OSG User's Meeting 2007

glideinWMS and gLExec

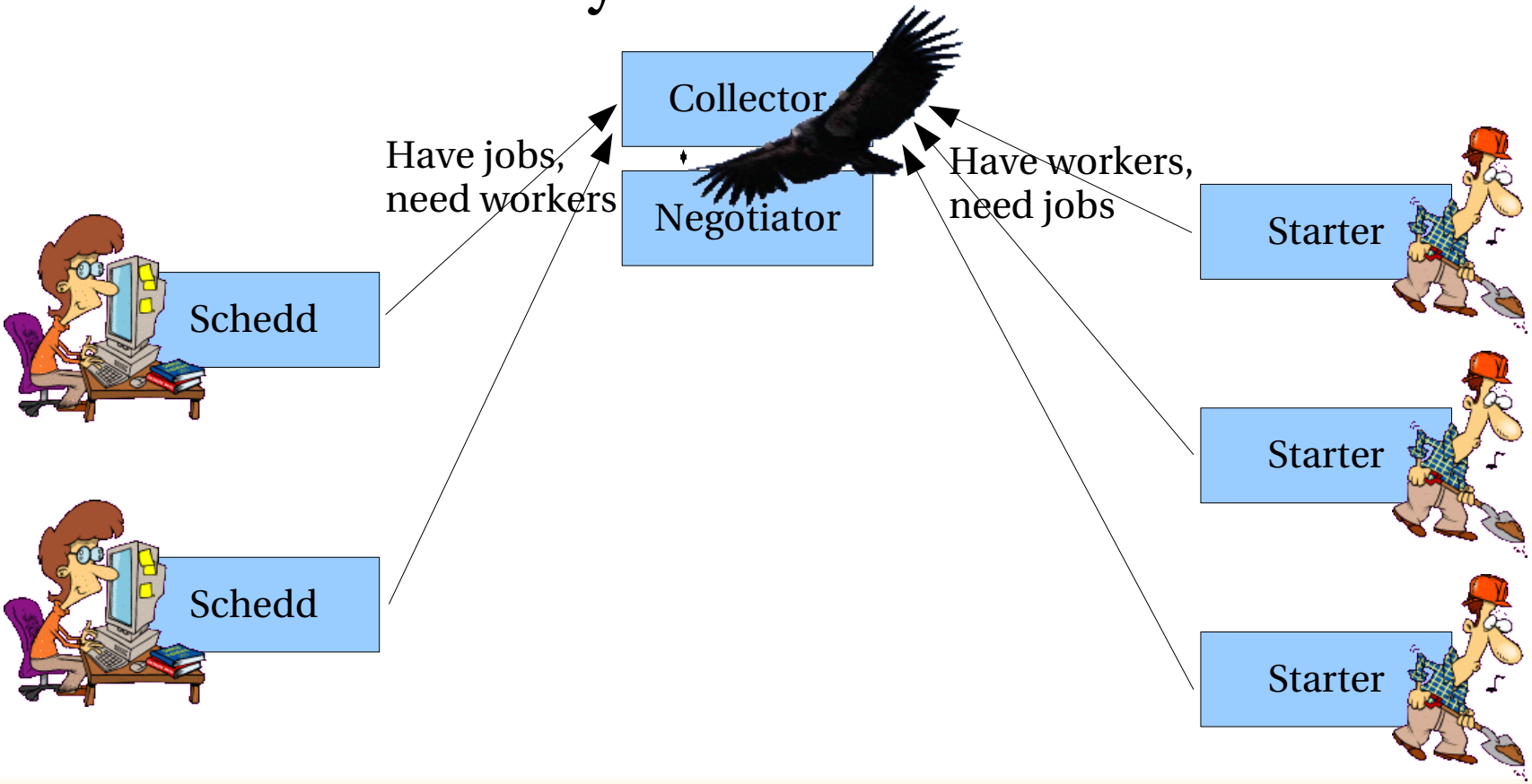
by Igor Sfiligoi (FNAL)

What is glideinWMS?

- A Condor glidein-based Workload Management System
- Developed by CMS for CMS, but generic enough to be used by other groups, too
- Available at: http://home.fnal.gov/~sfiligoi/glideinWMS/

