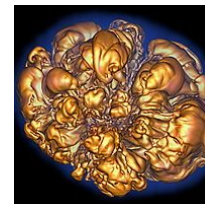
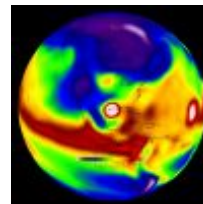
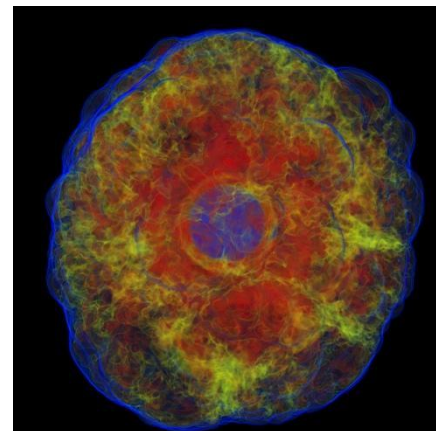
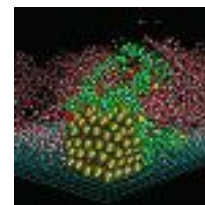
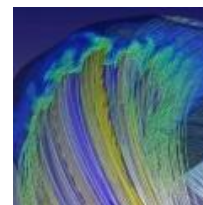
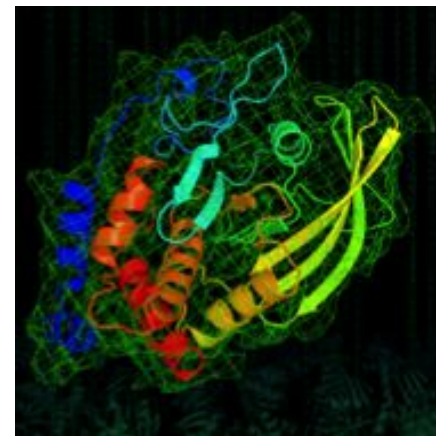
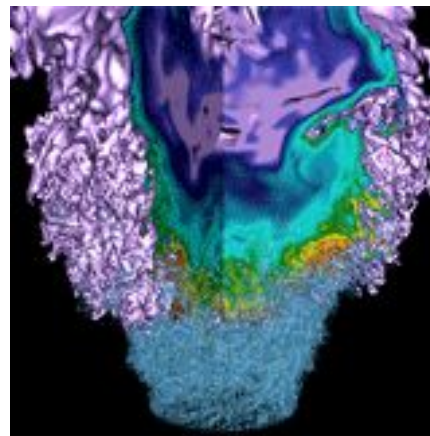
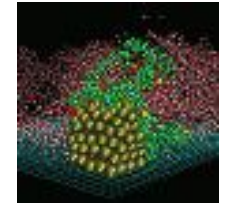
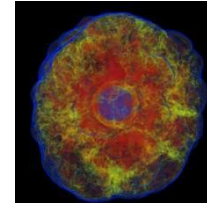
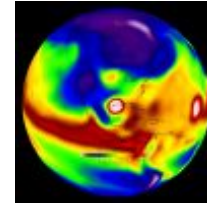
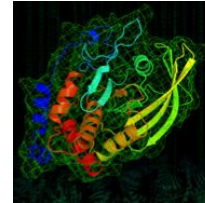
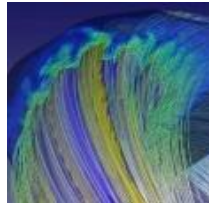
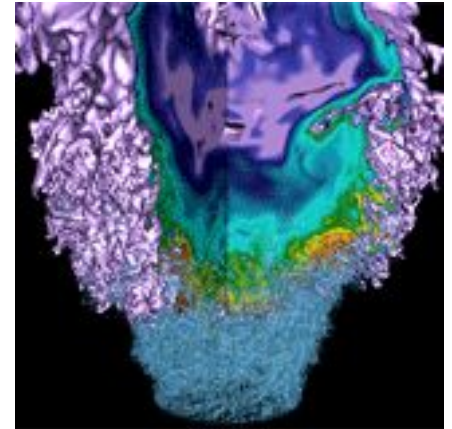


NUG Monthly Meeting



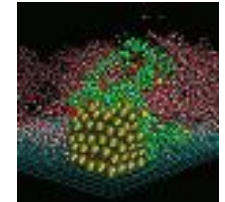
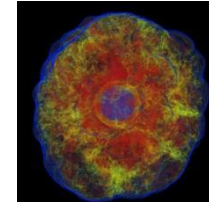
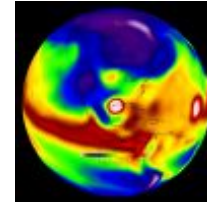
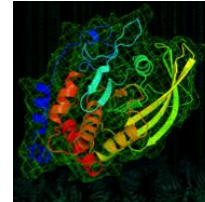
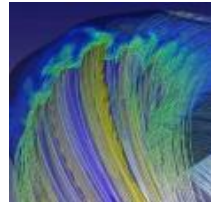
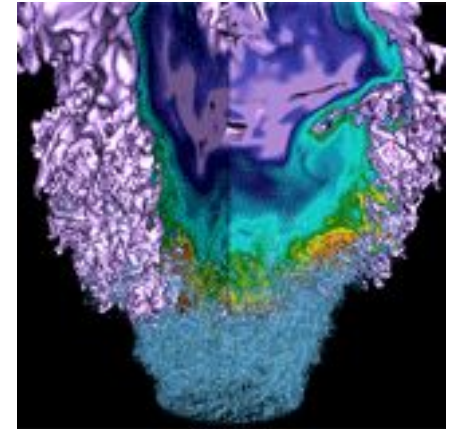
December 10th, 2015

CRT updates



Jeff Broughton
NUG Meeting, 12/10/2015

NERSC Move: Outage Schedule and User Impact



Helen He
NUG Meeting, 12/10/2015

Main Timeline



Event	Date
Cori Phase 1 available to all users	Already available!
Edison offline to move to CRT	Nov 30, 2015 (estimate to last 6 weeks), back online with SLURM batch scheduler
Hopper retires	Dec 15, 2015
Global project and global homes file systems migration to CRT	Nov – Dec 2015
JGI and PDSF file systems move to CRT	Jan - Feb 2016
Mendel moves to CRT (affects JGI, PDSF, Materials Project and Babbage)	Feb 2016

More details: <https://www.nersc.gov/users/move-to-crt/>

Global File Systems Move to CRT

- Migrate various file systems in phases over high speed network connection, expect NO file system downtime.
- During the migration, you may observe a reduction of file system performance.
- After the migration completes, some compute systems located at OSF will have decreased file system bandwidth.

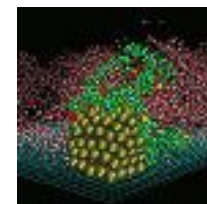
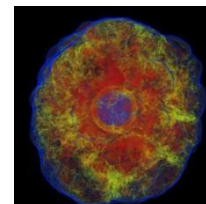
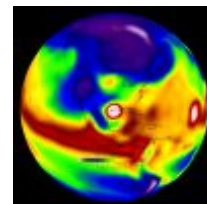
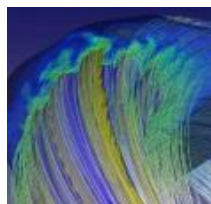
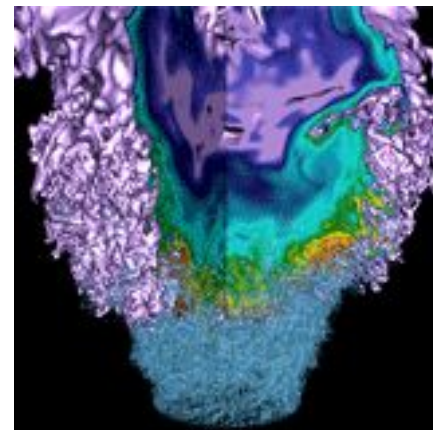
File System	Date
/global/projecta	Nov 13 - 16, 2015 (completed)
/global/project	Nov 30 – Dec 6, 2015 (completed)
/global/homes, /global/common, /global/syscom	Dec 14 - 16, 2015
JGI related file systems (/global/projectb, /global/dna/ and /global/seqfs)	Proposed start date Jan 11, 2016

More Moves in Feb 2016



- **Some PDSF file systems will be relocated to CRT**
 - About two weeks outage for these file systems
 - No overall PDSF downtime but some groups may not have access to all of their data
- **Mendel (a cluster providing resources to PDSF, JGI, the Materials Project, and Babbage) will move to CRT.**
 - PDSF and JGI: No outage. There are other compute components available. Total amount of resources reduced.
 - Material Project: expect 3 weeks outage.
 - Babbage may be offline for up to a month.

Hopper Retirement



Helen He
NUG Meeting, 12/10/2015

Hopper Retires on Dec 15, 2015 at Noon



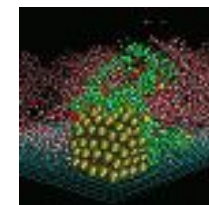
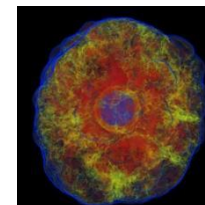
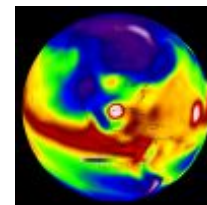
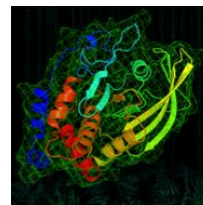
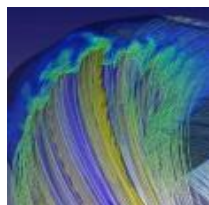
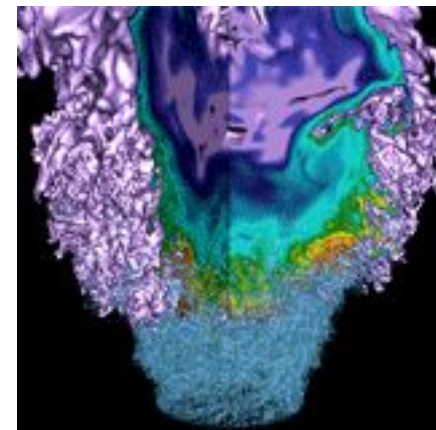
- **No further software changes will be made, except for security-related issues.**
- **At 12:00 PST (noon) on Dec 15, Hopper queues will be turned off. Any jobs running at that time will be killed, and all jobs in the queue will be deleted.**
- **Hopper login nodes and scratch file systems are planned to be available until 12:00 PST (noon) on Dec 22 for users to retrieve files.**
 - Availability not guaranteed for the full time since Hopper maintenance contract ends on Dec 15.
 - No file system recovery will be made after Dec 15.
 - File system may be taken down sooner (before Dec 22) if there are issues.

