



<http://fasterdata.es.net/performance-testing/2019-2020-data-mobility-workshop-and-exhibition/>

CC* Data Movement Workshop and Exhibition - Introduction

Jason Zurawski, Eli Dart

zurawski@es.net, dart@es.net

ESnet / Lawrence Berkeley National Laboratory

Dr. Jennifer M. Schopf

jmschopf@indiana.edu

Indiana University International Networks

CC/CICI PI Meeting Pre-Workshop
September 22nd 2019*



ESnet

ENERGY SCIENCES NETWORK



INDIANA UNIVERSITY

Outline

- *Who are We, Who are You?*
- Why are We Here?
- Context, in 2 Minutes
- Expectations
- And Away We Go ...

Who are We, Who Are You?

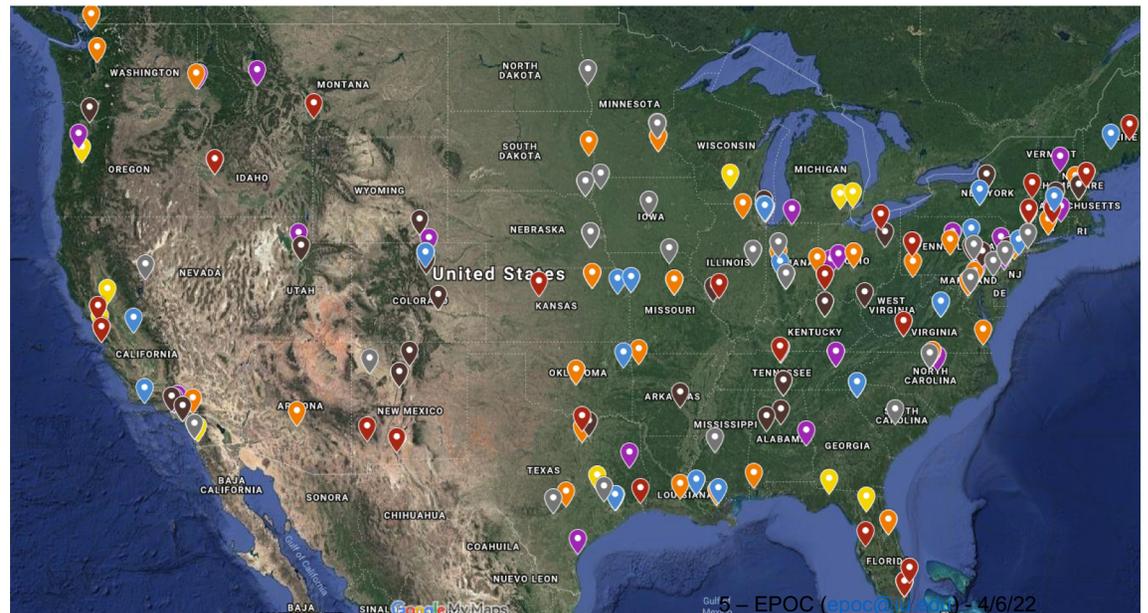
- Intros, etc.

Outline

- Who are We, Who are You?
- *Why are We Here?*
- Context, in 2 Minutes
- Expectations
- And Away We Go ...

This is CC*

- We “solved” technology support for the research community, right?
- Lets be honest, Kevin’s map is colorful:
 - CC* 250 Awards across 49 states/jurisdictions 2012-2018 (7 years)
 - \$100M+ investment
 - 130 (!) Campus Network upgrades
 - “More coming soon”



Network as Infrastructure *Instrument*



Connectivity is the first step – *usability* must follow



**To be cashed in @ Sept 2019 CC* PI Meeting, must be present to win.*

Measurable Outcomes

- What did the funding "do"?
 - "Build CI" – but was that all?
 - It was also supposed to improve scientific outcomes in some measurable way.
- Other Thoughts?
 - Did it all work as expected?
 - Are we all satisfied? (researchers are the survey population, not the IT org ...)
 - How do you know when you/we are done?
 - E.g. are you/we ever done?



