

CORAL

COLLABORATION
OAK RIDGE • ARGONNE • LIVERMORE



CORAL-A Collaboration of
Oak Ridge, Argonne, and
Lawrence Livermore to
Procure Their Next
Generation Leadership
Computing Systems

Presented to
52nd HPC User Forum
April 9, 2014
Santa Fe, New Mexico

Arthur Bland
Project Director, OLCF



Dept. of Energy's Leadership Computing Facility

- Mission: Provide the computational and data science resources required to solve the most challenging scientific & engineering problems.
- Two centers with two architectures to address diverse and growing computational needs of the scientific community.
- Highly competitive user allocation programs (INCITE, ALCC). 3x oversubscribed.
- Projects receive computational resources typically 100x greater than generally available.
- LCF centers partner with users to enable science & engineering breakthroughs (Liaisons, Catalysts).



- IBM Blue Gene/Q many-core architecture
- Peak Performance: 10 Petaflops
- 49,152 Compute Nodes, 768 TB system memory
- File system: 35 PB, 240 GB/s bandwidth
- System accepted December 2012
- Cray XK7 – Hybrid CPU/GPU architecture
- Peak Performance: 27 Petaflops
- 18,688 Compute Nodes, 710 TB system memory
- File System: 32 PB, 1 TB/s Bandwidth
- System Accepted May 2013

CORAL

Sequoia

- **Mission:** To ensure the safety, security, and reliability of the U.S. nuclear weapons stockpile today and into the future without underground testing by strengthening the foundations of predictive simulation.
- **IBM Blue Gene/Q many-core architecture**
- **Peak Performance: 20 Petaflops**
- **98,304 Compute Nodes, 1.6 PB system memory**
- **File System: 55 Petabytes, 1 TB/s**



CORAL: Joint DOE Office of Science and National Nuclear Security Administration Leadership Computing Acquisition Project

Objective - Procure 3 leadership computers to be sited at Argonne, Oak Ridge, and Lawrence Livermore National Labs in 2017-2018

Leadership Computers run the most demanding DOE mission applications and advance HPC technologies

Approach:

Competitive process - one RFP (issued by LLNL) leading to 2 Non-Recurring Engineering (NRE) contracts and 3 computer procurement contracts

For risk reduction and to meet a broad set of requirements,

2 architectural paths will be selected – one each at Argonne and Oak Ridge. Livermore will then select one of the two.

Once selected, multi-year, lab-awardee relationship to deliver the best performance

Both NRE contracts jointly managed by the 3 Labs

Each lab manages and negotiates its own computer procurement contract, and may exercise options to meet their specific needs

Understanding that **long procurement lead time may impact architectural characteristics and designs** of procured computers

Advantages of Diversity

Having system diversity provides many advantages.

- It promotes price competition, which increases the value to DOE.
- It promotes a competition of ideas and technologies, which helps provide more capable systems for DOE's mission needs.
- It helps promote a rich and healthy high performance computing ecosystem, which is important for national competitiveness and DOE's strategic plan.
- Leading HPC is a priority goal in DOE's strategic plan.
- It reduces risk that may be caused by delays or failure of a particular technology or shifts in vendor business focus, staff or financial health.

The CORAL team worked with the HPC vendor community to ensure that the responses had sufficient diversity to provide the above advantages.

CORAL RFP Terminology

- **RFP Documents**
 - Statement of Work: SOW
 - Proposal Evaluation & Proposal Preparation Instructions: PEPPI
- **SOW Requirement Categories**
 - Mandatory Requirements (MR) – essential and must be bid to be considered
 - Mandatory Options (MO) – Options that must be bid to be considered
 - Technical Options (TO-1) – Important, but not required to be responsive
 - Target Requirements (TR-1, TR-2, TR-3) – features, components, performance characteristics, or other properties that are important, but will not result in nonresponsive bid if omitted

High Level System Requirements

- Target speedup over current systems of 4x on Scalable benchmarks and 6x on Throughput benchmarks
- **Peak Performance ≥ 100 PF**
- **Aggregate memory of 4 PB and ≥ 1 GB per MPI task (2 GB preferred)**
- **Maximum power consumption of system and peripherals ≤ 20 MW**
- **Mean Time Between Application Failure that requires human intervention ≥ 6 days**
- Data centric capabilities
- Delivery in 2017 with acceptance in 2018

