CentralStation

zentralisierte Logsysteme für Rechenzentren
Aufbau, Fallgruben, Erfahrungen

SLAC, 26.06.2015, Berlin

Markus Manzke | 8ack GmbH

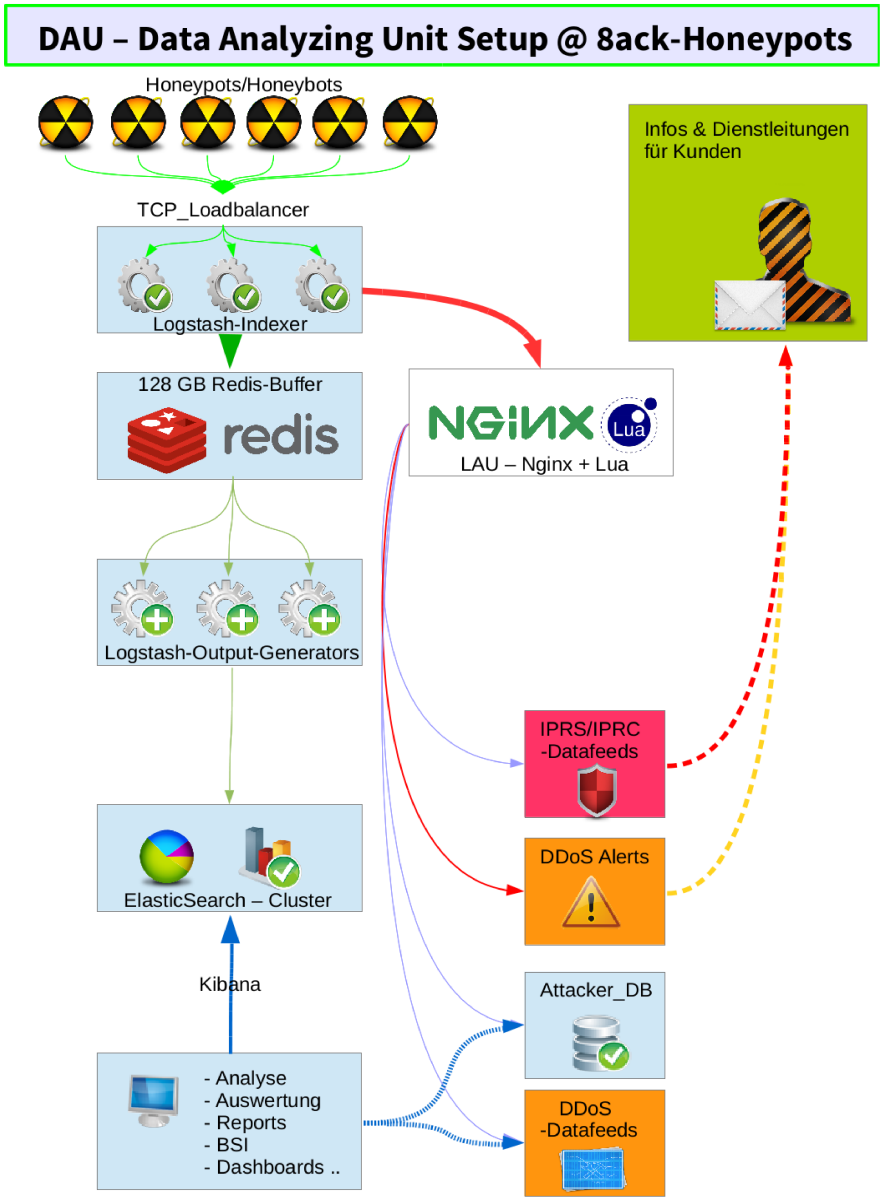
- Tools-Download
  https://8ack.de/slac-2015 (redirect zum Repo)

- Beispiele, Links etc

- Kontakt für PDF und Tools
  mm@8ack.de

Wo geht's hin?

- Log-Analyse-Setup @ 8ack
- 1 Mrd Events/Tag
  11.000 Events/Sek
- 1 TB Daten/Tag
- ELK-Stack + viel Kit



Warum zentrale Logauswertung mit ELK?

