Multi DC Zabbix Environment

Rights and Wrongs while building and operating a large, multi datacenter, Zabbix environment focusing on open-source software.

SENTIA | LEAD THE WAY

Julio C Hegedus

“Infra Continuity Engineer”

+30 years experience in general IT
+20 years with Linux and mixed environments
+18 years with service providers

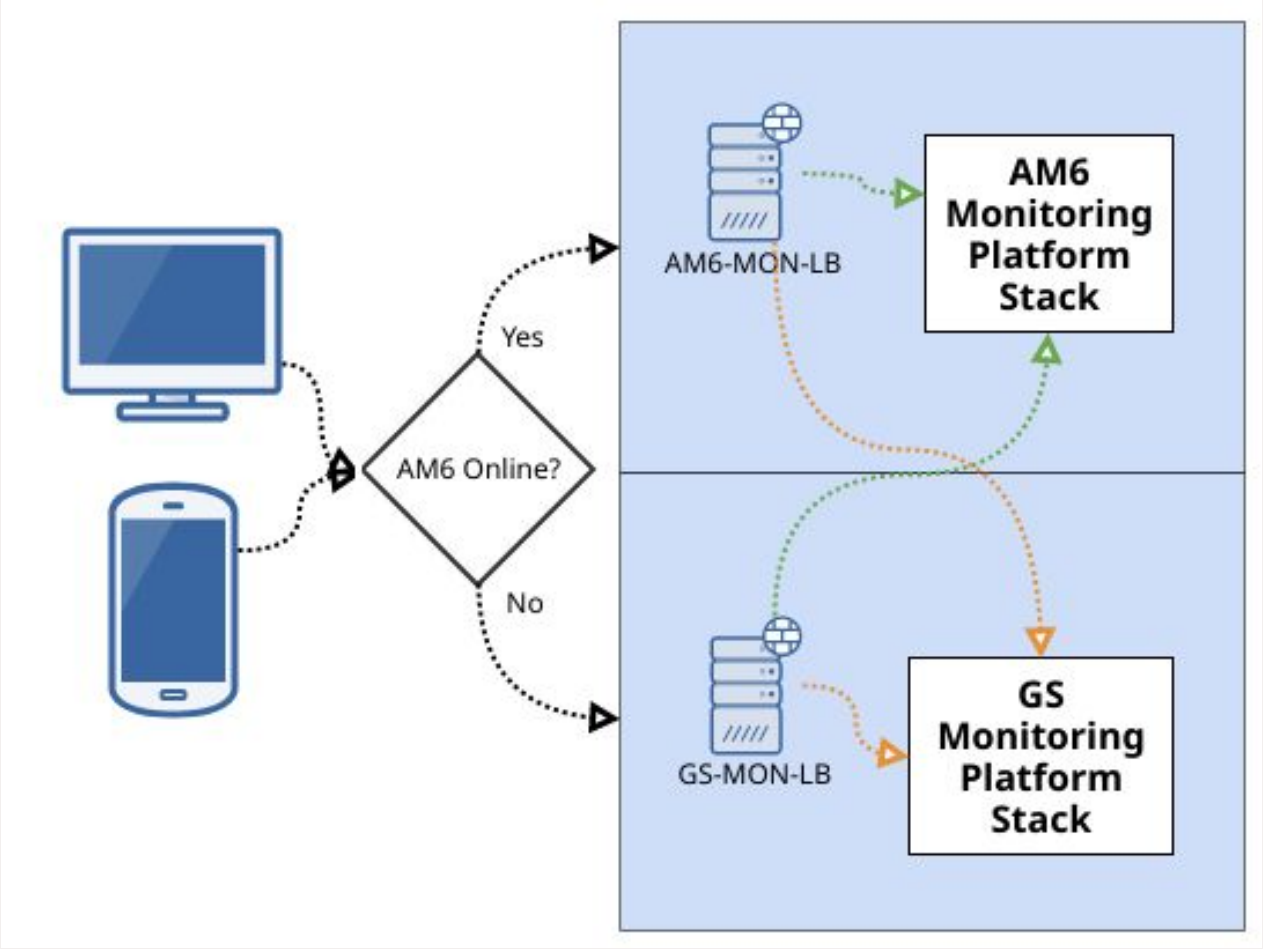
SENTIA | LEAD THE WAY

Project Goal

Maximum amount of availability depends on:

- **Uncompromised access**
- **Data duplication**
- **Multi component ability**
- **Server duplication**
- **Methodology of Failover and Failback**
- **Speed in communication**

99.99% Uptime seems attainable!



SENTIA | LEAD THE WAY

What Should (not) Happen?

Basic Discovery Phase:

- No Standards
- Zabbix, Nagios, PRTG, AWS, Azure, Custom, Others
- No Framework (Ticketing, Notification)
- Continuity
- Centralization

