

# Lemon for Quattor

I.Fedorko

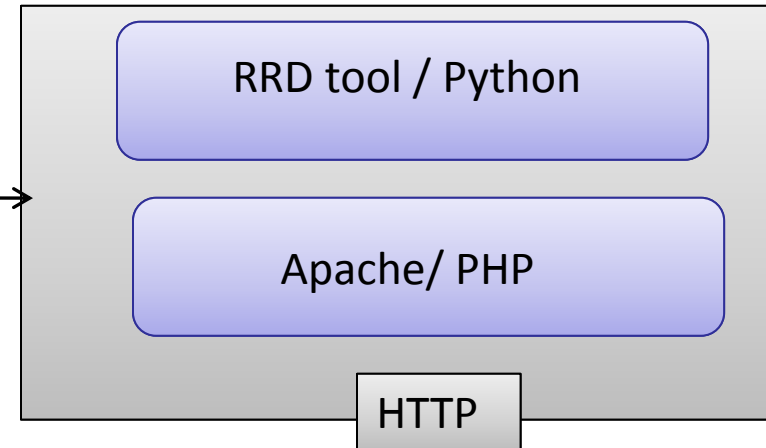
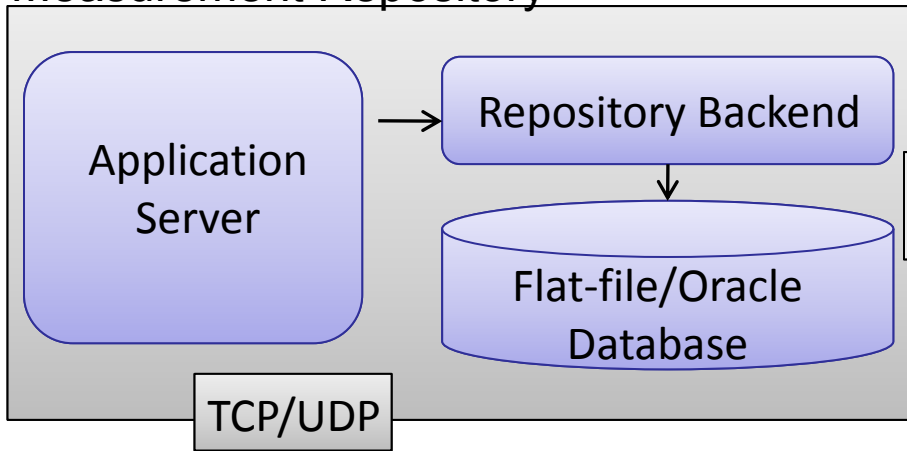
CERN CF/IT

16 March 2011

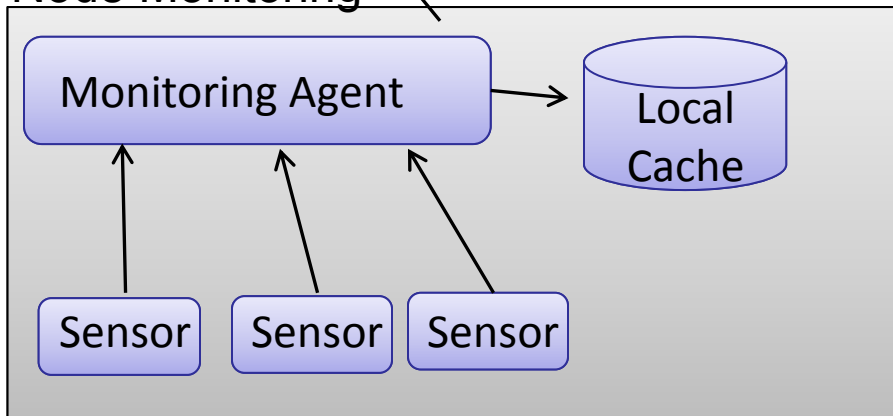
- Lemon Monitoring System
- Lemon in use
  - focus on CERN Computer Centre
- Current activities:
  - Federation
  - Monitoring of Virtualization
  - New Lemon Web
- Future perspective
- Monitoring Questionnaire



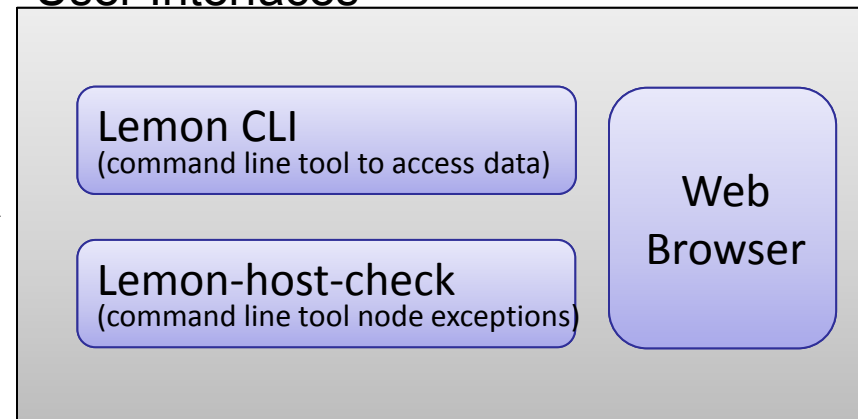
## Measurement Repository

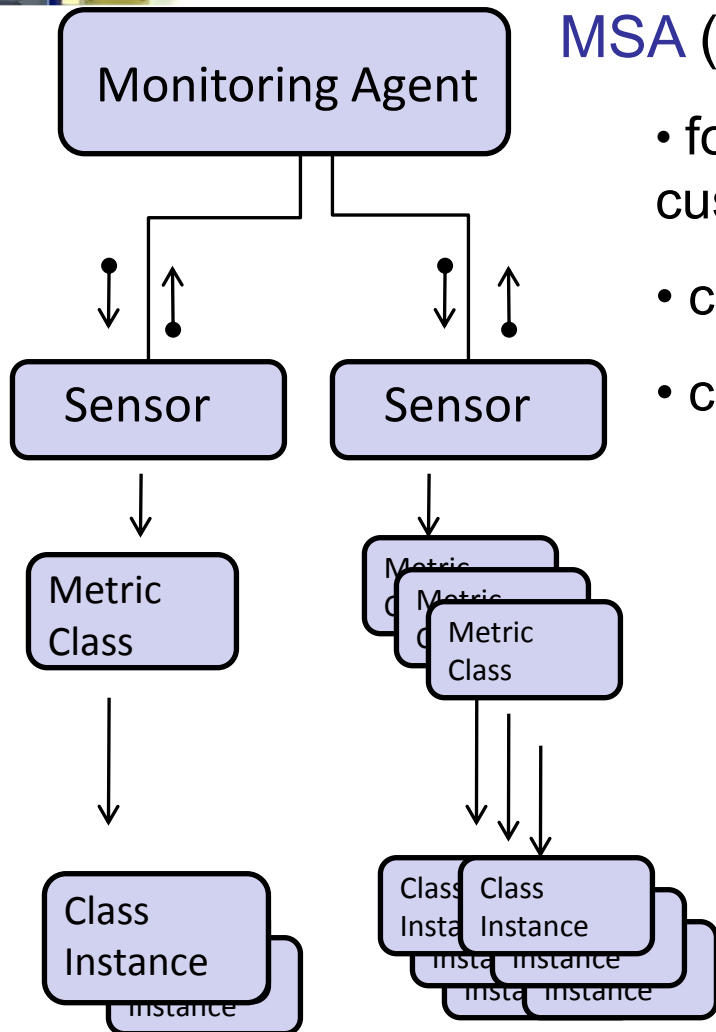


## Node Monitoring



## User Interfaces





## MSA (Monitoring Sensor Agent)

- forks sensors and communicate with them using custom protocol over a bi-directional “pipes”
- checks on status of sensors
- caches data locally ( e.g. /var/spool/lemon-agent/ )