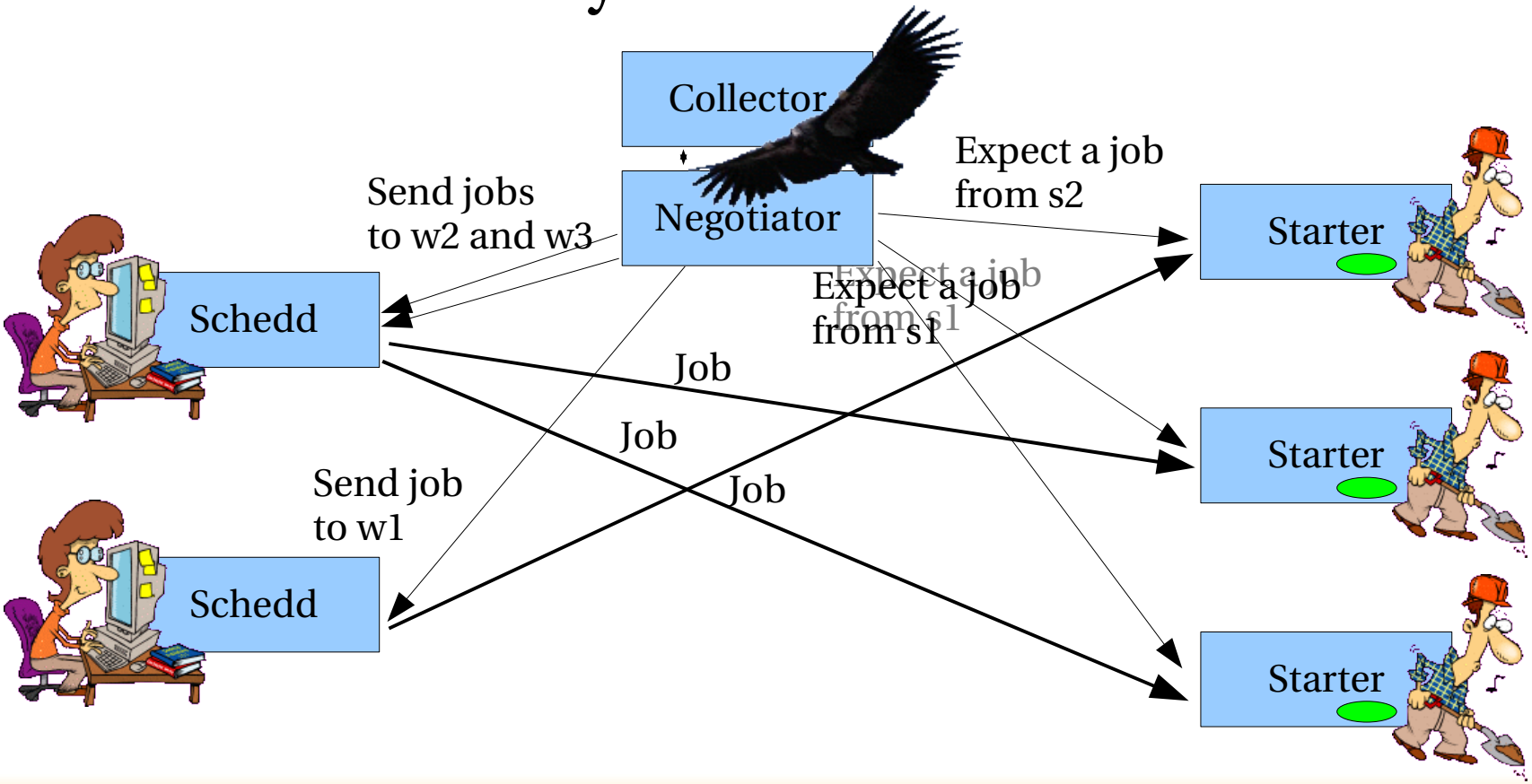
What is a Condor? (1)

- A widely used batch system
- Based on a fully distributed architecture



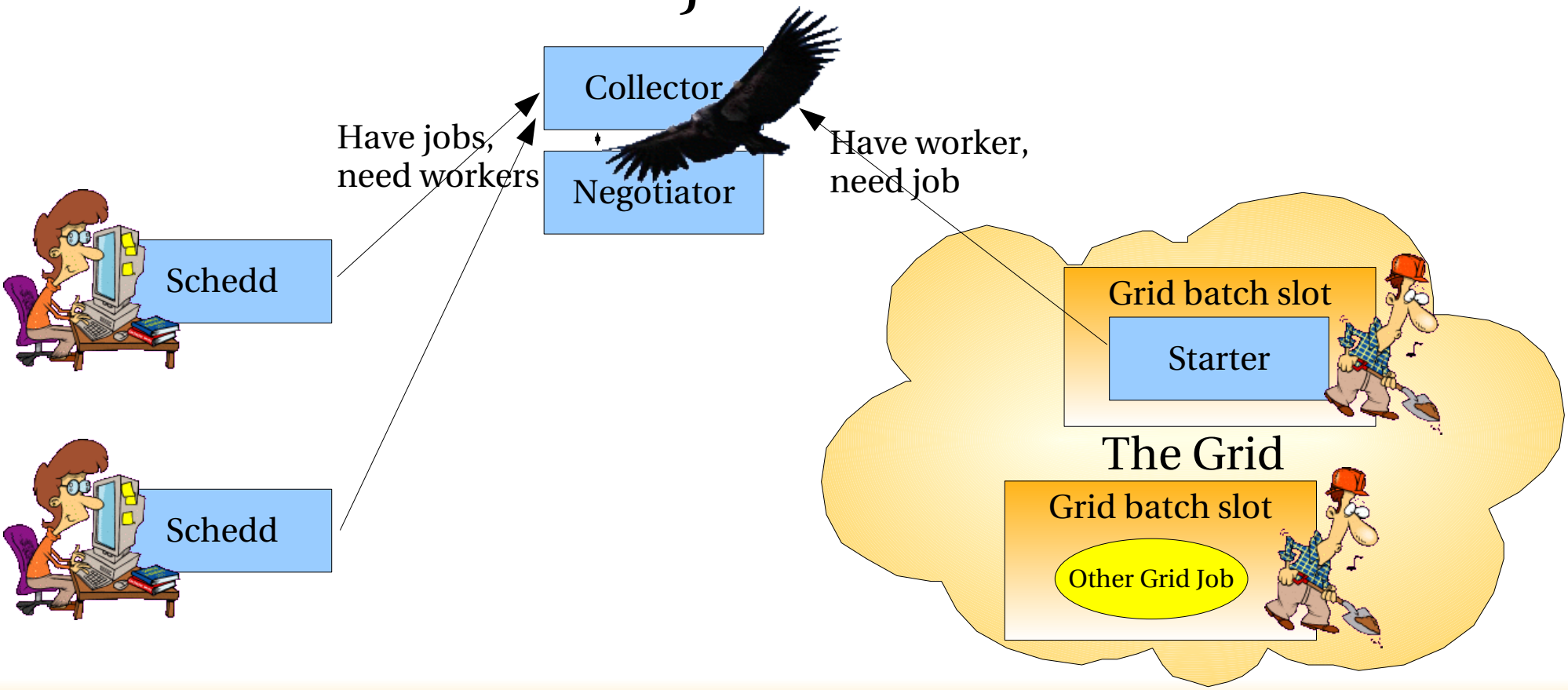
What is a Condor? (2)