Please Save Your Important Files Now



- **Save your important files from your /scratch and /scratch2 file systems to HPSS, your /project directory (notice it is 92% full now), or elsewhere.**
- **Do not count on retrieving your files after December 15. Backup your data ASAP.**
- **Use “htar” utility to concatenate small files when archiving to HPSS.**
- **Files in your global home and project directories can still be accessed from all NERSC systems after Hopper retirement.**

Cori Phase 1 Update



Wahid Bhimji

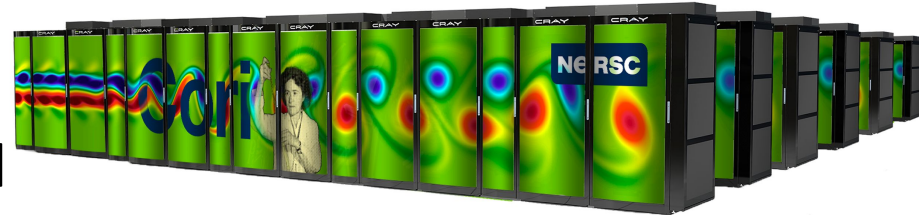
**NERSC Users Group
December 10th 2015**

Cori Overview (reminder)

- Phase 2 (coming mid-2016) - over 9,300 '[Knights Landing](#)' nodes

Phase 1 (installed now):

- **1630 Compute Nodes**
- Two x 16-core (2.3 GHz) Haswell
 - 128 GB DDR4 2133 Mhz memory/ node(some 512 /768 GB)
 - Cray Aries high-speed “dragonfly” topology interconnect
 - 12+ login nodes for advanced workflows and analytics
 - SLURM batch system
- **Lustre File system (also installed now)**
 - 28 PB capacity, >700 GB/sec peak performance
 - Ultimately will be mounted across other NERSC systems
- **Burst Buffer (NVRAM based)**
 - 750 TB now. Striped / Private / Persistent use-cases

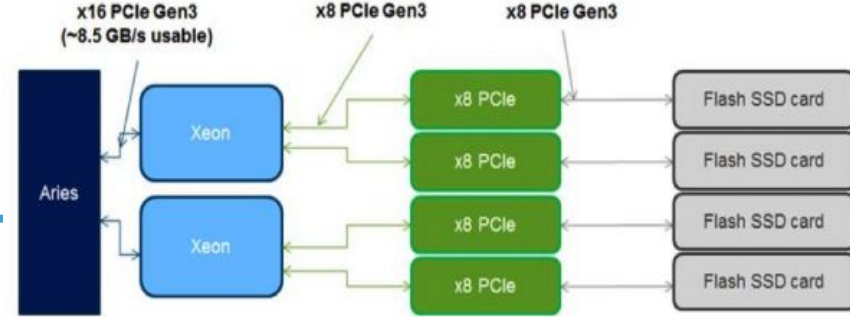


Cori P1 Genera Status

- **Fully open for all (mpp) users.**
- Still in acceptance (availability) period due to some node failures.
- Various changes to queue configurations to help jobs throughput
 - <https://www.nersc.gov/users/computational-systems/cori/running-jobs/queues-and-policies/>
 - ‘Realtime’ queue setup.
- Various issues resolved or have work arounds and are in progress (e.g. external login access; Intel compiler licences)
 - thanks for your patience.
- Please report any further problems to consult@nersc.gov

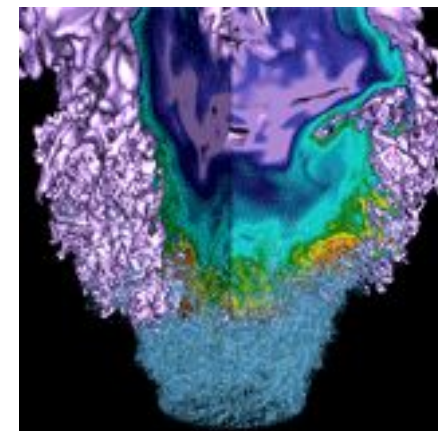
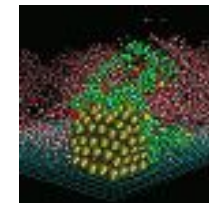
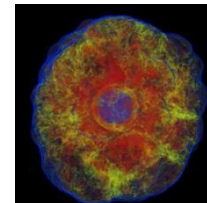
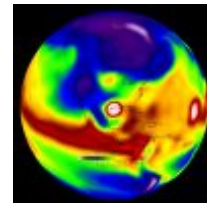
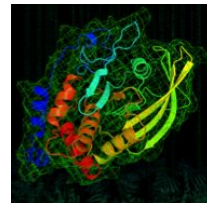
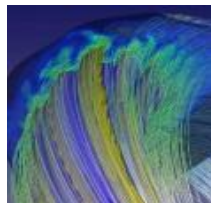
Full configuration details and user guides <https://www.nersc.gov/users/computational-systems/cori/>

Burst Buffer Status



- **Now available for early users (chosen by 'Early user program')**
- Downtime on 12/2 - installed patches which fixed various issues
- Many features now work:
 - 'scratch' per-job space - either in 'private' mode or 'striped' (across many BB nodes for performance)
 - 'persistent' reservation that can be used by multiple jobs/users
 - 'Stage in/out' - copies files on BB node before job start
 - Fully integrated with SLURM batch system. Example scripts at <https://www.nersc.gov/users/computational-systems/cori/burst-buffer/example-batch-scripts/>
- But still actively undergoing configuration and acceptance testing. Working with early users to shake out issues, improve performance.

Edison Updates



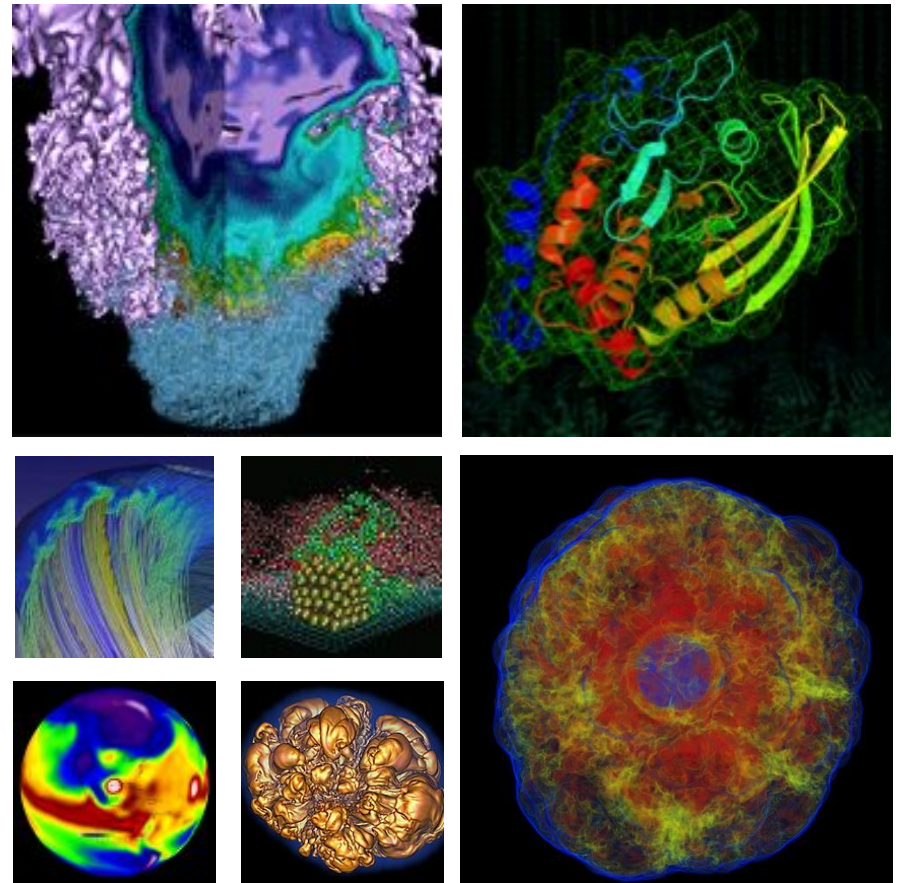
Zhengji Zhao
December 10, 2015

Edison Move Update



- **Edison was powered off on 11/30/2015 7:00 PST at the Oakland Scientific Facility building (Oakland, CA)**
 - Disconnecting/uncabling, packing, shipping ...
- **All cabinets arrived at CRT (Berkeley, CA) now**
 - Electricians are wiring Edison. Will have networking soon. Will switch batch system to Slurm. Edison will be ready to power up at end of the week.
 - Many more work on integration, hardware/software upgrades, configuration, verification, stabilizing, testing, ...
- **The scratch file system upgrade plan changed**
 - Only /scratch3 will be upgraded to Sonexion 2.0 with GridRaid, which requires reformat the /scratch3 file system.
 - The upgrade for /scratch1 and /scratch2 has been postponed
 - **All files on the /scratch3 on will be deleted.**
 - The files on the /scratch1, /scratch2 may remain intact, but no guarantee.
- **If everything goes smoothly, Edison may be back online for users as early as on 12/21.**

Allocations



Richard Gerber

December 10, 2015

Allocations



- **2016 awards will be announced Dec. 11, 2015**
- **Total allocated MPP Hours for 2016 will be the same as in 2015**
 - 2.4 billion for DOE Production
 - 5 billion requested
 - 300 million for ALCC
 - 300 million for NERSC Director's Reserve
- **Allocation Year 2016 begins on January 12, 2016**
 - Cori Phase 1 charging begins with machine_charge_factor=2.5 (units will be renormalized in 2017)