## Sensor:

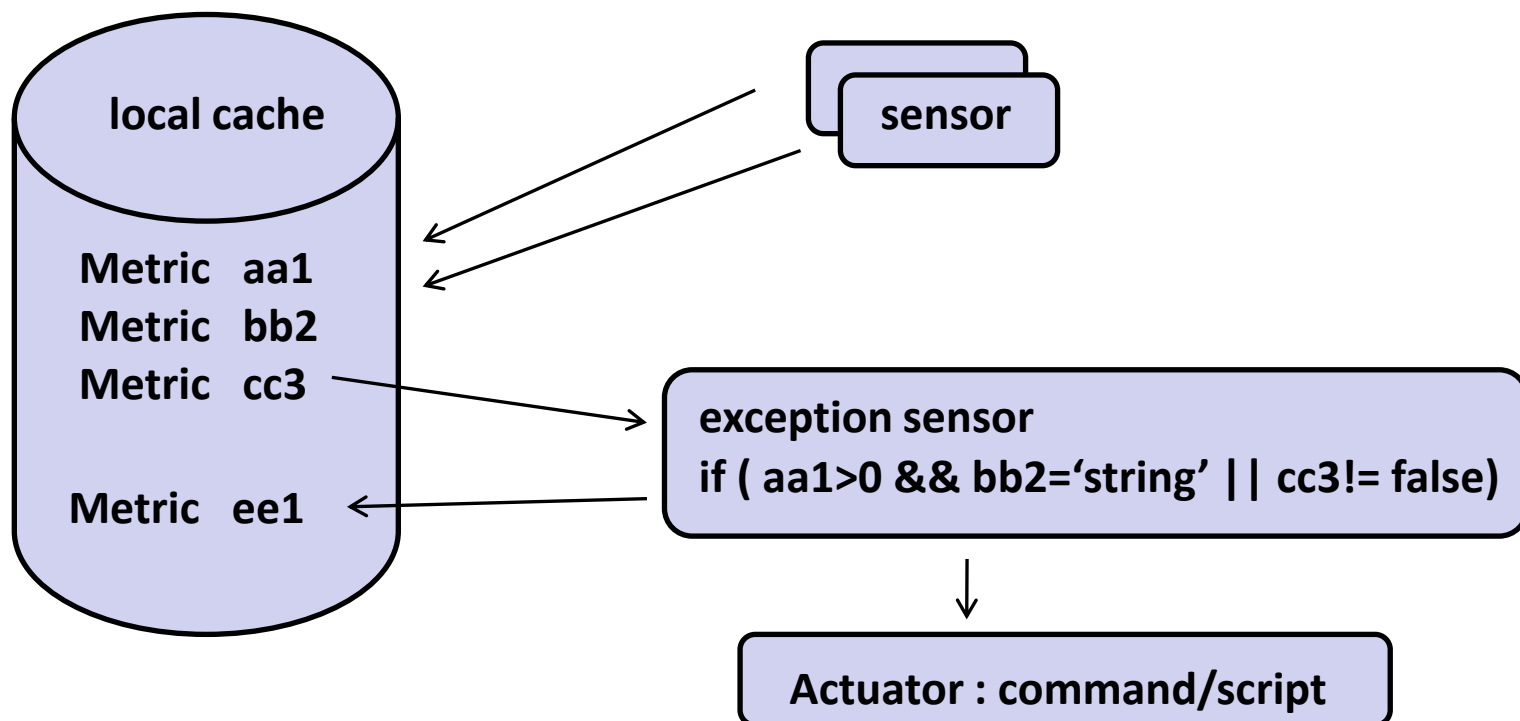
- process or script
- connected to the lemon-agent via a bi-directional pipe
- collects information on behalf of the agent

Example of linux sensor

```
for (i = 0, j = 0; i <= strlen(model); i++) {  
    if (model[i] == ' ')  
        r_model[j]= '_';  
    else  
        r_model[j]=model[i];  
    j++;  
}  
bogomips = bogomips * cpus;  
  
/* store sample */  
snprintf(value, sizeof(value) - 1, "%s %.0f %d", vendor, clockspeed, logical);  
storeSample01(value);
```

← Lemon API (perl, c++, for python a wrapper)

- an officially supported Lemon sensor coded in C++  
on the node



Full documentation at:

- <http://lemon.web.cern.ch/lemon/doc/sensors/exception.shtml>

- Objective:
  - To run corrective action when the occupancy of the /tmp partition is greater than 80%
- Involved Metrics
  - With ID 9104 (system.partitionInfo)
  - Field 1 = mountname, field 5 = percentage occupancy

```
[ifedorko@]~% lemon-cli -m 9104
[INFO] lemon-cli version 1.2.0 started by ifedorko at Thu Dec 16 09:06:48 2010 on lxadm11.cern.ch
[VERB] Nodes:    lxadm11
[VERB] Metrics:   9104
[VERB] Method:    Latest
[VERB]
[INFO] local: lxadm11 9104 Thu Dec 16 09:02:56 2010 / ext3 rw 30470176 17 -1 -1 2 1
[INFO] local: lxadm11 9104 Thu Dec 16 09:02:56 2010 /boot ext3 rw 2030736 3 -1 -1 1 1
[INFO] local: lxadm11 9104 Thu Dec 16 09:02:56 2010 /tmp ext3 rw 162392496 1 -1 -1 1 1
[INFO] local: lxadm11 9104 Thu Dec 16 09:02:56 2010 /usr/vice/cache ext3 rw 10153988 36 -1 -1 2 1
[INFO] local: lxadm11 9104 Thu Dec 16 09:02:56 2010 /var ext3 rw 15235040 16 -1 -1 1 1
[VERB]
[VERB] Total: 5 results
```

