Logging & Bookkeeping - Job monitoring, etc.

Daniel Kouřil, Michal Procházka

CESNET

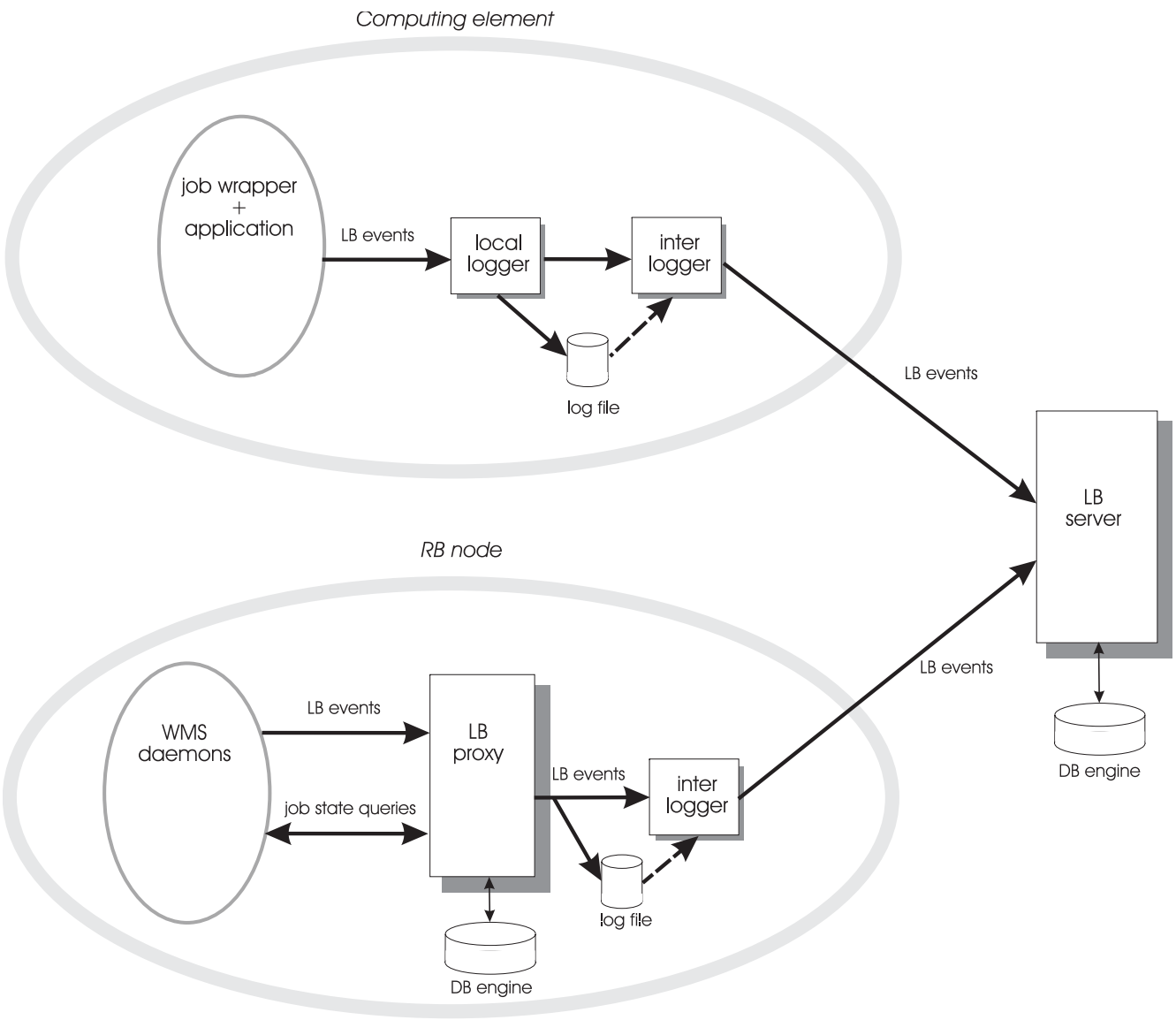
Logging and Bookkeeping

- **Monitoring system to track jobs in large grids**
  - in production for many years
  - capable of processing 1M jobs per day
    - tens of LB events per second
- **Two basic layers**
  - LB messaging infrastructure
  - LB server storing and processing job related data
- **Currently for jobs passing via WMS**
  - recently adapted to monitore PBS and Condor jobs, too
  - ongoing discussions with the CREAM developers
- **Query interface**
  - complex queries on jobs and their status
- **Notifications**
  - sent by LB server on changes

