

HPC@Uni.lu: my experience

Venkata Satagopam
Bioinformatics core, LCSB, Uni.lu

Bioinformatics activities

- Meta{genomics, transcriptomics, proteomics}
 - generates large amounts of data in the scale of TB
 - map the reads on reference sequence and calculate gene/transcript/protein abundance in different conditions
- Sequence similarity
 - widely use BLAST, BLAT algorithms

Bioinformatics activities

- HMM models
 - group the related sequences into families,
 - run multiple sequence alignment (MSA) (t-coffee, muscle, clustalw)
 - run hmmbuild to build HMM models from MSA
 - hmmscan - to scan the models with novel sequences

- Parsing large files
 - extraction and transformation

OAR usage

- Previously used following batch queuing systems
 - CODINE (Computing in Distributed Networked Environments), also called Sun Grid Engine (SGE), Oracle Grid Engine (OGE)
 - LSF (Load Sharing Facility)
 - PBS Pro (Portable Batch System Professional), free versions are OpenPBS and TORQUE

OAR usage

- Previously used following batch queuing systems
 - CODINE (Computing in Distributed Networked Environments), also called Sun Grid Engine (SGE), Oracle Grid Engine (OGE)
 - LSF (Load Sharing Facility)
 - PBS Pro (Portable Batch System Professional), free versions are OpenPBS and TORQUE
- OAR is very similar to PBS

OAR usage

- Previously used following batch queuing systems
 - CODINE (Computing in Distributed Networked Environments), also called Sun Grid Engine (SGE), Oracle Grid Engine (OGE)
 - LSF (Load Sharing Facility)
 - PBS Pro (Portable Batch System Professional), free versions are OpenPBS and TORQUE
- OAR is very similar to PBS
- Usage
 - classical usage - oarsub, oarstat, oardel
 - `oarsub -l /enclosure=1/nodes=1/core=4,walltime=24:00:00 ./iso00065.blat_job.sh`
[ADMISSION RULE] Modify resource description with type and ibpool constraints
OAR_JOB_ID=2828649
 - `oarstat -u vsatagopam`
 - `oarstat -j <job_id>` (e.g.: `oarstat -j 2828649`)
 - `oardel <job_id>` (e.g.: `oardel 2828649`)

OAR usage

- Many small jobs

- should be bundled together and submit as one job

- job arrays

- GNU Parallel with serial job

- https://hpc.uni.lu/users/use_cases/#launch-serial-jobs-in-parallel*

Monika



HPC @ Uni.lu

Chaos, Gaia, Nyx and Granduc clusters

<https://hpc.uni.lu>

Get Updates:

By RSS

On Twitter

Home Systems For Users Live Status HPC School Blog/News About

Search...

