

CRIBLCON²⁴

POWERED BY  Cribl

Navigating Transition: From Syslog and Logstash to Cribl.

CHANDA PULLIAM
Information Security Engineer, Synopsys





CHANDA PULLIAM

Information Security
Engineer, Synopsys

Portions Copyright Synopsys, Inc. 2024. All rights reserved. Used with permission

Agenda

1 WHAT WE DO

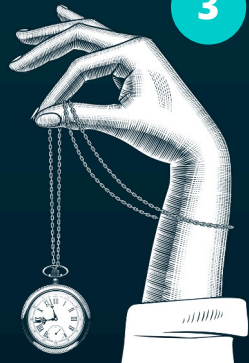
2 THE PROBLEM WITH SYSLOG

3 SIZING EVENTS IN ELASTIC

4 ZSCALER MIGRATION

5 STATS AND VALUE

6 RECOMMENDATIONS



In the InfoSec Command Center...

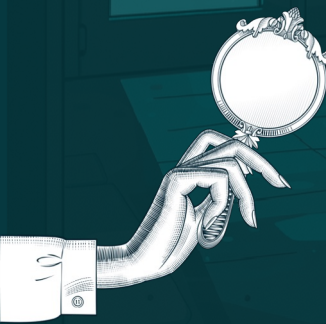
Who are we?

About Synopsys

- Catalyzing the era of pervasive intelligence, Synopsys, Inc. (Nasdaq: SNPS) delivers trusted and comprehensive silicon to systems design solutions, from electronic design automation to silicon IP and system verification and validation. We partner closely with semiconductor and systems customers across a wide range of industries to maximize their R&D capability and productivity, powering innovation today that ignites the ingenuity of tomorrow.

About me:

- 20+ years of experience
- 4 years of experience with Elasticsearch
- Started working with Cribl in October, 2023
- Two kids
- Afraid of horses

A hand holding a magnifying glass over a server room floor. The hand is wearing a white shirt cuff with a button. The magnifying glass is held over a checkered tile floor in a server room. In the background, there are several server racks and a door with a 'CLOSE' sign above it.

How many
audit log entries
does it take to
install syslog?

48....

The scene of the crime.

Opportunities for data optimization:

Too many moving parts:

- Beats agents.
- Logstash.
- Syslog server(s.)
- Docker containers.
- Ingest pipelines.
- Custom apps / scripts.

Too much noise:

- Duplicate data.
- Irrelevant events.
- Inconsistent field names.
- Incorrectly parsed logs.
- Unnecessary fields.

Goals and objectives:

- Improve event routing flexibility.
- Simplify onboarding process.
- Lessen workload and strain on SIEM.
- Improve data retention and retrieval

Sizing up the suspects.

AKA why you need to size Elasticsearch events.

Understand pain points :

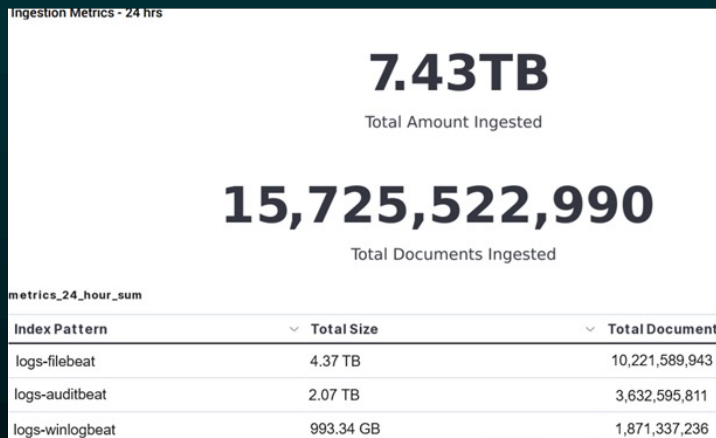
- Largest and smallest log sources.
- Percentage of ingest used by log sources.
- Surprise shifts in ingest amounts or sizes.

