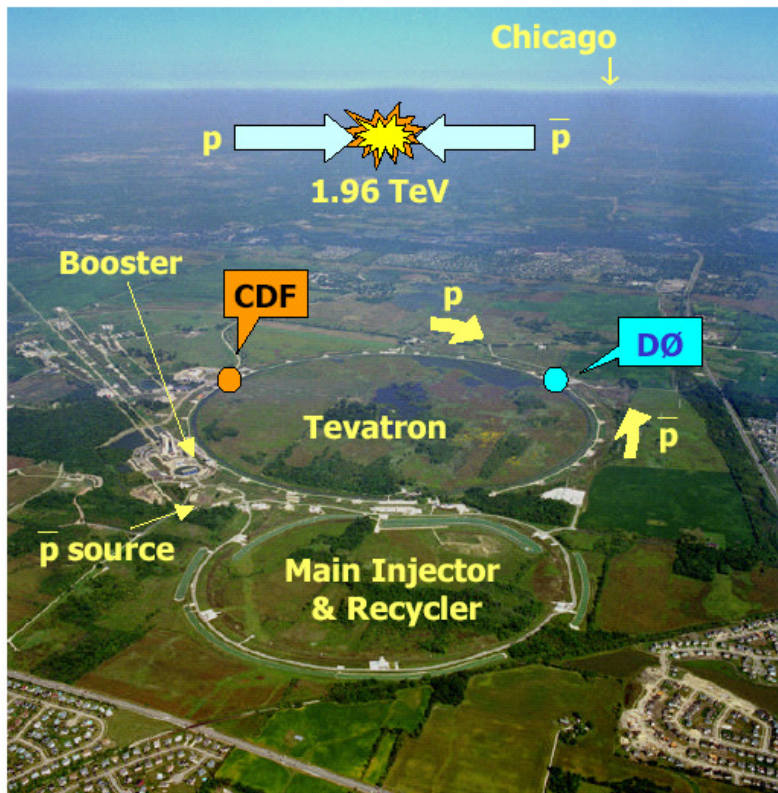
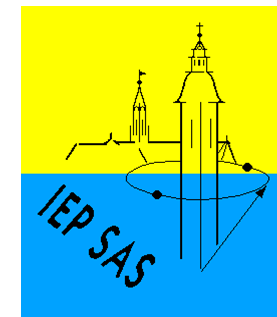


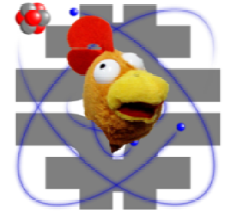
CDF GlueX WMS usage in Grid computing of High Energy Physics



Marian Zvada, CHEP 2009
(Fermilab/IEP SAS Kosice)

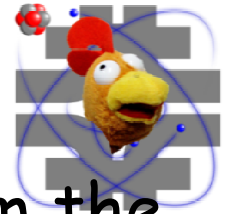


Outline



- > CDF CAF middleware
 - CAF middleware & Condor
 - Monitoring
- > GlideCAF
 - What are the glideins
- > CDF towards GlideinWMS
 - Why?
- > Computing facilities
- > GlideinWMS itself
- > Large scale tests
- > CDF GlideinWMS in production
- > Conclusions

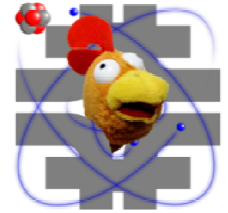
CDF CAF middleware



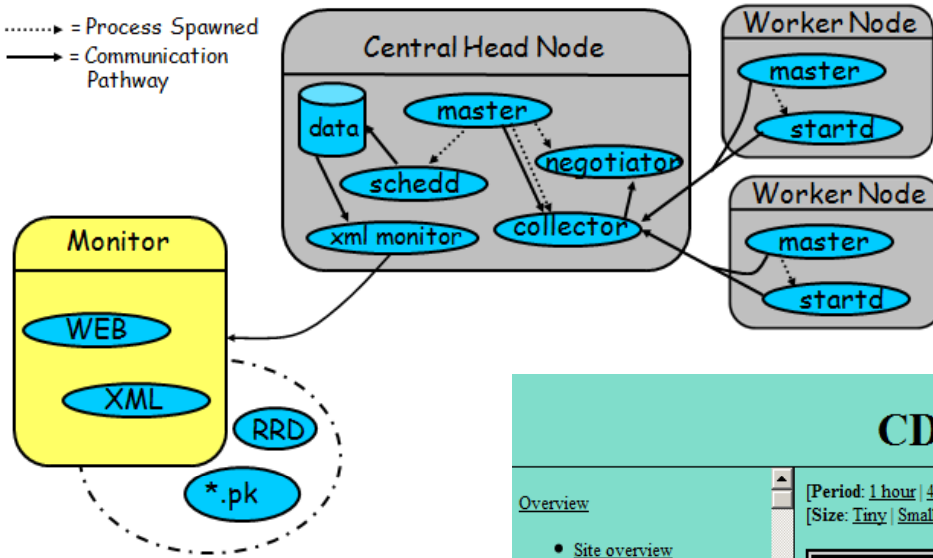
- > users can develop, debug and submit jobs from the desktop
- > authentication in secure way
 - kerberos user principal
 - X509 globus user proxy submitting to the grid
- > pseudo-interactive monitoring available
- > check the jobs status over the web-interface
- > no need to stay connected
- > notification and summary of the end of jobs via email



