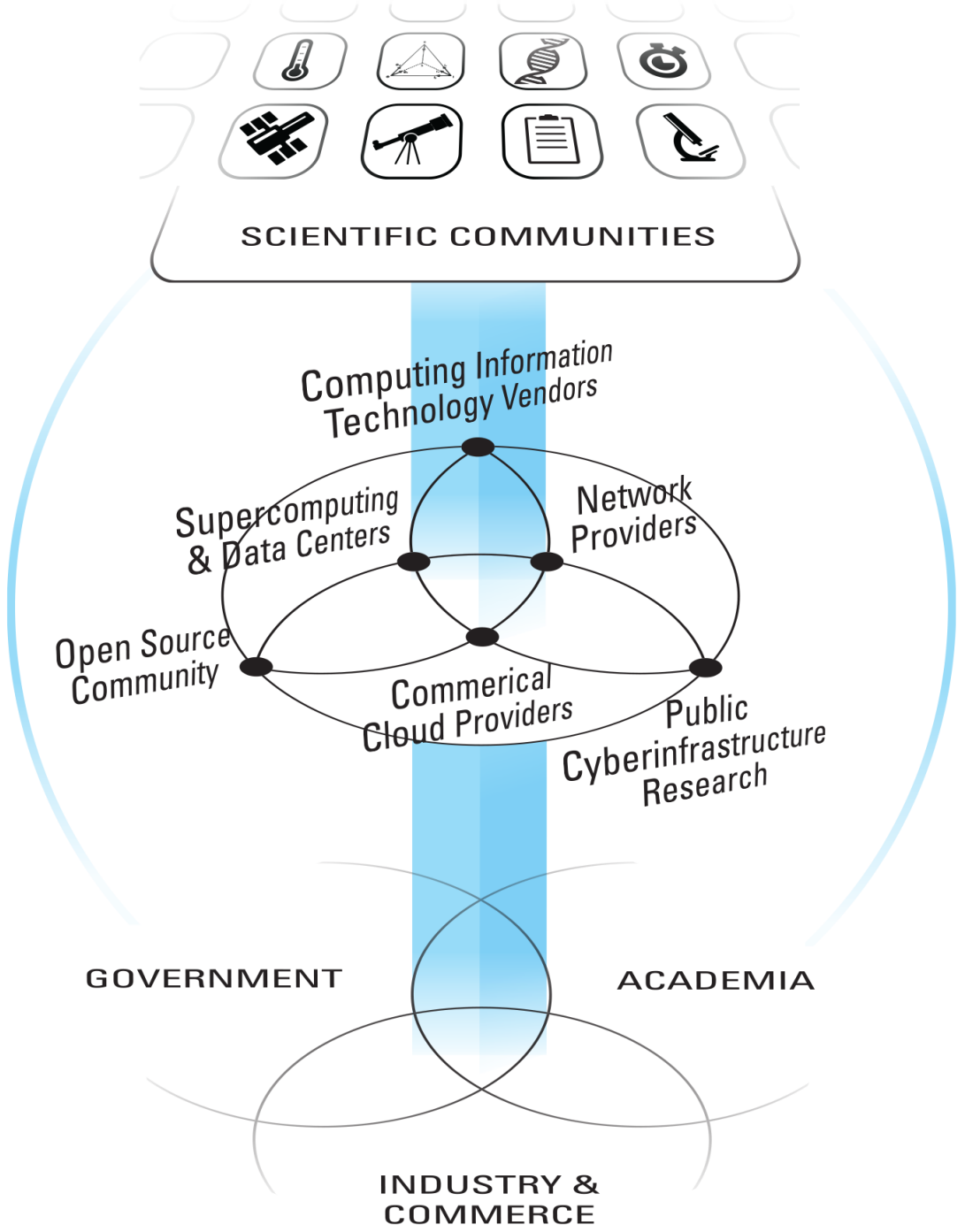
BDEC Wuxi 2017 Breakouts

Some suggested/possible question?

