

Tying together Zabbix and Elasticsearch/Logstash/Kibana (ELK) ... and Grafana, too!

Volker Fröhlich

19 Nov 2015, NLUUG

Who am I?

- Volker Fröhlich (volter)
- Geizhals Preisvergleich Internet Services AG
(<http://geizhals.at>)
- Zabbix frontend patches, conference, blog, book review
- Fedora packager, Openstreetmap contributor



What is this all about?

- 1 How logs are interesting and difficult

What is this all about?

- 1 How logs are interesting and difficult
- 2 Define what we want to achieve

What is this all about?

- 1 How logs are interesting and difficult
- 2 Define what we want to achieve
- 3 Explain the setup I am using

What is this all about?

- 1 How logs are interesting and difficult
- 2 Define what we want to achieve
- 3 Explain the setup I am using
- 4 How we can integrate it tighter

What logs can contain

- Operational messages
- Performance data
- Events
- Error messages, crashes
- Debugging information

Apache access log

```
10.0.0.137 - - [06/Nov/2015:01:01:07 +0100]  
"GET / HTTP/1.1" 200 33771  
"http://www.geizhals.at/"  
"Mozilla/5.0 (X11; Linux x86_64)  
AppleWebKit/537.36 (KHTML, like Gecko)  
Ubuntu Chromium/45.0.2454.101  
Chrome/45.0.2454.101 Safari/537.36"
```

- Message written to a file directly
- Custom timestamp, free-formish strings

Postfix

```
Nov  7 06:59:40 mailserver postfix/smtpd[29789]:  
C690912483F1: client=example.com[10.1.1.1]
```

```
Nov  7 06:59:59 mailserver postfix/smtp[32571]:  
C690912483F1: to=<root@geizhals.at>,  
relay=127.0.0.1[127.0.0.1]:10024, delay=18,  
delays=0.05/0.03/0/18, dsn=2.0.0,  
status=sent (250 2.0.0 Ok, id=26552-28,  
from MTA([127.0.0.1]:10025): 250 2.0.0 Ok:  
queued as 3155B1248447)
```

- A different timestamp format
- Syslog context
- Some timing information
- Queue ids that connect related messages

Cisco ASA

```
%ASA-1-105006: (Primary) Link status Up  
on interface interface_name.
```

```
%ASA-7-713204: Adding static route for  
client address: IP_address
```

interface_name and IP_address are placeholders

Apache 2.4 error logs

```
AH00940: %s: disabled connection for (%s)"  
AH01408: Zlib: %d bytes of garbage at the  
        end of "compressed stream."
```

IP tables

```
Oct  4 01:14:19 debian kernel: IN=ra0 OUT=  
MAC=00:17:9a:0a:f6:44:00:08:5c:00:00:01:08:00  
SRC=200.142.84.36 DST=192.168.1.2 LEN=60  
TOS=0x00 PREC=0x00 TTL=51 ID=18374 DF PROTO=TCP  
SPT=46040 DPT=22 WINDOW=5840 RES=0x00 SYN URGP=0
```

Mostly key/value, but not completely!

Jira backtrace

```
2015-11-07 01:11:00,026 Sending mailitem To='user@example.com' Subject='Some subject'
From='null' FromName='null' Cc='null' Bcc='null' ReplyTo='null' InReplyTo='null'
MimeType='text/plain' Encoding='UTF-8' Multipart='null' MessageId='null' ERROR anonymous
Mail Queue Service [atlassian.mail.queue.MailQueueImpl] Error occurred in sending e-mail:
To='user@example.com' Subject='Some subject' From='null' FromName='null' Cc='null'
Bcc='null' ReplyTo='null' InReplyTo='null' MimeType='text/plain' Encoding='UTF-8'
Multipart='null' MessageId='null'
    com.atlassian.mail.MailException: javax.mail.SendFailedException: Invalid Addresses;
    nested exception is:
    com.sun.mail.smtp.SMTPAddressFailedException: 550 5.1.6 <user@example.com>:
    Recipient address rejected: User has moved to somewhere else.
    For more information call Example at +43 123123 or e-mail info@example.com

    at com.atlassian.mail.server.impl.SMTPMailServerImpl.sendMessageId(SMTPMailServerImpl.java:44)
    at com.atlassian.mail.queue.SingleMailQueueItem.send(SingleMailQueueItem.java:44)
    ...
```

What do we want to achieve?

