IXPUG 2018 MIDDLE EAST MEETING

Argonne NATIONAL LABORATORY

ALCF SITE UPDATE

DAVID E. MARTIN
Manager, Industry Partnerships and Outreach
Argonne Leadership Computing Facility

MARK FAHEY
Director of Operations
Argonne Leadership Computing Facility

TIMOTHY J. WILLIAMS
Deputy Division Director,
Computational Science Division

ALCF SYSTEMS

Mira – IBM BG/Q	**Cetus** – IBM BG/Q	**Vesta** – IBM BG/Q	**Cooley** - Cray/NVIDIA
49,152 nodes 786,432 cores 786 TB RAM Peak flop rate: 10 PF	4,096 nodes 65,536 cores 64 TB RAM Peak flop rate: 836 TF	2,048 nodes 32,768 cores 32 TB RAM Peak flop rate: 419 TF	126 nodes 1512 Intel Haswell CPU cores 126 NVIDIA Tesla K80 GPUs 48 TB RAM / 3 TB GPU

Theta – Intel/Cray

4,392 nodes
281,088 cores
70 TB MCDRAM
679 TB DDR4
562 TB SSD
Peak flop rate: 11.7 PF

Production Storage Capability

HOME: 1.44 PB raw capacity

PROJECT:

- fs0 - 26.88 PB raw, 19 PB usable; 240 GB/s sustained
- fs1 - 10 PB raw, 7 PB usable; 90 GB/s sustained
- fs2 (ESS) - 14 PB raw, 7.6 PB usable; 400 GB/s sustained (not in production yet)

TAPE: 21.25 PB of raw archival storage [17 PB in use]

Argonne NATIONAL LABORATORY

THETA – ALCF'S KNL MACHINE

	Per node	18 cabinets June 2016	20 cabinets Jan 2017	24 cabinets Dec 2018
Cores	64	207,360	231,936	281,088
DDR4	192 GB	622,080 GB	695,808 GB	843,264 GB
HPL			5.88 PF	6.92 PF
HPL % peak			61.1%	59.2%
MCDRAM	16 GB	51,840 GB	57,984 GB	70,272 GB
Nodes		3,240	3,624	4,392
Peak		8.626 PF	9.648 PF	11.693 PF
SSD	128 GB	414,720 GB	463,872 GB	562,176 GB

NEW: NODE LOCAL SSDS AVAILABLE

- ALCF has enabled node local SSDs to provide temporary scratch storage for caching and processing workloads during job runs.
  - If you'd like to use this feature, project PIs can request access by contacting support@alcf.anl.gov. A use case will need to be provided.
  - For more information, visit: https://www.alcf.anl.gov/user-guides/xc40-file-systems and https://www.alcf.anl.gov/user-guides/running-jobs-xc40#requesting-local-ssd-requirements

NEW: SINGULARITY AVAILABLE

- Singularity 2.4.2, a virtualization framework, has been installed on Theta.
  - With Singularity, users can take their complete application environment and reproducibly run it anywhere Singularity is installed.
  - For information on using Singularity at the ALCF, visit: https://www.alcf.anl.gov/user-guides/singularity
- ALCF will create some known base images to help users
- Working with users to build more configurations

SINGULARITY ADVANTAGES

- Easy to install across diverse architectures and software stacks
  - x86, ppc64le, arm64
  - Cray, IBM, Torque, Moab, LSF, Cobalt, etc.
- Works well with old kernels (Titan is running kernel 3.0.x)
- Single file containers work with existing infrastructure
  - Share container using standard file techniques
- Interoperates well with Docker
- No root owned daemon processes required
- Portability with no performance penalty!

