

# A NERSC Primer for DES Data Analysis and Simulations or: How to meet your next DES publication deadline ...

DES Large Chicagoland Meeting, December 9, 2014

NERSC

40 YEARS at the FOREFRONT 1974-2014

Login  
Site Map | My NERSC | Share

search... >>

HOME ABOUT SCIENCE AT NERSC SYSTEMS FOR USERS NEWS & PUBLICATIONS R & D EVENTS LIVE STATUS NERSC@40

## National Energy Research Scientific Computing Center

1 > 2 > 3 > 4 > 5

**NOVEL ALGORITHMS HELP ILLUMINATE DARK ENERGY**  
Using NERSC supercomputers, Berkeley researchers developed novel computational tools to help the Dark Energy Survey identify and monitor Type 1a supernovae.  
» [Read More](#)

# Introduction

- **First question: Should I run at NERSC?**
  - ▶ If you run DES analysis tasks regularly, yes!
  - ▶ NERSC supports jobs of all sizes, from serial jobs to very large parallel jobs
  - ▶ If you have a large number of small jobs to run, you can bundle them up easily and run them as one large job (examples will follow)
  - ▶ Queuing times for jobs might seem annoying at first, but with some thought can be kept reasonably short (more on that to follow)
- **Second question: How do I get started?**
  - ▶ If you haven't done so yet, send an email to [habib@anl.gov](mailto:habib@anl.gov) and you will be added to the DES allocation at NERSC
  - ▶ Lots of information on NERSC webpages: <https://www.nersc.gov/users/getting-started/>



# DES Allocation at NERSC

- DES has this year an allocation of 4.5M core hours, requested more for next year
  - ▶ With your account you will get by default (controlled by Salman) 5% of time of that
  - ▶ If you need more, you can ask once you have used up your allocation
  - ▶ Salman (PI of DES allocation) can monitor usage
  - ▶ You therefore cannot “by mistake” use up all the DES allocation :)

Project Information	User Roles & Contact Info	User Status by Repo	MPP Usage & Quotas	HPSS Usage & Quotas	Transfer History	Project Access
<b>Project</b>	<b>Project Repos</b>	<b>Project Unix Groups</b>			<b>PI</b>	<b>Last Active</b>
DES	des desuser	des desuser			Salman Habib	2014
<i>Dark Energy Survey</i>						

Format: Read-only <--> Edit user percentages

NOTE: all hours displayed below are user hours, not repo hours.

des MPP Users, AY 2014 <--> Show users for prior AY

Login	Name	User Hrs Used	User Charged	Avg CF	% Used	% Allowed	User Balance	Repo	User Status	Base Repo?	Dflt Now?
aleksic	Aleksic, Jelena	454	454	1.0	0.0	5	237,046	Active	Y	Y	
amara	Amara, Adam	0	0		0	5	237,500	Active	Y	Y	
rarmst	Armstrong, Bob	0	0		0	5	237,500	Active	Y	Y	
kbarbary	Barbary, Kyle	0	0		0	5	237,500	Active	N	N	
battagli	Battaglia, Mattia	0	0		0	5	237,500	Active	Y	Y	
beckermr	Becker, Matthew	192,587	154,070	0.8	3	15	558,430	Active	N	N	



# Getting Started, File systems

<https://www.nersc.gov/users/data-and-file-systems/>

- **Two major machines: Hopper and Edison (Carver will retire next year), all machines share common file systems**
- **File systems:**
  - ▶ Your home directory: 40GB, backed up, good place to store your source codes, not a good place to run analysis or store data
  - ▶ Project space for DES: /project/projectdirs/des, accessible for everybody in the group, 40TB shared amongst all DES users, holds currently some DES data in /project/projectdirs/des/wl/desdata/DES, not backed up but will also not be deleted. DES can purchase more storage if needed
  - ▶ Scratch space: /scratch/scratchdirs/yourname (automatically added), 20TB, not backed up and will be purged after some time if the data is not touched, good for short-term use
  - ▶ HPSS: tape, to get there, type hsi, leads you to your tape partition, almost indefinite amount of space, not backed up but not purged, htar allows you to package small files and put them on tape