Focus on high-use, high-impact logs:

- Cut out the noise.
- Decrease Elasticsearch storage and compute requirements.

Execs love numbers:

- Storage Savings in Percentages.
- Cost savings in **actual dollar amounts**.





©2024 Synopsys, Inc. All rights reserved.


How to Size Elastic events.





Considerations.

- Elasticsearch compression matters.
- Mapper-size plugin not enabled by default.
- Logs need time to marinate.
- Self-managed Elasticsearch clusters may require a rolling restart to apply.

View: [Single document](#) [Surrounding documents](#)  

[Table](#) [JSON](#)

 Search field names

Actions	Field	Value
...	 _id	ABb46Y8B3tKeEGSuUoYX
...	 _index	filebeat-8.12.2-2024.06.05-
...	 _score	-
...	 _size	1,215

©2024. Synopsys, Inc. All rights reserved.

How to Size Elastic Events

Install the plugin

Self-Managed:

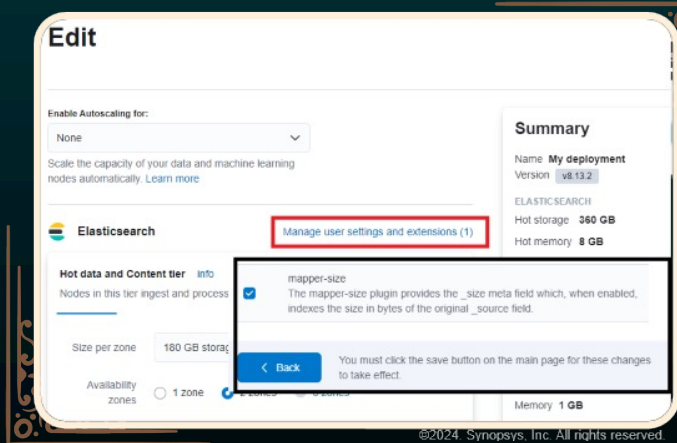
- Plugin must be installed on every node:
<https://www.elastic.co/guide/en/elasticsearch/plugins/current/mapper-size.html>
- Each node must be restarted:
<https://www.elastic.co/guide/en/elasticsearch/reference/current/restart-cluster.html>

```
[chanda@elastic1 elasticsearch]$ pwd
/usr/share/elasticsearch
[chanda@elastic1 elasticsearch]$ sudo ./bin/elasticsearch-plugin install mapper-size
→ Installing mapper-size
→ Downloading
[=====]
→ Installed m
→ Please rest
[chanda@elastic1 elasticsearch]$ sudo ./bin/elasticsearch-plugin install mapper-size
```

©2024 Synopsys, Inc. All rights reserved.

Elastic Cloud:

- Plugin installed as an extension:
<https://www.elastic.co/guide/en/cloud/current/ec-adding-elastic-plugins.html>
- No restart required.

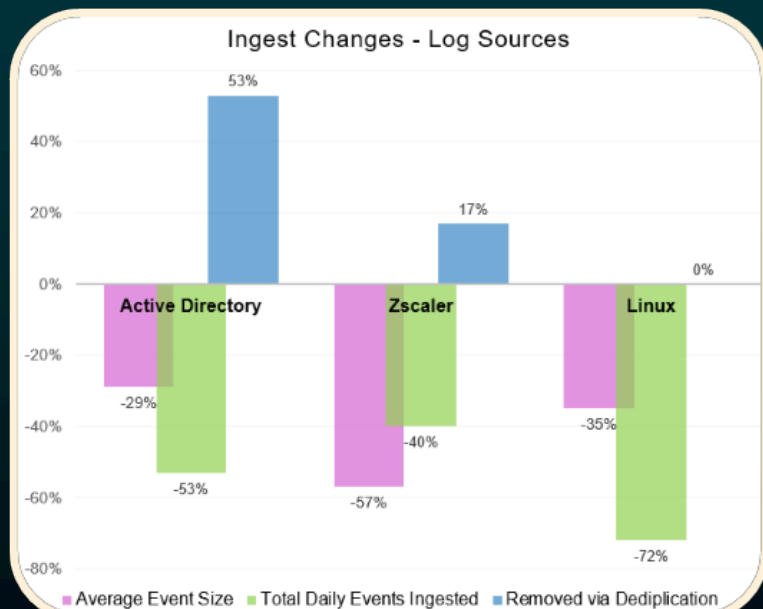


©2024 Synopsys, Inc. All rights reserved.

