

ASCR-BER Exascale Requirement Workshop Agenda

Monday, March 28

19:00 ** Chairs and Leads Pre-Meeting ** - Wilson Room

Tuesday, March 29

7:30 – 8:30: Registration

8:30 – 8:40: Welcome & Introduction, Steve Binkley, ASCR and Sharlene Weatherwax, BER

8:40 – 9:15: Genesis of this Meeting, Barb Helland, ASCR

9:15 – 9:45: Welcome — Todd Anderson and Gary Geernaert, BER

View from Biological and Environmental Research — Dorothy Koch and Ramana Madupu, BER

9:45 – 10:00: View from ESnet — Eli Dart, ASCR

10:00 – 10:30: Exascale Review Introductions — Chairs and Session Leaders

10:30 – 10:45: Break

10:45 – 12:00: ASCR Computing Facilities Presentation – Katherine Riley

12:00 – 13:00: Working Lunch — Charge to Working Groups – Chairs and Session Leaders

13:00 – 16:00: Breakout Sessions:

Climate: 13:00-14:00: **Climate and Environmental Science in the Exascale era (60 min) Leader: Ruby Leung**
What are the grand challenges for Climate and Environmental Science in the 5-10 year timeframe?
What are limiting factors to achieving results?
How would a 100x improvement in computational capabilities enhance research?
Plenary speakers: Dave Randall and Bill Collins

14:00-16:00: **Component-specific science and challenges (Leung) (2 hours; 3 parallel breakouts).**
For each component: What are the priorities and challenges in representing the system?
What advances are needed or anticipated?

- 1. Atmospheric research. Leaders: Minghua Zhang and Bill Gustafson**
- 2. Terrestrial research. Leaders: Peter Thornton and Dave Moulton**
- 3. Oceans and cryospheric research. Leaders: Todd Ringer and Wieslaw Maslowski**

Biology: 13:00-16:00: **From Sequence to Models of Organisms and Communities. Leaders: Dan Rokhsar and Rich Bonneau**

13:00-14:00: **From sequence to molecular structure and function**
Dan Rokhsar — Genome sequencing and diversity
Yang Zhang — protein structure prediction and annotation
Iddo Friedberg — functional annotation
Ingo Ruczinski — frontiers in primary data bioinformatics

14:00-15:30: **Integrative modeling of cells and networks**
Ilias Tagkopoulos — multi-omics network models

Marcus Covert — whole cell models: what's missing
Vassily Hatzimanikatis — metabolic modeling - a hybrid approach
Rich Bonneau — simultaneous estimation of networks and activities/function
Kerstin Kleese Van Dam — putting the genie back in the bottle

15:30-16:00: Discussions of sessions and white paper tuning

16:00 – 16:15: Break

16:15 – 17:30: Q&A Session with the BER and ASCR Associate Directors

17:30 – 18:15: Breakout leads present key questions and issues discussed in each session

18:30: Dinner on your own

Wednesday, March 30

8:00 – 8:30: Check-in

8:30 – 9:00: Summary from the chairs, outline of report

9:00 – 12:00: Breakout Sessions

Climate: 9:00-12:00: Coupled System Integration Challenges (3 hours; 1 hour in plenary, 2 hours in breakout) Leader: Collins
Working toward frameworks and hierarchies of modeling, e.g. ESM-IA-IAV parallel breakouts on:

1. ESMs. Leaders: Phil Jones and Nathan Urban
2. IA models. Leaders: Kate Calvin and Andy Jones

10:30-10:40: Break

Biology: 9:00-12:00: Big Data for Knowledge Discovery in Biology. Leaders: Kathy Yelick and Eoin Brodie

Big Data applications and challenges across scales

1. Assembly: Extreme scale metagenomic assembly (Buluc) 10 min
2. Multi-omics in complex environments: Metagenomics and proteomics of soil (Pan) – 10 min
3. Single cells- Big data: Understanding and utilizing cellular variability (Yosef) – 10 min
4. Discussion – 10 min

Combinatorial Analysis

1. Integration of multi-omics and interaction data for predictive modeling and analytics (Tagkopoulos) – 10 min
2. Combinatorial Applications for Systems Biology (Jacobson) - 10 min
3. BigData For Knowledge Discovery in Biology & Multiscale simulation of Biophysical Processes from Molecules to Biological Interfaces (Brown) – 10 min