- A widely used batch system
- Based on a fully distributed architecture



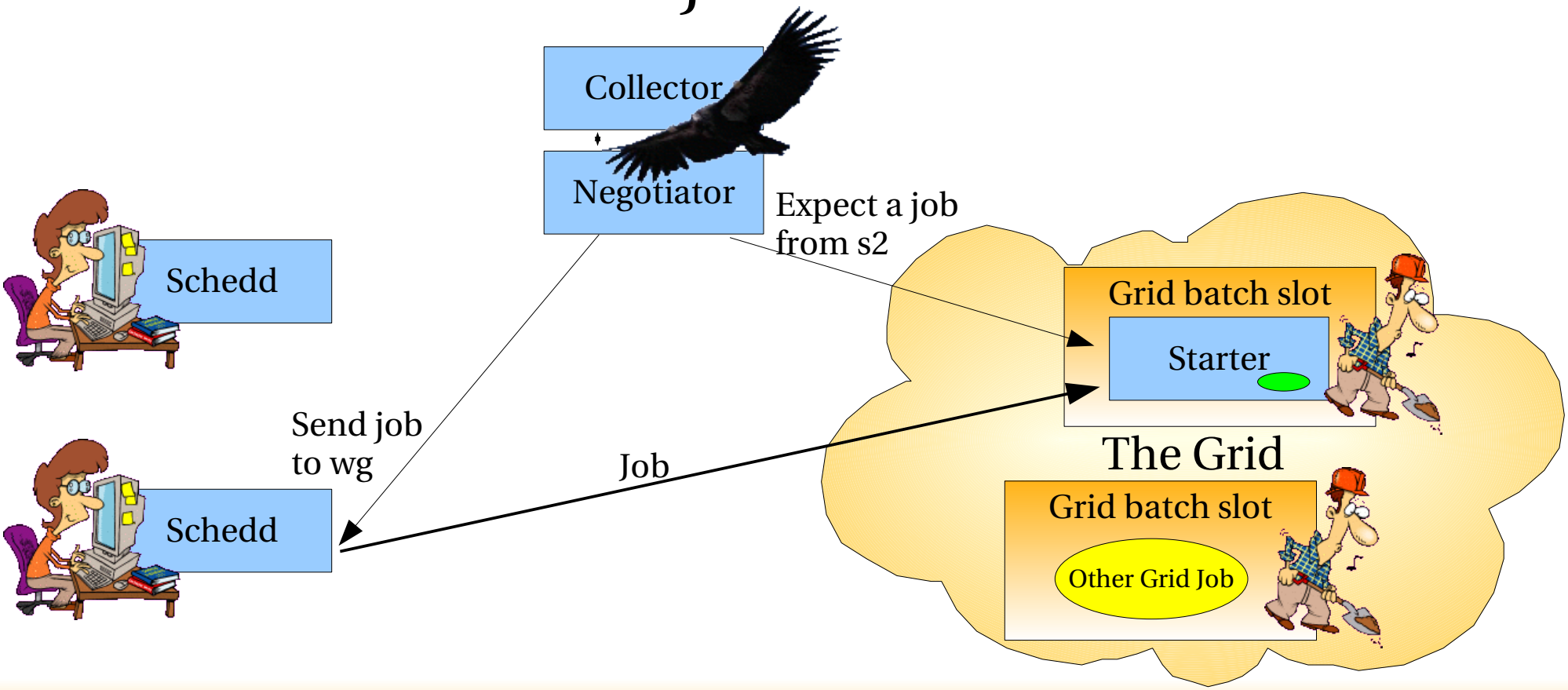
What is a glidein? (1)

- Just a regular starter
- Submitted as a Grid job



What is a glidein? (2)

- Just a regular starter
- Submitted as a Grid job



What else can a glidein do?

- Make sanity checks before fetching any job
- Discover and publish batch slot characteristics, like:
  - Site name
  - OS version
  - Availability of certain software
- Importing VO specific software
- Prepare the environment for the user jobs
  - Possibly putting the VO software in the path

Why using glideins? (1)

- If you are already using Condor
  - An easy way to extend the pool
    - Or to create one from scratch
  - Can hide all the grid stuff from your job
    - Can even run standard universe jobs on the Grid!
- If you want to use the Grid (even if you are not a Condor fan)
  - Protects the jobs from many obvious errors
    - A dead glidein will not pull a user job
  - Simplifies resource matching
    - A glidein can detect what is available on the worker and user jobs get sent only to complying workers

Why using glideins? (2)

- Get all the advantages of a local batch system
  - Locally set priorities between different users
    - Including group quotas
    - Or even priorities between jobs of the same user
  - Reliable, real time monitoring
  - Reliable file transfer
    - Full file encryption supported, too
- While still running on the Grid!

