

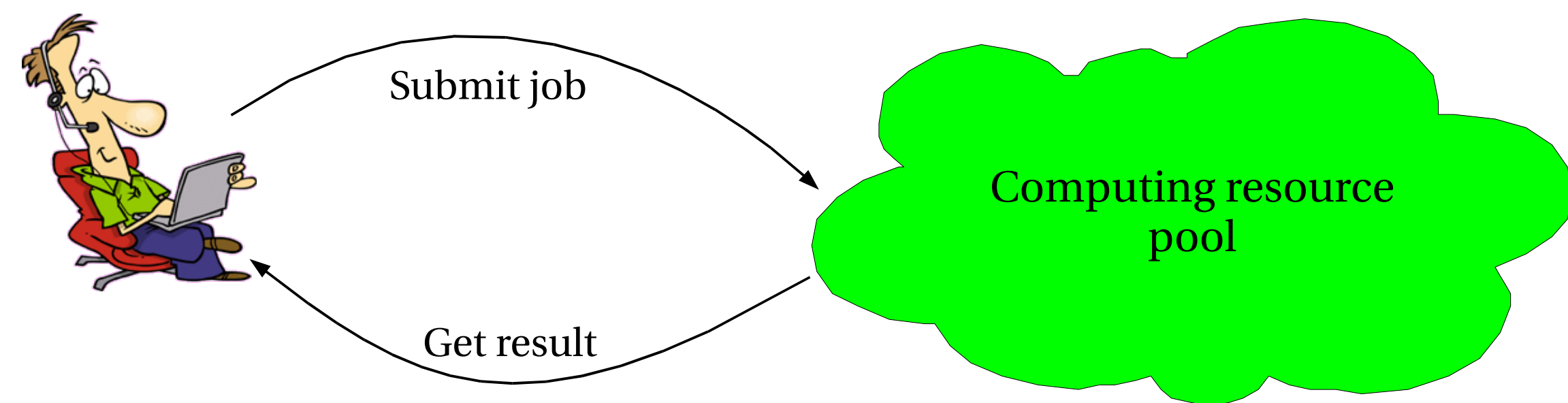
glideinWMS

A pilot-based Grid Workload Management System

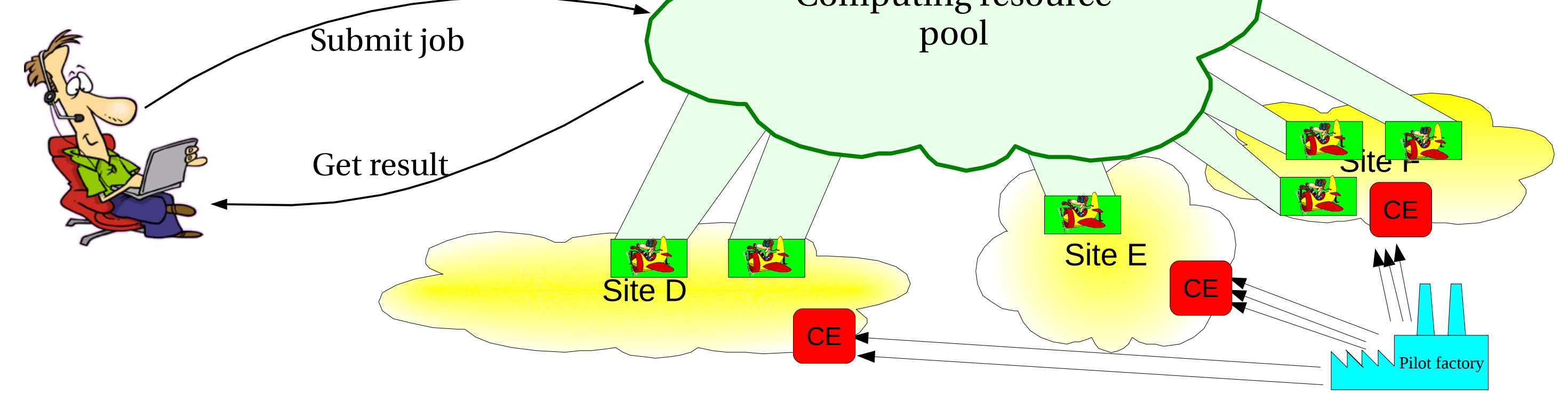
presented by Igor Sfiligoi

Why use a pilot-based WMS?