Terry Moore and Mark Asch

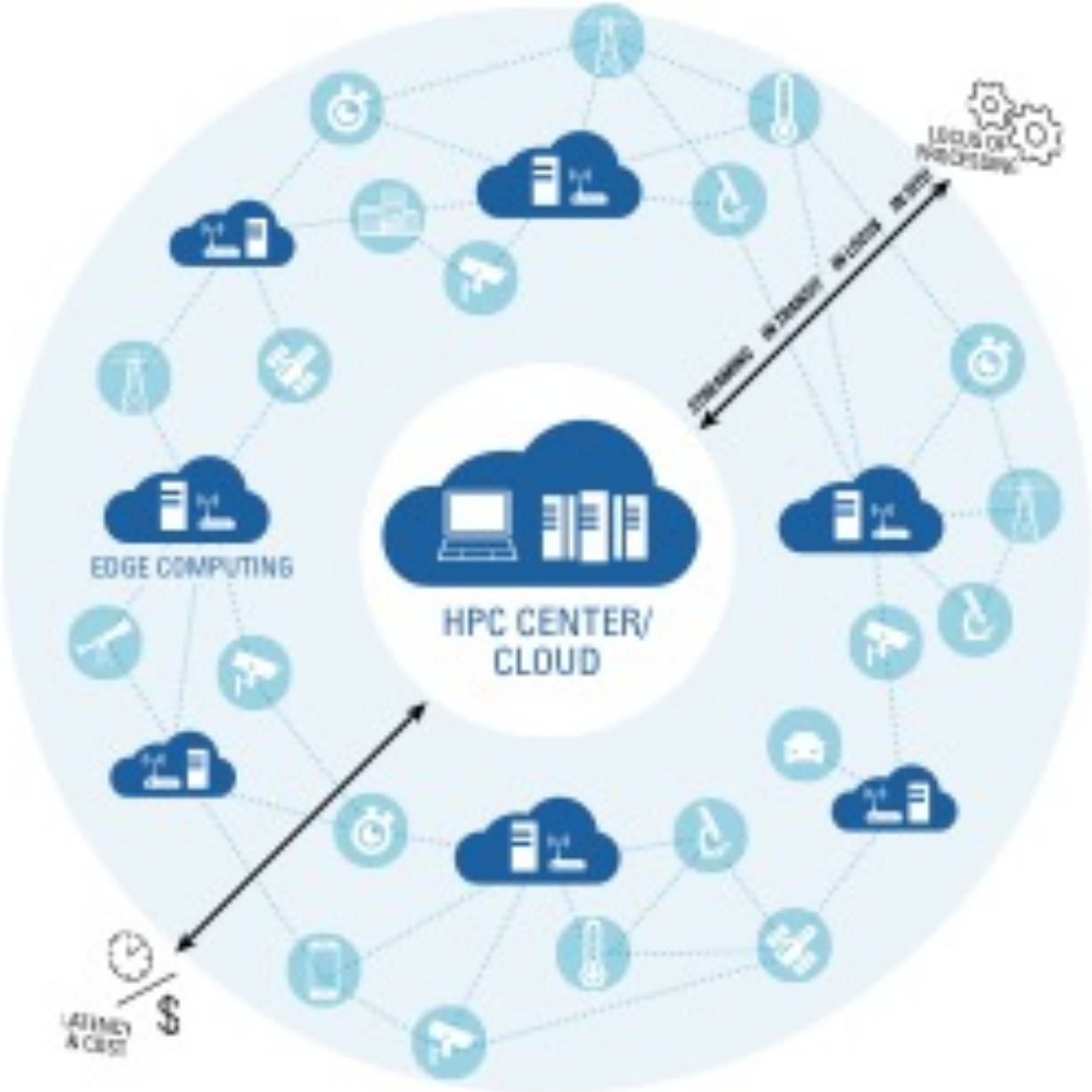
Who is the audience for this document?