DOE's Pre-Exascale and Exascale Systems

Pre-Exascale Systems

2013

Argonne
IBM BG/Q
Open

ORNL
Cray/NVidia K20
Open

LLNL
IBM BG/Q
Secure

2016

Argonne
Intel/Cray KNL
Open

LBNL
Cray/Intel Xeon/KNL
Open

LANL/SNL
Cray/Intel Xeon/KNL
Secure

2018

ORNL
IBM/NVidia
P9/Volta
Open

LLNL
IBM/NVidia
P9/Volta
Secure

2020

NERSC-9

LBNL
TBD
Open

Crossroads

LANL/SNL
TBD
Secure

Exascale Systems

2021-2023

Argonne
Intel/Cray TBD
Open

Frontier

ORNL
TBD
Open

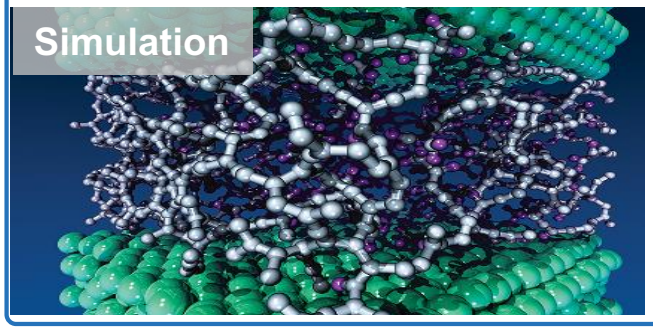
El Capitan

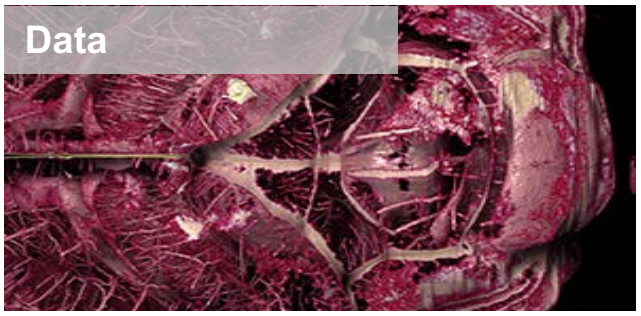
LLNL
TBD
Secure

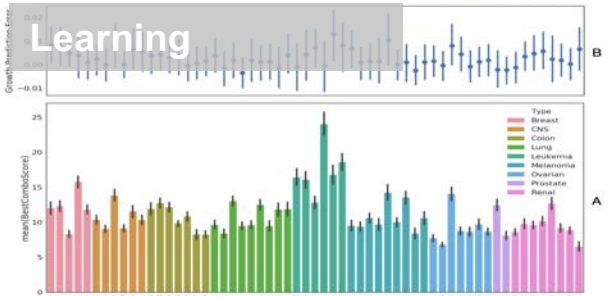
ALCF 2021 Exascale Supercomputer – Aurora

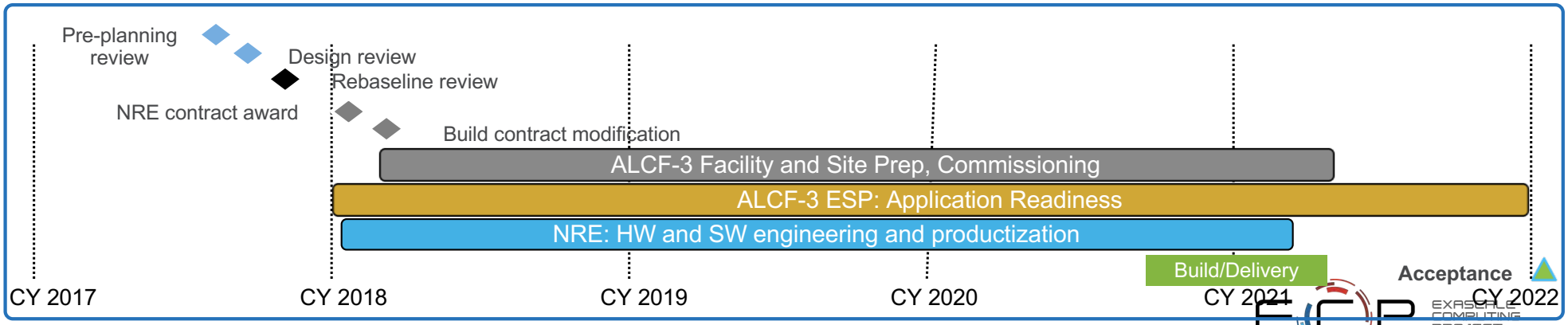
Intel Aurora supercomputer planned for 2018 shifted to 2021
Scaled up from **180 PF to over 1000 PF**

Support for three “pillars”







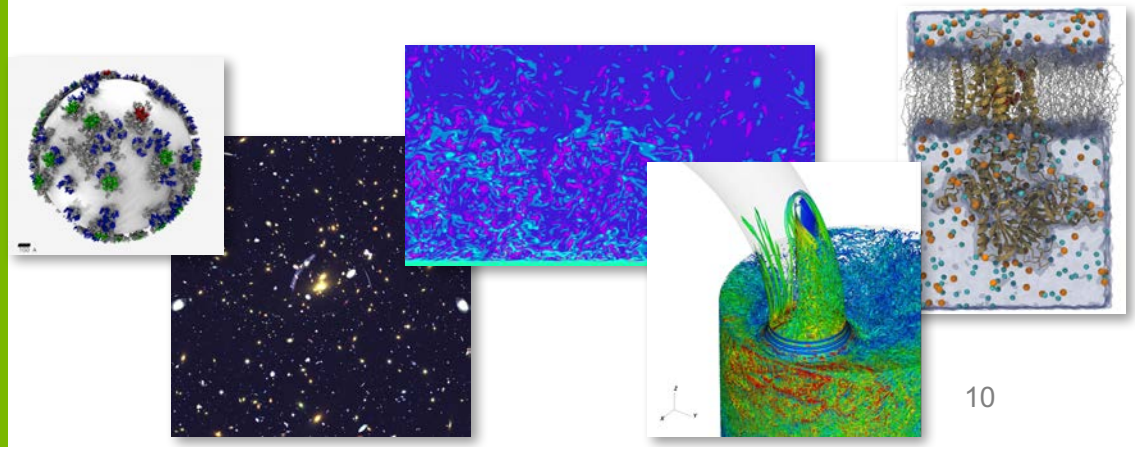


AURORA SYSTEM

- Nodes will have both high single thread core performance and the ability to get exceptional performance when there is concurrency of modest scale in the code
- Architecture optimized to support codes with sections of fine grain concurrency (~100 lines of code in a FOR loop e.g.) separated by serial sections
  - Degree of fine grain concurrency (e.g. number of loop iterations) that will be needed to fully exploit the performance opportunities is moderate. (~1000 for most applications)
  - Independence of these loops is ideal but not required for correctness
  - No limit on the number of such loops; overhead of starting/ending loops is very low
- Serial code (within an MPI rank) will execute very efficiently
- OpenMP 5 will likely contain the constructs necessary to guide the compiler to get optimal performance.
- The compute performance of the nodes will raise in a manner similar to the memory bandwidth
- The memory capacity will not grow as fast as the compute
  - The memory will all be high performance alleviating some concerns of explicitly managing multievel memory & data movement
  - The memory in a node will be coherent
- All compute will be first class citizens: equal access to all resources, memory and fabric etc.
- The fabric BW will be increasing similar to the compute performance for local communication patterns
  - Global communication BW will likely to not increase as fast as compute performance.

ALCF APPLICATIONS READINESS: EARLY SCIENCE PROGRAM

- Ten Years of Early Science have given us
  - Breakthrough science
  - Technical reports on code porting & tuning
  - Open community workshops (science & technology)
  - Synergy with Tools & Libraries project
  - Stable production platform (problems shaken out)
  - Persisting culture of apps readiness for next generation
  - Success stories for postdocs



ALCF AURORA EARLY SCIENCE PROGRAM

Applications Readiness

- Prepare applications for Aurora system:
  - Architecture
  - Scale - *Exascale*
- 5 Simulation (in progress), 5 Data, 5 Learning

Proposals

- Exascale science calculation
- Parallel performance
- Development needed for targeted science
- Appropriate team

CFP issued 10 January 2018
Deadline 8 April 2018

http://esp.alcf.anl.gov

Support

PEOPLE

- Funded ALCF postdoc
- Catalyst staff member support
- Vendor applications experts

TRAINING

- Training on HW and programming
- Community workshop to share lessons learned

COMPUTE RESOURCES

- Current ALCF systems
- Early next-gen hardware & simulators
- 3 months dedicated Early Science access
  - Pre-production (post-acceptance)
  - Large time allocation
  - Continued access for rest of year

CALL FOR PROPOSALS: A21 ESP DATA, LEARNING PROJECTS

- **CFP 10 January 2018**
  - Deadline 8 April 2018
- Selections June 2018
  - 5 Data projects
  - 5 Learning projects
- Two-year funded ALCF postdoc
- Cross-cutting proposals targeting the convergence of simulation, data and learning are very much encouraged.

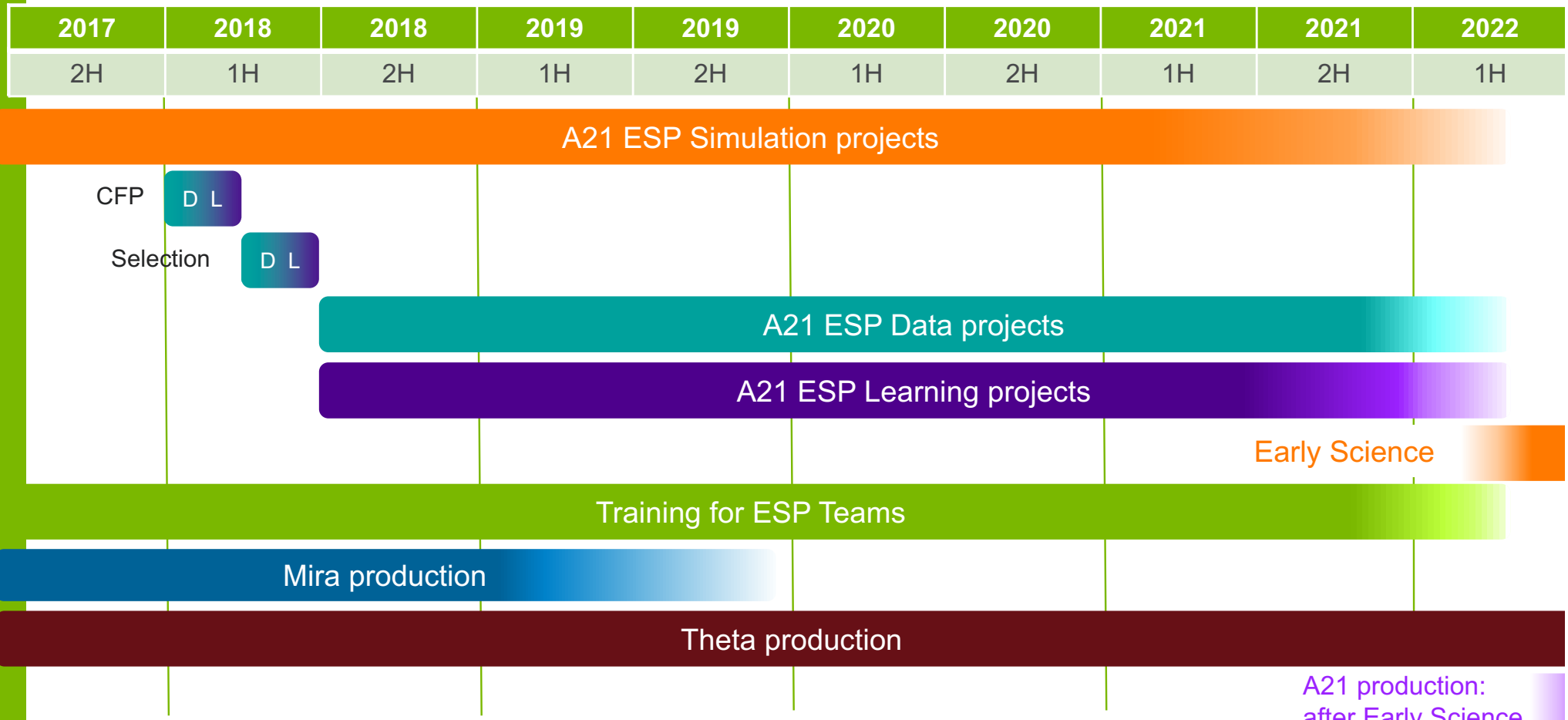
DATA

- Experimental/observational data
  - Image analysis
  - Multidimensional structure discovery
- Complex and interactive workflows
- On-demand HPC
- Persistent data techniques
  - Object store
  - Databases
- Streaming/real-time data
- Uncertainty quantification
- Statistical methods
- Graph analytics

LEARNING

- Deep learning
- Machine learning steering simulations
  - Parameter scans
  - Materials design
  - Observational signatures
- Data-driven models and refinement for science using ML/DL
- Hyperparameter optimization
- Pattern recognition
- Reduced model derivation
- Bridging gaps in theory

AURORA 2021 ESP TIMELINE



Argonne NATIONAL LABORATORY

THANK YOU

www.anl.gov

Argonne NATIONAL LABORATORY