Users want simple access to compute resources

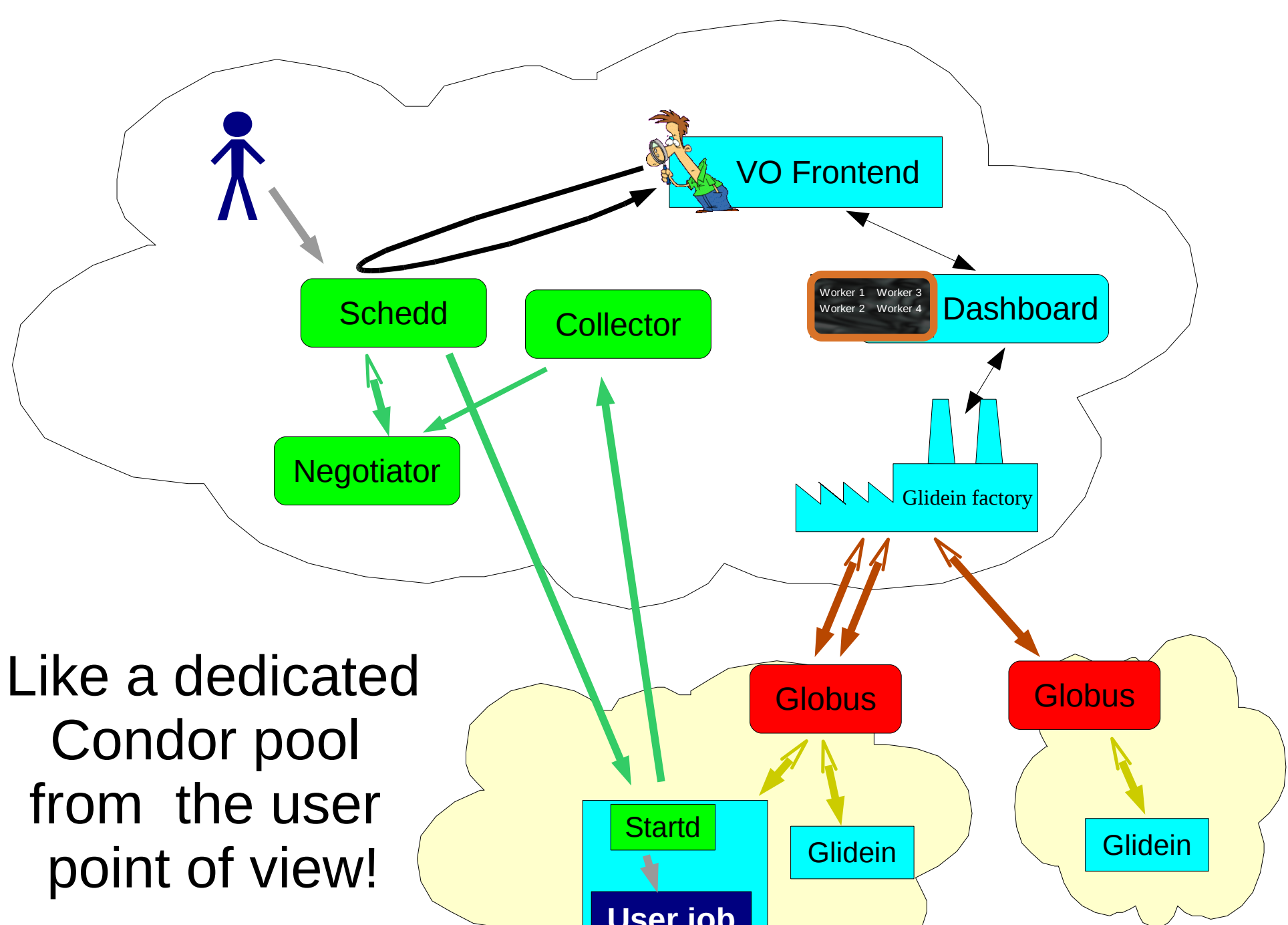


A pilot WMS creates a virtual private pool over Grid resources



glideinWMS architecture

Schematic view



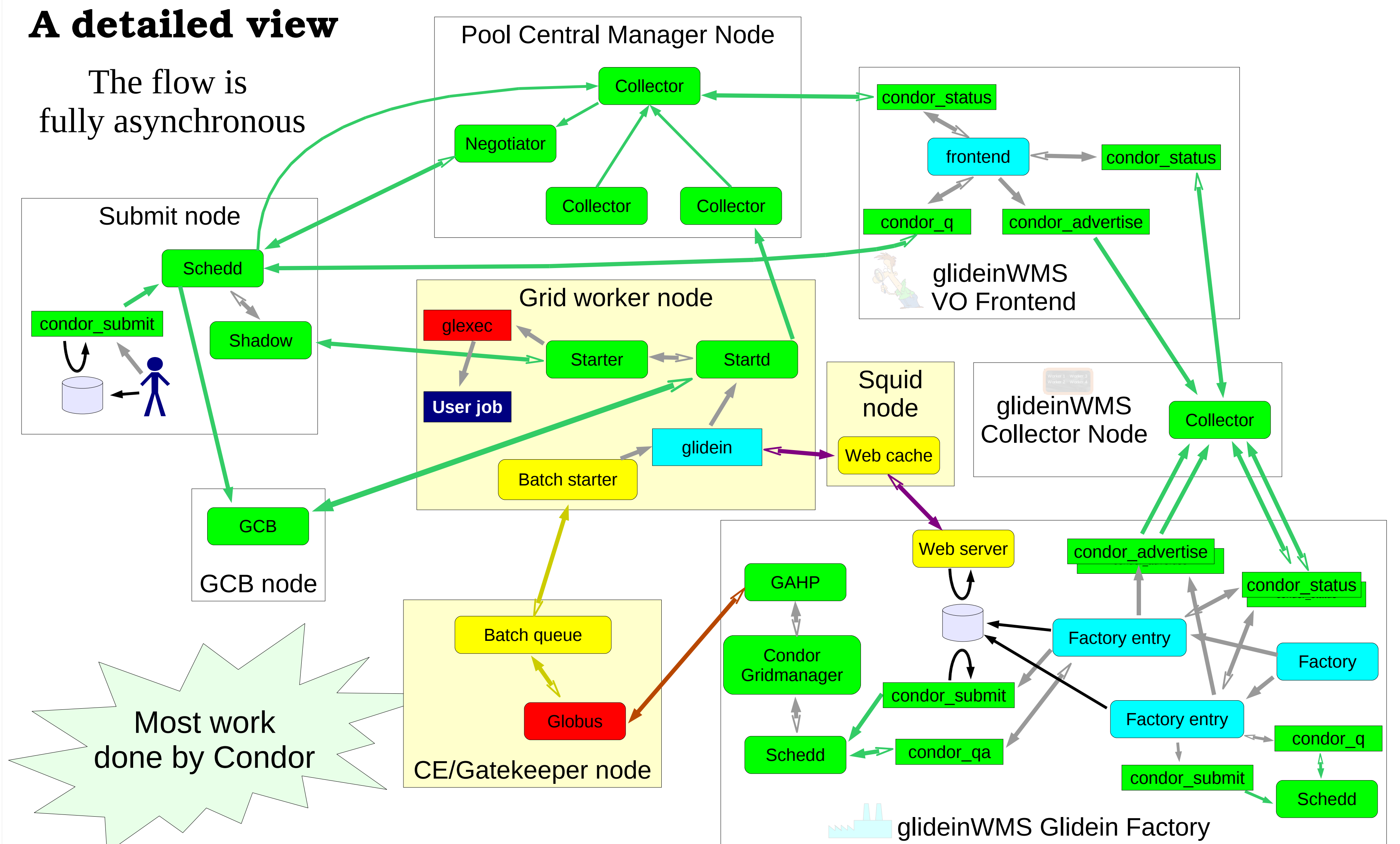
Like a dedicated Condor pool from the user point of view!

