BD|CESGA
Anomaly Detection at Scale

Javier López Cacheiro
Orquídea Seijas Salinas
Samuel Soutullo Sobral

Measurement is the first step that leads to control and eventually to improvement. If you can’t **measure** something, you can’t understand it. If you can’t **understand** it, you can’t control it. If you can’t **control** it, you can’t **improve** it.

H. James Harrington

Anomaly Detection

- Phase 1: Measure → Collect & Store
- Phase 2: Understand → Analyze&Visualize
- Phase 3: Control → Monitoring
- Phase 4: Improve → Anomaly Detection

Phase 1: Measure

Collect & Store

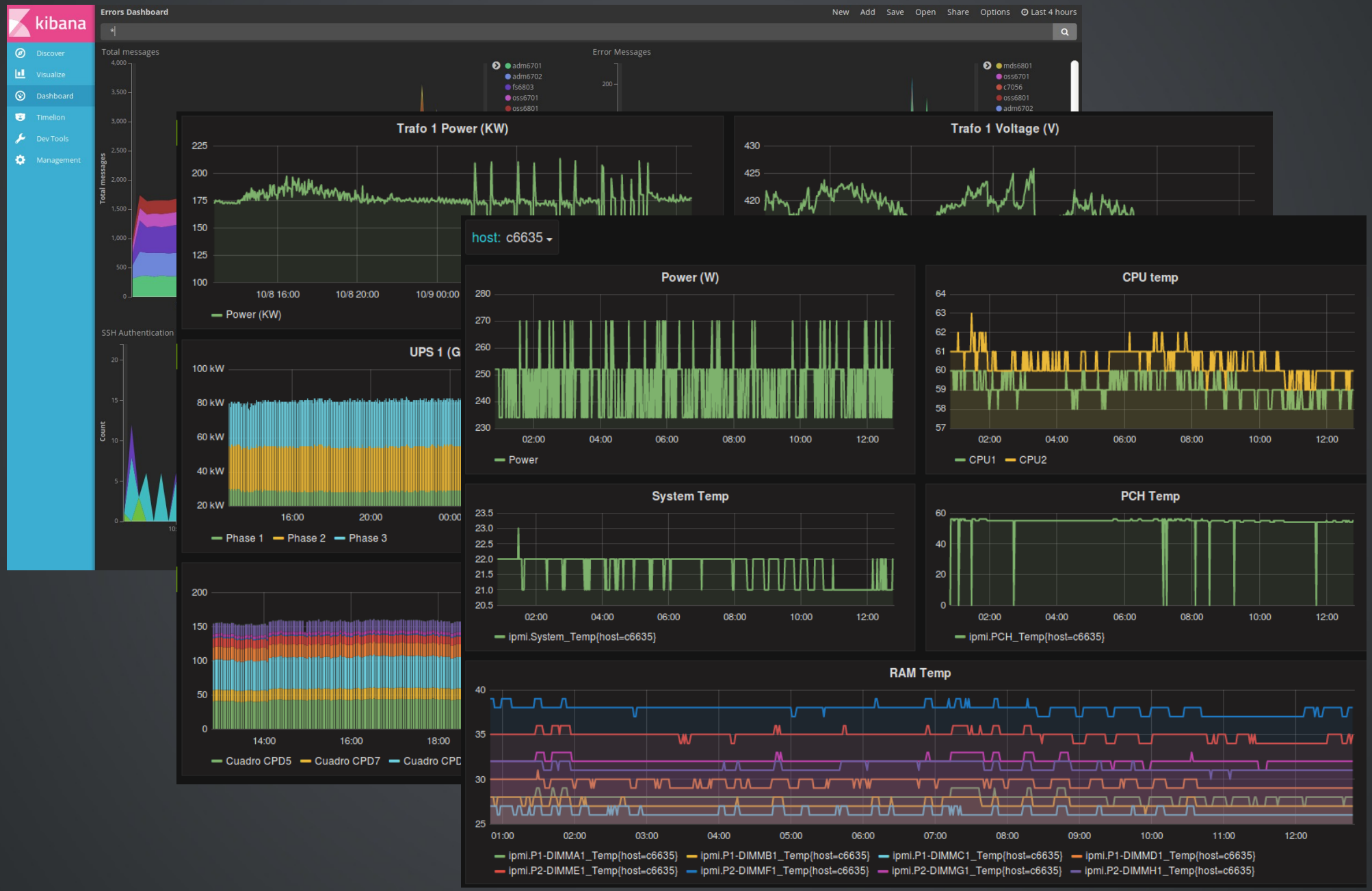
- System logs
- Process accounting
- Server metrics
- IPMI sensors
- SMART disk info
- Chillers & AHUs
- UPSs
- Power Meters
- Network traffic
- Application module usage
- SLURM Accounting
- Temperature/ Humidity sensors

