FLATIRON INSTITUTE

CHALLENGES IN COMPUTING FOR MULTI-DISCIPLINARY SCIENCE AT THE FLATIRON INSTITUTE

IAN FISK

March 21, 2018

Background

- The Simons Foundation has existed to support science for over 20 years
  - Currently Simons gives grants for over $300M annually across math and physical science, life science, large scale topic-driven science collaborations, and autism research
  - 5 years ago Simons decided they wanted to have an in house research group
    - Started with the Simons Center for Data Analysis (SCDA), which was expected to be about 50 people

FLATIRON INSTITUTE

Flatiron Institute

- Two and a half years ago it was decided to dramatically expand the scope
  - 250 scientists
  - 4 scientific centers
  - A dedicated Facility
- We moved into the building a year and a half ago
  - Construction should be over by summer

FLATIRON INSTITUTE

The Flatiron Institute is the research division of the Simons Foundation.

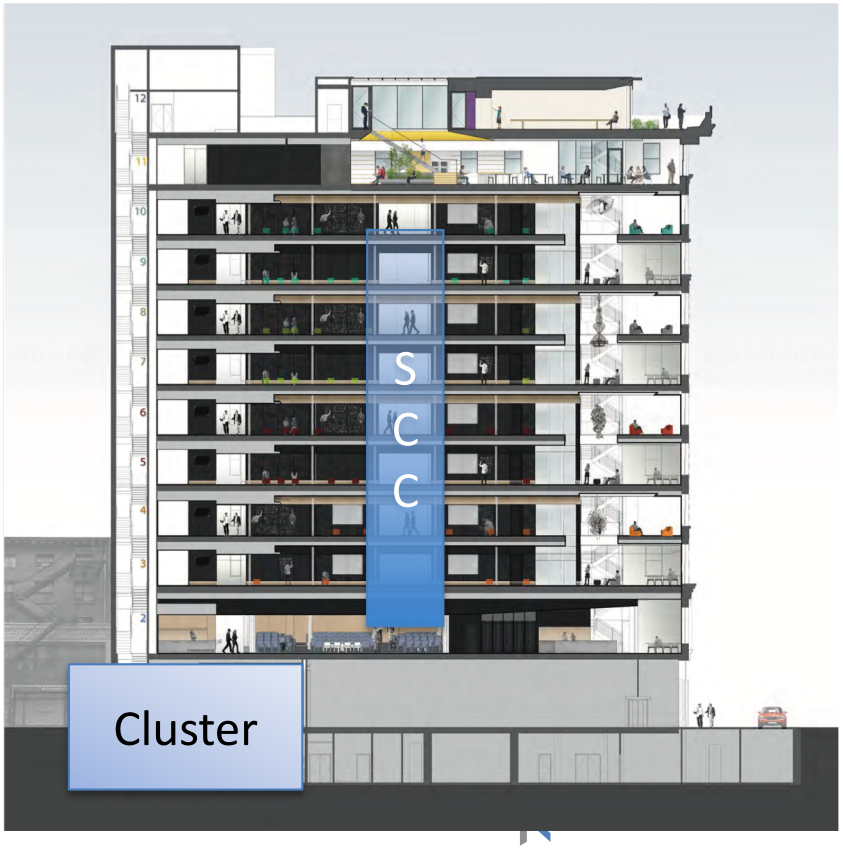
Its mission is to advance scientific research through computational methods, including data analysis, modeling and simulation.

- Center for Computational Biology (CCB)
- Center for Computational Astrophysics(CCA) and
- Center for Computational Quantum Physics (CCQ)
- CCX
- Scientific Computing Core (SCC)

FLATIRON INSTITUTE

Challenges

- We are in Manhattan
  - –Space is at a premium
    - Space used for computers and computing support is space not used for scientists
      - – Data Center is in the basement
      - – Currently Scientific Computing Core is 9 people (7 + 2 directors)
        - » Roughly balanced between hardware and software experts



LOCATIONS

- Flatiron has a data center in the basement
- A co-location facility at BNL
  - From the user perspective the two NY facilities are integrated
- A tape archival facility at Fermilab
- A leased super computer at the San Diego Computer Center
  - Gordon
- In addition there are opportunities in Massachusetts and Wyoming

