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

Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes
July 6, 2022, 12-2 pm ET

Virtual

Participants

Alan Sill (TTU)	Kate Evans (ORNL)
Alejandro Suarez (NSF)	Manish Parashar (NSF)
Arjun Shankar (ORNL)	Marcy Collinson (Oracle)
Ben Brown (ASCR)	Martin Swany (IU)
Bill Miller (DOE)	Mike Heroux (SNL)
Brian Lin (UW-Madison)	Miron Livny (UW-Madison)
David Martin (ANL)	Rama Vasudevan (ORNL)
Dhruv Chakravorty (Texas A&M)	Rich Carlson (DOE/SC)
Frank Wuerthwein (UCSD)	Seung-Jong Park (NSF)
Hal Finkel (DOE)	Sharon Broude Geva (University of Michigan)
Jack Wells (HPC)	Tevfik Kosar (NSF)
Jeff Conklin (NCO)	Todd Tannenbaum (UW-Madison)
Joseph Strathus (UW)	Tom Gibbs (Nvidia)

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

National Strategic Computing Reserve (NSCR): Background and Vision
Manish Parashar (NSF)

- Manish provided an overview of the COVID-19 High Performance Computing Consortium
  - 43 Consortium members
  - 600 Petaflops
  - 100 Active projects
  - 6.8 million CPU cores
  - 165K nodes
  - https://covid19-hpc-consortium.org
- Manish talked about what worked with the consortium
  - Collaboration across institutional and organizational
  - boundaries
    - Government, industry, and academia
  - Nimbly created and is operating a common portal
    - Meeting urgent computational resource requirements not easily met through other means
    - Developed and refined review, matching and onboarding processes for accessing resources

        - Set up a communications and user engagement framework for a worldwide community
    - Accelerated explorations in basic understanding of the SARS-CoV2 virus, its host interactions, strategies to mitigate its spread, and early-stage drug development.
    - Manish stated the consortium held Open Request for Information to help inform potential attributes and design of a NSCR (closed on 01/16/2021)
    - Deployment Scenarios: envisioned scenarios; design of triggers for activating and deactivating; testing readiness; barriers to activating.
    - Computational Resources: recruited, vetted, and sustained; incentives and mechanisms for compensation; assessing the suitability of resources; types of research supported.
    - NSCR Providers: resource providers' contributions; approaches for resource selection and allocation; resource providers opt in or opt out.
    - NSCR Users: principles for allocations of computational resources; eligibility restrictions/selection criteria for users and use cases.
    - Community Formation: community outreach and communications needed; organizations and services that NSCR should coordinate with.
    - Partnership Agreements: What are key aspects of partnership agreements (e.g., access to results, intellectual property rights) that can help sustain the NSCR over time?
    - Relationship to Other Strategic Reserves: other relevant strategic reserves and how to connect or interface with those reserves; lessons from other strategic reserves to inform NSCR.
- Manish described the National Strategic Computing Reserve Blueprint Report
    - Advanced computing cyberinfrastructure can be a strategic National asset in emergency response, if mobilized quickly.
    - Goals for an NSCR:
        - Ensure availability of a ready “reserve” of resources (computing, data, software, services) and expertise that can be leveraged nimbly in times of urgent need.
        - Establish policies, processes, and agreements to enable agile, effective, and impactful resource mobilization.
        - Build on continued longer-term strategic investments in resources (computing, data, software, services), expertise.
        - Coordinate across agencies, stakeholder communities, and other national reserves.
- Manish described the key elements of the National Strategic Computing Reserve (NSCR)
    - Resources
        - Computational and data resources
        - Systems expertise
    - Program office
        - Policies governing triggering and ramp down
        - Coordination mechanisms
        - Readiness testing
        - Systems and subject-matter expertise
        - Governance structures
        - Training and outreach
    - Cyberinfrastructure/Software framework
        - Cross agency, on-demand, federation, allocations, provisioning, operations; cloud integration

- Manish talked about NSCR next steps
    - Scenario-driven tabletop exercises
    - Community workshops (in coordination with other reserves/responders) to explore responses and role of NCSR in specific national emergencies (wildfires, hurricanes, floods, nuclear disasters, etc.).
    - Deep dives into key aspects of the NSCR, e.g., governance, incentives, triggering, sharing data, technical infrastructure (e.g., single sign-on) etc.