Legend:

- Condor process
- glideinWMS process
- Grid software
- System software
- glideinWMS domain
- Grid domain (OSG, EGEE and NorduGrid tested)

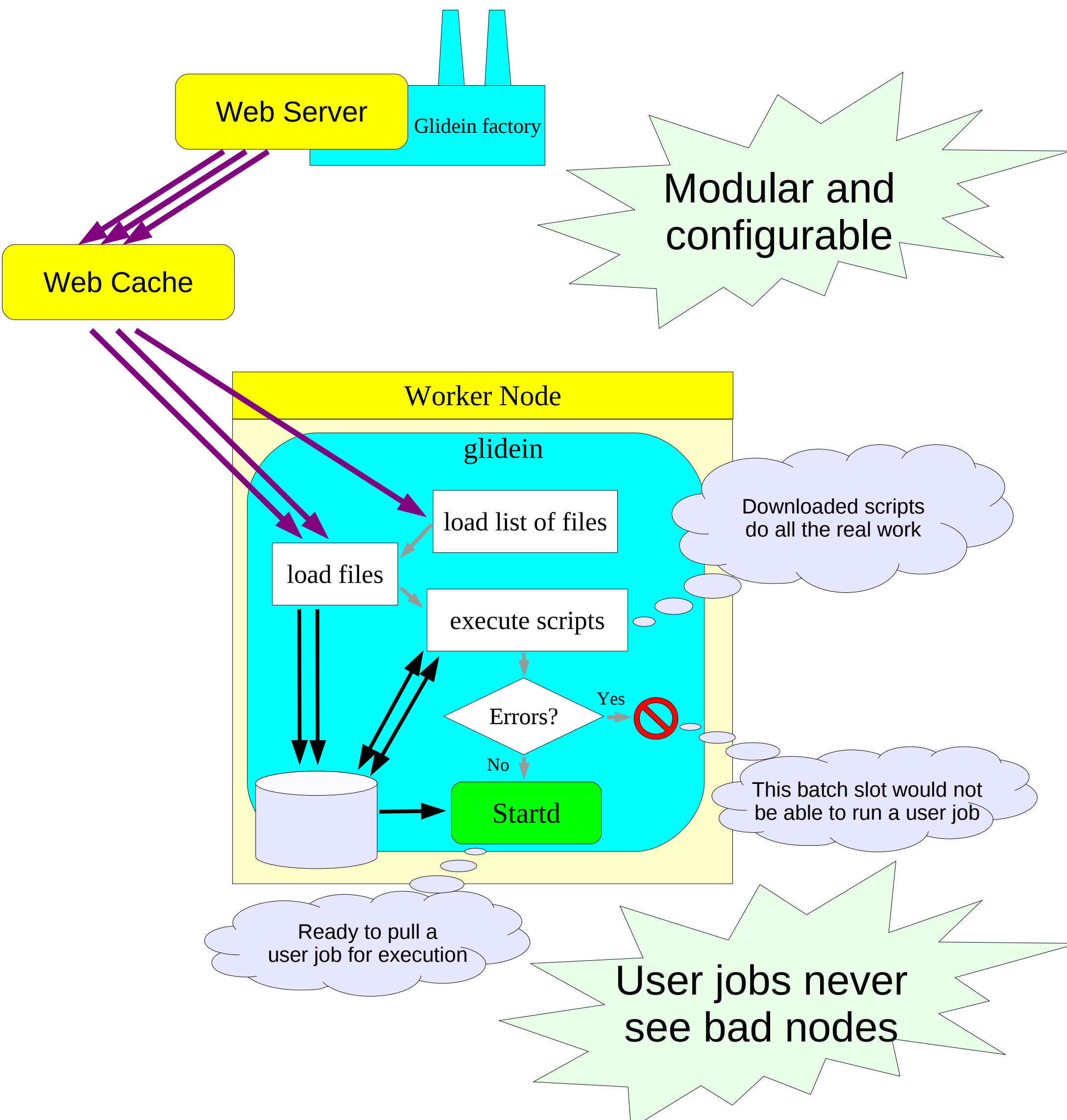
A detailed view

The flow is fully asynchronous



For more details see <http://www.uscms.org/SoftwareComputing/Grid/WMS/glideinWMS/doc.html>

Glidein startup logic



Monitoring

Standard Condor tools

```

sfiligoi@cmsrv37:~$ condor_q
-- Submitter: cmsrv37.fnal.gov : <131.225.204.208:60235> : cmsrv37.fnal.gov
ID OWNER SUBMITTED RUN TIME ST PRI SIZE CMD
214.0 sfiligoi 2/18 15:31 0+00:04:15 R 0 0.0 a.sh
214.1 sfiligoi 2/18 15:31 0+00:04:15 R 0 0.0 a.sh
214.2 sfiligoi 2/18 15:31 0+00:04:15 R 0 0.0 a.sh
214.3 sfiligoi 2/18 15:31 0+00:01:14 R 0 0.0 a.sh
214.4 sfiligoi 2/18 15:31 0+00:01:14 R 0 0.0 a.sh
214.5 sfiligoi 2/18 15:31 0+00:00:00 I 0 0.0 a.sh
214.6 sfiligoi 2/18 15:31 0+00:00:00 I 0 0.0 a.sh
214.7 sfiligoi 2/18 15:31 0+00:00:00 I 0 0.0 a.sh
214.8 sfiligoi 2/18 15:31 0+00:00:00 I 0 0.0 a.sh
214.9 sfiligoi 2/18 15:31 0+00:00:00 I 0 0.0 a.sh
10 jobs; 5 idle, 5 running, 0 held
sfiligoi@cmsrv37:~$ condor_status
Name OpSys Arch State Activity LoadAv Mem ActvtyTime
glidein_134140cmsw LINUX X86_64 Claimed Busy 2.520 2068 0+00:00:03
glidein_156070cmsw LINUX X86_64 Claimed Busy 2.120 2068 0+00:00:04
monitor_134140cmsw LINUX X86_64 Owner Idle 2.550 200 0+00:00:13
monitor_156070cmsw LINUX X86_64 Owner Idle 2.160 200 0+00:00:11
glidein_321300cmsw LINUX X86_64 Claimed Busy 1.820 3949 0+00:00:03
glidein_441000cmsw LINUX X86_64 Claimed Busy 1.820 3949 0+00:00:03
monitor_321300cmsw LINUX X86_64 Owner Idle 1.310 394 0+00:00:10
monitor_441000cmsw LINUX X86_64 Owner Idle 1.610 394 0+00:00:09
glidein_323240cmsw LINUX X86_64 Claimed Busy 2.940 3950 0+00:00:04
monitor_323240cmsw LINUX X86_64 Owner Idle 3.000 395 0+00:00:09
Total Owner Claimed Unclaimed Matched Preempting Backfill
X86_64/LINUX 10 5 5 0 0 0 0 0 0
Total 10 5 5 0 0 0 0 0 0
sfiligoi@cmsrv37:~$
    
```

Both for the users and glideinWMS administrators

Pseudo-interactive monitoring (to peek at Grid jobs in real time)