10:30-10:40: Break

Computational bottlenecks

1. Mass Spectrometry Imaging and LC-MS: Data Storage, Access and Analysis (Bowen) – 10 min
2. Computational Challenges with Multi-omic Association Analysis (Mostafavi) – 10 min
3. Discussion – 5 min

Cross cutting themes in Big Data

1. Genomes to predictive biology: cognitive integration of intra and inter-kingdom functional genomics datasets for dynamic, real-time, big data contextualization and interpretation for gene function discovery (Kleese van Dam) – 15 min

2. Extreme-scale Partial Correlation Estimation for High Dimensional Data Analysis (SangOh/Buluc) – 10 min
3. Data Fusion (Wiley) – 10 min
4. Discussion on general challenges in big data - 20 min

Biology/EMSL: 9:00-12:00: Microbes to Environment (Tim Scheibe & Lee Ann McCue)

- 9:00 – 10:30:**
1. Brief overview/summary of the white papers, set goals of the session
 2. Flash talks from the participants (session participants)

10:30-10:40: Break

- 10:40 – 12:00:**
- Group discussion:**
Multiscale modeling approaches and challenges to link molecular and microbial scale data to predictive models at higher scales:
- Algorithm and code development requirements
 - Uncertainty quantification requirements
 - Hardware requirements

12:00 – 13:00: Working Lunch

13:00 – 16:00: Breakout Sessions

Climate: 13:00-16:00: Computational Issues for Climate and Environment Leader: Mark Taylor
 (3 hours in 3 parallel sessions, 1h per subtopic in parallel sessions).

1. Hour 1: Modeling climate and the environment with Exascale architectures. 3 parallel breakouts:

- | | |
|--|----------------------------------|
| a. Algorithms. | Leader: Kate Evans |
| b. Programming models. | Leader: Esmond Ng |
| c. Software engineering for portability. | Leader: Pat Worley and Rob Jacob |

2. Hour 2: Unlocking scientific knowledge within Exabytes of data. 3 parallel breakouts:

- | | |
|---|--------------------------|
| a. Large-scale heterogeneous data management. | Leader: Dean Williams |
| b. Observational data processing: retrieval algorithms, instrument network simulation | Leader: Pavlos Kollias |
| c. Data Reduction and Analytics of observations and model results | Leader: Bert Debusschere |
| i. UQ, statistics, emulators | |
| ii. Data analytics of model-observation comparisons | |

3. Hour 3: Data-Model Fusion. 2 parallel breakouts:

- | | |
|--|--------------------|
| a. Model development test beds. | Leader: Hsi-Yen Ma |
| b. Data Assimilation, Model Initialization and Reanalysis. | Leader: Gil Compo |

Schedule

	Breakout Room 1	Breakout Room 2	Breakout Room 3
Hour 1	Algorithms (1a)	Observational Data processing (2b)	Data Management (2a)
Hour 2	Programming Models (1b)	(no breakout)	Testbeds (3a)
Hour 3	SE for portability (1c)	UQ (2c)	Data Assimilation (3b)

Biology: 13:00-16:00: Multiscale Simulation of Biological Processes from Molecules to Biological Interfaces
Leaders: Jeremy Smith and Matt Jacobson

13:00-13:10: Overview of Scientific and Computational Challenges and Opportunities Leaders: Smith and Jacobson

13:10-14:00: Lightning Talks, 5 minutes each. Speakers: Margaret Cheung, Michael Crowley, Xiaolin Cheng, Loukas Petridis, Sergio Wong, Gnana Gnankaran, Ron Dror, Leslie Loew, Roland Schulz, Marat Valiev

14:00-15:30: Deep Dive on Selected Topics from Whitepapers and Case Studies

15:30-16:00: Cross-cutting Requirements and Complementarity with Other Topics

16:00 – 16:20: Break

16:20 – 17:20: Reports on Wednesday Breakouts, Breakout Leads, 15 minutes each

17:20 – 17:45: Summary and Thanks from Chairs

** End for Most Participants **

Thursday, March 31

All Day: ** Co-chairs, Leads, Writers meet to continue working on report **