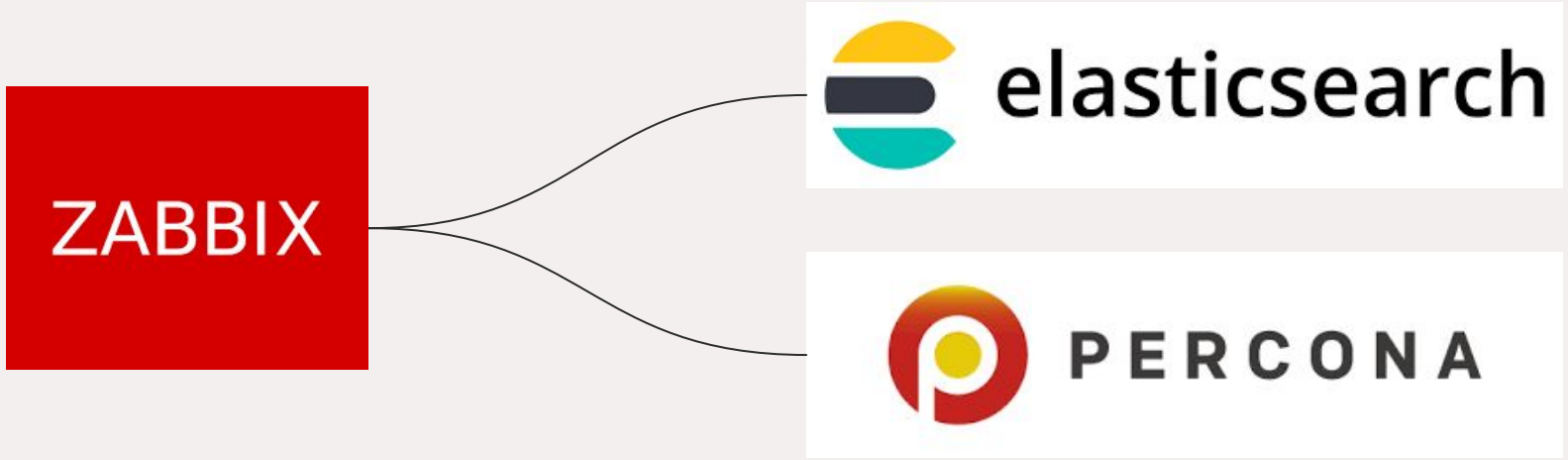
- > 4.500 Hosts
- > 650.000 Items
- > 3TB Year
- > 250 Users

SENTIA | LEAD THE WAY

Wishing / Targeting Phase

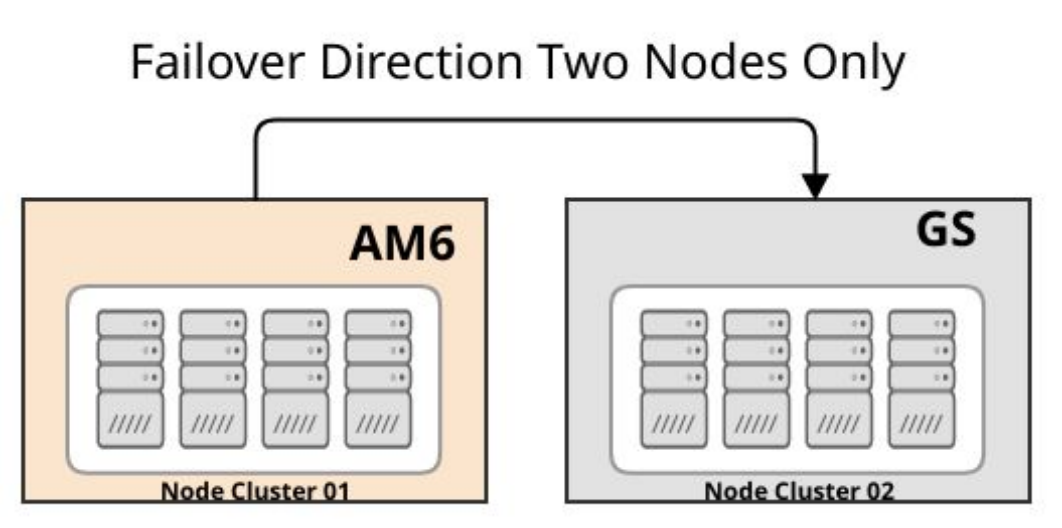
- Let’s do something fresh!
- Contains all the features everybody uses!
- Most Open-source as possible!
- Multi Datacenter is a MUST
- Multi Active Databases for failover
- How to keep 10 years of data?
- Notifications to multiple systems?
- Metrics from everything off the operations
- Dashboarding is never too much!
- Elasticsearch and Grafana on the mix
- In-House integration tooling
- Automation