33487 metrics

10 Million time series

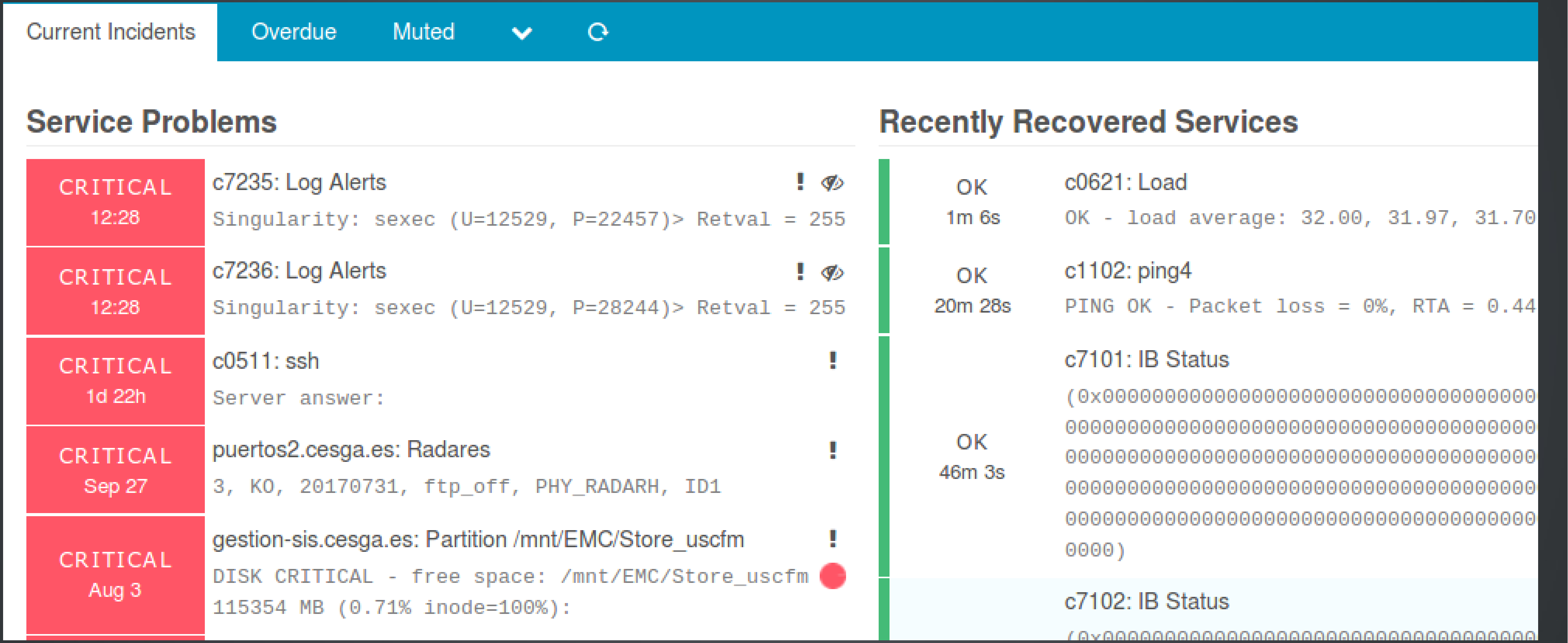
Phase 2: Understand

Dashboards


kibana
Errors Dashboard
New
Add
Save
Open
Share
Options
Last 4 hours
Discover
Visualize
Dashboard
Timelion
DevTools
Management
Total messages
Error Messages
SSH Authentication
Trafo 1 Power (KW)
Power (KW)
Trafo 1 Voltage (V)
UPS 1 (G
Phase 1
Phase 2
Phase 3
Cuadro CPD5
Cuadro CPD7
host: c6635
Power (W)
Power
CPU temp
CPU1
CPU2
System Temp
ipmi.System_Temp{host=c6635}
PCH Temp
ipmi.PCH_Temp{host=c6635}
RAM Temp
ipmi.P1-DIMMA1_Temp{host=c6635}
ipmi.P1-DIMMB1_Temp{host=c6635}
ipmi.P1-DIMMC1_Temp{host=c6635}
ipmi.P1-DIMMD1_Temp{host=c6635}
ipmi.P2-DIMME1_Temp{host=c6635}
ipmi.P2-DIMMF1_Temp{host=c6635}
ipmi.P2-DIMMG1_Temp{host=c6635}
ipmi.P2-DIMMH1_Temp{host=c6635}

Phase 3: Control

Monitoring



Service Problems

CRITICAL 12:28 — c7235: Log Alerts

```
Singularity: sexec (U=12529, P=22457)> Retval = 255
```

CRITICAL 12:28 — c7236: Log Alerts

```
Singularity: sexec (U=12529, P=28244)> Retval = 255
```

CRITICAL 1d 22h — c0511: ssh

```
Server answer:
```

CRITICAL Sep 27 — puertos2.cesga.es: Radares

```
3, KO, 20170731, ftp_off, PHY_RADARH, ID1
```

CRITICAL Aug 3 — gestion-sis.cesga.es: Partition /mnt/EMC/Store_uscfm

```
DISK CRITICAL - free space: /mnt/EMC/Store_uscfm
115354 MB (0.71% inode=100%):
```

Recently Recovered Services

OK 1m 6s — c0621: Load

```
OK - load average: 32.00, 31.97, 31.70
```

OK 20m 28s — c1102: ping4

```
PING OK - Packet loss = 0%, RTA = 0.44
```

OK 46m 3s — c7101: IB Status

```
(0x00000000000000000000000000000000000
00000000000000000000000000000000000000
00000000000000000000000000000000000000
00000000000000000000000000000000000000
00000000000000000000000000000000000000
0000)
```