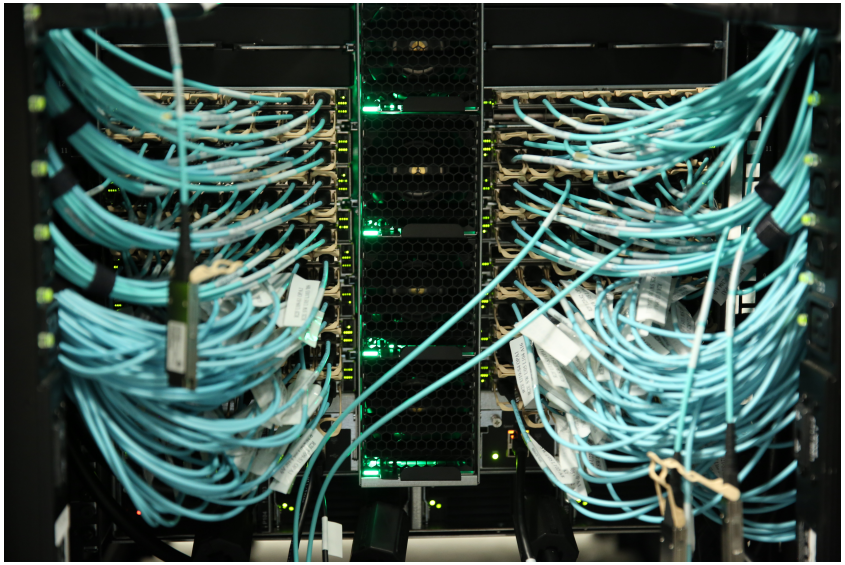
Technical Challenges

- Diversity of requirements
  - HPC/HTC
  - Finite vs. Infinite Computing
  - GPU and specialized hardware
- Dealing with Data
  - Quantity
  - Archiving
- Sharing
  - Exporting Data

FLATIRON INSTITUTE

Diversity of Requirements

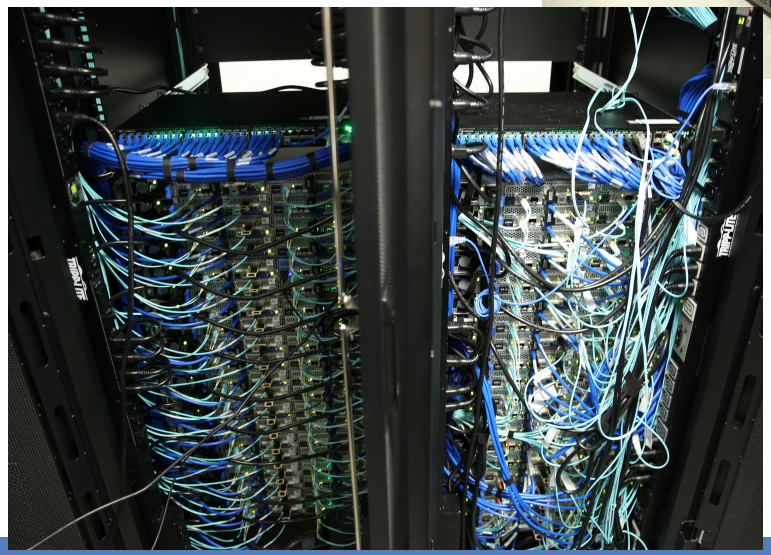
- Astronomy simulation is primarily HPC using a few hundred cores in a tightly coupled configuration
- Some portion of the Quantum systems work is also HPC
- Low latency interconnects like Omnipath are needed



FLATIRON INSTITUTE

High Throughput Computing

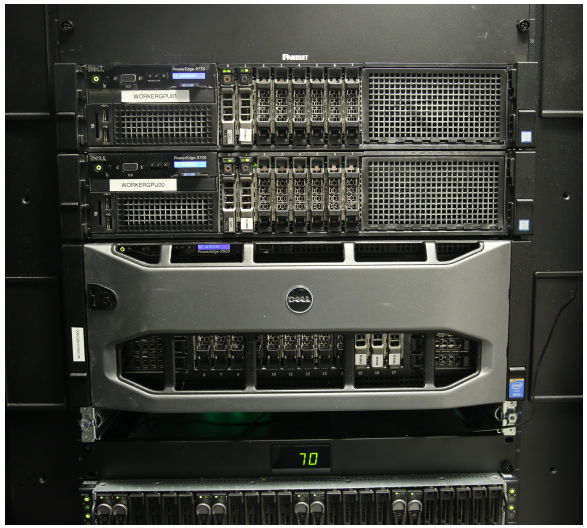
- Biology and half of Quantum systems are HTC
  - Processing data and validating code in the highly parallelizable infrastructure
    - Easier to support
    - Easier to schedule
- Nodes are common
  - 20GB RAM per core