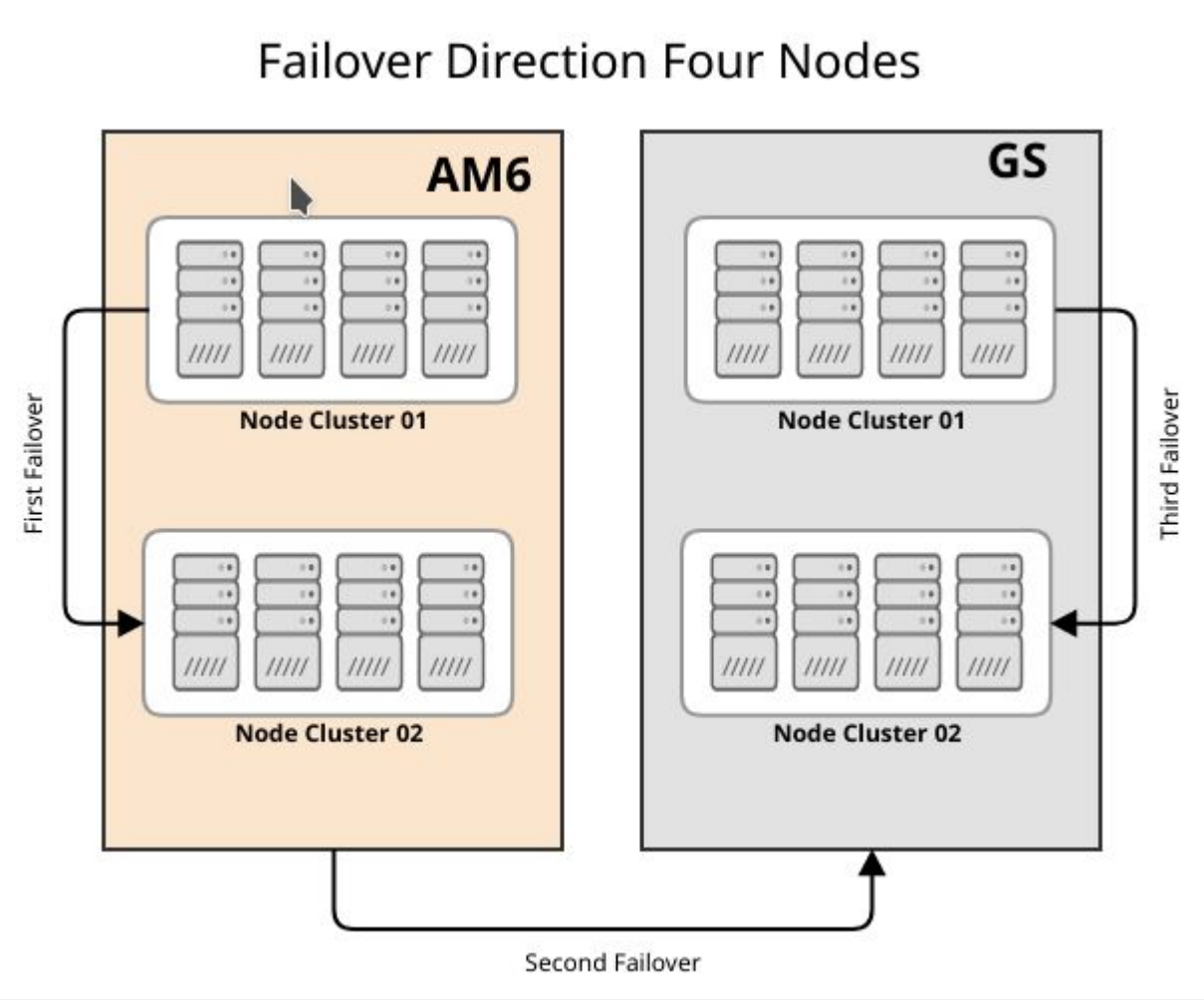
Elasticsearch as “Storage”



Testing Zabbix against Elasticsearch and Mysql is a breeze -- but there are limitations and broken promises

The plan on Storage and Failover

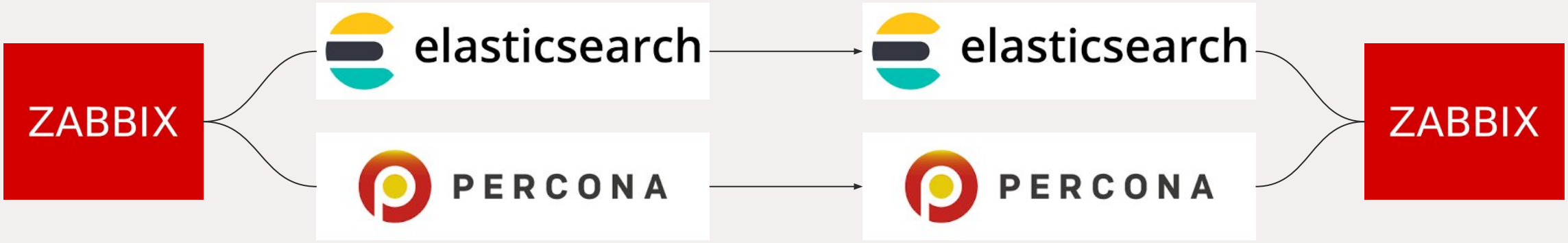




MySQL Percona Active/Active and Elasticsearch in a Cross Datacenter Setup is feasible

SENTIA | LEAD THE WAY

Data “Independency”