DEMO

What we were able to learn.

Here at Synopsys.

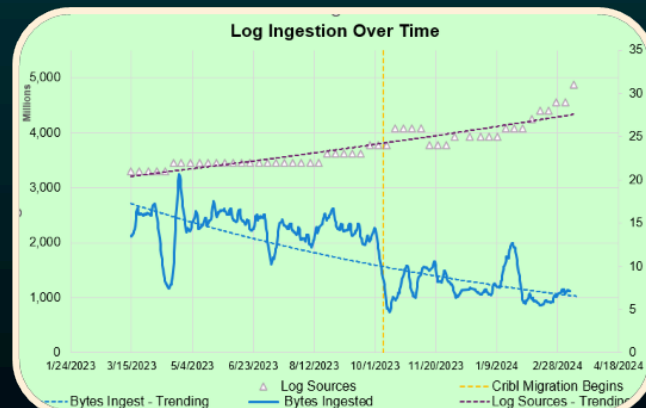


©2024. Synopsys, Inc. All rights reserved.

Potentially Reclaimed Resources	Storage (TB)
Hot Tier (5 days)	18.00
Warm Tier (60 days)	216.00
Cold Tier (300 days)	1,080.00
Total Storage Savings (TB)	1,314.00

*Assuming 10 TB/day ingest at 30% compression

©2024. Synopsys, Inc. All rights reserved.



©2024. Synopsys, Inc. All rights reserved.

Migrating from A to Zscaler

From Syslog, Filebeat, Logstash and chaos!



Tricky log source:

- Logs can only be shipped via TCP.
- Extremely high-volume logs.
- Original configuration not ideal.
 - 24-hour+ delays hitting SIEM.
 - Logs getting dropped.



The plan:

- Enable Cribl Syslog source.
- Redirect logs to Cribl.
- Parse Logs (Thank you, Zscaler Pack!).
- Send logs to SIEM.
- Go out for happy hour.

Migrating from A to Zscaler

The best-laid plans...

So what happened?

- TCP Pinning
- Backpressure!
- NSS Server hits capacity
- Logs start dropping!

What else did we try?

- ~~✗ • Load Balance TCP Traffic to multiple Cribl worker.~~
- ~~✗ • Zscaler to Syslog-ng to Cribl HTTP source.~~
- ~~✗ • Upgrade Syslog-ng to allow TCP multi thread processing.~~
- ✓ • Zscaler to Filebeat to Cribl Elasticsearch source.

Room for improvement:

- Cribl parsing too slow.
- Periodic parsing failures.
- Still experiencing 2-3 hour delay.

Migrating from A to Zscaler

Almost there!

Remove null or "None" values.

- Parser w/ Fields filter.
- 65.8ms per 100 events.
- Using Mask + Parser w/o filter.
- 9.88ms per 100 events.

The screenshot shows the configuration interface for a Mask and Parser. The Mask section is at the top, with a toggle for 'Mask' and a description 'Remove 'None' ... for to parsing' set to 'true'. Below it, the Parser section has a toggle for 'Parser' and a description 'Parse with fil... ve 'None' values' set to 'true'. The Filter section shows a filter set to 'true' with a description 'Parse with filter to remove 'None' values'. The Operation mode is set to 'Extract' and the Type is 'JSON Object'. The Source field is 'message' and the Destination field is 'Destination field na'. The List of fields is 'Enter field names'. The Fields to keep is 'Enter field names'. The Fields to remove is 'Enter field names'. The Fields filter expression is 'value !== "None"'. The interface includes 'Cancel' and 'Save' buttons at the bottom.