- Korellation von Events Debugging
- grafische Analyse von Trends
- Schnelle Volltextsuche Nadel im Heuhaufen
- Dashboards Dashboards Dashboards

- Individuelle Alarmsysteme

Logshipper

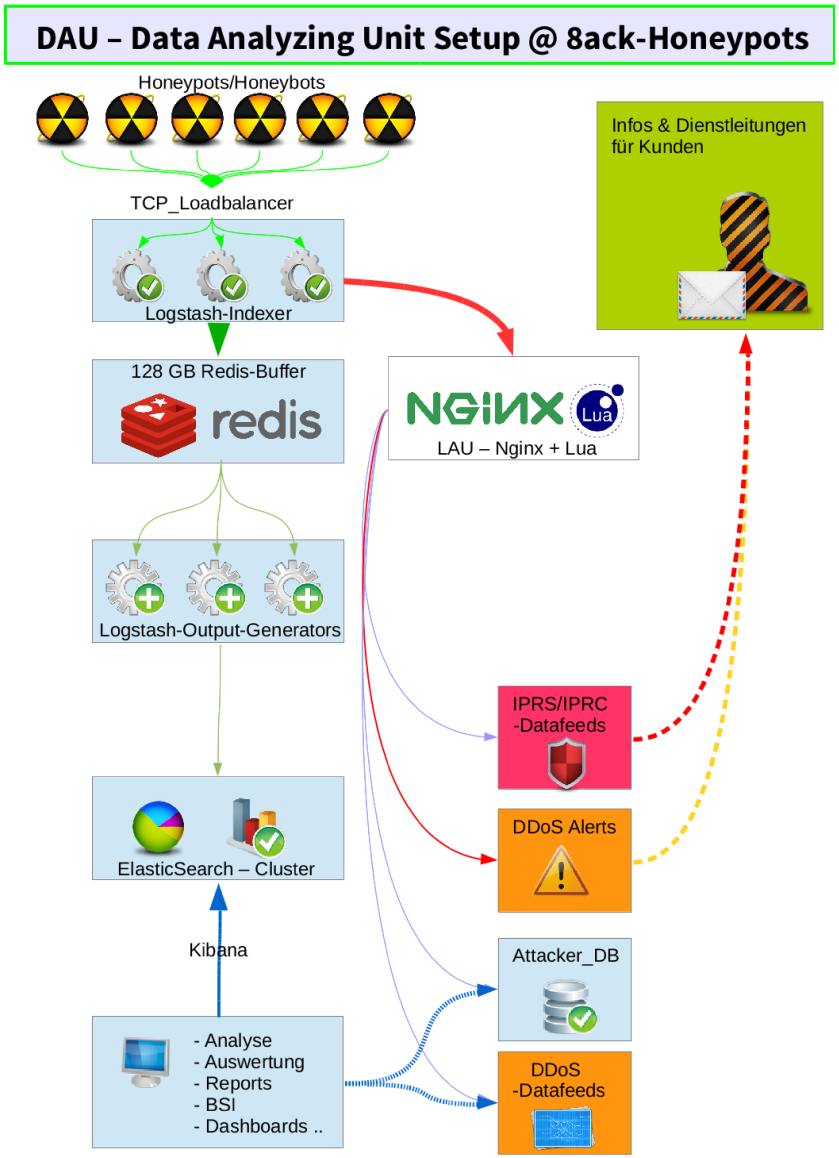
Indexer

Puffer

Output-Generator

Datenbank

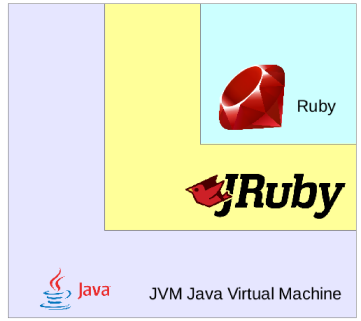
Analysetool



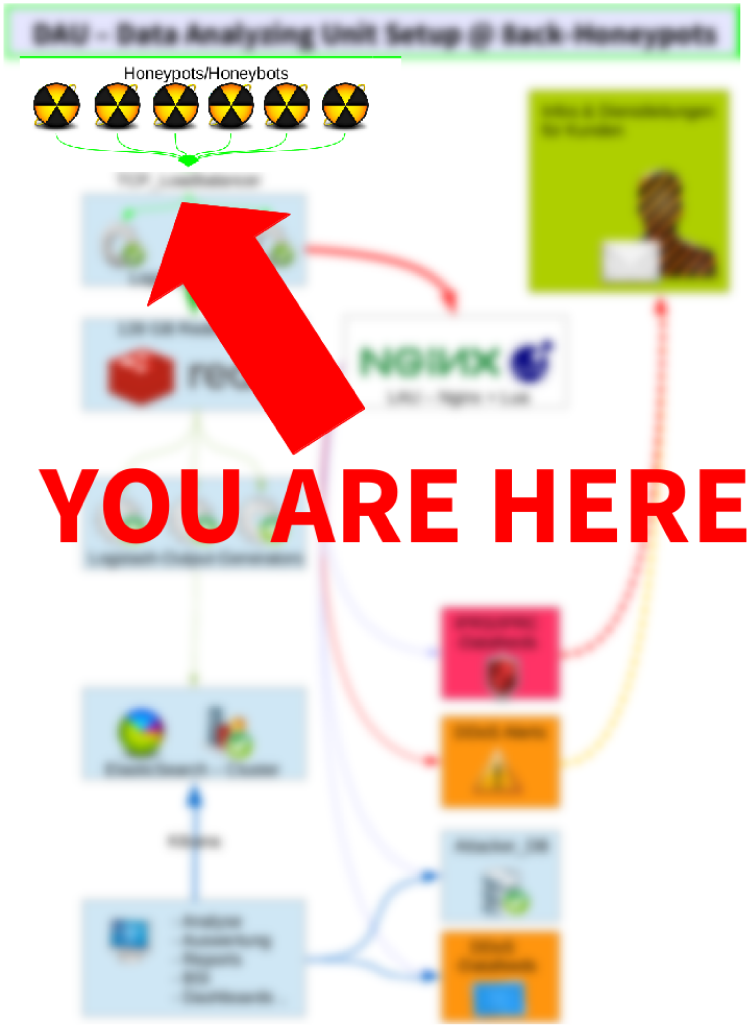
- Logshipper:
  Log-Courier
  https://github.com/driskell/log-courier
- Indexer
  Logstash
  https://www.elastic.co/

- Kit:
  Redis
- Database
  ElasticSearch
  https://www.elastic.co/

- Dashboards
  Kibana
  https://www.elastic.co/
- Layers
  on top of Layers
  on top of Layers



- Log-Courier
  - transportiert Logeinträge zu den Indexern
  - überwacht Logfiles
  - Logrotation
  - Monitoring- Schnittstelle Nagios-Plugin
  - Tagging
  - Beispiel + Live
  - Puppet
  - Alternativen



- Indexer
  - Loadbalanced
  - Input
  - GROK
    https://grokdebug.herokuapp.com
  - Output
  - Filter

Indexer – Inputs und Outputs

- couchdb_changes
- drupal_dblog
- elasticsearch
- exec
- eventlog

- file
- ganglia
- gelf
- generator
- graphite
- github
- heartbeat
- heroku

- irc
- imap
- jmx
- kafka

- log4j
- lumberjack
- meetup
- pipe
- puppet_facter

- relp
- rss
- rackspace
- rabbitmq
- redis
- snmptrap
- stdin
- sqlite
- s3
- sqs
- stomp
- syslog
- tcp
- twitter

- unix
- udp
- varnishlog
- wmi
- websocket

- xmpp
- zenoss
- zeromq

- boundary
- circonus
- csv
- cloudwatch
- datadog
- datadog_metrics

- email
- elasticsearch
- exec
- file
- google_bigquery
- google_cloud_storage
- ganglia
- gelf
- graphtastic
- graphite

