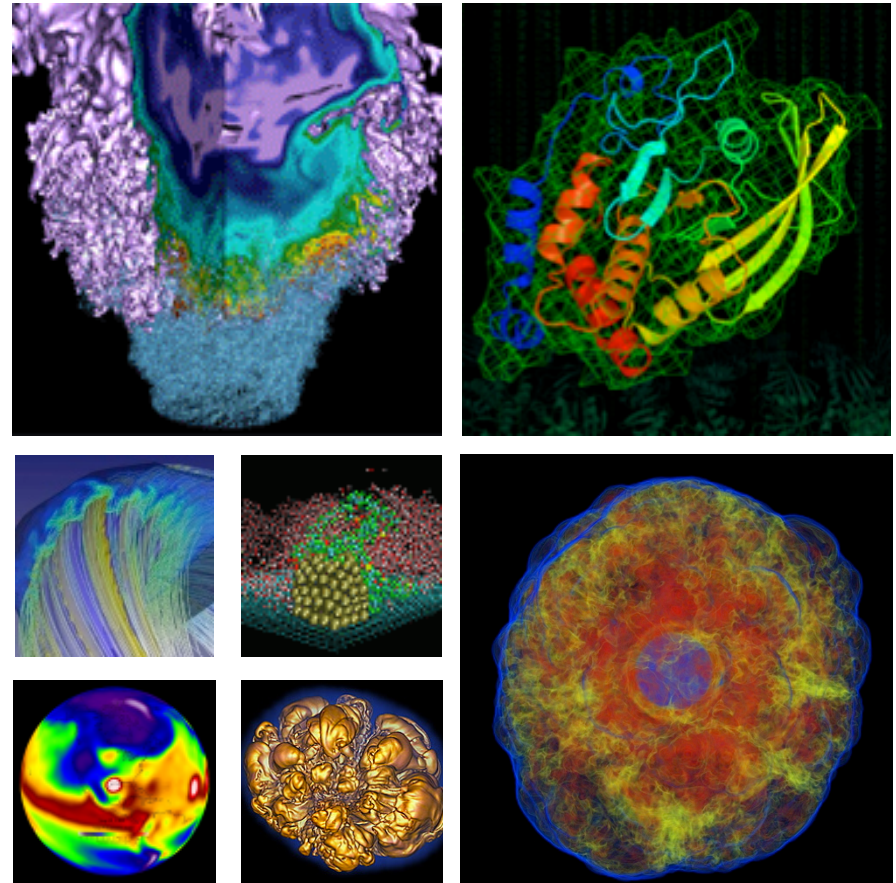


Job Analytics



Richard Gerber
NERSC User Services

November 8, 2012

The NERSC Ecosystem is Highly Complex



Intricate, first-of-a-kind systems and software

Unexplored scales

Innovative and experimental applications and workflows

6,000 scientists

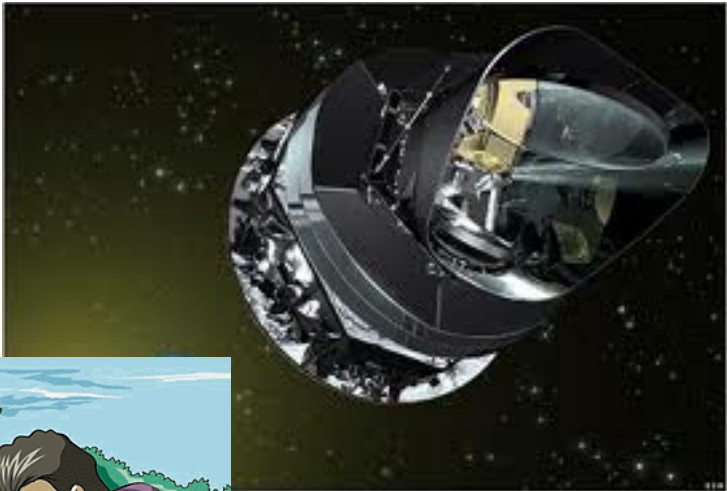
600 cutting edge codes



Observing the Ecosystem



- Through observation we can learn a lot about how scientists use our systems and how our systems react
- Scientists can likewise learn a lot about their applications



This is mostly an observational exercise so as to not perturb the systems.