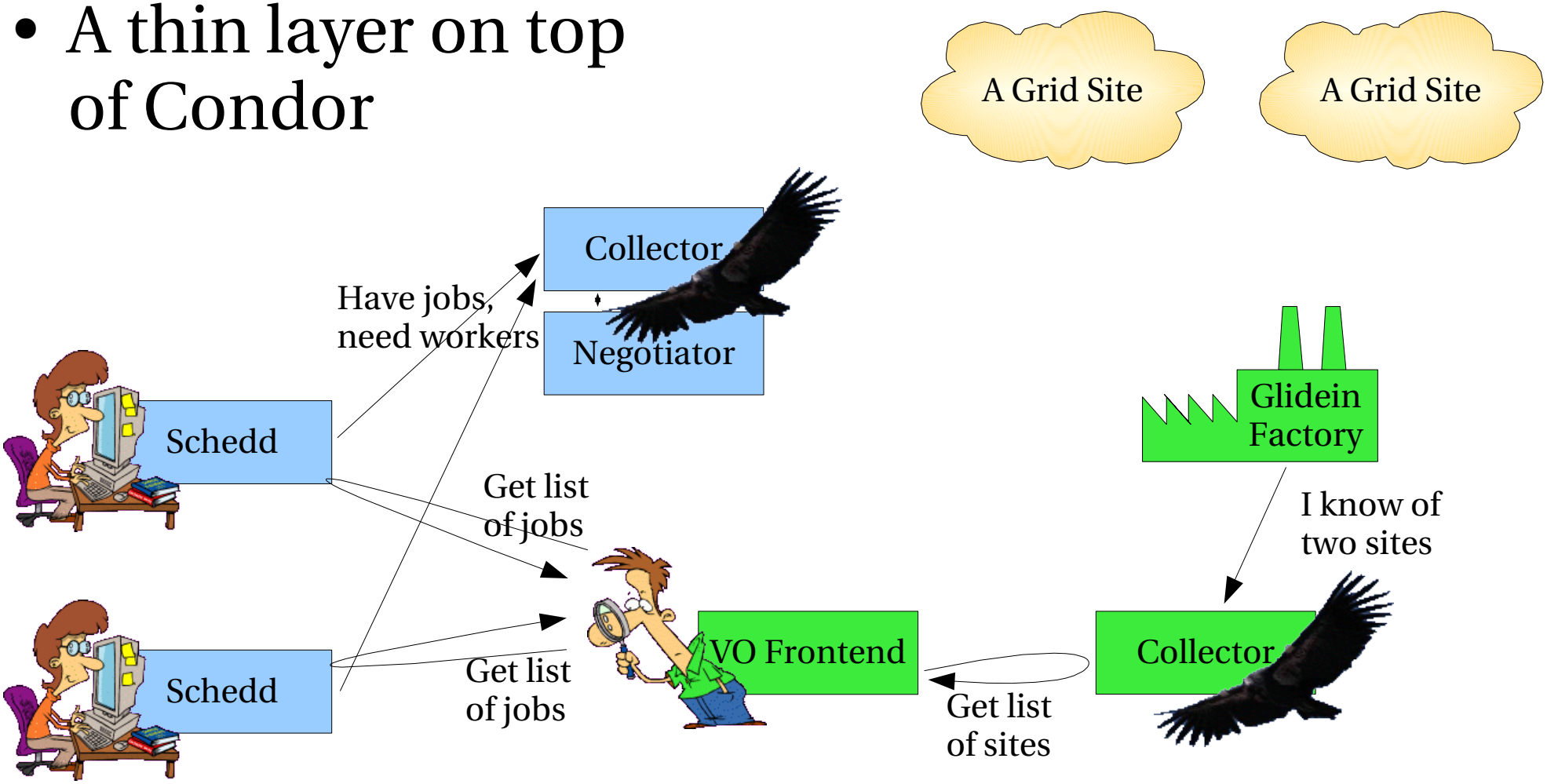
Any weak points will be presented at the end

How do I submit a glidein?

- Condor provides condor_glidein
  - Simple command line tool
  - **Useful when you have just a few jobs**
    - Will submit a single glidein x invocation
- Install a glideinWMS instance
  - Needs more resources and some initial effort to set it up
  - **Setup once, glideins will be lauched as needed**
    - Will look for jobs that need resources
    - Submit glideins as needed to sites that seem to match at least an idle job

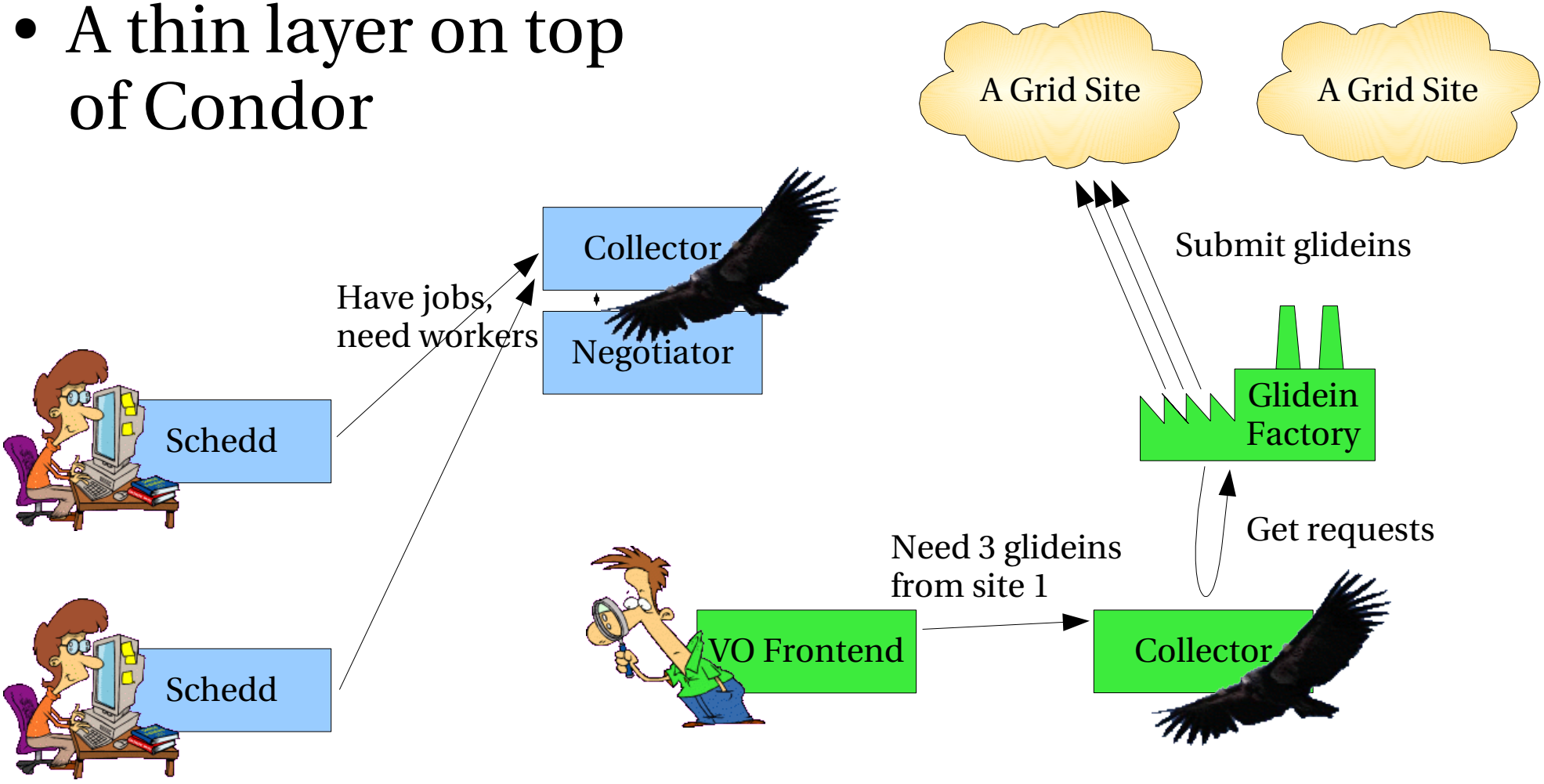
glideinWMS overview (1)