c7102: IB Status

```
(0x00000000000000000000000000000000000
```

Phase 4: Improve

Anomaly Detection

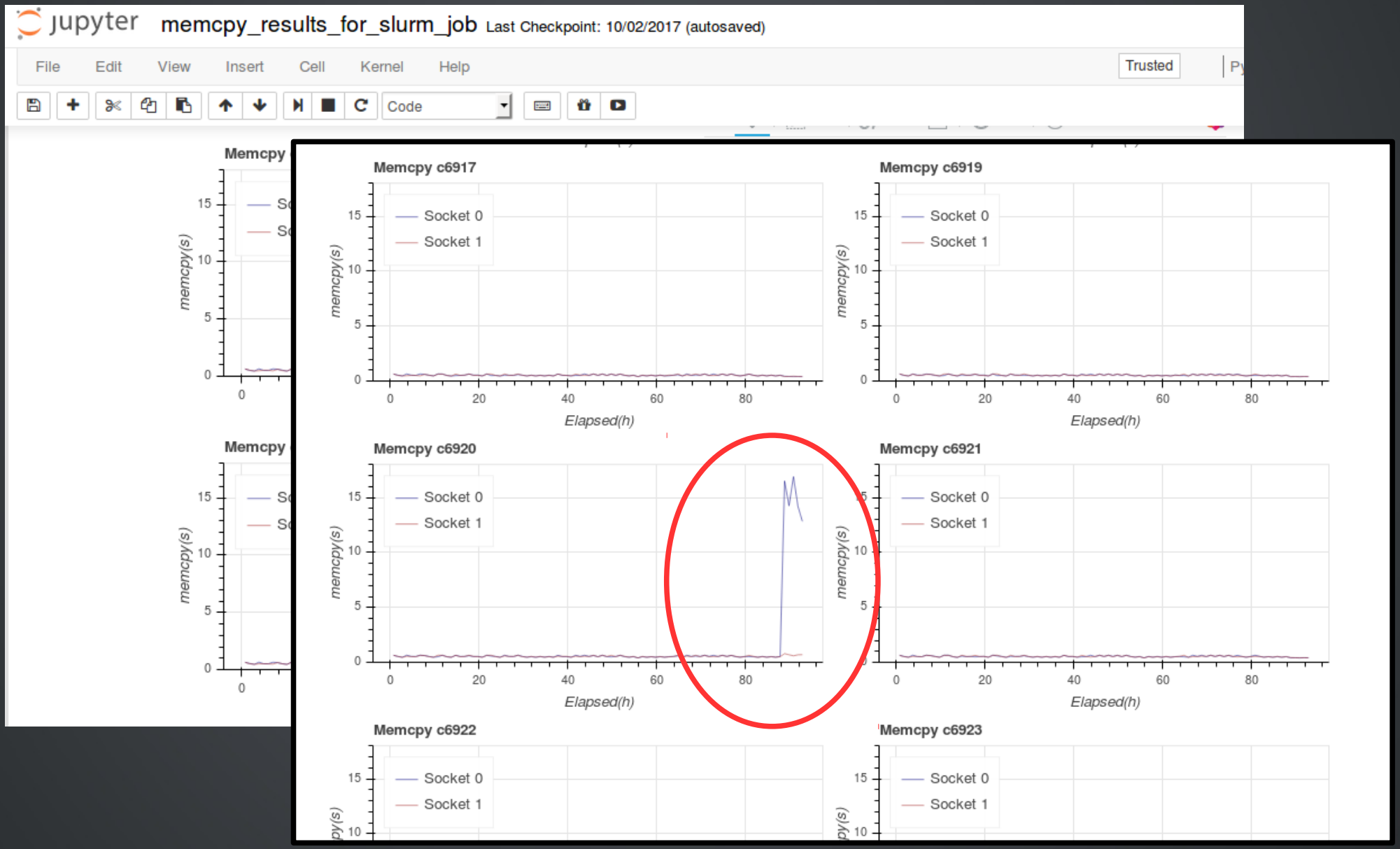
- Types of anomalies in time series
  - Outliers
  - Change points
  - Anomalous time series
- Generic Anomaly Detection Systems