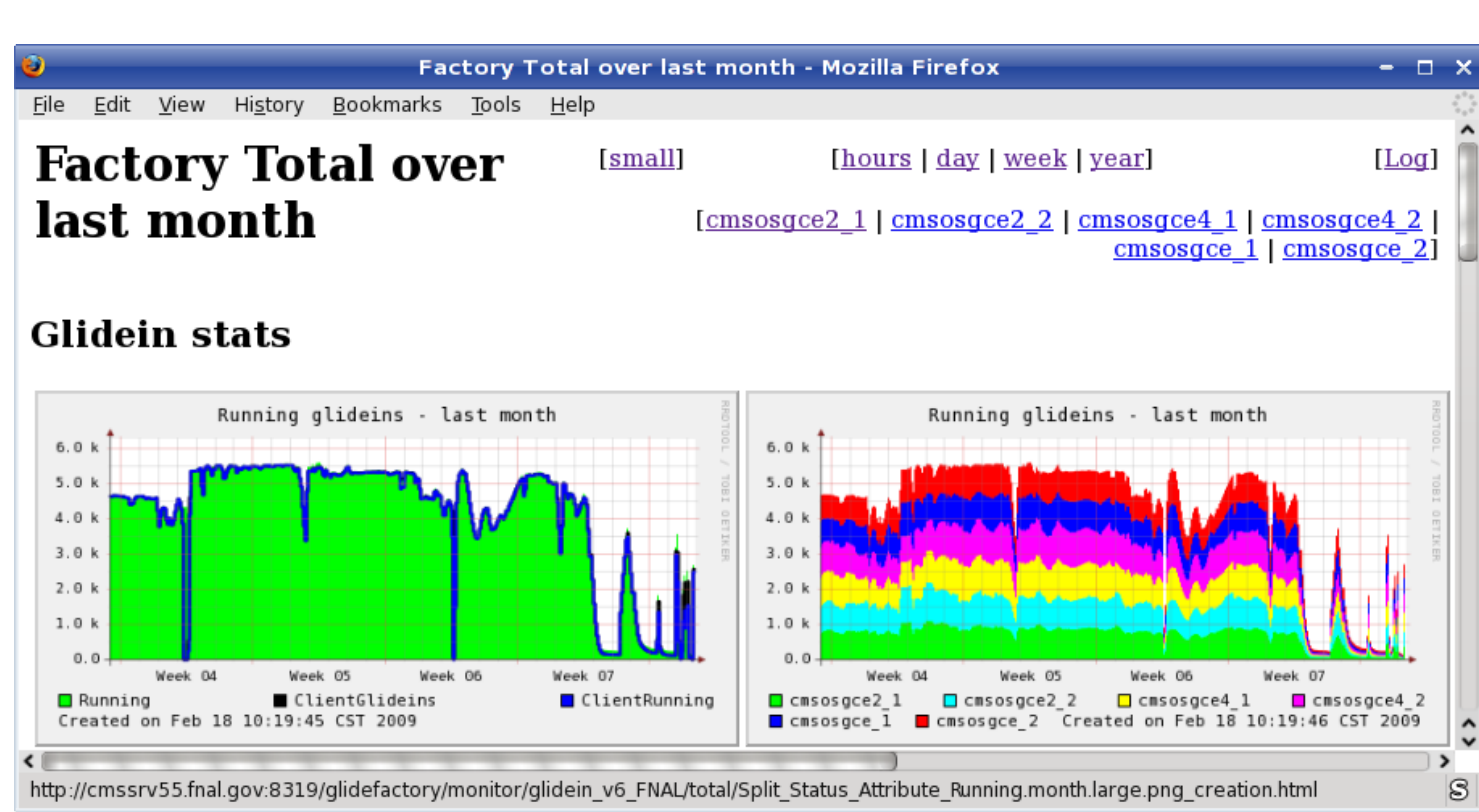
```

sfiligoi@cmsrv37:~/submit/glideinWMS/tools - Shell - Konsole
[ifiligo@cmsrv37 tools]$ ./glidein_top.py 214.2
top - 15:45:18 up 26 days, 23:21, 0 users, load average: 1.55, 1.28, 1.17
Tasks: 180 total, 3 running, 177 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.3% us, 1.6% sy, 87.5% ni, 8.8% id, 1.6% wa, 0.0% hi, 0.1% si
Mem: 4044744k total, 4007316k used, 37428k free, 22476k buffers
Swap: 33551744k total, 115768k used, 33435976k free, 2724068k cached

PID USER PR NI CPU TIME+ MEM VIRT RES SHR S COMMAND
4183 samgrid 35 10 100 308:16.78 0.6 110m 23m 572 R sprob-p20mfast
3323 root 34 10 2 234:11.21 0.0 0 0 0 R kmpir
18918 cmsprod 26 10 2 0:14.69 0.0 3536 1948 988 S condor_procd
27316 cmsprod 25 10 2 1:56.09 28.1 1097m 795m 105m S cmsRun
1 root 16 0 0 0:05.61 0.0 4756 352 320 S init
2 root RT 0 0 0:09.95 0.0 76 0 0 S migration/0
3 root 34 19 0 0:00.50 0.0 0 0 0 S ksoffirgd/0
4 root RT 0 0 0:08.03 0.0 0 0 0 S migration/1
5 root 34 19 0 0:00.46 0.0 0 0 0 S ksoffirgd/1
6 root RT 0 0 0:06.42 0.0 0 0 0 S migration/2
    
```

- Supports:
- ls
 - cat/tail
 - ps
 - top
 - ...