- A thin layer on top of Condor



glideinWMS overview (2)

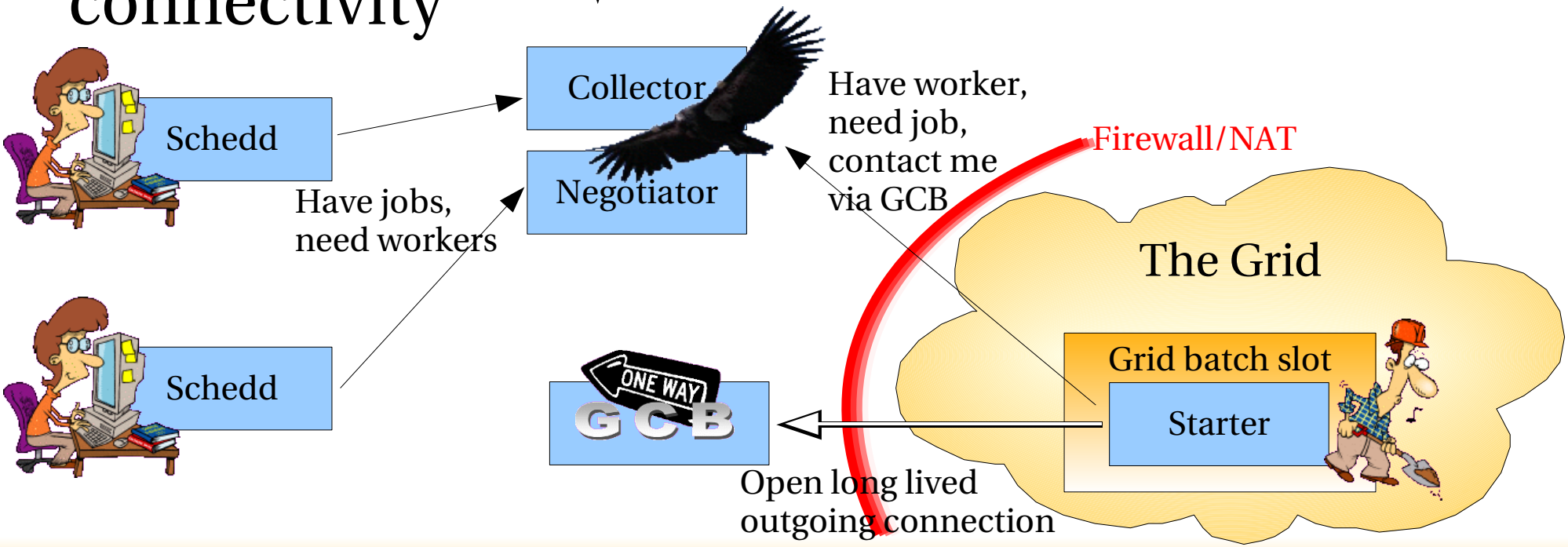
- A thin layer on top of Condor



More details at http://cd-docdb.fnal.gov/cgi-bin/ShowDocument?docid=2048

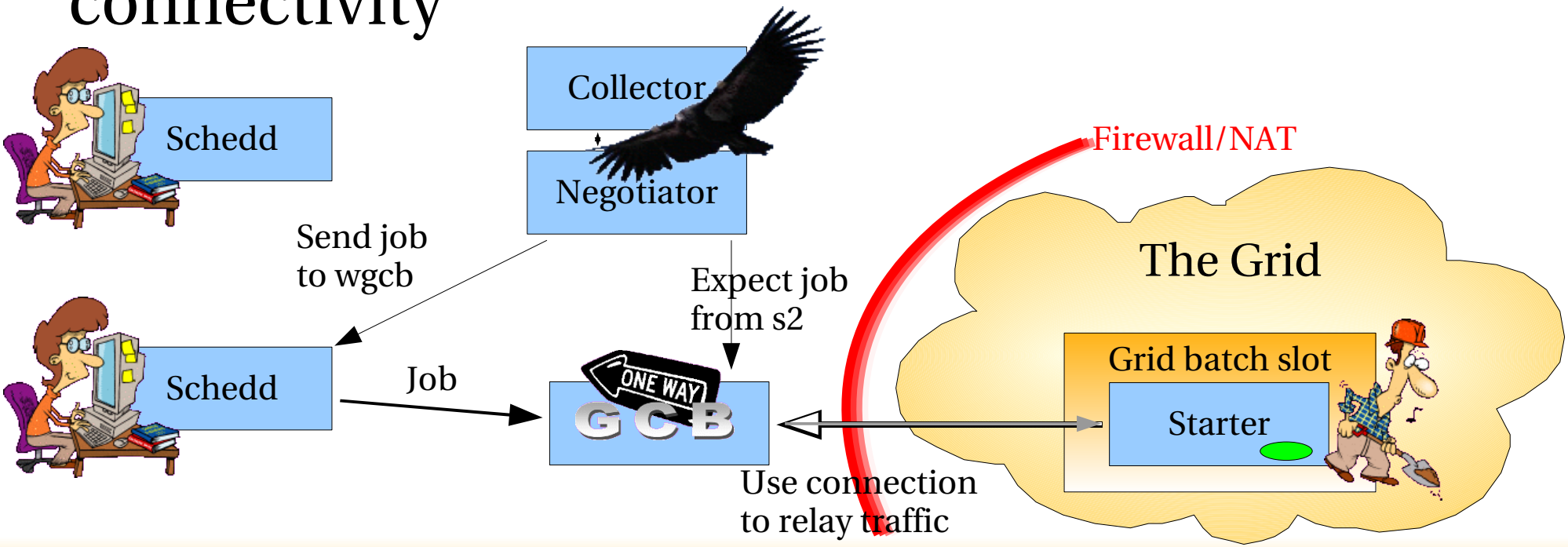
Does it really work over WAN? (1)

- Yes, but it needs GCB to work
  - A Condor proxy server
