Funded by the U.S. National Science Foundation, Grant #2127548

INTERNET2 COMMEX 25

APRIL 28-MAY 1
ANAHEIM CA

Cloud Adoption Survey in NSF funded Major Facilities and Mid Scale Institutes

Karan Vahi vahi@isi.edu
(CICompass Cloud Working Group)

May 1st, 2025

What are NSF Major Facilities and Mid-scales ?

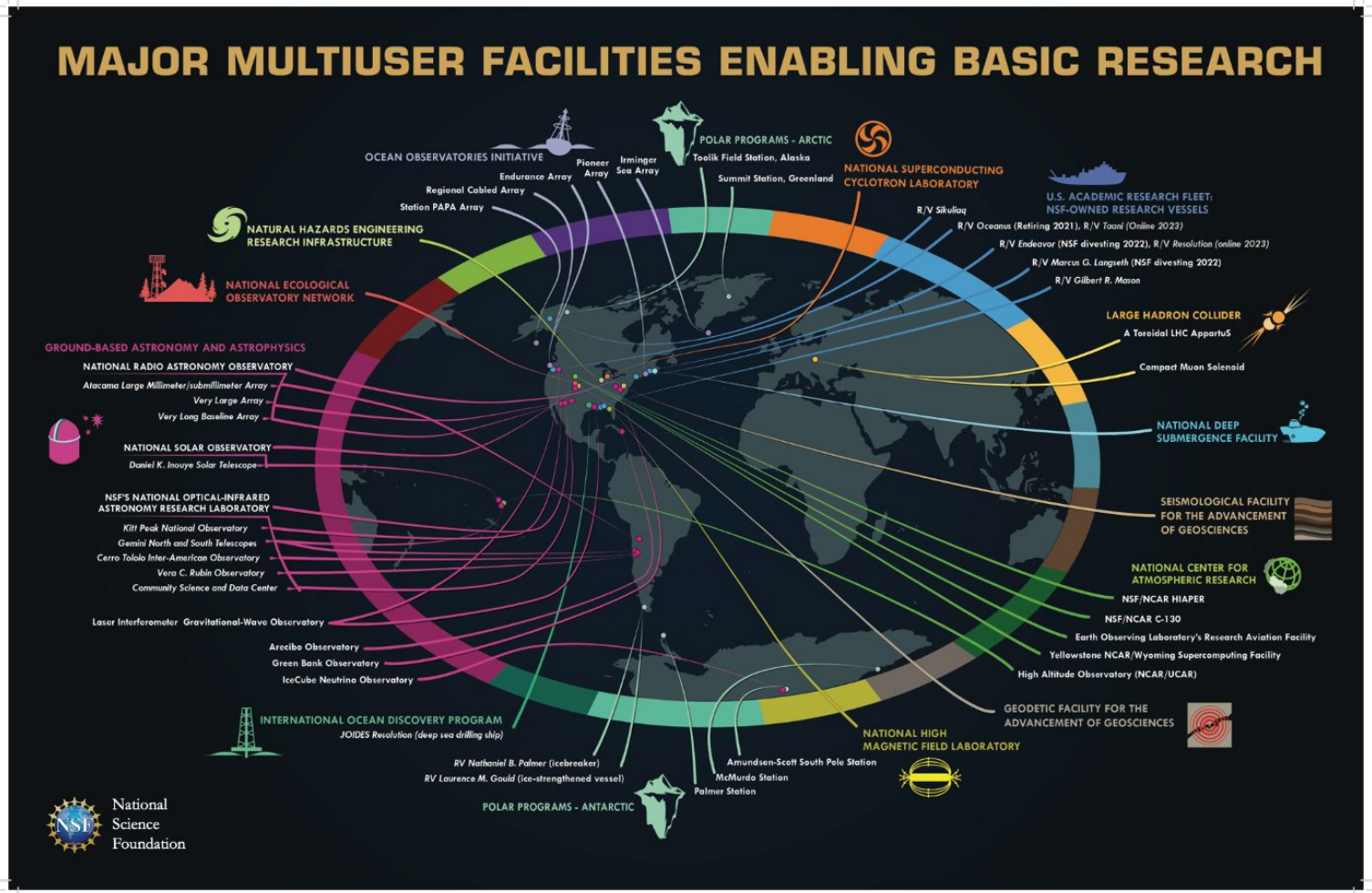


Image credit: National Science Foundation

Major Facilities (MFs) and Mid-scale Research Infrastructure (MRIs) are defined as *"shared-use infrastructure, instrumentation and equipment that are accessible to a broad community of researchers and/or educators. These facilities are generally intended to serve the science community that is critical to supporting innovation across the nation"*.

NSF Major Facilities **rely on complex cyberinfrastructure (CI)** to transform raw data into more interoperable and integration-ready data products.

Managed by NSF Research Infrastructure Office (RIO).

Mission

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

ci-compass.org

Funded by the U.S. National Science Foundation, Grant #2127548

USC Viterbi School of Engineering Information Sciences Institute · INDIANA UNIVERSITY · UNIVERSITY OF NOTRE DAME · THE UNIVERSITY OF UTAH · TEXAS TECH UNIVERSITY · renci

Topical Working Group: Cloud

Background:

- Started in 2020, to engage with Major Facilities about their Cloud usage.
- Need was recognized based on the CI CoE Pilot (Compass precursor) did with Major facilities about their data lifecycle

Charter

- Understand the current practices for Cloud Infrastructure 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 and how those can be adapted for specific facilities.

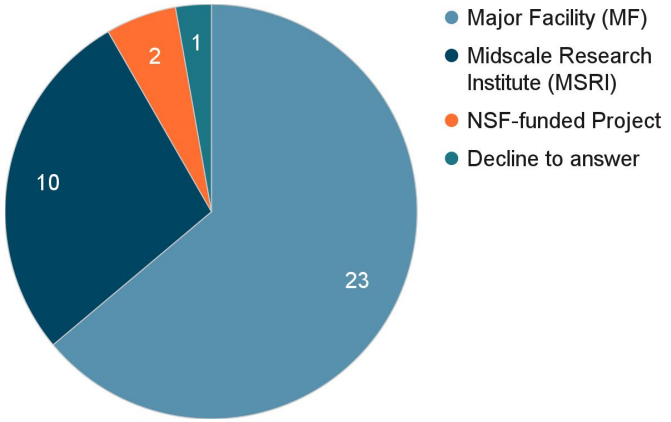
NSF Major Facilities Cloud Use Cases and Considerations

- Published January 2024 https://zenodo.org/records/10481410
