTB (DMZ)
Where We’re Heading
El Capitan

El Capitan vs the Rest

- Each bubble represents a cluster
- Size reflects **theoretical peak performance**, color indicates computing zone
- Only large user-facing compute clusters are represented

El Capitan
~ 2 EF (~ 2,000,000 TF)
system.syslog per hour
Logging Infrastructure
Logging Architecture

Cluster

Node Node Node Node Node Node

CSP Splunk

LC Splunk

splunk®

splunk®
Current Service to Hardware Allocations

~ 112TB NVMe (total)
~ 2PB HDD (total)
Cluster Stats Today

- Records per day
- Total Size of Indices
- Aggregate Log Counts by data_stream.dataset
- Index Growth Rate (WIP)
Log Sources Breakdown
Example: Audited log issue
Jan 1 – Jan 2 (1.25 seconds)
Jan 1 – Jan 8 (20 seconds)
Nov 1 – Jan 8 (42 seconds)
Monitoring
<table>
<thead>
<tr>
<th>Incident</th>
<th>Created on</th>
<th>Updated on</th>
<th>Short description</th>
<th>Recent Incident</th>
<th>Assigned to</th>
<th>Assignment group</th>
<th>State</th>
<th>Close code</th>
</tr>
</thead>
<tbody>
<tr>
<td>INC0406042</td>
<td>Mar 5, 2024 @ 21:50:52.000</td>
<td>Mar 5, 2024 @ 23:23:30.000</td>
<td></td>
<td>4</td>
<td>LC Operations</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC0406041</td>
<td>Mar 5, 2024 @ 21:14:29.000</td>
<td>Mar 5, 2024 @ 21:29:13.000</td>
<td></td>
<td>0</td>
<td>LC Operations</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC0406035</td>
<td>Mar 5, 2024 @ 19:17:53.000</td>
<td>Mar 5, 2024 @ 20:40:04.000</td>
<td></td>
<td>43</td>
<td>LC Operations</td>
<td>Resolved</td>
<td>Resolved Remotely</td>
<td></td>
</tr>
<tr>
<td>INC0406032</td>
<td>Mar 5, 2024 @ 19:13:34.000</td>
<td>Mar 5, 2024 @ 19:14:49.000</td>
<td></td>
<td>0</td>
<td>LC Production Systems</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC0406011</td>
<td>Mar 5, 2024 @ 18:40:16.000</td>
<td>Mar 5, 2024 @ 19:34:27.000</td>
<td></td>
<td>2</td>
<td>LC Operations</td>
<td>Resolved</td>
<td>Cancelled</td>
<td></td>
</tr>
<tr>
<td>INC0405988</td>
<td>Mar 5, 2024 @ 17:48:56.000</td>
<td>Mar 5, 2024 @ 18:54:36.000</td>
<td></td>
<td>0</td>
<td>LC Operations</td>
<td>Resolved</td>
<td>Software - Bounced</td>
<td></td>
</tr>
<tr>
<td>INC0405946</td>
<td>Mar 5, 2024 @ 16:53:50.000</td>
<td>Mar 5, 2024 @ 19:46:53.000</td>
<td></td>
<td>0</td>
<td>LC Operations</td>
<td>Resolved</td>
<td>Cancelled</td>
<td></td>
</tr>
<tr>
<td>INC0405938</td>
<td>Mar 5, 2024 @ 16:40:58.000</td>
<td>Mar 5, 2024 @ 18:58:26.000</td>
<td></td>
<td>1</td>
<td>LC Operations</td>
<td>Resolved</td>
<td>Solved</td>
<td></td>
</tr>
</tbody>
</table>
Operating System Versions
PAN Firewall Traffic Flows
SSH Authentications
Kibana Security Dashboards

Dissecting REMCOS RAT: An in-depth analysis of a widespread D2K4 malware, Part Two
2024-10-18
Part two: Digging into REMCOS recording capabilities, launch, and C2 communication.

New & updated prebuilt Elastic rules available
2024-09-18
The new out-of-the-box Elastic rules release includes new rules for Windows and AWS

Alert trend
Showing 11,785 alerts

Events
Showing 7,000,921 events

Host events
Showing 2,683,959 events

Network events
Showing 8,043,515 events

Kibana Security Dashboards
Monitoring Vision Going Forward

- Explore other offerings
  - ML / Anomaly Detection
  - Enterprise Search (unified search across web + confluence + gitlab)
  - Elastic Defend

- More automated alerts / processes
Security Baseline

- Security requirements often quite prescriptive
  - STIG > CIS Benchmark > Vendor Guideline > generic NIST 800-53 controls

- Developed a STIG for the TOSS operation system with DISA
  - Inspired by RHEL 8 STIG, which TOSS 4 is derived from
  - Small tweaks: adjust some DoD specific language to make compatible for other Gov agencies
  - Larger requests: no explicit allow-listing of software on TOSS, being a software development OS
  - HPC specific: RHEL STIG says 10 concurrent sessions for DOS reasons, TOSS STIG allows 256

- Need to regularly check and validation configuration
  - [https://github.com/llnl/toss-stig](https://github.com/llnl/toss-stig)
Community Work

- [https://github.com/llnl/cmvl](https://github.com/llnl/cmvl) (WIP)
  - Repository of Elastic, Splunk, etc queries, dashboards, and visualizations

- [https://github.com/LLNL/elastic-stacker](https://github.com/LLNL/elastic-stacker)
  - Export saved objects from Kibana for sharing

- [https://github.com/LLNL/toss-configs](https://github.com/LLNL/toss-configs) (WIP)
  - Configuration files and scripts for setting up and maintaining TOSS HPC systems

- [https://github.com/llnl/toss-stig](https://github.com/llnl/toss-stig)
  - Ansible implementation of the TOSS STIG
HPC Security Technical Exchange – August 2024

- August 5 – 8, Lawrence Livermore National Laboratory, CA
- Government focus; CUI up to TS//SCI
- Registration to open imminently, including Call For Topics / Prompts
  — Contact ian@llnl.gov for details / “wait list”
Thank you!

Happy to chat and answer questions!

ian@llnl.gov

@IanLee1521
What is Elastic?

- Elastic Stack is a collection of software for logging, reporting, and visualization of data
  - Formerly “ELK stack” or just “ELK”

- Consists of Elasticsearch, Logstash, Kibana, Beats, and more

- Open source components, commercial support, similar idea to GitLab
Why Elastic?

- Were already using parts of the Elastic stack before (Logstash, *Beats)

- Better integration / extension support
  - Enterprise Search, Machine Learning tools, Elastic Integrations via Agent

- Performance claims (should be significantly faster searches compared to Splunk)
  - Reality has been a bit mixed here, and there is definitely room to continue tuning our deployment.

- Compared notes with ORNL folks who moved (at least partially) from Splunk to Elastic
Deployment

- GitLab CI + Ansible configuration (separate from the TOSS Ansible repo)
  - James Taliaferro did a talk at S3C at NLIT going into detail on this

- Very fast to destroy and rebuild the Elastic clusters

- Straightforward to scale up the service to meet demand