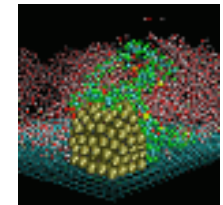
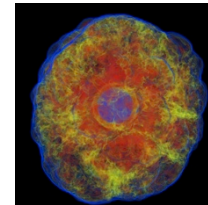
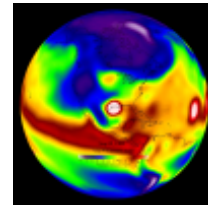
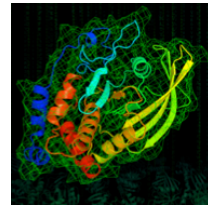
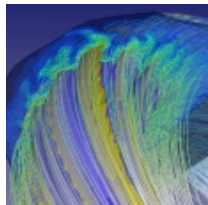
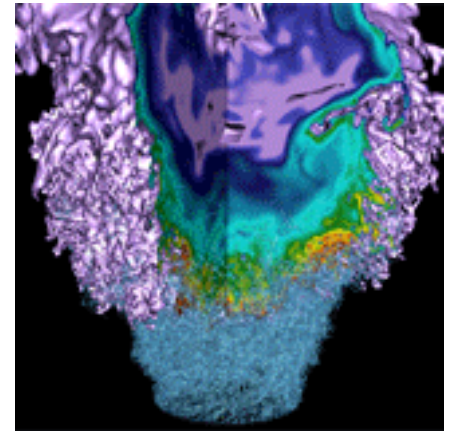
- **Data Collection**

- The interesting data is often not readily accessible or even measured: systems are not designed with this in mind
- NERSC currently collects available data from disparate sources
- NERSC is working with vendors to provide more data, more easily, accurately, and consistently

- **Data Analysis and Presentation**

- NERSC's goal is to make the interesting data and derived metrics readily available
- Web-based queries, reports, and analyses are the preferred method

Overview of Current Efforts



Job Queue Information



Users can estimate their jobs' progress through the queues.

NERSC can monitor the desired operation of the batch system.

Jobs that are “stuck” or submitted with incorrect specifications can be identified.

[https://www.nersc.gov/users + /live-status/global-queue-look/job-information/usage/summary/](https://www.nersc.gov/users/live-status/global-queue-look/job-information/usage/summary/)

The screenshot shows the NERSC Queue Status web application. The main heading is "GLOBAL QUEUE LOOK". Below this, there are several filter sections: "Display" (with checkboxes for Job ID, Job Name, User, Queue, Repo, Nodes, Cores, Wall Time Requested, Time Used, Time Remaining, Submit Time, Start Time), "Display Machine(s)" (with checkboxes for Hopper, Carver, Drac, Planck, Genepool), and "Display" (with checkboxes for Hide Running Jobs, Hide Queued Jobs, Hide Ineligible Jobs). There are also buttons for "Update Display", "Reset Form", and "Reload Default Display".

Below the filters, there is a table of scheduled outages and a section for "Query limited to 2000 results. Refine your search above if needed." The table below shows the following data:

#	Host	ID	Name	ST	User	Queue	Repo	Nds	Time Left	Seq
0	Hopper	2413419	test_ser...	R	dks	reg_small	mp27	18	42:06:54	
1	Hopper	2413425	test_ser...	R	dks	reg_small	mp27	24	42:06:10	
2	Hopper	2413591	test_ser...	R	dks	reg_small	mp27	12	45:08:03	
3	Hopper	2413596	test_ser...	R	dks	reg_small	mp27	12	45:21:56	
4	Hopper	2415503	test_ser...	Q	dks	reg_small	mp27	12		148
5	Hopper	2421592	test_ser...	Q	dks	reg_small	mp27	12		850

