BLOG POSTS

	Content
Post Title	Should You Use an ELK Stack for Syslog?
Post Date	
Attributed To	Peter Welcher
Written By	Peter Welcher
Reviewed By (Name & Date)	Dave Donati (11/13)
Reviewed By (Name & Date)	

Meta Title (55 characters including spaces)	Should You Use an ELK Stack for Syslog?
Meta Description (156 characters including spaces)	
Target Keywords	
Categories	Technology
Tags	N/A
Call to Action	N/A
Image	Put a brief description of what the image should be or note that a file is attached. DO NOT paste the image into this Word doc; send it as a separate file.

Note: Naming convention for files as they go back and forth

- Original writer names file with "_V1" at the end (e.g., blogtitle_V1)
- First reviewer, makes edits and renames with initials at end (e.g., blogtitle_V1_af)
- If another reviewer, again add initials to end to keep the string of reviewers (e.g., blogtitle_V1_af_pw)
- When original writer gets it post back with edits, she makes revisions and saves the file as V2 (e.g. blogtitle_V2) – then reviewers continue as above with initials
- When post is complete, it is saved with "Final" and the post date at the end (e.g. blogtitle_FINAL_022012)

COPY FOR POST:

I've just posted a five (5) part blog tutorial on an ELK stack, with log data as the primary data explored. The last of the five blogs is here, and has links to its predecessors. The basis for those blogs was a pre-built ELK container, for simplicity. Do note that you can run ELK as one or more VM's as well.

Commented [PW1]: Add link before posting

If you don't know what an ELK stack is, please skim those first!

This blog considers using an ELK stack to manage syslog. It compares the ELK stack to the prior approach that Terry Slattery and I have been recommending for years, and that many sites use.

In particular, this blog compares pros and cons of an ELK stack versus syslog-NG!

By the way, if you don't centrally collect and monitor syslog, you're missing some good information, both for pro-active response and for alerts that you won't get any other way!

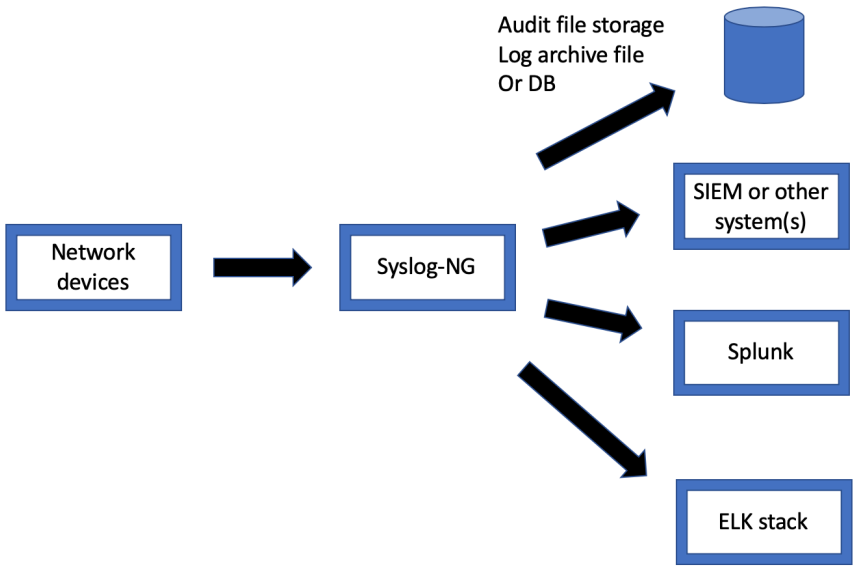
CURRENT SYSLOG SOLUTION

We've been recommending that people point network syslog at a Linux system running syslog-NG. Then use syslog-NG to archive audit trail records that won't be further processed or looked at unless needed, filter out "noise" (frequent messages of little value), and then multiplex the remaining syslog stream to various consuming products, including pagers and other forms of notification, with per-consumer filters as desired.

Even when sites have Splunk, you can use syslog-NG to reduce the volume of data going into Splunk, possibly reducing your licensing costs.

Syslog-NG has been evolving with the times. One thing to note is that syslog-NG can act as a filter that uses Elasticsearch (the 'E' in 'ELK') for storage.

Here's a diagram showing what we generally recommend:



By the way, many network management (NM) tools come with a syslog collector and display. Most of them are after-thoughts and rather pathetic. They can also consume CPU and disk on your NM tool server. Dedicated server(s) and storage for syslog are best.

I'm also a firm believer that you should never store syslog by device. It should all go into one log system, so that you can easily see time-correlated events surrounding a performance problem or outage.

ELK Stack for Syslog?

If you already have Splunk, using an ELK stack is probably of little interest.

You might be interested in an ELK stack if you're not fond of Splunk and its costs, or are not heavily using Splunk's reporting capabilities.

If you're using a network management tool for syslog, and its syslog support rather limited, or if you're using syslog-NG, writing syslog to file, and using scripts or occasional manual scans to monitor syslog, then an ELK stack might help you up your game.

Ownership and control might also be relevant, if Splunk control, permissions, or other issues keep it from meeting your needs.

If you think an ELK stack might be useful, bear in mind that ou would need budget and time to stand up an appropriately-sized ELK server, VM, or cluster and related storage.

If your organization is already using an ELK stack for app / server functions, then having one tool might be attractive, although ownership / scaling / cost sharing / admin access might be factors. The good news is that if you will bring your own hardware (or budget) to run on, there might be in-house resources to provide advice!

