Detection of erratic behavior in load balanced clusters

Martin Adam
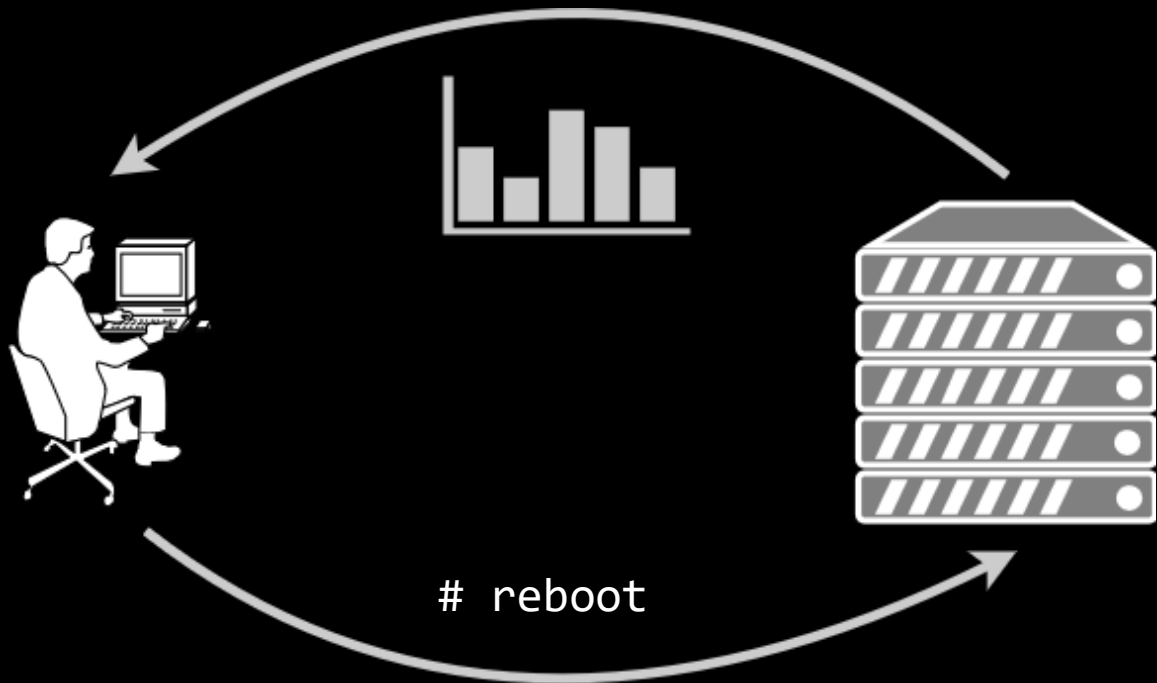
L.Magnoni, D.Adamová, M.Pilát
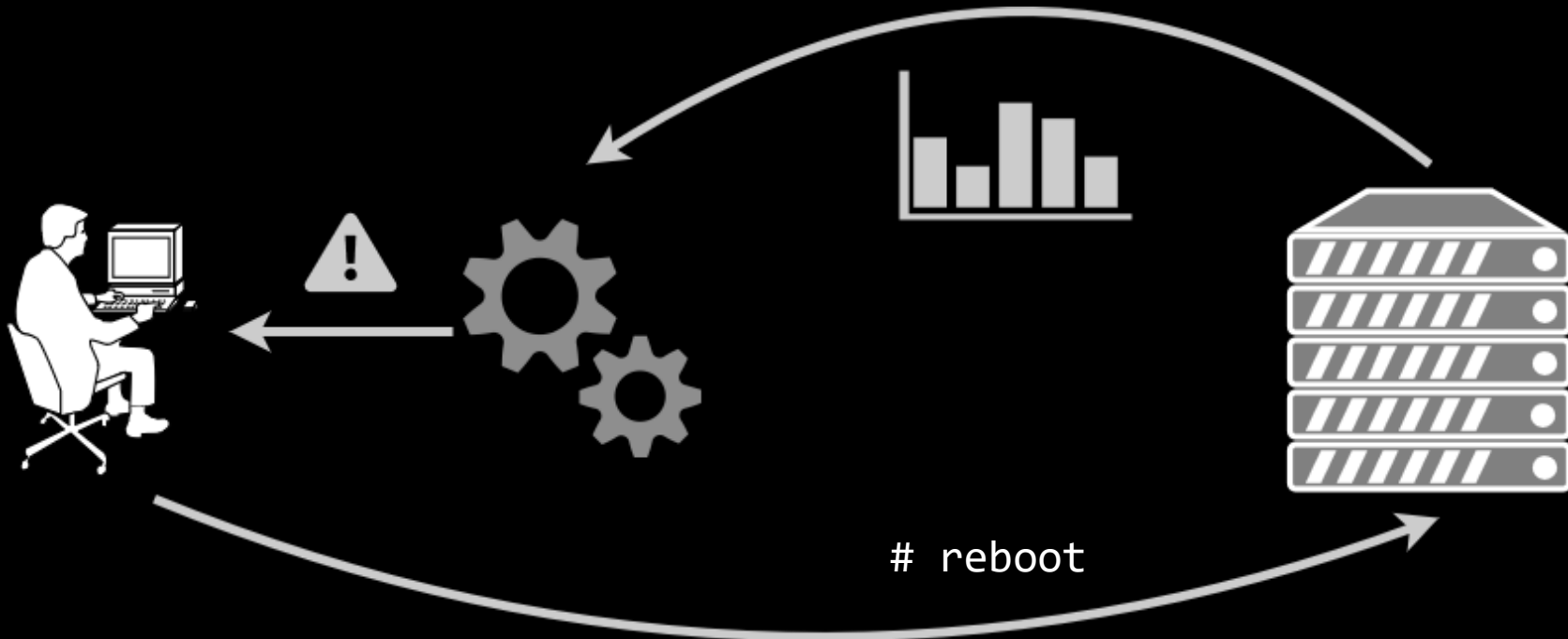
Monitoring Data Everywhere!