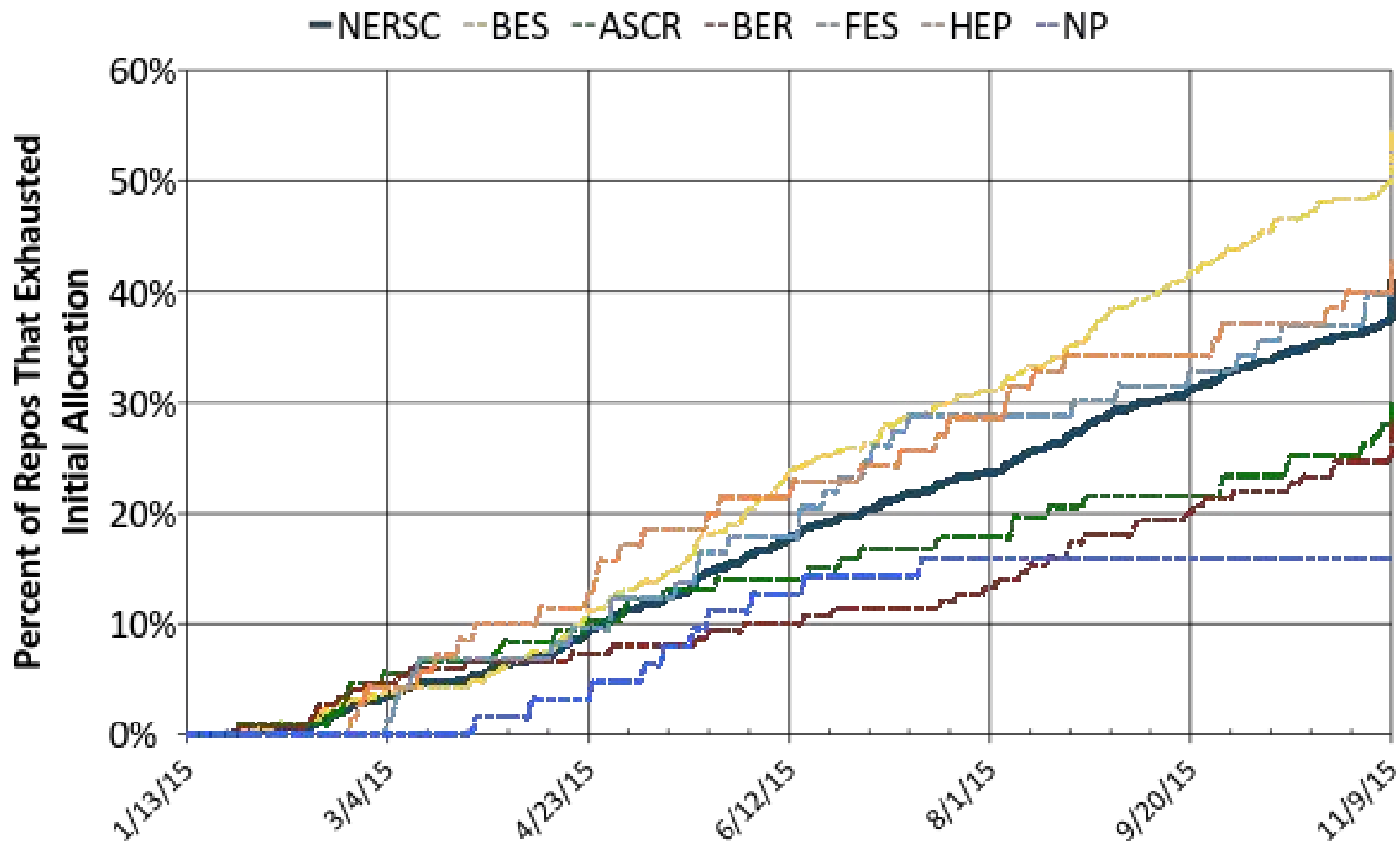
Allocations Changes for 2016



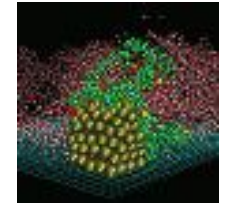
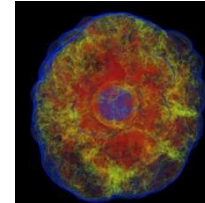
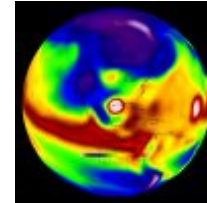
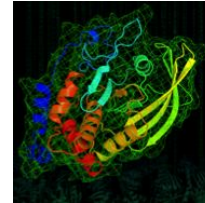
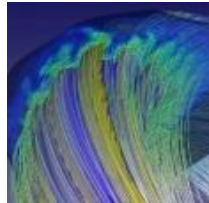
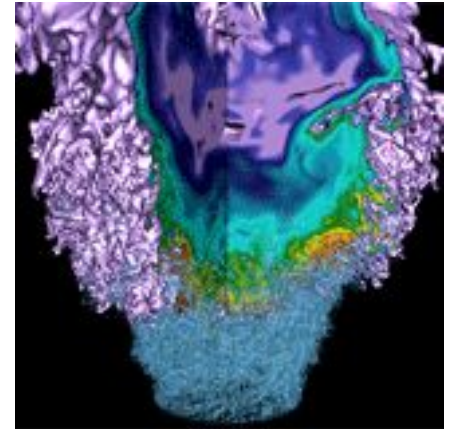
- **Repos that run out of time during the year will still be able to run jobs, but at a low priority (scavenger) on both Cori and Edison**
 - Provides priority access for repos that budget their usage based on their allocation, while not cutting off projects completely when they run out of time
 - Hope it will greatly reduce large number of allocations transfers and requests for additional time
 - Assures high system utilization even as repos run out of time
 - Starting configuration, which we'll modify if needed:
 - Premium (+5 days of aging priority), regular (0), low (-7 days), scavenger (-15 days)
- **First quarterly decrement will be eliminated**

Changes
made in
collaboration
with DOE
program
managers
and NUGEX

As year progresses, more repos run out of time and scavenger turnaround time improves.



NUGEX elections



Frank Tsung
NUG Meeting, 12/10/2015

NUGEX is made up of 21 members, 3 members from each of the DOE's 6 offices of sciences, and 3 at-large members.

In December 2015, there are 13 openings in NUGEX, including:

- **ASCR: 3**
- **BER: 3**
- **BES: 3**
- **FES: 1**
- **NP: 1**
- **At large: 2**

NUGEX Elections (cont)



The responsibilities of NUG includes serving on various committees (e.g. the queue sub-committee, NUGEX meeting sub-committee), this monthly teleconference, and the annual NUGEX meeting.

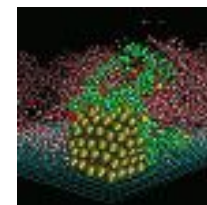
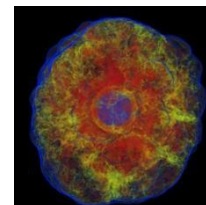
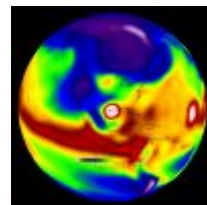
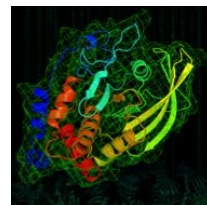
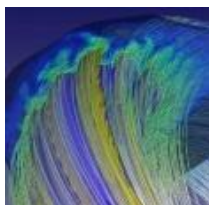
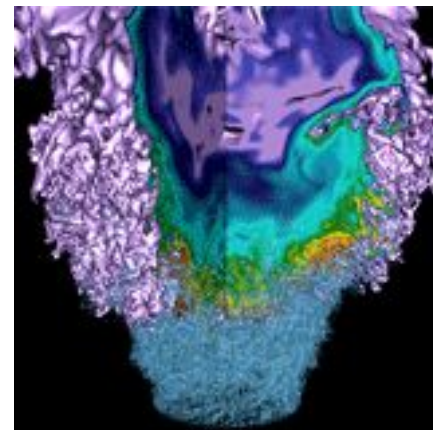
If you know anyone (including yourself) who is qualified please e-mail

Frank Tsung tsung@physics.ucla.edu

Anubhav Jain ajain@lbl.gov

as soon as possible, we would like to start the election in January.

User Survey



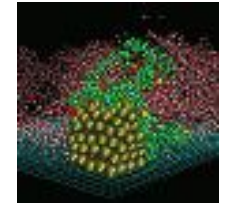
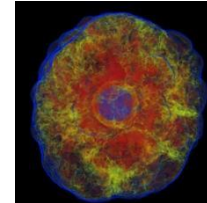
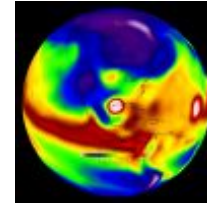
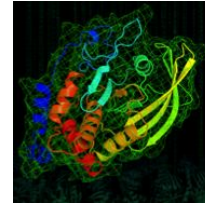
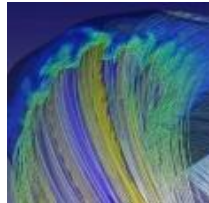
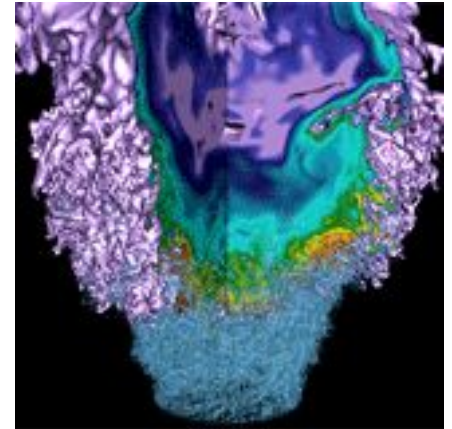
Debbie Bard
NUG Meeting, 12/10/2015

Please fill out NERSC Annual User Survey!



- Your feedback is important because it allows us to judge the quality of our services, give DOE information on how we are doing, and point us to areas in which we can improve.
<https://www.nersc.gov/news-publications/publications-reports/user-surveys/2015/>
- If you are not already logged in to the NERSC web site, you will be prompted for your NERSC (NIM) password when you start the survey.
- If you need assistance with your password, please contact our account support office at 1-800-66-NERSC (or 510-486-8600), menu option 2.
- ***The survey covers the allocation year 2015.***

Share Your Research



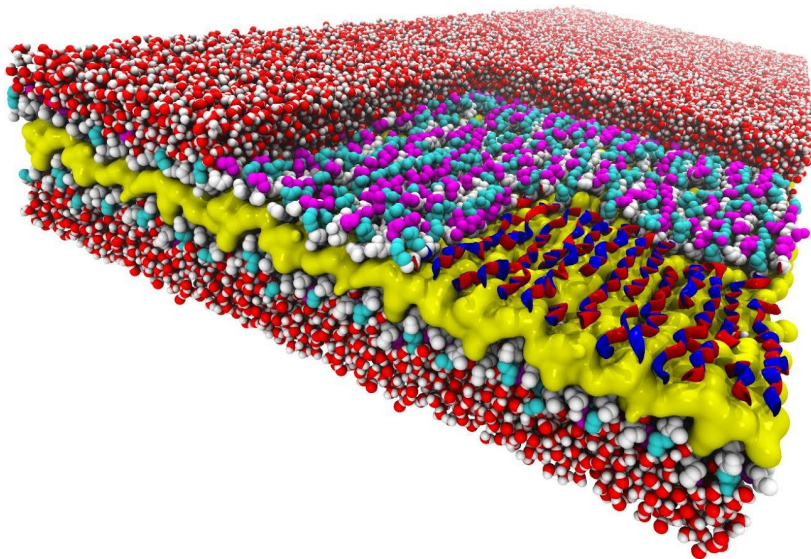
Debbie Bard
NUG Meeting, 12/10/2015

Share Your Research



- We want to hear about your success stories!
- <https://www.nersc.gov/science/share-your-research/>
- Publication reference, NERSC resources, brief description.
- Use this to highlight your research in various places...
- May lead to NERSC science story!
- We look forward to your submissions...