- The only requirement is that there is outgoing connectivity



Does it really work over WAN? (2)

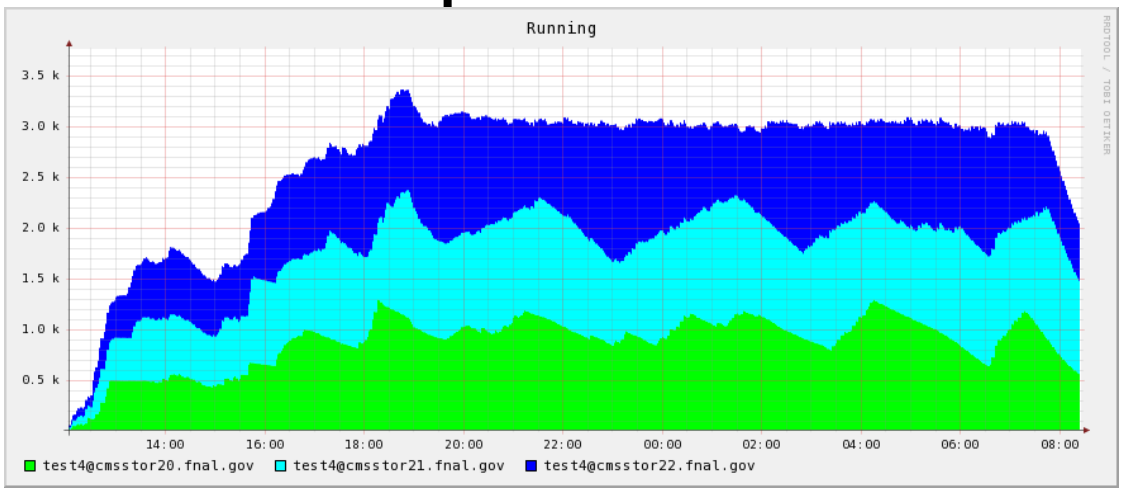
- Yes, but it needs GCB to work
  - A Condor proxy server
- The only requirement is that there is outgoing connectivity

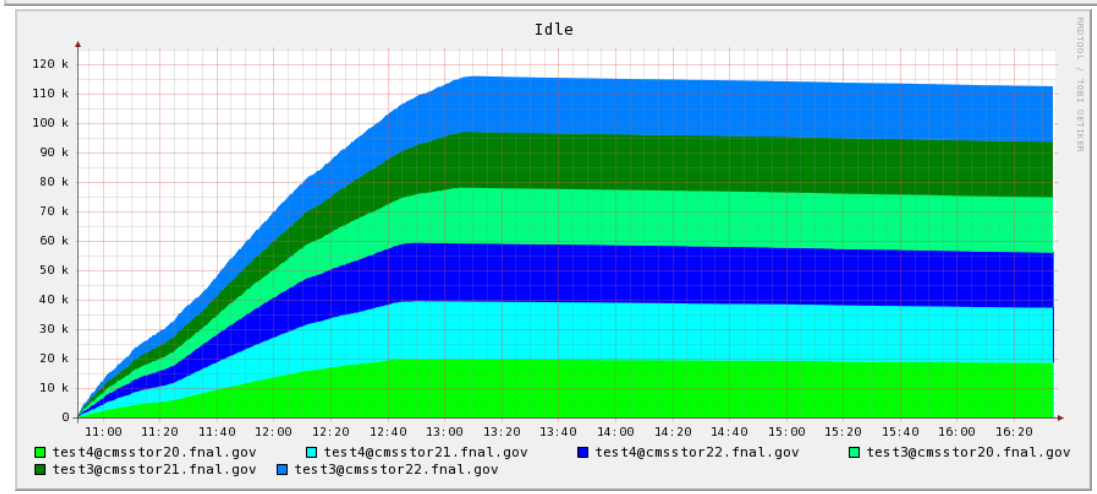


Does it scale? (1)

