Cyberinfrastructure for Major Facilities Workshop

Getting Together, Working Together

In person in Redondo Beach, CA and Online

Ewa Deelman
University of Southern California
CI Compass, PI

March 1, 2022
Cyberinfrastructure for Major Facilities Workshops

2022: Getting together, working together

- Making data FAIR
- Cloud migration
- Future of CI for MFs
- CI workforce:
  - Developing and retaining talent
  - Developing resilience
**2017 NSF Workshop on CI for Large Facilities**

- **Establish a center of excellence** (following a model similar to the NSF-funded Center for Trustworthy Scientific Cyberinfrastructure, CTSC) as a resource providing expertise in CI technologies and effective practices related to large-scale facilities as they conceptualize, start up, and operate.

---

**“NSF’s Blueprint for a National Cyberinfrastructure Ecosystem” April 2019**

Funded by NSF under grant #1742969
NSF Major Facilities rely on complex cyberinfrastructure (CI) to transform raw data into more interoperable and integration-ready data products

49 MF + mid-scale RI-2 participants
14 NSF PDs
12 Students
NSF CI Ecosystem From the CI Calling Cards

32 Participants
Mission

CI Compass provides expertise and active support to cyberinfrastructure practitioners at NSF Major Facilities in order to accelerate the data lifecycle and ensure the integrity and effectiveness of the cyberinfrastructure upon which research and discovery depend.
Overall CI Compass Strategy

1. Recognize the expertise, experience, and mission-focus of Major Facilities
2. Contribute knowledge and expertise to the MF Data Lifecycle (DLC) CI and enhance the overall NSF CI ecosystem
3. Build expertise, not software - Leverage existing knowledge, tools, community efforts
4. Build partnerships to leverage community expertise
5. Broker connections and share knowledge, lessons learned, best practices with MFs, Partners, CI community

Partners: Trusted CI, OSG/PaTh, SGCI, EPOC, RCD Nexus, Fabric, Chameleon

Advisory Committee: S. Anderson (LIGO), A. Bolton (NOIRLab), B. Hurwitz (UofA), M. Livny, (OSG/PaTh) V. Welch (Trusted CI) M. Zentner (SGCI)
CI Compass Team: Who we are

Deep expertise in several CI areas critical to the MFs
- Data management, data processing, visualization, archiving, semantic technologies
- Automation, resource management, workflows, sensors
- Networking, clouds, systems and infrastructure
- Large-scale CI deployment and operations, IdM
- Social science

Experience in the management of CI projects
- Conceptualization, from design phase to broad adoption
- Project Management and Evaluation
- Organizational science
- Communications & Outreach

Highly collaborative, strong history of working together
- Many diverse community connections in astronomy, earth science, physics

Dedicated to the advancement of CI for science, engineering, and education
CI Compass services: What we can help with.

Evaluate CI Plans, Help Design New Solutions, Develop Proofs of Concept, Assess Applicability/Performance of Existing Solutions, Help Leverage CI Solutions

**Data Capture**: Discuss sensor data annotations, help apply community sensor data models to the MFs data, explore messaging systems, support interoperability across MFs, and connect to industry standards and to similar communities of practice within scientific domains.

**Central Processing**: Help leverage existing testbeds to evaluate new software stacks and configurations during CI design/enhancements.

**Data access, dissemination & visualization**: Assist in designing data visualization solutions (tools, formats, proofs of concept, etc.) for enhancing data discoverability and accessibility.

**Identity management**: Assist in developing IdM solutions for managing user data access to data, help in using IdM for data usage reporting and tracking.

Offered services’ list (non-exhaustive) is published on CI Compass website: https://ci-compass.org/assets/453214/cicompass_services_table.pdf/
CI Compass activities: What we do.

Engagement

Topical Working Groups

Workshops

MF CI & Topical

Student Program

Other

Webinars, surveys, reports

Internal

Planning, project mgmt

Funded by the National Science Foundation, Grant #2127548
Examples of 1-1 Engagements

<table>
<thead>
<tr>
<th>Major Facility</th>
<th>NEON / NCAR</th>
<th>ARF/RCRV</th>
<th>ARF</th>
<th>NOIRLab</th>
<th>SAGE / GAGE</th>
<th>Arecibo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Exchange, Clouds, Visualization

Shipboard CI/network plan review

Identity Management

Identity Management

Cloud platform design

Data management

For details of these engagements, please visit our website [https://ci-compass.org/](https://ci-compass.org/) and view Anirban’s webinar from February 2022.

TACC’s Ranch supercomputer, a long-term data mass storage system, is safely preserving over three petabytes of data from the Arecibo radio telescope. Ranch is an allocated resource of the Extreme Science and Engineering Discovery Environment (XSEDE) funded by the National Science Foundation (NSF). (Credit: TACC)
CI Compass activities: Topical Working Groups

Identity Management Topical WG
Disseminate IdM information
- Quarterly meetings with speakers and discussions on topics relevant to MFs: e.g. CILogon
- Issues of identifying data usage and enabling reporting

Cloud Infrastructure Topical WG
- Understand the current practices for clouds used by MFs
- Research alternative solutions and keep up to date with emerging cloud technologies
- Develop a general set of best practices that can inform the MFs

Send email to engage@ci-compass.org to participate in any of these or future Topical Working Groups or propose new ones!

