PathForward DRAFT Technical Requirements

March, 2016

Note: Proposals are not due at this time. The formal date for the submittal of proposals will be identified in the request for proposals (RFP). The RFP will be issued later at a later date.

Contents

Draft

1 INTRODUCTION

The U.S. faces serious and urgent economic, environmental, and national security challenges based on energy, climate, and growing security threats. High performance computing (HPC) is a requirement for addressing such challenges, and the need for the development of *capable* exascale computers has become critical for solving these problems.

- In 2016, the Department of Energy (DOE) began the Exascale Computing Project (ECP). ECP is a joint DOE Office of Science and DOE National Nuclear Security Administration (NNSA) effort focused on advanced simulation through a capable exascale computing program emphasizing sustained performance on relevant applications and data analytic computing to support their missions. The hardware efforts in ECP seek to realize capable exascale systems in the 2023-2025 timeframe by building on the existing DOE-SC and NNSA/ASC investments in the FastForward and DesignForward programs, software R&D, and application development and readiness projects.

- Separate from the Exascale Computing Project, DOE Office of Science and NNSA HPC laboratories will begin acquisitions for HPC exascale systems in 2019 with delivery occurring in 2023.

The ECP plan is structured into four focus areas:

Application Development: The exascale application development effort will create and/or enhance important DOE applications through development of models, algorithms, and methods; integration of software and hardware using co-design methodologies; systematic improvement of exascale system readiness and utilization; and demonstration and assessment of effective software/hardware integration.

Software Technology: To achieve the full potential of exascale computing, the software stack on which DOE SC and NNSA applications rely will be enhanced to meet the needs of exascale applications and evolved to utilize the features of exascale hardware architectures efficiently.

Hardware Technology: The Hardware Technology focus area supports vendor and lab hardware R&D activities required to design at least two capable exascale systems with diverse architectural features in support of HPC exascale system acquisitions.

Exascale Systems: This area bridges the gaps between the usual scope of the DOE HPC facilities and the extra resources required to field the first exascale systems. This focus area includes funding for non-recurring engineering (NRE) work beginning in 2019, supplemental acquisition funding, additional site preparations, and funding for prototypes and testbeds for application development and software testing. System procurement activities will be coordinated with the DOE HPC facility's existing 2023 system acquisitions.

The DOE laboratories and their associated companies participating in ECP are as follows:

- Argonne National Laboratory (ANL) managed by UChicago Argonne, LLC

- Los Alamos National Laboratory (LANL) managed by Los Alamos National Security, LLC,

- Lawrence Berkley National Laboratory (LBL), managed by the University of California

Draft

- Lawrence Livermore National Laboratory (LLNL), managed by Lawrence Livermore National Security, LLC

- Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC

- Sandia National Laboratory (SNL) managed by Sandia Corporation.

Instructions to Offerors:

This document defines draft technical requirements for the PathForward program, the central element of the Hardware Technology effort. PathForward is the follow-on to FastForward and DesignForward and runs through 2019.

PathForward is schedule constrained. PathForward seeks solutions that will improve application performance and developer productivity while maximizing energy efficiency and reliability of an exascale system. PathForward responses should describe R&D that will;

1. Substantially improve the competitiveness of the Offeror's HPC exascale system proposals in 2019, where application performance figures of merit will be the most important criteria.
2. Improve the Offeror's confidence in the value and feasibility of aggressive advanced technology options that they are willing to propose for the HPC exascale system acquisitions.
3. Identify the most promising technology options that would be included in 2019 proposals for the HPC exascale systems.

The period of performance for any subcontract resulting from this RFP will be through 2019.

Proposals shall include descriptions of independently priced work packages that focus on one or more specific component-level enabling technologies related to system or node design. Offerors must describe the path by which their system or node R&D could intersect an HPC exascale system delivered in the 2022-2023 timeframe. Offerors must also detail their vision for a complete conceptual system design and how the proposed research would accelerate or close gaps in their technology roadmap and provide capability to address DOE mission needs that existing market forces would otherwise not ensure. Offerors should also discuss how the proposed R&D will impact their commercialization/business strategy for their company's HPC and high performance data analytics projects.