# Finding Software of Interest

<https://www.nersc.gov/users/software/>

- **NERSC uses modules to help you:**
  - ▶ module avail: shows you all the software packages available
  - ▶ module list: shows you what is loaded by default in your user environment
  - ▶ module load xxx (where xxx can be idl, python, any number of compilers, gsl, mathematica....) will load software package xxx which you can then use
  - ▶ module avail xxx lets you search for a specific software package

```
heittmann — ssh — 80x24
hopper05 h/heittmann> module avail python

----- /usr/common/usg/Modules/modulefiles -----
python/2.7.1          python_libs/2.7.5
python/2.7.3          python_base/2.7.3
python/2.7.5(default) python_base/2.7.5(default)
python/2.7.8          python_base/2.7.8
python_libs/2.7.3
hopper05 h/heittmann> module avail numpy

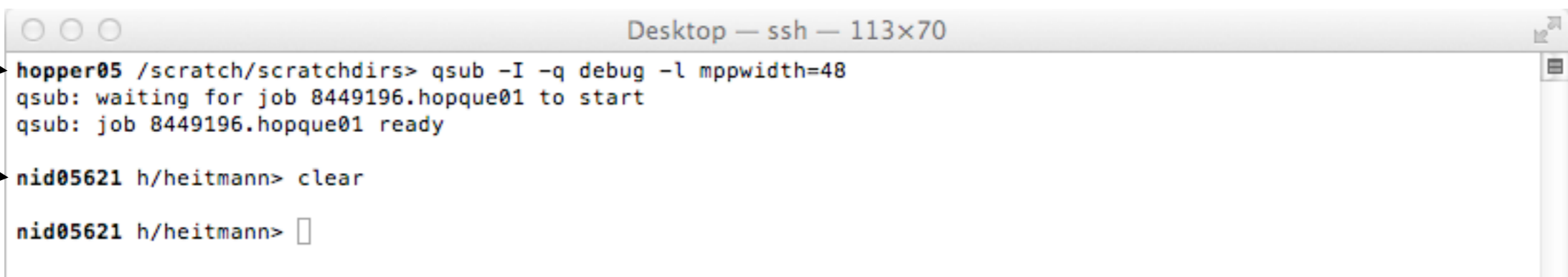
----- /usr/common/usg/Modules/modulefiles -----
numpy/1.6.1          numpy/1.7.1(default)
numpy/1.6.2          numpy/1.8.2
hopper05 h/heittmann> 
```



# Queues and Job Submission

<https://www.nersc.gov/users/queues/>

- When you log into Hopper or Edison, you will be on a front-end node, here you can compile your code
- DO NOT run your jobs on the front-end, this would slow down everybody else
- Instead, login in via an interactive queue or submit jobs to the back-end nodes
  - ▶ Interactive job: `qsub -I -q debug -l mppwidth=48`
  - ▶ On hopper, this gets you 48 cores (or 2 nodes) for 30 minutes
  - ▶ To run your job: `aprun -n 48 ./a.out`
  - ▶ Or: submit a job via `qsub run.sh` (example for submission script next)



```
Desktop — ssh — 113x70
hopper05 /scratch/scratchdirs> qsub -I -q debug -l mppwidth=48
qsub: waiting for job 8449196.hopque01 to start
qsub: job 8449196.hopque01 ready
nid05621 h/heitmann> clear
nid05621 h/heitmann> █
```

The image shows a terminal window titled "Desktop — ssh — 113x70". The first part of the terminal shows a user on a hopper node submitting an interactive job with the command `qsub -I -q debug -l mppwidth=48`. The output shows the job is ready. The second part shows a user on a nid node logging in and clearing the screen with `clear`.



# Example Submission Script

Queue, depends on number of cores and length of the job, system will pick correct one automatically

Number of cores, divisible by 24 (number of cores per node)

Sends you an email when job begins/ends/aborts

Guess for run time

```
hopper05 Coyote_small_new/M002> more run_gad.bat
#PBS -q regular
#PBS -l mppwidth=264
#PBS -o gadget.log
#PBS -e gadget.err
#PBS -A cosmosim
#PBS -m bea
#PBS -l walltime=15:00:00
#PBS -N gadget
#PBS -j oe
```

Log file  
Error file

Repo used for run, if you are only in DES, automatically set

Job name

Puts you into your work directory (without it, job will try running in your home directory)

```
cd $PBS_O_WORKDIR
aprun -n 256 ./Gadget2_xxs M002.param
```