Science Byte: nature-inspired nanostructures



Atomic-resolution simulation of 2D peptoid nanosheet reveals a new snake-like structure.

red/blue: backbones

yellow: water repelling core

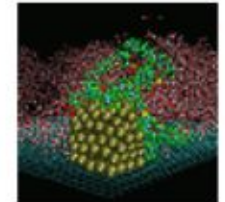
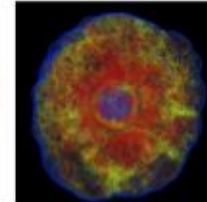
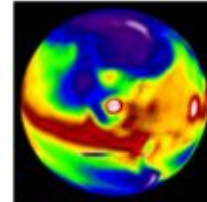
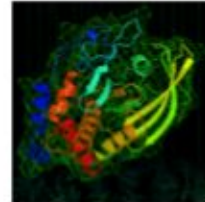
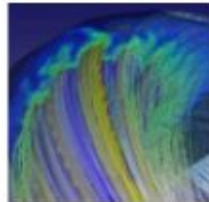
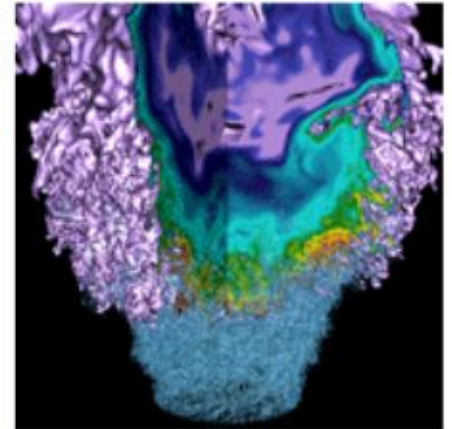
white: peptoid backbone

magenta/cyan: charged sidechains

red/white: surrounding water

- **Discovered design rule for new material: peptoid nanosheet**
 - synthetic molecules similar to protein-forming peptides
 - rule controls how polymers join to form “backbones”: snake-like pattern allows backbones to remain linear and untwisted → larger/flatter sheets than any natural material.
- **Combined simulations with x-ray scattering/imaging methods.**
- **New way to design biomimetic structures - design principles other than nature.**
 - “molecular paper”
 - sensitive chemical detectors? long-lasting catalysts?

Storing SSH public keys in NIM

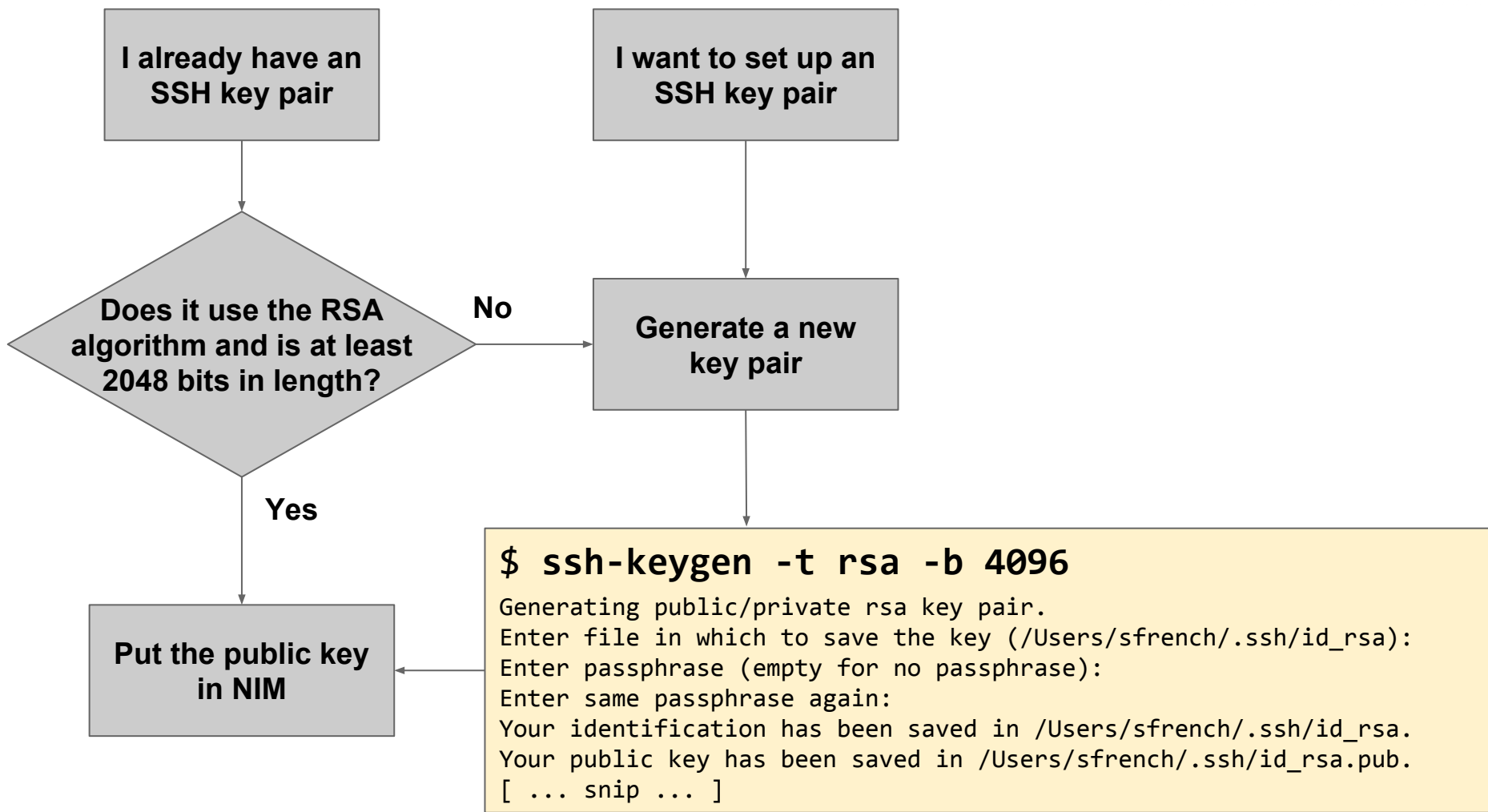


Scott French
NUG Meeting, 12/10/2015

SSH keys in NIM: Background

- **What's happening?**
 - Starting January 12, 2016, authentication via public keys in `~/.ssh/authorized_keys` is **disabled**
- **Who is affected by this change?**
 - **Affected:** If you use public key authentication to access NERSC systems via SSH
 - **Not affected:** If you authenticate using your NIM password (or connect via another mechanism)
- **Why is NERSC doing this?**
 - Security improvements that require authentication from a single source (LDAP)

So, what should I do?

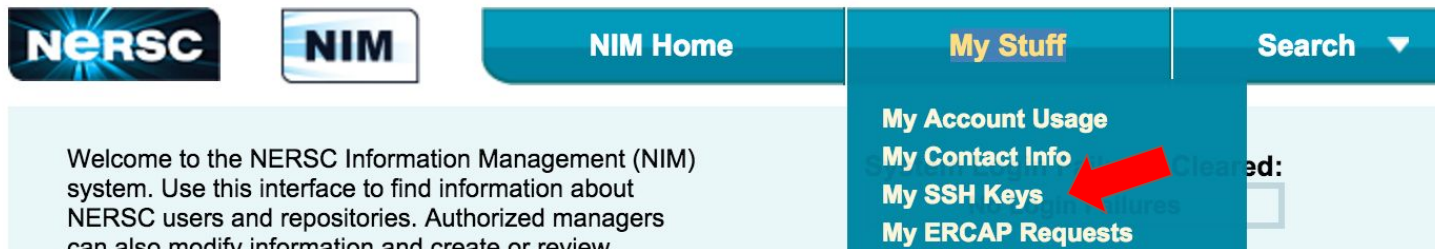


Adding SSH public keys to NIM (1)

Login to NIM and navigate to My Stuff → My SSH Keys:

Scott French | [Logout](#) | [Change My Password](#)

Q



You should see something like this:



Scott French



Adding SSH public keys to NIM (2)

Copy and paste your **public key** into the text box; Click Add:

Account Usage	Logins by Host	Unix Groups	Roles	Contact Info	Grid Certificates	SSH Keys
Scott French						
<div> <div>sfrench</div> <div> ssh-rsa AAAAAB3NzaC1yc2EAAAADAQABAAQCAQC9rSsk3/kSOez8YyYTIImxeAgPr++4CMxfS3/fRoIQiubW 8QdcxgU3et/bRWDEYM88rYwGWCQpZhlmqB2vZIPL/NhWkkK9d9AwFflxtSfMTJT14sJyd0u6SkM mNK7f4THRO3dc6ch9x88GLIOIIAI7xjvAhsQ5qKTgjb7OrGpmC7Kls2xnKkOA+mmzRERD2IAFUFE ZRipJd3eyNs+q/uDUttoPDQAlgeJM6aGmpHRaU8pbeR2OeEFs6UIJ5BzAUH02BiNa2i8ovLzCoUT 26KVkmC87D15cNj7WjN8F1s3lGrBHCzjOdkVtpHroKg5JTb0jU3MRu9MUJQpeotRQG6f/9L4+XZzu e1tepY7BZbvejH3QpLeF8lCu0PmrcTIm3Ax1LZXaHNpUuz6E7hZCf9zYvnKztVN4d7znCHQhdq0O eQyL4XiCud1V5vd/we6Qi3QmnNB7LNjzThPnKDsE9tgek5qkGOp6jA3i28BrtwYMn9fGxuADbSwW1 </div> </div> <div>Add</div>						

***Easy ways to
copy the public
key in its entirety***

Mac

```
$ pbcopy < ~/.ssh/id_rsa.pub
```

Linux

```
$ xclip < ~/.ssh/id_rsa.pub
```


Adding SSH public keys to NIM (3)

When you're done, you'll see something like this:

SSH Public Keys

 Uname	SSH Public Key	Enabled
<input type="checkbox"/> sfrench	<pre>ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCAQC9rSsk3/kSOez8YuYTIImxeAgPr++4CMxfS3/fRoIOiubW 8QdcxgU3et/bRWDEYM88rYwGWCQpZhImqB2vZIPL/NhWkkK9d9AwFIfxtSfMTJT14sJyd0u6SkM mNK7f4THRO3dc6ch9x88GLIOIIAI7xjvAhsQ5qKTgjb7OrGpmC7KIs2xnKkOA+mmzRERD2IAFUFE ZRipJd3eyNs+q/uDUttoPDQAlgeJM6aGmpHRaU8pbeR2OeEFs6UIJ5BzAUH02BiNa2i8ovLzCoUT 26KVkmC87D15cNj7WjN8F1s3IGrBHCzjOdkVtpHroKg5JTb0jU3MRu9MUJQpeotRQG6f/9L4+XZzu e1tepY7BZbvejH3QpLeF8lcu0PmrctTim3Ax1LZXaHNpUuz6E7hZCf9zYvnKztVN4d7znCHQhdq0O eQyL4XiCud1V5vd/we6Qi3QmnNB7LNjzThPnKDsE9tgek5qkGOp6jA3i28BrtwYMn9fGxuADbSwWj</pre>	Y

Tip: To delete an existing public key, mark the checkbox and click the red **X**

Tip: You can add multiple public keys to NIM

But, will it work for me?