Offerors may submit more than one proposal if they have more than one exascale system architecture in mind. In this case, the Offeror shall discuss how each architecture will impact their commercialization/business strategy for their company's exascale- and HPC-related projects.

For the purposes of PathForward a capable exascale system is defined as a supercomputer that can solve science problems 50X faster (or more complex) than on the 20PF systems (Titan, Sequoia) of today in a power envelope of 20-30 MW and is sufficiently resilient that user intervention due to hardware or system faults is on the order of a week on average.

While the focus of the R&D should be on exascale systems, if there are interim benefits that could feed into DOE's upcoming pre-exascale platform procurements, these should be described. While technology demonstrations are not required, they are desirable and can be proposed as optional work packages. Examples of a technology demonstrations include any demonstrable

hardware solution, perhaps with a minimal software layer, that can be used to show viability and/or measurable performance of the hardware or software technology.
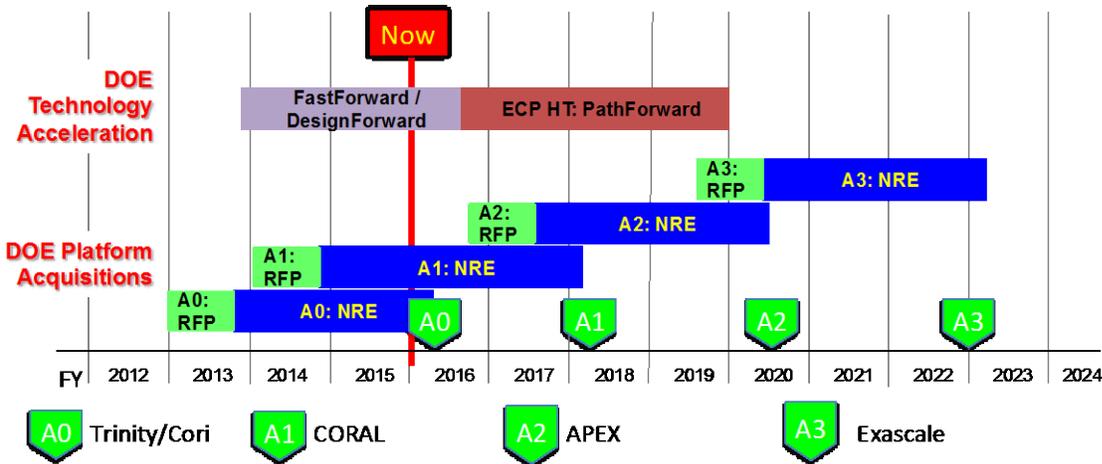

Figure 1: Path Forward RFP and its relationship to DOE system acquisitions roadmap and NRE.

Personnel from ANL, LANL, LBL, LLNL, ORNL, and SNL and their respective companies will have access to the proposals submitted for the PathForward requirements.

2   MISSION DRIVERS

2.1   The Department of Energy Office of Science Mission Drivers

The Department of Energy Office of Science (SC) is the lead Federal agency supporting fundamental scientific research for energy and the Nation's largest supporter of basic research in the physical sciences. The SC portfolio has two principal thrusts: direct support of scientific research and direct support of the development, construction, and operation of unique, open-access scientific user facilities. These activities have wide-reaching impact. SC supports research in all 50 States and the District of Columbia, at DOE laboratories, and at more than 300 universities and institutions of higher learning nationwide. The SC research portfolio covers a broad spectrum: basic energy sciences, biology and environment, fusion energy sciences, nuclear physics, high energy physics, and scientific computing. The SC user facilities provide the Nation's researchers with state-of-the-art capabilities that are unmatched anywhere in the world.

DOE's strategic plan calls for promoting America's energy security through reliable, clean, and affordable energy**,** strengthening U.S. scientific discovery, economic competitiveness, and improving quality of life through innovations in science and technology. In support of these themes is DOE's goal to advance simulation-based scientific discovery significantly, which includes the objective to "provide computing resources at the petascale and beyond, network infrastructure, and tools to enable computational science and scientific collaboration." All research programs within DOE SC depend on Advanced Scientific Computing Research (ASCR)

Program to provide the advanced facilities needed as the tools for computational scientists to conduct their studies.

