Condor-G: Condor and Grid Computing

Condor Project
Computer Sciences Department
University of Wisconsin-Madison

Condor-G

- Condor for the grid
  - Same job management capabilities as a local Condor pool
  - Use other scheduling systems' resources

Job Management Interface

- Local, persistent queue
- Job policy expressions
- Match-making
- Job activity logs

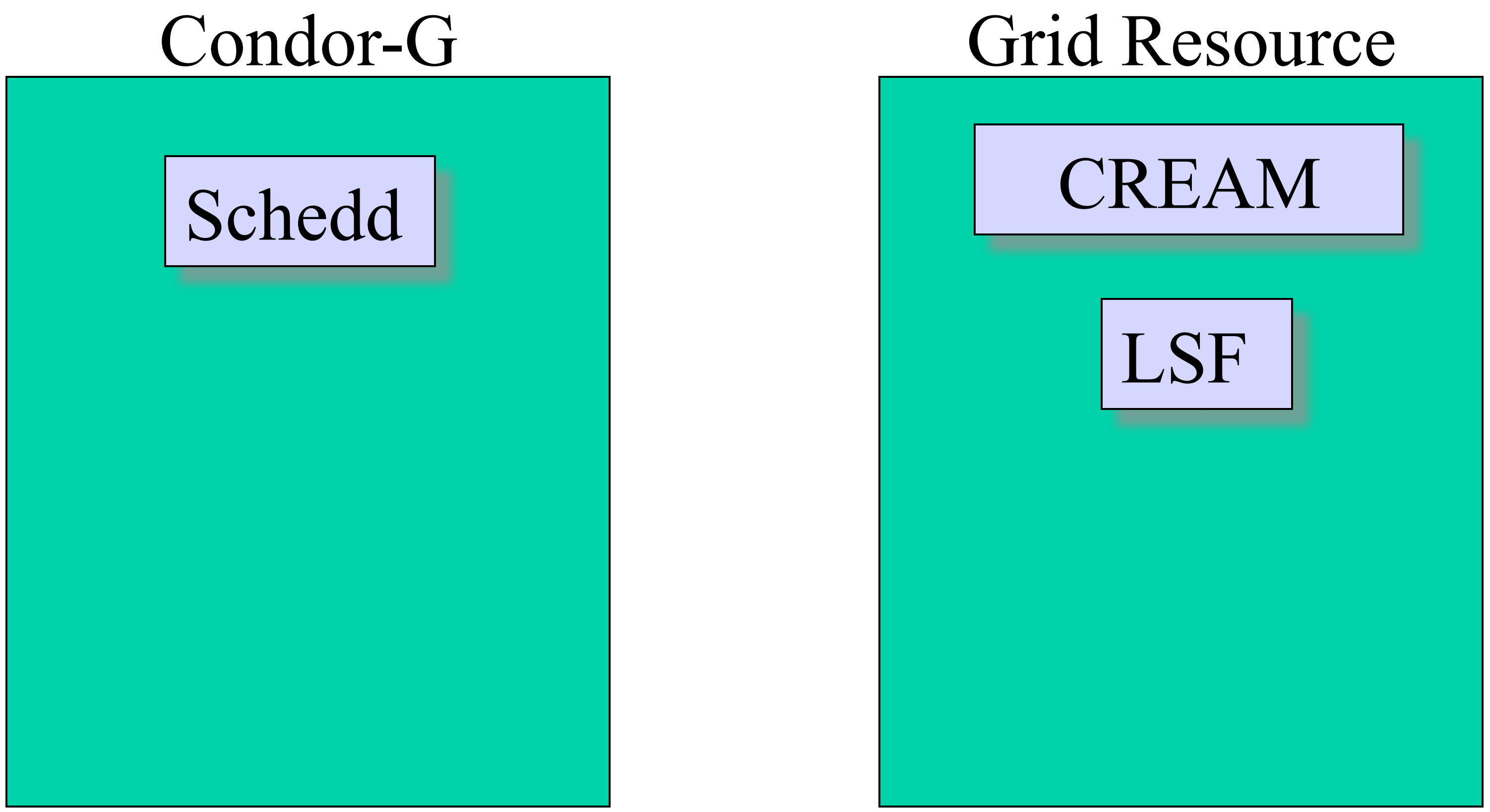
Gridmanager Daemon

- Runs under the schedd
- Similar to the shadow
- Handles all management of grid jobs
- Single instance manages all grid jobs for a user