FLATIRON INSTITUTE

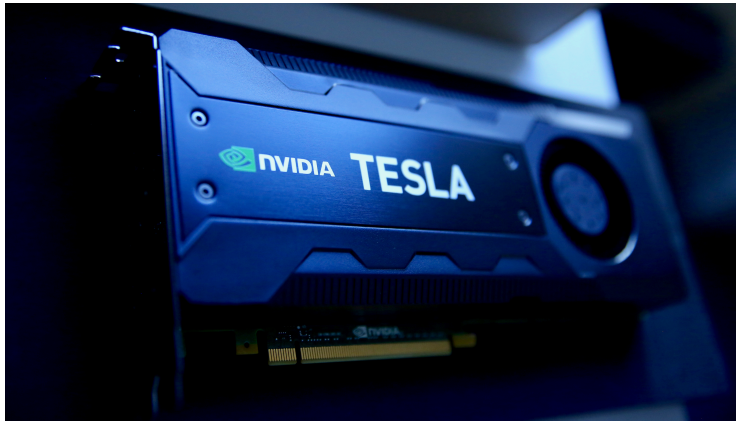
Infinite >> Finite

- Science that is done by comparing physics driven simulation to observation has no end point
  - One can always adjust a parameter add a second order effect etc.
- Science that processes data and tries to draw a conclusion tends to have an end of the iterations



Specialized Hardware

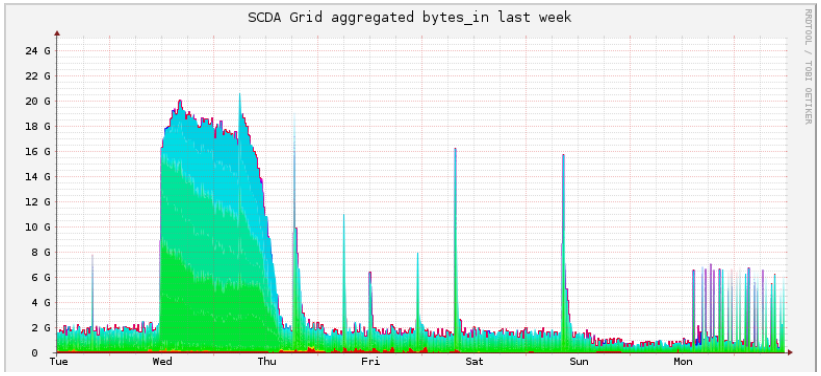
- We have seen some GPU requests
  - Mostly for canned machine learning applications
- The quantum system group has been hiring people who can actually program GPUs so we are watching if the requests go up