Between 2008 and 2010, program offices within the DOE held a series of ten workshops[1] to identify critical scientific and national security grand challenges and to explore the impact that exascale modeling and simulation computing will have on these challenges. The extreme scale workshops documented the need for integrated mission and science applications, systems software and tools, and computing platforms that can solve billions, if not trillions, of equations simultaneously. The platforms and applications must access and process huge amounts of data efficiently and run ensembles of simulations to help assess uncertainties in the results. New simulations capabilities, such as cloud-resolving earth system models and multi-scale materials models, can be effectively developed for and deployed on exascale systems. The petascale machines of today can perform some of these tasks in isolation or in scaled-down combinations (for example, ensembles of smaller simulations). However, the computing goals of many scientific and engineering domains of national importance cannot be achieved without exascale (or greater) computing capability.

2.1.1 Advanced Scientific Computing Research Program

Within SC, the mission of the ASCR Program is to discover, develop, and deploy computational and networking capabilities to analyze, model, simulate, and predict complex phenomena important to DOE. A particular challenge of this program is fulfilling the science potential of emerging computing systems and other novel computing architectures, which will require numerous significant modifications to today's tools and techniques to deliver on the promise of exascale science.

2.2 National Nuclear Security Administration Mission Drivers

The NNSA is responsible for the management and security of the nation's nuclear weapons, nuclear non-proliferation, and naval reactor programs and generally ensuring America's nuclear security. It also responds to nuclear and radiological emergencies in the United States and abroad. Additionally, NNSA federal agents provide safe and secure transportation of nuclear weapons and components and special nuclear materials along with other missions supporting the national security.

The NNSA Stockpile Stewardship Program (SSP) requires much higher performance computational resources than are currently available. The current predictive capability of codes used by the NNSA Advanced Simulation and Computing (ASC) Program is a result of both scientific and engineering advances and the extraordinary increases in computing capability over the past two decades. While these codes support most of today's missions, they need to have greater predictive capability to support future missions. Aging of weapon components, advanced and additive manufacturing techniques, and changes resulting from alterations and Life Extension Programs (LEPs) are moving the stockpile further from the basis established via data collected in underground nuclear tests. Predictive capability is currently limited by approximations in the physics models used to simulate complex physical phenomena, the

---

[1] http://science.energy.gov/ascr/community-resources/workshops-and-conferences/grand-challenges/

inability to resolve critical geometric and physics features at very small length scales, and the need to further quantify margins and uncertainties. Making progress on these limitations requires NNSA to move beyond today's computer systems to *usable* exascale computing systems, which will allow systematic removal of some approximations, allow simulations to resolve substantially smaller length scales, and enable more accurate quantification of margins and uncertainties.

2.2.1   Advanced Simulation and Computing Program

Established in 1995, the ASC Program supports the NNSA SSP shift in emphasis from test-based confidence to simulation-based confidence. Under ASC, high-performance simulation and computing capabilities are developed to analyze and to predict the performance, safety, and reliability of nuclear weapons and to certify their functionality. As the nuclear stockpile moves further from the nuclear test base through either the natural aging of today's stockpile or introduction of component modifications, the realism and accuracy of ASC simulations must further increase through development of improved physics models and methods, requiring ever greater computational resources.

3   EXTREME-SCALE TECHNOLOGY CHALLENGES

Recent and ongoing analyses of capacity requirements across SC and NNSA establish an aggregate mission need of 2-10 exaflops of capability by the mid 2020s. Reaching this capability level has significant challenges because of the physical limits of existing computing technology, both hardware and software.

A 2008 DARPA report[2] described how energy consumption was a major impediment to building an exascale computer. Initial estimates noted such a system would require hundreds of megawatts (MW) of power to operate. Research investments by DOE and architecture redesigns by vendors have made significant strides to improve performance, but a number of key challenges[3] remain that must be addressed to bridge the capability gap and achieve exascale:

- *Parallelism:* System design must enable applications to exploit the extreme levels of parallelism effectively that will be necessary at exascale.
- *Resilience:* System-level resilience to both permanent and transient faults and failures must enable applications software to "work through" these problems to achieve successful, accurate, reliable execution and completion.
- *Energy Consumption:* Energy efficiencies must enable the entire system to operate within affordable power budgets when run at the targeted computational rates.
- *Memory and Storage Challenge*: Memory and storage architectures must enable applications to access/store information at high capacities and with low latencies to support anticipated computational rates.