An old Silicon Valley Chestnut

- "Move fast and break things. Unless you are breaking stuff, you are not moving fast enough."

- Mark Zuckerberg

- An alternative view:
 - E.g. understand the impact of technology on the people being forced to use it. Maybe ask that they are ok with the changes.

Tristan Harris @tristanharris
👍 We need a new field of "Society & Technology Interaction" (or STX) as @aza and I have said, since "Human-Computer Interaction" is limited to individual human + machine. Getting alignment right = "Humane Social Systems" (per invitation to @HumaneTech_ community in 4/23 event) twitter.com/metaviv/status...

Aviv Ovadya @metaviv
Everyone talks about UX: User Experience

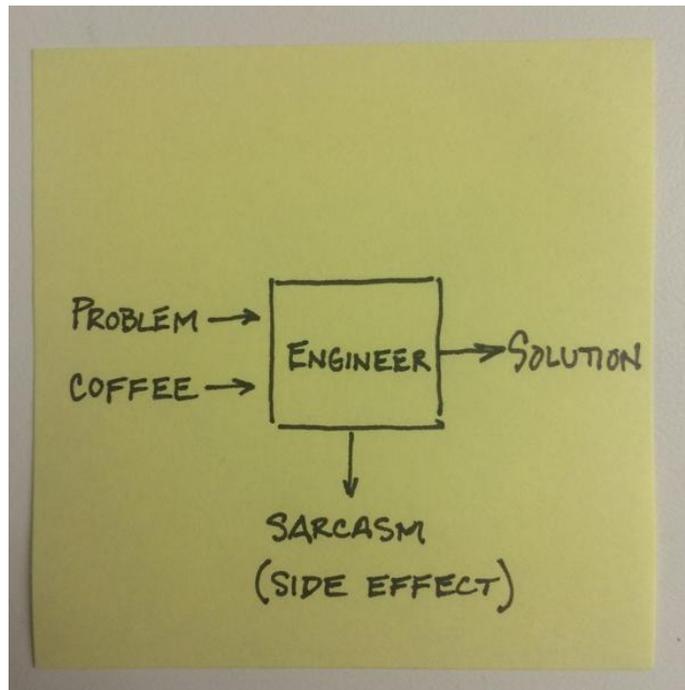
But what about CX?
- Collective Experience
- Community Experience

The usability and impact of a product across a *collection of people* within a society.

301 10:52 AM - Jun 10, 2019
207 people are talking about this



(Bad) Science Engagement in a Nutshell



(Bad) Science Engagement in a Nutshell



Move Slow, and Build Relationships

- Scientific Community needs (based on talking with real scientists):
 - **Stability**: What I use should continue to work tomorrow in the way it worked yesterday
 - **Predictability**: Low standard deviation from expectations
 - **Plug and Chug**: Fits the process of science, since that is what its all about
 - **'Make it So'**: If changes need to be made, do them in some way that doesn't impact the 3 items above.
- Technology implementation community
 - **'New and Improved'**: Change is frequent, disruptive, and 'expected'
 - **'Value Add'**: Optimizing on the wrong axis of need
 - **'I know what you need'**: Things you didn't know you needed, or wanted, in the sake progress

Outline

- Who are We, Who are You?
- Why are We Here?
- ***Context, in 2 Minutes***
- Expectations
- And Away We Go ...

The

Consi

- “Fri
-
-
-
-
- Dec
-
-
- Per
-
- Eng

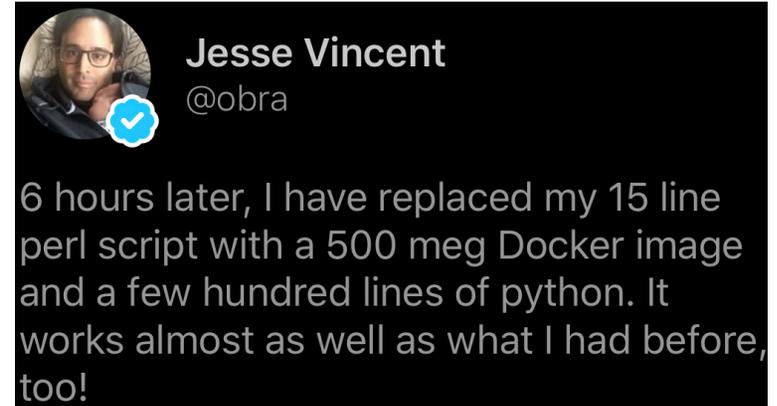


NAR

z/

My friend, I'll say it clear (e.g. JZ Rant™)