Caution: Alert overload

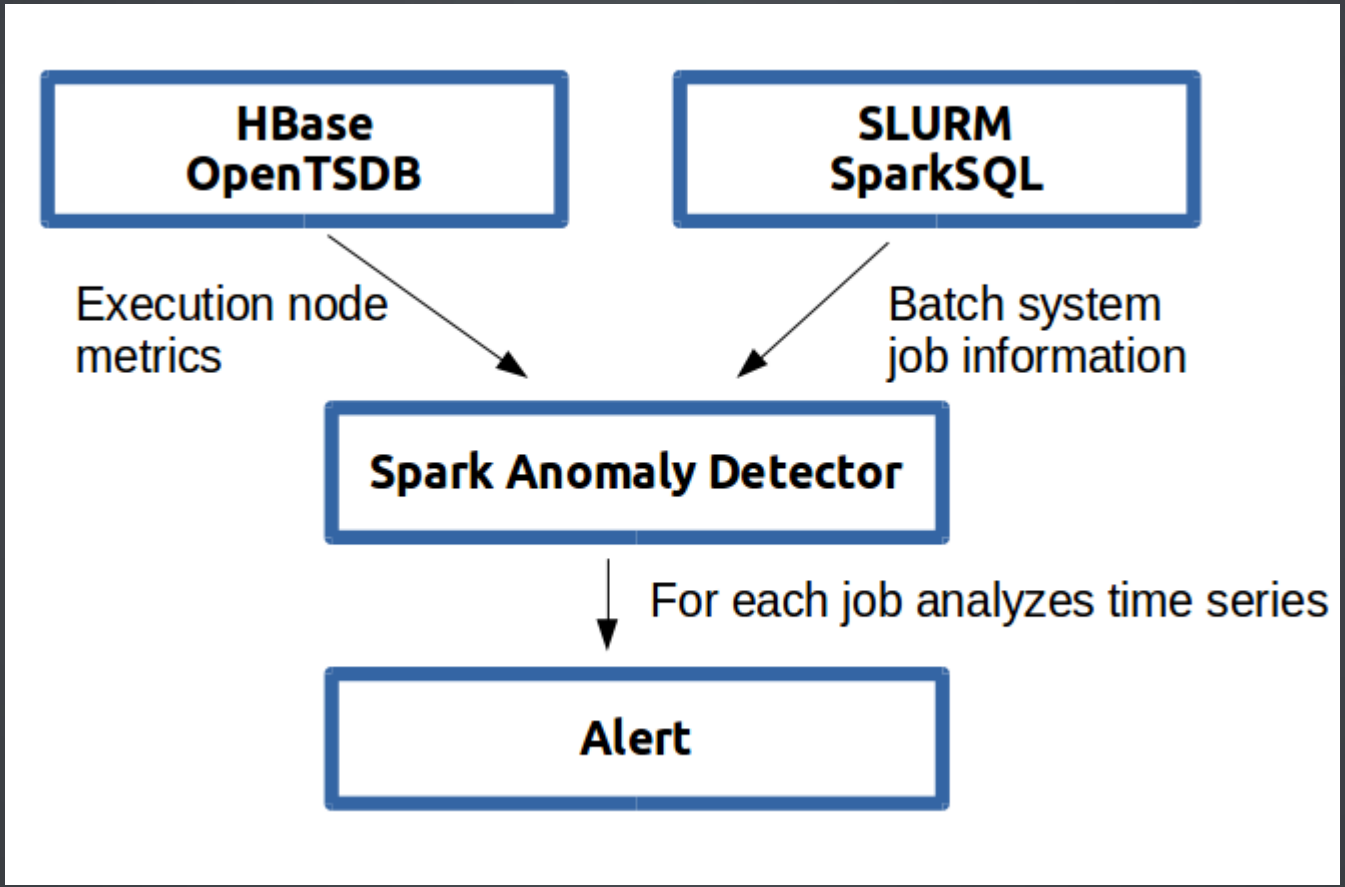
Server Anomalous Performance

- **Problem:** Parallel jobs are cancelled because some of the nodes have poor performance. Computation is lost.
- **Detection:** Analyze & visualize server metrics to spot the anomalous node
- **Objective:** Automatically detect low performance nodes

Analyze & Visualize



Anomalous Performance Detection



Results

- 6 months
- 11965 jobs
- 22 anomalies detected
  - Precision: 100%
  - Recall: 96%
  - F-score: 0.98