reboot

reboot

CRIT	aplex20.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 07:04:06	36 min
CRIT	aplex21.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 05:08:51	31 min
CRIT	aplex22.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 05:10:58	29 min
CRIT	aplex23.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:48:19	51 min
CRIT	aplex24.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:53:04	47 min
WARN	aplex25.farm.particle.cz	load			WARNING - load average: 33.72, 33.63, 33.49		105 min	66 sec
CRIT	aplex25.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:57:49	42 min
CRIT	aplex26.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 05:02:34	37 min
CRIT	aplex27.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 05:07:18	32 min
CRIT	aplex28.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:42:01	58 min
CRIT	aplex29.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 07:11:17	28 min
CRIT	aplex30.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:51:31	48 min
WARN	aplex31.farm.particle.cz	EDAC			WARNING - mc1: csrow0: CPU_SrcID#0_Ha#0_Channel#0_DIM: 1436 Corrected Errors		2018-06-30 14:15:50	23 hrs
CRIT	aplex31.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 06:56:16	43 min
WARN	aplex32.farm.particle.cz	load			WARNING - load average: 59.13, 58.67, 58.73		2 hrs	8 min
CRIT	aplex32.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 05:01:01	39 min
CRIT	aquaklima.monitor	AQUAKLIMA			ERROR: 1 - Common alarm		8 hrs	8 min
CRIT	arc1.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (3.10.0-862.3.2.el7) is older than installed (3.10.0-862.3.3.el7). Reboot needed.		2018-06-19 14:06:03	12 min
CRIT	arc2.farm.particle.cz	disk			DISK CRITICAL - free space:		2018-07-03 20:03:13	7 min
CRIT	arc2.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (3.10.0-862.3.2.el7) is older than installed (3.10.0-862.3.3.el7). Reboot needed.		2018-06-19 13:54:31	5 min
WARN	arc2.farm.particle.cz	load			WARNING - load average: 0.32, 2.42, 3.01		17 min	3 min
CRIT	arc-it4i.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (2.6.32-696.28.1.el6) is older than installed (2.6.32-696.30.1.el6). Reboot needed.		2018-06-29 23:36:41	33 sec
CRIT	arc-it4i.farm.particle.cz	YUM			13 Security Updates Available. 21 Non-Security Updates Available		2018-07-03 05:07:20	32 min
CRIT	argus.grid.cesnet.cz	YUM			27 Security Updates Available. 2 Non-Security Updates Available		2018-07-03 10:43:38	56 min
CRIT	auger-sdeu.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (2.6.32-696.28.1.el6) is older than installed (2.6.32-696.30.1.el6). Reboot needed.		2018-06-28 18:35:43	8 min
CRIT	auger-sdeu.farm.particle.cz	YUM			4 Security Updates Available		2018-07-03 04:46:49	53 min
CRIT	backup.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (2.6.32-696.30.1.el6) is older than installed (2.6.32-754.el6). Reboot needed.		2 hrs	16 min
CRIT	bdii1.egee.cesnet.cz	YUM			22 Security Updates Available. 1 Non-Security Update Available		2018-07-03 06:37:45	62 min
CRIT	bdii1.farm.particle.cz	CONFIGURATOR			cached catalog used since: 07.06.2018 10:33:34		2018-06-10 00:12:53	12 min
CRIT	condor.farm.particle.cz	KERNEL_VERSION			ERROR: Running kernel version (3.10.0-862.3.2.el7) is older than installed (3.10.0-862.6.3.el7). Reboot needed.		2018-06-26 10:49:47	5 min
CRIT	cream1.grid.cesnet.cz	KERNEL_VERSION			ERROR: Running kernel version (2.6.32-696.28.1.el6) is older than installed (2.6.32-696.30.1.el6). Reboot needed.		2018-05-24 12:59:19	9 min
CRIT	cream1.grid.cesnet.cz	YUM			31 Security Updates Available. 7 Non-Security Updates Available		2018-07-03 05:08:57	31 min
CRIT	cream2.grid.cesnet.cz	YUM			31 Security Updates Available. 7 Non-Security Updates Available		2018-07-03 04:42:06	58 min
CRIT	db2.farm.particle.cz	YUM			17 Security Updates Available. 29 Non-Security Updates Available		2018-07-03 04:43:41	56 min
CRIT	dfc-auger.grid.cesnet.cz	KERNEL_VERSION			ERROR: Running kernel version (2.6.32-696.28.1.el6) is older than installed (2.6.32-696.30.1.el6). Reboot needed.		2018-06-29 23:04:36	15 min
CRIT	dfc-auger.grid.cesnet.cz	YUM			9 Security Updates Available		2018-07-03 04:43:41	56 min
CRIT	dpm1.egee.cesnet.cz	YUM			23 Security Updates Available. 1 Non-Security Update Available		2018-07-03 04:50:01	50 min
CRIT	dpmpool3.grid.cesnet.cz	YUM			21 Security Updates Available. 1 Non-Security Update Available		2018-07-03 04:59:35	40 min
CRIT	dpmpool4.grid.cesnet.cz	YUM			21 Security Updates Available. 1 Non-Security Update Available		2018-07-03 05:07:29	32 min