- 1 Solve real-world problems

What do we want to achieve?

- 1 Solve real-world problems
- 2 Keep it simple

What do we want to achieve?

- 1 Solve real-world problems
- 2 Keep it simple
- 3 Collect in one place

What do we want to achieve?

- 1 Solve real-world problems
- 2 Keep it simple
- 3 Collect in one place
- 4 Search and analyze

What do we want to achieve?

- 1 Solve real-world problems
- 2 Keep it simple
- 3 Collect in one place
- 4 Search and analyze
- 5 React upon things automatically

What do we want to achieve?

- 1 Solve real-world problems
- 2 Keep it simple
- 3 Collect in one place
- 4 Search and analyze
- 5 React upon things automatically
- 6 Improve our current monitoring system

What is Zabbix?

- Classic monitoring system
- Relation database backend for config and data
- Mostly C and PHP
- Server, proxy, agent
- Has complex concepts; Permission model
- Item, trigger, event, action, operation, ...
- Supports trapping mechanisms
- Versatile, but weak with visualization
- SOAP JSON API
- Can be extended and hacked



Zabbix 3.0 frontend

Monitoring
Inventory
Reports

Dashboard
Overview
Web
Latest data
Triggers
Events
Graphs
Screens
Maps
IT services

Dashboard

Favourite graphs

Zabbix.org: MySQL queries per second

Zabbix.org: Users in #zabbix on freenode (3h forecast)

Electricity in Latvia: Households w/o electricity in Latvia

Zabbix.org: Free disk space on / (and timeleft)

Zabbix.org: Zabbix busy poller processes

Zabbix.org: Used diskspace

Zabbix.org: Disk usage

Zabbix.org: Disk read/write bytes per second

Zabbix.org: Memory usage

Zabbix.org: Network utilisation

Zabbix.org: CPU utilisation

Zabbix.org: Amount of unsupported items

Zabbix.org: Cached memory

Graphs

System status

HOST GROUP	DISASTER	HIGH	AVERAGE	WARNING	INFORMATION	NOT CLASSIFIED
Linux servers	0	1	1	1	0	0
Misc	0	0	0	0	0	0
Pseudo-hosts	0	0	0	0	0	0
Web pages	0	0	0	0	1	0

Updated: 02:43:25

Last 20 issues

HOST	ISSUE	LAST CHANGE	AGE	INFO	ACK	ACTIONS
OpenStreetMap.org	Testtrigger	2015-11-13 16:48:51	5d 9h 54m		No	
Zabbix.org	Lack of free swap space on Zabbix.org	2013-11-05 09:12:27	2y 13d		Yes	
Zabbix.org	Inetd is not running on Zabbix.org	2013-11-05 09:12:13	2y 13d		Yes	
Zabbix.org	Some items are not supported	2012-04-16 21:50:08	3y 7m 6d		Yes	

4 of 4 issues are shown

Updated: 02:43:26

Favourite maps

Local network

Maps

Web monitoring

HOST GROUP	OK	FAILED	UNKNOWN
Linux servers	2	1	0
Web pages	1	0	0

Updated: 02:43:26

Favourite screens

Zabbix translation status

empty

Zabbix.org

Zabbix server

Web monitoring

Screens Slide shows

Host status

HOST GROUP	WITHOUT PROBLEMS	WITH PROBLEMS	TOTAL
Linux servers	1	1	2
Misc	2	0	2
Pseudo-hosts	1	0	1
Web pages	1	1	2

Updated: 02:43:26

How can we solve the transport problem?

Challenges

- There are many different sources and devices we should cover
- We must not stall operations
- We should not lose a lot of messages

Possible solutions

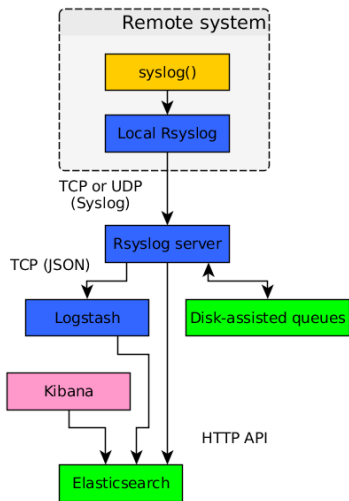
- A transport abstraction layer like fluentd
- Some special agent and shipping protocol
- Process on the host and store remotely
- Some messaging system (Kafka, ...)
- Zabbix

Why don't we use Zabbix' capabilities?