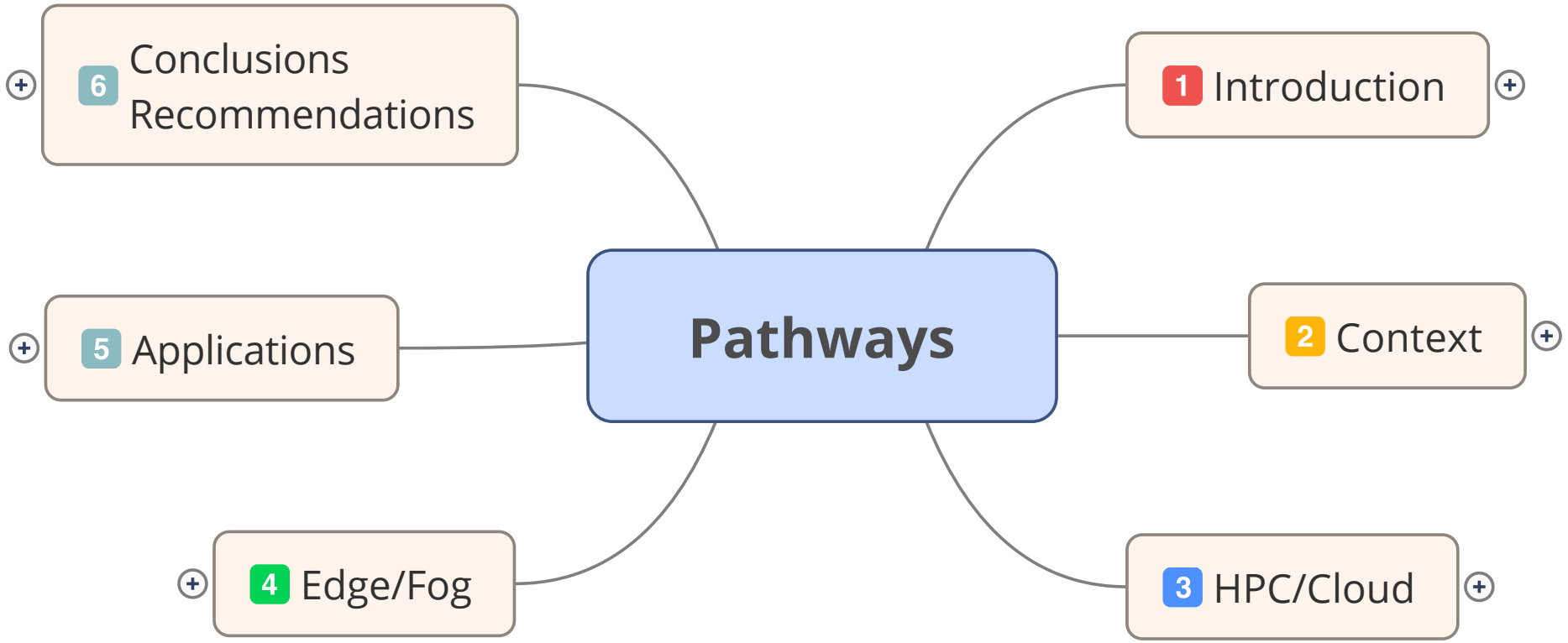
Here is one image of the BDEC stakeholders



Three breakouts

- HPC/Cloud Infrastructures (Physically or logically centralized resources)
- Edge/Cloud Infrastructures (Middle)
- Applications (Left)




Pathways
1 Introduction
2 Context
3 HPC/Cloud
4 Edge/Fog
5 Applications
6 Conclusions
Recommendations

Breakout - Table View

1. HPC/CLOUD	2. EDGE/FOG	3. APPLICATIONS
Architecture	Architecture	Workflows
Software	Software	HPC/Cloud
Hardware	Hardware	Edge/Fog
Data	Data	

HPC/Cloud: Kate, Satoshi, Thomas;
Edge: Ewa, Gabriel, Geoffrey;
Apps: Anshu, David, Jean-Pierre

Possible architecture questions

- What are the main workflow patterns? "Follow the data.”
- Visualization is part of the inference cycle. Should we think of it as part of the workflow?
- Are containers/virtualization a key technology for managing software complexity and the ”spanning layer” (aka ”narrow waist”)?
- Do we need to support resource federation? How?

Possible software questions

- What new types of standard libraries for need to be developed? (e.g., data reduction, etc.)
- How do traditional libraries need to be adapted for new application/workflow challenges?
- What software issues are involved in the major workflow strategies: streaming, in situ, in locus, in transit?
- How do we write programs for “the edge”?
- What can we learn from AWS Lambda functions and other new cloud services?