Comparison

There are several factors to consider when deciding syslog-NG versus ELK stack. Here are those that come to mind:

Complexity and Learning Curve

Syslog-NG seems a bit simpler: fewer moving parts. When I've looked at it in the past, it seems to be well documented.

Concerning ELK, there's a lot of web resources, the documentation is passable, not bad, maybe a bit thin in spots, left me scratching my head a few times. Of course, the problem there might be my head, not the documentation.

Platform and Support Costs

Most people start syslog-NG with a Linux server and a fair-sized disk drive, and play it by ear from there.

Some considerations: how robust does your syslog system have to be? Dual servers and disks, versus RAID and more costly storage? Backed up? How long do you want to retain the syslog data?

Most sites I've discussed this with aren't too worried about loss of syslog data. As far as keeping the data, audit trail data that is separated out might have to be retained. Other than that, maybe 2-3 months' worth of syslog data is plenty. It can be handy to compare or do trending on volume of alerts, broken out by alert code.

On the ELK side, I'd probably consider starting with installing the ELK stack on a single server with big disk, as with syslog-NG. In researching ELK sizing, I found one expert whose comment was roughly that he's never gotten it right the first time — although that was for much larger systems. My interpretation of that is that it may be best to do something, get sizing data based on your actual syslog volume and other needs, and then adjust if necessary.

This all changes if you work for a large organization. Up-front sizing may be more of an issue. Also, if your organization is large and / or wants a more robust solution, clustered ELK might be appropriate. If the organization is already doing Docker / Kubernetes, you might be able to get assistance bringing up clustered ELK.

Filtering capabilities

This seems like a wash, but I lack hands-on data. Note: organizations rarely hire us as consultants to set up syslog monitoring for them, they tend to do it in-house, but often lack time to do it well. That's where bringing in someone to up your network management and syslog game might be useful.

Parsing

Syslog-NG and Logstash can both transform messages in different formats. I will note that the massive use of log data in the server / app world, plus a much larger market than networking, suggests that ELK may benefit from a larger user base and community.

Storage

Both syslog-NG and Logstash can store to Elasticsearch! And that really suggests that the comparison or positioning should be syslog-NG versus Logstash. This comparison, however, is based on the notion the syslog-NG and ELK are syslog gathering and reporting systems.

Multiplexing capabilities

Both seem good for multiplexing the log streams. I'd think forwarding to a file and forwarding to another service via a network connection would be the two big capabilities, and both support that. How flexibly and how many options they support certainly might be something to consider, which I have not delved into.

Log tail monitoring

Kibana is a nice way to watch the event stream, and its ad hoc query capability just adds to that value!

Reporting

Clear win for the ELK stack. You can build tail functionality into a dashboard, with a severity level filter for only the more severe events. The dashboard can also provide bar charts etc. (number of messages at the various severity levels, top N messages, IP addresses, and interfaces per the blog series).

Elasticsearch can perform other reporting, as shown in the prior pivot table blog.

Queries

I give ELK a big win for the ability to use the Kibana Discover functionality to quickly do ad hoc queries and generate visualizations.

Syslog-NG doesn't do queries per se. It does let you store data into a SQL database. While I've committed SQL in the past, one of my life goals is to do as little SQL as

possible. I'd say ELK wins hands-down, in terms of ease of generating filters and extracting data for simple queries.

Sizing ELK

I haven't found a simple answer.

If you're doing simple syslog-NG, you can just use ordinary syslog to capture say one day's worth of syslog and do the math. With ELK, a similar approach might help you get started. Volume of events (average and peak messages per second) and storage consumption are the two obvious driving factors. Events per second impacts CPU, and total bytes of messages impacts storage. CPU scales by bigger processors then by adding cluster members. With ELK, there is in-memory indexing for fast queries, so storage not only impacts disk space, but RAM.

I've included some reference links below.

If you're looking for basic syslog functionality like most sites do with syslog-NG, sizing may not be a big problem. Put differently, unless your requirements drive a need for ELK clustering, don't go looking for trouble.

You might also consider getting professional help!

Conclusion

My personal feeling here is that the Kibana Discovery display is a win, and graphics (visualizations) on top of that is a major plus.

If you like your syslog-NG setup, especially if you're getting paging, alerts, audit logging, etc. from it, then you might consider having it feed the ELK stack as storage / query engine and reporting tool.

If you have Splunk, my guess is that you probably haven't read this far, due to lack of interest in ELK stack for syslog. Splunk is a powerful toolset. Not having to manage the datastore and being able to leverage others' Splunk skills are other potential benefits.

References

Syslog-NG:

- https://www.syslog-ng.com

ELK Sizing:

- https://thoughts.t37.net/designing-the-perfect-elasticsearch-cluster-the-almost-definitive-guide-e614eabc1a87
- http://alexander.holbreich.org/elasticsearch-configuration/

COMMENTS

Comments are welcome, both in agreement or constructive disagreement about the above. I enjoy hearing from readers and carrying on deeper discussion via comments. Thanks in advance!

Hashtags: #CiscoChampion #TechFieldDay #TheNetCraftsmenWay

Twitter: @pjwelcher

Disclosure Statement

[INSERT the usual IMAGES HERE: 20 Year CCIE and Cisco Champions **2019** as per recent blogs]

NETCRAFTSMEN SERVICES

Did you know that NetCraftsmen does network /datacenter / security / collaboration design / design review? Or that we have deep UC&C experts on staff, including @ucguerilla? For more information, contact us at <<insert suitable link here>>.

SOCIAL MEDIA:

Facebook: Like, comment or share our status using this link.

Twitter: Like and RT our tweet using this link.

LinkedIn: Like, comment or share our status using this link.