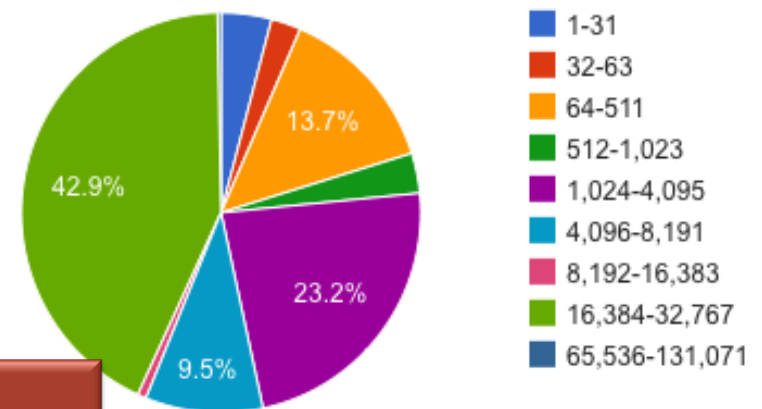
Workload Characterization



Important for procuring systems that work well with the NERSC workload

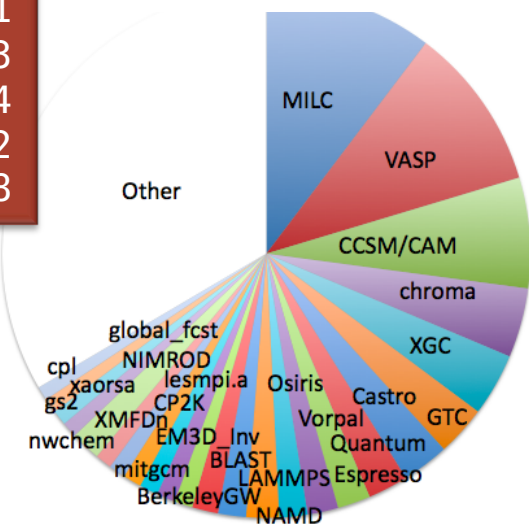
We can identify jobs and/or codes that will benefit from consultations with NERSC User Services

Raw Hours (in Millions) by Cores Used



Threads per task	Threads	Count	Percent
	1	6191	81.11
	2	10	0.13
	3	3	0.04
	6	1414	18.52
	24	14	0.18

<https://www.nersc.gov/users/job-information/usage/summary/job-information/>



System performance monitoring



NERSC regularly runs suites of benchmark codes on its systems.

Deviations from past performance often indicate performance or functional problems that are not uncovered by unit tests of individual system components.

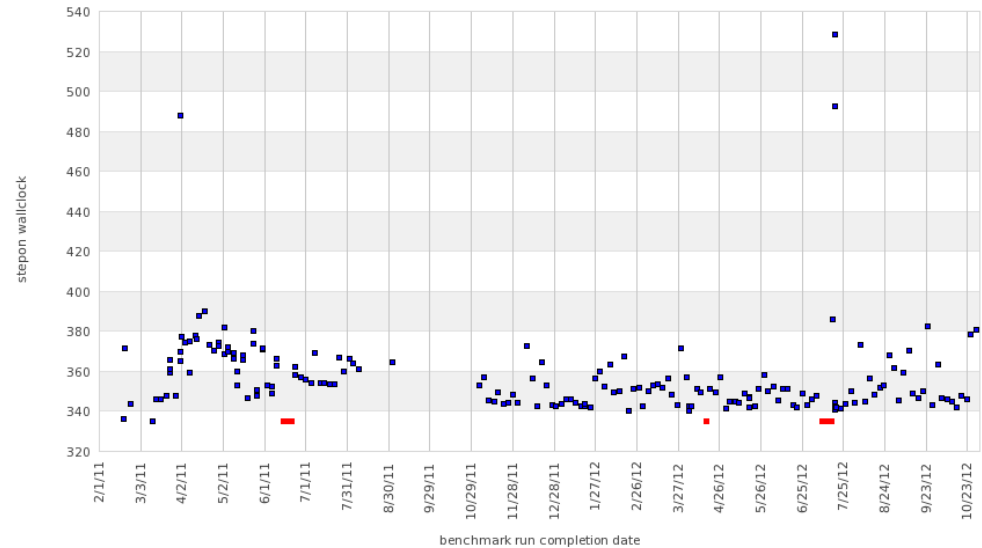
HOPPER BENCHMARK MONITORING

Benchmark Results

Select Benchmark

CAM

Displaying results for CAM:



<https://www.nersc.gov/users/job-information/hopper-benchmark-monitoring/>

Accounting and Repo Usage



NERSC charges usage against banks of computer time allocated by DOE.