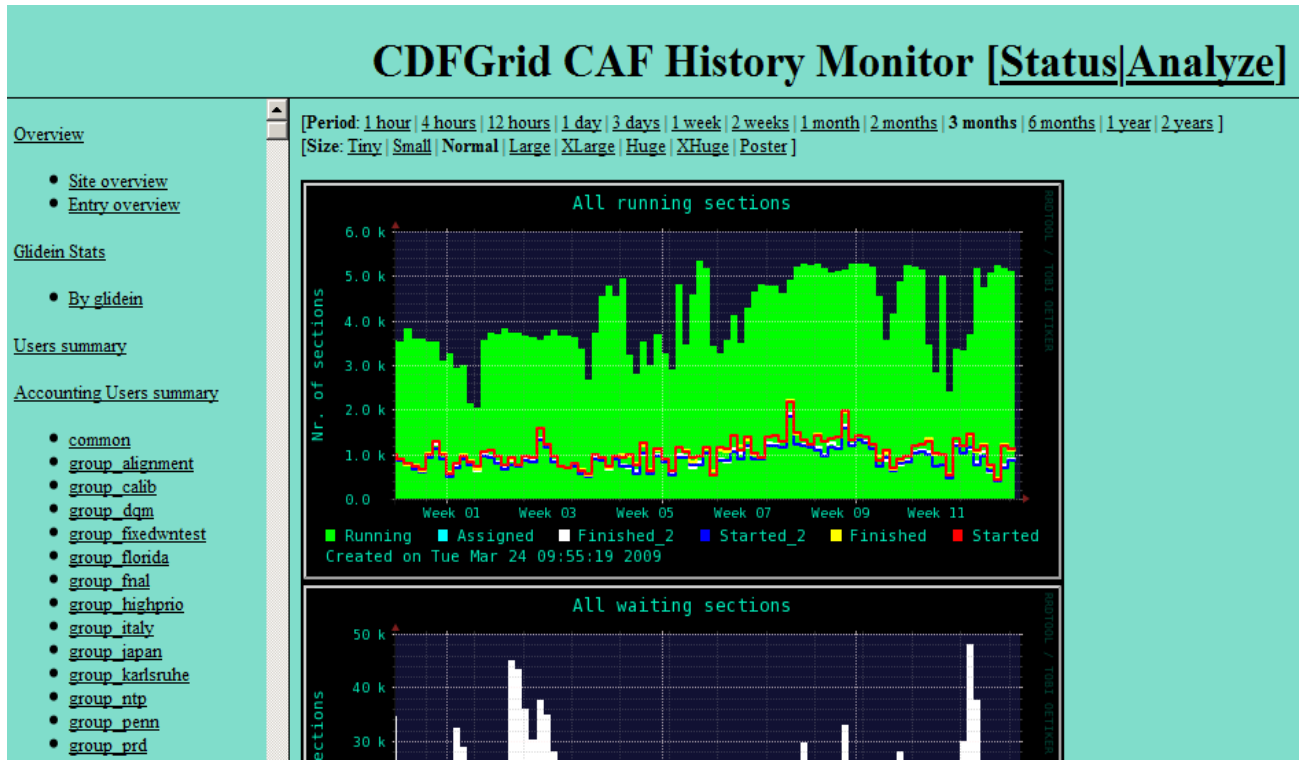
CDF CAF User Monitoring

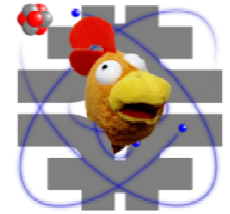


.....> = Process Spawned
 → = Communication Pathway



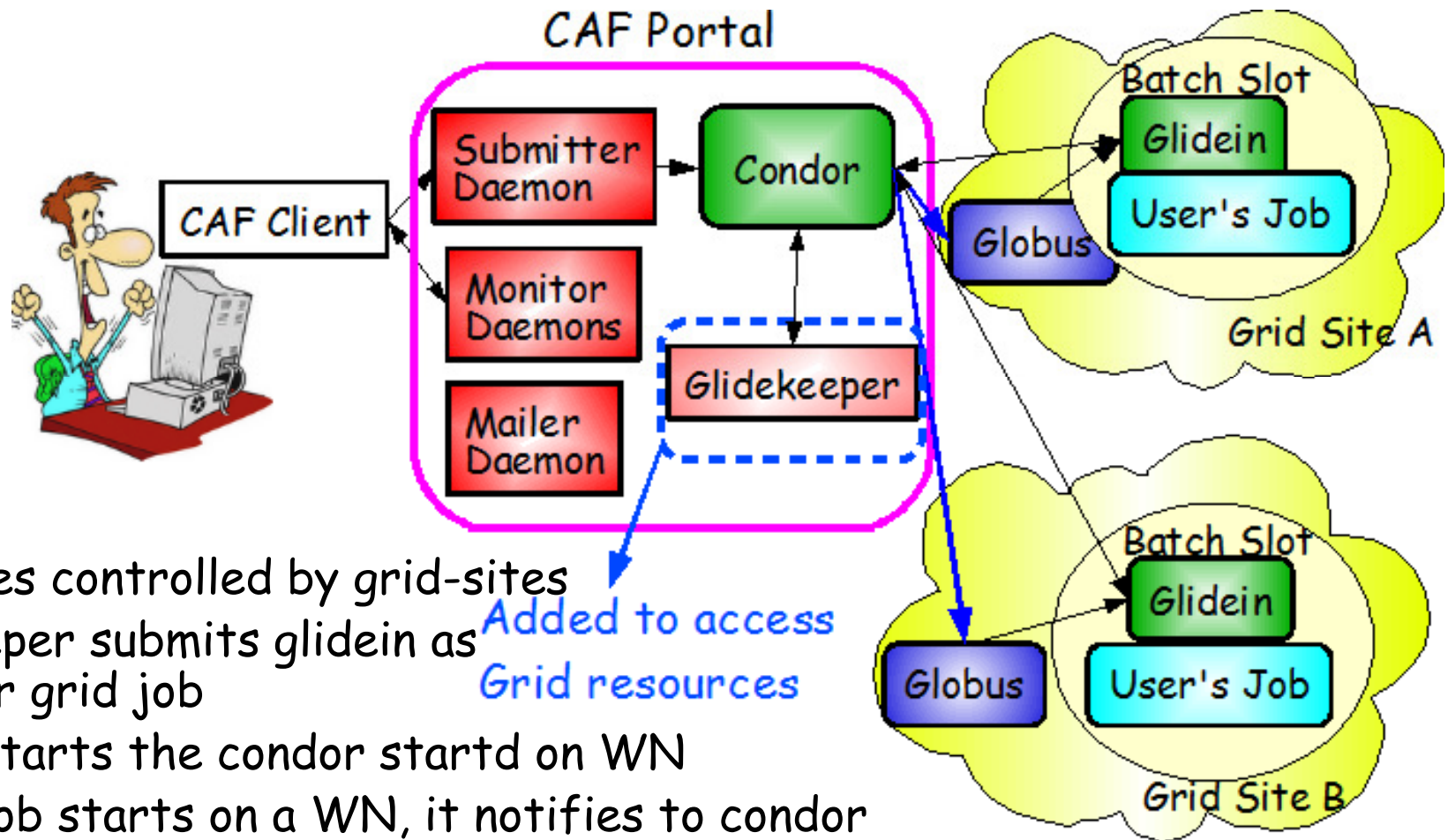
- > Pseudo-interactive monitor
- > User web monitoring



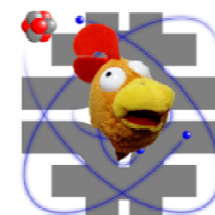


CAF towards to Grid (GlideCAF)

- > using just GRID resources

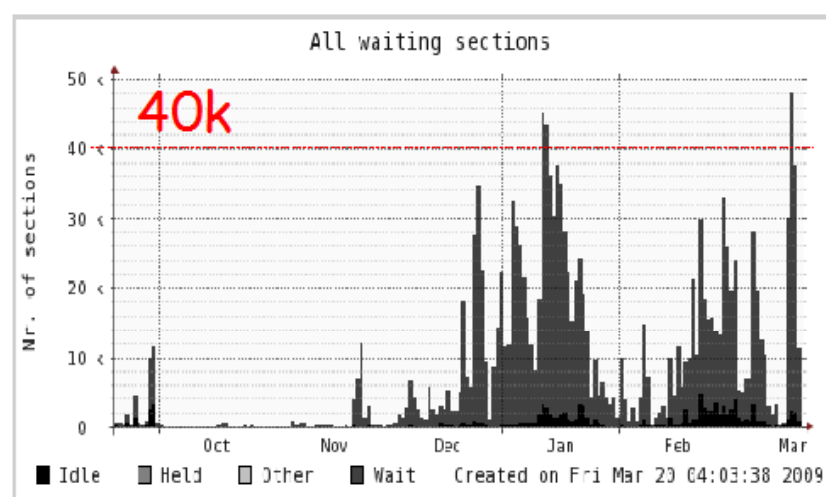
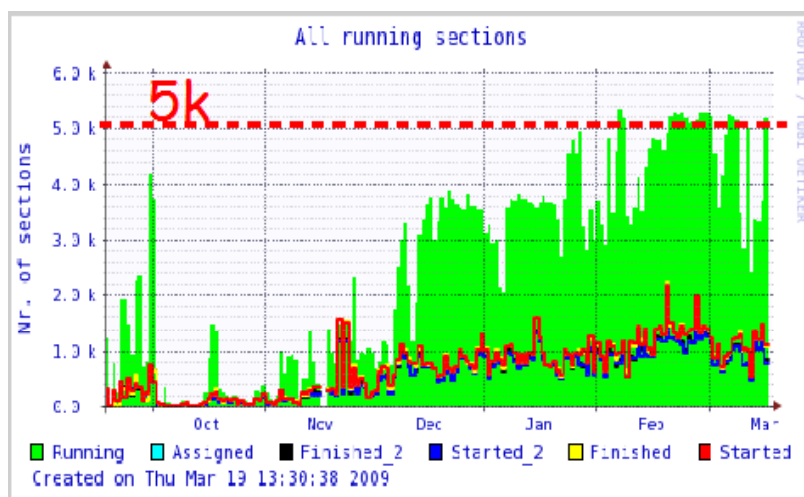


- resources controlled by grid-sites
- glidekeeper submits glidein as a regular grid job
- glidein starts the condor startd on WN
- once a job starts on a WN, it notifies to condor and join the pool as a new VM
- glidekeeper is first pilot towards to glidenWMS



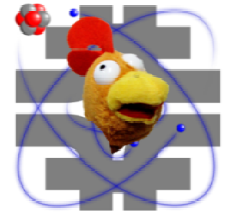
Towards CDF GlideinWMS

- Why going from glidekeeper to glideinWMS?
 - hit scalability issues



- no CDF manpower to improve glidekeeper
- new product, glideinWMS, already available on the market
- better workload management over the pilots (glideins) and their monitoring
- less memory/cpu consumption while running large number of jobs/glideins concurrently

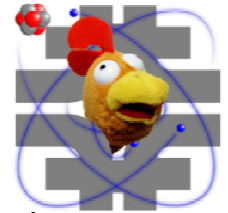
Computing facilities



Type	CPU type	GHz/CPU	Cores
1	XEON 1x2	2.6	2
2	XEON 1x2	3.06	2
3	2*XEON 1x2	3.0	2
4	2*DUAL CORE OPTERON 265 2x2	1.8	4
5	2*LOW VOLT XEON 2x2	2.33	4
6	DUAL QUAD CORE XEON X5355 2x4	2.66	8