FLATIRON INSTITUTE

Dealing with Data

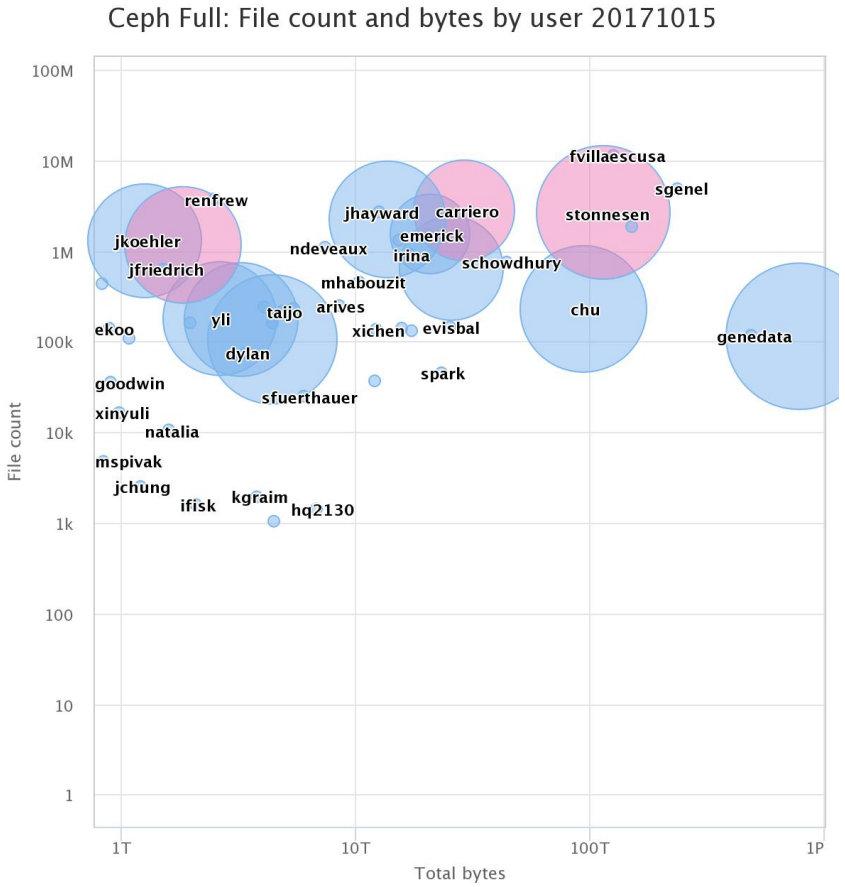
- We have been using cephfs
  - Currently we have almost 20PB of raw disk space
    - Mostly triple replicated in ceph
    - It's run with about 0.75FTE



