Visualization of dCache accounting information with state-of-the-art Data Analysis Tools

Tigran Mkrtchyan
for DESY dCache operating Team
Outline

Tigran Mkrtchyan | Visualization of dCache accounting information | Date | Page 4
The Flow

Collector

Parser

Processor

Selector

Visualizer
The Flow (typical)

Collector

Parser

Selector

Visualizer

cat | awk | grep | gnuplot
Result

Bytes Read (365 days since Thu Mar 06 00:00:00 CST 2014)

- **Gbytes/day**
  - 10^6
  - 10^5
  - 10^4
  - 10^3
  - 10^2
  - 10^1
  - 10^0
  - 10^-1

**Time**

- April 2014
- May
- June
- July
- August
- September
- October
- November
- December
- January 2015
- February
- March

**Legend**
- DCache Reads
- HSM Reads
Scaling problems

> ~20GB billing files/day
> ~50.000.000 records/day
  - ~500 records/sec
> 4 dCache instances
> need to adopt scripts for different needs
> need for a 'State at Glance'
The Flow

Collector

Parser

Processor

Selector

Visualizer

logstash

elasticsearch

kibana
Logstash

> Collect logs from any source
> parse them
> gets the right timestamp
> index them
> and move it into a central place
input {
  # read log events
}

filter {
  # parse, fix formats, mutate
}

output {
  # store processed events
}
Logstash, single liner

```bash
$ echo "hello logstash" | logstash -e 'input { stdin{} } output { stdout {codec => rubydebug} }'

{
    "message" => "hello logstash",
    "@version" => "1",
    "@timestamp" => "2015-03-06T22:49:37.797Z",
    "host" => "dcache-lab"
}
```
Parse

filter {
  grok {
    match => [ "message", "%{TRANSFER_CLASSIC}" ]
    remove_field => [ "message" ]
  }

  date {
    match => [ "billing_time", "MM.dd HH:mm:ss" ]
    timezone => "CET"
    remove_field => [ "billing_time" ]
  }
}

Parse

> Regexp like syntax
> Lot of ready patterns for common cases
> supports labels and types
[00009A23BB6D280F46A7A6C12AC67F5EA897,59419220]
[003800000000000000559888,46305280]

PNFSID_NEW (?:[A-F0-9]{36})
PNFSID_OLD (?:[A-F0-9]{24})
PNFSID %{PNFSID_OLD}|%{PNFSID_NEW}

PNFSID_SIZE [%{PNFSID:pnfsid},%{NONNEGINT:size:int}]
Parser, example

\{DCap-3.0,131.169.74.175:34232\}

PROTO (?:%{DATA}-[0-9]\.[0-9])

PROTOCOL \{%{PROTO:proto}(\{%{IPORHOST:remote_host}\})(\{%{NONNEGINT:remote_port:int}\})
TRANSFER_CLASSIC %{BILLING_TIME:billing_time} %
{CELL_AND_TYPE} %{PNFSID_SIZE} %{PATH} %{SUNIT}
{TRANSFER_SIZE} %{TRANSFER_TIME} %{IS_WRITE}
{PROTOCOL} %{DOOR} %{ERROR}
Real Life example

```json
{
    "@version" => "1",
    "@timestamp" => "2015-05-02T06:35:49.000Z",
    "type" => "dcache-billing",
    "host" => "ani",
    "path" => "/var/lib/dcache/billing/2014/05/billing-2004-05-02.log",
    "pool_name" => "dcache-desy23-05",
    "bill_type" => "transfer",
    "pnfsid" => "00009A23BB6D280F46A7A6C12AC67F5EA897",
    "size" => 59419220,
    "file_path" => "/pnfs/desy.de/desy/dcache.org/2.1/dcache-server_2.1.1-1_all.deb",
    "sunit" => "desy:generated@osm",
    "transfer_size" => 90112,
    "transfer_time" => 1195,
    "is_write" => "false",
    "proto" => "Http-1.1",
    "remote_host" => "dcache-infra03.desy.de",
    "remote_port" => 0,
    "payload" => ":WebDAV-dcache-door-desy13:webdav-dcache-door-desy13Domain:",
    "initiator_type" => "door",
    "initiator" => "WebDAV-dcache-door-desy13@webdav-dcache-door-desy13Domain:1399012548236-1399012548243",
    "error_code" => 0
}
```
output {
  elasticsearch {
    host => "elastic-search-master-node"
    index => "logstash-%{+YYYY.MM.dd}"
  }
}