Number of cores

Executable

Input

This line can be more complicated, e.g. include threads



# End of a Log file

```
-  
Peano-Hilbert done.  
Start force computation...  
Starting periodic PM calculation.  
done PM.  
Tree construction.  
Tree construction done.  
Begin tree force.  
tree is done.  
force computation done.  
type=1 dmean=0.449815 asmth=0.266309 minmass=0.924878 a=1 sqrt(<p^2>)=705.788 dlogmax=0.00943305  
displacement time constraint: 0.00943305 (0.025)  
  
writing snapshot file...  
done with snapshot.  
Application 40296478 resources: utime ~11522683s, stime ~190693s, Rss ~579516, inblocks ~747127570, outblocks ~2\  
65117292
```

```
+ -----  
+ Job name: gadget  
+ Job Id: 8422229.hopque01  
+ System: hopper  
+ Queued Time: Tue Nov 25 21:03:22 2014  
+ Start Time: Sun Nov 30 05:04:36 2014  
+ Completion Time: Sun Nov 30 17:47:27 2014  
+ User: heitmann  
+ MOM Host: nid03934  
+ Queue: reg_small  
+ Req. Resources: mppnodect=11,mppnppn=24,mppwidth=264,walltime=20:00:00  
+ Used Resources: cput=00:00:06,mem=5968kb,vmem=58364kb,walltime=12:42:52  
+ Acct String: cosmosim  
+ PBS_0_WORKDIR: /global/project/projectdirs/hacc/Coyote_small_new/M001  
+ Submit Args: run_gad.bat  
+ -----
```





## Other useful commands

- `qdel job-id`: delete your job
- `qstat -f job-id`: status of the job
- `qhold job-id`: puts job in the queue on hold (in case you want to change something but still have to think, you don't have to delete the job and lose your position in the queue)
- `qalter -t`: allows you to change the time for which you have submitted the job while the job is waiting in the queue
- `showq`: shows all the jobs (`qstat` by itself works too)
- `showq | grep heitmann`: shows my jobs
- `showstate`: shows which jobs are running on which racks, if nodes are down etc.



# CLASSIC QUEUE LOOK

## Queue Wait Times

**Display**

Job ID   
  Job Name   
  User   
  Queue   
  Repo  
 Nodes   
  Cores   
  Wall Time Requested   
  Time Used   
  Time Remaining  
 Submit Time   
  Start Time

User (Use "all" for all users)  Repo  Submit Queue  Execution Queue

**Display Machine(s)**

Hopper   
  Edison   
  Carver   
  Dirac   
  Planck   
  Matgen   
  Genepool

Hide Running Jobs   
  Hide Queued Jobs   
  Hide Ineligible Jobs

Data updated every 5 minutes. Display updated: Wed, 03 Dec 2014 22:24:01 -0800  
 Data updated: Hopper: 2014-12-03 22:22:39; Edison: 2014-08-19 08:21:53; Carver: 2014-12-03 22:23:51; Genepool:

Query limited to 2000 results. Refine your search above if needed.

Column one (#) signifies the row in the table below. The value in the "Seq" column is the order in which jobs were scheduled to start (per machine) at the time the snapshot was taken.

Total nodes in use: Hopper 6,375, Edison 5,521, Carver 1,381, Planck 4, Planck 132, Genepool 530  
 Total nodes free: Hopper: 36; Edison: 29; Carver: 245; Planck: 0; Genepool: 0;

Running (R)    Queued (Q): Eligible to be scheduled    System Hold (HS)    User Hold (HU)    Batch Hold (HB)    Blocked (B): Ineligible to be scheduled    Not Queued (NQ): Ineligible to be scheduled    Complete (C): Entry will be removed from queue display

Show  entries Search:

#	Host	ID	Name	ST	User	Queue	Nds	Time Left	Time Req.	Submit Time	Seq
0	Hopper	8442110	gadget	Q	heitmann	reg_small	11	15:00:00	15:00:00	12-02 21:02:33	1157
1	Hopper	8442115	gadget	Q	heitmann	reg_small	11	15:00:00	15:00:00	12-02 21:05:34	1158
2	Hopper	8442125	gadget	Q	heitmann	reg_small	11	15:00:00	15:00:00	12-02 21:15:08	1162

Showing 1 to 3 of 3 entries

First Previous 1 Next Last

**Host Status**

Edison **Up**  
 Hopper **Up**  
 Carver **Up**  
 PDSF **Up**  
 Genepool **Up**  
 HPSS **Up**

**Queue Policies**

[Hopper](#)  
[Edison](#)  
[Carver](#)



Current data ends at 11:00 p.m. Pacific yesterday. Users may select only (1) all jobs, (2) their jobs, or (3) jobs for their repository.