Questions:

- Marcy Collison asked if Manish's org was dovetailing efforts in terms of utilizing HPC in order to do analysis across populations to track COVID dispersion.
    - Manish stated that for some projects they did in California, NSCR provided resources for relevant research and development and operational activities to support the response to the emergency.
- Miron Livny wanted to know what were the next steps to get the NSCR moving forward.
    - Manish answered by saying there are a few different dimensions so they setup an entity that can coordinate this. He stated that a lot of this is bringing the pieces that already exist in a lot of different areas together and integrate them to meet a particular need. How do they then set up the right structure incentives ad hoc? What would it take for a resource provider to be part of this? What would it take to get subject matter experts engaged and what kind of agreements does one need in place? Those are the type of things that we are hoping will come back to the Community and get input on, as well as volunteers to help put those things together and then to understand how it gets funded.
    - Miron then stated that this is the open question – the relationship between the governance and the infrastructure – so that will be the new thing to add today, because we don't operate in a nationally coordinated way today.
    - Manish followed up by agreeing and stating, for example how do you go about having single sign on between agencies – that requires a lot of policies and decisions and then engineering to make that work. He talked about piloting some of these activities.
- Alan Sill stated he was coming to the point of view that it's just a variation on long standing patterns of distributed and resilient computing that the Community has pursued to a number of projects services, including open science grid. He stated he would like to understand this proposal in that context. For non-urgent computing if it is understood this as part of a spectrum of computing is perhaps necessary for a complete picture. One way to pull the Community in is not just to focus on the urgent projects, but in fact to think of the spectrum of providers. This gives you a number of things, first of all, it broadens the set of resources. One might think, well, not all of those are top end high performance computing resources but it helps to address government and in collaboration issues and understanding what are the unique characteristics of variations concerning the model of distributed computing.
    - Manish answered affirmatively, but also added he thought the focus was seeing how you operationalize design, there are two things, one is getting the solutions to actually work at the scale, but also understanding what the government structure and policy needs are so that they can engage with industry to get the resources they need, or they can engage with academia. For COVID 19 everyone wanted to do something right and contribute resources but there were no policies or mechanisms in place for this to happen.
- Marcy wanted to know if there were any publications he could point to around this effort.

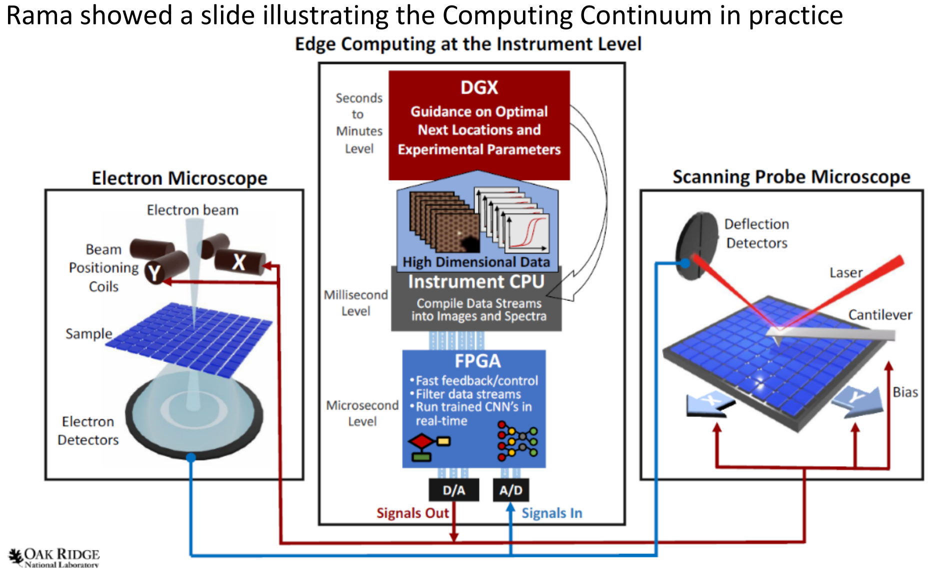
    - Manish mentioned there was this blueprint he talked about here and he also referenced a paper that came out in IEEE. He also mentioned an article that came our in Scientific American that might be relevant here.