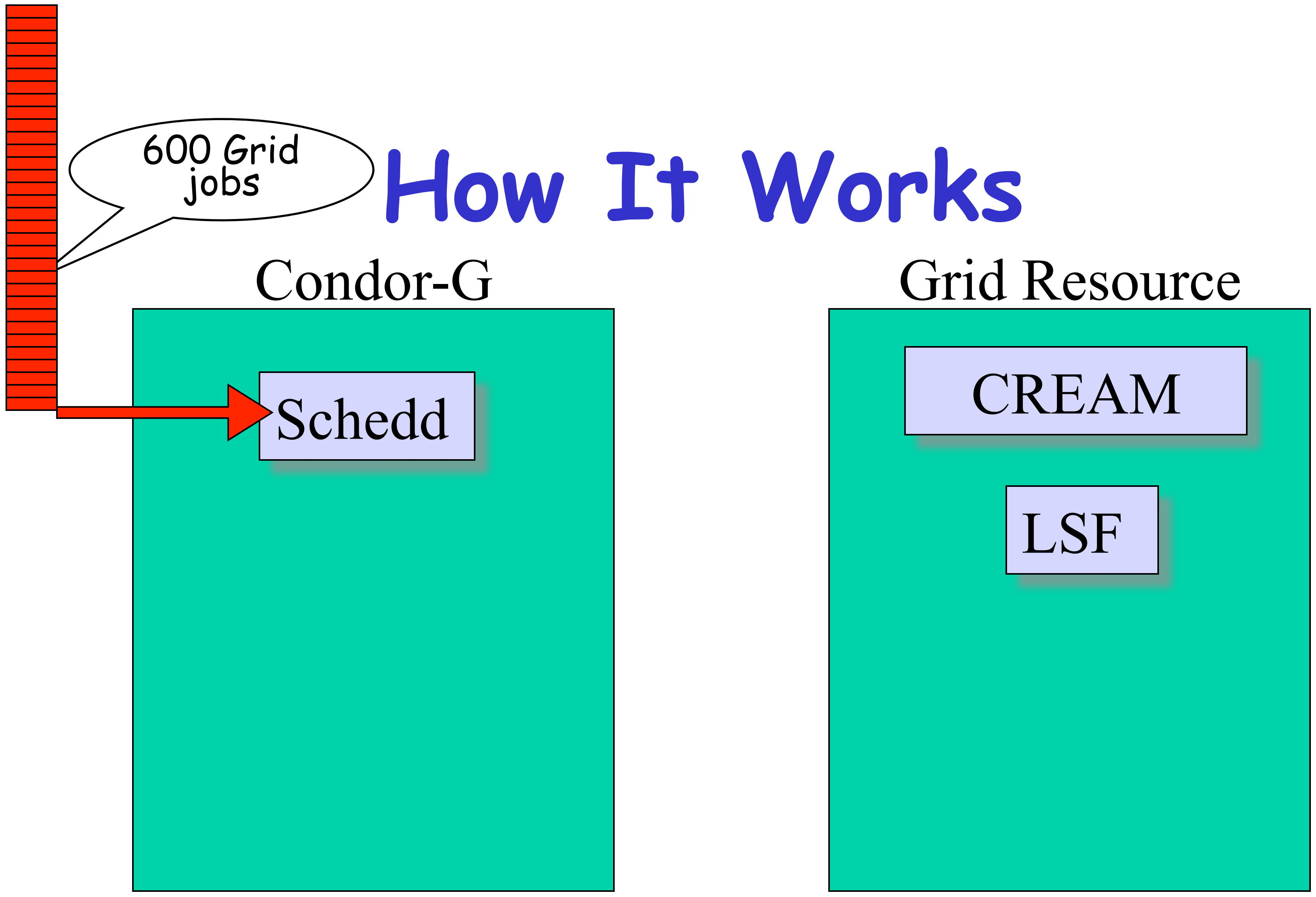
Grid ASCII Helper Protocol (GAHP)

