



The government seeks individual input; attendees/participants may provide individual advice only.

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes¹
August 5, 2020, 12-2 pm ET

Virtual

Participants

Valentine G. Anantharaj (ORNL)	Miron Livny (UW-Madison)
Lisa Arafune (CASC)	David Martin (ANL)
Eric Burger (OSTP)	Deep Medhi (NSF)
Devin Casey (NARA)	Donald Petravac (NCSA)
Richard Carlson (DOE/SC)	Steve Petruzza (Utah)
Martin Doczkat (FCC)	Lavanya Ramakrishnan (LBL)
Sharon Broude Geva (UMich)	Stefan Robila (NSF/CISE/OAC)
Katherine Evans (ORNL)	Birali Runesha (UChicago)
Dan Fay (Microsoft)	Suhas Somnath (ORNL)
Margaret Johnson (NCSA)	Todd Shechter (UW-Madison)
Joyce Lee (NCO)	Kuang-Ching Wang (Clemson)

Proceedings

This meeting was chaired by Richard Carlson (DOE/SC) and Stefan Robila (NSF).

MAGIC Tasking: FY2022

Background

MAGIC reports to the Large Scale Networking IWG (LSN). Each year, MAGIC proposes tasks at LSN's annual planning meeting for the next FY. In past, multi-month meetings with common theme or single topics/months with wide-ranging topics with varying number of presentations. need to meet needs of MAGIC community. How structure meeting to be productive for MAGIC members.

FY21 Tasking:

- **Task 1: Speaker Series:** Speaker(s) from NITRD IWGs (AI, CSIA, HCSS, HEC, and/or SPSQ) to joint sessions on topics of mutual interest
- **Task 2: Single Sessions on a range of topics**
 - Data integrity
 - Data confidentiality
 - Implications of AI to science
 - ROI and cost efficiency- academic and lab-based computing;

¹ Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program.

- Direct integration of energy sources and computing facilities
- Networking infrastructure underlying middleware
- **Academic Roundtables**
 - Continue information exchange on events and topics of interest (workforce development)
- **Data Life Cycle**
 - Data life cycle issues in the context of distributed computing environments

Potential Topics for FY 22

- **Cloud** - Introduction
 - Cloud: Last meeting addressed cloud access (Cloud Bank) – coordinate access to cloud resources provided by commercial providers. Discussed value of project and flexibility to research PIs. Network infrastructure and value of putting single point of access for cloud resources. NSF released another solicitation CSSI (cyberinfrastructure for sustained scientific innovation). how to best use cloud resources (e.g., through single mechanism;) and impact on network infrastructure (e.g., centralized resource allocations).
 - Wish to hear more from user/researcher side re: concerns/limitations with current computing environment and what can be solved by cloud
 - Strategy for using commercial cloud and its impact on NSF’s research cloud activities – need to separate? HTC – no pressure from researchers to use cloud; not address needs
 - Clouds can give elasticity unlike in most other infrastructures – scientific value? How CI perspective reconciles with researcher’s perspective on cloud? Scenarios preventing scientific discovery; can clouds play a role?
- **Cloud Speakers** (Begin Dec 2020 or February 2021) – 4-6 month series
 - Speakers: Stakeholders’ perspectives; managing investment portfolio (academic campuses)
 - large science community – Stefan Robila, lead
 - investigator – Stefan Robila, but also other agencies, participants
 - instructor (understand students’ role in scientific computing) - Stefan
 - agencies: how encouraging/discouraging use of cloud; buying resources (Rich Carlson)
 - academic campuses: management of investment portfolio; decision making process
 - See “Cloud and on-premises data center usage, expenditures, and approaches to return on investment: A survey of academic research computing organizations” <https://dl.acm.org/doi/pdf/10.1145/3311790.3396642> \ Share findings; pointers to potential academic speakers.
 - Access: network infra and campus infra leveraged to facilitate cloud access and use
 - research community –need seamless access to machines, thereby reducing to workflow manager deciding on resources, not individual scientists
 - industry – commercial cloud use perspective (Dan Fey, Microsoft)

- **Speakers to address:**
 - Definition (NIST, commercial cloud)
 - Role of cloud in scientific discovery; benefits of cloud to research - focus on research problems
 - Side-by-side comparison of cloud offerings
 - Research needs: availability, elasticity (scale & moldability), changing prompts to fit use case
 - How to address: bring in end users, agencies (strategies/when support cloud)
 - NSF expectations of researchers obtaining needed resources?
 - Academic campus strategies to encourage cloud use vs. buying hardware; large scale resources (expectations re: attaining resources)
 - Not much discussion on use (vs. cost) or bridging needs and implementation (CloudBank bringing people who understand research needs to work on implementation)
 - Ongoing challenge: amount of investment in shared computing resource on campus? important to provide certain level of computing? Also, how Interface with on-prem or cloud (viewed as costly)
 - Lack development for cloud vs. on-prem; lack end user perspective. Adoption difficult due to lack of understanding of how elasticity add value? Abstract complexity to end user akin to in-prem? See <https://cloud.rcc.uchicago.edu> re: ongoing work. People use if there is credit and not use if no credit. Agencies should run experiments with campuses to incentivize provision of services/ensure at par services, then address cost, security, IDC issues.
- **Other topics:**
 - National Research Cloud – direction; impact on NSF, NIH and DoD
 - Appears to have arisen from research side of Stanford: look at “means”; need speakers
- **Workforce development**
 - OAC/NSF - significant investments in workforce development; note workforce development workshop
 - Campus Research Computing Consortium (CaRC)- much activity in CI workforce development.
 - Lauren Michaels (UW-Madison); Centers of Excellence
 - John Cheatam, CaRC chair (Sharon, POC)
 - Research Software Engineers and the associated new career paths being developed in universities (as part of the workforce development topic (Dan Katz)
 - Speaker:
 - Focus: How to move forward/solutions to implement (training, inclusion)
 - Challenges/roundtable format

- Restructuring to enable a seat at the table
- Brainstorming/collaboration; relationship with HBCUs
- Campus, research POV
- Early career mentorship
- Lead to collaboration

Speaker Schedule

- **September** – Potential speakers: Kevin Thompson (NSF), Dan Katz (Illinois), Mekisha Marshall (ODNI)
- **October** – Diversity and inclusion in HPC
 - HBCU program/approach towards underrepresented groups (ask Mekisha)- research computing resources/infra/training challenges; may connect to cloud perspectives, launch relationship
 - While funding always an issue, what can we do to improve?
 - Linda Hayden (Stefan Robila, POC) – science gateways curriculum and workforce development
 - Jason Arviso (Navajo Technical Univ)
 - University of Puerto Rico programs (Sharon, POV)
 - Science domain inclusion: Dark Energy; Climate survey (Don Petravick), computing infra (Miron Livny); reaching underrepresented groups (pipeline)
- **November** - Lauren Michaels (UW-Madison), Thomas Cheatham (CaRC) (Sharon) -collaboration paper or Jim Wilgenbusch (MN)

Next meeting: September 2 (12 pm ET). Workforce development.