- Delves into various Cloud use cases for Major Facilities such as increased availability ; scalability , capacity and flexibility.
- Features case studies, such as those NCAR, IceCube, OOI and SAGE GAGE facilities.
- Overview of cloud providers and spectrum of deployment options available.
- Explores challenges such as cost considerations and training of the workforce.

2024 Cloud Survey

NSF-funded Major Facilities (MFs) and Mid-scale Research Infrastructure (MRIs) handle vast amounts of data (i.e., Petabyte or PByte+ scale), requiring significant storage, processing, and serving resources.

- Published March 2025
  - https://zenodo.org/records/15192261
- Data collected September 2024
- Sent to 288 personnel across 21 MFs and 10 MSRIs,
- yielding 36 responses from 14 facilities and 7 mid-scales.
- **Goals**
  - Assess their current and future use of cloud services
  - Identify barriers to adoption.



Pie-chart representing the affiliations breakdown of the survey respondents.

Methodology

Questions asked to respondents were around

- their role in the facility
- perceived benefits of cloud adoption for their organization
- challenges in cloud adoption including cost concerns
- identification of cloud providers (commercial and academic both)
- share their cloud experiences

Role in Facility

Most respondents identified as technical leaders or managers

Role	Count
Project Management	6
Data Manager	**2**
DevOps / IT	5
Data Scientist	5
Software Engineer	3
Domain Researcher	1
Technical Director / Technical Leadership/ Product owner (architect, chief engineer, chief tech office, data services lead, software and computing coordinator, IT Director , leadership, Technical product owner)	**7**
Management (Director, Director of operations, Leadership, IT Director, PI)	**6**
Systems Engineer	1

Cloud Benefits - What benefits do you think the cloud provides/could provide for your facility?