- Runs under gridmanager
- Encapsulates grid client libraries in separate process
- Simple ASCII protocol
- Easy to use client libraries when they can't be linked directly with gridmanager

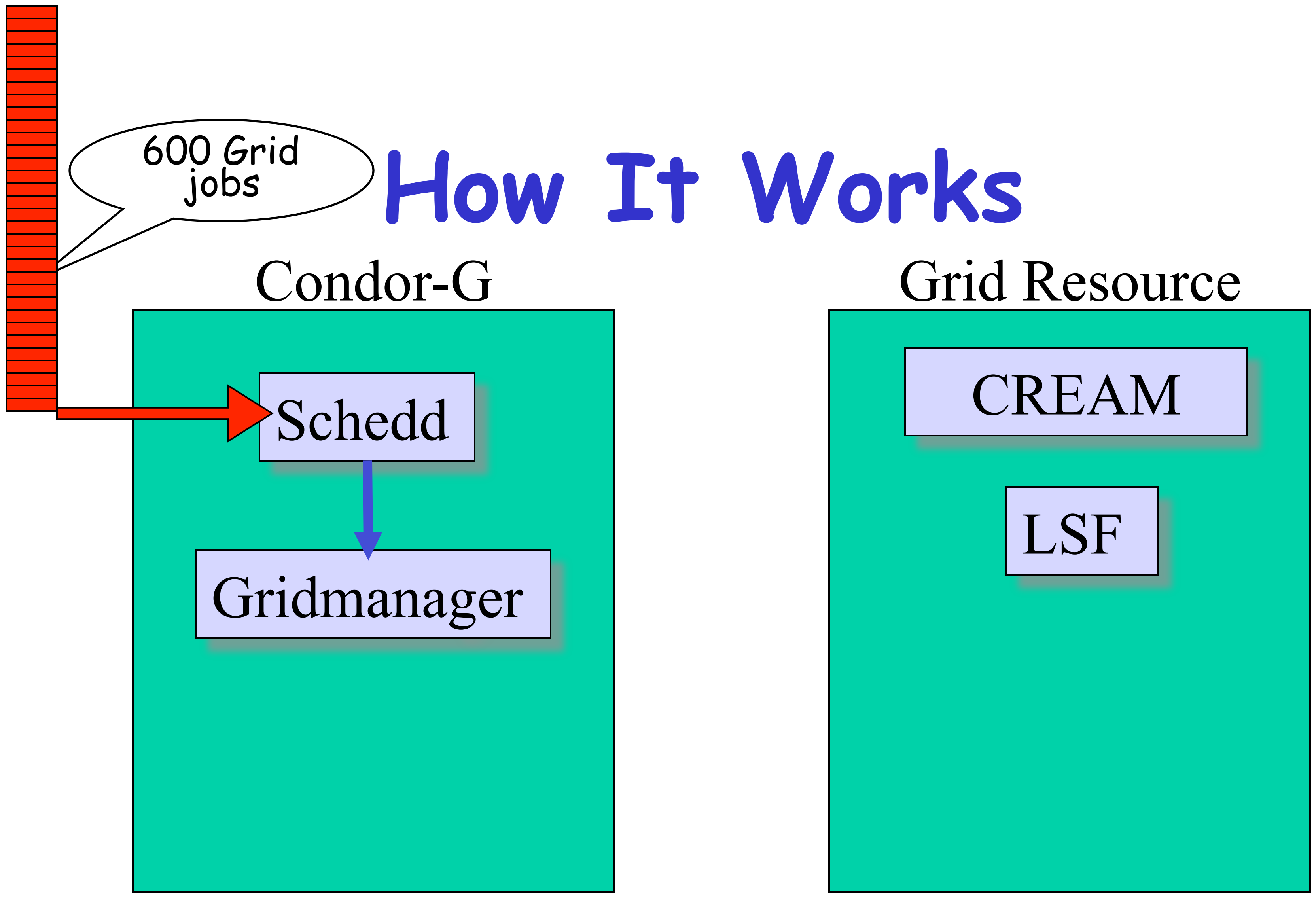