Tactical Overview

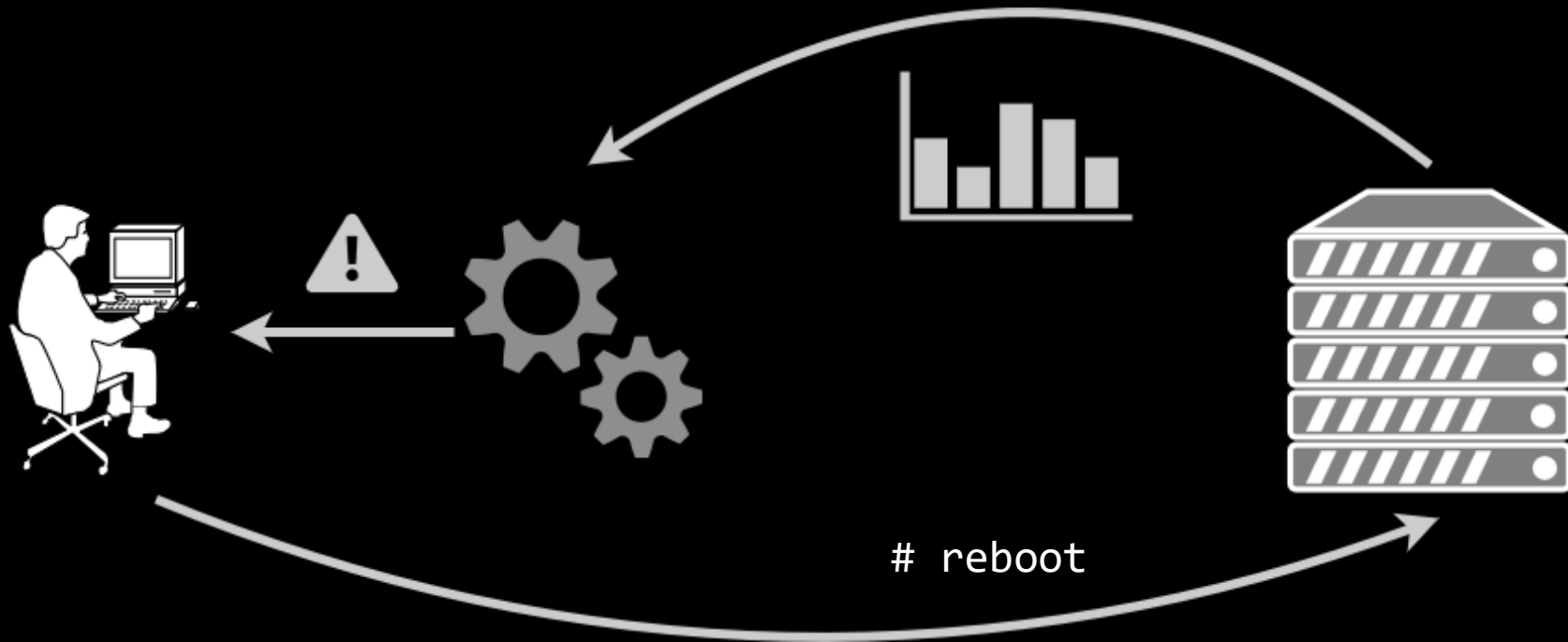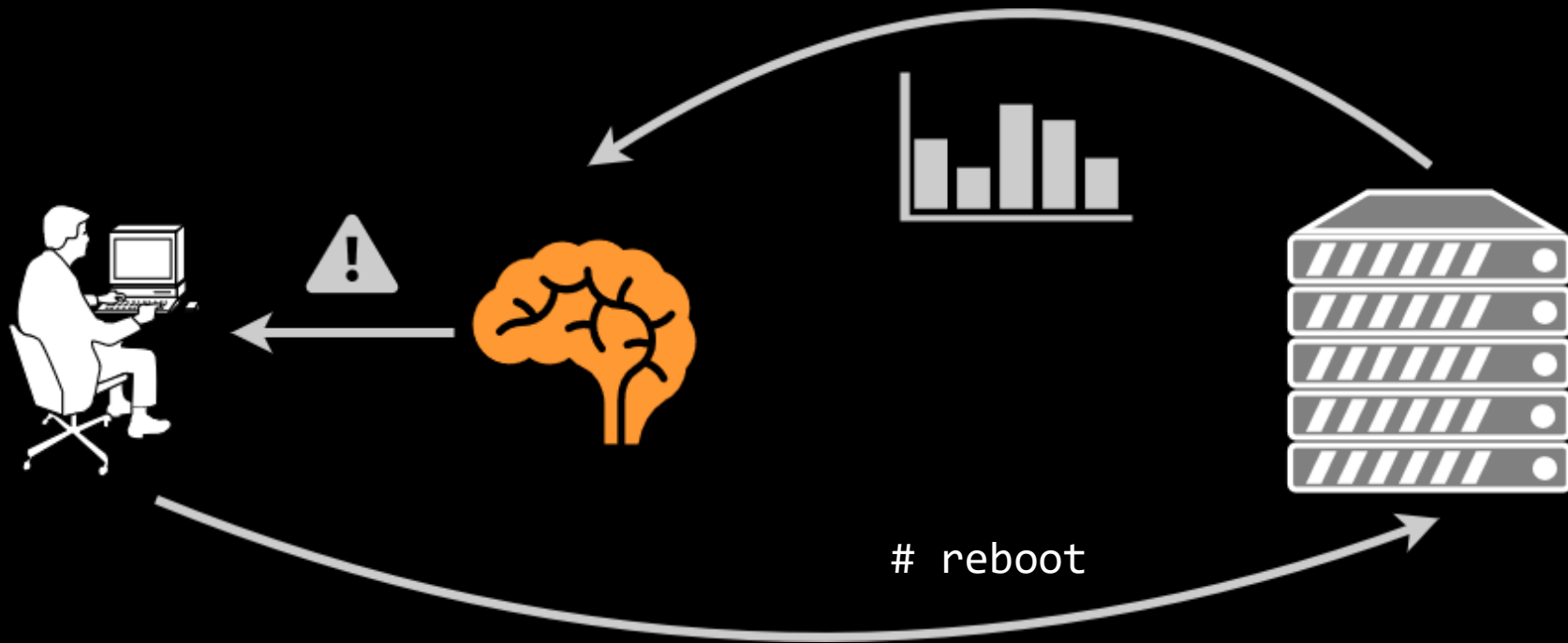| Hosts	Problems	Unhandled
416	2	2
Services	Problems	Unhandled
7550	409	409