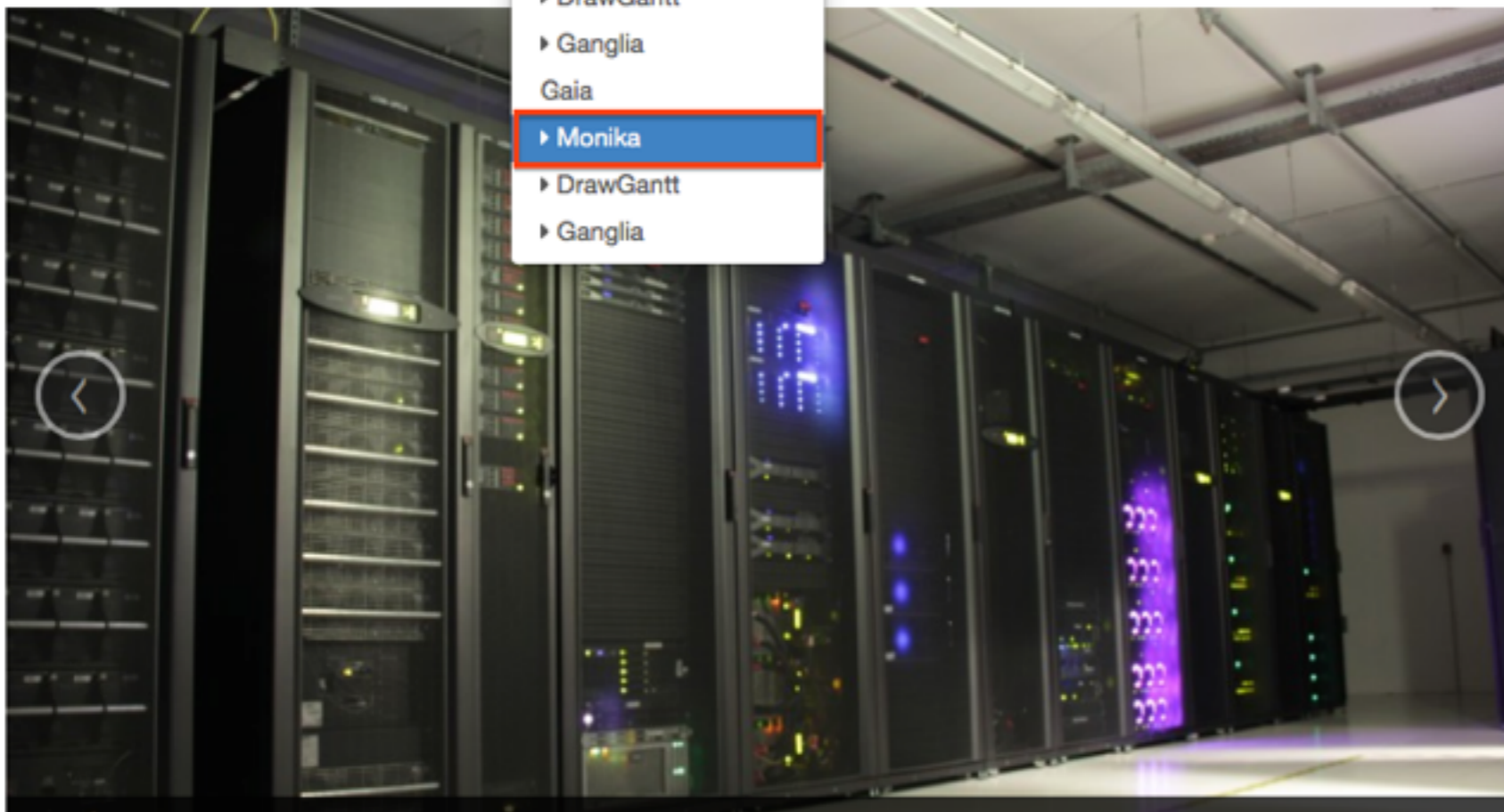
Welcome to the HPC @

This is the official website of HPC @ Uni.lu p
University of Luxembourg and the organizati

- Overview
- Monika
- DrawGantt
- Ganglia
- CDash
- Chaos
 - ▶ Monika
 - ▶ DrawGantt
 - ▶ Ganglia
- Gaia
 - ▶ Monika
 - ▶ DrawGantt
 - ▶ Ganglia

m !

Information about the computing clusters operated by the
out-computes will be the one that out-competes.
— The Council on Competitiveness



Recent Posts

- [HPCS'2014 article presenting the UL HPC platform](#)
- [UL HPC School 2014](#)
- [Understanding Git Subtree](#)
- [Create a sandboxed python/ruby environment](#)
- [XFS & inode64](#)

GitHub Repos

- tutorials
- dotfiles
- qualif
- launcher-scripts
- documents
- viridis-bootstrap
- ganglia_infiniband_module

Tweets

Follow



ULHPC @ULHPC

19 Apr

Save the date! The UL HPC School is planned on May 6th and 7th, 2014. More info on hpc.uni.lu/hpc-school/

Gaia nodes

default summary

	Free	Busy	Total
network_address	74	75	151
resource_id	968	1008	2000

Reservations:

gaia-1	2828366	2828366	2828366	2828366	2828366	2828366	gaia-2	2828371	2828371	2828371	2828371	2828371	2828371	2828371
	2828366	2828366	2828366	2828366	2828366	2828366		2828371	2828371	2828371	2828371	2828371	2828371	2828371
gaia-3	2828202	2828202	2828202	2828202	2828202	2828202	gaia-4	2828202	2828202	2828202	2828202	2828202	2828202	2828202
	2828202	2828202	2828202	2828202	2828202	2828202		2828202	2828202	2828202	2828202	2828202	2828202	2828202
gaia-5	2828015	2828015	2828015	2828015	2828015	2828015	gaia-6	2828194	2828194	2828194	2828194	2828194	2828194	2828194
	2828015	2828015	2828015	2828015	2828015	2828015		2828194	2828194	2828194	2828194	2828194	2828194	2828194
gaia-7	2828194	2828194	2828194	2828194	2828194	2828194	gaia-8	2828353	2828353	2828353	2828353	2828353	2828353	2828353
	2828194	2828194	2828194	2828194	2828194	2828194		2828353	2828353	2828353	2828353	2828353	2828353	2828353
gaia-9	2828354	2828354	2828354	2828354	2828354	2828354	gaia-10	2828202	2828202	2828202	2828202	2828202	2828202	2828202
	2828354	2828354	2828354	2828354	2828354	2828354		2828202	2828202	2828202	2828202	2828202	2828202	2828202
gaia-11	2828137	2828137	2828137	2828137	Free	Free	gaia-12	2828650	2828650	2828650	2828650	2828650	2828650	2828650
	2828339	2828294	2828295	2828296	2828349	Free		2828650	2828650	2828650	2828650	2828650	2828650	2828650
gaia-13	2828340	2828340	Free	Free	Free	Free	gaia-14	2828372	2828372	2828372	2828372	2828372	2828372	2828372
	Free	Free	2828264	2828265	2828266	2828341		2828372	2828372	2828372	2828372	2828372	2828372	2828372
gaia-15	2828440	2828440	2828440	2828440	2828440	2828440	gaia-16	2828367	2828367	2828367	2828367	2828367	2828367	2828367
	2828440	2828440	2828440	2828440	2828440	2828440		2828367	2828367	2828367	2828367	2828367	2828367	2828367
gaia-17	2828341	Free	Free	Free	Free	2828349	gaia-18	Free	Free	Free	Free	Free	Free	Free
	2828349	2828349	2828349	2828349	2828349	2828349		Free	Free	Free	Free	Free	Free	Free
gaia-19	2828251	2828251	2828251	2828251	2828251	2828251	gaia-20	2828251	2828251	2828251	2828251	2828251	2828251	2828251

Issues

- Scratch file system
 - lot of I/O issues
- Initially had difficulty with max no. of jobs and time limitation per job

- We need an API or other solution to submit jobs from remote node for following services
 - Galaxy
 - GenePattern

Conclusions

- Fine with current HPC setup given the available resources
- Only 2 personal keeping it up and running

Thank you



venkata.satagopam@uni.lu