Charge for wallclock*nodes (sometimes a queue factor)

May want to charge for memory, disk, power

<https://nim.nersc.gov>

<https://www.nersc.gov/users/job-information/usage/summary/>

The screenshot shows the NIM (NERSC Information Management) system interface. At the top, there's a navigation bar with 'NERSC' and 'NIM' logos, and buttons for 'NIM Home', 'My Stuff', and 'Search'. Below this, a welcome message states: 'Welcome to the NERSC Information Management (NIM) system. Use this interface to find information about NERSC users and repositories. Authorized managers can also modify information and create or review allocation requests. Staff-only options are in gold.' It also shows the user's last login on 11/07/2012 13:54:32 and that their password will expire on 02/19/2013. A 'NERSC System Status' section shows a green indicator for 'Available' and a table of system components: Carver, Dirac, Euclid, Hopper, HPSS, NGF, and PDSF. Below this is a navigation menu with tabs for 'Account Usage', 'Logins by Host', 'Unix Groups', 'Roles', 'Contact Info', 'Grid Certificates', 'Summary', 'Logins List', and 'Aliases'. The 'Account Usage Summary' table is displayed, showing details for user Richard Gerber across different resource categories.

Resrc	Repo	User_id	Login	Login Type	User Hrs Used	User Charged	Avg CF	% Used	% Allowed	User Balance	Last Charged On	
MPP	mpccc	12721	ragerber	Authorized	504	504	1.0	0.0	15	1,199,496	31-OCT-12	Y
MPP	m1266	12721	ragerber	Authorized	0	0		0	10	12,500		N

Resrc	Repo	User_id	Login	Login Type	User Charged	SRU Files	SRU Space	SRU I/O	Proj %	% Used	% Allowed	
HPSS	m1266	12721	ragerber	Authorized	0	0	0	0	0	0	100	
HPSS	mpccc	12721	ragerber	Authorized	9,182	515	8,603	65	100	0.5	10	1

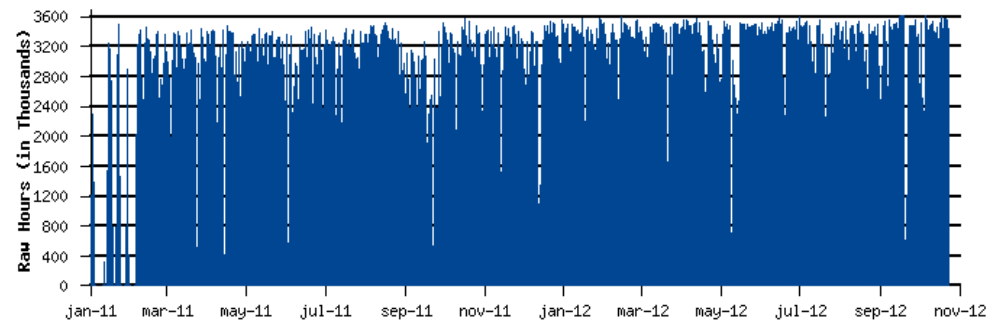
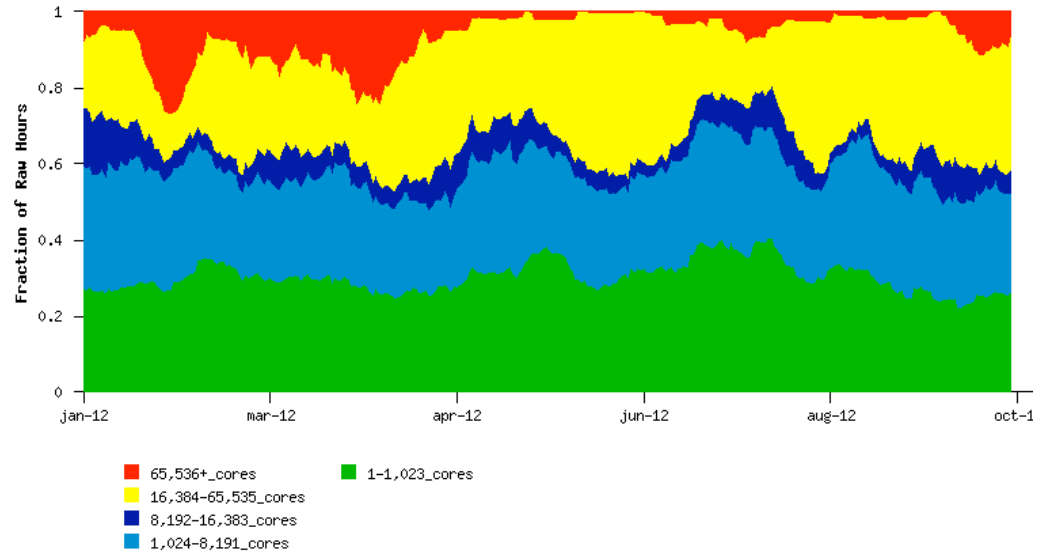
Queue Config Optimization and Metrics



NERSC uses trend analysis of queue wait times, job run sizes, queue utilization to determine whether queue configurations can be improved.