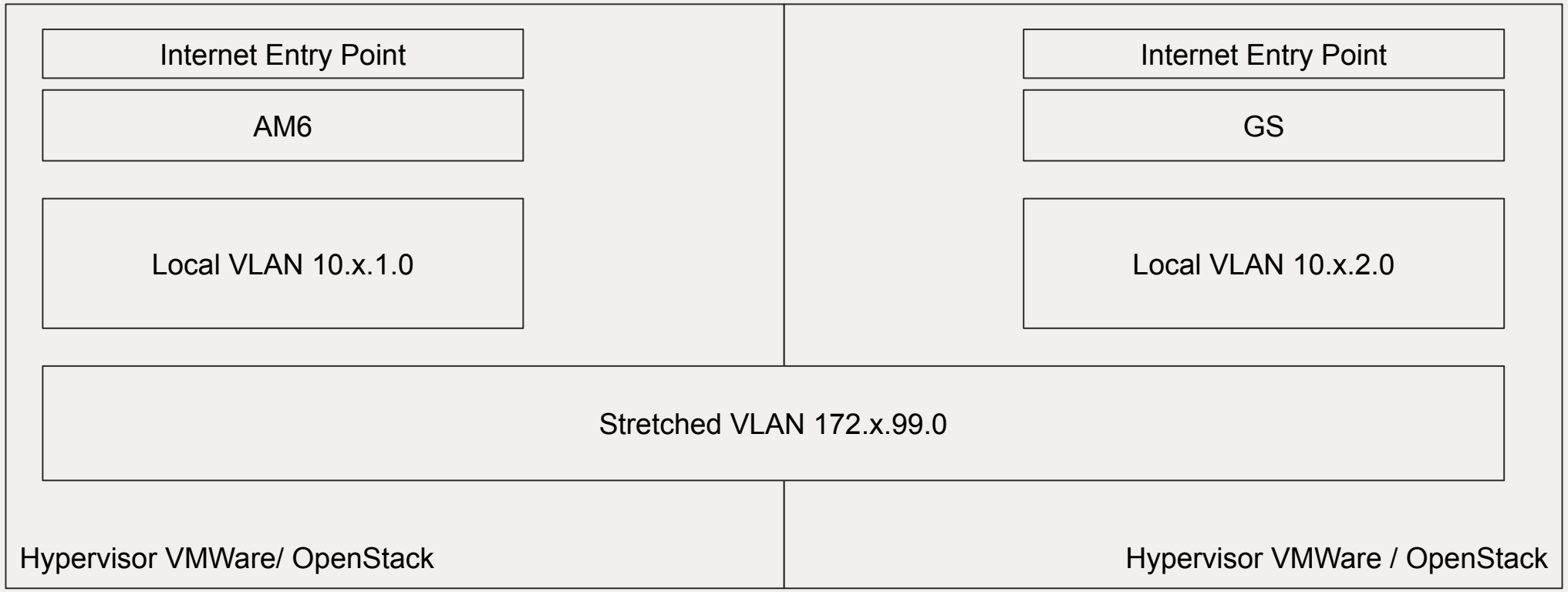
Datacenter AM6

Datacenter GS

When catastrophe strikes one should be able to continue without the other and re-synchronization should be easy

SENTIA | LEAD THE WAY

Final Setup per DC - Networking



SENTIA | LEAD THE WAY

Final Setup per DC - Zabbix

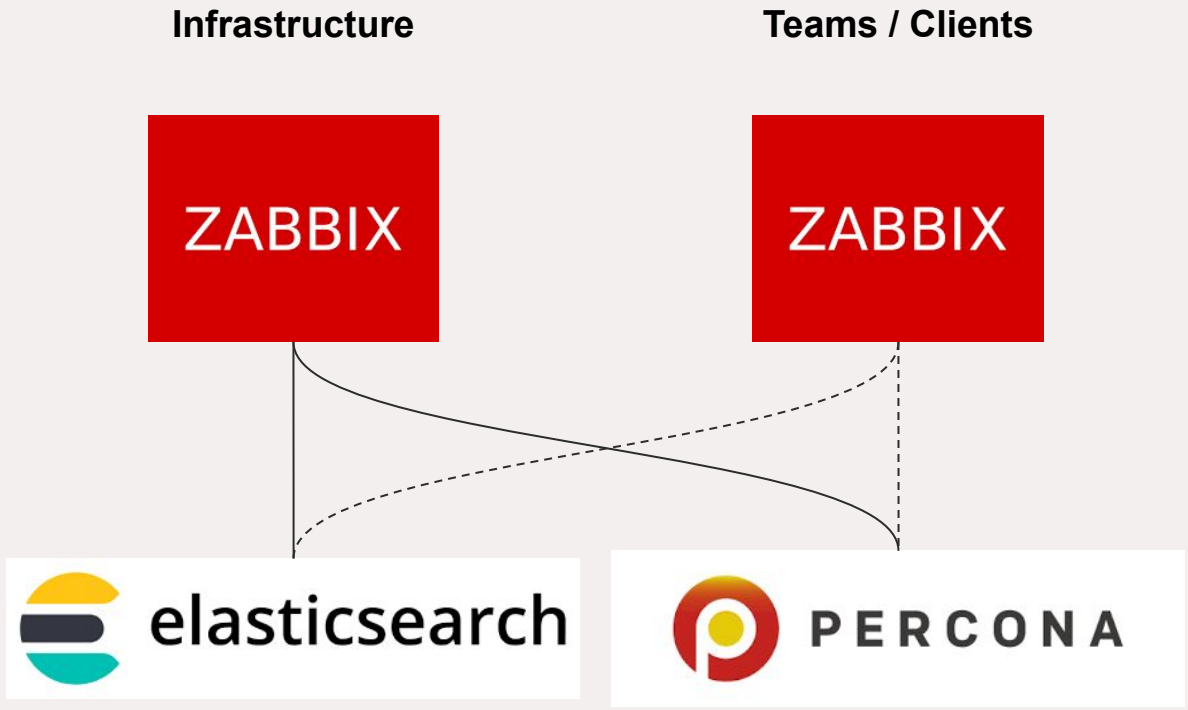
Zabbix Mysql CFG accepts:

- Custom MySQL DB names
- Custom Ports
- Custom Hosts

Zabbix Elasticsearch accepts:

- Only one server, one index

Two Zabbix Servers should be running per Datacenter -- One for Infra and one for Clients / Teams of operations



SENTIA | LEAD THE WAY

Elasticsearch Framework Fix

PATCH: https://support.zabbix.com/browse/ZBXNEXT-4968

Elasticsearch Problems

```
14661:20190308:092001.514 cannot get values from elasticsearch, HTTP status code:
503
14679:20190308:092032.654 cannot get values from elasticsearch, HTTP status code:
429
14709:20190308:          cannot get values from elasticsearch, HTTP status code:
404
```