Benefit	Choice Count
Increased reliability and availability	23
Better options for storage and archiving	23
On-demand use	23
Dynamic Scalability	21
Flexibility	20
Rapid innovation	15
Other, please specify	9

- Most respondents agree that cloud adoption provides tangible benefits
- Much lower proportion of respondents (43%) agreed that cloud adoption spurs rapid innovation

- **Other Responses**
  - highlighted better physical and technical security.
  - emphasized the availability of Kubernetes and the ability to do Infrastructure as Code (IaC) deployments.
  - one of the respondents responded that they were unconvinced about whether there were any benefits.

Scope of Cloud Usage

Scope	Number
None	3
Not anymore (we have used it in the past, but currently not used)	3
Curious (don't use Cloud currently, but curious about possibilities)	0
Experimental (exploring, but nothing in production)	3
Emerging (small amount of cloud usage)	10
Hybrid	10
Cloud First	6

- 26 respondents (74%) indicate that their organizations are using cloud for part or all of their operations
- 46% (16 respondents) characterize their organizations using the cloud in a meaningful way, either as hybrid or being Cloud First
- Worth noting that 6 respondents (17%) from 5 different organizations indicated their facilities are Cloud First

Integrate the Cloud - If you are not already a cloud user, do you plan to integrate cloud for your operations in the future?

Already in the process of doing it	10
In the current funding cycle	0
In the next funding cycle	0
Long Term	9
Not at all	4
Not applicable	12
Total	35

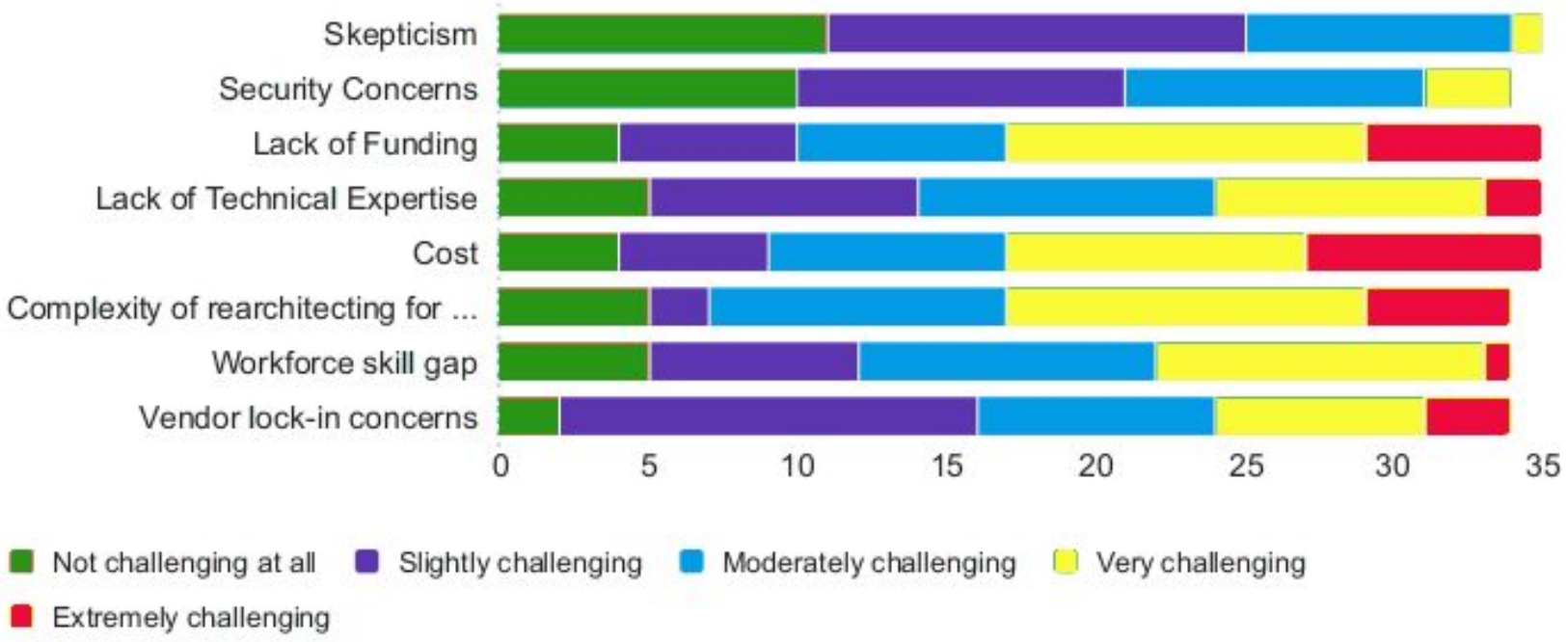
- "Not applicable" indicates that the respondents' facilities are already cloud users. Removing those leaves 23 responses. Of those,
- 10 (43%) respondents indicated that they are already in the process of adopting cloud,
- 9 (39%) have plans to do it in the long term