Ignore the colors

- Synthetic tests scaled well up to
  - 3000 running jobs
  - 100k queued jobs

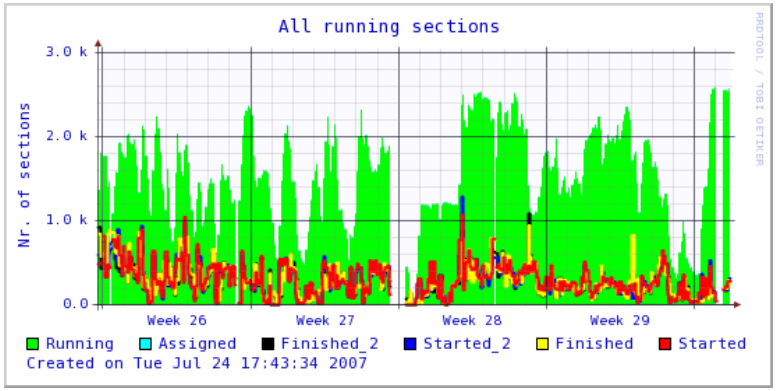


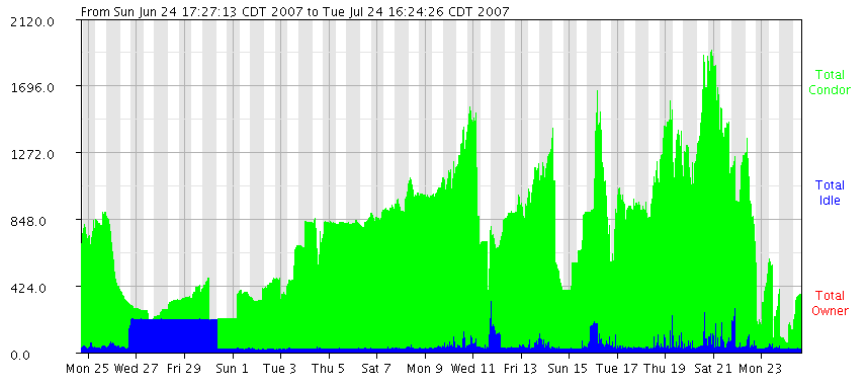


For more details, see http://home.fnal.gov/~sfiligoi/glideinWMS/presentations/osg_lajolla0703/OSG_WMS_0703_final.pdf

Does it scale? (2)

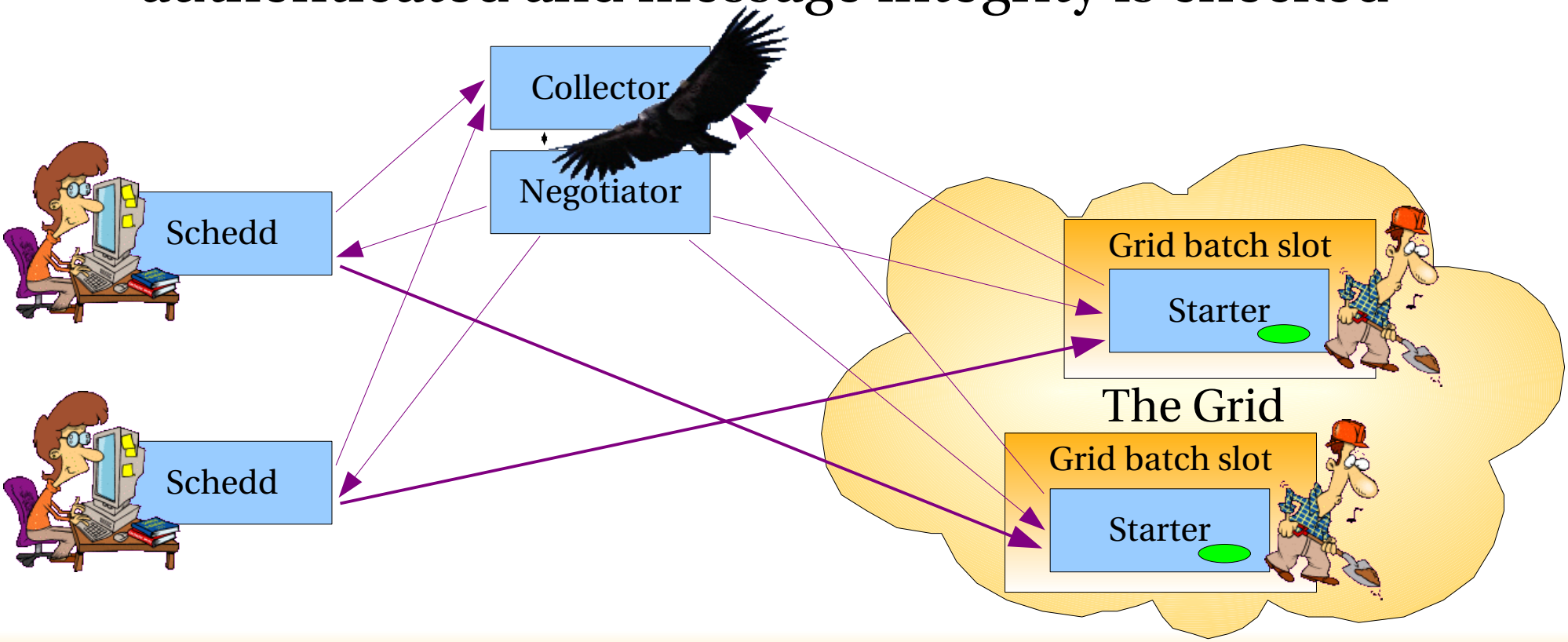
- CMS is running MC production using glideinWMS with up to 900 jobs in parallel
  - No graph, sorry
- CDF running production GlideCAF up to 2.5k
- ATLAS Cronus running up to 1.8k