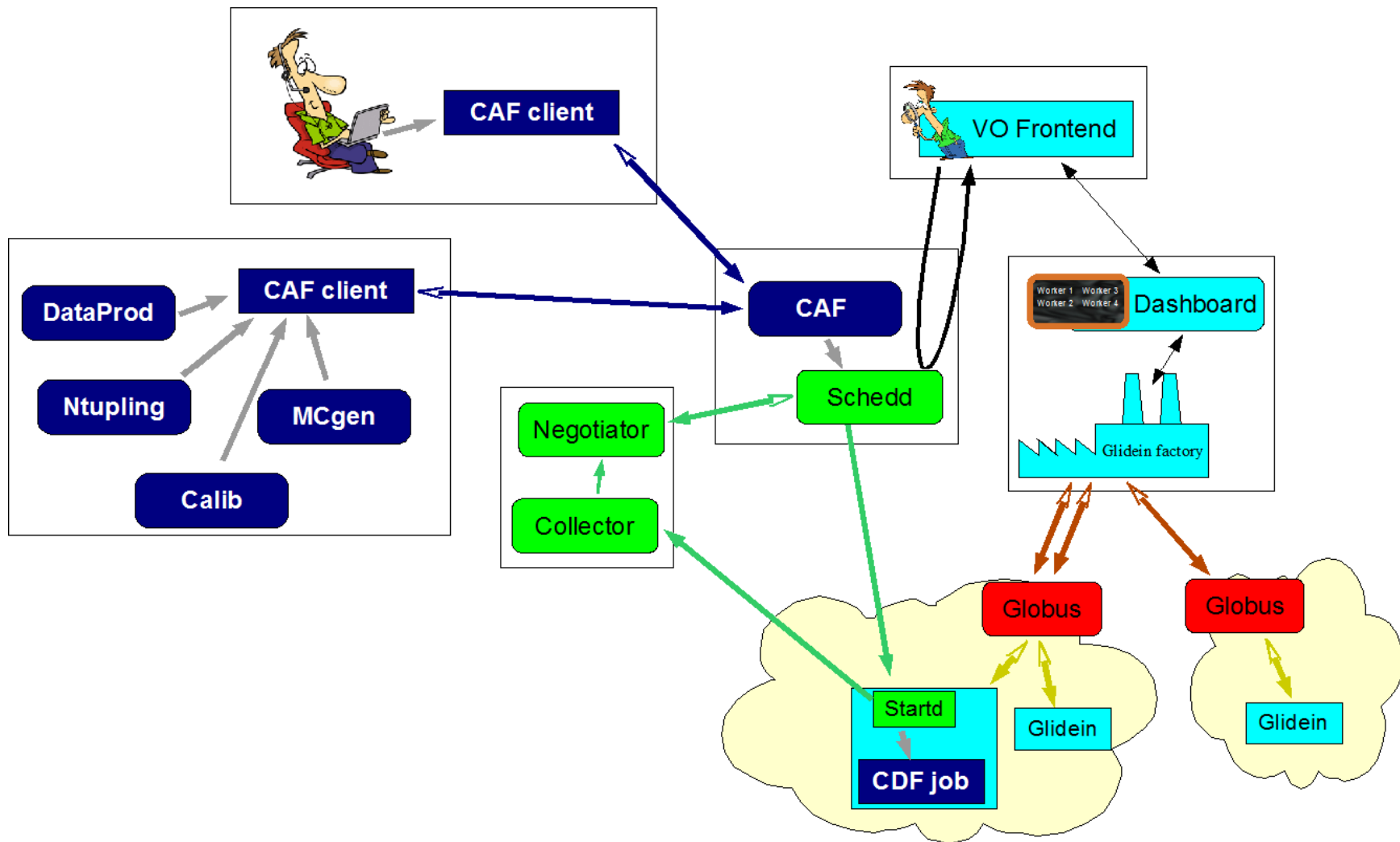
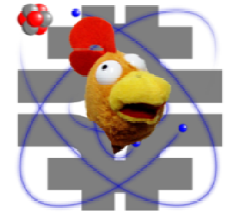
- > for CDF production data processing and user analysis
- > currently ~5000 slots at FNAL for CDF
- > Very powerful head nodes, currently serves all the CAF services for running jobs including condor batch system
 - 32GB RAM, DUAL Quad Core w/ 8 cores

The GlideinWMS overview

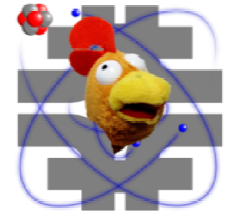


- > Generic pilot-based workload management system
- > glideinWMS is generalized version of the glidekeeper developed by USCMS@FNAL
- > www.uscms.org/SoftwareComputing/Grid/WMS/glideinWMS/
- > glideinWMS system is composed of **several elements**
 - > **Condor central manager** machine (collector/negotiator)
 - > **Condor submitter machine** (run the condor user schedds and keep the job queues + CDF CAF daemons)
 - > **Glidein Factory machine** (run glidein factory daemon that will submit the pilot jobs to a set of Grid pools)
 - > **VO Frontend machine** (frontend daemon monitors the schedd queues, matches them to the glidein factories, and decides which factory submit the pilots and how many)
 - > **WMS collector machine** (used for communication between the glidein factory daemons and the VO frontend daemons)

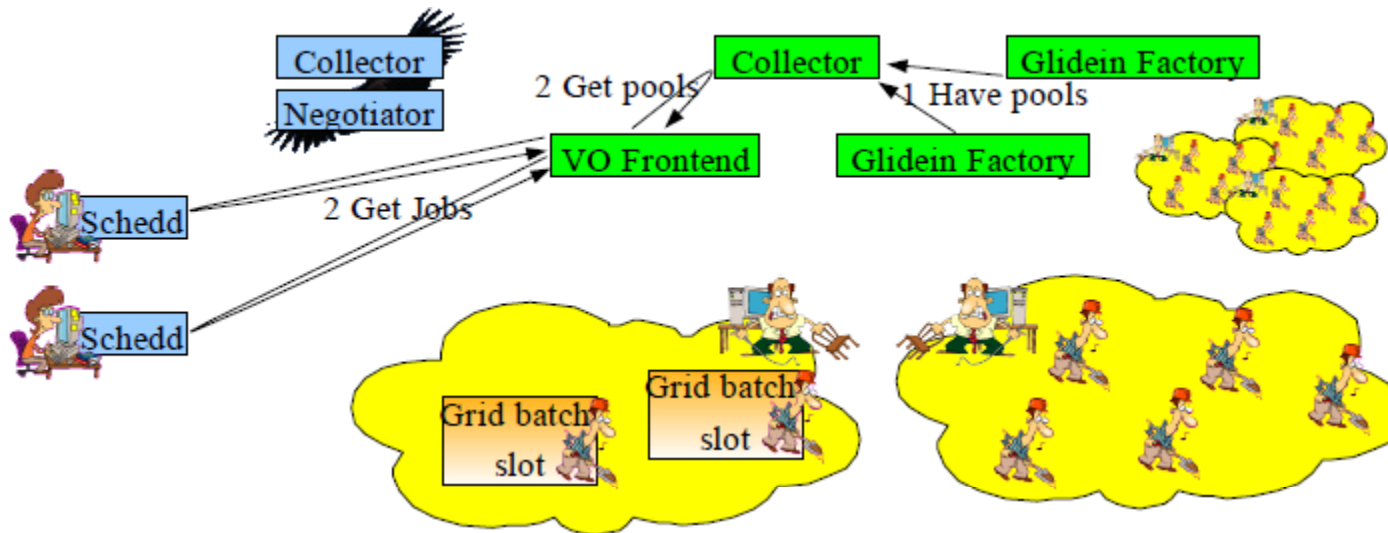
CDF glideinWMS overview (1)



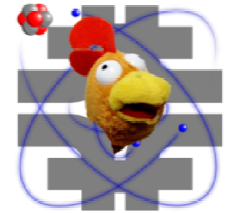
glideinWMS overview (2)