- hipchat
- http
- irc
- influxdb
- juggernaut
- jira

- kafka
- lumberjack
- librato
- loggly
- mongodb
- metriccatcher

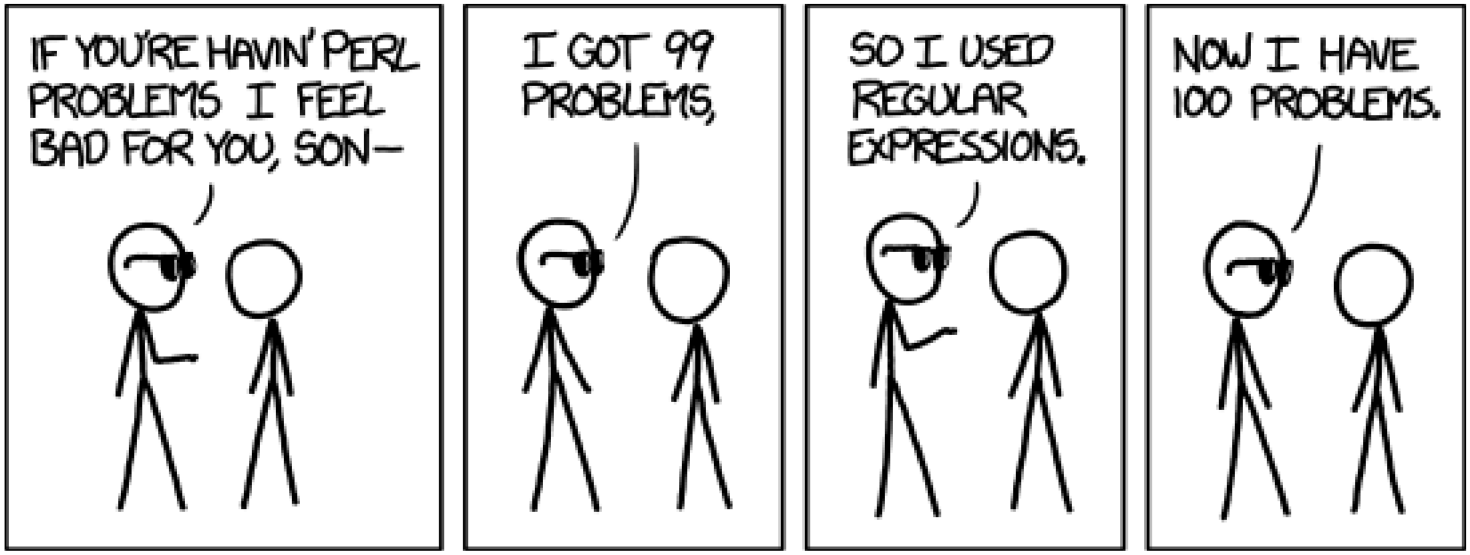
- nagios
- null
- nagios_nsca
- opentsdb
- pagerduty
- pipe

- riemann
- redmine
- rackspace
- rabbitmq
- redis
- riak
- s3
- sqs
- stomp
- statsd
- solr_http
- sns
- syslog
- stdout
- tcp

- udp
- websocket
- xmpp

- zabbix
- zeromq

- GROK: Logstash-Filter zum Zerlegen der Logeinträge
- REGEX-ähnliche Syntax
- Test Early, Test Often
- GROK-Pattern können zum Bottlenek werden
- Beispiele



- Loadbalancing
- Tuning
  - 4 GB RAM / IDX
  - Limits anpassen
- 5% Ausfall am Tag
- Traffic/Durchsatz kontrollieren
- Monitoring
- Redis-Puffer
  12-24h