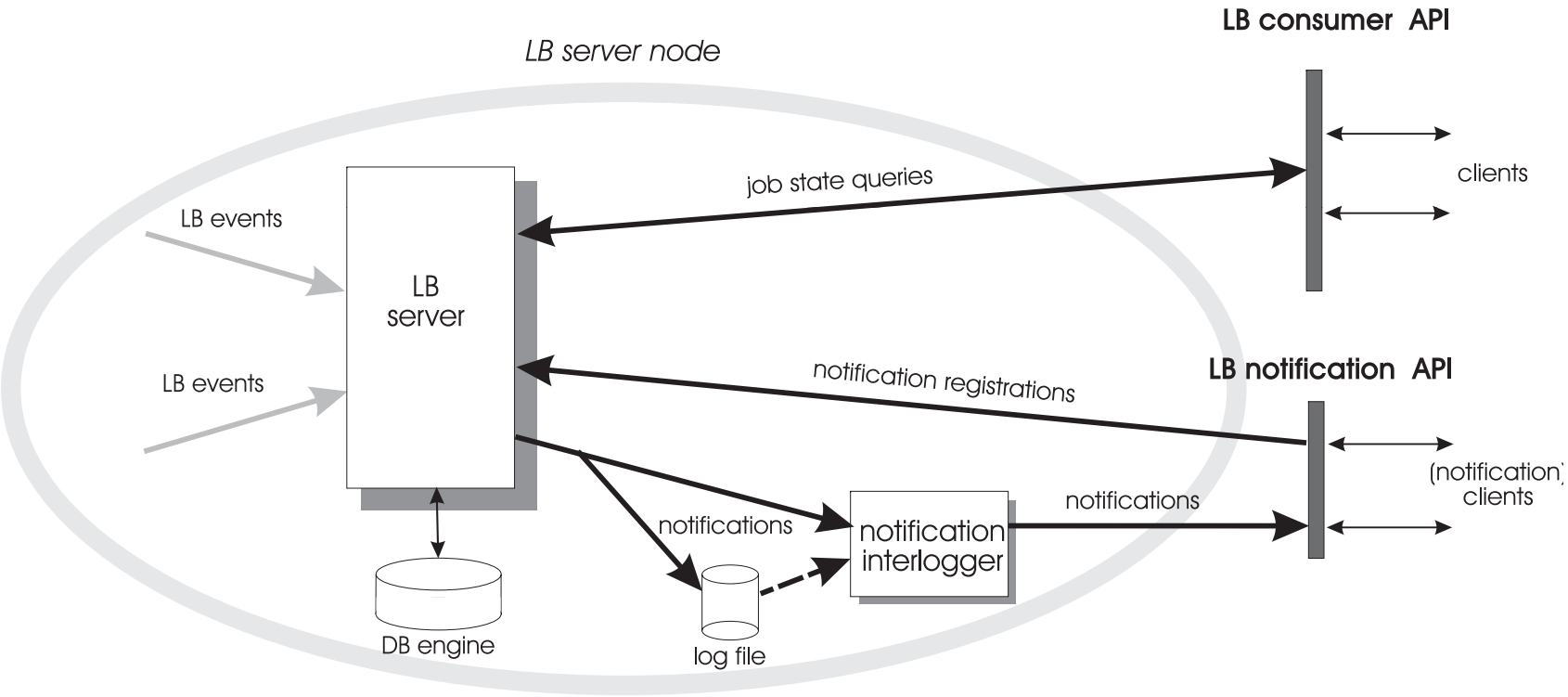
Gathering L&B data

- **LB collects *events* from individual Grid components**
  - information about a important point in the job‘s lifetime
    - transfer between components, start runnning, done, ...
  - events sent as messages to the LB server
  - own messaging infrastructure
    - secure (protection, authN) and reliable (fault-tolerancy)
    - notifications use this messaging infrastructure too
  - events are tied with job (using the jobid)
    - job registration
- **Push model**
  - events are sent by the components (mostly WMS) upon changes
  - instrumented components or reading log files
  - no useless polling
- **Trust model**