Web monitoring



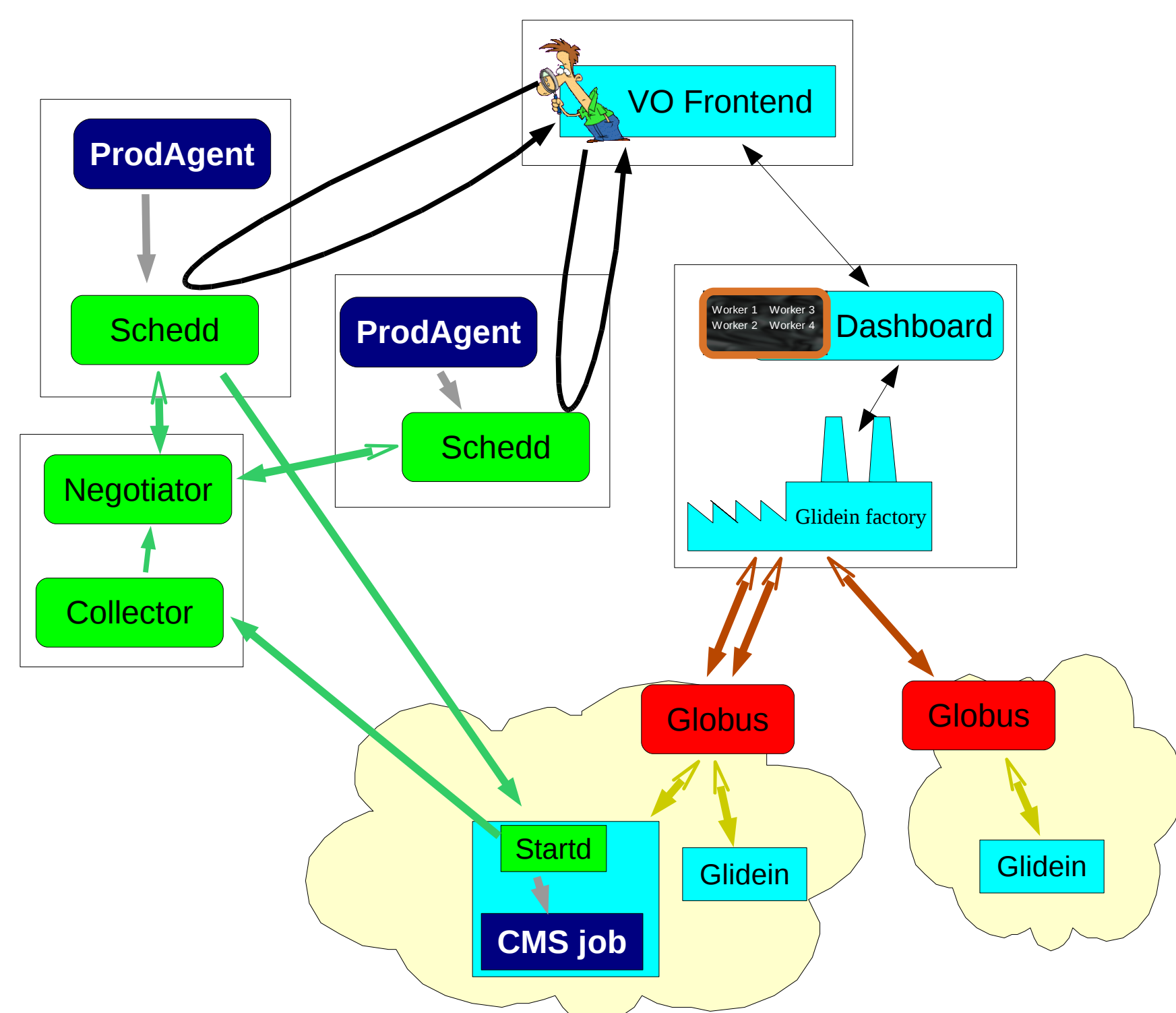
Glidein factory only (being improved)