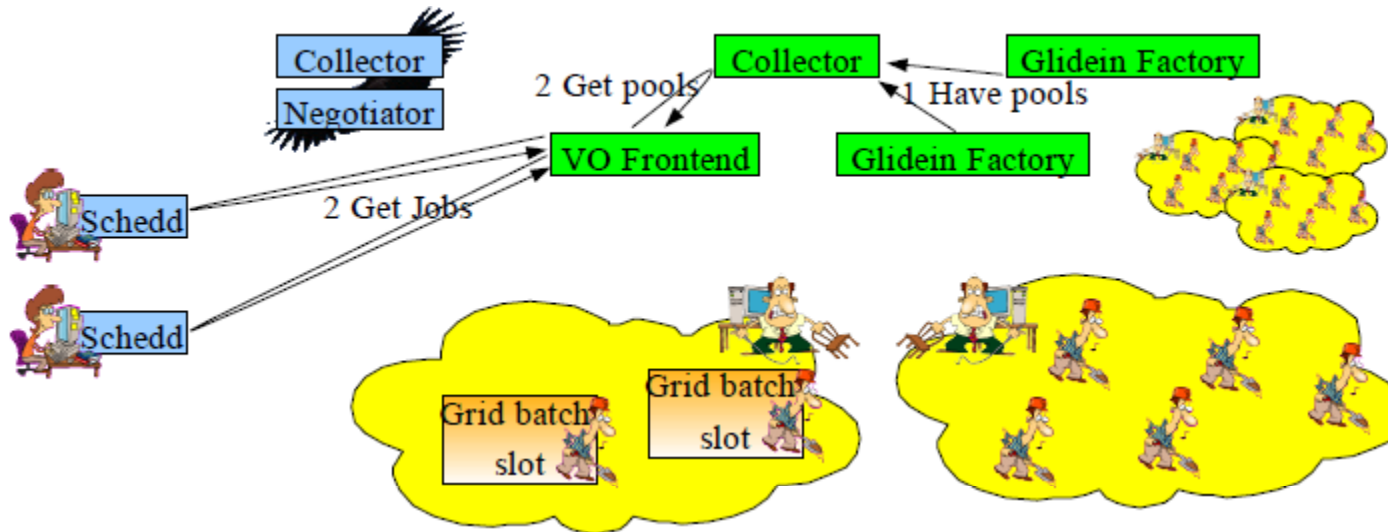
- > glideinWMS daemon information flow



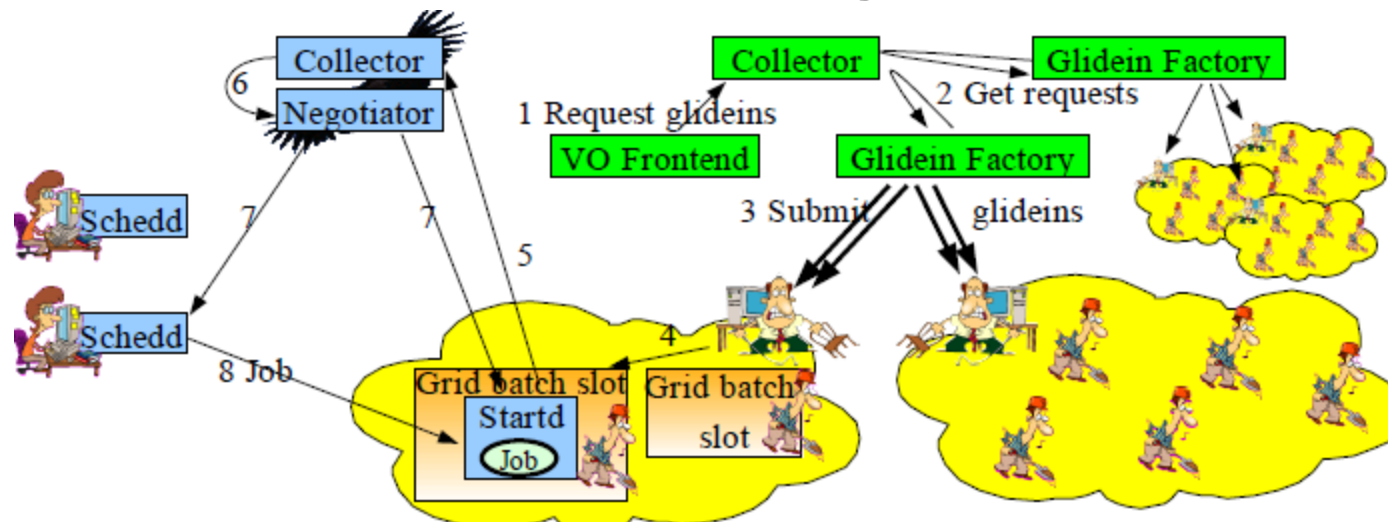
glideinWMS overview (2)



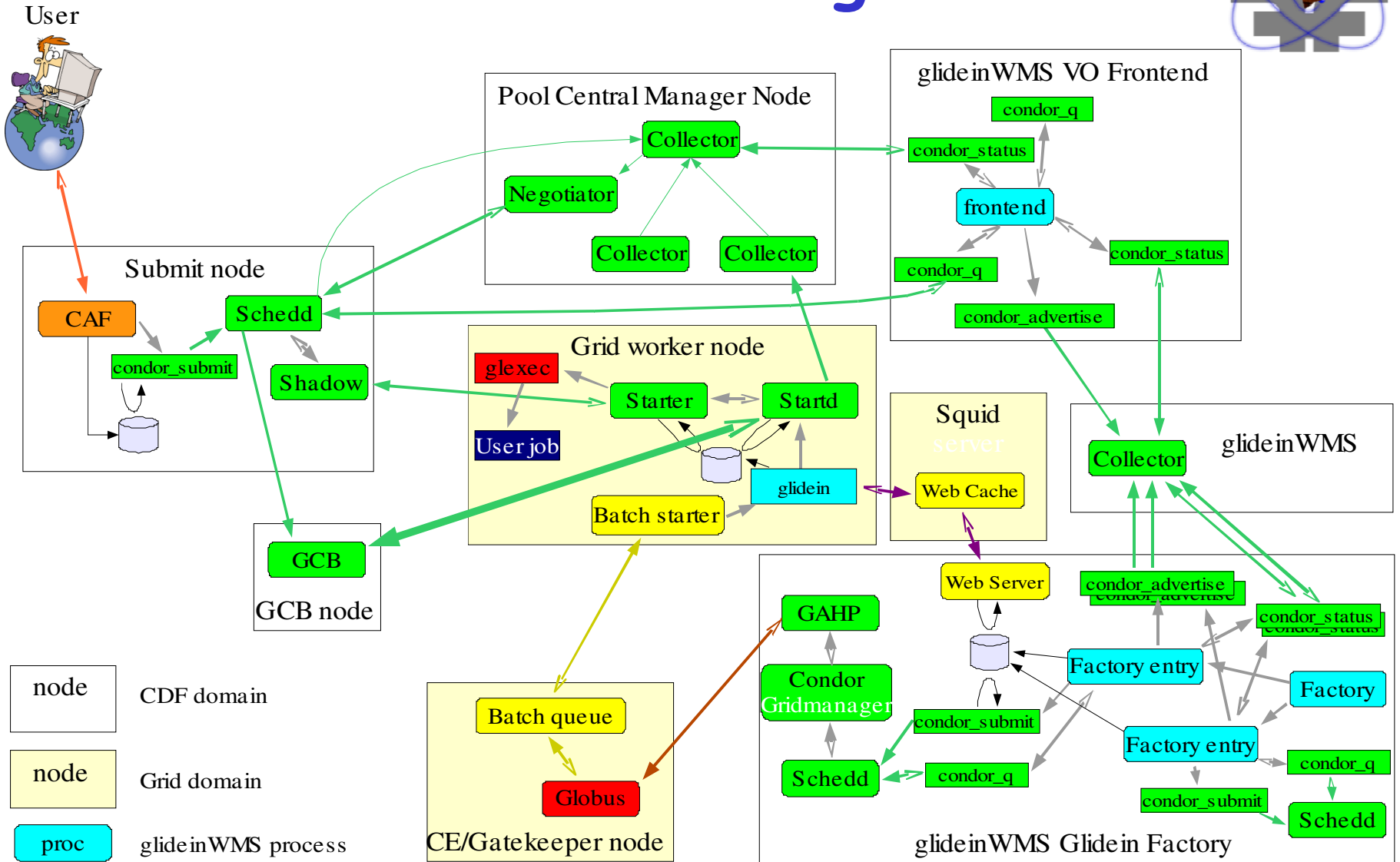
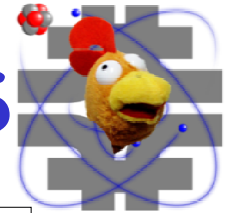
- > glideinWMS daemon information flow