The screenshot shows the 'Pipeline diagnostics - TEST-parser' window. It has three tabs: 'Statistics', 'Pipeline Profile', and 'Advanced CPU Profile'. The 'Pipeline Profile' tab is active, showing a table with columns: Function, Bytes In, Bytes Out, Bytes Out-In, Events In, Events Out, and Process Time. The table lists two functions: '1. Mask (disabled)' and '2. Parser'. The 'Parser' function shows a process time of 65.80ms. A red arrow points to the 'Total: 65.80ms' at the bottom right of the table.

Function	Bytes In	Bytes Out	Bytes Out-In	Events In	Events Out	Process Time
1. Mask (disabled)	---	---	---	---	---	0.00ms
2. Parser	231.53KB	353.91KB	↑52.85%	100	100	65.80ms
SUMMARY	0.00B	353.91KB	↑Infinity%	0	100	Total: 65.80ms

©2024. Synopsys, Inc. All rights reserved.

©2024. Synopsys, Inc. All rights reserved.

Migrating from A to Zscaler

Going strong!

I comb

Help [?] Session

3. Eval

80.64KB

37.98KB

↓-52.90%

100

100

5.57ms

FAST!

9%

Final [?] No

Evaluate Fields [?]

Name [?]	Value Expression [?]	Enabled [?]
event	{"action": action, "outcome": status, "reason": a...	Yes [?]

Add Field

Keep Fields [?]

Enter field names

Remove Fields [?]

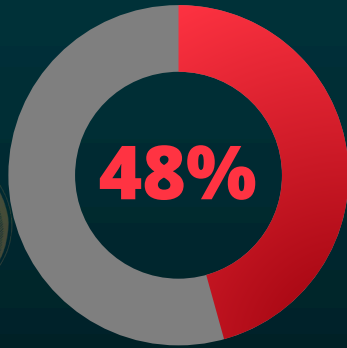
host X p4tress.cribl_breaker X _raw X action X

Migrating from A to Zscaler

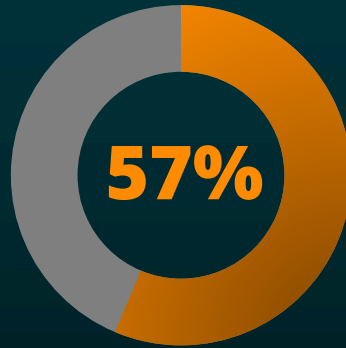
Final Verdict

- Used Filebeat to handle initial TCP-based workload
- Discovered some events dropping at NSS server level
- **Via Cribl :**
 - Removed unnecessary fields to decrease total event size
 - Used Suppression & Redis to drop duplicate events
 - Decreased parsing time per 1000 events from 2.3 seconds to 400 ms.

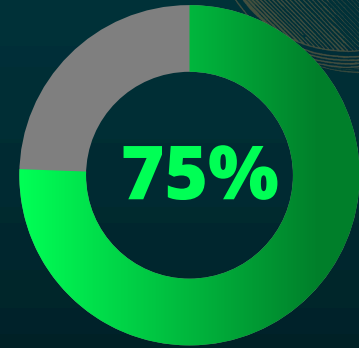
So how did we do?



48% fewer
Zscaler events
sent to SIEM.



57% average
smaller
event size.



75% less
storage required
for Zscaler logs.

Recommendations:

Do this, not that.



Understand TCP pinning.

- Cribl Syslog source now improved.



Get rid of the noise!

- Drop unnecessary fields and events.
- Focus on actual useful data.



Don't be afraid of Redis!

- Duplication reduction.
- Accurate aggregation.
- Easy tool to configure!



Thank you!

@Chanda Pulliam in Community Slack.