- Correlation

**Correlation ((9104:1 eq '/tmp') && (9104:5 > 80))**

**Actuator /usr/local/sbin/clean-tmp-partition -o 75**

## Lemon instances

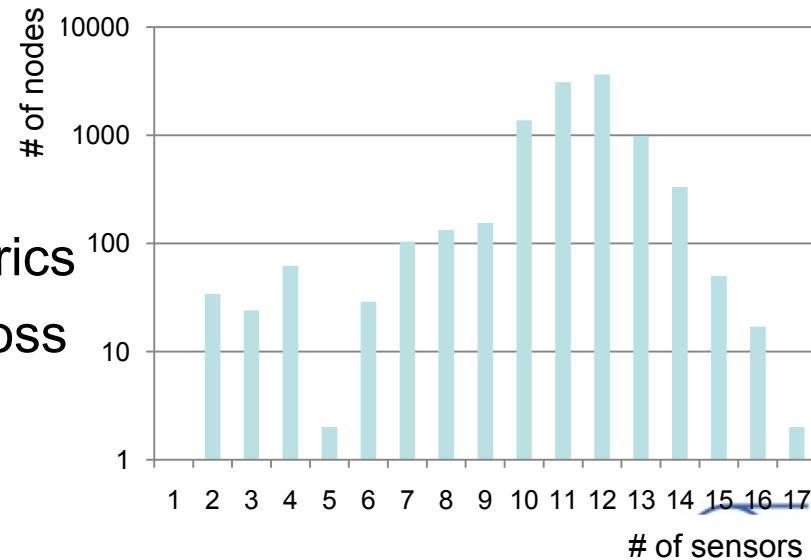
- CNAF, Morgan Stanley (Oracle backend)
- CERN:
  - Alice DAQ, LHC control (flat-file backend)
  - Computer Center (Oracle backend)
    - Lemon Alarm System
    - Service Level status

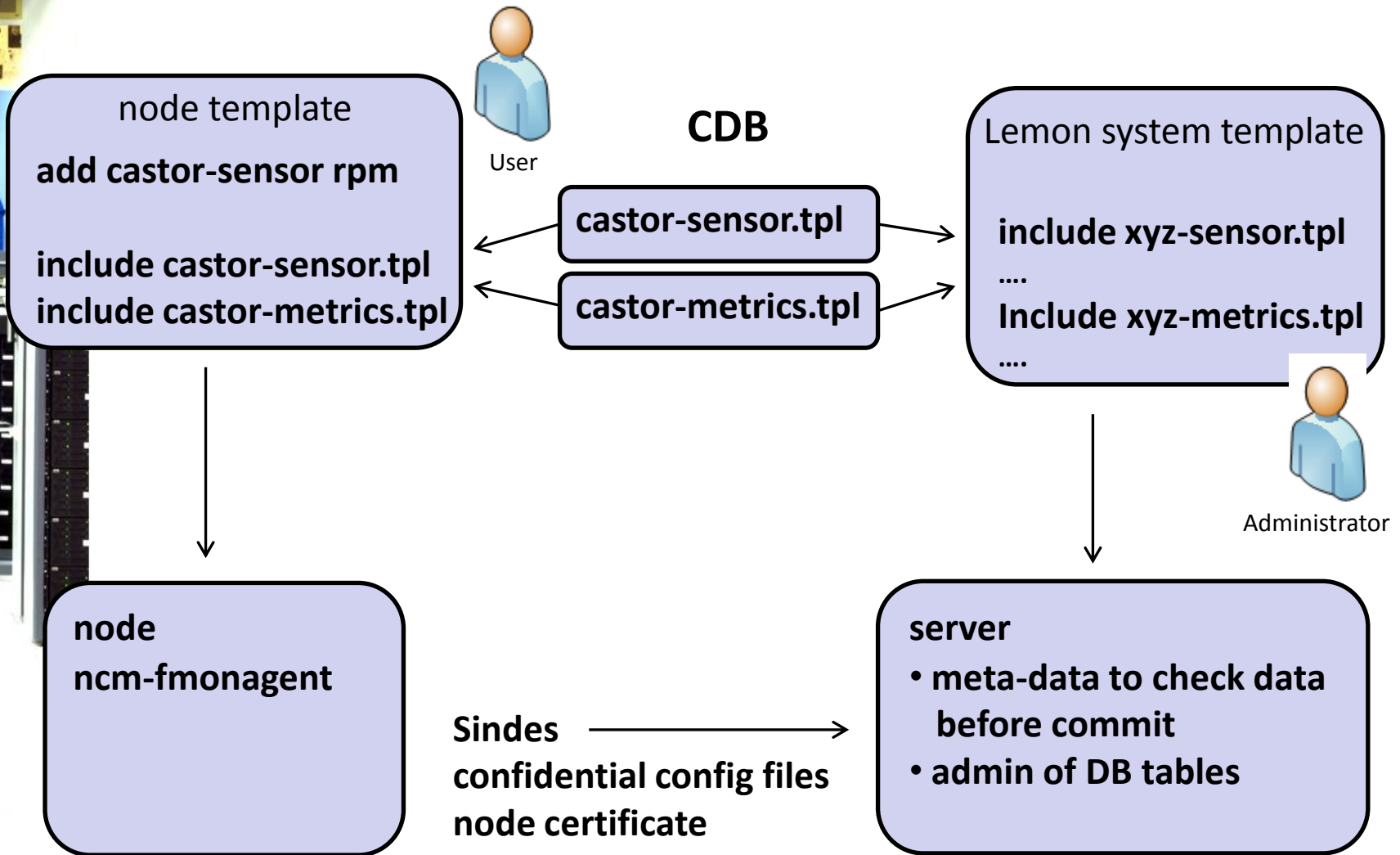
Missing MySQL?