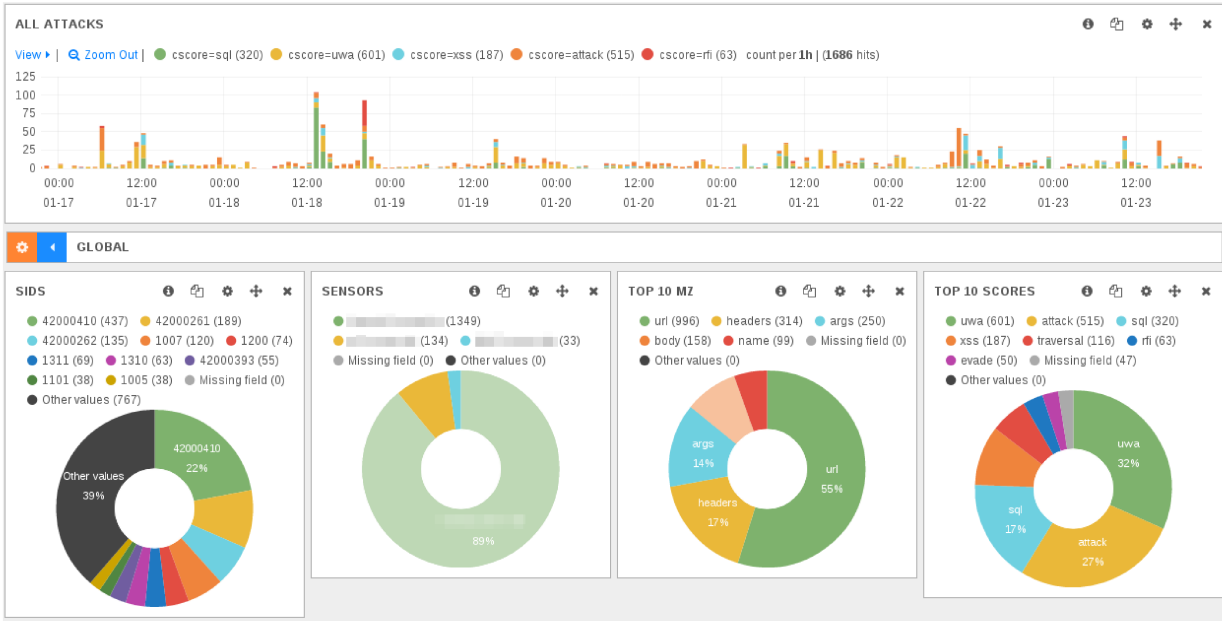
- Probleme bei Peaks (JVM?)

- Input:
  Lesen aus dem Redis-Puffer
- Keine GROK – Filter
- Output:
  eine oder mehrere ES-Instanzen
- 10 IDX ~ 1 OUT

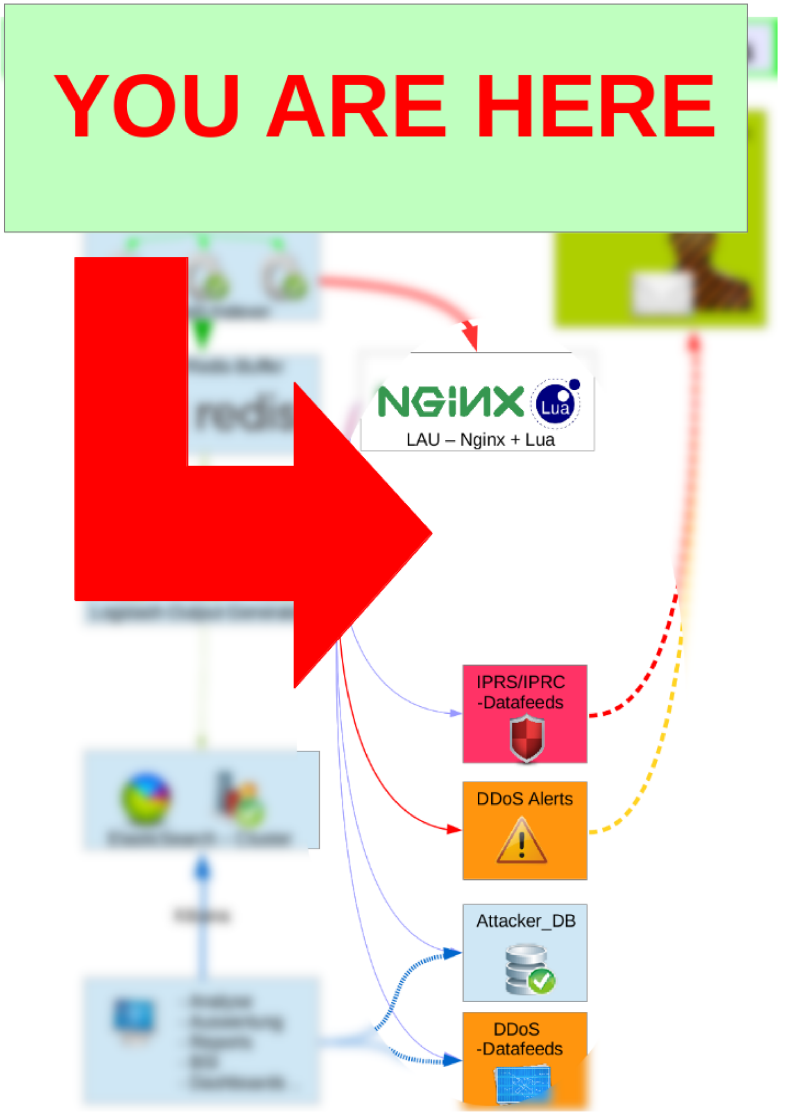
- ES → Cluster Build-In
- Nodes
  - Master
  - WorkHorse
  - Search
- Limits
  - 32 GB RAM
  - Verläßlichkeit
  - Platz
- REST-API