Metrics are reported to DOE and NERSC's users, both formally and informally.

<https://www.nersc.gov/users/job-information/>



Job Execution Metrics for Debug/Opt



Maximum memory used and usage during the run.

I/O statistics are valuable for optimizing data access.

MPI statistics can locate communication inefficiencies and/or errors

Hardware chip counters for CPU optimization info.

Job placement

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

/live-status/global-queue-look/
/software/debugging-and-profiling/ipm/

NERSC Completed Batch Jobs Listing

https://www.nersc.gov/users/job-informatio

Richard Gerber (rgerber) | Logout
My NERSC | Site Map | Share

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

COMPLETED JOBS

Select a time period

Show jobs that completed after: Oct 31 2012 @ 0 : 0 :00 Pacific Time
and completed on or before: Nov 7 2012 @ 16 : 8 :59 Pacific Time

Limit the number of jobs to display
Max. Number of results to show: 50 100 500 1,000 5,000 10,000

Refine search

Host: All User: Repo: Not Exec Queue: All
Charge Class: All JobID: Node Name:

Select columns to display

Hostname
 Job ID Job Name User Exec Queue Charge Class Repo
 Status Nodes Cores Submitted Start Complete
 Wallclock Raw Hours MPP Hours Wall Request Wait Time
 Repeat column headings IPM jobs only

Update Display Update with CSV Export Reset Form Load Default Display

Displaying 500 jobs that completed between Oct-31-12 00:00 and Nov-07-12 16:08

#	Host	JobID	Job Name	User	Nds	Complete	Wall hrs	Raw hrs
0	carver	3704001	K2500i10WJ... [Full Name]	u15013	1	11/07/12 16:04	0.246	0.25
1	carver	3703783	K17875d9WJ... [Full Name]	u15013	1	11/07/12 16:04	1.372	1.37

View Page in: CMS Draft Site Published Site Logged in as Richard Gerber - Log out

User Input



- We want to know what you want. What would be useful or interesting?
 - Data
 - Analysis
 - Presentation and Communication

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

The screenshot shows the NERSC user portal interface. The main content area is titled "MY NERSC" and includes several sections:

- My Accounts:** A table showing accounting data updated daily at 05:00 Pacific Time. The table has columns for Login, Repo, Alloc Type, Used, User Remain, Repo Alloc, and Repo Remain.
- System Status:** A table showing the status of various hosts.
- My Current Jobs:** A section indicating that data is updated every 5 minutes and that there are no queued jobs.
- My Completed Jobs:** A table showing the last 50 completed jobs for user ragerber through midnight last night. The table has columns for #, Host, Job ID, Job Name, Nds, End, Wall Hrs, and Raw Hrs.

Login	Repo	Alloc Type	Used	User Remain	Repo Alloc	Repo Remain
ragerber	m1266	MPP	0	12,500.0	125,000	47,222
ragerber	mpccc	MPP	504	1,199,500.0	8,000,000	1,944,666
ragerber	m1266	STR	0	6,493.3	10,000	6,499
ragerber	mpccc	STR	9,152	190,848.0	2,000,000	719,766

Host	Status
Hopper	Up
Carver	Up
PDSF	Up
Genepool	Up
Euclid	Up
HPSS	Up

#	Host	Job ID	Job Name	Nds	End	Wall Hrs	Raw Hrs
0	Carver	3565421	STDIN	1	10/23/12 15:30	0.04	0.30
1	Carver	3565399	STDIN	1	10/23/12 15:26	0.01	0.08
2	Hopper	2249969	STDIN	1	09/24/12 14:54	0.02	0.55
3	Hopper	2206579	STDIN	1	09/13/12 16:46	0.54	12.98
4	Hopper	2205504	STDIN	2	09/13/12 11:32	0.06	2.80



National Energy Research Scientific Computing Center