Cloud Challenges - How challenging are the following in relation to using cloud resources?

Respondents were asked to rate the following concerns

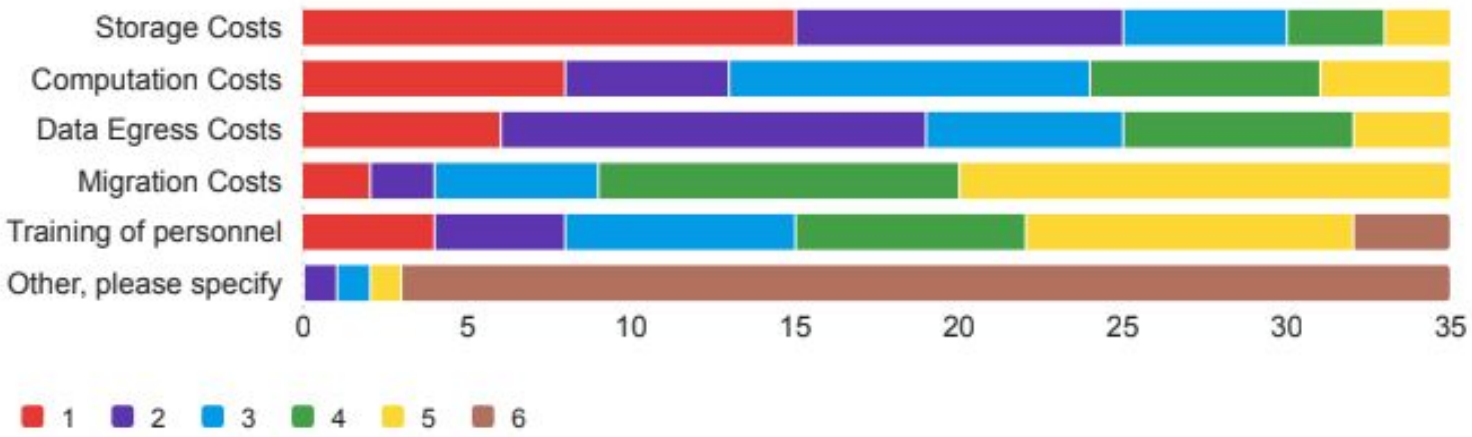
- Skepticism
- Security Concerns
- Lack of funding
- Lack of Technical Expertise
- Cost
- Complexity for rearchitecting for migration
- Workflow Skill gap
- Vendor Lock in concerns

Cloud Challenges



- Most respondents see the value of cloud adoption, as indicated by the low level of cloud skepticism
- “Lack of Funding”, “Complexity for rearchitecting their deployments,” and a “Workforce Skills Gap/ Lack of Technical Expertise” are significant barriers to Cloud adoption
- Costs are a significant challenge

Cost Concerns Ranked



Note One indicates the most concerning, and 6 the least concerning.