Quicksearch

Views

Overview
- Host & Services Problems
- Main Overview

Hosts
- All hosts
- All hosts (Mini)
- All hosts (tiled)
- All hosts by status
- Favorite hosts
- Host search

Host Groups
- Host Groups
- Host Groups (Grid)
- Host Groups (Summary)

Services
- All services
- CONFIG
- Favorite services
- KERNEL-farm.particle.cz
- Recently changed services
- Serv. by host groups
- Service search

Service Groups
- Service Groups
- Service Groups (Grid)
- Service Groups (Summary)
- Services by group

Business Intelligence
- All Aggregations
- Hostname Aggregations
- Problem Aggregations
- Single-Host Aggregations
- Single-Host Problems

Problems
- Alert Statistics
- Host problems
- Pending Services
- Service problems
- Stale services

Addons
- NON OK - Search Graphs
- Search Graphs

Inventory
- CPU Related Inventory of all Hosts
- Software Package Search

Other
- Comments
- Downtimes
- History of scheduled downtimes
- Host- and Service events

© Mathias Kettner

Can we do it better?

reboot

Can we do it better?

reboot

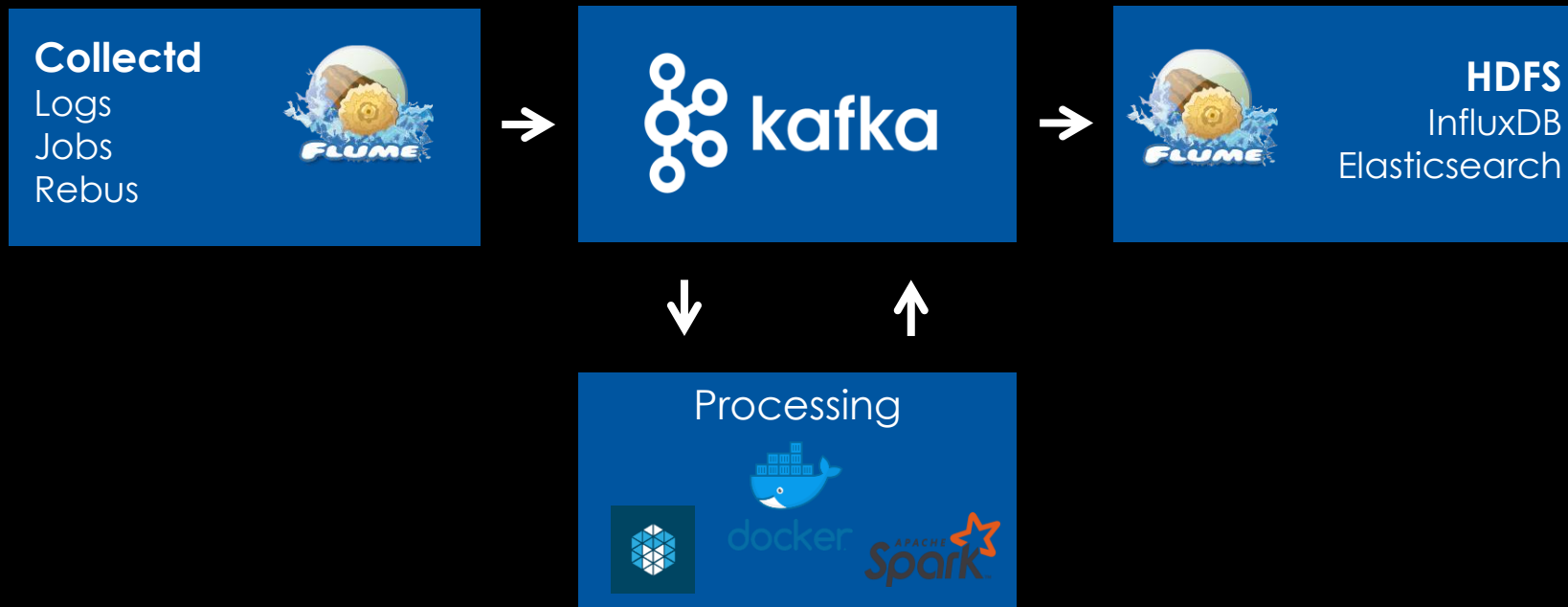Fixed threshold

Uncorrelated data
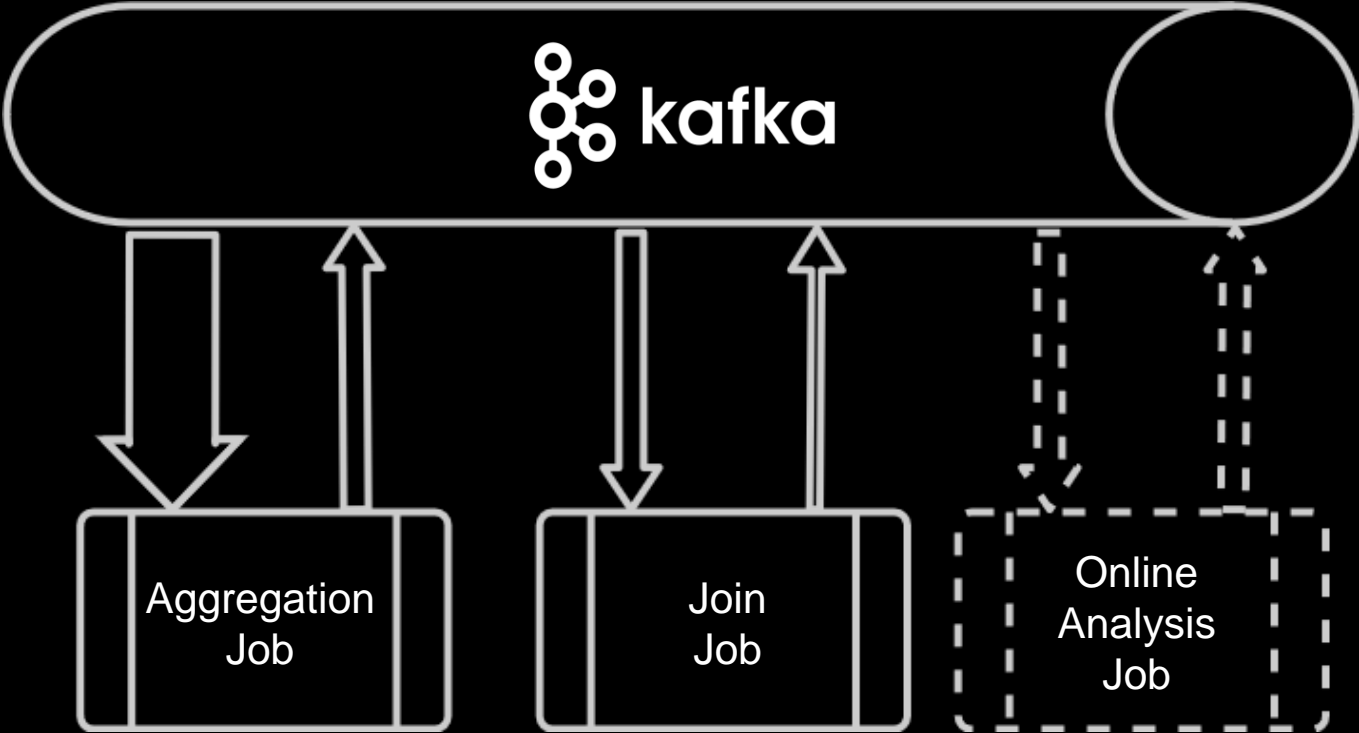
~~Fixed threshold~~

~~Uncorrelated data~~

Workflow and workload aware analysis

The CERN MONIT Infrastructure

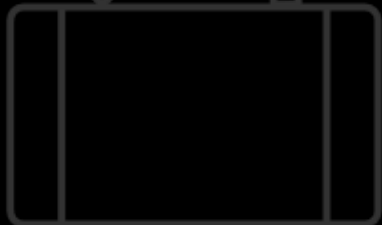# Implementation

The Data – Raw Collectd Event

```
{
  "metadata": {...},
  "data": {
    "host": "host.cern.ch",
    "value": 5.570256057198357,
    "plugin": "cpu",
    "type_instance": "idle",
    ...
  }
}
```

kafka

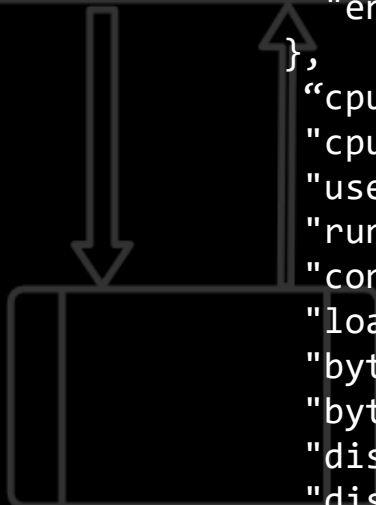# The Data – Aggregated Collectd Event

```
{
  "metadata": {...},
  "data": {
    "host": "host.cern.ch",
    "window": {
      "start": "2018-05-21T16:40:00",
      "end": "2018-05-21T17:00:00"
    },
    "value": 5.570256057198357,
    "plugin": "cpu",