ZABBIX

503 = Service Unavailable
429 = Too Many Requests
404 = Not Found
400 = Bad Request


```
failed to execute bulk item (index) index
source[{"itemid":27534,"ns":397382014,"clock"
:1554673134,"value":"18446744073709551615","t
tl":604800}]}
```

Elasticsearch CFG Update

Error 400 Fix

http.max_initial_line_length: 16kb

Defaults to 4kb

Header tweaks also might help

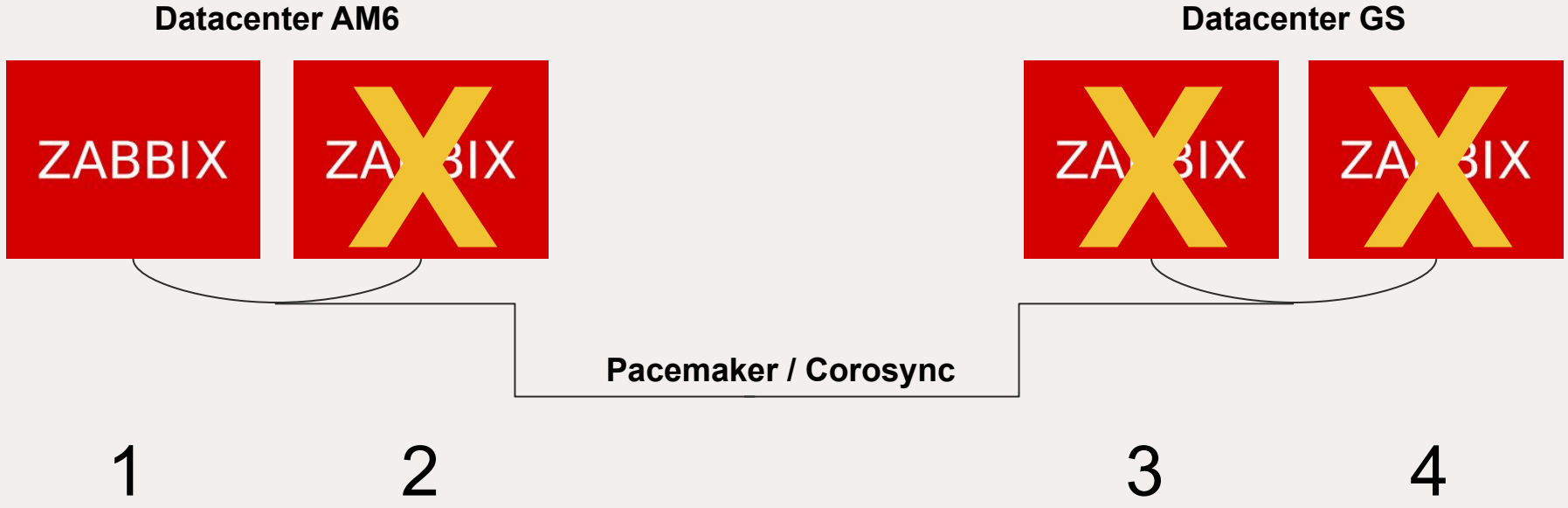
```
DELETE
/_search/scroll/DnF1ZXJ5VGhlbkZldGNoBQAAAAABhkGqFmJ5Vm1xLXpO
UWYteEJzeE1iWjk2MUEAAAAAAjqF4hZCWkhadGlZcVRwNmFreFluR1VfRz
Z3AAAAAAGGQasWYnlWbXEtek5RZi14QnN4TWJaOTYxQQAAAAAC_fPCF
mRlamRiU083U0N1UFp1M0tlSUpvNEEAAAAAAgNDERZXdjYzaVFlZFRfTzF
OelRyeWdFY2dB HTTP/1.1
Host: 172.27.84.10:9200
Accept: */*
Content-Type: application/json
Content-Length: -1
Expect: 100-continue

HTTP/1.0 400 Bad request
Cache-Control: no-cache
Connection: close
Content-Type: text/html

<html><body><h1>400 Bad request</h1>
Your browser sent an invalid request.
</body></html>
```