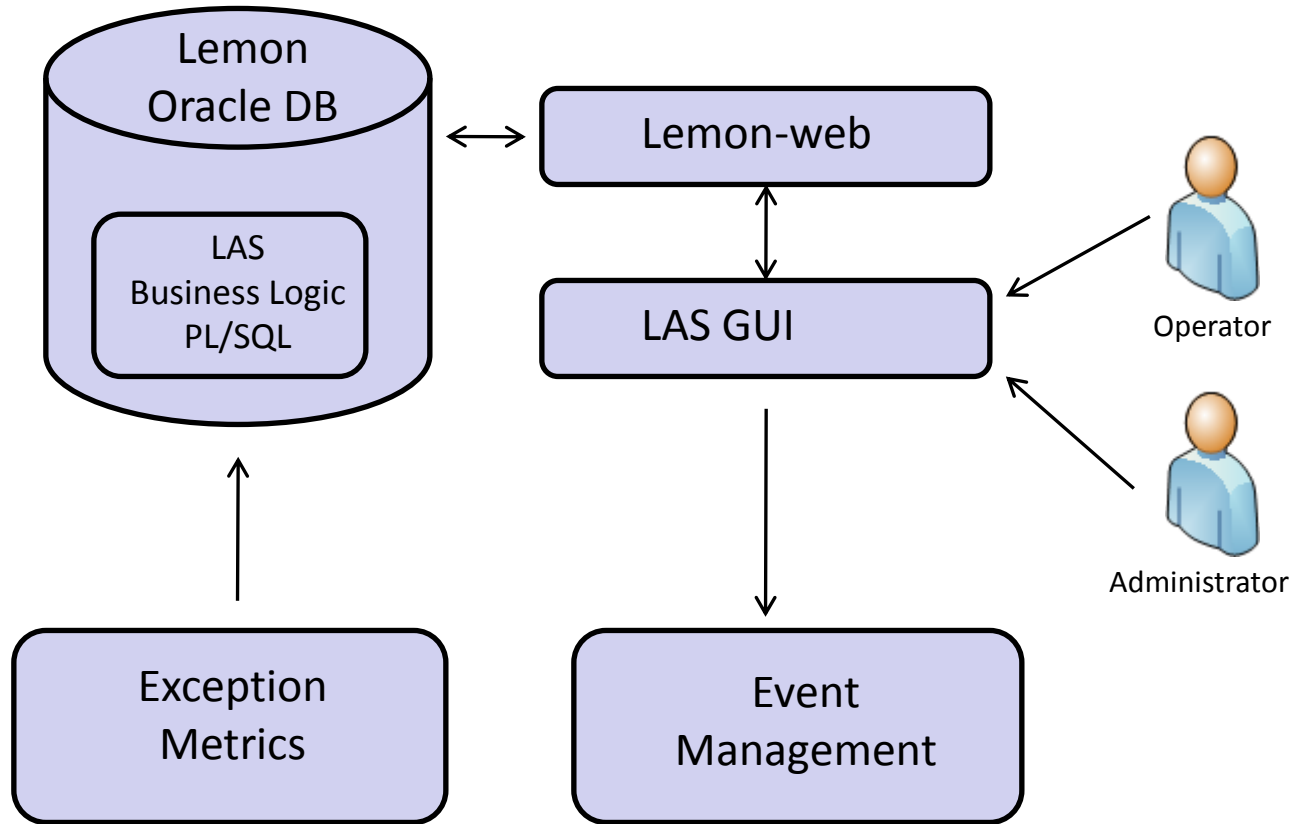
Current development is reflecting DB flavor independent policy  
→ to implement functionality is up to the external user



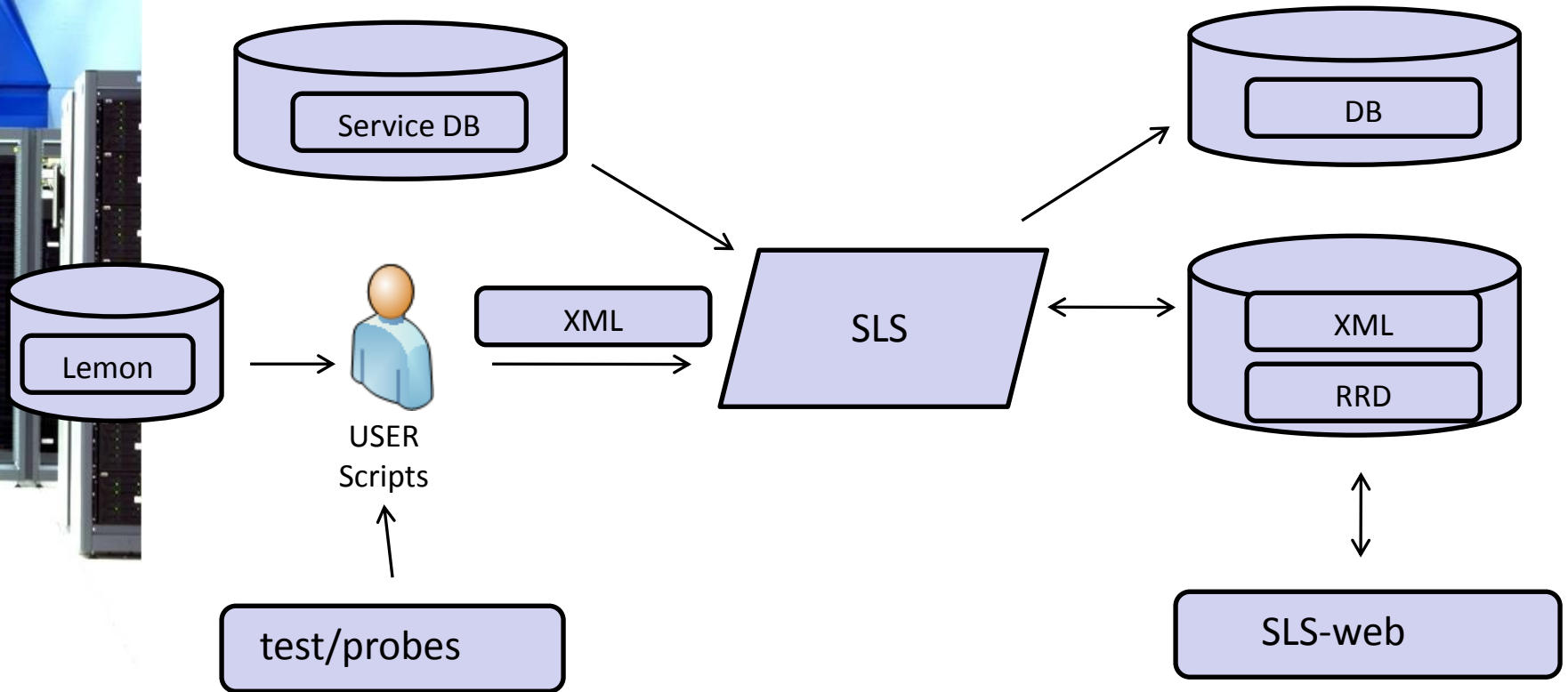
- ~11k monitored entities (~8k nodes)
  - performance monitoring
    - CPU Load, partitions information
  - application monitoring
    - File, log parsing
  - power, temperature
  - remote (ping, http, snmp)
- 5 core sensors covering ~60% of performance and application monitoring
- ~30 misc sensor
  - hw\_scan, snmp, castor
- >5000 nodes with 150-200 metrics
- ~1.7 M monitored 'services' across CERN CC
- Oracle backend
  - throughput 300GB/month





