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

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes[1]
May 5, 2020, 12-2 pm ET

Virtual

Participants

Alan Sill (TTU)	Keith Beattie (LBL)
Alison Derbenwick Miller (Oracle)	Kent Foster (NC State)
Bryan Lyles (ORNL)	Mallory Hinks (NCO)
Chase Cotton (UD)	Michael Corn (UCSD)
David Martin (ANL)	Miron Livny (OSG)
Dhruva Chakravorty (TAMU)	Misha Ahmadian (TTU)
Donald Petravick (Illinois)	Richard Carlson (DOE)
Elizabeth Bruce	Saswata Hier-Majumder (DOE)
Eric Lancon (BNL)	Sharon Broude Geva (U of Michigan
H Birali Runesha (U of Chicago)	Stefan Robila (Montclair State U)
Haining Wang (Virginia Tech)	Terril Frantz (Harrisburg University)
Jack Wells (NVIDIA)	Tevfik Kosar (NSF)
Kate Evans (ONL)	Weijia Xu (UT Austin)
Kathy Yelick (UC Berkeley)	William Pentland (GenBright)
Kathy Austin (TTU)	

Introductions: This meeting was chaired by Richard Carlson (DOE/SC) and Tevfik Kosar (NSF)

Cloud Speaker Series – Cloud for Teaching and Training

CloudBank – Managed Services to Simplify Cloud Access for Computer Science Research and Education
Kathy Yelick, Co-PI CloudBank, UC Berkeley

- Gave background on CloudBank and described the advantages: No indirect cost, resource management, training and user support, easy switching across vendors, community support, data hub for education
- Described trends in CS and DS education
    - CS majors grew by 291% in 10 years
    - Growth in Masters of Information and Data Science and Masters of Engineering growing
    - PhD programs relatively flat – number of faculty has not grown much
- Teaching CS
    - IPython/Jupyter Notebook

---

[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.

- - - Tool for creating interactive textbooks
      - o Otter Graders: Scalable autograder - runs on the cloud
      - o Jupyterhub on Kubernetes
        - Deploy JupyterHub on a cloud cluster
        - Uses Docker for student Environments
        - Platform agnostic
      - o Gave an overview of the Datahub by the numbers
  - Building Community
    - o 2i2c.org – The International Interactive Computing Collaboration
      - Mission: Make interactive computing more accessible
    - o CloudBank Computing for Data Science Education
      - El Camino College – HSI
      - Howard University – HBCU
      - Skyline – HSI
    - o Went over CloudBank workflow

Q&A
- Saswata Hier-Majumder – I was at the University of London before this and we launched a similar program, called "Digital Geosciences". Have you had any issues with merging the Jupyter environment with a learning system environment like Moodle or Blackboard?
  - o It is integrated into Canvas and Grade Scope. That integration has certainly happened.
- Sash also asked about industry collaboration.
  - o There is certainly interest. It is something we are trying to figure out how to deal with. But bringing in industry would be great.
- Rich Carlson – You talked a lot about data science. What about your other computational side, operating systems, programming languages)? Are you using similar mechanisms to teach those?
  - o In the computer science courses, they do use them for things like the autograder and things like that. I don't think the upper division courses like the compiler courses and operating systems courses are currently using the cloud environment. But they are using some of these kinds of infrastructure.
- Rich – Do you have to wipe the environment each semester?
  - o Yes, they do wipe the environment at the end of the semester. I think they create a different hub instance for each new class and the old ones disappear.
- Kathy mentioned at the end that thinking about how to finance this when it gets to something more expensive, whether you're doing parallel computing or machine learning, we need to think about our national infrastructure.
- Elizabeth Bruce (Microsoft): What do you see as the biggest barriers or blocker for scaling cloud?
  - o People. Scaling the support (personnel to support the infrastructure and build it). This is what 2i2c is set up to try to handle. Also, the cost issue which comes up in certain kinds of computationally intensive courses.

Integrating Cybersecurity Education with Cloud Computing
Haining Wang (Virginia Tech) and Chase Cotton (University of Delaware)

- Objective is to introduce recent security research activities in cloud computing to the classroom and to enhance the undergraduate and graduate curriculum in cybersecurity.
- Working on developing a new senior/1 year graduate level course
- Described topics covered and lab exercises
- Labs Activities

- o Cache Template Attack
  - ▪ Shared resources used for side-channel attacks: Caches, Memory, Buses
  - ▪ Described 3 classic cache-based attacks: Prime & Probe, Flush & Reload, Flush & Flush
- Chase reflected on the use of cloud computing infrastructure in the classroom
- Chase shared some thoughts on cloud use in the classroom
  - o Mentioned that they roll a lot of their own infrastructure

Q&A
- Tevfik Kosar – When the students use these different attacks you described in the cloud environment as part of this course, it may be possible that they access regular users' data. Is that correct?
  - o One of the features we use in cybersecurity is we don't have conventional labs. Students all build their own Linux machines and run them on their own laptops. Most of these labs can in fact be run under virtual machines. The memory deduplication lab, which takes advantage of the fact that a commercial hypervisor like ESXI, when it finds two clients that use the same memory, they put them in the same space. Once it was showed that they could make this attack, that has become turned off by default. But if you were running a large infrastructure, you would turn deduplication back on. We don't have to send them back out to the cloud to practice this.
- Dhruv Chakrovorty – What was your university's IT enterprises reaction to the kind of things you're trying to accomplish, especially with the on-premise solution you had on your last slide?
  - o Chase said that he was good friends with them.
- Keith Beattie – Could you post that link to those videos?
  - o Posted in the chat: https://drive.google.com/file/d/1eXlydAaW6-aPZwxJDMoci52q3C_shzf-/view

Facilitating Remote Resources Usage via User Driven Web Applications
Weijia Xu (UT Austin)
- Background
  - o Data-driven AI analytics help organization gain new insights
  - o Due to increasing data storage requirement, computing complexity, there is an increasing trend to move to remote advanced computing resources
- Described observations of new user needs on cyberinfrastructure
- Described traditional access models and the associated concerns
- Described the web portal model the associated concerns
- Project goals:
  - o User driven model that empowers users to easily set up and deploy applications on supercomputers with web UIs for interactivities
  - o Enable users to share the session dynamically with collaborators and/or students
  - o Hide resources and allocation details from end users
  - o Portable and extensible
  - o Easy access control
- Gave overview of TASK Classes
- Described support for education activities
- Gave overview of credential management design

Q&A

- Tevfik Kosar – What major functionalities did you see missing in the existing tools mentioned?
  - We want to lower the access barrier for people to use the HPC system

Wrap Up
- Rich said that this meeting would wrap up the cloud speaker series. Next, we need to roll up all the presentations and do a summary report. He asked that anyone interested in helping us put together or review the report once it starts getting put together, please email Mallory.
- Rich said that next month we are going to start a new series on software sustainability. Next month we will kick it off with a set of discussions, no formal presentations. We can hash out as a group what we want to do with software sustainability.
- Rich invited attendees that are not members of MAGIC to join us. This is an open public meeting. Contact Mallory to join the mailing list.

Next Meeting
June 2 (12 pm ET)