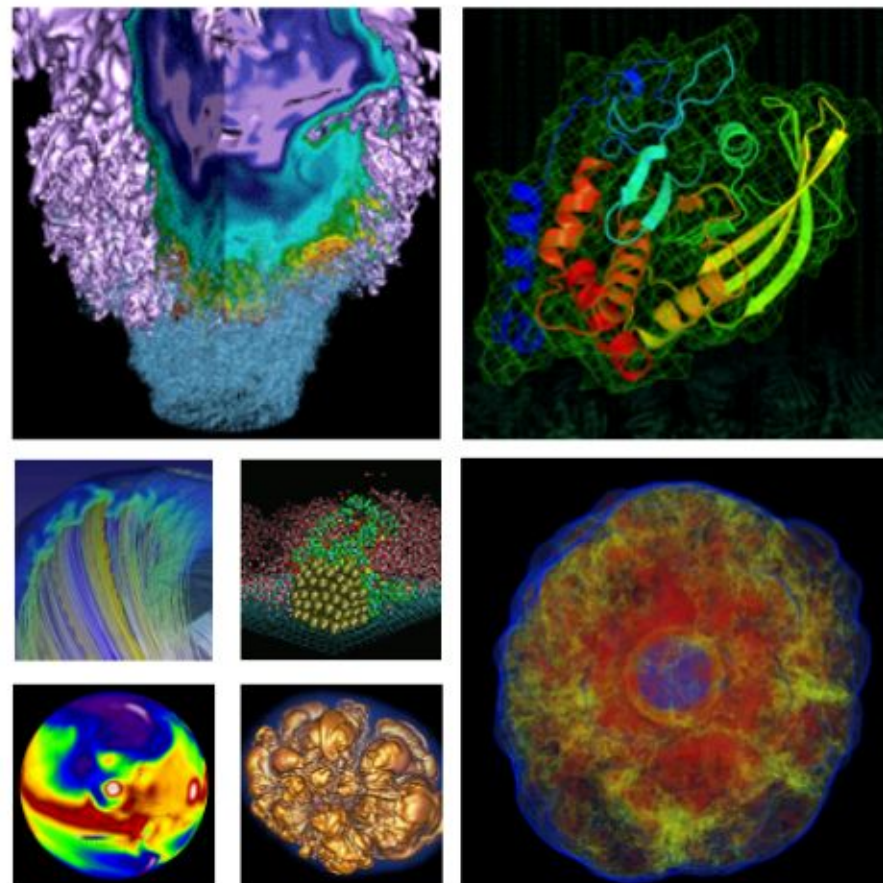
- **You can test this on Cori! After saving in NIM:**
 - Rename your `~/.ssh/authorized_keys` file to something else (e.g. `authorized_keys.old`)
 - Try logging into Cori
- **What if I ssh between nodes of the same system?**
 - This should “just work” (server-side configuration)
- **What if I ssh between NERSC systems?**
 - Option 1: ssh agent forwarding (`ForwardAgent`)
 - Option 2: Generate a new key pair at NERSC, and add that public key to NIM as well

What if I need help?

- **As always, the consultants are here to help!**
- **If you have a use case or workflow for which this change could pose problems, we'd like to know**
 - There is still time to come up with solutions or workarounds for unusual cases
- **Bottom line:** If you use public key authentication when ssh-ing to NERSC, please give this a try soon on Cori
- **See also:** <https://www.nersc.gov/users/connecting-to-nersc/connecting-with-ssh/>

WARP/PICSAR

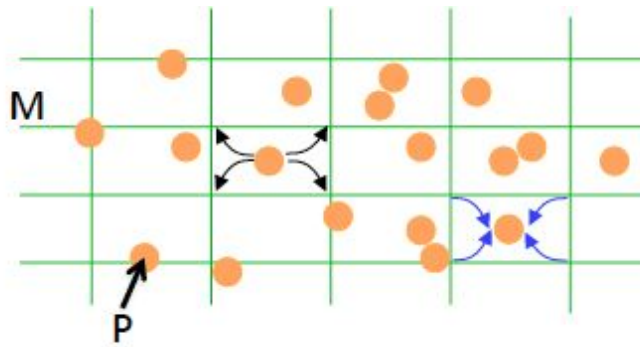
Update from Intel dungeon session Oct 5-7



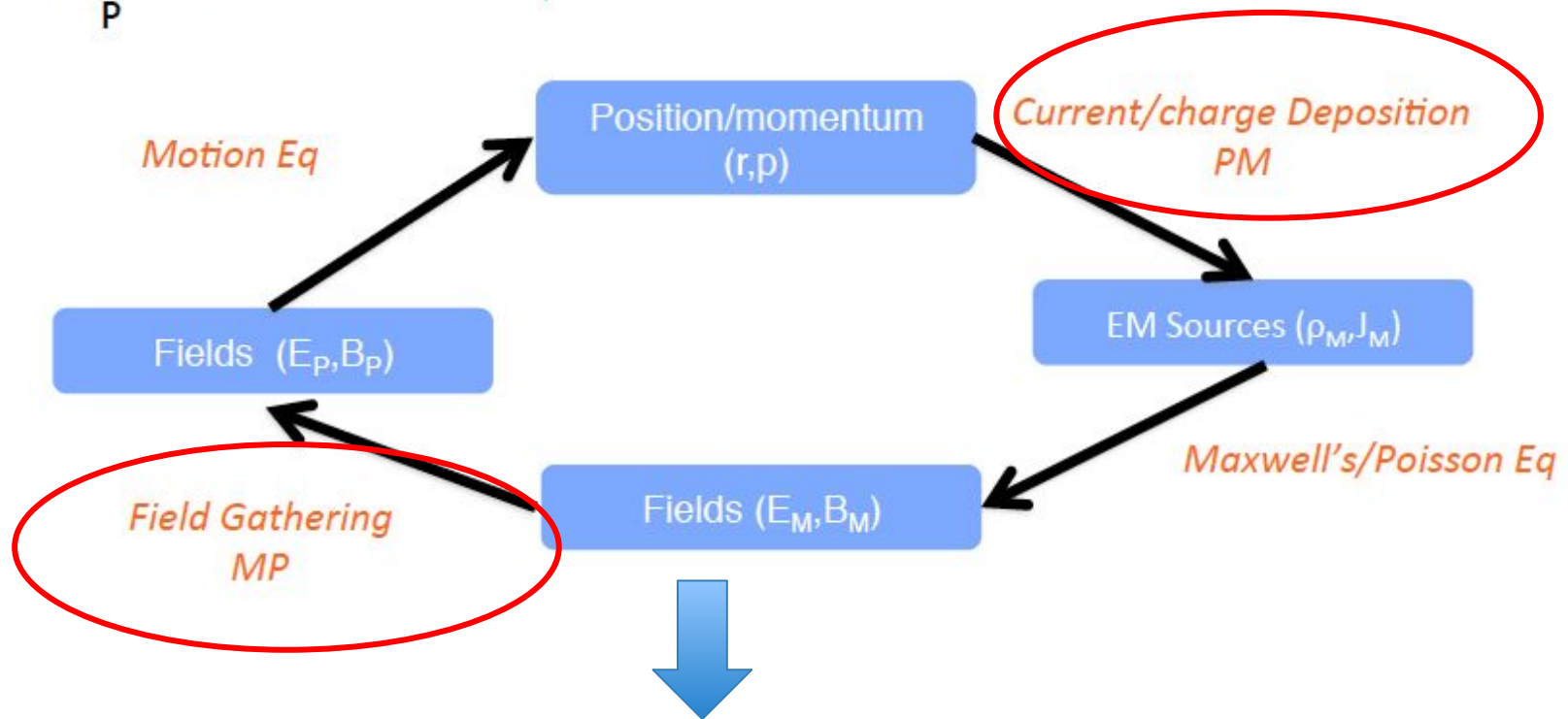
Ankit Bhagatwala
NERSC User Services

12/10/2015
NUG Telecon

Main steps in a PIC accelerator code



Charged macro-particles (p^+ , e^- , ions, ...) interacting through fields on grid



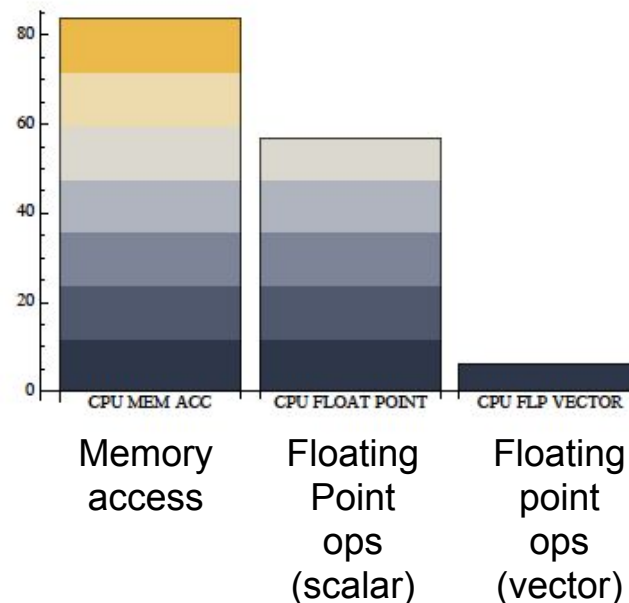
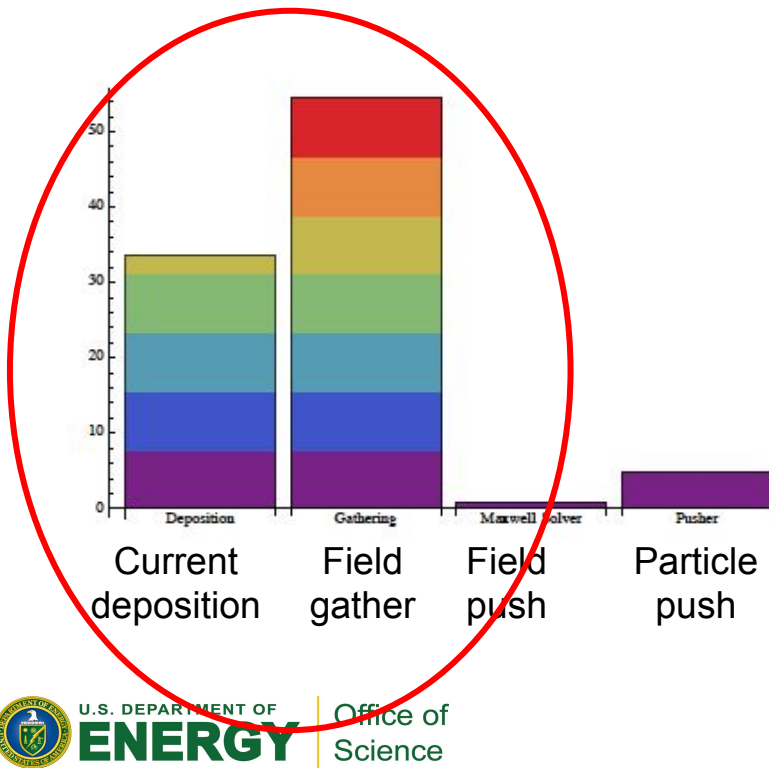
PICSAR → Particle-In-Cell Scalable Application Resource



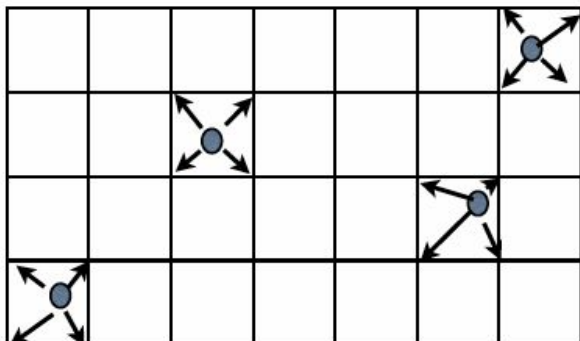
Single thread performance on Babbage



- Current deposition (particle-to-grid) and Field gather (grid-to-particle) most time consuming subroutines
- Large time spent in memory accesses
- Low vectorization