- Building CI for the sake of spending money/doing fun upgrades is folly
- Unless you F*#@d up at least once (and will admit it), the project didn't succeed



- Tearing down and building again should be an expectation – this is how you learn what is really needed. No one can “Carnac the Magnificent™” the answer*.
- Its not a success until the researchers are happy
 - “Is this what you want/is this useful” should be step 0 before a PO is issued on anything
- “Is it being used”? “Why or why not”?

* Maybe Matt Z, but he's usually busy.

I'll state my case, of which I'm certain (</rant>)

- A large population of CC * sites miss the mark (*prove me wrong*):
 - Network Infrastructure. I bet that 'data center' switch was super cheap; good thing its in front of your DTN.
 - Computation. Who went out and asked the researcher(s) how they computed? If the answer was 'on my laptop' was that addressed?
 - Storage (!) I will just leave this here.
 - Science Engagement: Touching base with the scientists/research community should be routine and regular – not just at the start and the end of the build
 - Research Office: Not just mechanism to justify the grant, the pipeline for users of the infrastructure. You should be talking to them as frequently as the known users. What you are building is a strategic asset for the institution.
 - “Ease of use”
- Are we better off?



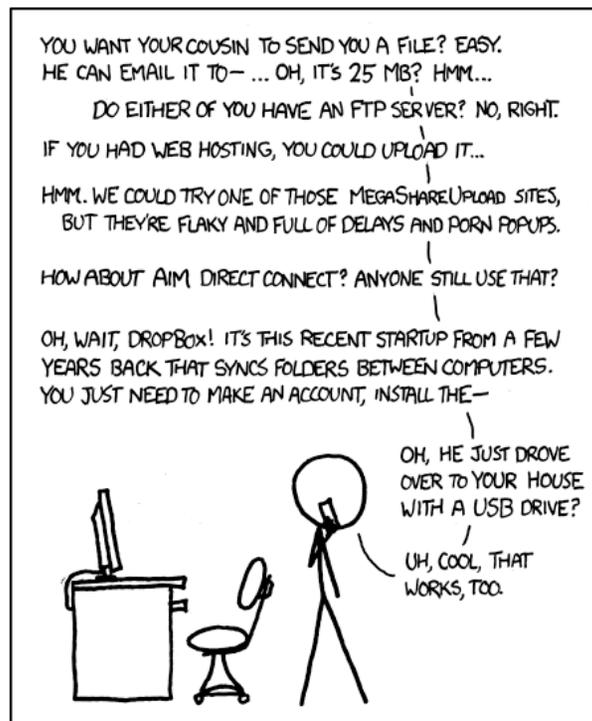
Outline

- Who are We, Who are You?
- Why are We Here?
- Context, in 2 Minutes
- ***Expectations***
- And Away We Go ...