SENTIA | LEAD THE WAY

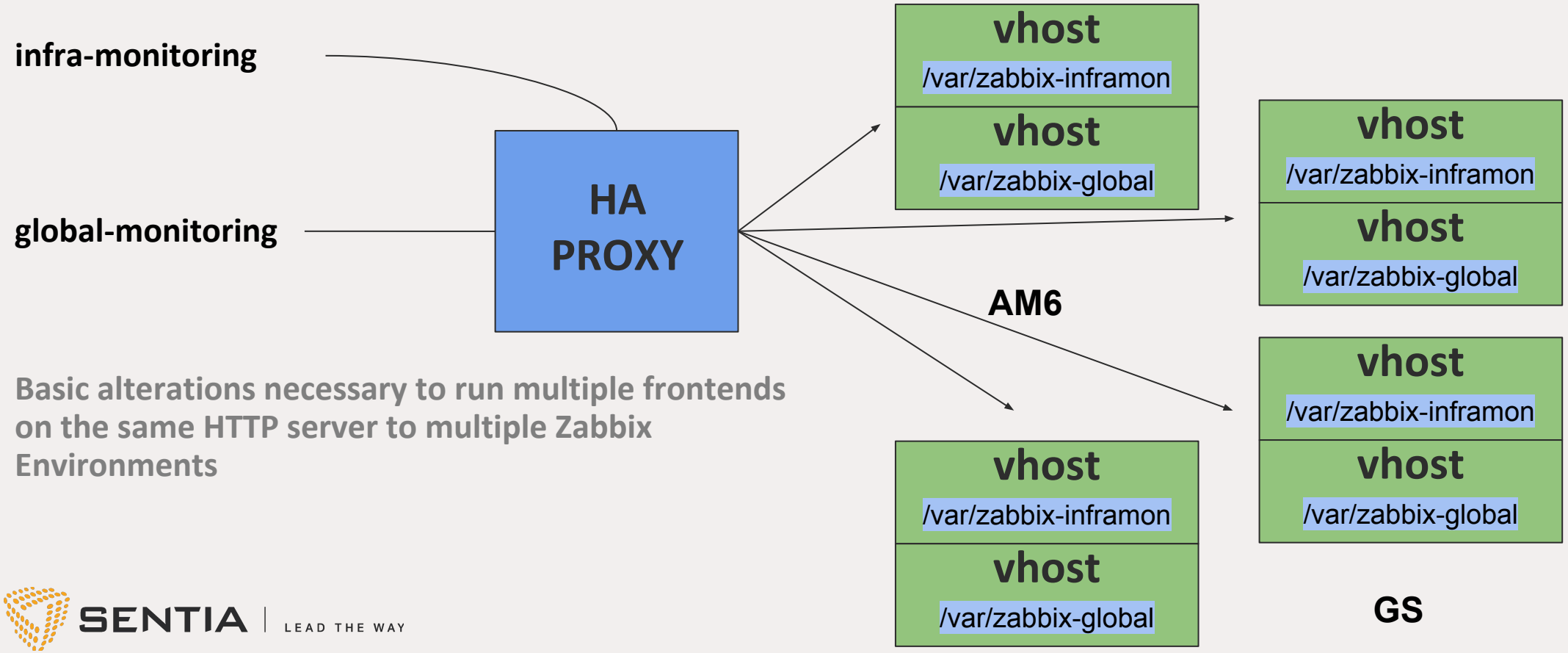
PaceMaker, Only One Server Please



Zabbix Servers in standby cannot run simultaneously, or they will “corrupt” data and create confusion

SENTIA | LEAD THE WAY

Multiple Front End / Same Server



Basic alterations necessary to run multiple frontends on the same HTTP server to multiple Zabbix Environments

SENTIA | LEAD THE WAY

Zabbix Front End Changes v:4.0.3

PHP FE: cp /usr/share/zabbix/ /usr/share/zabbix-infra/

CFG: cp /etc/zabbix /etc/zabbix-infra

edit **include/classes/core/ZBase.php:276**

path to maintenance.php -- /etc/zabbix-infra/

edit **include/classes/core/CConfigFile.php:27**

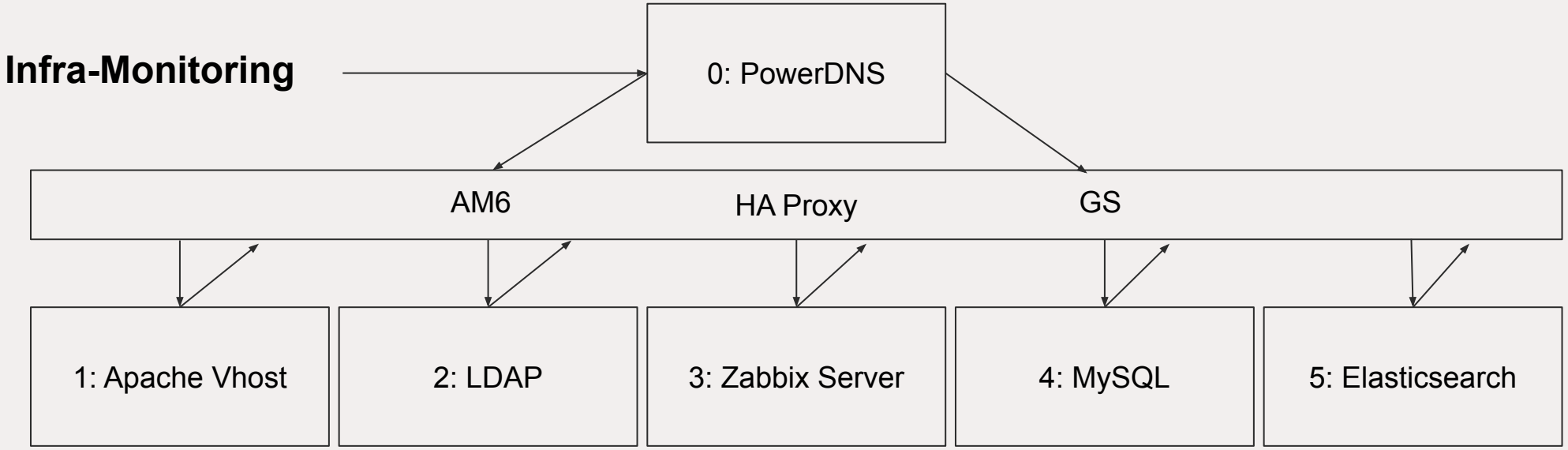
path to zabbix.conf.php -- /etc/zabbix-infra/