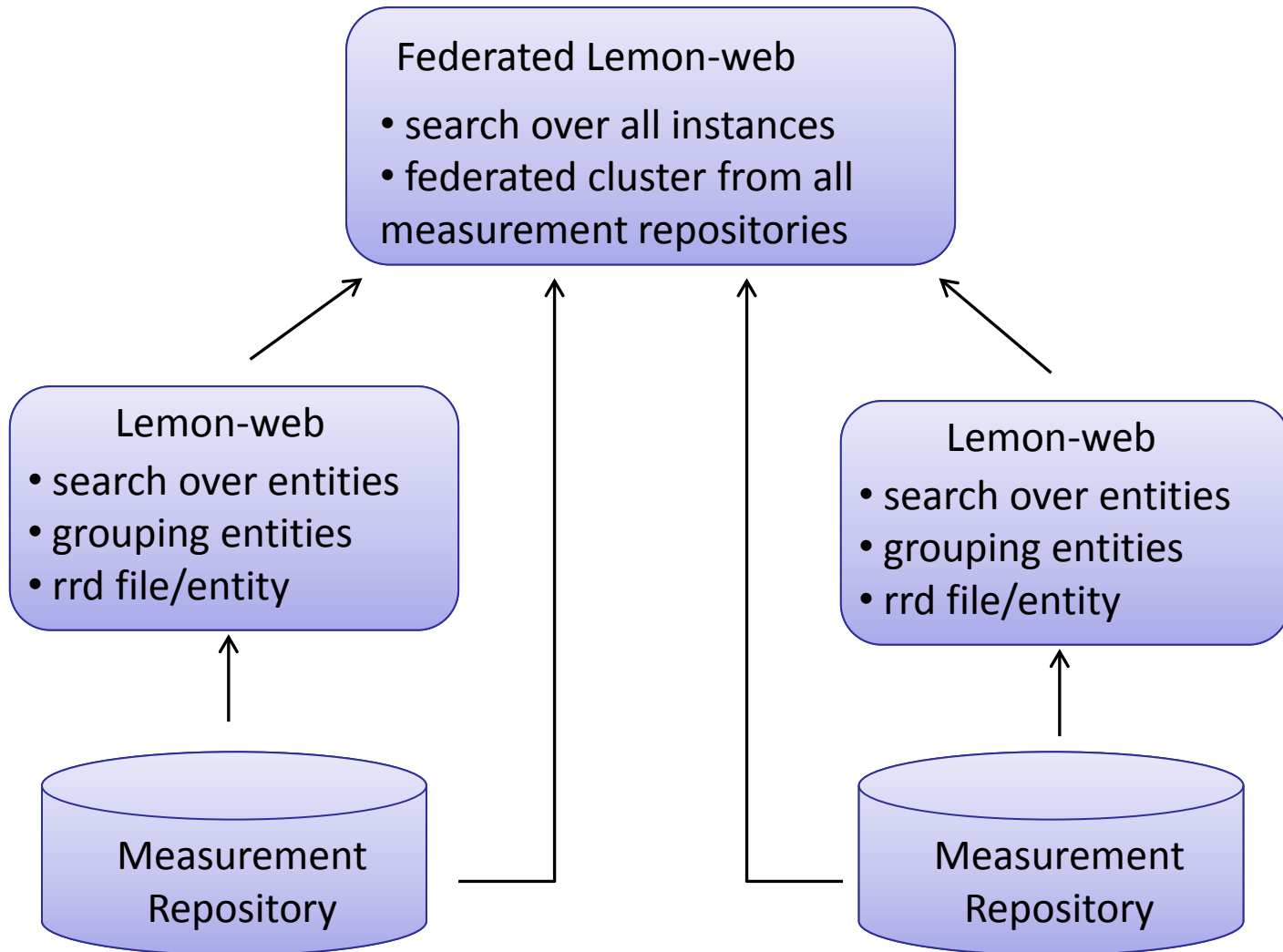


<https://sls.cern.ch/sls/>



Used by RAL and MS

- Federation of instances
- Monitoring Virtualization @CERN
- New Lemon-web



- Quattor managed nodes are authenticated in Lemon by Sintes certificates
  - Lemon fetches certificates from Sintes
- virtualized batch nodes generated from golden node
  - golden node is Quattor managed
  - 'golden' certificate is shared by generated nodes
- addressed by lemon admin tool → no new code needed
- monitoring of hypervisors addressed by customized sensors

More details:

<https://twiki.cern.ch/twiki/bin/viewauth/CloudServices/>

Don't hesitate to contact <Ulrich.Schwickerath at cern.ch>

http://lemonweb.cern.ch/lemon-web/info.php?time=2&offset=0&entity=lxmred0803&type=host



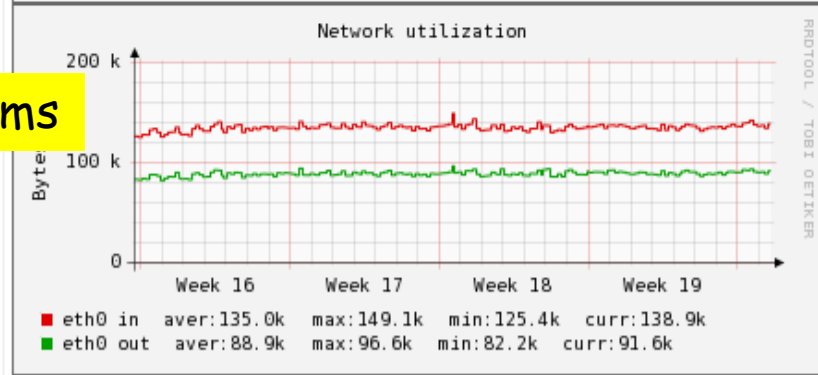
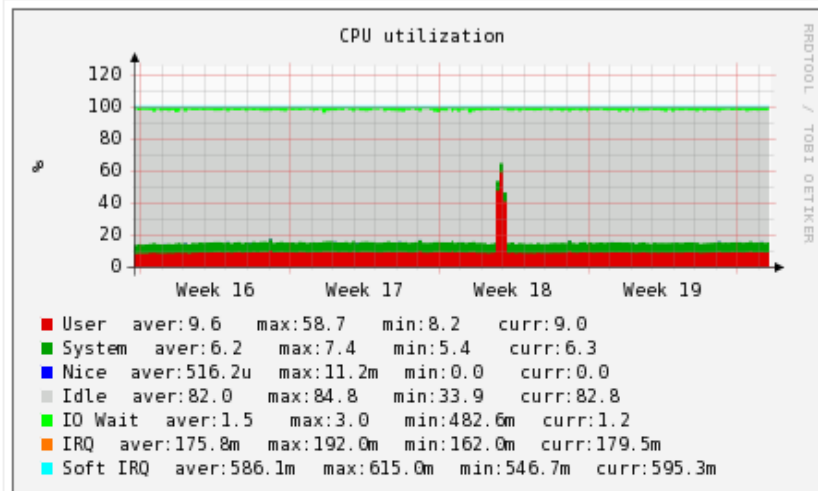
Lemon Monitoring