    "type_instance": "idle",
    ...
  }
}
```

The Data – DAM Event

```
{
  "metadata": {...},
  "data": {
    "host": "host.cern.ch",
    "window": {
      "start": "2018-05-21T16:40:00",
      "end": "2018-05-21T17:00:00"
    },
    "value": 5.570256057198357,
    "plugin": "cpu",
    "type_instance": "idle",
    ...
  }
}
```

```
{
  "metadata": {...},
  "data": {
    "host": "host.cern.ch",
    "window": {
      "start": "2018-05-21T16:40:00",
      "end": "2018-05-21T17:00:00"
    },
    "cpu_idle": 5.570256057198357,
    "cpu_iowait": 0.005315354980072367,
    "used_memory": 2673355161.6,
    "running_processes": 0.25,
    "connections_established": 2716.25,
    "load": 0.9026250000000001,
    "bytes_sent": 259170.90712136906,
    "bytes_received": 978519.553353911,
    "disk_write": 16923.2423603067,
    "disk_read": 0   }
  }
}
```
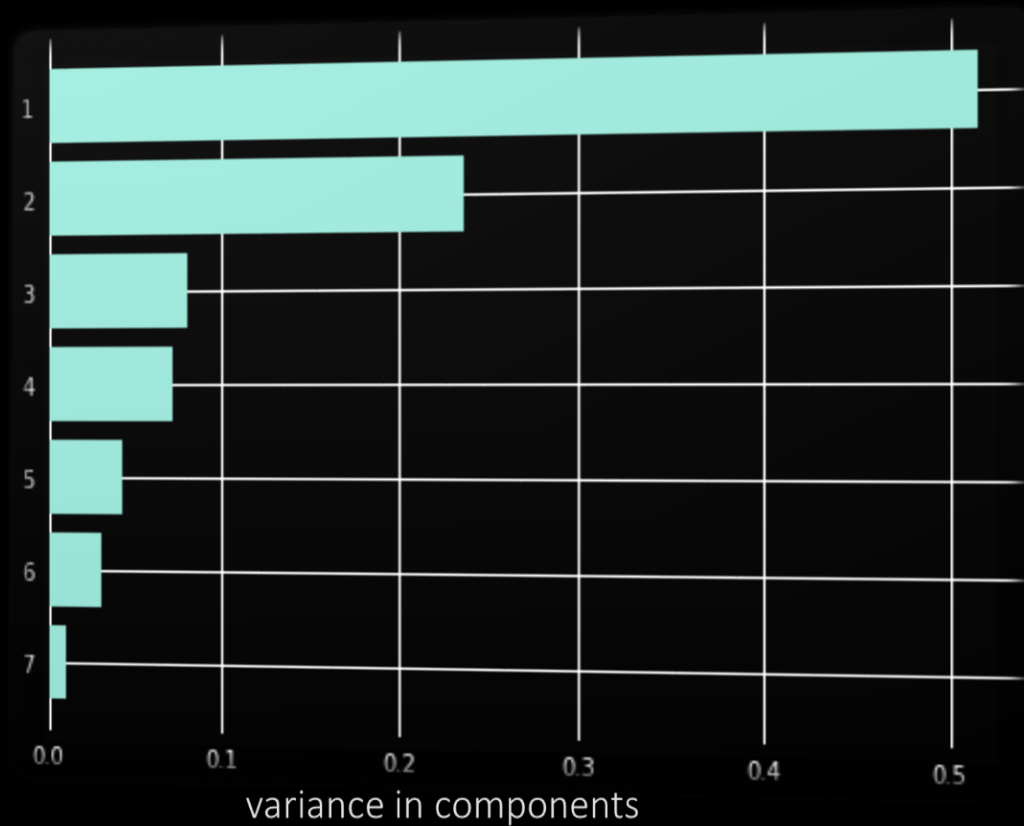
The Metrics