Current users of glideinWMS

presented by Igor Sfiligoi

CMS Processing

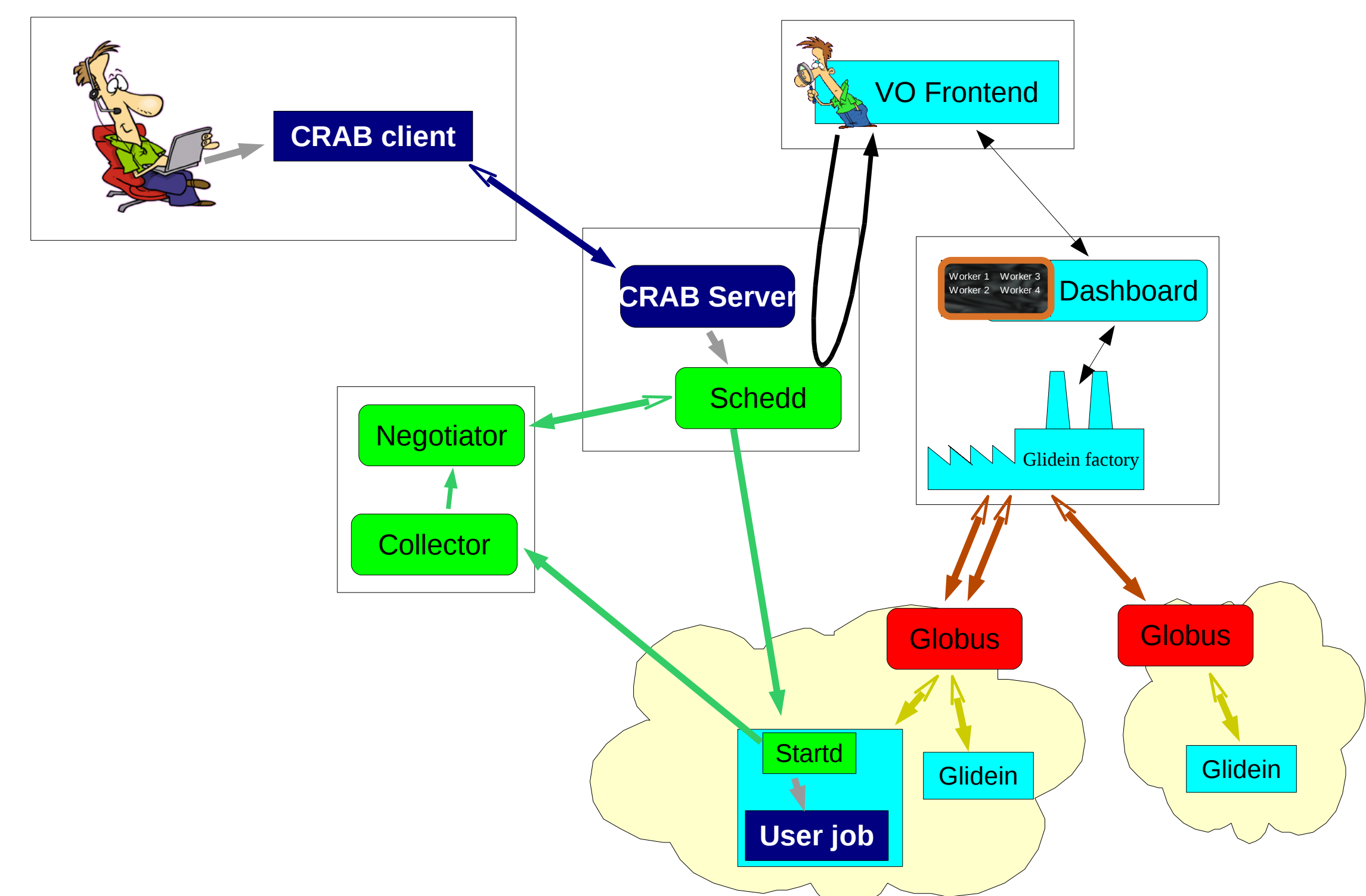
Fully automated, process driven



