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

**Middleware and Grid Interagency Coordination (MAGIC) Meeting Minutes**
August 3, 2022, 12-2 pm ET

Virtual

**Participants**

| | |
|---|---|
| Alan Sill (TTU) | Miron Livny (UW-Madison) |
| Alejandro Suarez (NSF) | Pete Beckman (ANL) |
| Geoffrey Fox (UV) | Rich Carlson (DOE/SC) |
| Glenn Lockwood (Microsoft) | Seung-Jong Park (NSF) |
| Gregor von Laszewski | Shantenu Jha (Rutgers) |
| Jack Wells (HPC) | Sharon Broude Geva (University of Michigan) |
| Kevin Thompson (NSF) | Tom Gibbs (Nvidia) |
| Marcy Collinson (Oracle) | Diana Weber (NCO) |
| Martin Swany (IU) | Eric Blanco (NCO) |

**Introductions:** This meeting was chaired by Rich Carlson (DOE/SC) and Seung-Jong Park (NSF). Rich added a topic to today's agenda to address what the MAGIC group will do for FY23.
Diana Weber (NITRD NCO) was standing in for Jeff Conklin during today's meeting.

***Advancing Science with AI from the Edge to the Datacenter***
*Geoffrey Fox (University of Virginia)*

- Geoffrey provided a link to his presentation: https://docs.google.com/presentation/d/11ajkTovSLPzIcIV5Pcf_-0Y0uqwkNJXS05aEyPq31uI/edit?usp=sharing
- Geoffrey stated he would be talking about how science can be advanced or is being advanced by using AI with compute systems from the edge to the data center.
- Geoffrey began his discussion with an abstract of his presentation:
    - We distinguish systems used to support AI and AI supporting data centers (supercomputers, clouds), fog, and edge.
    - Today most AI advances are in deep learning which inevitably uses HPC and stretches from edge to data center.
    - AI will permeate all parts of the computing continuum software stack, systems and application software
    - AI surrogates replace compute intensive libraries by learnt surrogates.
    - This support will be at the system level where it will both improve performance and remove the distinction between supercomputers and clouds by carving out application specific compute resources from a heterogeneous networked hyper-hyper systems.
    - Meta or Foundation Models could generalize across science domains

    - Thanks: This work is partially based on collaborations with Tony Hey, Jack Dongarra, Shantenu Jha, Pete Beckman, Madhav Marathe, Judy Qiu, Gregor von Laszewski.
- Geoffrey then moved onto the next couple of slides which described different areas AI is used in engineering and science:
    - Experimental Physics (accelerators of anything, microscopes, telescopes. Detect gravitational waves): Control and monitor apparatus, remove noise, design optimal apparatus and then analyze resultant data – edge devices
    - Theoretical Physics: Speed up simulations – AI surrogates learns results; new AI-based methods of performing integrals (QCD)
    - Astronomy: Identify galaxies, stars, black holes, asteroids from observations across multiple wavelengths
    - Material Science: Predict properties of material from its specification and inversely predict specification from desired properties. Data from simulations and measurements such as Xray and light scattering
    - Drug Discovery: Similar but properties are not "strength" or "weight" but rather effectiveness against a disease.
    - Also predict protein structure from specification (AlphaFold2)
    - Life Sciences: AI to analyses omics, patient records, pathology images, epidemiology (Covid), AI+" mechanistic models" to build virtual humans
    - Earth, Agricultural and Environmental Science: analyze images from satellites and time series from distributed sensors
    - Climate and Weather Science: AI controls ensembles of simulations and integrates observational and simulation results; AI learns microstructure (clouds)
    - AI for "net zero" will make world green
    - Critical Infrastructure: Electrical power grids; AI will control the new distributed mix of renewable and traditional power sources
    - Fusion: Tokomaks use an electron plasma to create "clean nuclear fuel" that needs real time AI control from observation and simulation (surrogates)
    - AI for Computing Infrastructure
    - Schedule and control computers and networks
    - Learn results of searches, data analyses and simulations (surrogates)
    - Visualization
    - Lots of social media and e-commerce commercial applications
    - AI Above Everything: AI will read all the books and gobble all the data and advise/tell us what to do?
- Geoffrey then talked about Features of AI Use in Science in Computing Continuum:
    - Synergistic with industry work e.g., computer vision nets starting point for remote sensing, pathology; industry time series starting point for science sensor analysis
    - Simulation surrogates used everywhere from material science, climate and QCD to visualization (open questions such as uncertainty quantification)
    - Surrogates often only replace part of simulation
    - ~95% of new AI development is deep learning with decreasing use of classical ML but substantial data engineering before and after the core AI
    - Data gathered and (partially) analyzed at the edge
    - Realtime sometimes essential (web shopping, wildfires) and other times data not all streamed (some monitored online) but rather batched and analyzed fully at data center.