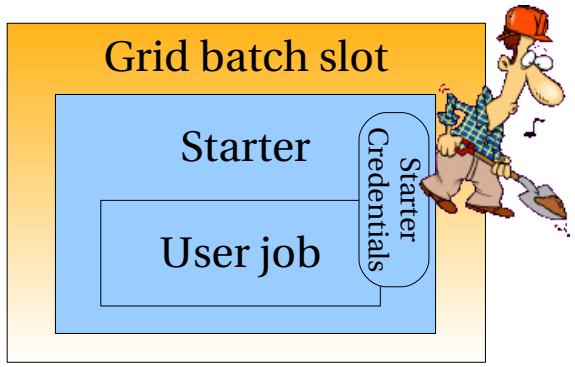
What about security? (1)

- glideinWMS glideins use GSI for authentication
  - All daemon to daemon communication is fully authenticated and message integrity is checked



What about security? (2)

- However, starter does not run as root!
  - This means that the user and the starter have to run under the same account!
    - The malicious user job can use starter privileges
  - A glidein factory should thus only run jobs from trusted users; i.e.
    - the factory owner, or
    - people you would entrust with a blank check

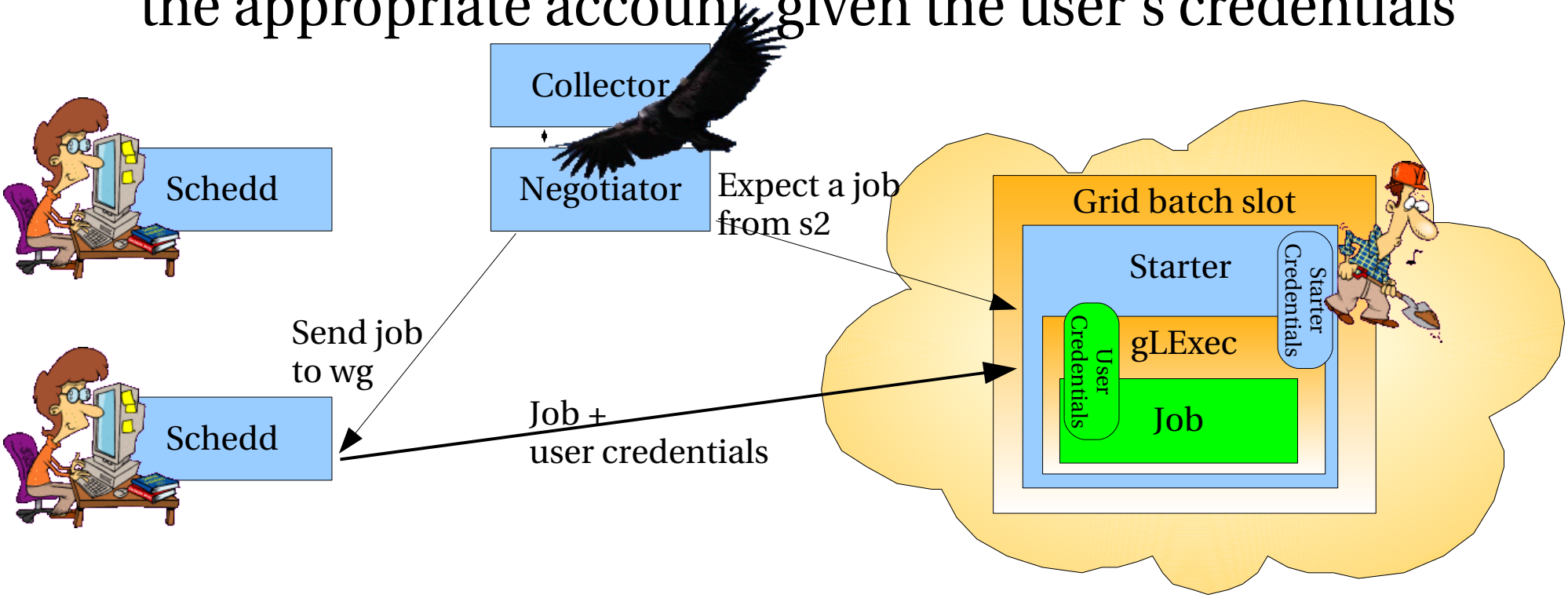


Running jobs from multiple users (1)

- The true power of glideins is however in running jobs from multiple users
- gLExec tries to solve the security problem
  - By allowing the starter to run the job under the user account
  - It also makes Grid sites happy, by letting them know who the real users are
- gLExec will be part of the upcoming OSG 0.8
  - Already installed and in use at Fermilab since Nov'06

Running jobs from multiple users (2)

- gLExec is a privileged executable, similar to Apache suexec
  - The starter can use it to execute the user job under the appropriate account, given the user's credentials



Any other drawbacks?

- If anything goes wrong with the setup, the debugging can be challenging
  - GCB logs are difficult to understand
    - Better logging promised for future versions
  - Starter log files are returned only when the job finishes
    - May not get them back, if it never ends
    - This is a fundamental Grid limitation, not much that can be done about it
- GCB is currently the weak point
  - If it dies, all jobs served by that GCB will be restarted

Conclusions

Condor Glideins

- Can shield user jobs from the Grid
- Give you total control over your jobs
- Allow you to have more control over the jobs scheduling

GlideinWMS

- An automatic way to create glidein pools on the fly
- Needs some initial effort, but then it operates on its own

glideinWMS home page

http://home.fnal.gov/~sfiligoi/glideinWMS/

sfiligoi@fnal.gov