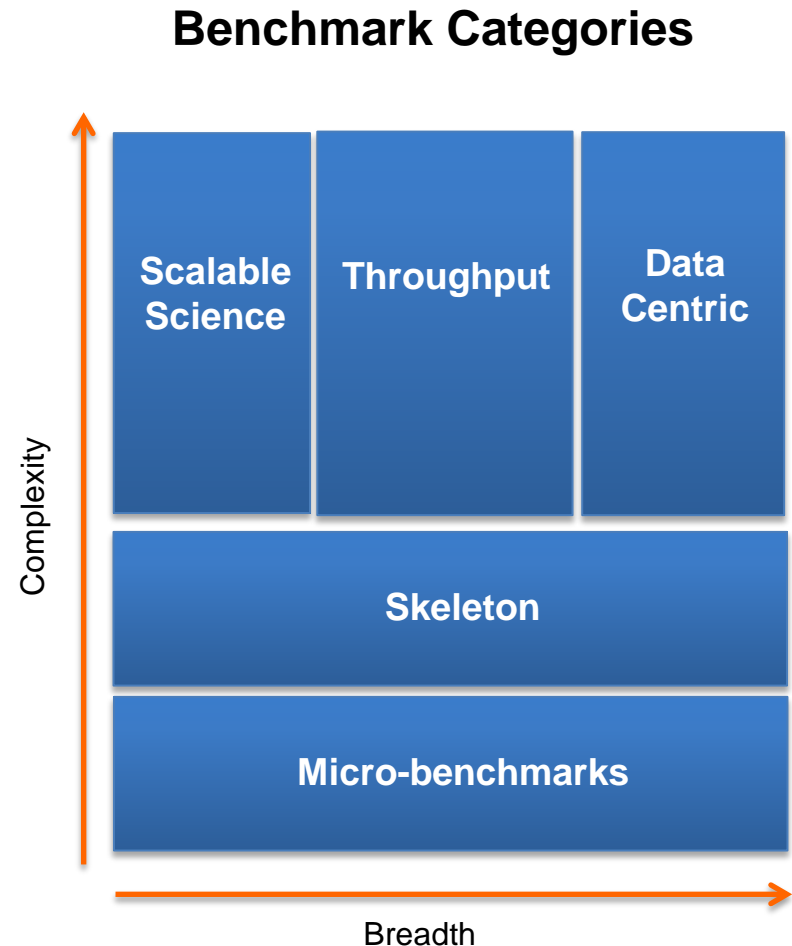
Technical requirements guiding principles

- Minimize number of Mandatory Requirements (MRs) and allow consideration of widest range of architectural solutions.
 - Word those requirements to allow architecturally diverse solutions
- Focus on requiring science and throughput performance. Avoid overly prescriptive explicit speeds and feeds.
- Agree on common technical requirements across all three Labs – not three separate sets of requirements.
 - Teams concur on requirements in each technical area
- Require vendors to describe available options to adjust system size and configuration to meet individual site needs and/or budgetary constraints.



CORAL benchmark categories represent DOE workloads and technical requirements

- Scalable science benchmarks
 - Expected to run at full scale of the CORAL systems
- Throughput benchmarks
 - Represent large ensemble runs; may be subsets of full applications
- Data centric benchmarks
 - Represent emerging data intensive workloads
 - Integer operations, instruction throughput, indirect addressing
- Skeleton benchmarks
 - Investigate various platform characteristics including network performance, threading overheads, I/O, memory, memory hierarchies, system software, and programming models
- Micro benchmarks
 - Small code fragments that represent expensive compute portions of some of the scalable science and throughput applications
 - Useful for testing programming methods and performance at the node level & for emulators and simulators



CORAL benchmarking suite uses mini-apps and a few larger applications

Categories	Scalable Science	Throughput	Data Centric	Skeleton
Marquee (TR-1)	LSMS QBOX NEKbone HACC	CAM-SE UMT2013 AMG2013 MCB	Graph500 Int sort Hashing	CLOMP IOR CORAL MPI Memory CORAL loops
Elective (TR-2)		QMCPACK NAMD LULESH SNAP miniFE	SPECint_ peak2006	Pydynamic HACC I/O FTQ XSBench miniMADNESS
Elective Micro-Benchmarks (TR-3)	NEKbonemk HACCmk	UMTmk AMDmk MILCmk GFMCmk		

Application performance requirements are the highest priority to CORAL

The goal is an average performance improvement over today's systems of:

- 4-8x for scalable science apps
- 6-12x for throughput apps

Two Step Evaluation Process

Step 1: Technical Evaluation

We created **eight teams of technical experts** with three people from each of the CORAL labs plus one from either LANL or Sandia to represent their users of the Livermore system.

1. Project Management
2. System Hardware
3. System Software
4. System Performance
5. Programming Environment
6. File System
7. Facilities and Operations
8. NRE

These 8 technical teams reviewed the proposals for three weeks ahead of a two day face-to-face meeting at ORNL to assess each of the proposals against the DRAFT SOW and PEPPI criteria

Two Step Evaluation Process

Step 2: Buying Team

A **Buying team** consisting of the management, technical, and procurement leadership of the three CORAL labs met to select the set of two proposals that provided the best value to the government

Evaluation Criteria:

- DOE mission requirements - the best overall combination of solutions
- Technical proposal excellence; projected performance on the applications is the single most important criterion
- Feasibility of schedule and performance
- Diversity
- Overall Price
- Supplier attributes
- Risk evaluation

Procurement approach designed specifically for acquiring leadership computers

- Long-term contractual partnership with vendors
- CORAL is a partnership for the long-term. All three labs are collaborating on both NRE contracts and Build contracts.
- Few mandatory requirements; many targets
 - Targets are performance levels that both parties reasonably believe can be achieved depending on the NRE; converted to traditional mandatory requirements after GO/NO-GO
 - Evaluate the level of achievement of the system as a whole rather than an assessment of performance of each target individually
- NRE contracts coupled to build contracts
 - NRE results are basis for GO/NO-GO decisions in build contracts
 - Modeled on the successful Sequoia/Mira partnership
- **Announcement of Evaluation Results cannot be made until contracts have been negotiated (expected late FY14)**

CORAL RFP

- <https://asc.llnl.gov/CORAL>
 - Provides all documents for the bidders
 - PEPPI – Proposal Evaluation & Proposal Preparation Instructions
 - DRAFT CORAL Build Statement of Work
 - Specific requirements for each laboratory
 - Worksheets for bidders to fill in to make it easier to compare information in a standard way
 - Bidder Questions and our Answers
- <https://asc.llnl.gov/CORAL-benchmarks/>
 - Provides all information on the benchmark codes

Questions?