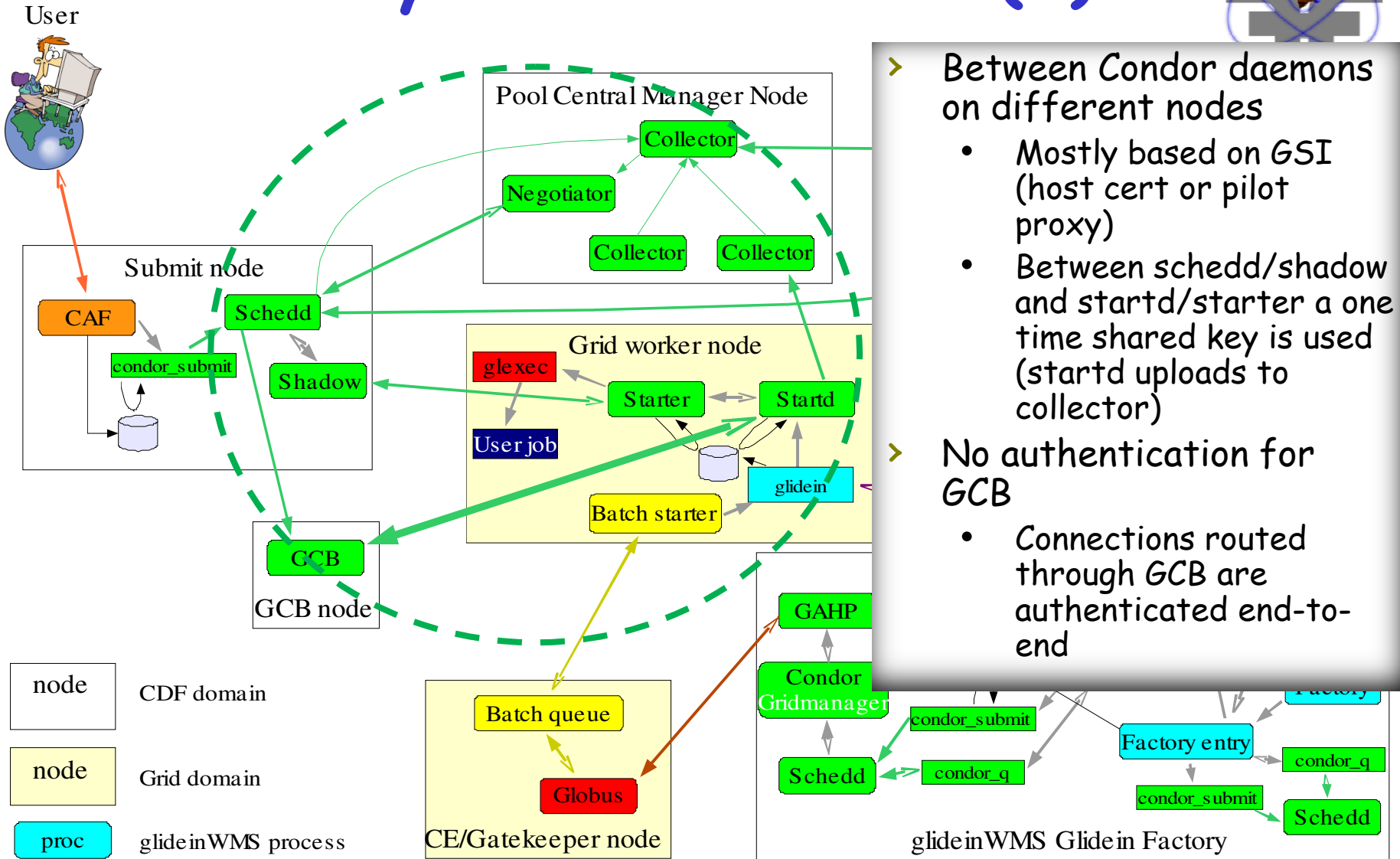
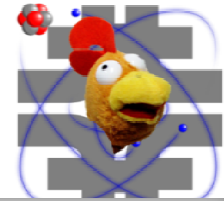
- > glideinWMS resource matching



Schematic view of CDF glideinWMS

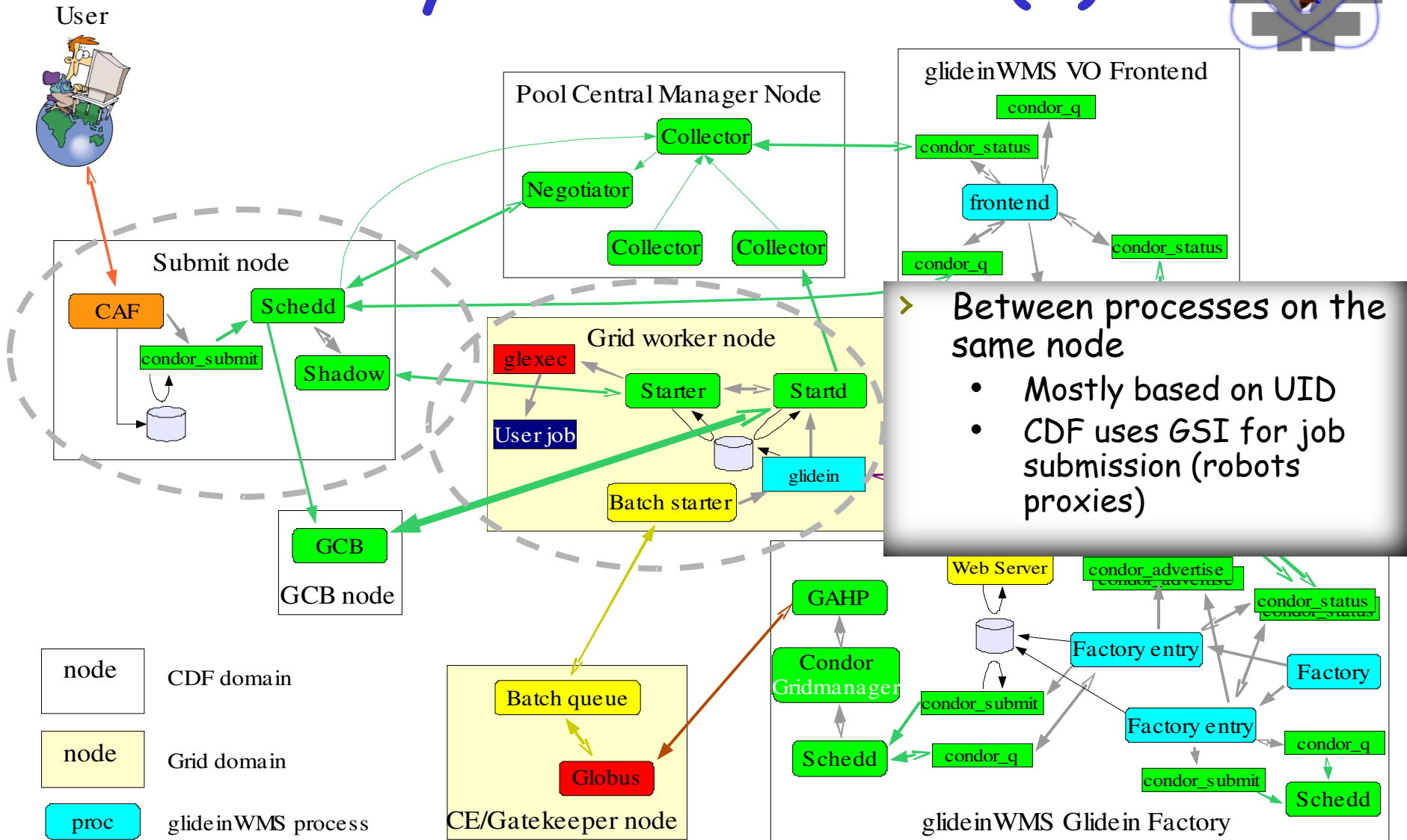
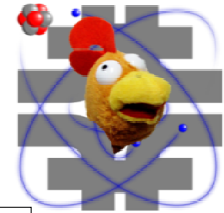


Security considerations (1)