- Needs an agent, an active one even!
- Is file-based (efficiency, permissions)
- Can only grab complete lines or one single value
- Is not very flexible with date formats
- Is exclusively POSIX-regex-based
- Can not be graphed, except for those single values
- Can not be searched through
- Becomes even less interactive and sufficient when crossing hosts

Why not syslog?

- Syslog is ubiquitous
- Syslog has limitations
- 90% of them are probably irrelevant for you or can be worked around
- No new technologies, easy to set up
- Little resource consumption, robust



What is Rsyslog?

- Journald?
- Modern syslogd implementation
- TCP, RELP
- Queues
- Supports various output modules
- Also exists for Windows
- Structured logging? CEE-enhancement!

```
Nov 17 12:37:31 x250 volker: @cee:{"key":value,  
    "key2":"utf-8", "key3":{"subkey":value}}
```

What is Logstash (LS)?

- JRuby-based "processing pipe"
- File based configuration with if-clauses
- Input – tcp
- Codec – json_lines
- Filter – grok, kv, csv, geoip, ...
- Output – elasticsearch, zabbix
- JSON



What is Elasticsearch (ES)?

- Java-based document storage
- Built on Lucene
- Meant to easily scale horizontally
- No pre-configured schema necessary
- REST HTTP JSON API
- Permissions can be difficult

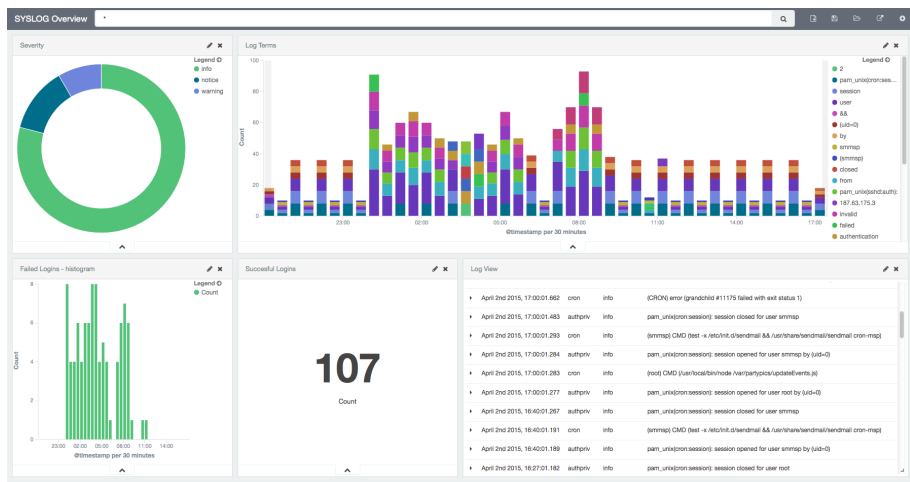


What is Kibana (4)?

- NodeJS-based web frontend
- Only data source is ES
- Allows to search with Lucene queries
- Exposes some of ES' capabilities
- Attempts to break request length limits
- Has no permission model



Example Kibana dashboard



Graylog2, Heka, Splunk?

Graylog2

- Java and NodeJS
- Offers processor and frontend
- Offers live configuration changes and streams
- Offers an API and stats
- Uses ES as the backend

Heka

- Go and Lua

What is Grafana?