Has been running for over a year now
Using both OSG and EGEE CMS Tier-1's

CMS User Analysis

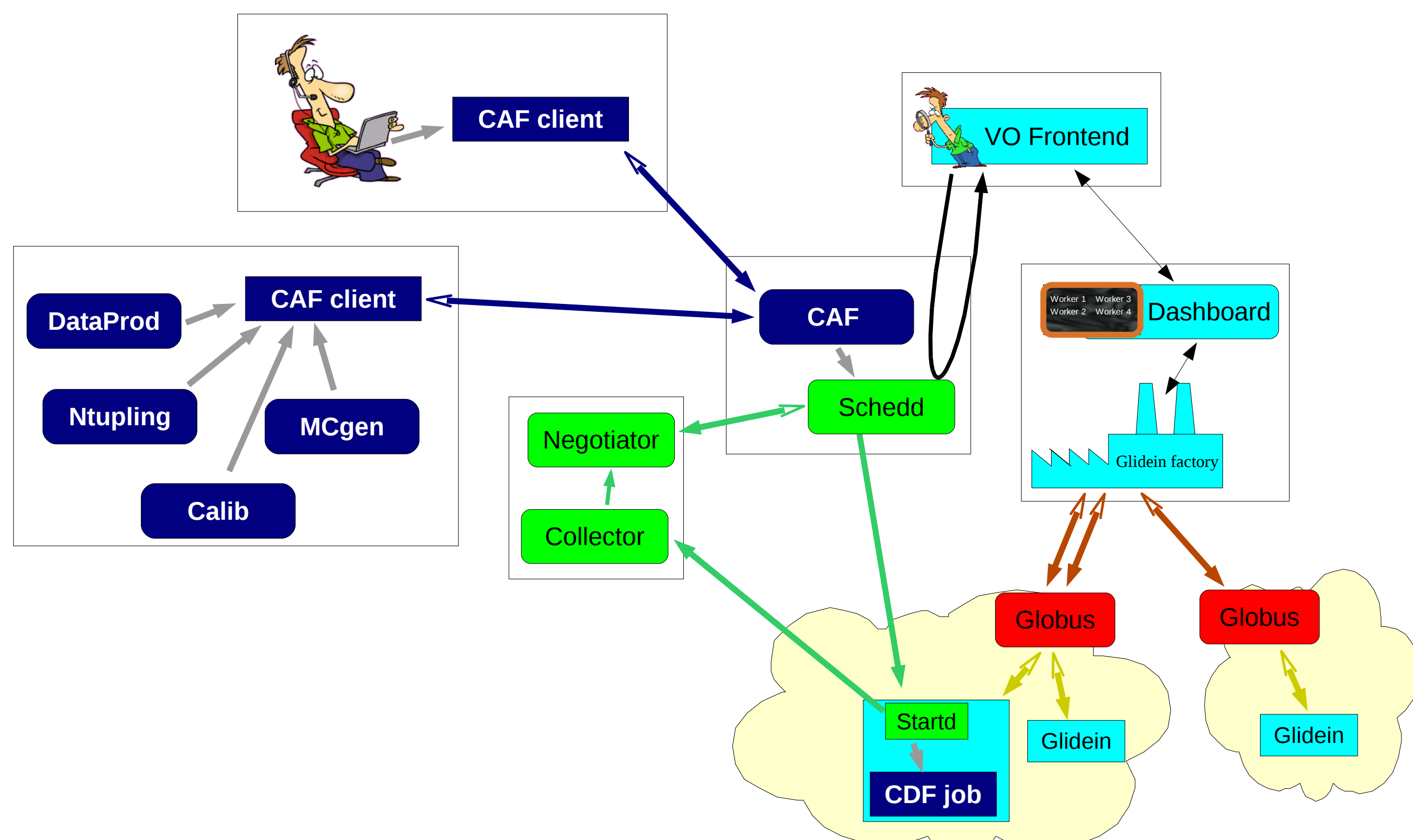
Portal based
Users don't see the glideinWMS