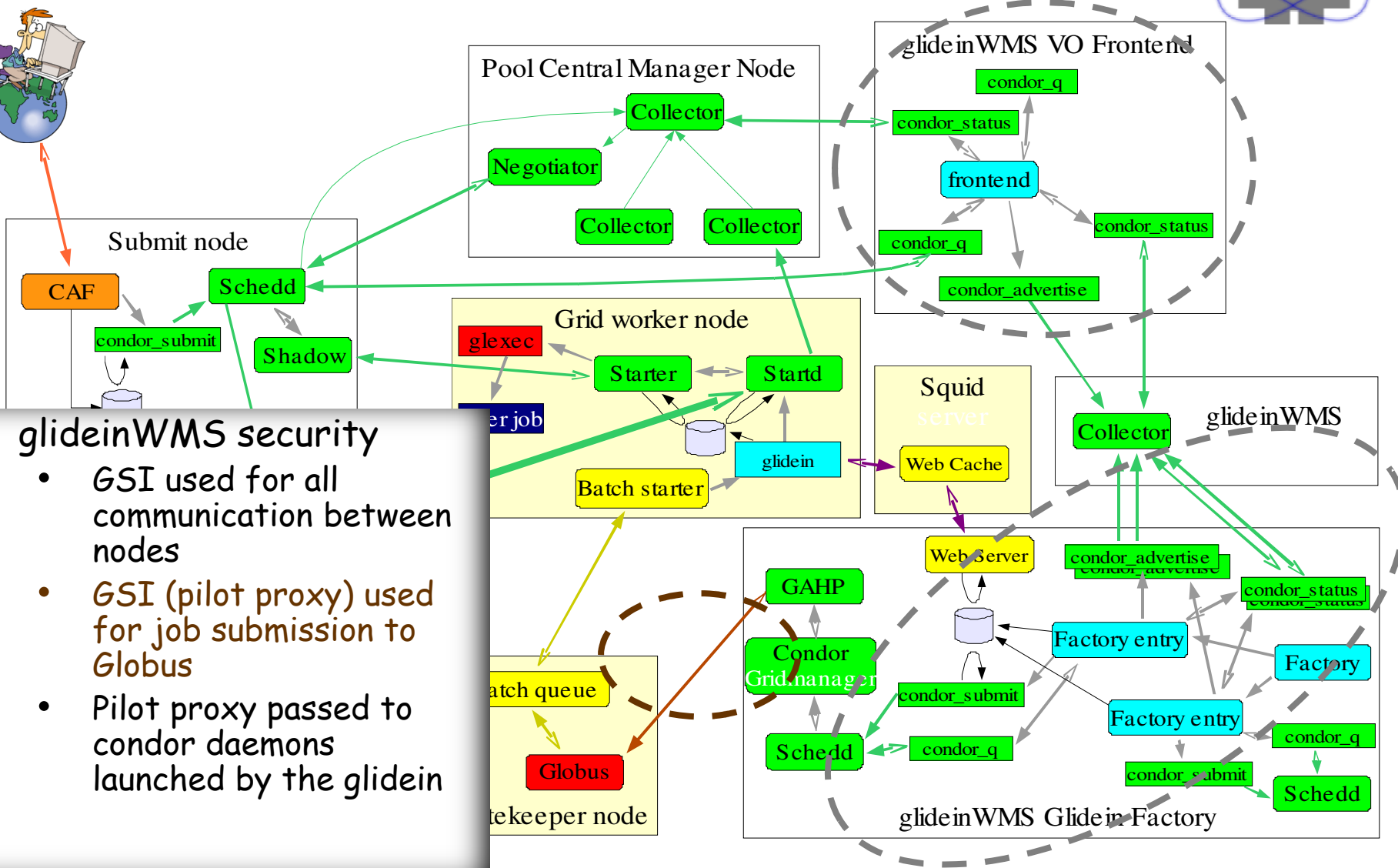
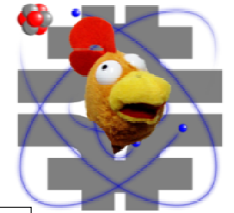


- > Between Condor daemons on different nodes
 - Mostly based on GSI (host cert or pilot proxy)
 - Between schedd/shadow and startd/starter a one time shared key is used (startd uploads to collector)
- > No authentication for GCB
 - Connections routed through GCB are authenticated end-to-end

Security considerations (2)

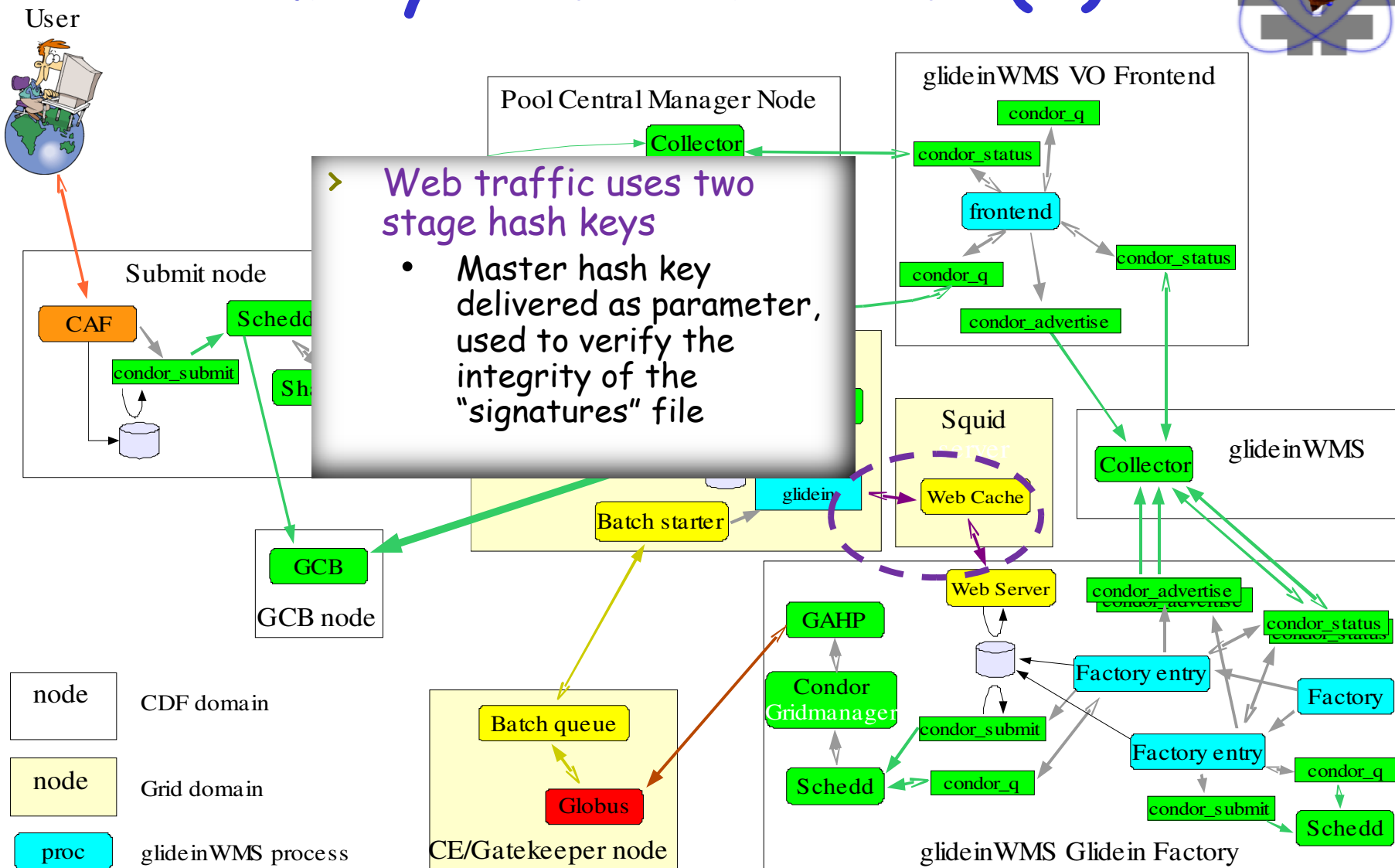
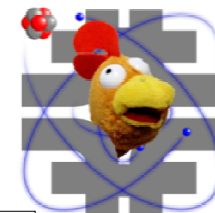


Security considerations (3)

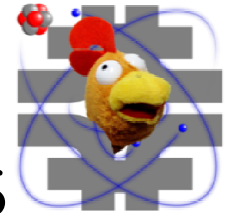


- > glideinWMS security
- GSI used for all communication between nodes
 - GSI (pilot proxy) used for job submission to Globus
 - Pilot proxy passed to condor daemons launched by the glidein

Security considerations (4)



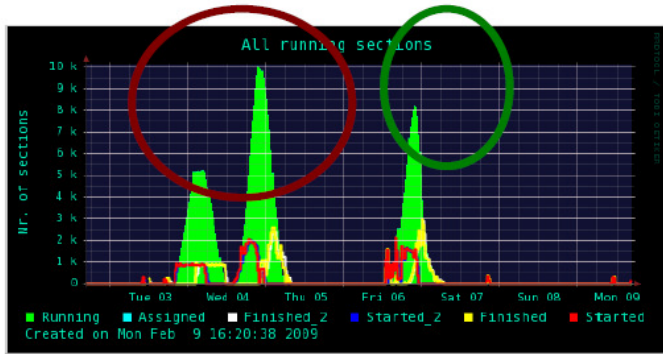
CDF large scale tests



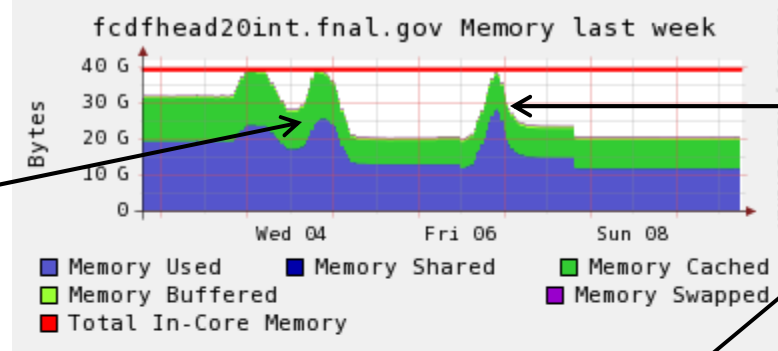
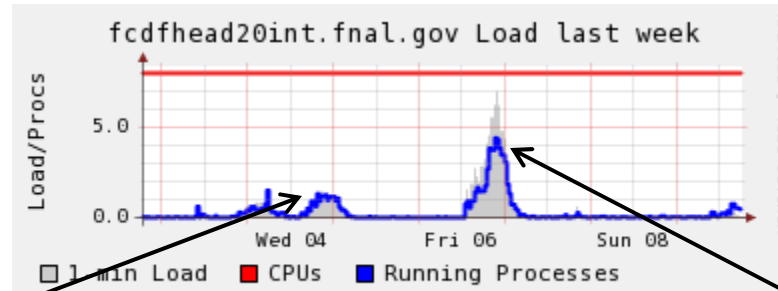
- > cpu/memory usage: glidekeeper vs. glideinWMS