Regrets, I've had a few

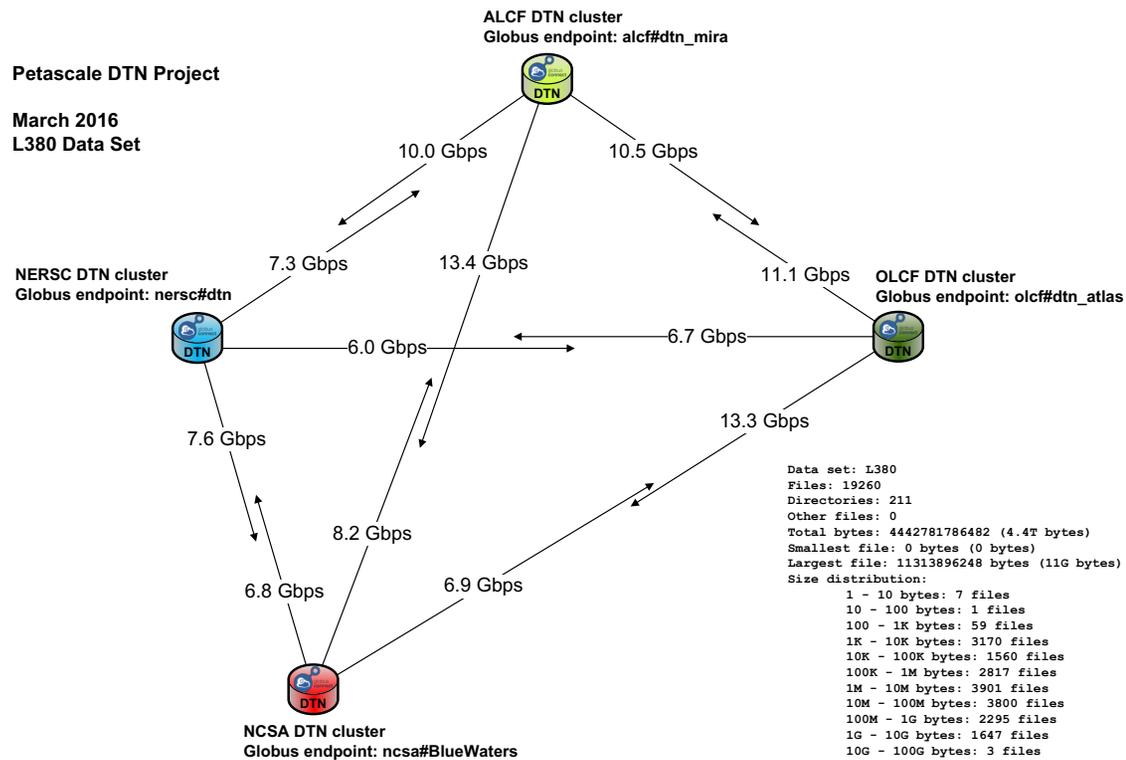
- We need to redefine the mission, approach, solutions, and success metrics
 - Supporting science has to be a social conversation between tech and science
 - “Look at this cool thing” isn’t helpful, it distracts from establishing trust and building useful infrastructure
 - **Core mission**: handling data, from creation to curation (and all steps in the middle)
- Why this, why now?
 - We are not all the way there.
 - We have the foundation, but we’re not done yet

I did what I had to do



I LIKE HOW WE'VE HAD THE INTERNET FOR DECADES, YET "SENDING FILES" IS SOMETHING EARLY ADOPTERS ARE STILL FIGURING OUT HOW TO DO.