10 <u>Hardware</u> Metrics

```
{
  "metadata": {...},
  "data": {
    "host": "host.cern.ch",
    "window": {
      "start": "2018-05-21T16:40:00",
      "end": "2018-05-21T17:00:00"
    },
    "cpu_idle": 5.570256057198357,
    "cpu_iowait": 0.005315354980072367,
    "used_memory": 2673355161.6,
    "running_processes": 0.25,
    "connections_established": 2716.25,
    "load": 0.9026250000000001,
    "bytes_sent": 259170.90712136906,
    "bytes_received": 978519.553353911,
    "disk_write": 16923.2423603067,
    "disk_read": 0  }
}
```

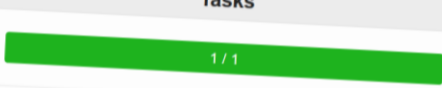# The Metrics

PCA projection to 3

variance in components

SWAN – Offline Analysis

```python
In [1]:  from pyspark.sql.functions import from_unixtime, when, col
         from pyspark.sql.types import *
         from pyspark.sql.functions import from_json
         from pyspark import SparkContext
```

```python
In [2]:  %matplotlib inline
         import pandas as pd
         import matplotlib.pyplot as plt
         import numpy as np
         import seaborn as sns
```

```python
In [3]:  df_raw = spark.read.json("/project/monitoring/archive/adam/agg/os/2018/06/30/*")
         df_raw.count()
```

▼ Apache Spark: 3 EXECUTORS   6 CORES   Jobs: 2 COMPLETED

Job ID	Job Name	Status	Stages	Tasks	Submission Time	Duration
▶ 0	json	COMPLETED	1/1	1 / 1	24 minutes ago	5s
▶ 1	count	COMPLETED	2/2	2 / 2	23 minutes ago	1s

```
Out[3]:  4421
```

SWAN – Offline Analysis

```
In [1]:   from pyspark.sql.functi
          from pys

In [6]:   from pyspark.ml import Pipeline
          from pyspark.ml.feature import VectorAssembler, StandardScaler, PCA

          NUM_OF_PCA_COMPONENTS = 7
          NUM_OF_CLUSTERS = 2

          metadata_columns = ["host", "window", "hostgroup", "environment", "availability_zone"]
          metrics = [m for m in list(df_pandas.columns.values) if m not in metadata_columns]

          vecAssembler = VectorAssembler(
              inputCols=metrics,
              outputCol="features")

          scaler = StandardScaler(
              inputCol=vecAssembler.getOutputCol(),
              outputCol="scaledFeatures",
              withStd=True, withMean=False)

          pca = PCA(k=NUM_OF_PCA_COMPONENTS, inputCol=scaler.getOutputCol(), outputCol="pcaFeatures")

          preprocessing_pipeline = Pipeline(stages=[vecAssembler, scaler, pca])
          preprocessing_fitted_pipeline = preprocessing_pipeline.fit(df)
          proc_df = preprocessing_fitted_pipeline.transform(df)

Out[3
```