- Web time series graphing solution
- Go and NodeJS
- Various data sources, including ES, from 2.5 on
- grafana-zabbix by Alexander Zobnin
- Highly customizable graphs
- Templated and scripted dashboards
- Has a permission model



Example Grafana dashboard



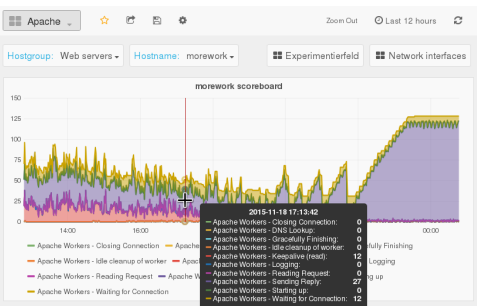
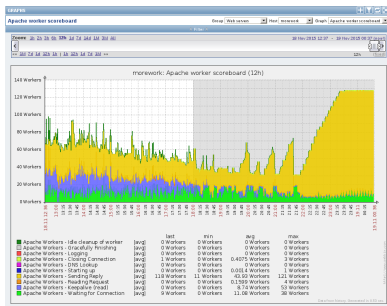
What can be done?

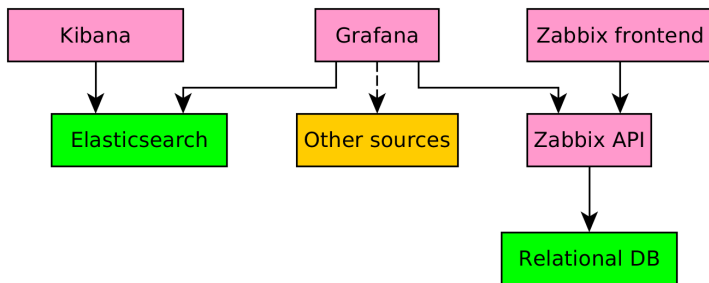
- 1 Graphing things together
- 2 Navigating with context
- 3 Tagging logs with Zabbix context
- 4 Sending data from LS
- 5 Polling data from ES
- 6 Sending Zabbix events to LS
- 7 Sending deployment events to LS
- 8 Zabbix daemon logs

Graphing things together

- Shortcomings in Zabbix graphing and screens
- Kibana only supports ES
- Grafana has a plugin for ES and Zabbix
- None of the three offers a complete sub-set of another
- It is not a trivial task to "include" one into another

Zabbix versus Grafana

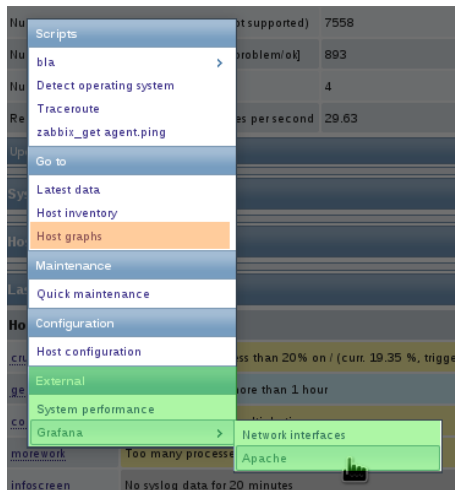




Navigating with context

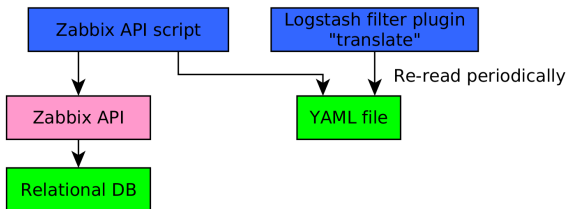
- No interface can handle all your needs
- Make it easy to navigate between frontends
- Use and extend the Zabbix JS menu
- Use templated and scripted dashboards in Grafana

Early stage of JS menu navigation













































Tagging logs with Zabbix context

- Assume that Zabbix host groups are relevant
- Optionally ignore some of them
- Periodically poll host group data from API
- Use LS "translate" filter plugin
- http://zabbix.org/wiki/Tagging_logstash

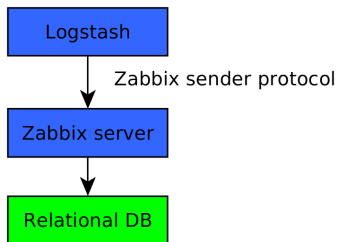


Zabbix host groups added

t @source_host	   morework
🕒 @timestamp	   November 19th 2015, 01:10:25.000
t @version	   1
t _id	   AVEdE_RDsB7YtUgQ1fs5
t _index	   logstash-2015.11.19
t _type	   logs
t component	   qmgr
t message	   AD6A81248430: removed
t processid	   12053
t qid	   AD6A81248430
t syslog_facility	   mail
t syslog_program	   postfix/qmgr
t syslog_severity	   info
t zbx	   Web services, Linux servers