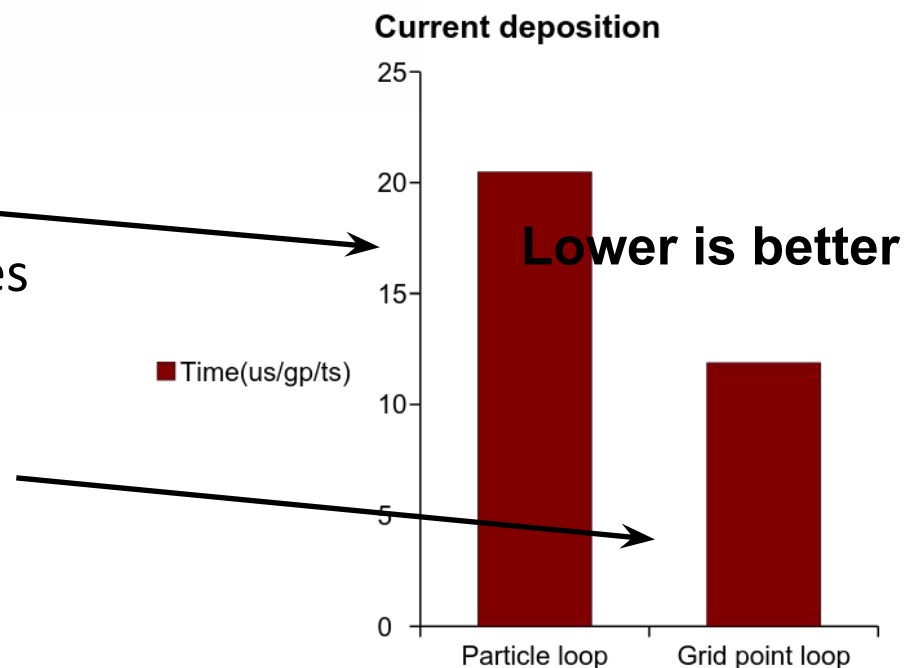


Optimization 1: Vectorized current deposition

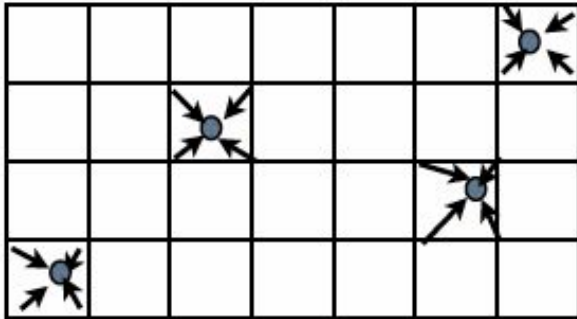


`mesh = mesh + particle_value`

- **Vectorize over particles**
 - Non-contiguous memory accesses over neighboring grid points
- **Vectorize over 8 neighboring grid points**
 - 8x Memory overhead but substantial speedup

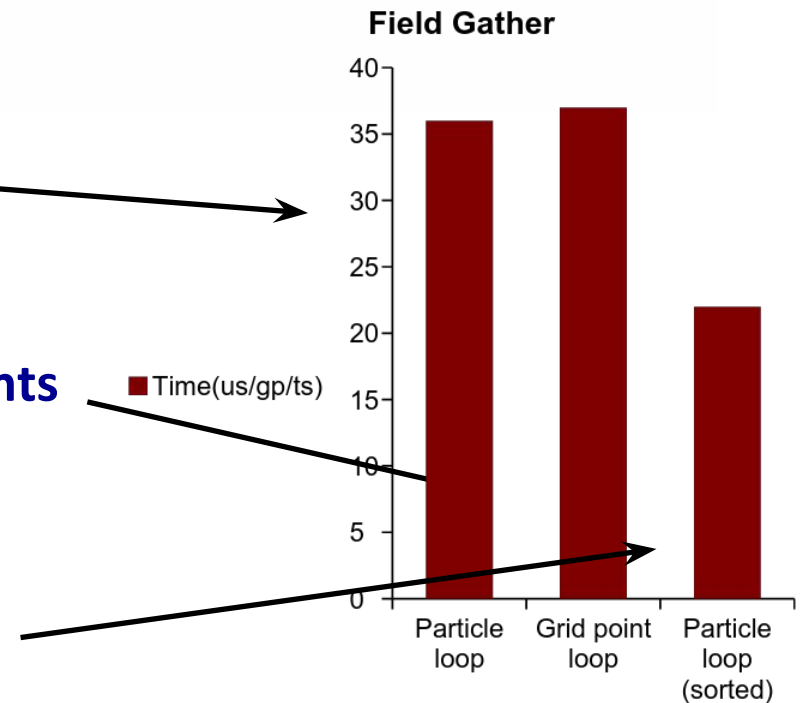


Optimization 2: Vectorized field gather



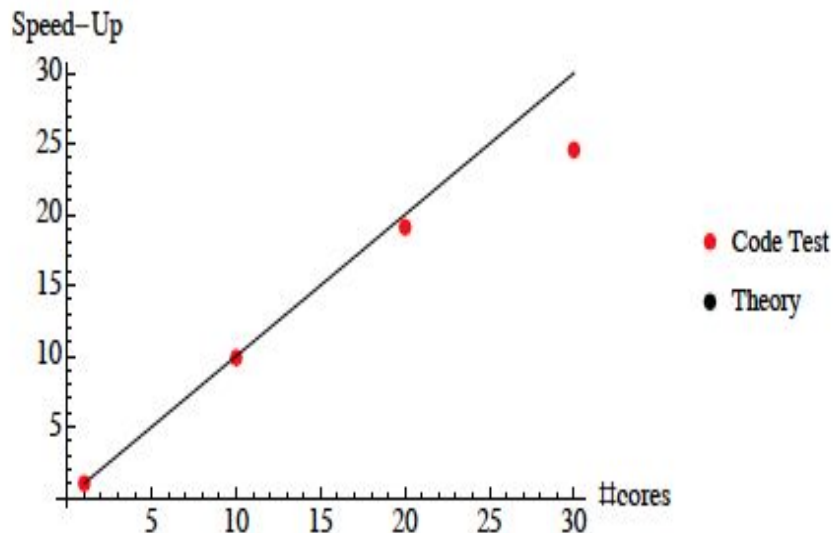
$\text{particle_value} = \text{particle_value} + \text{mesh}$

- **Vectorize over vertices**
 - Non-contiguous memory accesses over neighboring grid points
- **Vectorize over 8 neighboring grid points**
 - Reduction operation: no improvement
- **Vectorize over sorted particles in one cell** – Sorting overhead

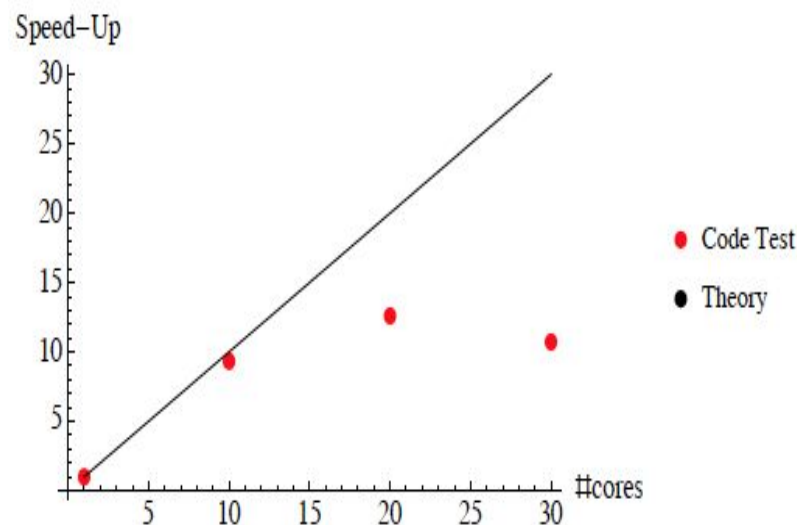


Lower is better

Optimization3: OpenMP scaling

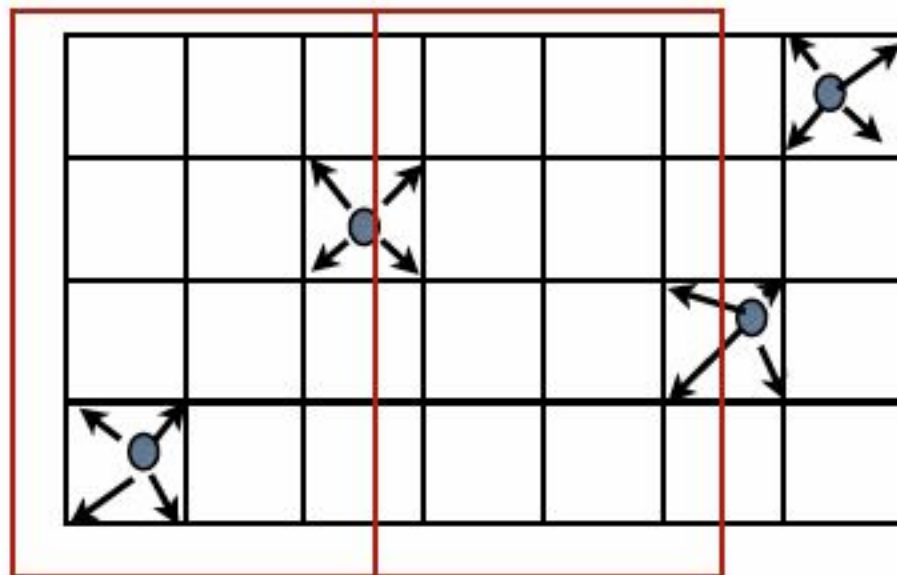


Field gather



Current deposition

- **Field gather** routine shows good thread scaling on Babbage
- **Current deposition** suffers from serialization bottleneck due to reductions across multiple threads updating same grid points – need to rethink parallelization approach