Conclusions

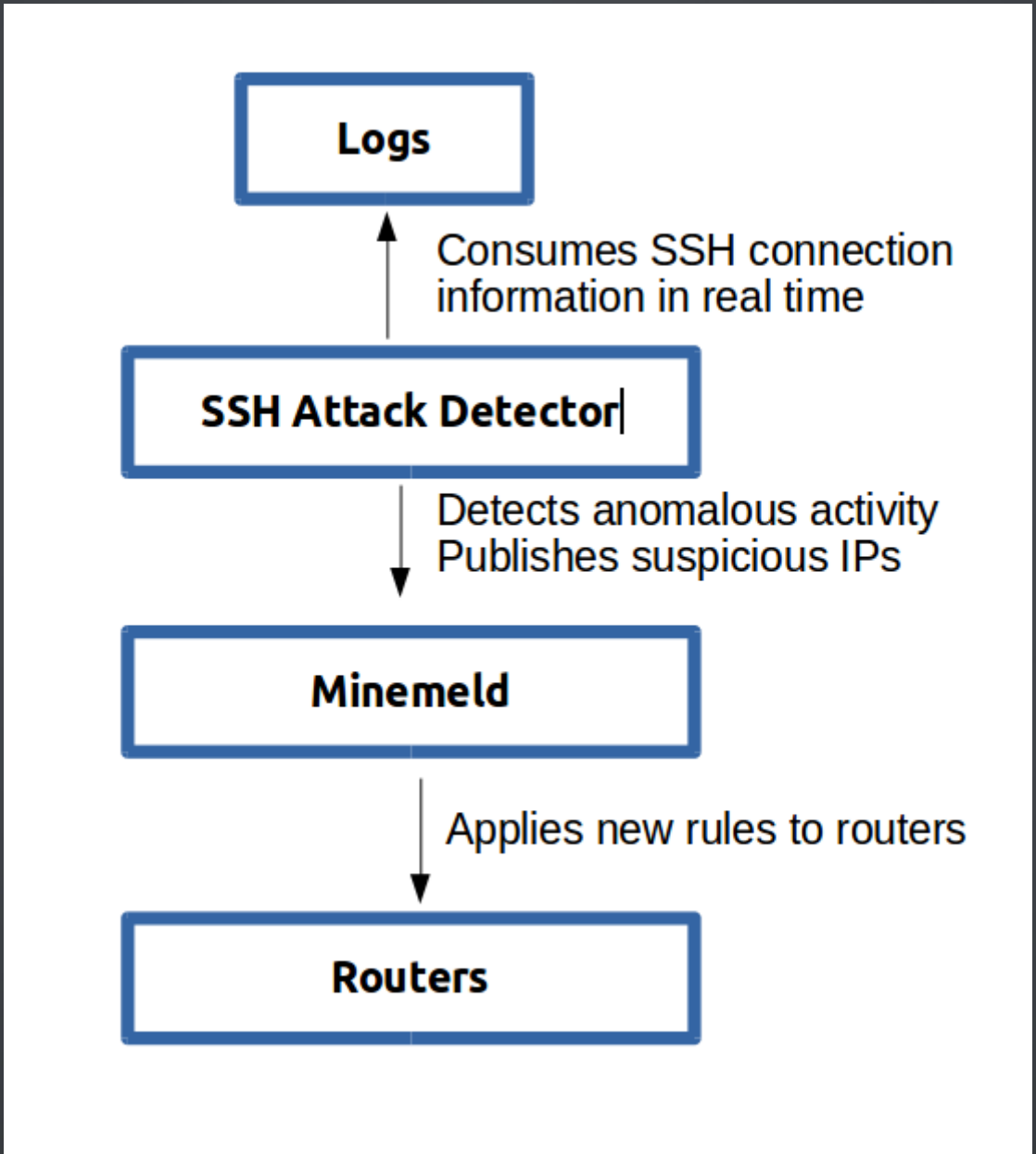
- No longer needed to delete old data
- Understand your data
- Generic anomaly detecion systems generate too many alerts
- Target specific use cases
- Maintain number of alerts low

Thanks!

SSH Attack Detection

- **Problem:** Daily our public servers are scanned and attacked
- **Detection:** Correlate real-time SSH connection information to detect attacks
- **Objective:** Automatically update router configuration to stop the attacks