These four challenges require investments in new architecture design and application development. The overriding objective of the ECP is to ensure hardware and software research

---

[2] http://www.cse.nd.edu/Reports/2008/TR-2008-13.pdf

[3] http://www.isgtw.org/feature/opinion-challenges-exascale-computing

and development, including applications software and system deployments, are completed in time to meet the scientific and national security mission needs of 2023.

4   APPLICATIONS

The ECP exascale applications, the initial set of which will begin development in earnest in 2016, must address key DOE strategic priorities in programs within SC, the applied energy offices, and NNSA Defense Programs. The development of applications targeting the mission space of other federal agencies, e.g., NIH, NSF, NOAA, and NASA, may also be supported by ECP. ECP application development is expected to create or enhance predictive capability through targeted development of requirements-based models/algorithms/methods, integration of appropriate software and hardware co-design technologies, impactful improvement of exascale system readiness and utilization, and demonstration of effective software/hardware integration and challenge problem capability. While the key challenges in many fields have been effectively articulated in the workshop reports referenced in Section 1, the development of models, algorithms and software is dependent on how hardware and software stack designers respond to the challenges of building an exascale system and vice versa. The ECP plans to resolve many possible trade-offs in the space of applications, software and architectures by using the co-design methodology described in the next section. The ECP will weigh the potential benefits of a new hardware feature against the costs of software impacts (e.g., the need to rewrite or to refactor code) that are incurred to make use of the new hardware feature.

5   ROLE OF CO-DESIGN

5.1   Overview

The R&D funded through this acquisition is expected to be the product of a co-design process. Co-design refers to a system-level design process where scientific problem requirements influence architecture design and technology, and architectural characteristics inform the formulation and design of algorithms and software.

Co-design methodology requires the combined expertise of vendors, hardware architects, software stack developers, domain scientists, computer scientists, applied mathematicians, and systems staff working together to make informed decisions about the design of hardware, software, and underlying algorithms. The *transformative* co-design process is rich with trade-offs, and give and take will be needed from both the hardware and software developers. Understanding and influencing these trade-offs is a principal co-design requirement.

The ECP has a strong co-design theme that runs through all its focus areas. In addition, the ECP will establish several new co-design centers around common mathematical motifs, in contrast to the ASCR and ASC co-design centers, which were established around specific applications. The PathForward acquisition for hardware R&D is ECP's process for supporting vendor engagement in our co-design effort.

6 REQUIREMENTS

6.1 Description of Requirement Categories

Requirements are either mandatory (Mandatory Requirements - designated MR) or target (Target requirements - designated TR-1, or TR-2), and are defined as follows:

- MRs are essential requirements. An Offeror must satisfactorily address *all* MRs to have its proposal considered responsive and eligible for further evaluation.

- TRs are important but will not result in a nonresponsive determination if omitted from a proposal. TRs add value to a proposal and are prioritized by dash number. TR-1 is more desirable than TR-2.

TR-1s and MR are of equal value. The aggregate of MRs and TR-1s form a baseline solution. TR-2s are goals that boost a baseline solution. Taken together as an aggregate, MRs, TR-1s, and TR-2s form an enhanced solution.

6.2 Requirements for Research and Development Investment Areas

Examples that define the scope of possible PathForward R&D investment areas are provided in Attachment A. Attachment A provides examples of items that are in scope, but recognizing the unique nature of each Offeror solution, the Offerors are requested to self-identify areas that require research investment to improve their Exascale system offerings for DOE. Each proposal shall address all of the MRs listed below, in the order listed.

The various R&D work packages within a proposal are all expected to contribute to their vision for an overall exascale architecture.