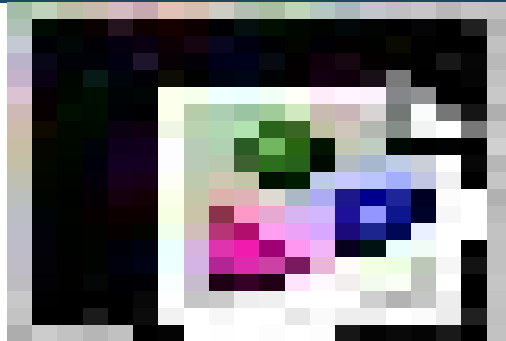
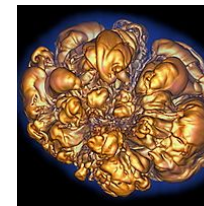
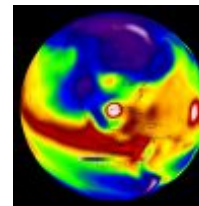
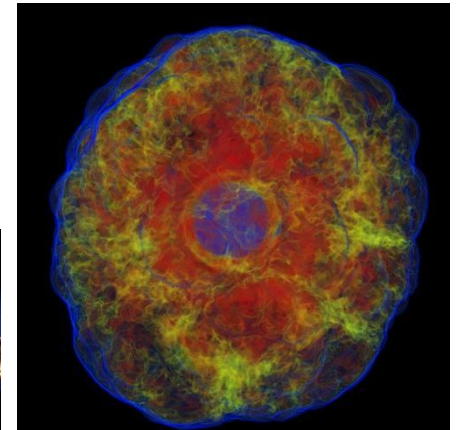
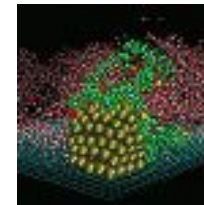
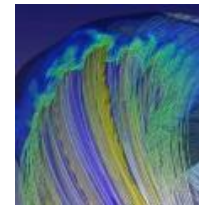
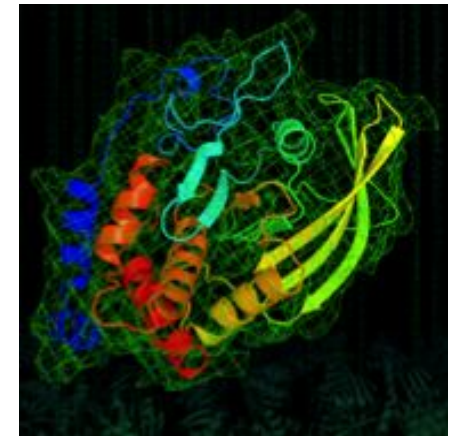
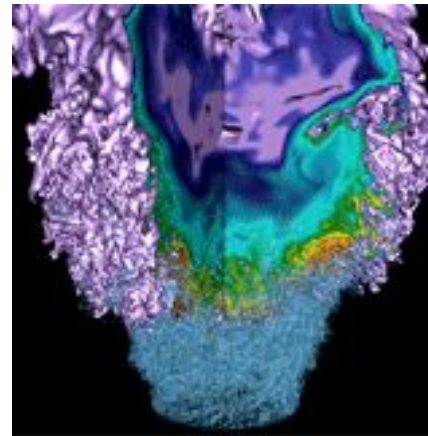
Thread 0

Thread 1

- Multiple threads updating same grid points
- To avoid race issues, each thread maintains private copy of tile
- Reduction operation to merge each threads contribution to shared array induces serialization

- Introduce optimizations from PICSAR to WARP
- Particle sorting to enable vectorization of **field gather** routine
- Merge routines for greater reuse of cached tile data
- Include guard cells (for **current deposition**) into the shared memory arrays
 - Get rid of the private copies for each thread.

Accelerating Analysis with SciDB



Lisa Gerhardt
NERSC, LBNL

December 10, 2015

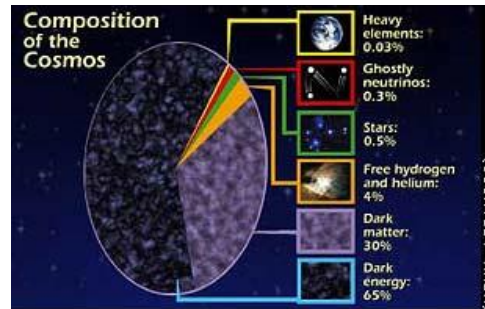
- **Vast majority of experimental data is array-like**
 - Time series data from detectors (LHC, IceCube)
 - Image data from satellites and light sources
- **TBs to PBs of data, that is written once and read many times**
- **Reduced, aggregated, analyzed**
- **Strong need for a tool that can handle this kind of heavy workload**

- **Array DB management system and analytic platform**
- **Parallel architecture that is transparent to the user**
- **Python and R interface**
- **Filtering and aggregating**
- **SQL-like joins**
- **Lots of optimized functions**
 - Sampling
 - Matrix operations (ScaLAPACK)
 - Statistics
 - User designed functions

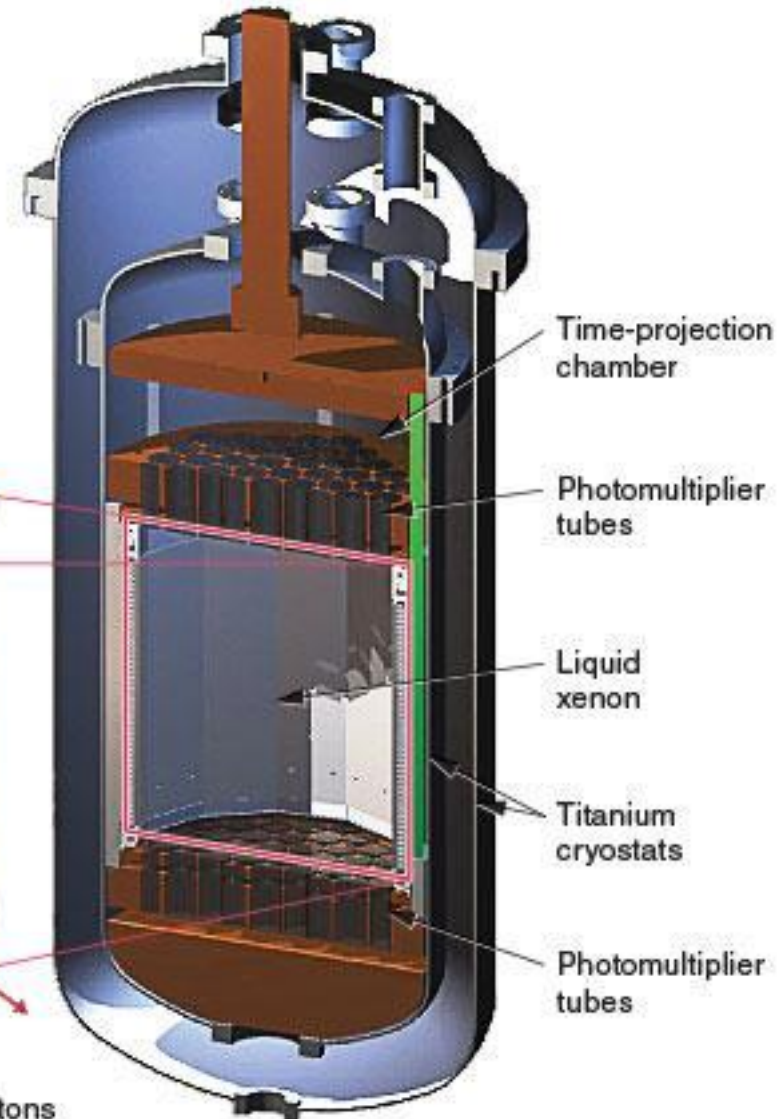
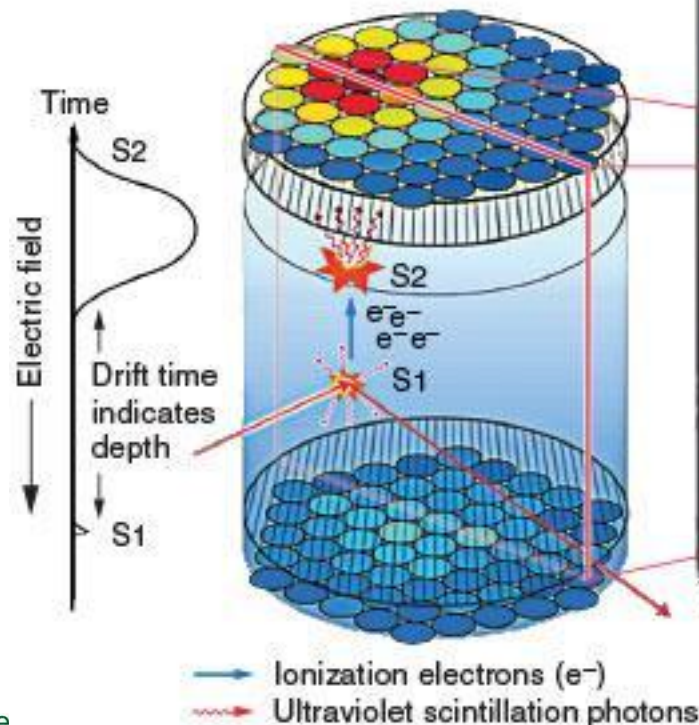
Case Study 1: LUX Detector (HEP)

NERSC

Searching for dark matter with 370 kg of liquid Xenon



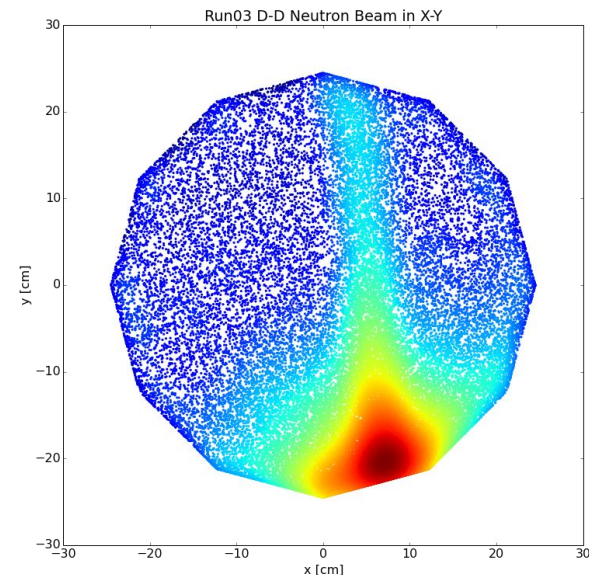
Precursor to LZ, chosen by DOE as one of two next generation dark matter detectors



LUX's SciDB Setup at NERSC



- **Proof of concept with 32 instances on test cluster**
- **Uploaded 32 TB of inaugural data**
- **600 Million events**
- **Dimension of array:**
 - Unique timestamp
 - Pulse type
 - Pulse number
- **Roughly 50 attributes**



Problem 1: Detector Stability



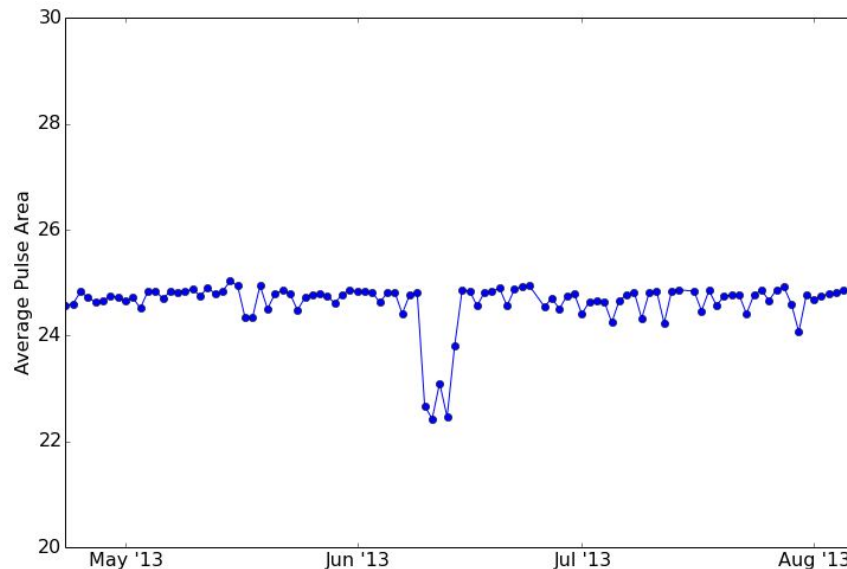
- **Look at variation in measured electron pulse size over time**
 - First, pull out all single electron pulse types and toss the rest of the attributes:

```
filter(project(pulse_array,pulse_size),  
       pulse_type==single_electron)
```
 - Gaussian fit needs a histogram, so we can leverage SciDB regridding functionality
 - Regrid to clump time into 5-hour window size and to get a count of pulse_size in each pulse_size bin
 - Automatic count functionality is folded into regridding