I faced it all and I stood tall

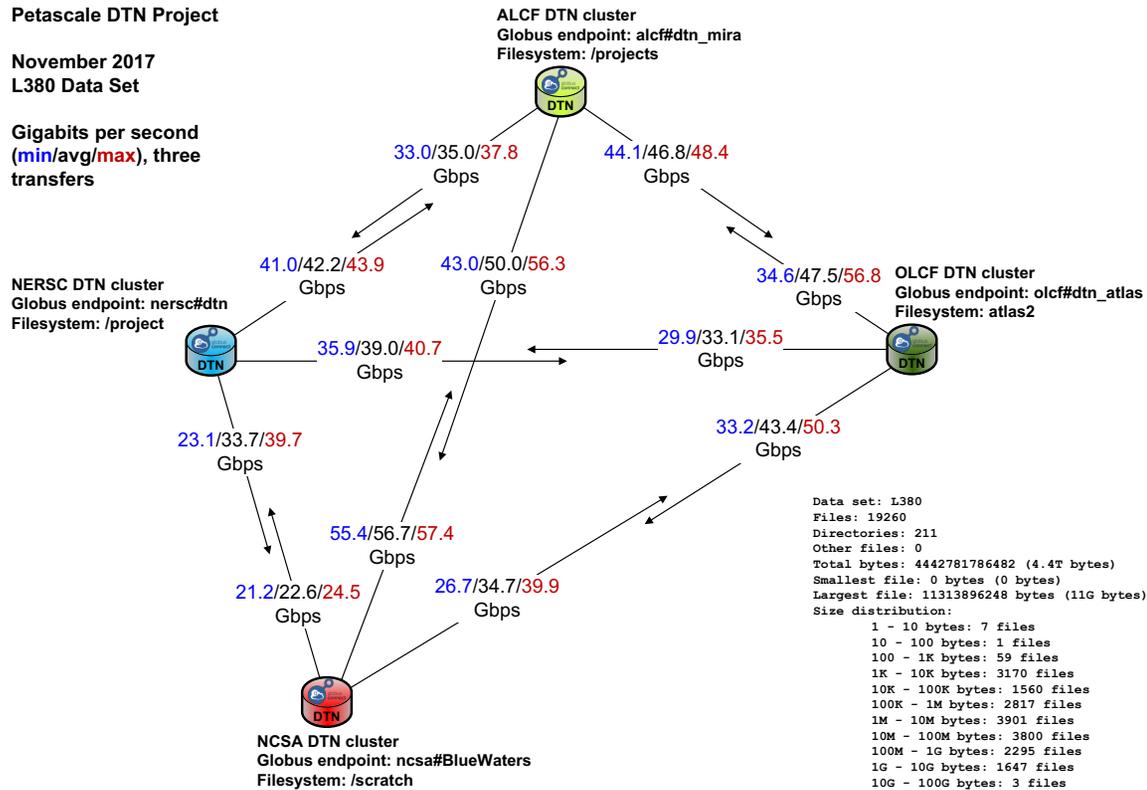


And did it my way (even if it took a year)