- Home
- Documentation
- Alarms
- Metrics
- Misc
- Help

## Information for host: lxmred0803

### Host information

<b>operating system(s)</b>	Scientific Linux CERN SLC release 4.8 (Beryllium)
<b>architecture (kernel)</b>	x86_64 (2.6.9-89.0.20.EL.cernsmp)
<b>up time (since)</b>	68 days, 0h:43m (Thu, 11 Mar 2010 14:19:17 +0100)
<b>CPU (count/logical)</b>	Intel(R) Xeon(TM) CPU 3.00GHz (1/2)
<b>memory (swap)</b>	3946 MB (4095 MB)
<b>cluster (subcluster)</b>	<b>lemonsrv (lemonsrv/server)</b>
<b>IP address(es)</b>	137.138.128.195 (eth0)
<b>state</b>	production
<b>status</b>	Available



alarms RSS

host exceptions & alarms

CDB templates

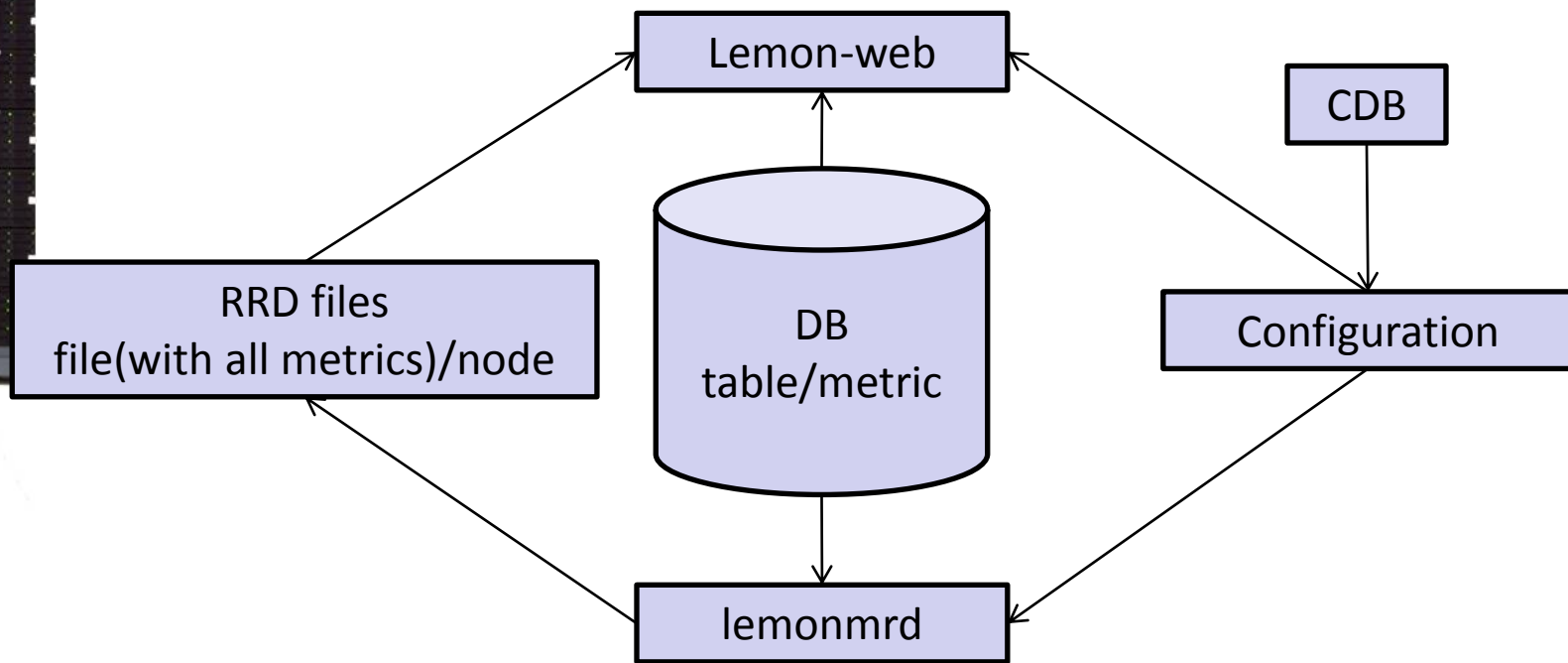
metrics raw data graphing

List of all metrics on host and latest values stored in central DB



It has two parts:  
Lemonmrd  
Lemon-web

Retrieve, Store, Display information



## Lemon-mrd

- enhanced configuration
- **new math operations (-, \*, / ) for dynamic cluster data aggregation**
  - in current version only sum and average
- data aggregation from multiple DB backends

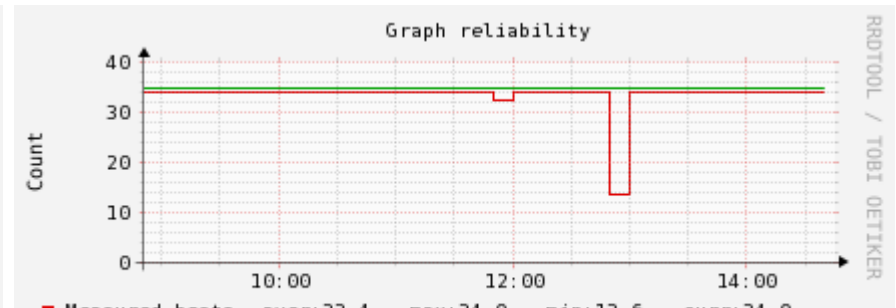
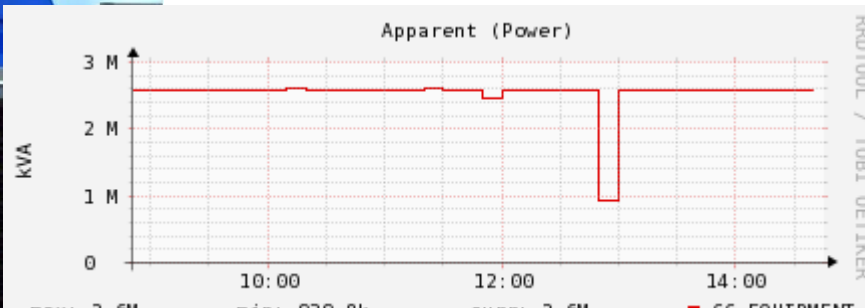
## Lemon-web

- flexible menu structure
- personalized views
- url pointing at graphs, can be embedded in any pages
- connecting to multiple DB sources

Demo available on demand

Graph reliability:

How many entities are reported from the expected?



Lemon Monitoring

Home | Documentation | Alarms | Metrics | Misc | Help

### Information for Power B513 / POWER Totals / CC EFFICIENCY

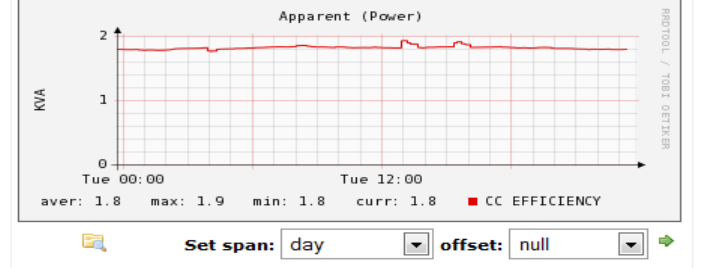
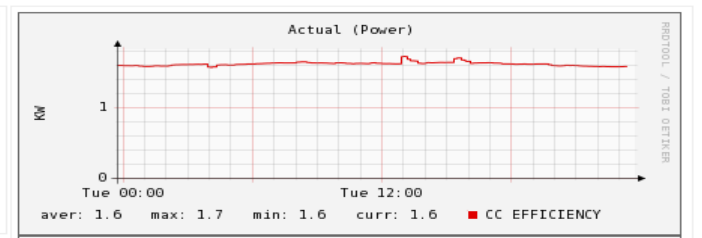
**Cluster information**

select from power menu:  None selected

parent: **POWER Totals,**

Formula(kW): "CC TOTAL"(3872kW) / "CC IT EQUIPMENT" (2444kW)

Formula(kVA): "CC TOTAL"(4646kVA) / "CC IT EQUIPMENT" (2580kVA)



Search entities:  Virtual Clusters | Clusters | Racks | Hardware types | KVM | Services



## Consolidation of monitoring activities

@ CERN IT Department

→ grid, windows, computer centre, experiment support, etc.