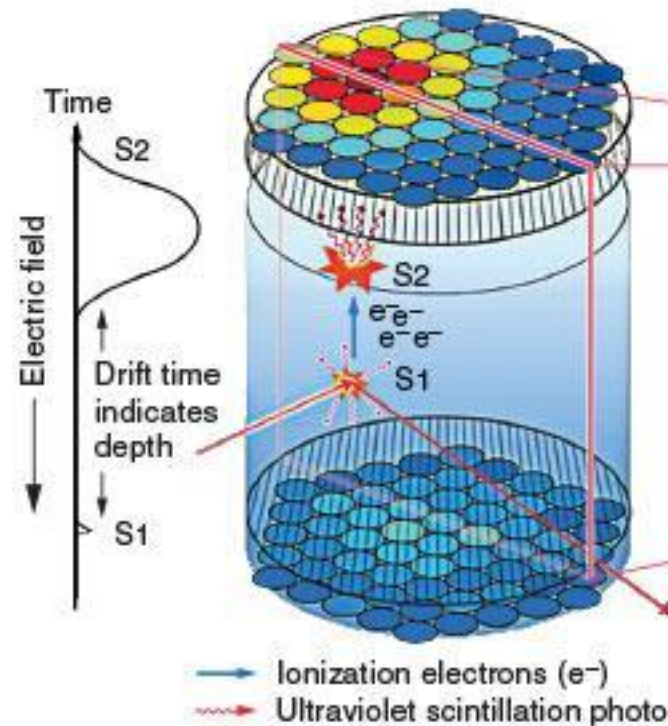
Results



- Result is ~2000 histograms, one for each time window
- Spun over 200 M entries (including scipy Gaussian fits for each slice) in 7 minutes



Problem 2: Search for Signal



- Task: Find every event with a small initial pulse followed by larger pulse within 1 ms

- **Produces an array with every possible combination of the two input arrays**

```
cross(left_array, right_array)
```

- **Split array into two sub arrays one**
 - 300,000 entries with small peaks
 - 1,000,000 entries with large peaks
- **Use cross to join these two arrays**
- **Filter events that have a time difference of 1 ms or less**
- **Took only 4 hours to search 100 days of data for signal candidates**

Case Study 2: Metabolite Atlas (BER)

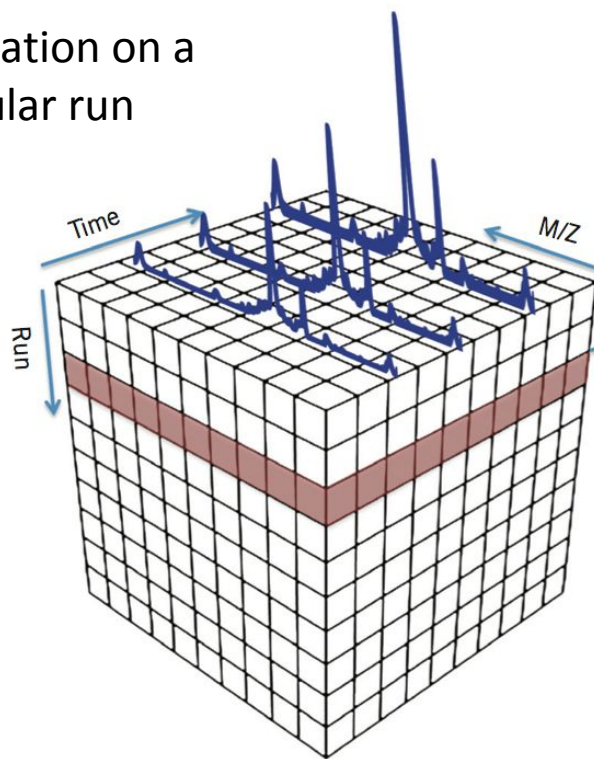


- **Profile molecular composition of metabolic by-products**
- **Many different metabolites and apparatus produce 10s of GBs of data a day**
 - At each time step the intensity of mass spectrum reading for each M/Z is recorded
- **Used SciDB to compile an atlas of measurements that is searchable via web portal**

Common Uses

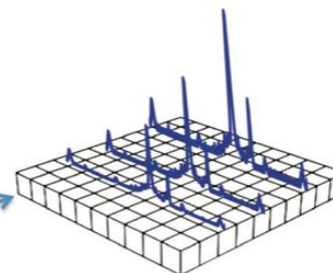


Pull all
information on a
particular run
(filter)

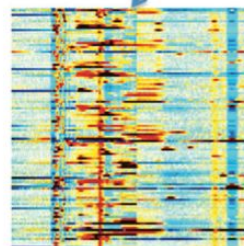


Data cube

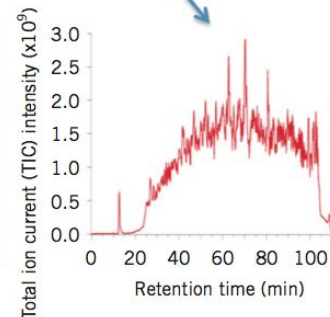
Sub-selection



Aggregate
operations



Heat map



TIC/XIC

Sum over all M/Z
values
(aggregate)

Plot time, M/Z
heat map (regrid)

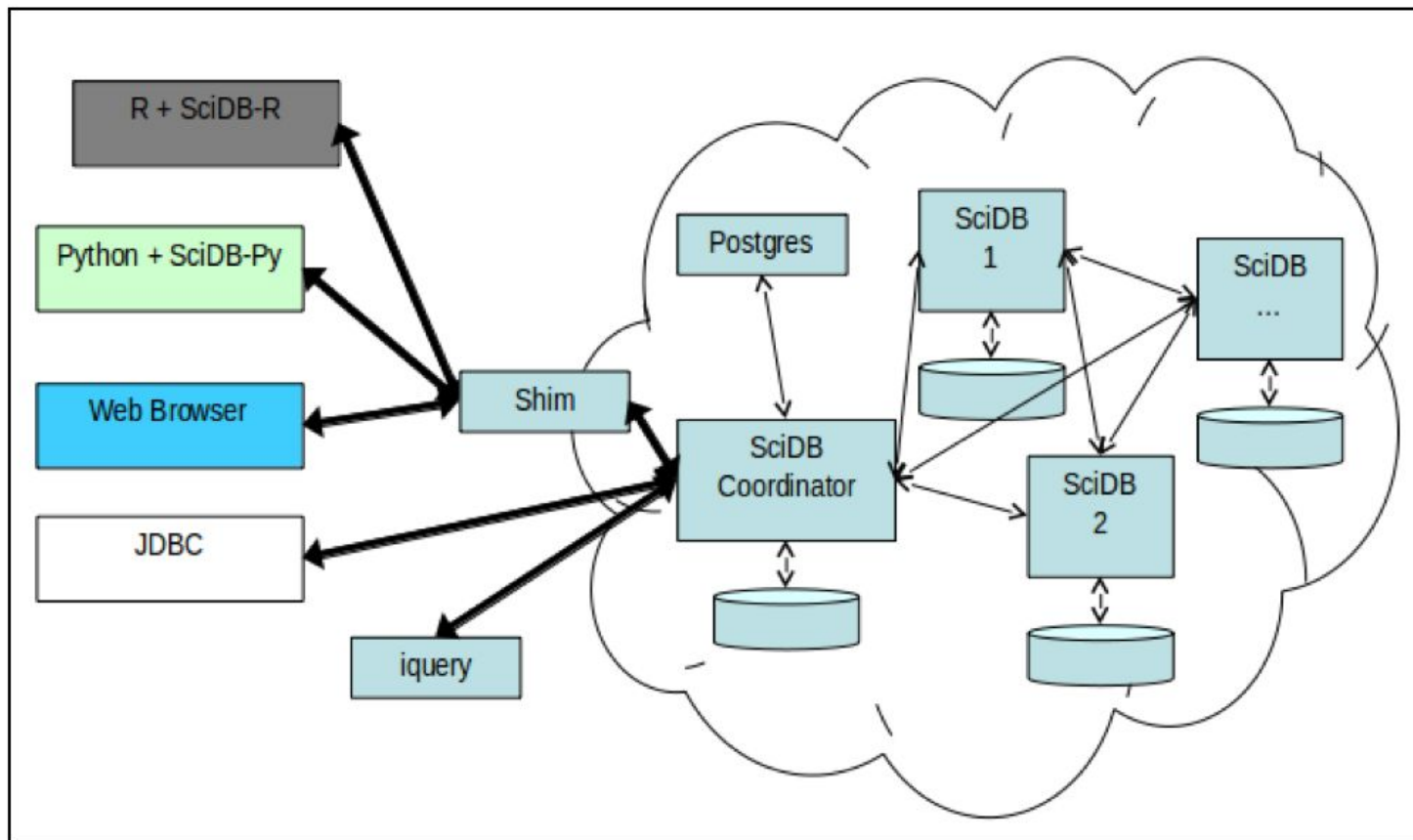
Picture from
Y. Yao et al., Computing in Science & Engineering, May 2015

- **SciDB is good for**
 - Easily using parallel resources
 - Write once, read many analyses
 - Daily updates of reasonable size okay (~300k entries)
 - Filtering, aggregating, joining
 - Can easily create new array shapes for new analysis ideas
 - Can easily handle 10 – 100s TB data
- **If you're interested in giving SciDB a try, please contact consult@nersc.gov**



National Energy Research Scientific Computing Center

SciDB Architecture



SciDB Structure



Arrays are chunked

Chunk 1

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

Chunk 2

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

Chunk 3

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

Chunk 4

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

Each chunk is stored on a different instance

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02

0.02	0.01	0.01	0.02
0.01	0.01	0.5	0.02
0.01	0.02	0.01	0.01
0.02	0.01	0.02	0.02



Data is scanned from array storage on all instances and streamed into and out of each operator one chunk at a time. "Hot" chunks are kept in memory.

SciDB Test Cases at NERSC



Big Data Tasks in SciDB@NERSC Projects

	Bio-Informatics	Climate Simulation	Odetta	Variable Stars	OpenMSI	SDSS Spectrum Analysis
Querying	✓	✓	✓	✓	✓	✓
Density estimation						
Regression				✓		✓
Classification				✓		
Dimension reduction	✓		✓		✓	
Outlier detection						
Clustering	✓					
Time series analysis		✓		✓		
Feature selection & causality						
Fusion and matching					✓	✓

Creating an Array



Create array ZOO <monkey:int64 zebra:int64 elk:
int64> [cagenum=0:100,10,0, distemper=0:*,2,0]

Attributes: monkey, zebra, elk

Dimensions: cagenum, distemper (can be non-ints)

Dimension size: 0:100 (* unbounded)

Chunk size: 10

Chunk overlap: 0