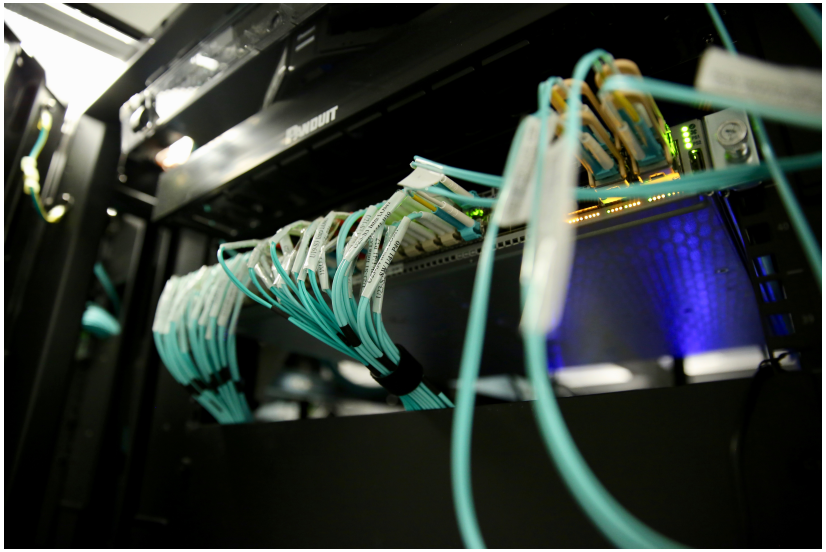
FLATIRON INSTITUTE

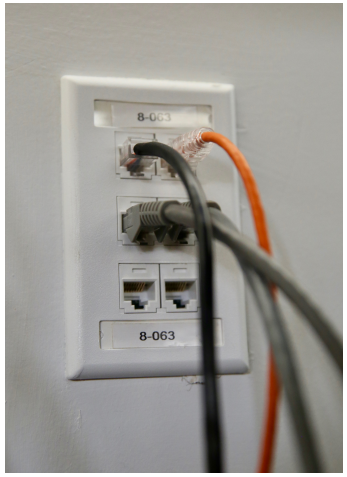
Dealing with Data

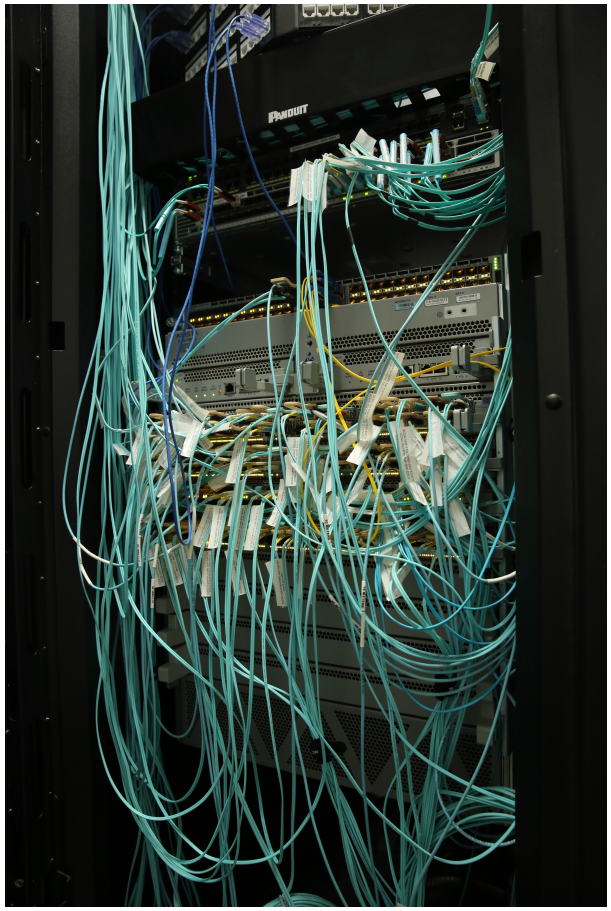
- About 50% of our space is used genomic sequencing data
  - 200GB per person in one file
- The astronomy simulation uses about 30%
  - Many millions of small files



Delivering Data







- Flatiron is a purpose built facility
  - All office floors and offices are optimized for data access

FLATIRON INSTITUTE

Archive

- If we had implemented our storage system on a non-rewritable media, we would not notice a difference in usage
  - Almost nothing is ever removed
  - We were ingesting data at 400TB a month during the Fall
- The mix of files ranges from millions of tiny files to thousands of 100GB files
  - Makes traditional tape solutions for HEP sub-optimal
    - We are looking for something that would slower and cheaper than our ceph system to check data into

FLATIRON INSTITUTE

Sharing

- A problem that has come up in all the centers is the desire to share and collaborate on big data samples
  - 100TB to 3PB sized
- We have seen a variety of tools, which generally fall into two camps
  - Tools designed to move groups of files around, which are good for replicating entire datasets but little knowledge of the contents
  - Tools that are so tightly coupled to the contents that they can select a subset of data, but couldn't be used for anything than the sample they were designed to handle

FLATIRON INSTITUTE

Tools

- We exposed people to bare globus-url-copy tools
  - Data was exported from the FNAL archive directly
    - Biggest hassle was handling grid certs
- We moved people to globus-online
  - Globus sharing tools delegate a lot of authorization to globus so fewer cert issues
    - People like it because it's easy and the performance has been OK
      - Has enough functionality to get people into trouble

FLATIRON INSTITUTE

Challenges of Sharing

- The biology community has PBs of datasets that they want to share with 40 labs
  - Chunks of data run into the hundreds of TB
    - Data samples went up by a factor of 20 in a couple of years
  - Most of the destinations are behind firewalls
  - Few of the places have any experience with moving large quantities of data
    - Transition to trusting the network hasn't happened yet

FLATIRON INSTITUTE

Challenges of Sharing

- The Astronomers have PBs of simulation that they want to share widely
  - It's some collection of files and often a database for the meta-data
  - Each simulation is it's own ecosysyem
    - Every project essentially gets it's own portal
    - Many of the portals are designed to process locally and give distributions
      - Rather than export data to process locally
    - Even in those that can export the information is chopped into new files

COLLABORATION

- Data access and management is an area we would like to get involved in with the the OSG
  - Data into stash cache, wide area access to samples, data federation for distribution
    - If we had this then expanded into OSG opportunistically would be interesting
  - Cross community discussion on data management
    - How to think about datasets, trade-offs of immutable files, moving samples around, peer serving of data

FLATIRON INSTITUTE

OUTLOOK

- Supporting multiple sciences is hard
- Nearly any amount of computing can be used
- I think there are a number of interesting areas for activity thinking about data distribution and access

FLATIRON INSTITUTE


FLATIRON
INSTITUTE