    - Edge instruments (from sensors to LHC/SKA/LIGO/Light sources) are designed, monitored, and controlled by AI while data analyzed with AI
    - High Performance Computing HPC essential whether at instrument or at data center and whether latter is a cloud or a supercomputer
    - Networking from edge to data center controlled by AI as are both data center and edge/fog computing systems
- Geoffrey went onto discuss MLCommons – a consortium of 62 companies, academia
    - MLCommons aims to accelerate machine learning innovation to benefit everyone. Benchmarking, Datasets, Best Practices
    - Major effort of 62 companies to produce benchmarks with ongoing challenges
    - Training at V2.0 (sixth release)
    - Fox set up Science Working Group with co-chair Tony Hey who has a significant benchmarking group SciML.
    - Some relevant working groups:
        - Training
        - Inference (Batch and Streaming)
        - TinyML (embedded)
        - Power
        - Datasets
        - HPC (Supercomputer Implementations)
        - Research (Academic-Industry Links)
        - Science (AI for Science)
        - Storage (new Research Group)
        - Best Practice (Software)
        - Logging/Infrastructure (metadata)
- Geoffrey talked about the MLCommons (MLPerf) Consortium Activity Areas:
    - Benchmarking – provides consistent measurements of accuracy, speed and efficiency. Consistent measurements enable engineers to compare innovations and choose best ideas to drive the solutions of tomorrow.
    - Datasets – are the raw materials for all machine learning. Models are only as good as the data they are trained on. Academics and entrepreneurs depend on public datasets to create new technologies and companies.
    - Best Practices – empower researchers and engineers to more easily exchange models, reproduce experiments and build applications that leverage machine learning. Improving best practices accelerates progress in and market for machine learning.
- Geoffrey talked about the benchmarks produced by the science working group and showed a slide illustrating benchmarks for:
    - Satellite data benchmarks
    - Earthquake data
    - Electron microscopy
    - Drug discovery
- Geoffrey went on to show a slide of MLPerf HPC v1.0 Benchmarks – with contributions from a number of laboratories:
    - CosmoFlow
        - 3D CNN regression on cosmology simulations
        - Originally published at SC18 total size 10.2 TB
    - DeepCAM

        - 2D CNN segmentation, identifying weather phenomena in climate simulations
        - 2018 Gordon Bell prize paper total size 8.8 TB
    - OpenCatalyst
        - Graph NN predicting energy and forces in atomic catalyst systems (material surface + molecule)
        - Dataset: Open Catalyst 2020 (OC20), variable system size, 300GB total size
        - Reference model: DimeNet++, 1.8M parameters
- Geoffrey showed a slide illustrating an example - State of the Art AI Models are increasing in Model and Data size – showing a graph about large scale models.
- Geoffrey showed another slide illustrating the increase in performance of AI models
- Geoffrey presented some remarks on MLCommons:
    - Highest performance "cloud" systems are from Google (4096 TPU-v4, fastest) and NVIDIA (4216 A100's) but these are surely HPC "clusters"
        - In fact, larger than supercomputer partitions used in responses to HPC Training MLPerf benchmarks
        - Not very efficient at maximum size!
    - MLCommons has as much emphasis on open datasets as on models
    - Industry participants typically HPC experts with sometimes University or DOE lineage
    - Inference has datacenter, edge, mobile and tiny(embedded) benchmarks
        - MLCommons spans Computing Continuum
    - Power, Performance and Science Discovery metrics
    - Looking at using open data/benchmarks to produce an AI for Science Gymnasium to provide tutorials and examples that can be transferred across domains
- Geoffrey presented information on AI for HPC:
    - Traditionally we did HPC Systems for AI -- run deep learning on a GPU BUT AI for HPC Systems is probably more interesting (stems from Jeff Dean 2017 NeurIPS)
    - Currently in science the AI is used to enhance simulations (not data analytics) and dominantly the AI used is Deep Learning
    - Both SC20 Gordon Bell winners used AI enhanced simulations a) Learn multi-particle potential b) Learn to explore phase space
    - With Shantenu Jha, introduced a set of major categories of AI/ML for HPC and 8 subcategories but Shantenu usefully simplified to 3.
    - AI in HPC is learning and replacing traditional HPC software components, surrogates
    - AI about HPC is "in charge" of simulation and optimizing execution e.g., path through phase space
    - AI out HPC is AI running synergistically with an HPC job, such as post processing to find particle trajectories
- Geoffrey talked about an example, up to two billion times acceleration of scientific simulations with deep neural architecture search
    - January 23, 2020, https://arxiv.org/pdf/2001.08055.pdf Updated October 2020
    - 10 scientific cases including astrophysics, climate science, biogeochemistry, high energy density physics, fusion energy, and seismology, using the same super-architecture, algorithm, and hyperparameters.
    - Approach also dynamically choses deep network and provides uncertainty estimation, adding further confidence in their use.
- Geoffrey presented other examples of AI for HPCSys