Still a prototype
Using both OSG and EGEE CMS Tier-2's

CDF's Combined User Analysis and Centralized Activities

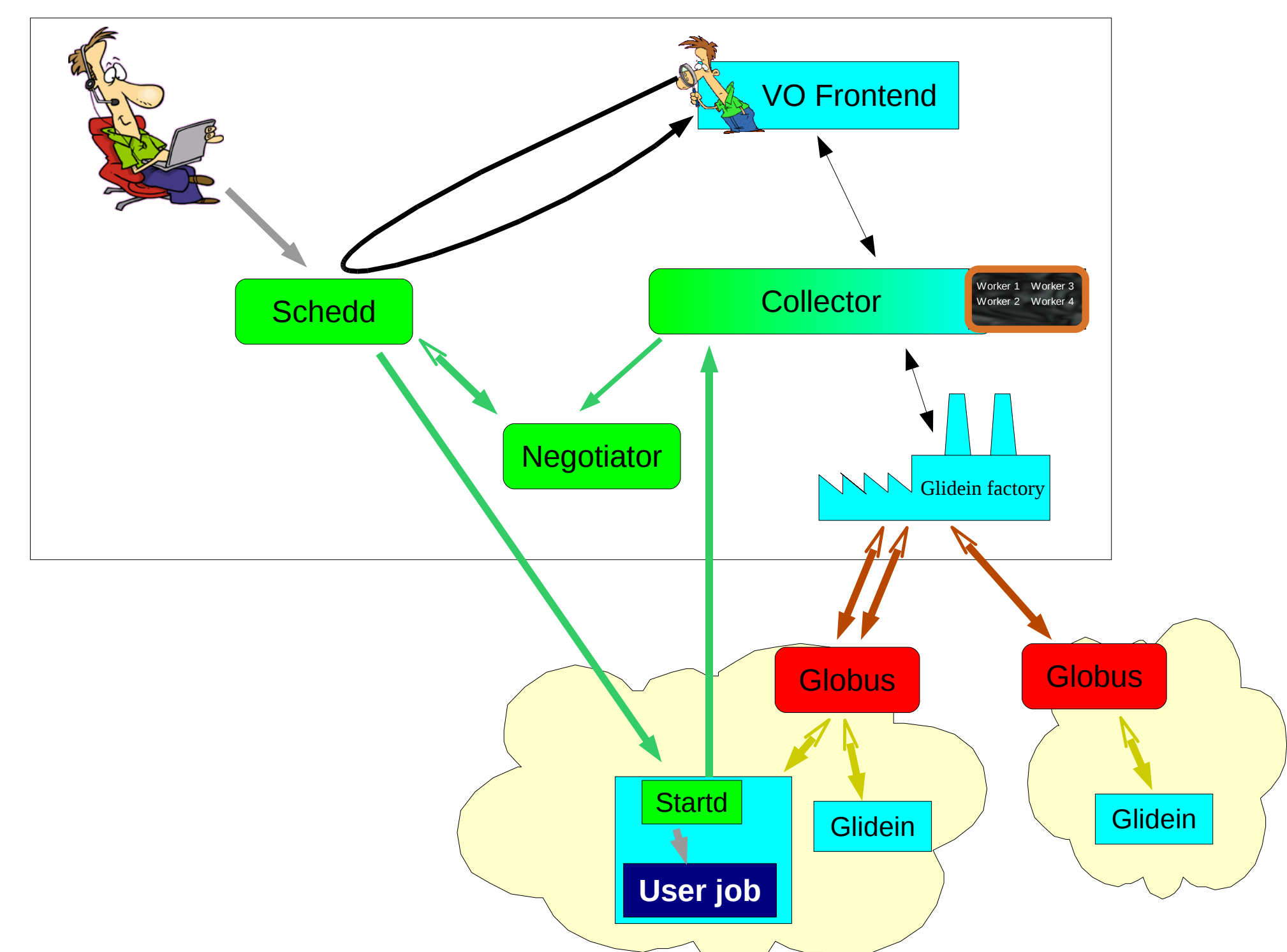
Portal based
Users and automated tasks don't see the glideinWMS



Used glidein-based solution for several years.
Currently moving to glideinWMS.
Using both CDF-owned and opportunistic OSG resources.

MINOS User Analysis

Users submit to a central schedd
All processes share a single machine



Has been running for a year now
Using both MINOS-owned and opportunistic OSG resources at Fermilab.

Scalability tests

Condor scalability tested

Collector scalability

- 11k on a single collector, on LAN
- 11k using a tree of 1+20 collectors over WAN (between Europe and USA)

Schedd scalability

- 11k running jobs on a single schedd **current limit** with 16GB of memory
- 200k idle jobs on a single schedd

GCB scalability

- 8k glideins on a single GCB
- 11k glideins using 2 GCBs

Memory limited

glideinWMS scalability tested

VO Frontend scalability

- 200k idle and 11k running user jobs
- 20 schedds

Glidein factory scalability

- 40 Grid sites **current limit**
- 100 Grid sites with some effort
- 5 VO frontends
- 11k running glideins

Prototype that scales higher available