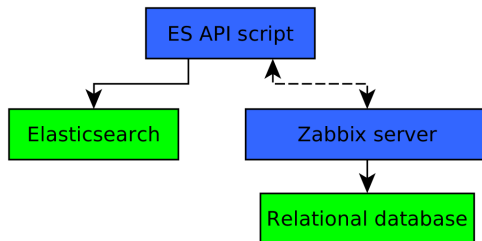
Sending data from LS

- LS output plugin "zabbix"
- Implements Zabbix sender protocol
- Allows to submit arbitrary data on arbitrary events
- You must know the Zabbix host name
- You must know the key of an existing trapper item
- No fallback item?
- Create a trigger with "multiple problem events" and hysteresis?



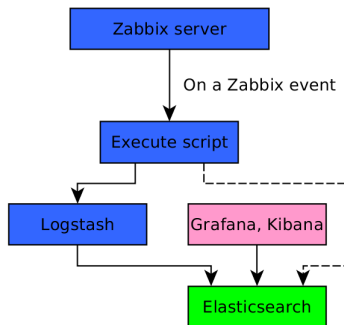
Polling data from ES

- Query using the ES HTTP API
- Write a script that accepts a reference to a JSON object
- Set up an according "Simple script" item
- Set up a trigger



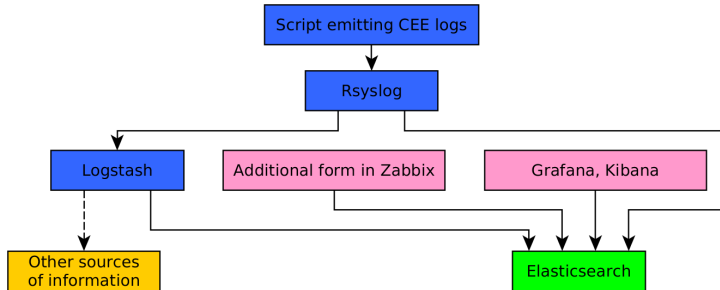
Sending Zabbix events to LS

- Set up a custom script
- Set up an action
- Neither Kibana 4 nor Zabbix can visualize them
- None of the systems is offering Gantt charts



Sending deployment events to LS

- Free-form information with Zabbix context from UI
- Or deployment hook elsewhere
- Neither Kibana 4 nor Zabbix can visualize them
- http://zabbix.org/wiki/Docs/comment_for_logstash

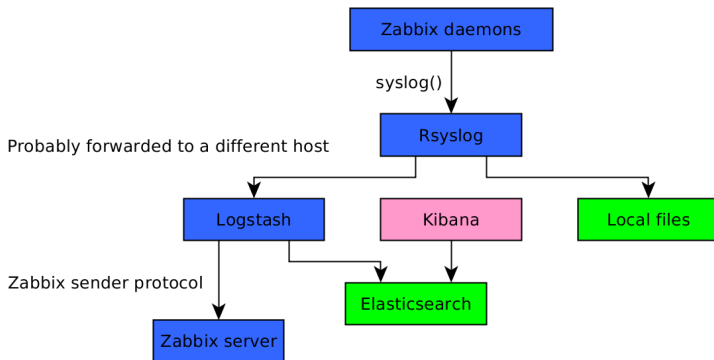


Event markers in Grafana showing Git commits



Zabbix daemon logs

- Don't set a log file
- Set up syslog daemon and log rotation
- Could we have monitored Zabbix logs with Zabbix?
- Works for all components, except JMX gateway?



Summary and outlook

- Great benefit
- Great potential for improvement
- Tests, automatisms
- Will everything become easy soon?
- Will any single interface be enough?
- Do we need a meta-interface?

Contact information and readings

- volter on Freenode IRC
- volker.froehlich@geizhals.at

Resources

- #zabbix, #logstash, #elasticsearch, #kibana, #grafana
- <http://www.zabbix.org>
- <https://github.com/alexanderzobnin/grafana-zabbix>
- <http://www.logstashbook.com>
- <https://github.com/coolacid/GettingStartedWithELK>
- http://geofrogger.net/zabbix_elk_nluug.pdf