- Tuning
  - Loadbalancing Hot-Standby-Master
  - RAM-O-RAMA
  - Schnelle HDs
  - Limits hochsetzen
  - Setup nach Last
  - Search-Nodes

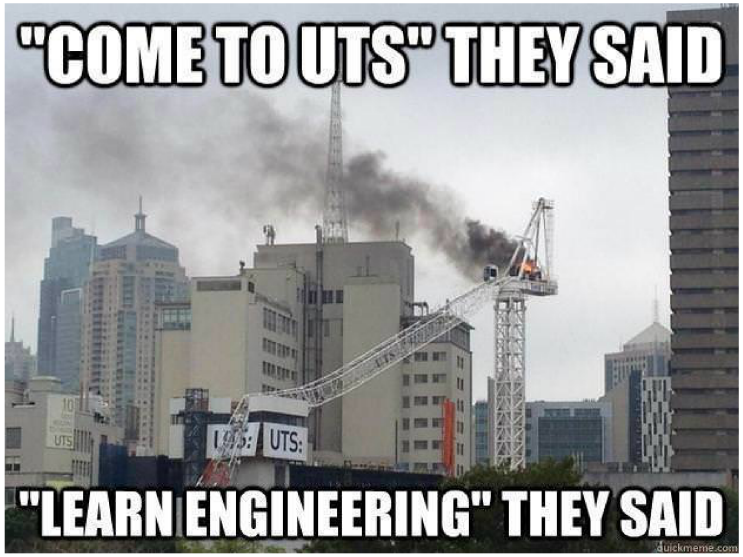
- Kibana
  3.x – Angular-App (statisch)
  4.x – Node.js (eigener Server)
- Klick- und speicherbare Dashboards für ES
- grafische Analyse
- keine Alarm-funktion



- NGINX
- LUA
- GO-Worker
- customized $Zeug



- Erst den Stack aufbauen
  dann ausrollen,
  Step-by-Step
  Multi-Instanz-LS-Setup
- Langsam in Schritten hochskalieren
- Loadbalancing (LS)
- Clustering (ES)
- Puffer als Auffangbecken
- Monitoring



- ES + Kibana absichern
  - IP-Restriktion
  - DMZ
  - ReverseProxy + AUTH
- Offen per default
  - Listen 0.0.0.0
  - Keine Authentifizierung
  - sensitive Informationen sehen oder manipulieren
  - RCE - Kandidat
- LS + Redis absichern
  - IP-Restriction
- Lässt sich sehr einfach DoSen

Limits

- Offiziell 5000 Eevents/s, bei uns 1000 Events/s
- LS bricht häufig weg
- ES bleibt manchmal stehen

- LS kann ES überfluten

- Workarounds:
  - Viel hilft Viel: mehr Blech Monitoring Loadbalancing ES-Cluster

- Redis als Puffer