How It Works



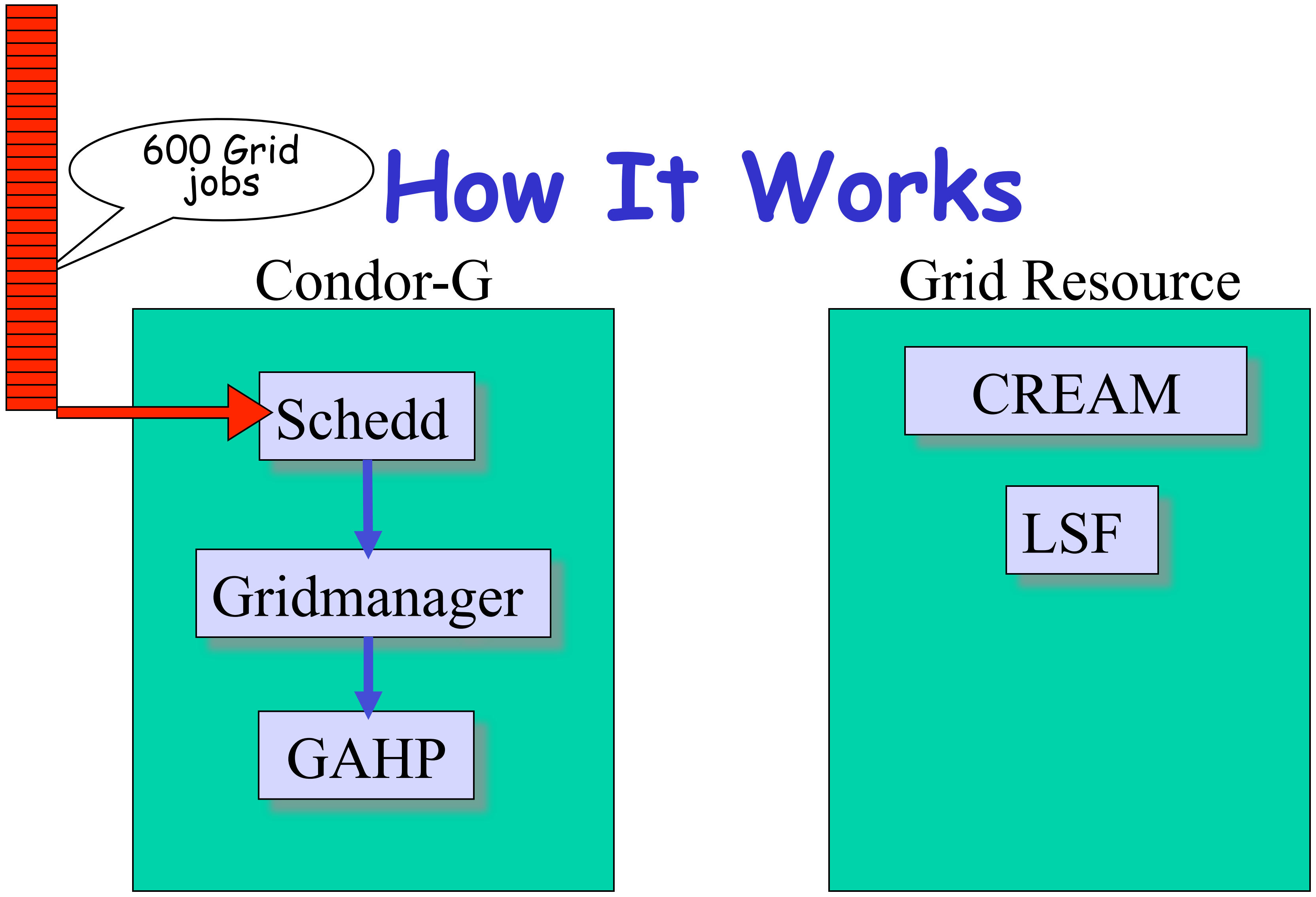
How It Works



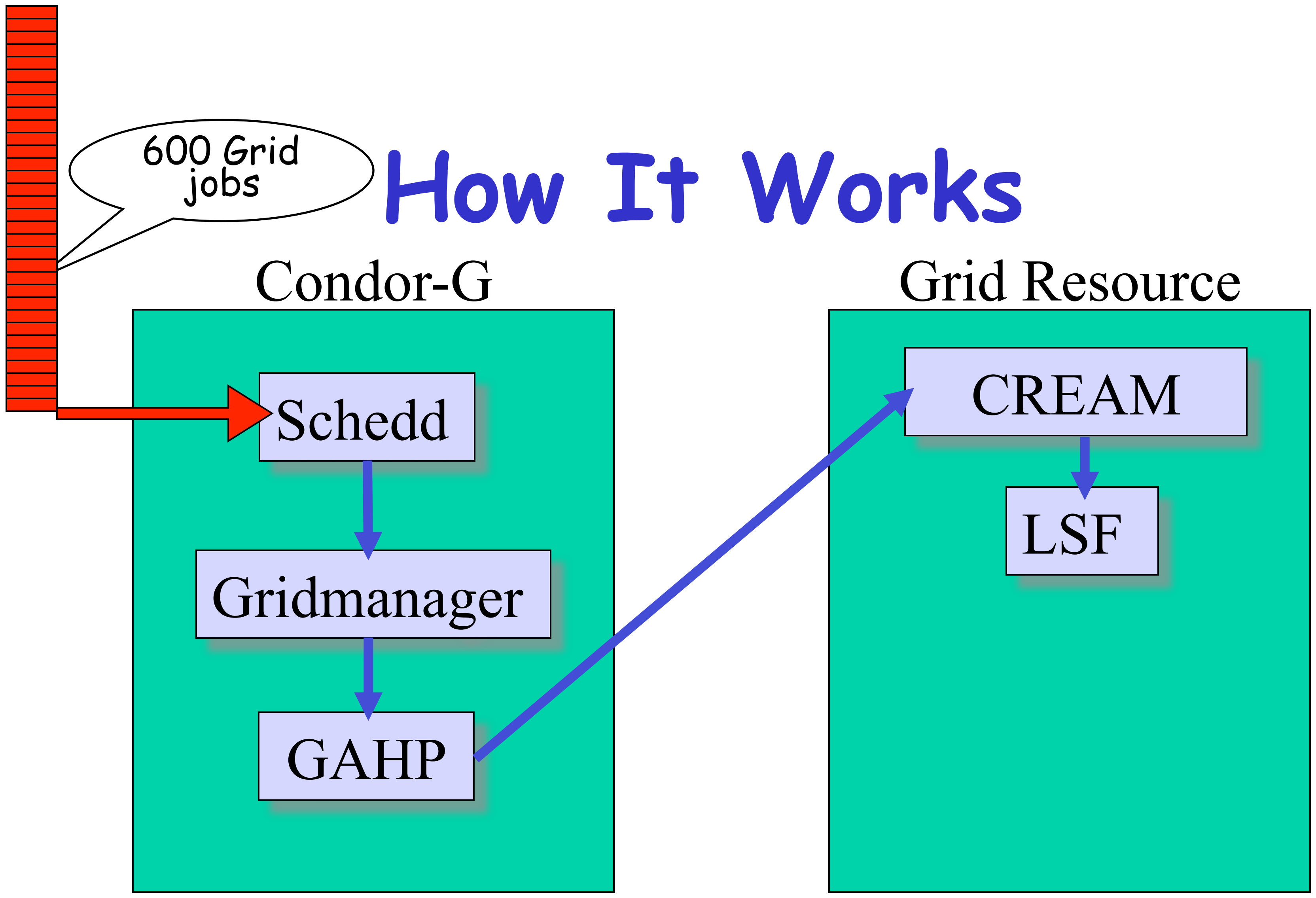
How It Works


600 Grid jobs
Condor-G
Schedd
Gridmanager
Grid Resource
CREAM
LSF

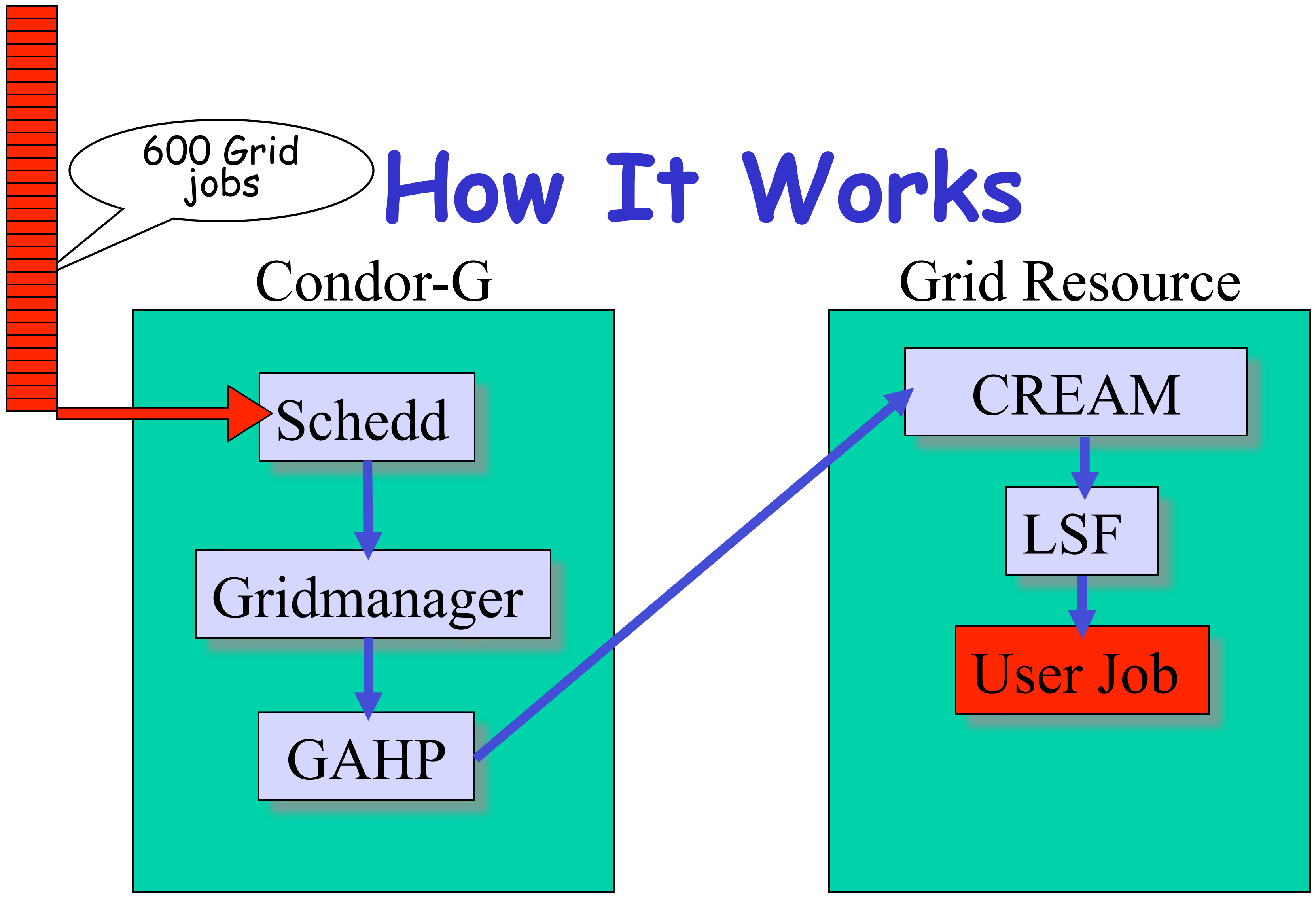
How It Works


600 Grid jobs
Condor-G
Schedd
Gridmanager
GAHP
Grid Resource
CREAM
LSF

How It Works


600 Grid jobs
Condor-G
Schedd
Gridmanager
GAHP
Grid Resource
CREAM
LSF

How It Works


600 Grid jobs
Condor-G
Schedd
Gridmanager
GAHP
Grid Resource
CREAM
LSF
User Job

"Grid" Universe

- All handled in your submit file
- Supports a number of "back end" types:
  - Globus: GT2, GT4, GT5
  - CREAM
  - NorduGrid
  - UNICORE
  - Condor
  - PBS
  - LSF
  - EC2
  - Deltacloua