Edge Computing and Microscopy: Autonomous and Automated Experiments Leveraging Advances in Computing, Control Systems and Algorithms *Rama Vasudevan*, *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory*

- Rama described a graphic focusing on the Information flow in SPM
    - 'All information gathered from volume of tip surface junction
    - Chemistry and physics merge at the nanoscale
    - Access to wide range of bias, thermal, and pressure induced phenomena
    - Discovering and differentiating phenomena requires multiple, complex spectroscopies
    - Advanced analysis to convert data to understanding

- Rama displayed an article entitled Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy
- He also talked about abstraction and automating the microscopy world
    - Chemputation – digital chemistry - A complete programming language for chemistry that can run on open hardware
    - PycroManager
    - Hardware abstraction using FPGA
    - Software abstraction using Python
        - Call low level functions to control tip position, scanning (e.g., raster, spiral, move etc.), specify voltage waveforms, collect data all in Jupyter notebook
        - Enables design of complicated automated and autonomous experiments, hardware independent
    - Putting it all together – User -> Software layer -> Hardware connection layer -> instrument
- Rama had a few illustrations on real-time automated experiments in SPM
    - Streaming functional fitting during scanning on the microscope
    - Instrument -> Gateway -> GPU Server
- Rama showed a slide focusing on "Smart" Microscopy Experiments
    - Techniques
        - Adaptive sampling
        - Structure-triggered events
        - Streaming analysis with model comparisons "on the fly"
    - Yields:
        - Enables more efficient and more complex measurements
        - Engineering new metastable states
        - Better hypothesis testing for physics discovery
- Rama showed a slide focusing on Real-Time Automated SPM
    - Provide list of new measurement conditions
    - Measure a batch of hysteresis loops
    - Perform data preprocessing
    - Upload to dgx for Bayesian Optimization
    - Repeat

- For long spectroscopies, measuring every pixel across a grid is not feasible. Bayesian optimization enables sampling only those points that will maximize a property of choice, enabling new experiments that would otherwise not be possible
- Rama showed a slide on Autonomous workflows
  - Materials design problem in microscopy: how we can modify materials intelligently
  - Workflow
    - Write automated script
    - Collect data and state transitions with policy
    - From data, train a surrogate model (Domain knowledge simulations)
    - Train autonomous agent, obtain policy
    - Update policy, go back to Collect data step
- Rama talked about changing the Autonomous workflow in the future:
  - Workflow
    - Write automated script
    - Collect data and state transitions with policy
    - From data, train a surrogate model (***Massively parallel simulations***)
    - Train autonomous agent, obtain policy
    - Update policy, go back to Collect data step
- Rama displayed a slide illustrating their computational needs using a Streaming, near edge, and HPC computing continuum approach
  - Streaming edge
    - About 64mbps
    - FPGA
    - Jetson Nano
  - GPU "Near Edge"
    - MB-GB per run
    - DGX-2 16 GPUs
  - HPC- simulation and modeling refinements
    - 10+ GB per job - leadership class
    - 27000 GPUs
- Rama showed a slide illustrating the Computing Continuum in practice



Questions:

- Marcy wanted to know if this could be utilized for other domains of research.
  - Rama stated that this solution is generic enough to be utilized in other domains. If one has a system with automated data collection, one could collect a large amount of data in a reasonable amount of time.
- Arjun Shankar wanted to know what his sense was of how the data planes and workflow planes work together. How does he see the data for experiments being captured somewhat separately than the control workflow.
  - Rama stated this is a key question – he stated when they develop the system they run into this issue. For example, if their system was running an optimization on where to sample next, they could leverage previously captured data on similar material systems to be more efficient. The model could already be "pre-trained for this purpose.
- Bill Miller was curious about the layer of data above the workflow, the metadata, and being able to trace back to the original characteristics of the data.
  - Rama stated this was an important question as they don't necessarily know what Metadata will be important down the line, so they capture as much as they can. They are trying to develop a system where, no matter how you captured your data, all those steps would be visible down the line. When you did go back to that data you would know exactly what workflow was used to run it, what processing algorithms were there.
- Bill had a follow-up question about the instrument itself, how they used to be sold with proprietary firmware and this could prevent one from doing certain things that today we would call open science (being able to get in there and program the instrument). He wanted to know whether this was still an issue.
      - Rama stated thing are improving. He can do a lot of API calls in Python that handle things at that level, but he would still say vendors are reluctant to give them full control.
- Tom Gibbs wanted to know if they had done any experiments on which simulations you would want to run based on whether you know what that part of the workflow looks like.
  - Rama stated they ran simulations and that there were certain unknown parameters within those simulations. So, if they run massively parallel simulations they could find the optimal parameters, to best match what they are seeing on experiments. Once they have that then they could train a surrogate model based on those simulations.

<u>Roundtable</u>

- The co-chairs talked about the planning sessions coming up in October which focuses on what the group would like to accomplish over the coming year.

<u>Next Meeting</u>

August 3rd (12 pm ET)