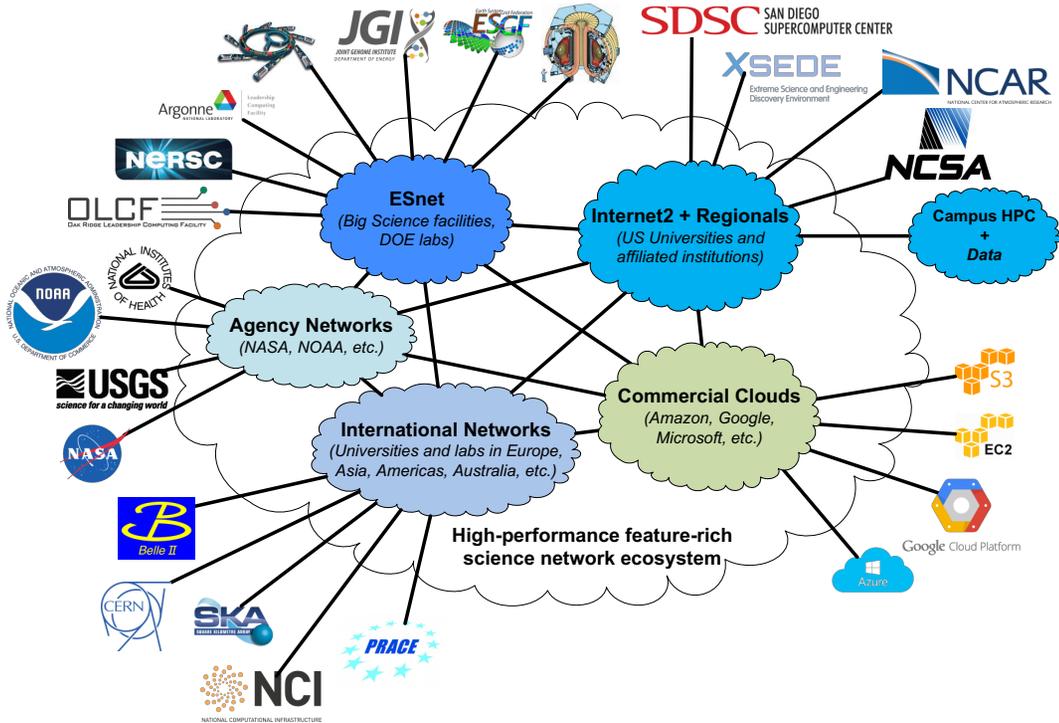
Petascale DTN Project

November 2017
L380 Data Set

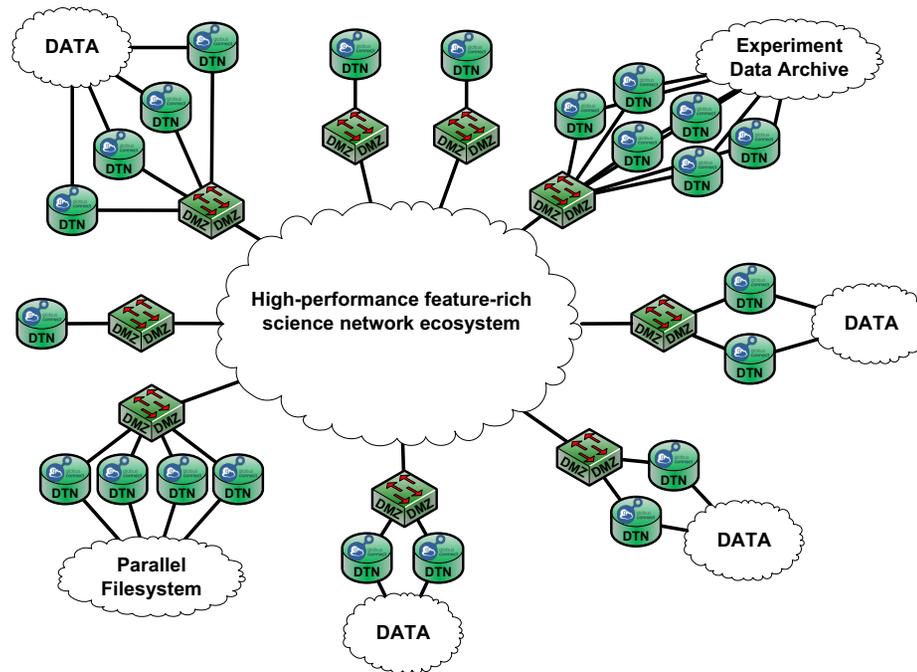
Gigabits per second
(min/avg/max), three transfers