glideinwms

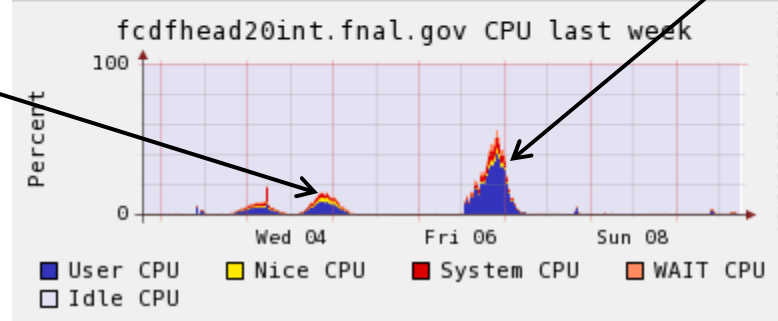
glidecaf



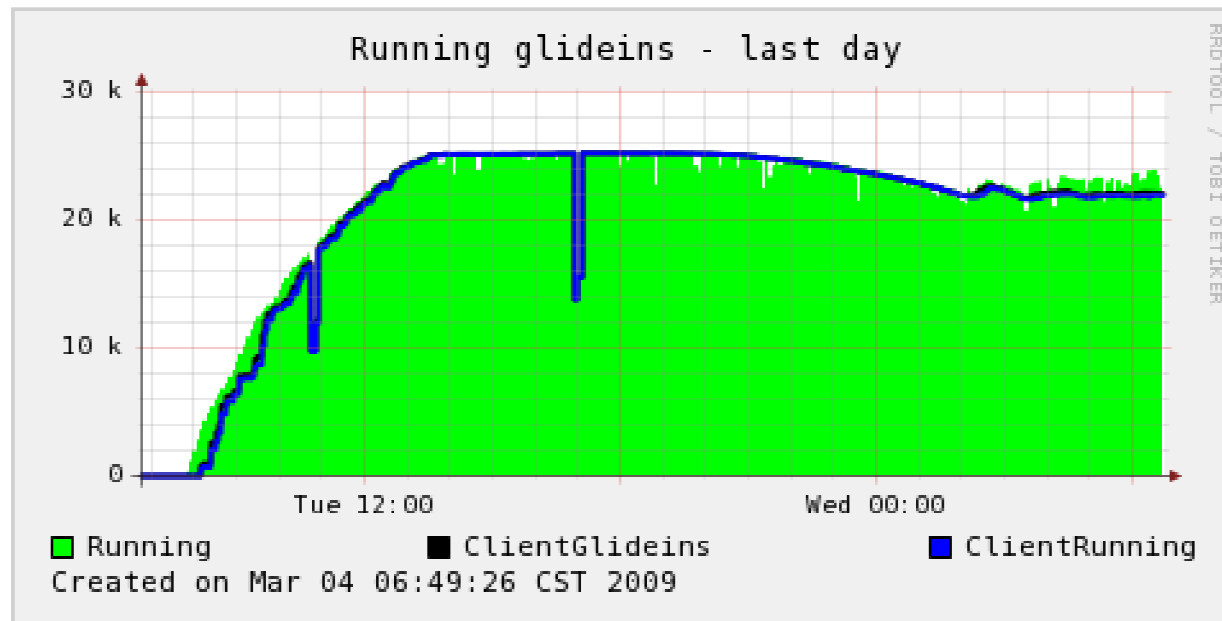
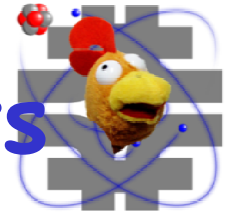
10k jobs running,
using *glideinWMS*



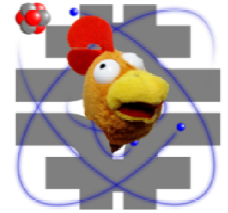
8k jobs running,
using *glidekeeper*



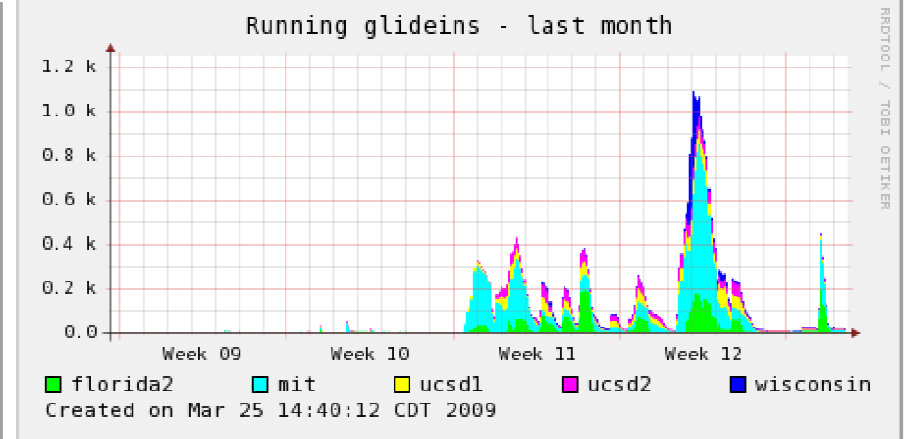
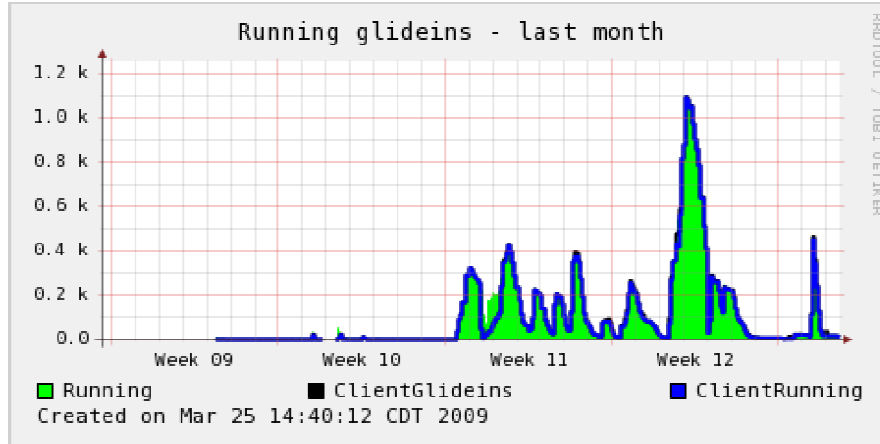
glideinWMS developers latest tests



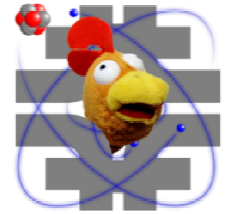
gWMS in production at CDF



- > NAMGRID cluster up almost a month
 - Using currently few OSG resources
 - Factory total over last month



- Plan on to add more resources soon



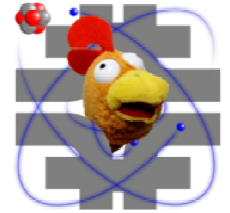
Conclusions

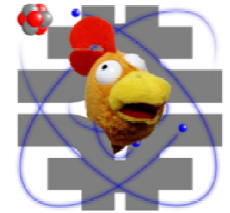
- > CDF has been successfully using Grid resources through glideins for the past 4 years
 - but we reached the scalability limits of the home grown software
- > CMS has developed a more scalable glidein solution (glideinWMS)
 - general purpose, so we can use it
 - very similar to the CDF glidekeeper since glideinWMS borrowed heavily from the CDF experience
- > CDF is migrating to glideinWMS
 - experience up to now very positive