- Extraction of ionic structure in electrolyte solutions confined by planar and spherical surfaces.
- Classic HPC code written with C++ and accelerated with hybrid MPI-OpenMP.
- Uses quite small multi-layer perceptron AIP to predict 150 observables from 5 input parameters (~5000 in training set)
- AIP outperforms other AI choices
- Deployed on nanoHUB for education (an attractive use of surrogates so students get answers fast)
- General Electric uses similar approach to give interactive Engine design options (200 in training set)

- Learn Newton's laws with Recurrent Neural Networks
  - Deep Learning is revolutionizing (spatial) Time series Analysis
  - Good example is integrating sets of differential equations
  - Train the network on traditional 5-time step series from (Verlet) difference equations
  - Verlet needs time step .001 for reliable integration but Learnt LSTM network is reliable for time steps which are 4000 times longer and also learn potential.
  - Speedup is 30000 on 16 particles interacting with Lennard-Jones potentials
  - 2 layer-64 units per layer LSTM network: 65,072 trainable parameters
  - 5000 training simulations
- Geoffrey talked about Futures of Surrogates and Computational Science
  - Is Surrogate-based Computational Science the future?
    - Traditional simulations used to train very flexible surrogates
    - Some production examples (as in GE's use in engine design) clear
    - Value of surrogates in education clear
    - Challenge is cases where a simulation set (mapping to a single surrogate) is small
    - Surrogates run well on GPU's; traditional simulations may run well on GPU's
  - Need a neural architecture (meta-surrogate or Foundation model) which can both describe individually many different simulation sets and for each set be as broad as possible So a computational scientists can base their work around one (or a few) surrogates

| | | | |
|---|---|---|---|
| Training samples in simulation factory -> | Optimized by Computer science Surrogate Architecture | Training | Computational Science Inference/Computational Results |

- Geoffrey presented the Process for Science Foundation or meta-Models
  - Data Creation
  - Data Curation
  - Training
  - Adaptation
  - Deployment
- Geoffrey presented Research Plan: AI for Science Patterns
  - Currently, deep learning plays a dominant role in science, from data analytics and simulation surrogates to policy decisions.
  - We can group the system structure into a "handful" (two hands ≤10) of patterns such as
  - an Overall Reasoner based on reinforcement learning and large language models to learn the world's knowledge and control experiments.
  - Image-based systems for astronomy, pathology, microscopy, and light scattering.

- Graph-based systems such as in social media and traffic studies; represent molecular and other structure.
- Dense systems to map structure to properties as in drug discovery.
- Time series and sequence (Transformer) models as in language, earth, and environmental science.
- GAN/Diffusion models to generate scenarios as in datasets to test experimental system.
- All Network types can be mixed together as in text to image systems DALL-E and Imagen.
- We imagine taking good examples of each pattern and supporting them with high-performance, easy-to-use environments in end-to-end systems, including data engineering.
- This would cover parallelism, storage and data movement, security, and the user interface. Then we explore multiple examples of each pattern, possibly leading to "Foundation" models for each of them.
- Provide AI for Science Gymnasium (in MLCommons) for tutorials, demos