Generate Report Reset to Defaults

Report Period: Dec-02-2014 00:00:00 to Dec-02-2014 23:59:59

Found 18,866 jobs.

Print this page

Wait Time in Hours

This "heat map" shows the queue wait times for jobs as a function of Nodes Requested and Hours Requested for Machine all, Queue: all. The columns represent the number of hours requested: "3" is in the range 3:00:00 to 3:59:59.

Heatmap table with columns: Nodes, Hours Requested (<1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 48+). Rows represent node ranges from 1 to 6144-9531.



## Submission of Multiple Jobs

<https://www.nersc.gov/users/computational-systems/hopper/running-jobs/example-batch-scripts/>

- **Several options:**

- ▶ TaskFarmer: software that helps to launch a number  $t$  of independent small jobs on  $n$  cores. Within the submission script use `tf -t 100 -n 2 ./task.sh` instead of `aprun -n 2 ./a.out`, example for `task.sh` on NERSC webpage (small script)
- ▶ CCM (Cluster Compatable Mode), include module load `ccm` in submission script and `ccmrun ./runtask.sh` where `runtask.sh` holds a list of jobs (example on NERSC webpage)
- ▶ Python: `mpi4py`, again explained on NERSC webpage
- ▶ Using MPI code to bundle up calls

Vinu's DEMO



# Moving Large Data Sets to and from NERSC

- Data transfer nodes dedicated to data transfer: dtn01.nersc.gov ... dtn04.nersc.gov
- If possible use Globus Online for data transfers

The screenshot shows the Globus Online web interface for managing data transfers. The browser address bar displays `https://www.globus.org/xfer/StartTransfer#` and the search bar contains `nersc data transfer nodes`. The main navigation bar includes `Manage Data`, `Groups`, `Support`, and `heitmann`. Below the navigation bar, there are links for `Transfer Files`, `Activity`, `Manage Endpoints`, and `Dashboard`. The `Transfer Files` section is active, showing two endpoint configurations side-by-side.

The left endpoint configuration is for `nersc#dtn` with the path `/project/projectdirs/des/`. Below the path input, a list of folders is displayed: `heitmann`, `jkwan`, `tcp`, `test`, `vinu`, `wl`, and `www`. An arrow points from the text `DES project directory` to the `Path` input field.

The right endpoint configuration is for `bencouver#datastar` with the path `/media/star2/`. Below the path input, a list of folders is displayed: `apope`, `benjamin`, `galacticus`, `heitmann`, `kovacs`, `lost+found`, `nanli`, `phosim`, `rbiswas`, and `simulations`. An arrow points from the text `Local workstation at Argonne` to the `Endpoint` input field.

## Useful Links

- Getting started: <http://www.nersc.gov/users/getting-started/>
- Queues: <http://www.nersc.gov/users/queues/>
- Python tools: <http://www.nersc.gov/users/software/development-tools/python-tools/>
- Serial jobs: <https://www.nersc.gov/users/computational-systems/edison/running-jobs/serial-queue/>
- Example batch scripts: <https://www.nersc.gov/users/computational-systems/hopper/running-jobs/example-batch-scripts/> (including examples to package up several jobs into one)
- How to use modules to manage your software environment: <https://www.nersc.gov/users/software/nersc-user-environment/modules/>
- Links to the different supercomputers: <https://www.nersc.gov/users/computational-systems/>



# Installing Software without Root

<https://theodorekisner.com/software/hpcports/>

- HPCPorts is a tool to manage software (I have not used it myself)
- For DES NERSC, a specific module has been created, to load it, follow instructions below (again, I have not used these so can't comment on them much)

```
source /global/project/projectdirs/des/w1/setup/setup.sh
```

```
source /project/projectdirs/cmb/modules/carver/hpcports.sh
```

```
module load python
```

```
module load des-nersc
```