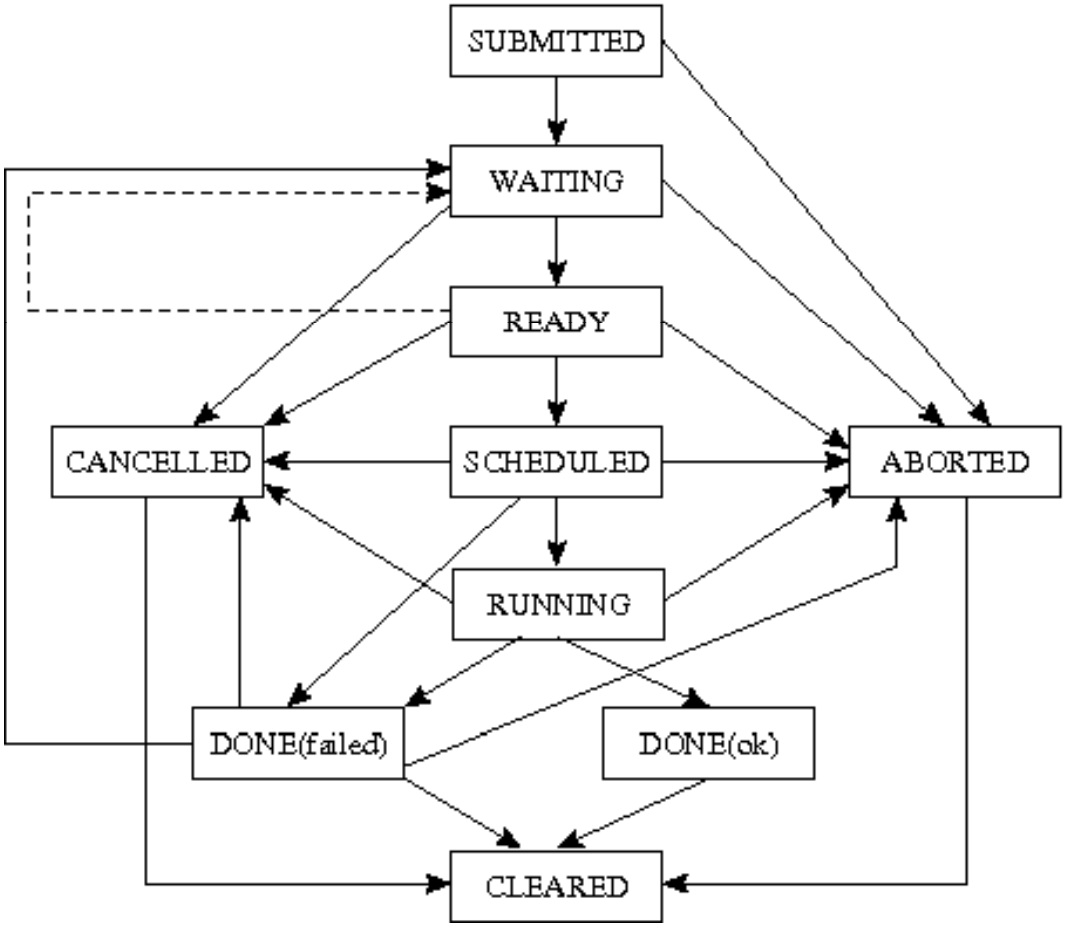
L&B Infrastructure



L&B Architecture



Job State Diagram



Accessing L&B data

- **Events are processed on the LB server**
  - LB defines Job state diagram
  - Each event could trigger a change in a job state (computed on the fly)
  - strict authZ
- **purging of completed jobs (Job Provenance)**
- **Users make queries against LB server**
  - glite-job-status, glite-job-logging-info
  - C/C++/WS/(http) interface
- **and/or subscribe for notifications**
  - send by LB server upon changes in job state
  - a simple client-side application needed

Examples

- **What are the jobs executed on CEs ce.foo.org or ce.bar.org since yesterday?**

Examples

- **What are the jobs executed on CEs ce.foo.org or ce.bar.org since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = NULL
destination=ce.foo.org destination=ce.bar.org
time > 2008-18-03 00:00:00; Running
```

Examples

- **What are the jobs executed on CEs ce.foo.org or ce.bar.org since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = NULL
destination=ce.foo.org destination=ce.bar.org
time > 2008-18-03 00:00:00; Running

jobId: https://scientific.civ.zcu.cz:10330/BG8AS0hPXsG603gnP9VepQ
jobId: https://scientific.civ.zcu.cz:10330/2fGIUBlK9Sctx9XFPCqtFX
```