- Geoffrey presented a slide on AI and Systems Interaction
    - The System hosts the AI software
    - AI writes the System Software (Software 2.0)
    - AI designs the Systems CPU, Storage, memory etc.
    - AI designs the Network Architecture/Hardware
    - AI monitors and manages system: DevOps AIOps (including AI for Data center)
    - AI in real time optimizes system communication on wide area network
    - AI in real time optimizes job scheduling; Current heuristics, New DL/RL Based?
    - AI in real time optimizes I/O, heterogeneous compute, local area network
    - AI learns all possible functions/services/job itself inside a job and supplies very fast results (surrogates)
    - All of these capabilities are integrated together with new distributed operating system
    - Initially runs on single nodes, single clusters, single supercomputer, single cloud, then multiple clouds, clusters, supercomputers
- Geoffrey presented his conclusions
    - AI is promising pervasive approach to continue increases in system performance
        - The largest factors come from the inner layers but are application specific
        - outer layers (scheduling, interoperability of HPC Clouds) give useful increases
    - AI for HPC Systems very promising where we could aim at
        - First Zettascale effective performance in next year or so
        - Hardware/Software aimed at general AI assisted speedup of computation
        - Your health can be engineered with AI-assisted personalized nanodevices designed based on the AI-assisted digital twin of disease in your tissues
    - Global AI and Modeling Supercomputer good framework with HPC Cloud linked to HPC Edge
        - Training on cloud; Inference and some training on the edge
        - AI will design, monitor, control experiments; generate apparatus simulations; analyze its data and propose new theories
    - Need to develop meta-models (Foundation models) valid across many domains
        - Broad use of deep and reinforcement learning

<u>Questions</u>

- Rich asked how well the operating systems, run-time environments. and workflow systems are developed to allow these applications to run effectively across this range of computing resources that they are being presented with.
  - Geoffrey stated that the applications are pretty new, being developed in the last three years. He stated the current systems are promising but not optimal. There needs to be more development.
  - Rich asked if this was on the application side or the distributed computing side.
  - Geoffrey answered both sides. He stated the thing that needs to be done for these dynamic large scale hyper parameter searches has different challenges for workflow systems.
  - Miron stated they are working with a group from Stanford on running worker manager AI application that can operate in a more dynamic distributed tightly coupled environment. They are working not only on the AI side but also on the GPU side when talking about their simulation, they are running on 16 A100s. But if you go to checkpointing you can break things. Some of these simulations can run for weeks, even months and better checkpointing in a distributed environment may help.
  - Geoffrey stated they have a good start but also some big challenges when looking at intense computing and storage and networking, because they are trying to optimize the use of accelerators and a lot of the simulations are accelerator dominant.
- Jay Park stated they are developing software for workflow and science gateways for different applications and now they have this AI based application where AI demands from many different domains. He asked Geoffrey what he suggested for helping the community to change their ways to help this new demand on AI based applications especially for the HPC community.
  - Geoffrey stated that with patterns of compute that would have different optimizations he recommends the usual type of collaboration between domain scientists and infrastructure/computer scientist people like to do. He went onto reference a radical pilot system and some work in Chicago. He stated its harder to train these models from an academic researchers' perspective because of resources whereas industry has compute resources that are far greater in scale and power.

**<u>From the Instrument to the Supercomputer: How Edge Computing is Transforming Science</u>**

*Pete Beckman (ANL)*

- Pete stated he would be talking about a project funded by NSF about cyber infrastructure and it's specifically to build what Geoffrey alluded to, which is this pipeline between AI at the edge going to the HPC.
- Pete showed examples of both large (telescopes) and small instruments (neon sensors) out on the edge and stated this is where the data pipeline begins. Instrumentation is distributed.
- He described the data pipeline historically as instrument collects data and then goes into something more static for analysis. He said that now we have very sophisticated instruments like IoT, and we want to inject computation right at the point where the data is first sensed. Part of the reason is that the newest of sensors can sense more data that can possibly be sent.
- The question becomes what data to keep as there is just too much data to keep it all. You need to analyze the data at the edge and then connect it across the digital continuum. To do that we need to add computing across these points of entry.
- Pete then talked about the SAGE project – Cyberinfrastructure for AI at the Edge sagecontinuum.org