6.3 Proposal Page Limit

The total length of a proposal, excluding cover letter, cover page, table of contents, references and staff curricula vitae (CVs), shall not exceed fifty pages, with a minimum text font size of 11 points and margins no smaller than one inch on all sides. Tables, figures, appendices, and attachments to the multiple work packages in the proposal are included in the 50 page limit.

If an Offeror has more than one exascale system conceptual design, then an Offeror can submit one proposal for each design.

6.4 Mandatory Requirements

The following items are mandatory for all proposals.

6.4.1 Exascale System Description (MR)

Offeror shall describe a complete exascale system design including the conceptual node design that would be proposed for the HPC exascale system acquisitions in 2019.

Offerors shall discuss the innovative nature of the proposed exascale conceptual system design R&D and describe where it differs from their company roadmaps. Work that funds a company's current roadmap is not acceptable. The primary intent is to fund long-lead-time R&D objectives

that overcome the extreme-scale technology challenges that were described in Section 3 of this document.

6.4.2 Prioritized list of work-packages (MR)

Provide prioritized list of independent work-packages that can address the Extreme-scale Technology Challenges identified in Section 3. For each item in the list, provide the following summary information;

1. **Technology Area**: Aspect of the system that this technology targets (e.g., parallelism, resilience, energy consumption, memory and storage);

2. **Area of innovation**: Short (a few words) description of the technology innovation that is proposed in this work package;

3. **Cost Summary**: Estimated cost and timeline for this work package;

4. **Impact:** Estimated probability that an investment in this area will significantly increase the competitiveness of your technology offering for a 2019 exascale systems procurement RFP response;

5. **Risk:** The likelihood of success/failure of this technology option.

6.5 Target Requirements

6.5.1 Work Packages (TR-1)

Offeror shall provide independently priced work packages for conducting the proposed R&D, including timelines, milestones, and proposed deliverables. Deliverables shall be meaningful and measurable. **Pricing shall be assigned to each milestone and deliverable**. A schedule for periodic technical review by the DOE laboratories shall also be provided. A work breakdown structure shall be provided.

Quantitative measures of design innovations are desired. Deliverables will include simulations, analyses, or hardware demonstrators that assess the impact (or feasibility) of a proposed innovation in the conceptual system design.

For each work package, describe the following:

- **Describe the technical challenge**: Performance opportunity or risk in conceptual design that requires DOE investment.

- **Value proposition**: In recognition of the diversity of potential R&D options, the Offeror shall define a Figure of Merit (FOM) and describe how much the FOM can be improved through PathForward investments. The FOM must be *specific*, *measurable*, and *attainable* with quantitative goals. The selection committee will not be able to assess a poorly defined FOM or a metric with weak or difficult to evaluate performance targets. A value proposition that is not compelling to DOE cannot be selected.

- **Remedy**: Provide a description of the specific technological remedy to the identified challenge.

- **Work Plan:** What concrete steps are required in development of the technological solution that meets the FOM and describe any co-design required.

- **Cost and schedule:** Please provide quarterly milestones, dependencies on other work packages or technologies, exit conditions, and costs for this R&D package.

6.5.2 Prioritized List of Work Packages across Proposals (TR-1)

If the Offeror has more than one Exascale system design and submits more than one proposal, then a separate document shall prioritize all work-packages across the proposals, which shall not count against the page-count of the Offeror's proposals.

6.5.3 Impact and Risks (TR-2)

The following additional information would be useful to include in your Work Package.

- The Offeror shall describe the impact of their technology work package on the programming environment and/or describe what the programming interface is to access the new technology feature if relevant.

- The Offeror shall describe how the Work Package will achieve the goals of increasing the performance of key DOE extreme-scale applications relative to energy usage while maintaining or increasing reliability and maintaining or decreasing runtimes.

- The Offeror shall describe for any hardware technology that requires software changes, the level of effort required to move existing applications and components of the software stack to the new system design. For these cases, the Offeror should also discuss the impact of staying with existing programming models (primarily in terms of performance but perhaps other metrics). Software engineering effort and sacrifice are important parts of our evaluation of the potential improvement enabled by new hardware that requires software modifications.