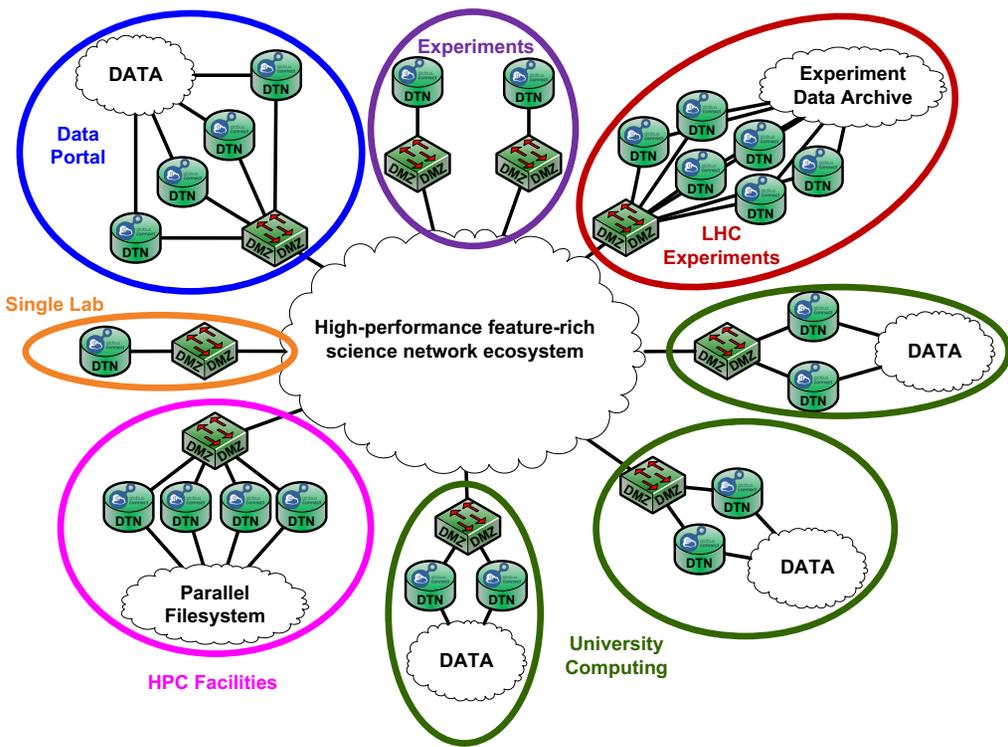
Our World



Our World (Through Eli's Eyes)

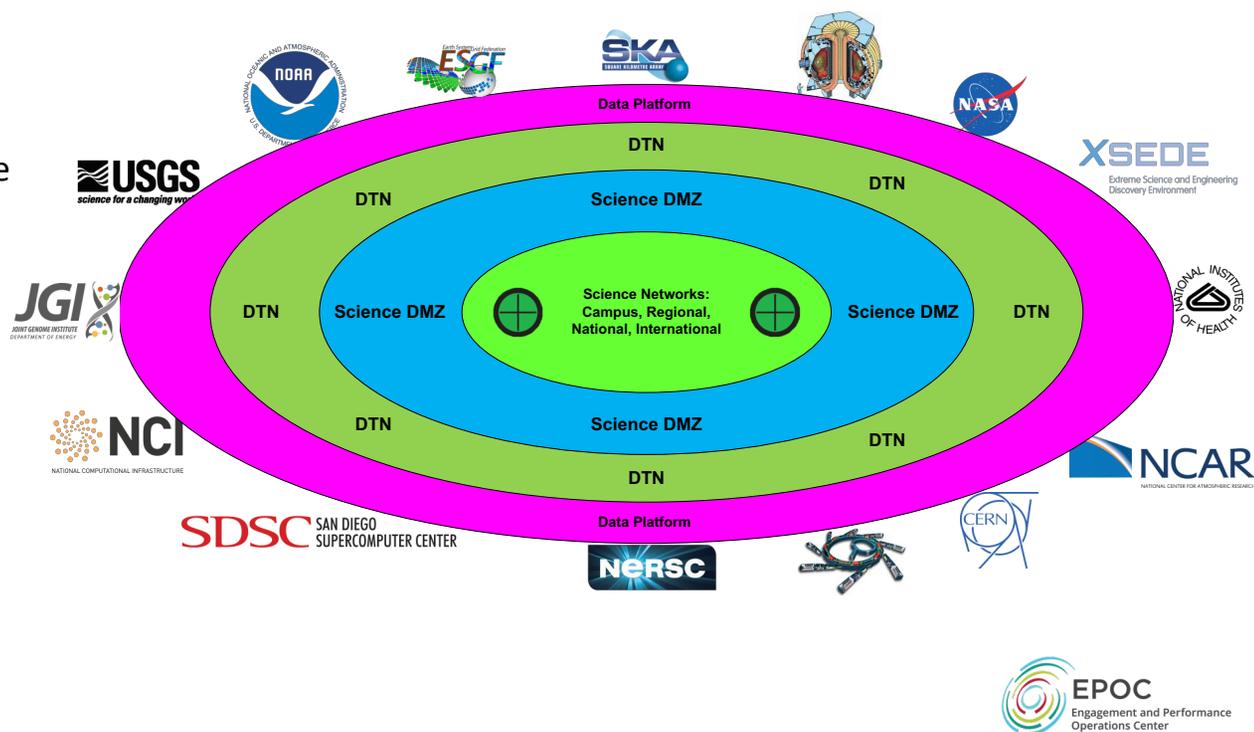


Our World (Where we need to be)