- MFs and MSRIs handle vast amounts of data, reaching the PByte+ scale
  - cost concerns related to storage, compute, and data egress are quite understandable
  - storage stands out, with nearly half of the respondents (43%) ranking it as their top concern.
- Migration and training costs remain important to respondents
  - secondary to the ongoing operational costs associated with storage, compute, and egress

Cost Concerns - If commercial cloud providers were to make data egress free, would it increase your willingness to use commercial cloud for data storage?

Definitely	**11**
Most probably	**9**
Maybe	**11**
Not at all	**3**

- With egress ranking high as a concern; this is a very expected response distribution.
- Commercial providers do provide reduced egress costs, it can often be difficult to navigate and introduce uncertainty in costing models.

Cloud Providers Used

Amazon Web Services (AWS)	24
Google Cloud (GCP)	21
Microsoft Azure	11
Oracle	1
Others	4

Commercial cloud providers used by the respondents

"Other" responses specified the use of Wasabi, Druva, and Digital Ocean. One respondent used the cloud only for prototypes and spin-off projects.

IU Jetstream2 (ACCESS)	10
Fabric	1
Chameleon	1
Self-hosted cloud (OpenStack for example)	7
OSG	4
Other ACCESS Resources	4
Other, please specify	11

Academic cloud providers used by the respondents

"Other" responses specified a variety of resources: Open Storage Network, HPC Centers, Kubernetes, OSF.io, Globus, CyVerse, and TACC.

Experiences Shared

- The survey included 20 valid responses (out of 36 total), which were analyzed to identify recurring themes related to experiences and challenges in cloud adoption
    - ownership and long-term cost implications
    - financial barriers to full cloud adoption,
    - long-term commitments and maintenance, and
    - others like politics!

Experiences Shared: Ownership and Long-Term Cost Implications

- Concerns about losing ownership of infrastructure when transitioning to cloud services.
  - *“After construction funding runs out, we don’t own the infrastructure like we would for at least some physically purchased in-house infrastructure, therefore there are greater operational costs.”*

- Disparity in cost-effectiveness for specific use cases
  - *“It’s really not cost competitive with owned resources the way HEP uses them.”*

Experiences Shared: Financial Barriers to Full Cloud Adoption

- 30% of respondents identifying financial barriers as a primary hurdle
  - *“Cost is the only reason we are not fully in the cloud.”*
  - *“Costs are challenging. If you do not replace the need for people in IT Operations, thus lowering costs on the other side, cloud can be prohibitive.”*

Experiences Shared: Long-Term Commitments and Maintenance

- One respondent raised this issue, long-term contractual commitments and maintenance requirements were still a concern for some organizations.
    - *“Concerns of long-term commitments and maintenance, and speed.”*
- Highlights the need for cloud providers to offer flexible contracts that adapt to needs of the clients

Experiences Shared: Others

- 10% of respondents expressed concerns about governance and evolving vendor cost structures:
  - *“Politics. On-premises NSF/DOE-funded data centers are fighting the cloud and undermining it by appearing to undercharge for on-prem computing.”*
  - *“Vendors changing cost structure for data archival and data egress in the future. Ability to share data via cloud platform to other countries and remote geographical locations.”*

Conclusions

- Respondents generally recognized cloud benefits, including improved reliability, scalability, on-demand access, and better storage services.
- However, respondents also cited major barriers to adoption, including lack of funding, complexity in re-architecting deployments, and workforce skill gaps.
- Cost concerns, often lead organizations to adopt hybrid (i.e., on-premise and cloud) strategies.
- Many on-prem data centers were found to undercharge for computing, often lacking proper cost models.

Future Plans

The Cloud working group plans to set up the interviews in the coming year with respondents that agreed for a follow up regarding cloud usage at their institution.

Thank You !

To learn more about CI Compass services, news, upcoming events and our resource library,
please visit ci-compass.org

Cloud Working Group Mailing List:
cloud-discuss@ci-compass.org

Contact the CI Compass Team with questions or requests by emailing
contact@ci-compass.org

Funded by the U.S. National Science Foundation, Grant #2127548