SSH Attack Detection



Results

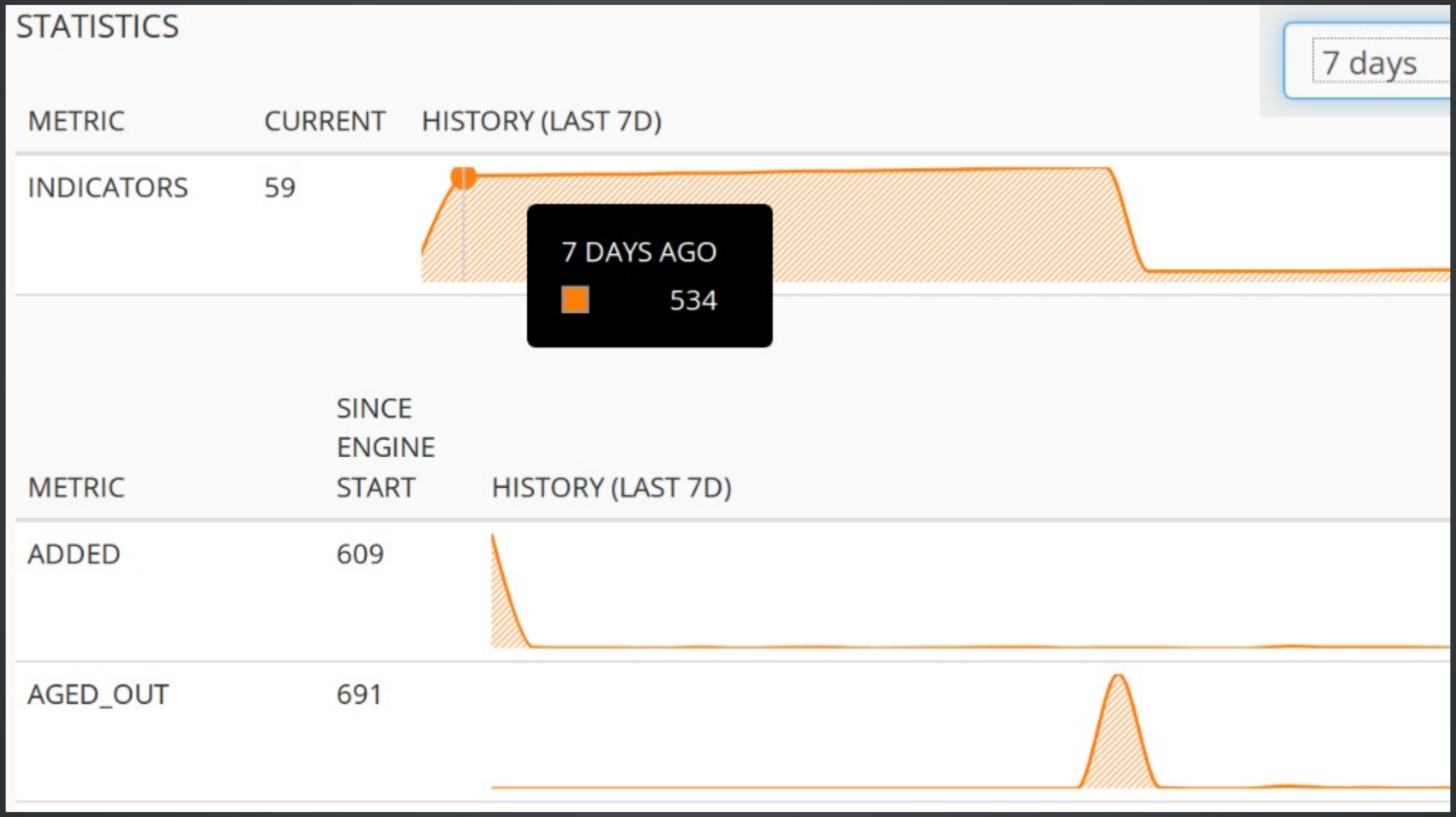
STATISTICS

24 hours

METRIC	CURRENT	HISTORY (LAST 24H)
INDICATORS	59	

METRIC	SINCE ENGINE START	HISTORY (LAST 24H)
ADDED	609	
AGED_OUT	691	

Results


STATISTICS
7 days
METRIC
CURRENT
HISTORY (LAST 7D)
INDICATORS
59
7 DAYS AGO
534
METRIC
SINCE ENGINE START
HISTORY (LAST 7D)
ADDED
609
AGED_OUT
691