Platform View Of Our World

- Interoperable computing and data
- Scientists should not have to reason about anything other than the platform layer
- Science DMZs connected to fast networks are necessary but not sufficient
- Need an interoperable data orchestration platform too



Networks Cannot Do This Alone

- We need a whole-community effort
 - Networks
 - HPC facilities
 - Data repositories / Data portals
 - Experimental facilities
 - Campus CI
 - Science collaborations
 - Science programs
- Networks can help, and must be part of the conversation
 - Heavy lifting is now at the network edge, in collaboration with the network core
 - Interoperability and usability are key



Outline

- Who are We, Who are You?
- Why are We Here?
- Context, in 2 Minutes
- Expectations
- *And Away We Go ...*

Let's Do it "This Way"

- What is the ask of all of you?
- Three-part event at the upcoming NSF CC and CICI PI meeting in September of 2019:
 - ***An introductory workshop – Level the Playing Field***
 - ***An advanced workshop – Integrate the technology with the use via the software stack.***
 - A data movement exhibition. Lets prove that we (as a community) can improve data movement outcomes.

Workshop(s)

Agenda

Time	Topic	Instructor(s)	Expected Outcomes
1:00pm	Arrivals, Informal Discussion		
1:00pm – 1:15pm	Welcome & Introduction	All	Set context for day/event
1:15pm – 1:45pm	Science DMZ Architecture/Security & Context	TBD <u>ESnet</u> /EPOC	Background on scientific network use requirements and approaches
1:45pm – 2:15pm	TCP & Performance Monitoring	TBD <u>ESnet</u> /EPOC	TCP, performance testing, and how it relates to DMZs
2:15pm – 2:45pm	Data Transfer Hardware	TBD <u>ESnet</u> /EPOC	Construction/use of data movement infrastructure
2:45pm – 3:15pm	Program Break (30 Min)		
3:15pm – 3:45pm	Discuss Data Movement Exhibition	All	Talk about the purpose and expected outcomes
3:45pm – 4:15pm	Globus Functional Overview	TBD Globus	Targeted more toward early users, with some down level information
4:15pm – 4:45pm	Overview of the Globus Platform	TBD Globus	emphasis on the CLI and automation (e.g. instrument data egress)
4:45pm – 5:15pm	Building Modern Research Data Portals (MRDPs)	TBD Globus	emphasis on <u>Auth</u> and integrating Globus and other custom services, e.g. in a data distribution use case
5:15pm – 5:30pm	Conclusions, Next Steps	All	Solidify future plans
5:30pm	Adjourn		

Intro Workshop

- Think 'Old OIN', except shorter and focused purely on the higher level ideas.
 - Some DMZ, some DTN, some perfSONAR
- Goal is for brand-new PIs (and their engineers) to develop a vocabulary and learn the landscape
- Materials will be available offline, and can be used back at home institutions
- Most likely will also show up as a webinar, here, in the future.

Advanced Workshop

- Partnering with Globus, we want to provide a 'blueprint' for the development of Modern Research Data Portals (MRDPs)
- Go over the ins/outs of the Globus Software, APIs
- Learn how the software can be constructed
- Learn how the data sets from research groups can be integrated
- Learn how you can validate/verify the performance results.

Data Movement Challenge

- More Later ...



<http://fasterdata.es.net/performance-testing/2019-2020-data-mobility-workshop-and-exhibition/>

CC* Data Movement Workshop and Exhibition - Introduction

Jason Zurawski, Eli Dart

zurawski@es.net, dart@es.net

ESnet / Lawrence Berkeley National Laboratory

Dr. Jennifer M. Schopf

jmschopf@indiana.edu

Indiana University International Networks

CC/CICI PI Meeting Pre-Workshop
September 22nd 2019*



ESnet

ENERGY SCIENCES NETWORK



INDIANA UNIVERSITY