Globus GRAM2

- Used for a Globus GT2 back-end
  - "Condor-G"
- Format:

```
Grid_Resource = gt2 Head-Node
Globus_rsl = <RSL-String>
```

- Example:

```
Universe = grid
Grid_Resource = gt2 beak.cs.wisc.edu/jobmanager
Globus_rsl = (queue=long)(project=atom-smasher)
```

Globus GRAM4

- Used for a Globus GRAM4 back-end
- Format:

```
Grid_Resource = gt4 <Head-Node> <Scheduler-Type>
Globus_XML = <XML-String>
```

- Example:

```
Universe = grid
Grid_Resource = gt4 beak.cs.wisc.edu Condor
Globus_xml = <queue>long</queue><project>atom-
   smasher</project>
```

Globus GRAM5

- Used for a Globus GRAM5 back-end
  - More scalable version of GRAM2
- Format:

```
Grid_Resource = gt5 Head-Node
Globus_rsl = <RSL-String>
```

- Example:

```
Universe = grid
Grid_Resource = gt5 beak.cs.wisc.edu/jobmanager
Globus_rsl = (queue=long)(project=atom-smasher)
```

CREAM

- Used for a CREAM back-end
- Format:

```
Grid_Resource = cream <CREAM service>
Cream_Attributes = <JDL attributes>
```

- Example:

```
Universe = grid
Grid_Resource = cream foo.edu/cream-pbs-
  normal_queue
Cream_Attributes = CpuNumber=5
```

Condor

- Used for a Condor back-end
  - "Condor-C"
- Format:

```
Grid_Resource = condor <Schedd-Name> <Collector-Name>
Remote_<param> = <value>
```

  - "Remote_" part is stripped off
- Example:

```
Universe = grid
Grid_Resource = condor beak condor.cs.wisc.edu
Remote_Universe = standard
```

NorduGrid ARC

- Used for a NorduGrid back-end

```
Grid_Resource = nordugrid <Host-Name>
```

- Example:

```
Universe = grid
Grid_Resource = nordugrid ngrid.cs.wisc.edu
```

UNICORE

- Used for a UNICORE back-end
- Format:

```
Grid_Resource = unicore <USite> <VSite>
```

- Example:

```
Universe = grid
Grid_Resource = unicore uhost.cs.wisc.edu vhost
```

PBS

- Used for a PBS back-end
- Format:

```
Grid_Resource = pbs
```

- Example:

```
Universe = grid
Grid_Resource = pbs
```

LSF

- Used for a LSF back-end
- Format:

```
Grid_Resource = lsf
```

- Example:

```
Universe = grid
Grid_Resource = lsf
```

Credential Management

- Condor will do The Right Thing™ with your X509 certificate and proxy
- Override default proxy:
  - `X509UserProxy = /home/einstein/other/proxy`
- Proxy may expire before jobs finish executing
  - Condor can use MyProxy to renew your proxy
  - When a new proxy is available, Condor will forward the renewed proxy to the job

Configuration Files

- One main config file and multiple additional files
- Can have multiple Condor installations on a machine, each with a different set of config files

Finding the Config Files

- Condor looks for the main config file in
  - $CONDOR_CONFIG
  - /etc/condor/condor_config
  - /usr/local/etc/condor/condor_config
  - ~condor/condor_config
  - $(GLOBUS_LOCATION)/etc/condor_config

Finding the Config Files

- Main config file can specify additional files
- `LOCAL_CONFIG_FILE=foo,bar`
  - List of additional files
- `LOCAL_CONFIG_FILE=wget http://foo.edu/config|`
  - Execute program to produce config setting
- `LOCAL_CONFIG_DIR =/etc/configs/`
  - Directory containing config files