## Lemon enhancements under consideration

- New Lemon DB schema
- Lemon repository data export
  - export to data files
- Service monitoring with alarming and classification
- Remote instances
- Interfacing with Nagios → any experience to share is welcome
- Interfacing with Windows monitoring (both sides)



## 3 answers:

- instance size: 150-1500 nodes
- preferred solution: Ganglia + Nagios
- effort 0.2 to 0.5 FTE
  - additional effort for sensor/probes enhancement
- alarming by Nagios
- up to 15GB of data (per year ?)
- no history beyond rrd
- Lemon used for raid controllers (Belgium grid)

From now on backup

- an officially supported Lemon sensor coded in C++

## Main characteristic

- access agent local cache
- **correlation engine** allows to evaluate 1 or more metrics values to detect problem on a node
- supports reporting on behalf of other monitored entities
- allows corrective actions → **actuators**
- output is an exception

Full documentation at:

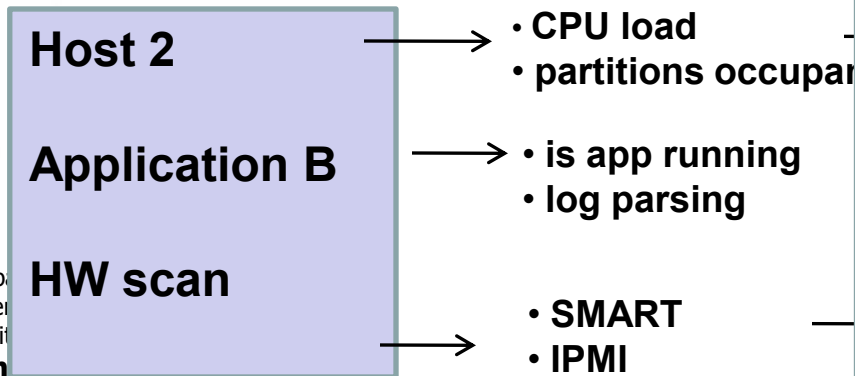
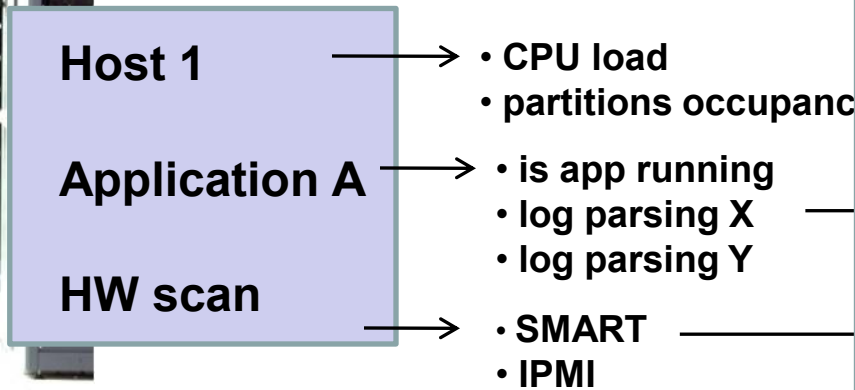
- <http://lemon.web.cern.ch/lemon/doc/sensors/exception.shtml>