Acknowledgments for CDF glideinWMS

Igor Sfiligoi, Doug Benjamin, Donatella Lucchesi

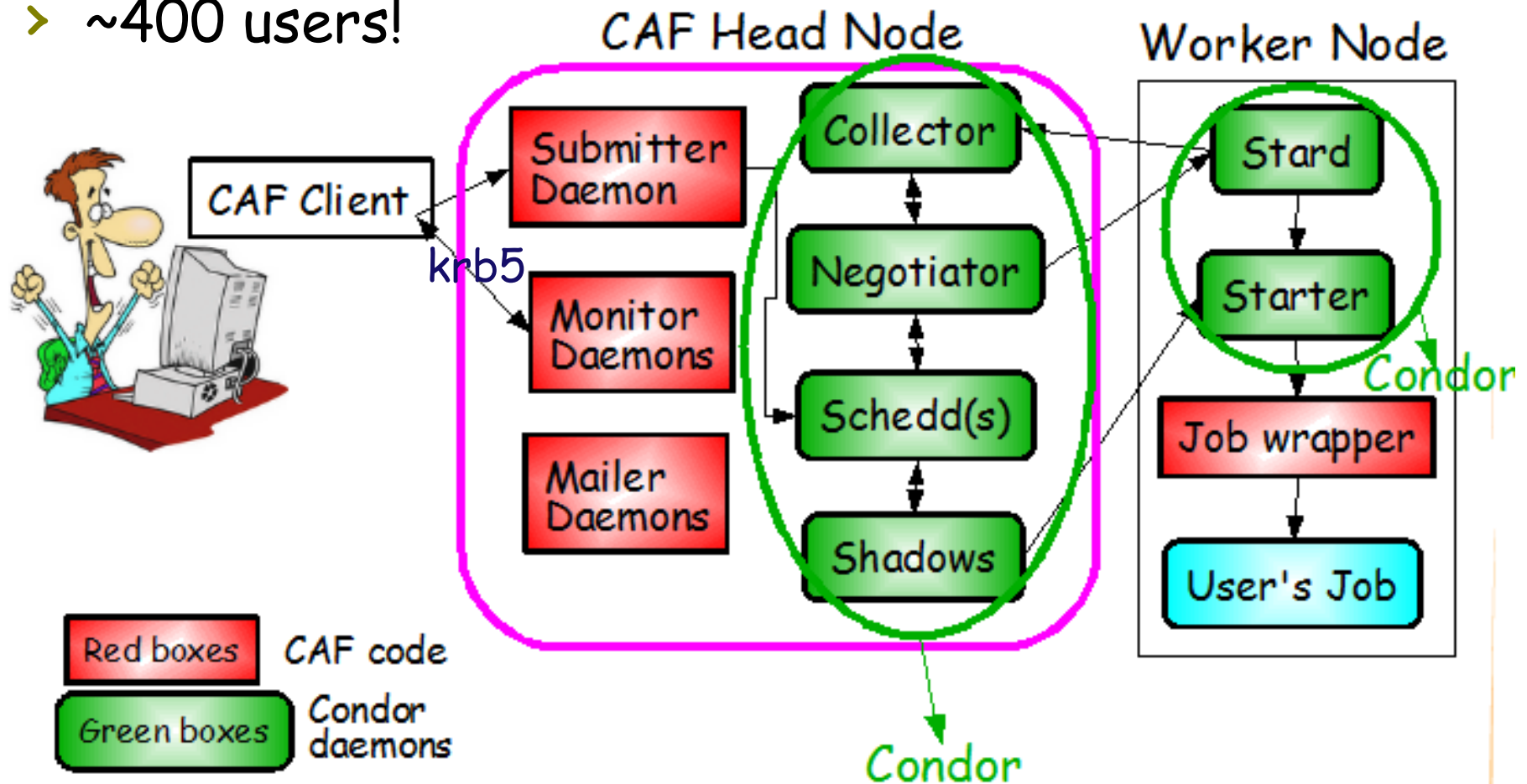
...BACKUP slides...



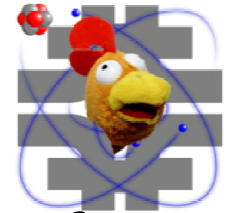


CAF middleware & Condor

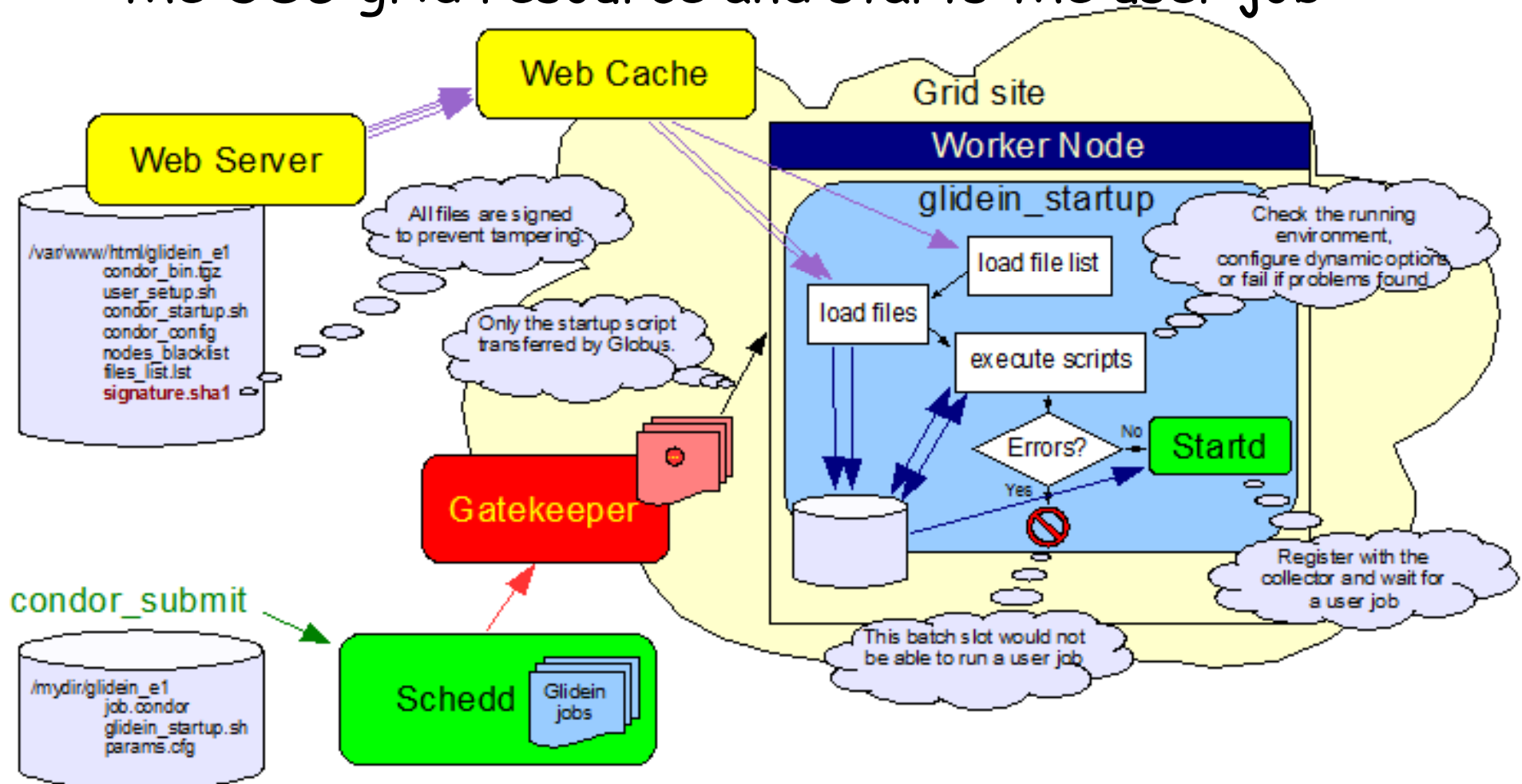
- > CDF runs DH and MC jobs, submission via CAF
- > Same infrastructure for both types of jobs
- > ~400 users!



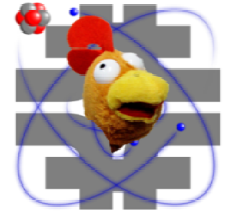
Glidein startup script overview



- > Glidein delivers job environment across the CE of the OSG grid resource and starts the user job

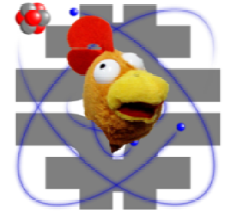


Security considerations (1)



- > Between Condor daemons on different nodes
 - Mostly based on GSI (host cert or pilot proxy)
 - Between schedd/shadow and startd/starter a one time shared key is used (startd uploads to collector)
 - No authentication for GCB
 - Connections routed through GCB are authenticated end-to-end
- > Between processes on the same node
 - Mostly based on UID
 - CDF uses GSI for job submission (robots proxies)

Security considerations (2)



- > glideinWMS security
 - GSI used for all communication between nodes
 - GSI (pilot proxy) used for job submission to Globus
 - Pilot proxy passed to condor daemons launched by the glidein
- > Web traffic uses two stage hash keys
 - Master hash key delivered as a parameter
 - Used to verify the integrity of the "signatures" file
 - All other files have a hash key in the "signatures" file