Organizing a panel tomorrow!
Let’s Work Together

To explore potential opportunities for an engagement with CI Compass regarding a CI issue you might have, please email engage@ci-compass.org

To respond to our Call for Engagement, please click the following link: https://ci-compass.org/services/apply-for-engagement-with-ci-compass/

Please feel free to reach out to us, we are here to work with you!
Student Internship Program (Pilot)

Undergraduate Students in CS
- 5 at USC, 1 at UND
- Spring with CI Compass directed research
- Summer Internships: working on projects, ultimately with MFs

Led by Angela Murillo, Laura Christopherson, Jarek Nabrzyski

Directed by Ryan Tanaka, Loïc Pottier

Check out their calling cards!

February 28, 2022, Topic: Software Best Practices II, Containers!
Lesson Topics

- data
- research
- distributed
- practices
- notebooks
- workflows
- jupyter
- major
- development
- fair
- presentations
- professional
- computing
- parallel
- software
- cloud
- best
- lifecycle
- python

Research lectures notes

Help us connect with you and your programs: contact@ci-compass.org
Workshop Goals:

Bring together the MFs, CI community, and NSF to

- Share best practices in use of CI
- Discuss opportunities in leveraging the NSF CI ecosystem
- Brainstorm solutions to the challenges we face today and in the future

Outcomes:

- A webinar summarizing the workshop discussions (including breakouts) **April 4, 2022**
- A report summarizing the findings and recommendations, including the analysis of the CI Calling Cards -- thank you!
- We will reach out during and after the workshop to those that indicated interest in talking with us
## Tuesday’s Agenda

<table>
<thead>
<tr>
<th>Time (PST)</th>
<th>Session Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-8:30</td>
<td>Workshop Welcome and Overview of CI Compass</td>
</tr>
<tr>
<td>8:30-9:15</td>
<td>NSF’s Perspectives on Cloud Computing</td>
</tr>
<tr>
<td>9:15-9:35</td>
<td>Thematic Lightning Talks (Cloud, CI Challenges)</td>
</tr>
<tr>
<td>9:35-9:55</td>
<td>Brainstorming: Developing and Retaining a Vibrant Team of Workers (for day 2)</td>
</tr>
<tr>
<td>9:55-10:00</td>
<td>Group Zoom picture</td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>Break</td>
</tr>
<tr>
<td>10:30-11:00</td>
<td>NSF CI Ecosystem Lightning Talks</td>
</tr>
<tr>
<td>11:00-12:00</td>
<td>Developing Resilience and Managing Uncertainty during the Pandemic (Breakout)</td>
</tr>
<tr>
<td>12:00-1:30</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>1:30-3:15</td>
<td>Making the Major Facilities Data Lifecycle FAIR to Provide AI-Ready Data (Group discussion)</td>
</tr>
<tr>
<td>3:15-3:45</td>
<td>Office Hours (remote only)</td>
</tr>
<tr>
<td>4:00-6pm</td>
<td>Reception - in person</td>
</tr>
</tbody>
</table>

Please contribute notes, comments [https://tinyurl.com/ci4mfs2022](https://tinyurl.com/ci4mfs2022)
<table>
<thead>
<tr>
<th>Time (PST)</th>
<th>Session Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00-10:00</td>
<td>Cyberinfrastructure for Major Facilities: Challenges and a path forward (breakout)</td>
</tr>
<tr>
<td>10:00-10:30</td>
<td>Break</td>
</tr>
<tr>
<td>10:30-12:00</td>
<td>The March Toward the Clouds: MF Perspectives (Topical Cloud WG)</td>
</tr>
<tr>
<td>12:00-1:20</td>
<td>Lunch Break</td>
</tr>
<tr>
<td>12:45-1:20</td>
<td>Office Hours (online only) Ilya Baldin, Ewa Deelman, Anirban Mandal, Valerio Pascucci</td>
</tr>
<tr>
<td>1:20-2:20</td>
<td>Brainstorming: Developing and Retaining a Vibrant Team of Workers (breakout)</td>
</tr>
<tr>
<td>2:20-3:00</td>
<td>NSF CI Ecosystem Lightning Talks</td>
</tr>
<tr>
<td>3:00-3:05</td>
<td>Closing</td>
</tr>
<tr>
<td>3:05-4:00</td>
<td>Social Coffee Hour - in person</td>
</tr>
</tbody>
</table>
Recording

- The whole workshop is being recorded
- Only formal presentations will be edited and posted online
- Other recordings will be used to clarify notes
Virtual Breakout Group Assignments
(For all breakouts except those for the Brainstorming session.)