- Pete defined the goals of SAGE:
    - Build new reusable cyberinfrastructure
        - High level resilient, well documented software
        - Leverage best Open-Source frameworks – PyTorch, OpenCV, Tensorflow, Kubernetes, Docker
    - Build a community of AI@Edge scientists
        - New AI-based measurements
        - New AI Algorithms at edge
    - Deploy experimental testbed into production facilities
    - Provide new capabilities for live data and triggered responses
    - Each and train students, explore new ideas
- Pete said they are building on an array of things that came out of the last NSF project, they're using a lot of the ideas and some of the basic core components to build this new concept that is more distributed and more capable.
- In his testbed there are two endpoints.
    - The first is pure software design using a commodity server (this is in Colorado) called SAGE Blade – Sage software stack + commodity server – installed in a little hut next to the outside instrumentation.
    - The second is a Wild Sage Node, and it is ready for mounting outside (more weatherproof) and any PoE sensor can be easily added.
- Pete showed a video snippet of the deployment of the second node in the city od Chicago. These are an extension from the original project array of things, but they have much bigger CPUs and more sensors. These give them the ability to run computation at the edge.
- Pete showed a slide that described the different computations that can be done at the edge. Examples include cloud cover, traffic state, solar irradiance, wildfire-smoke detection, etc. These computations run inside docker containers that can be shipped out to the edge. Cyber infrastructure has a tremendous system software component to make this happen.
- Pete next provided some examples of what he means by processing at the edge in these kinds of containers:
    - Avian Diversity Monitoring – uses the Cornell deep learning module – can be run in docker containers and shipped out to nodes
        - Bird diversity changes as a metric to track the current environmental conditions
        - Automated by using a DNN called Birdnet, capable of identifying 984 bird species.
        - Correlated with human point count observations
    - Cloud motion vectors with hemispheric sky camera
        - Has application in meteorological analysis, nowcasting and short-term prediction of solar irradiance.
        - The use of AI/ML with sensors produce valuable products
        - Allows understanding about what is happening at a cloud level.
        - 
    - NEON project measuring water and snow depth
        - Computer vision (CV) based
        - Machine learning algorithms
            - UNet, ResNET
            - Self-supervised learning
        - Camera contamination by rain and snow is identified by ML algorithm and reported

- Surface water detection – linked with HPC can be used to build hydrology models and predictive capabilities. There is a project currently at a ballfield in Chicago that predicts if the ballfield will close based on surface water and predicts if future weather will affect it.
- Wildfire detection and prediction – exploring wildfire at the edge linked to HPC simulations
  - Using cameras and AI to detect fires in real-time.
  - Two approaches to improve predictions are: Use of thermal cameras, and incorporating the motion of smoke in the DL models
  - Prescribed burns and real wildfire data is needed to train the AI models.
  - Cloud temperatures can be used to estimate cloud-based heights and cloudiness
- Pete talked about using two instruments to train the other, in this case using a visual model to train the infrared model.
- Pete then talked about a specific example using a NEON Mobile Deployment Platform with Sage at the Konza Prairie in Kansas for a controlled burn (Wildfire Detection) in April 2022
  - This was a controlled burn on undisturbed tall grass and his instruments were put on a tower to sense during the burn.
  - They monitor in real time and then create these kinds of graphs but then do the AI again to train the system so that they can build this edge to cloud composition, infrared data and visual data to train.

- Pete then directed the audience to a page of URLs for SAGE resources as follows:

**Getting started with Sage!** - https://docs.sagecontinuum.org
**Sage AI@Edge Apps** - https://portal.sagecontinuum.org/apps/explore
**Sage Data** - https://portal.sagecontinuum.org/data
**Sage Konza MDP Campaign** - https://mdp.sagecontinuum.org
**Overall Sage system status** - https://admin.sagecontinuum.org/status

Questions

- Gregor von Laszewski asked how Pete's projects deal with the power for these remote instruments.
  - Pete stated that in the cities they have line power, and the sensors are PoE (power over ethernet). They started working on a deployment using solar panels – Oregon Hazard Labs has a lot of experience with that so he can leverage that. He also has starlink. The prairie burn used AT&T cellular and it ran up a pretty high bill.
- Marcy Collinson asked if Pete used research and education networks to tackle the power problem.
  - Pete said they have done some of that, for example they are working with a native American tribe in Wisconsin, and they are providing the hosting, including power.
- Marcy also asked if the node technology could track pollinator patterns.
  - Pete stated that a few years back they visited a marine biological lab where they talked at length about the pollinator, and they talked about things like tracking individual bees and the like.
- Rich asked a similar question he asked of Geoffrey about operating system runtime environments. He noticed Pete mentioned docker containers and wanted to know if there were other things one needs to do to move data around this continuum.
  - Pete commented that they as scientists are of the supercomputing world, where for job scheduling and multi-tenancy they are used to owning nodes then giving them over to another user when done. But at the edge you need to share all the time – no one wants to wait for data. You have to balance these goals dynamically. You need to be connected

from the edge to the supercomputer and there are a lot of system software and networking problems that live in that space. There are also points in the middle, between edge and HPC that they need to focus on too.