Querying the Configuration

- `condor_config_val -config`
  - Prints list of config files being used
  - May not be same config used by daemons
- `condor_config_val LOG`
  - Prints value of LOG parameter
- `condor_config_val -v LOG`
  - Prints value of LOG parameter
  - Also prints where it's set

Log Files

- User log
  - Records significant events in a job's life
- Event log
  - Like user log, but for all jobs
- History file
  - Copy of job ad when it leaves the queue
- Gridmanager log
  - Gridmanager's record of activity

User Log

- Can find with:

```
condor_q -format '%s\n' UserLog 17.0
```

- Set with "log" in the submit file
- Contains the life history of the job
- Often contains details on problems
- Never rotates

Event Log

- Like the user log, but used for all jobs
- Set via **`EVENT_LOG`** in the config file

History File

- Record of all jobs that leave the queue
- Snapshot of job state
- Can view via **`condor_history`**
- Can use **`PER_JOB_HISTORY_DIR`** to save non-rotating history

Gridmanager Log

- Find via **`GRIDMANAGER_LOG`** param
  - Contains job owner's username
- When debugging, can enable more verbose logging
  - **`GRIDMANAGER_DEBUG = D_FULLDEBUG`**
  - **`MAX_GRIDMANAGER_LOG = 50000000`**

HELD Status

- Jobs will be held when Condor-G needs help with an error
- On release, Condor-G will retry
- The reason for the hold will be saved in the job ad and user log

Hold Reason

- condor_q -held

```
 161.0   jfrey              2/13 13:58 CREAM_Delegate
   Error: Received NULL fault;
```

- cat job.log

```
012 (161.000.000) 02/13 13:58:38 Job was held.
        CREAM_Delegate Error: Received NULL fault; the
   error is due to another cause…
```

- condor_q -format '%s\n' HoldReason

```
CREAM_Delegate Error: Received NULL fault; the error is
   due to another cause…
```

Common Errors

- Authentication
  - Hold reason may be misleading
  - User may not be authorized by CE
  - Condor-G may not have access to all Certificate Authority files
  - User's proxy may have expired

Common Errors

- CE no longer knows about job
  - CE admin may forcibly remove job files
  - Condor-G is obsessive about not leaving orphaned jobs
  - May need to take extra steps to convince Condor-G that remote job is gone

Nonessential Jobs

- Jobs can be marked nonessential in the submit file
  - `+nonessential = true`
- This makes Condor-G more willing to leave orphaned jobs and files on the CE
- Use with caution

More Detail on Errors

- More details on errors can be found in the gridmanager log
- You'll probably want to increase the debug level and log file size
  - `GRIDMANAGER_DEBUG = D_FULLDEBUG`
  - `MAX_GRIDMANAGER_LOG = 5000000`

Machines Down

- If a remote server is down, Condor-G will wait for it to come back up
- The time it went down is kept in the job ad
  - `GridResourceUnavailableTime = 1297628439`
- And in the user log

```
026 (163.001.000) 02/13 14:20:39 Detected Down
   Grid Resource
     GridResource: gt2 chopin.cs.wisc.edu/
   jobmanager-fork
```

Throttles and Timeouts

- Limits that prevent Condor-G or CEs from being overwhelmed by large numbers of jobs
- Defaults are fairly conservative

Throttles and Timeouts

- `GRIDMANAGER_MAX_SUBMITTED_JOBS_PER_RESOURCE = 1000`
  - You can increase to 10,000 or more
- `GRIDMANAGER_MAX_JOBMANAGERS_PER_RESOURCE = 10`
  - GRAM2 only
  - Default is conservative
  - Can increase to ~100 if this is the only client

Throttles and Timeouts

- `GRIDMANAGER_MAX_PENDING_REQUESTS = 50`
  - Number of commands sent to a GAHP in parallel
  - Can increase to a couple hundred
- `GRIDMANAGER_GAHP_CALL_TIMEOUT = 300`
  - Time after which a GAHP command is considered failed
  - May need to lengthen if pending requests is increased

Network Connectivity

- Outbound connections only for most job types
- GRAM requires incoming connections
  - Need 2 open ports per <user, X509 DN> pair