Elasticsearch

- Open-source full-text search engine
- Schema-free JSON documents
- Powerful JSON based REST-API
- Distributed
  - data can be divided into shards
  - each shard can have zero or more replicas
- Node can be Master-node, Data-node or both
- Can be used as a NoSQL database
Document, Index and type

> Document is a basic unit of information
> Documents are expressed in JSON
> Each log entry corresponds to a document
> Index is a collection of documents
> An index is identified by a name
> Name is used to refer to the index when performing actions
> Type is a logical category/partition of an index
> Type is defined for documents that have a set of common fields

(something like DATABASE (index), ROW(document) and TABLE(type) in RDBMS)
Shards and Replicas

> Index can be subdivide into multiple pieces

> Each piece called shard

> Each shard is an independent "index" and can be hosted on any node in the cluster.
  - allows horizontally split/scale data volume
  - allows distribute operations across shards

> You can make one or more copies of index’s shards called replicas
  - provides high availability in case a shard/node fails
  - allows to scale out search volume/throughput since searches can be executed on all replicas in parallel
> REST API

- **POST** – create document, index
- **GET** – search/read document
- **PUT/PATCH** – update document
- **DELETE** – delete document, index
Kibana

> Flexible analysis and visualization platform
> Real-time summary and charting of streaming data
> Intuitive interface for a variety of users
> Instant sharing and embedding of dashboards
Get started

- Dump data into elasticsearch
- Create a simple dashboard (discovery in Kibana4)
- Play with data
Get started

Type to filter...

- @timestamp
- @version
- _id
- _index
- _type
- bill_type
- billing_time
- error_code
- file_path
- host
- initiator
- initiator_type
- is_write
- message
- path
- payload
- pfnsid
- pool_name
- proto
- remote_host
- remote_port
- size
- suinit
- tags
- transfer_size
- transfer_time
- type

_message":"03:07 15:25:31 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}

_message":"03:07 15:25:30 [door:NFS-dcache-dir-cloud01@nts4Domain.request] ["48:48:131.189.71.129] [00006BCB38CF5510DFEBBA226932E6C8F883FA74] [d] cacherereplca@osm 5 0 [0:0:0:0]."],@version":1,"@timestamp":"2015-03-07T14:25:31.052Z","type":"dcache-billing","host":"dcache-se-cloud01.desy.de","..."}
Get started

HISTOGRAM

TABLE

Fields

All (1) / Current (27)

_type
_billing_time
_bill_type
_error_code
_file_path
_host
_initiator
_initiator_type
_is_write
_message
_path

_source (select columns from the list to the left)

0 to 100 of 500 available for paging

Tigran Mkrtchyan | Visualization of dCache accounting information | Date | Page 28
And you start to explore

EVENTS OVER TIME

- GridFTP (18036)
- Dcap (38167)
- NFSv4.1 (10458)
- Http (4834)
- Xrootd (6428)

count per 10m | (77723 hits)
And you start to explore

![Graph showing transferred bytes over time]

- **View**: Multiple options for viewing the data.
- **Zoom Out**: Possibly to view the data at a higher level.
- **Writes (5566)**: 5566 write operations.
- **Reads (72157)**: 72157 read operations.
- **transfer_size total per 10m**: Total transfer size per 10 minutes.
- **77723 hits**: Number of hits on the data.
And you start to explore
And you start to explore
And you start to explore
And you start to explore
Transfers at glance

- Transfers at glance
- Transferred Bytes
- Status
- Read vs. Write
- Top 10 MUnts
- Top 10 Pools
- Bytes Total

- Views:
  - [View 1]
  - [View 2]

- Date: [Page 35]
Too good to be the truth

> Elasticsearch is very giddy component

> Number of active indexes is limited by file descriptors
  - #indexes * #shards * #replicas * #segments
  - In production we can have max ~180 days

> Can't be used to analyze historic data

> But good enough for live monitors
  - not a reporting tool
Looking back (how to organize index)

> index => "logstash-%{+YYYY.MM.dd}"
  
  ▪ typical search will use 1-7 indexes
  ▪ typical search data overhead one day
  ▪ limited number of indexes
  ▪ discard granularity one day

> index => "logstash-%{+YYYY.MM}"
  
  ▪ typical search will use 1 index
  ▪ typical search data overhead one month
  ▪ discard granularity one month

> your 'live view' defines which type of index you need
Who-is-Who

> No Authentication/Authorization
> All data available as soon as you can get access to ES REST-API
> Different projects to solve this issue
> We are looking at Kibana_Auth
  - a proxy between and user and Kibana
  - modifies user requests
  - integrated with ldap
Kibana4

> New concepts
> Allows to have saved searches, visualization
> Allows to crate Dashboards with data from different indexes
> We still collect experience
Questions?