- Miron stated that he has the same issues in reverse, as he is focused on the HPC and has the problems getting data out to the edge, where Pete is doing processing at the edge and needs to go towards the HPC. He stated that Oregon should look at what his group has done to target these issues. Miron also wanted to know how Pete established a trust between all these different pieces.
  - Pete stated the kinds of things that Open Science Grid are doing are very applicable in these end-to-end situations, so this is something to follow up on. With respect to trust they started having students build their own node and then volunteer it into Sage. So, for nodes you own you've established a trust system – know the certificates, the encryption, the end-to-end pieces.
  - Miron followed up with stating that someone might take it over and steal the certificate because there is no physical control.
  - Pete stated that these nodes are unprotected, someone could climb a pole and steal a node and imitate that node, sending sensor data. Pete admitted they are limited here unless they use more sophisticated security equipment.
  - Pete said in a nutshell, because you have the public in images it can't be disbursed to the public but can be available to government and academic agencies if the right waivers are signed. It's an agreement with the city of Chicago.
- Tom Gibbs wanted to ask about the ability to process streaming data.
  - Pete stated they are using an Nvidia jetson annex that runs in the 50-watt range. Indoors it just depends on how many servers they stack up. But for edge devices such as a telescope would it be an open science grid kind of thing or a multi-tenant interface?
  - Miron commented about how do you manage the capacity of all these things. So, it comes down to how do we allocate capacity.
  - Pete finished with is that this is a perfect R&D thing.

**Additional Agenda Item**

- Rich began the discussion by stating that the group needs to start planning for what we're going to do next year starting in October, and we will have to submit a slide to our parent organization Large Scale Networking Interagency Working Group to say, this is what we want you to task us with doing in fiscal year 2023. Rich went on to say that historically we have done multiple different things; we have done one off session each month, we have done things like we're doing now, which is a multi-session, virtual workshop scattered over several months.
- Starting at next month's meeting we need to come up with ideas and put names to them like we did with Geffrey Fox and Pete Beckman. He stated today he is putting a bug in the participant's ear so that they can think about ideas and presenters associated with ideas and provide discussion at the next meeting. Rich started off with a few preliminary ideas, such as a series on AI, machine learning, and deep learning in the various technologies and he wanted feedback from the group on what we could do in these areas.
- Rich also stated that in November there would be an in-person meeting at supercomputing in November in Dallas (referring to SC22) and he was expecting the MAGIC meeting to be run at this event. Rich then put it back to Pete and Geoffrey to comment.

  - Pete stated that think the programming model of this edge to the digital continuum is a big hard problem and the topic of AI programming model for that is very worthwhile.
  - Tom Gibbs offered the topic of multi-tenancy, based on some of the discussions today led by Miron.
  - Geoffrey provided AI for the instruments and felt that would be interesting.
  - Miron offered the subject of troubleshooting of software as for the providers it might go to the wrong place in a multi-tenant environment.
  - Pete added to that saying that in a “normal” HPC environment you know who the owner of that facility is but if you are in a remote place there is a little more trust needed, like what does that code do, and is it doing the right thing and using the data correctly. For example, for example, we don't allow X filtration of data, all the data has to go through established pipelines, so you always can track what how the data is being used.
- In a nutshell these are the ideas mentioned above: a series on AI/Deep Learning/ML, Programing Modules, Multitenancy, AI for the Edge Instruments, troubleshooting challenges.
- Rich stated that now MAGIC has two or three potential topics, so the next task is come up with a few more to discuss at the September meeting.

**Roundtable**

- The November MAGIC meeting is historically set to run at Supercomputing meeting, which will be in Dallas this year and will be in person most likely.
- Marcy Collinson: Oracle Research Fellowship just opened, and link is: https://blogs.oracle.com/research/post/five-reasons-to-apply-for-an-oracle-for-research-fellowship

**Next Meeting**

September 7th (12 pm ET)