Examples

- **What are the jobs executed on CEs ce.foo.org or ce.bar.org since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = NULL
destination=ce.foo.org destination=ce.bar.org
time > 2008-18-03 00:00:00; Running

jobId: https://scientific.civ.zcu.cz:10330/BG8AS0hPXsG603gnP9VepQ
jobId: https://scientific.civ.zcu.cz:10330/2fGIUBlK9Sctx9XFPCqtFX
```

- **What are the jobs submitted by me since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = "/DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril"
time > 2008-18-03 00:00:00; Submitted
```

Examples

- **What are the jobs executed on CEs ce.foo.org or ce.bar.org since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = NULL
destination=ce.foo.org destination=ce.bar.org
time > 2008-18-03 00:00:00; Running

jobId: https://scientific.civ.zcu.cz:10330/BG8AS0hPXsG603gnP9VepQ
jobId: https://scientific.civ.zcu.cz:10330/2fGIUBlK9Sctx9XFPCqtFX
```

- **What are the jobs submitted by me since yesterday?**

```
$ glite-lb-query_ext -m scientific.civ.zcu.cz
owner = "/DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril"
time > 2008-18-03 00:00:00; Submitted

jobId: https://scientific.civ.zcu.cz:10330/GsUwK82soKKw-sKK3sskCw
```

Events output

```
Event: RegJob
- arrived                 =    Sun Jun  3 13:50:32 2007 CEST
- host                    =    skurut4.cesnet.cz
- ns                      =    skurut67-6.cesnet.cz:7772
- nsubjobs                =    0
- seed                    =    uLU0BArrdV98O41PLThJ5Q
- source                  =    UserInterface
- timestamp               =    Sun Jun  3 13:50:31 2007 CEST
- user                    =    /DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril
        ---
Event: Transfer
- arrived                 =    Sun Jun  3 13:52:33 2007 CEST
- dest_host               =    skurut67-6.cesnet.cz
- dest_instance           =    skurut67-6.cesnet.cz:7772
- destination             =    NetworkServer
- host                    =    skurut4.cesnet.cz
- result                  =    START
- source                  =    UserInterface
- timestamp               =    Sun Jun  3 13:52:32 2007 CEST
- user                    =    /DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril
        ---
Event: Transfer
- arrived                 =    Sun Jun  3 13:54:48 2007 CEST
- dest_host               =    skurut67-6.cesnet.cz
- dest_instance           =    skurut67-6.cesnet.cz:7772
- destination             =    NetworkServer
- host                    =    skurut4.cesnet.cz
- result                  =    OK
- source                  =    UserInterface
- timestamp               =    Sun Jun  3 13:54:47 2007 CEST
- user                    =    /DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril
        ---
Event: Accepted
- arrived                 =    Sun Jun  3 13:54:39 2007 CEST
- from                    =    UserInterface
- from_host               =    skurut67-6.cesnet.cz
- host                    =    skurut67-6.cesnet.cz
- source                  =    NetworkServer
- src_instance            =    7772
- timestamp               =    Sun Jun  3 13:54:39 2007 CEST
- user                    =    /DC=cz/DC=cesnet-ca/O=Masaryk University/CN=Daniel Kouril

.......
```

Tracing jobs using LB

- **primarily aimed at honest users**
- **Issues:**
  - users can select their LB
    - multiple LB must be checked
    - not all can allow „super-user“ access
  - only jobs passing via WMS are logged to LB
  - users can distort the LB records
    - only add events, not change logged ones
  - strict access control for accessing data
- **several remedies possible**
  - Operational policy/configurations, ...

Syslog over LB

- **Collecting syslog data from distributed resources**
  - Czech NGI
  - Kerberos-based infrastructure
- **Current syslog doesn‘t protect messages**
  - even syslog-ng requires ssl tunnels
- **LB messaging layer can be utilized**
  - already used to distribute CRLs
  - detailed knowledge of internals
- **„keep it simple“ approach**
  - usual networking API
    - accept/connect, read/write, close
  - on top of GSS-API
    - currently using GSS/GSI from Globus
    - any GSS API implementation can be plugged in
  - strict timing out
    - any network operation can take indefinite time

Collecting Syslog data

- **LB loggers installed on each machine**
  - messages can be logged via cluster head-nodes, too
- **Added a client-side daemon**
  - takes data from local syslogd
  - puts it into the LB infrastructure
- **An „LB“ server for central syslog server**
  - reads data from the loggers
  - pass it to standard syslog server
    - adds timestamps and clients‘ ids
- **No changes to syslog needed**
- **Transparent security and reliability gained**