					Submission Time	Duration
▼ Apache Spark:	3 EXECUTORS	6 CORES	Jobs:	5 COMPLETED		12s
				Tasks	23 minutes ago	1s
Job ID	Job Name	Status	Stages			
▶ 7	treeAggregate	COMPLETED	3/3	214 / 214		

```python
from pyspark.ml.clustering import KMeans, GaussianMixture, BisectingKMeans

gmm = GaussianMixture().setK(NUM_OF_CLUSTERS).setFeaturesCol(vecAssembler.getOutputCol()).setPredictionCol("gmm_predicti
kmns = KMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(vecAssembler.getOutputCol()).setPredictionCol("kmeans_prediction_un
bkm = BisectingKMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(vecAssembler.getOutputCol()).setPredictionCol("bkmeans_pred

gmm_scaled = GaussianMixture().setK(NUM_OF_CLUSTERS).setFeaturesCol(scaler.getOutputCol()).setPredictionCol("gmm_predict
kmns_scaled = KMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(scaler.getOutputCol()).setPredictionCol("kmeans_prediction")
bkm_scaled = BisectingKMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(scaler.getOutputCol()).setPredictionCol("bkmeans_pre

gmm_pca = GaussianMixture().setK(NUM_OF_CLUSTERS).setFeaturesCol(pca.getOutputCol()).setPredictionCol("gmm_prediction_pc
kmns_pca = KMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(pca.getOutputCol()).setPredictionCol("kmeans_prediction_pca")
bkm_pca = BisectingKMeans().setK(NUM_OF_CLUSTERS).setFeaturesCol(pca.getOutputCol()).setPredictionCol("bkmeans_predictio

pipeline = Pipeline(stages=[gmm, kmns, bkm, gmm_scaled, kmns_scaled, bkm_scaled, gmm_pca, kmns_pca, bkm_pca])

proc_df = proc_df.where("environment = 'production' and hostgroup = 'monitoring/kafka'")
fitted_pipeline = pipeline.fit(proc_df)
out_df = fitted_pipeline.transform(proc_df)
out_pd = out_df.toPandas()
```
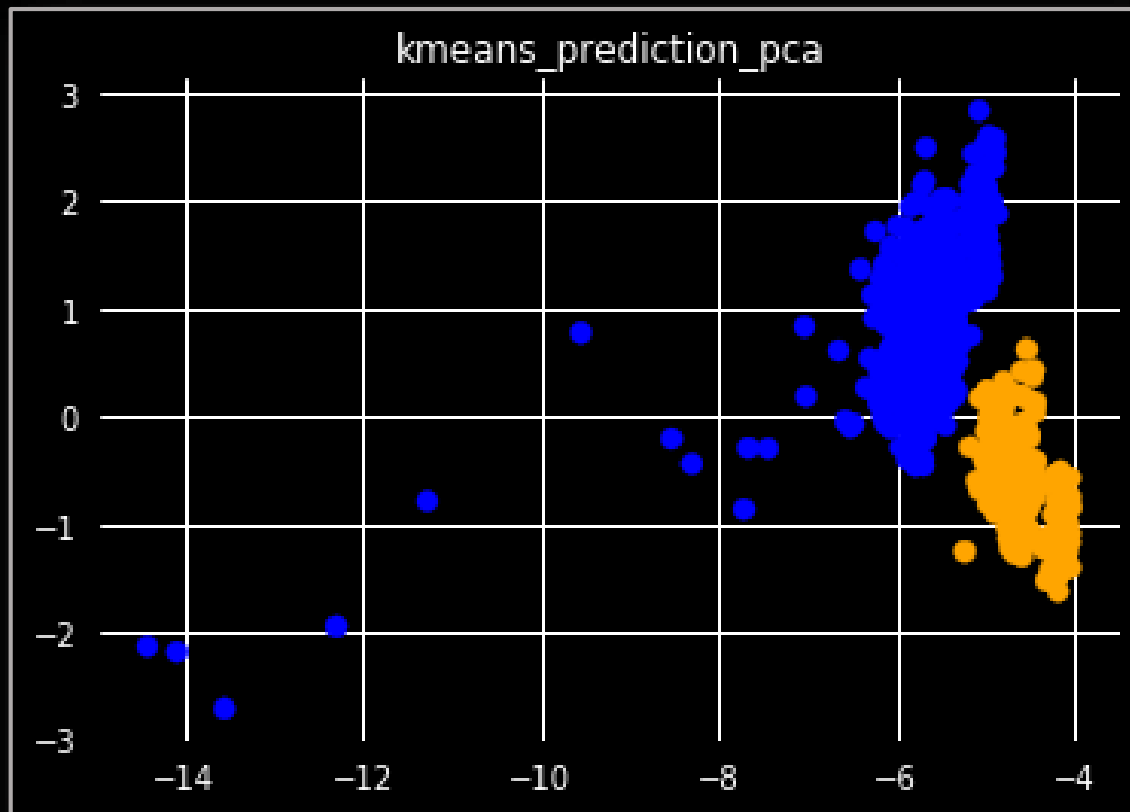
Apache Spark: `3 EXECUTORS` `6 CORES` **Jobs:** `156 COMPLETED`

Job ID	Job Name	Status	Stages	Tasks	Submission Time	Duration
12	first	COMPLETED	2/2	2 / 2	24 minutes ago	3s
13	takeSample	COMPLETED	1/1 (1 skipped)	200 / 200	24 minutes ago	4s
14	takeSample	COMPLETED	1/1 (1 skipped)	200 / 200	24 minutes ago	1s
15	treeAggregate	COMPLETED	2/2 (1 skipped)	213 / 213	24 minutes ago	4s
7	treeAggregate	COMPLETED				

The Data Analysis (so far)

kmeans_prediction_pca

Roadmap

Time series model (trained on cluster, applied per server)

Model with taking inputs from the whole cluster and predicting whether one server belongs

Exploring other workloads

Summary

- Streaming data aggregation and consolidation

- Offline (for now) analysis based only on HW metrics

- PCA for more compact data

- Clustering spots outliers → There is value in the data

- More complex models to be investigated and applied in online analysis

Special Thanks To:

Luca Magnoni and the CERN IT-CM-MM section

Prasanth Kothuri and the Spark & Hadoop team

D.Adamová & M.Pilát

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS