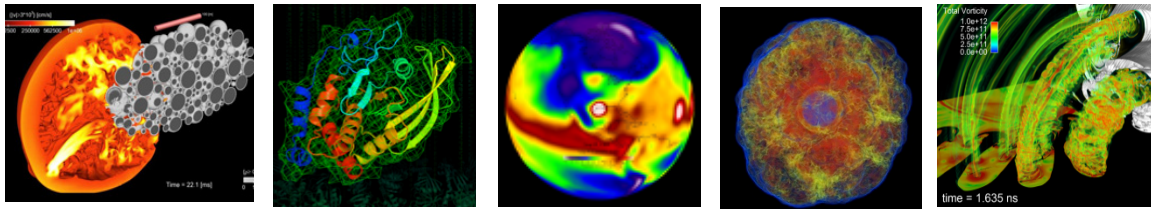
Trinity Advanced Technology System Overview



Manuel Vigil

Trinity Project Director

Douglas Doerfler

Trinity Chief Architect

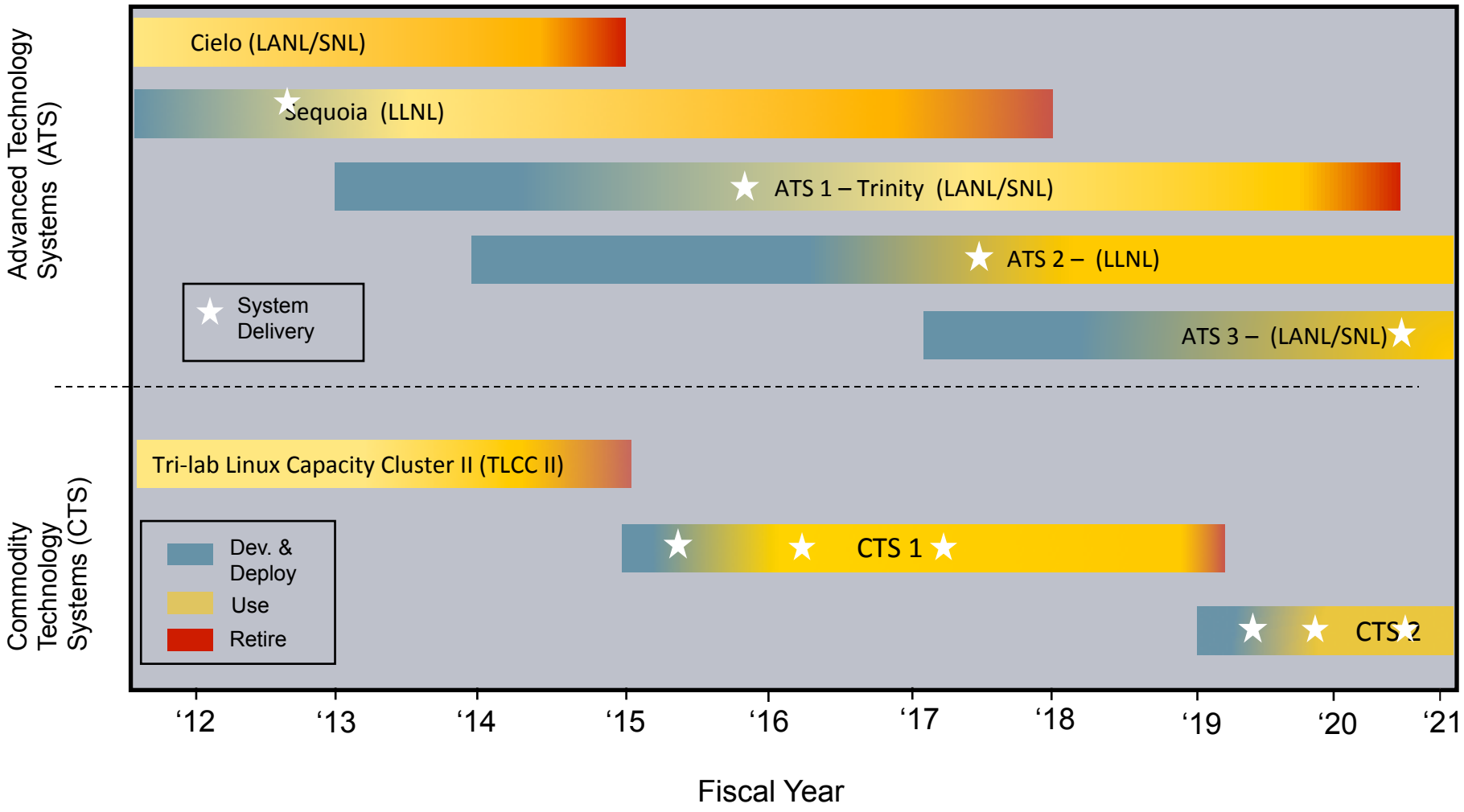
Outline

- ASC Computing Strategy
- Project Drivers and Procurement Process
- Platform Architecture Overview
- Schedule and Status
- Questions, and maybe some answers

ASC computing strategy

- Approach: Two classes of systems
  - Advanced Technology: First of a kind systems that identify and foster technical capabilities and features that are beneficial to ASC applications
  - Commodity Technology: Robust, cost-effective systems to meet the day-to-day simulation workload needs of the program
- Investment Principles
  - Maintain continuity of production
  - Ensure that the needs of the current and future stockpile are met
  - Balance investments in system cost-performance types with computational requirements
  - Partner with industry to introduce new high-end technology constrained by life-cycle costs
  - Acquire right-sized platforms to meet the mission needs

ASC Platform Timeline



Advanced Technology Systems

- Leadership-class platforms
- Pursue promising new technology paths with industry partners
- These systems are to meet unique mission needs and to help prepare the program for future system designs
- Includes Non-Recurring Engineering (NRE) funding to enable delivery of leading-edge platforms
- Trinity (ATS-1) will be deployed by ACES (New Mexico Alliance for Computing at Extreme Scale, i.e. Los Alamos & Sandia) and sited at Los Alamos
- ATS-2 will be led by LLNL, ATS-3 by ACES, etc

Trinity Project Drivers

- Satisfy the mission need for more capable platforms
  - Trinity is designed to support the largest, most demanding ASC applications
  - Increases in geometric and physics fidelities while satisfying analysts time to solution expectations
  - Foster a competitive environment and influence next generation architectures in the HPC industry
- Trinity is enabling new architecture features in a production computing environment (ATS Components)
  - Tightly coupled solid state storage serves as a "burst buffer" for checkpoint/ restart file I/O & data analytics, enabling improved time to solution efficiencies
  - Advanced power management features enable measurement and control at the system, node and component levels, allowing exploration of application performance/watt and reducing total cost of ownership
  - Trinity's architecture will introduce new challenges for code teams: transition from multi-core to many-core, high-speed on-chip memory subsystem, wider SIMD/vector units

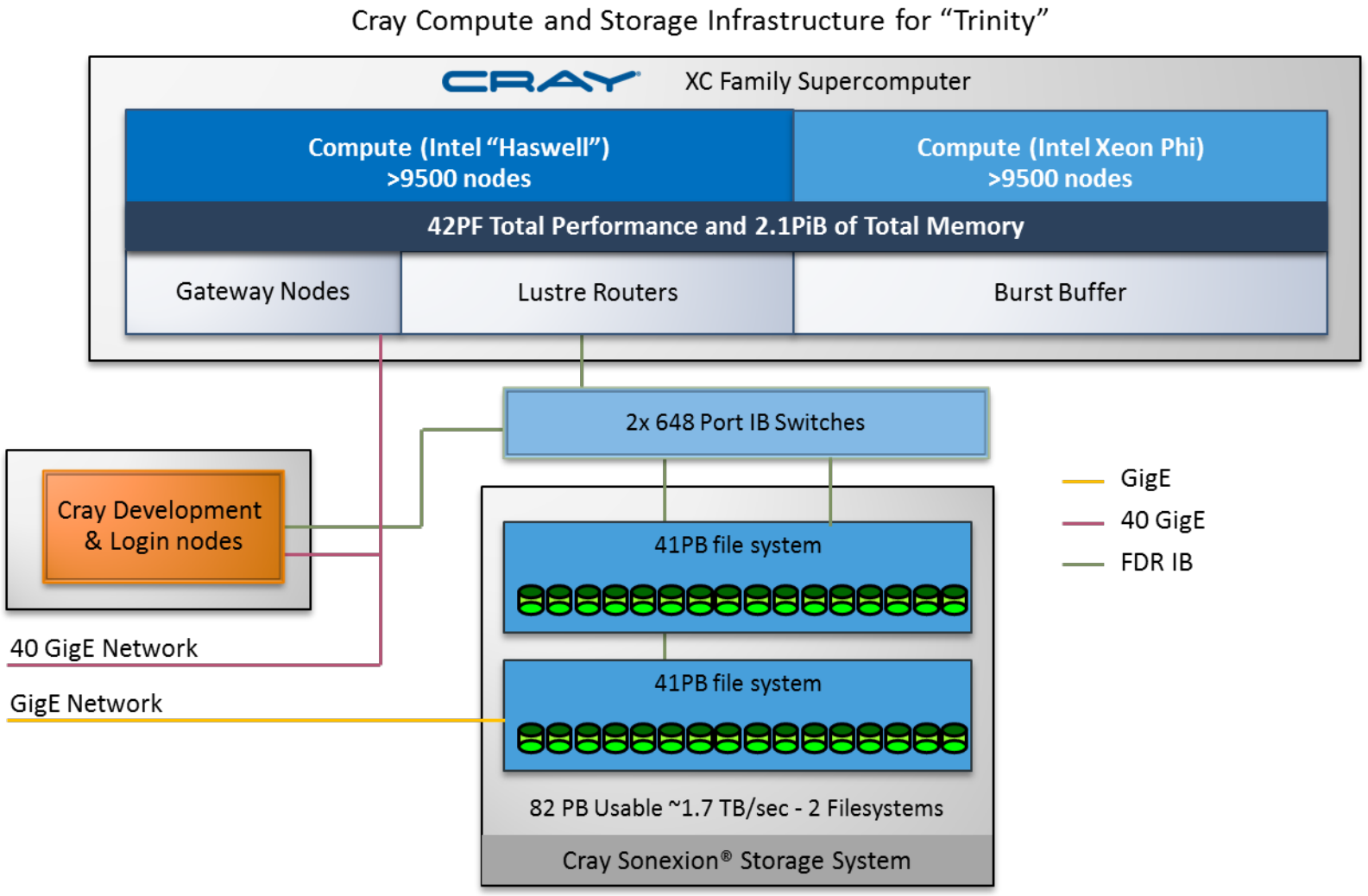
Trinity/NERSC8 Procurement Process Timeline

- ACES (LANL/SNL) Project started November 2011
- Market Survey started January 2012
- Partnered with LBL/NERSC on RFP (NERSC 8) March 2012
- CD-0, Draft Technical Requirements and RFI issued December 2012
- Formal Design Review completed April 2013
- Independent Project Review (Lehman) completed May 2013
- CD-1, Trinity/NERSC8 RFP issued August 2013
- Technical Evaluation of the proposals completed September 2013
- Initial negotiations for both systems completed November 2013
- NNSA Independent Cost Review completed Jan 2014
- CD-2/3, NERSC8 awarded April 2014
- CD-2/3, Trinity awarded July 2014 after Best and Final Offer (BAFO)

Trinity Platform Solution

- Cray has been awarded the contract, July 2014
  - Based on mature Cray XC30 architecture
- with Trinity introducing new architectural features
  - Intel Knights Landing processor
  - Burst Buffer storage nodes
  - Advanced power management system software enhancements
- A single system that contains both Intel Haswell and Knights Landing (KNL) processors
  - Haswell partition satisfies FY15 mission needs (well suited to existing codes) and fits the FY15 budget profile.
  - KNL partition delivered in FY16 results in a system significantly more capable than current platforms, provides the application developers with an attractive next generation target, and fits the FY16 budget profile.
- Managed Risk
  - Cray XC30 architecture minimizes system software risk and provides a mature high-speed interconnect
  - Haswell partition is low risk as technology is available Fall CY14
  - KNL is higher risk due to new technology, but provides a good path for codes teams to transition to many-core architecture

Trinity High-Level Architecture



Trinity Architecture Details

Metric	Trinity			
Node Architecture	KNL + Haswell		Haswell Partition	KNL Partition
Memory Capacity	2.11 PB		> 1 PB	>1 PB
Memory BW	>6 PB/sec		> 1 PB/s	>1PB/s + >4PB/s
Peak FLOPS	42.2 PF		11.5 PF	30.7 PF
Number of Nodes	19,000+		>9,500	>9,500
Number of Cores	>760,000		>190,000	>570,000
Number of Cabs (incl I/O & BB)	112			
PFS Capacity (usable)	82 PB usable	> 8x Cielo		
PFS Bandwidth (sustained)	1.45 TB/s	> 10x Cielo		
BB Capacity (usable)	3.7 PB			
BB Bandwidth (sustained)	3.3 TB/s			

Compute Node Specifications

	Haswell	Knights Landing
Memory Capacity (DDR)	2x64=128 GB	Comparable to Intel® Xeon® processor
Memory Bandwidth (DDR)	136.5 GB/s	Comparable to Intel® Xeon® processor
# of sockets per node	2	N/A
# of cores	2x16=32	60+ cores
Core frequency (GHz)	2.3	N/A
# of memory channels	2x4=8	N/A
Memory Technology	2133 MHz DDR4	MCDRAM & DDR4
Threads per core	2	4
Vector units & width (per core)	1x256 AVX2	AVX-512
On-chip MCDRAM	N/A	Up to 16GB at launch, over 5x STREAM vs. DDR4

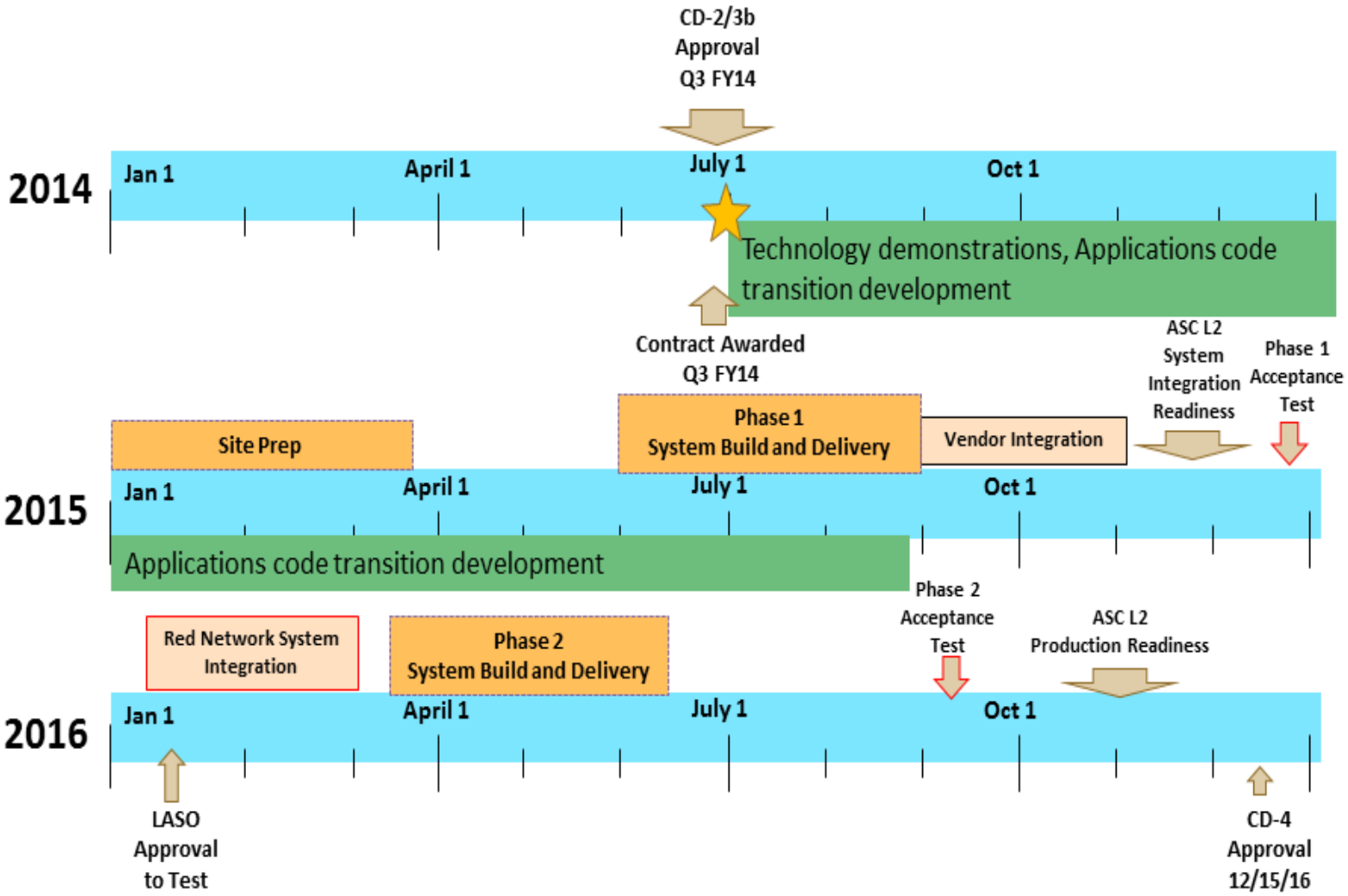
Trinity Capabilities

- Each partition will accommodate 1 to 2 large mission problems (2 to 4 total)
- Capability relative to Cielo
  - 8x to 12x improvement in fidelity, physics and performance
  - > 30x increase in peak FLOPS
  - > 2x increase in node-level parallelism
  - > 6x increase in cores
  - > 20x increase in threads

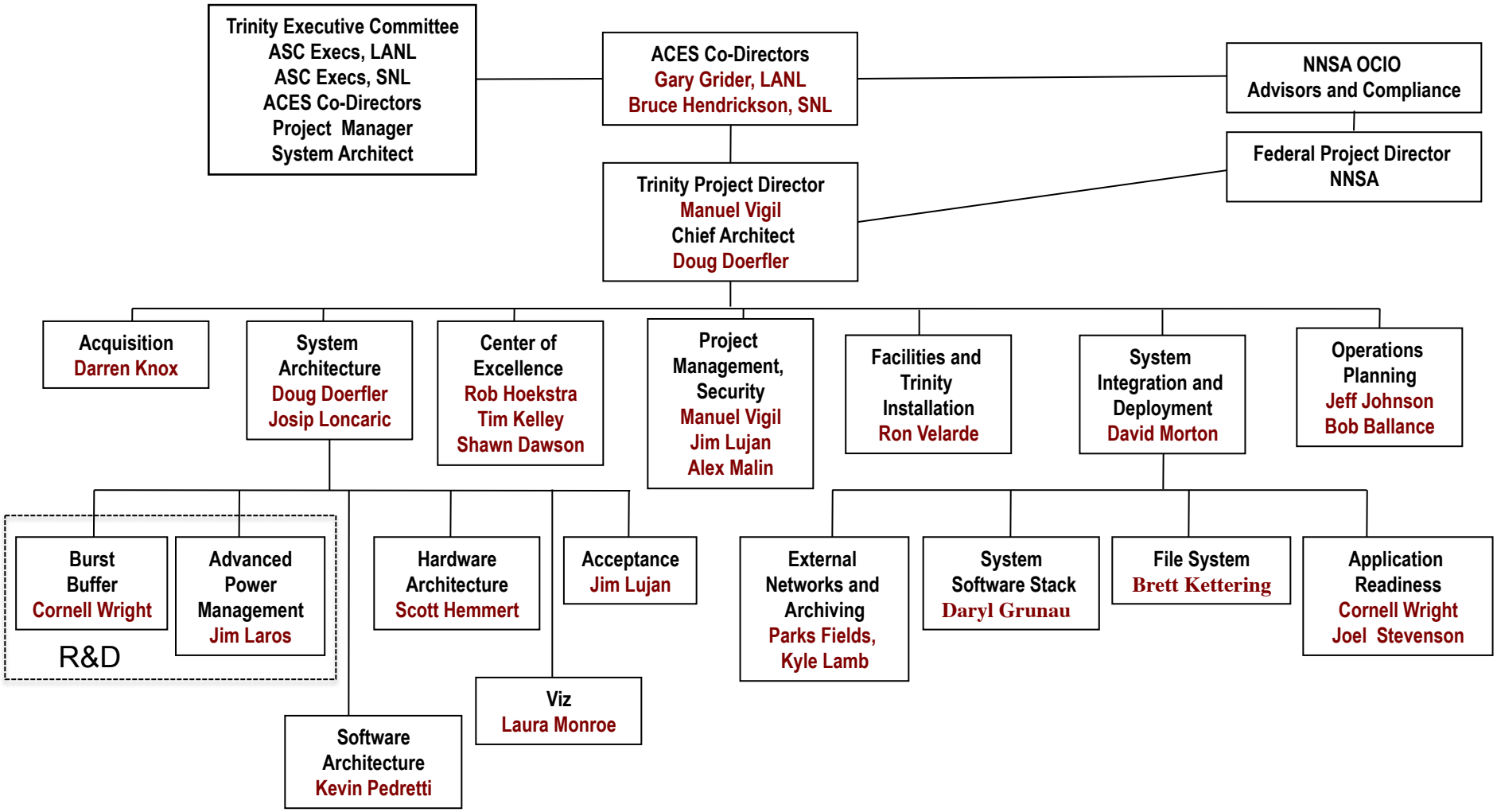
The Trinity Center of Excellence & Application Transition Challenges

- Center of Excellence
  - Work with select NW application code teams to ensure KNL Partition is used effectively upon initial deployment
  - Nominally one application per laboratory (SNL, LANL, LLNL)
  - Chosen such that they impact the NW program in FY17
  - Facilitate the transition to next-generation ATS code migration issues
  - This is NOT a benchmarking effort
- Intel Knights Landing processor
  - From multi-core to many-core
  - > 10x increase in thread level parallelism
  - A reduction in per core throughput (1/4 to 1/3 the performance of a x86-64 core)
  - MCDRAM: Fast but limited capacity (~5x the BW, ~1/5 the capacity of DDR4 memory)
  - Dual AVX-512 SIMD units
- Burst Buffer
  - Data analytics use cases need to be developed and/or deployed into production codes
  - Checkpoint/Restart should “just work”, although advanced features may require code changes

Trinity Platform Schedule Highlights 2014-2016



Trinity Project Team



Questions

Manuel Vigil

mbv@lanl.gov

Douglas Doerfler

dwdoerf@sandia.gov