<table>
<thead>
<tr>
<th>Breakout Room 1</th>
<th>Breakout Room 2</th>
<th>Breakout Room 3</th>
<th>Breakout Room 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brad Barber</td>
<td>Maggie Benoit</td>
<td>Brian Dobbins</td>
<td>Kathy Benninger</td>
</tr>
<tr>
<td>Devan Bougie</td>
<td>Adam Brazier</td>
<td>James Edson</td>
<td>Bruce Berriman</td>
</tr>
<tr>
<td>Doug Ertz</td>
<td>Jerry Carter</td>
<td>Jeremy Fischer</td>
<td>Alisdair Davey</td>
</tr>
<tr>
<td>Jeannette Dopidehe</td>
<td>Lauren Clay</td>
<td>Stephen Jacobs</td>
<td>Lee Ellett</td>
</tr>
<tr>
<td>Ken Feldman</td>
<td>Jeannette Jackson</td>
<td>Bogdan Mihaila</td>
<td>James Holik</td>
</tr>
<tr>
<td>Sharon Broude Geva</td>
<td>David Kratz</td>
<td>Giri Prakash</td>
<td>Miron Livny</td>
</tr>
<tr>
<td>Margaret Levenstein</td>
<td>Greg Madden</td>
<td>Roland Roberts</td>
<td>Margaret Martonosi</td>
</tr>
<tr>
<td>Kevin Porter</td>
<td>Brandi Murphy</td>
<td>Dan Stanzione</td>
<td>Craig Risien</td>
</tr>
<tr>
<td>Paul Redfern</td>
<td>Joy Pauschke</td>
<td>Alex Szalay</td>
<td>David Schultz</td>
</tr>
<tr>
<td>Jennifer Schopf</td>
<td>John West</td>
<td>Von Welch</td>
<td>Adam Shepherd</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Breakout Room 5</th>
<th>Breakout Room 6</th>
<th>Breakout Room 7</th>
<th>Breakout Room 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Christopher Cameron</td>
<td>Kay Avila</td>
<td>Steve Barnet</td>
<td>Peter Beckman</td>
</tr>
<tr>
<td>Robert Casey</td>
<td>Shafaq Chaudhry</td>
<td>Philip Gates</td>
<td>Gordon Bonan</td>
</tr>
<tr>
<td>Dru Clark</td>
<td>Chris Davis</td>
<td>John Haverlack</td>
<td>Patrick Brady</td>
</tr>
<tr>
<td>Josh Drake</td>
<td>Clint Dawson</td>
<td>Mark Krenz</td>
<td>Mark W. Coles</td>
</tr>
<tr>
<td>Jeff Glatstein</td>
<td>David Halstead</td>
<td>Allen Pope</td>
<td>Douglas Fils</td>
</tr>
<tr>
<td>Jeff Hoch</td>
<td>Christine Laney</td>
<td>Kelli Shute</td>
<td>Terry Fleury</td>
</tr>
<tr>
<td>Dawn Lenz</td>
<td>Sarah Ruth</td>
<td>Laura Stolp</td>
<td>Nirav Merchant</td>
</tr>
<tr>
<td>Richard Oram</td>
<td>Mike Simpson</td>
<td>Werner Sun</td>
<td>Eric Palm</td>
</tr>
<tr>
<td>Vladimir Papitashvili</td>
<td>Bob Tawa</td>
<td>Ross Thomson</td>
<td>John Palm</td>
</tr>
<tr>
<td>Benedikt Reidel</td>
<td></td>
<td>Michael Zentner</td>
<td></td>
</tr>
</tbody>
</table>

The questions will be the same for each breakout room.

During breakouts, please go to your assigned room.
We have note takers, but please add your thoughts!

No report back right after breakout.

tinyurl.com/ci4mfs2022
Workshop materials

Welcome packet, Notes, Slides, Calling cards: https://tinyurl.com/ci4mfs2022

Please contribute to discussions and notes!

To ask questions or make a comment

- Please raise your hand in Zoom or use Zoom chat
- The moderator will call on you

Please do not use zoom chat for other discussions as this makes it hard to see the questions.
To chat with other participants
You can view session in Whova or popped out into Zoom

To get Help

1. You can view session in Whova or popped out into Zoom

2. Email ci4mfs@ci-compass.org
Best Home Office Companion Photo Contest

- **Before COB Tue**, post your scaly/feathery/furry work-from-home companion photo
- **Before lunch break on Wed**: vote for photos you like best: click the thumbs up icon
- *Winners announced after lunch*

The 3 photos with the most votes win! Prizes are a gift certificate, a box of chocolates, or a CI Compass T-shirt.
Whova Ice breakers

Self introduction
Hi all. My ice breaker is that I literally just got off a short 10 day cruise on an actual ice breaker. Granted no ice was broken during...

John Haverlack
IT Manager

Questions?