Possible hardware questions

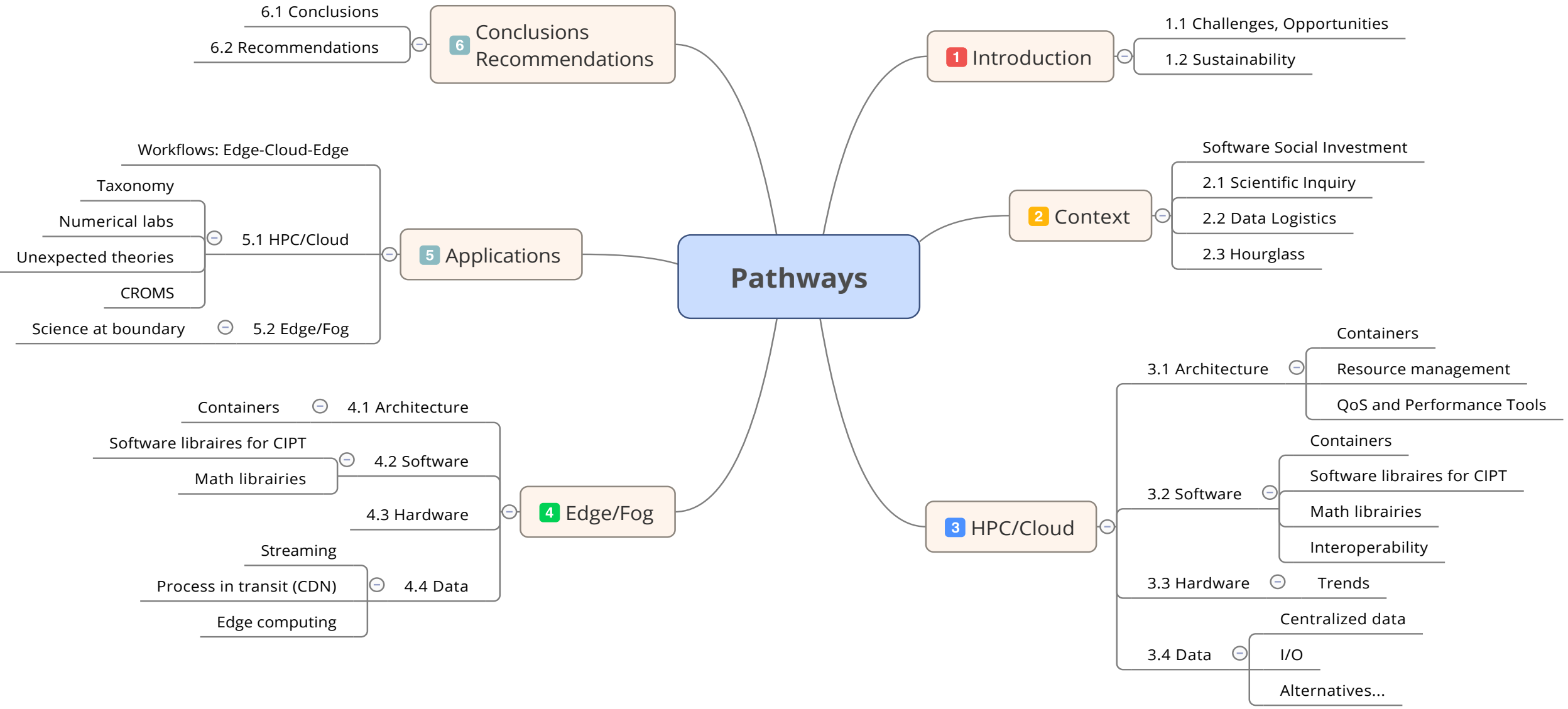
- Are there hardware designs which provide more balanced support for HDA and HPC?
- What are the hardware trends that will have the greatest impact on the design of new infrastructure and the applications they can support.
- How will DL hardware be integrated?

Possible workflow questions

- How does data staging and buffering fit in?
- What, if any are the characteristic ways in which workflows combine the use of resources in HPC/Cloud infrastructure, and resources in edge environments?
- How can we support workflows that move across different resource ecosystems (e.g. from edge, to Cloud, to HPC center)?
- Should we redefine "performance" to encompass the entire workflow, including all of data movement and intermediate processing involved?
- How does visualization fit into the data workflow model?

Questions for Deep Learning

- What are the key frameworks and workloads for DL?
- Is DL becoming a major element of scientific computing applications?
- What hardware and systems architectures are emerging for supporting DL?
- Is DL a distinct class worthy of its own software stack in the BDEC universe?


Pathways
1 Introduction
1.1 Challenges, Opportunities
1.2 Sustainability
2 Context
Software Social Investment
2.1 Scientific Inquiry
2.2 Data Logistics
2.3 Hourglass
3 HPC/Cloud
3.1 Architecture
Containers
Resource management
QoS and Performance Tools
3.2 Software
Containers
Software libraires for CIPT
Math librairies
Interoperability
3.3 Hardware
Trends
3.4 Data
Centralized data
I/O
Alternatives...
4 Edge/Fog
Containers
4.1 Architecture
Software libraires for CIPT
Math librairies
4.2 Software
4.3 Hardware
Streaming
Process in transit (CDN)
Edge computing
4.4 Data
5 Applications
Workflows: Edge-Cloud-Edge
Taxonomy
Numerical labs
Unexpected theories
CROMS
5.1 HPC/Cloud
Science at boundary
5.2 Edge/Fog
6 Conclusions Recommendations
6.1 Conclusions
6.2 Recommendations