Run-time data aggregation →

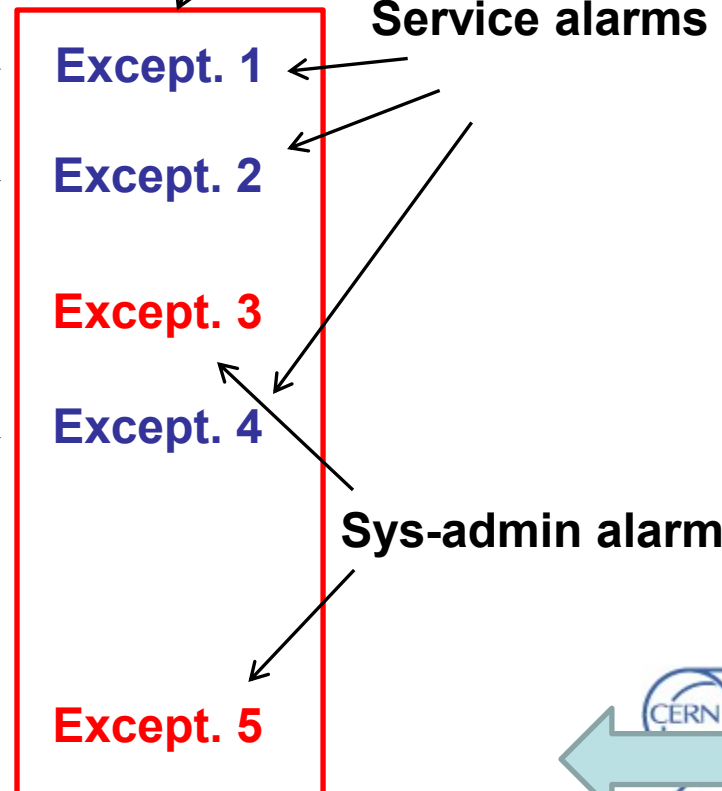
- CPU load over cluster

Resource utilization by service →

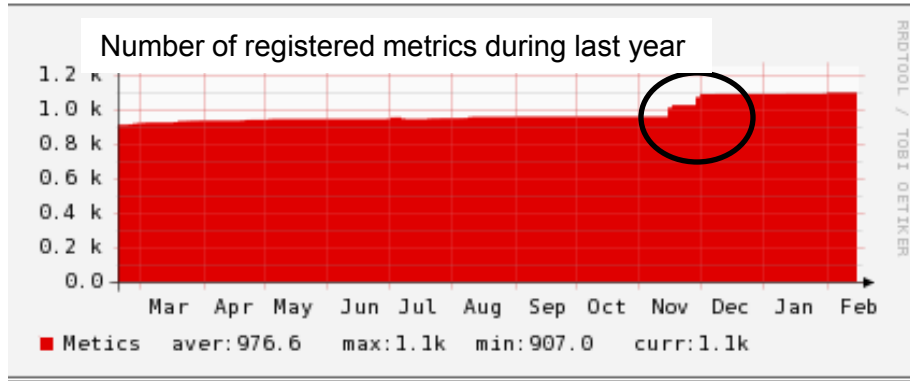
- Castor load by experiment



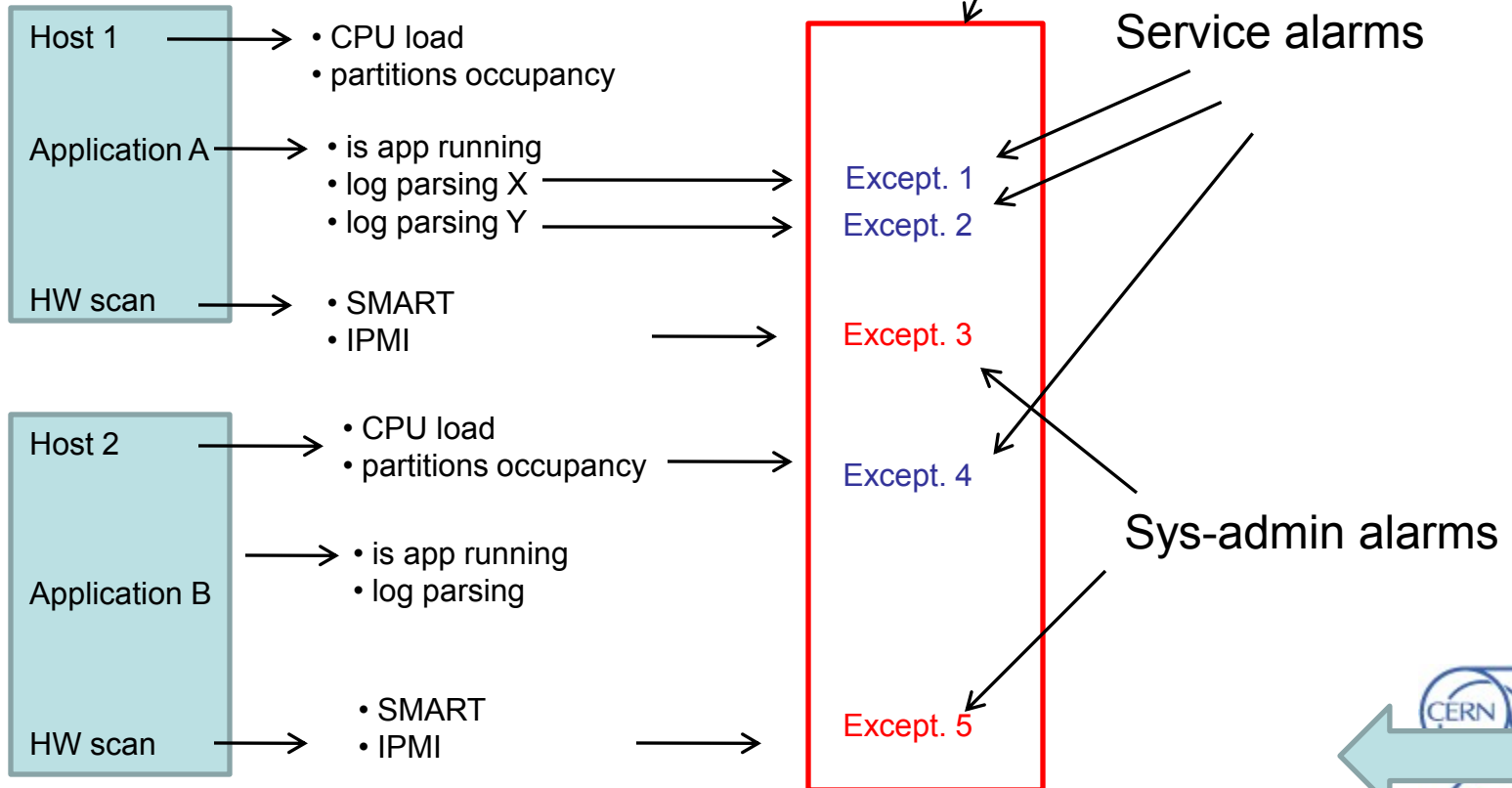
Current LAS view







Current LAS view



## Supported Lemon Sensors

<http://lemon.web.cern.ch/lemon/doc/sensors.shtml>

- [MSA](#) - checks the health of the Lemon sensor agent (built in)
- [Linux](#) - provides standard performance monitoring of the system
- [file](#) - provides various file-related utilities
- [parse-log](#) - provides log parsing metric
- [exception](#) - exception handling with support for correlation between metrics.
- remote (ping, http), SURE, DIP, PLC

## Misc Lemon Sensors

<https://svnweb.cern.ch/trac/lemon/browser/trunk/sensors/misc>

- Hwprobpe (memory error)
- Hwscan (scan hw and compare with CDB conf)
- ipmi (measured locally)
- Snmp-get (no for trap because async alarm not supported)

- an officially supported Lemon sensor coded in C++
- developed in collaboration between CERN and BARC
- implements the Lemon alarm protocol

## Main characteristic

- access agent local cache
- correlation engine allows to evaluate 1 or more metrics values to detect problem on a node
- supports reporting on behalf of other monitored entities
- allows corrective actions → actuators
- output is an exception

## Full documentation at:

- <http://lemon.web.cern.ch/lemon/doc/sensors/exception.shtml>

- Information:
  - Run as forked processes.
  - Are connected to the sensor via a pipe.
  - All information written to stdout or stderr by the actuator is caught and recorded in the agents log file.
- Running shell style actuators:
  - The system call used to run actuator doesn't provide shell style conveniences.
  - To use shell style syntax like \*, &&, | etc you must define you actuator like this:

```
"/system/monitoring/exception/_30010" = nlist(
    "name", "tmp_full",
    "descr", "tmp utilization exceeds limit",
    "active", true,
    "latestonly", false,
    "importance", 2,
    "correlation", "((9104:1 eq '/tmp') && (9104:5 > 80))",
    "actuator", nlist("execve", "/usr/local/sbin/clean-tmp-partition -o 75",
        "maxruns", 3,
        "timeout", 300,
        "window", 900,
        "active", true)
);
```

Actual value of correlation objects are accessible for actuator

```
/bin/sh -c \"/bin/echo '$act_value_01 $act_value_02'\\"
```

```
/bin/sh -c \"/bin/echo 'Died lemonmrd daemon $act_value_01 ' | /bin/mail -s 'Lemon RRD
Daemon problem' 0041xyz@mail2sms.cern.ch\\"
```

1. Use basic object for correlation

`<metric_id>:<field_position>`

2. Combine exceptions

`10004:1 > 600 && (10004:7 > 10 || (10004:8 > 150000 && 4109:3 eq 'i386') || (10004:8 > 600000 && 4109:3 regex '64') || 10007:2 > 50 || 10007:3 > 10 || 10007:4 > 0)`

4. Join the metrics

`(9200:1 == 9208:1)`

3. You can collect information on behalf, you can define exception on behalf

`[entity_name]:<metric_id>:<field_position>`

e.g. `(*:9501:5 != 200) && (*:9501:5 != 301)`

5. "correlation", "4109:2 ne 'symlink('/system/kernel/version)'"