SENTIA | LEAD THE WAY

CFG File excerpts:

```
global $DB, $HISTORY, $HISTORY_PREFIX;
$DB['DATABASE']          = 'zabbix_infra';
$ZBX_SERVER              = '172.27.84.10'; // haproxy
$HISTORY['url']   = [
        'uint' => 'http://172.27.84.10:9200',
        'dbl' => 'http://172.27.84.10:9200',
        'str' => 'http://172.27.84.10:9200',
        'log' => 'http://172.27.84.10:9200',
        'text' => 'http://172.27.84.10:9200'
];
// Value types stored in Elasticsearch.
$HISTORY['types'] = ['uint', 'dbl', 'str', 'log', 'text'];
$HISTORY_PREFIX='infra'; //  **## SENTIA PATCH ONLY ##**
```

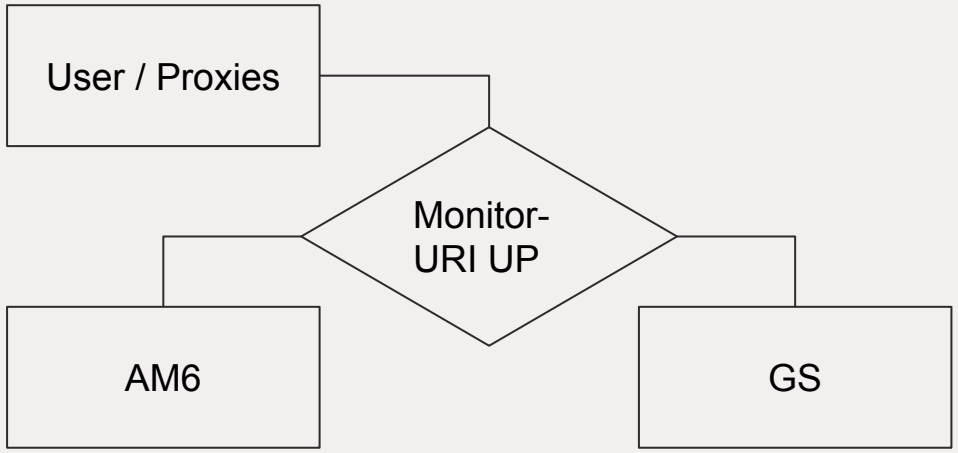
Front End Access Layer



Apache VHost configuration has no secrets, but HA Proxy becomes the basic layer in between everything

SENTIA | LEAD THE WAY

PowerDNS with LUA




```
inframon-sentia.net      1     IN    LUA  A
 "ifurlup('https://infra-monitoring/site-alive', {
   {'185.133.296.111'}, {'213.264.142.222'}
 })"
```

Notes on External DNS failover:

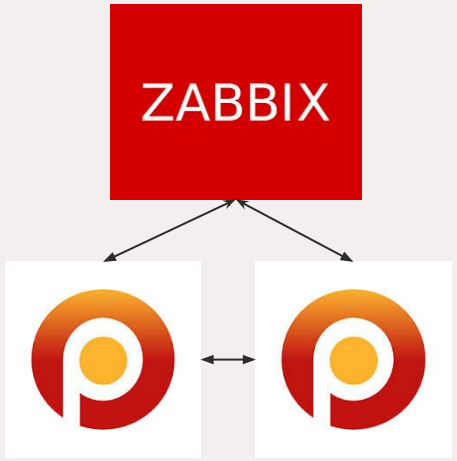
- DNS should be **outside** of both DCs
- Test **monitor-uri** technique from haproxy
- **ifurlup** have orderable targets

Notes on Internal DNS failover:

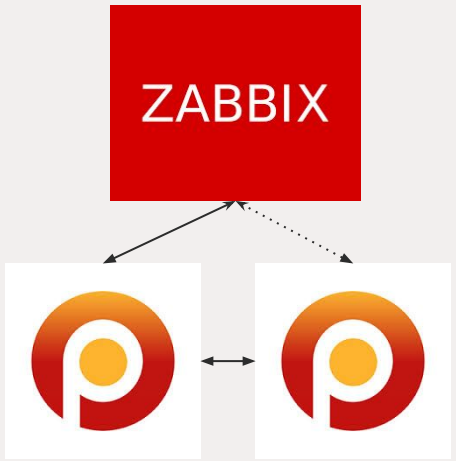
- DNS Should be **Internal** on both DCs
- **Same monitor-uri** component as test source
- Pointing to **internal addresses** for sync

SENTIA | LEAD THE WAY

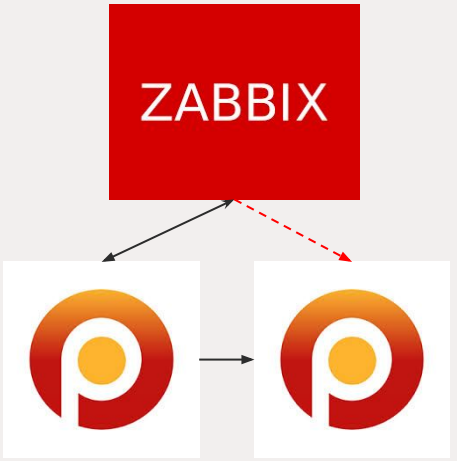
MySQL Percona Active/Passive/Slave



Active/Active



Active/Passive



Active/Slave

The more automatic the better, if no manual action required is perfect.

SENTIA | LEAD THE WAY

MySQL Percona XTraDB Problems

General Notes:

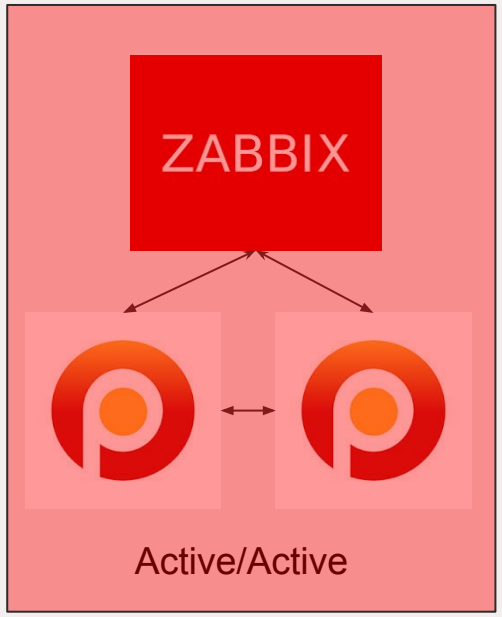
- **Multi Database write is not an advantage**
- **Async failures stop the whole cluster**

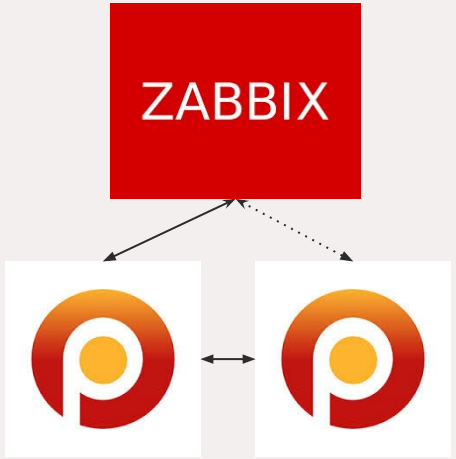
What can be a problem?

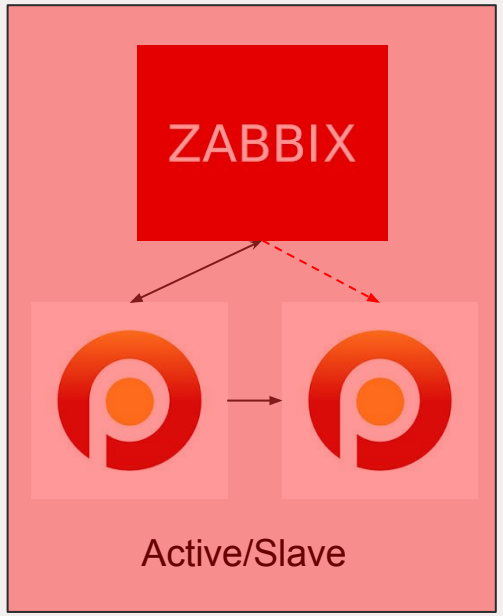
- **Disk Speed**
- **Data Volume**
- **Latency**

2019-03-21T01:01:38.719314Z 8 [Note] InnoDB: **BF-BF X lock conflict**,mode: 1027 supremum: 0
2019-03-21T01:01:38.719325Z 8 [Note] InnoDB: conflicts states: my: 0 locked: 0
RECORD LOCKS space id 951 page no 459 n bits 152 index PRIMARY of table **`zabbix_infra`.`problem`** trx id 9444233 lock_mode X locks rec but not gap
2019-03-21T01:01:38.719356Z 8 [ERROR] Slave SQL: Could not execute **Delete_rows event on table zabbix_infra.problem; Can't find record in 'problem', Error_code: 1032; handler error HA_ERR_KEY_NOT_FOUND;** the event's master log FIRST, end_log_pos 8314, Error_code: 1032

MySQL Percona Active/Passive/Slave







The more automatic the better, if no manual action required is perfect.

SENTIA | LEAD THE WAY

Other Observations

- **Percona Requires an Arbiter, one per datacenter, in a second Pacemaker cluster**
- **Pacemaker can run stretching both Datacenters or two separated Pacemaker clusters controlled by a Booth Cluster Ticket Manager**
- **Elasticsearch can run in CCS (Cross Cluster) or if the latency is very very low, one cluster can be stretched**
- **Kibana, Grafana, Zabbix FE, other web services can run all from the same servers as multiple instances**
- **All the communication for the internal components are only done via the stretched VLAN**
- **Servers and Configurations can be all deployed using Ansible**
- **MySQL load balancing via haproxy requires advanced techniques for node availability checks**
- **Ingestion pipelines can help on new methods of indices rotation**
- **Grafana has compatibility problems with Elasticsearch**

Migration steps for our case

Export configurations from client / group

First Steps

Preparation

Import / Test all the configuration on the new server

Connectivity

Firewalls

Proxy configuration to new destination

No historical data

Manual work

Notification tests
ACL Cascading
Harmonization

Observation

Finalization

Conclusions

- **Multi Datacenter is feasible and works!**
- **Elasticsearch may be better to environments with less volume of data (up to 4.0.x)**
- **Active / Active with Percona is promising but not perfect, difficult to say where the problem really comes from**
- **With the multiple different forms of implicitly changing the configuration, this setup can have multiple forms of failback… If one scenario doesn't work for you, small changes can be done to recover from architecture failures**
- **Increasing all proxy buffers will guarantee to avoid data loss with changes on the server, failing over or doing changes for updates and unforeseen load changes**
- **Keeping each DC as much independent as possible the better**

THANK YOU