6.5.4 Productization Strategy (TR-2)

Offeror shall describe how the proposed conceptual system design will be commercialized, productized, or otherwise made available to customers and in what timeframe.. Offerors shall include identification of target customer base/market(s) for the technology. Offerors shall describe impact specifically on the HPC market as well as the potential for broad adoption. Solutions that have the potential for broader adoption beyond HPC are highly desired (e.g., for data centers, or other commercial markets that might benefit from this technology). Offerors shall indicate a projected timeline for productization.

7 EVALUATION PROCESS

7.1 Evaluation Team

The Evaluation Team, which will make recommendations to ECP, includes representation from six DOE laboratories: Argonne National Laboratory, Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Los Alamos National Laboratory, Oak Ridge National Laboratory, and Sandia National Laboratories. The evaluation team will also include procurement and contract representatives from the entity awarding subcontracts as a result of this RFP on behalf of DOE SC and NNSA, one of whom will act as the source selection official.

Draft

7.2 Evaluation Factors and Basis for Selection

The Offeror's proposal should identify and discuss the performance features and supplier attributes that will be important to the Offeror's successful performance and the attainment of the PathForward project objectives. ECP intends to select work packages from the responsive and responsible Offerors whose proposals contain the combination of price, performance features, and supplier attributes offering the best overall value to DOE. The Evaluation Team will determine the best overall value by comparing differences in performance features and supplier attributes offered with differences in price, striking the most advantageous balance between expected performance and the overall price. Offerors must, therefore be persuasive in describing the value of their proposed performance features and supplier attributes in enhancing the likelihood of successful performance or otherwise best achieving DOE's objectives for extreme scale computing. The primary aspect of this overall evaluation will be the extent to which the proposed activities will directly improve offerings available in response to the anticipated exascale systems procurement.

7.3 Performance Features

The Evaluation Team will validate that an Offeror's proposal satisfies the MRs. The Evaluation Team will then assess if, and how well, an Offeror's proposal addresses the TRs. An Offeror is not solely limited to discussion of features described in TRs. **An Offeror may propose other features or attributes if the Offeror believes they may be of value.** If the Evaluation Team agrees, consideration may be given to them in the evaluation process. In all cases, the Evaluation Team will assess the value of each proposal as submitted.

The technical/management proposal should contain a comprehensive discussion of how the Offeror will fulfill the MRs and TRs and successfully perform the Subcontract. The Evaluation Team will evaluate the following performance features as proposed:

- The extent to which the proposed hardware architecture solution will impact offerings available in response to an exascale systems procurement RFP;

- The degree to which the technical proposal meets or exceeds any TR;

- The degree of innovation in the proposed R&D activities;

- The extent to which the proposed R&D achieves substantial gains over existing roadmaps of the Offeror in particular and the industry in general;

- The extent to which the R&D requires software changes, and the level of effort required to move existing applications to the new system.

- The extent to which the proposed R&D will impact HPC and the broader marketplace;

- Credibility that the proposed R&D will achieve the stated results;

- Credibility of the productization plan for the proposed technology;

- Realism and completeness of the project work breakdown structure.

7.4 Feasibility of Successful Performance

The Evaluation Team will assess the likelihood that the Offeror's proposed research and development efforts can be meaningfully conducted and completed within the subcontract period of performance. The Evaluation Team will also assess the risks to successfully fielding an Exascale system that meets DOE mission needs, to both the Offeror and the DOE laboratories, associated with the proposed solution. The Evaluation Team will evaluate how well the proposed approach aligns with, accelerates, or closes gaps with the Offeror's corporate roadmap and the level of corporate commitment to the project.

7.5 Supplier Attributes

The Evaluation Team will assess the following supplier attributes.

- The Offeror's experience and past performance engaging in similar R&D activities. The Offeror may include in its proposal a written description of up to three recent contracts that the Offeror successfully completed, similar in type and complexity to the scope of the proposed Subcontract.

- The Offeror should have adequate financial resources to perform the Subcontract or the ability to obtain them.

- The Offeror's demonstrated ability to meet schedule and delivery promises.

- The alignment of the proposal with the Offeror's product strategy.

- The expertise and skill level of key Offeror personnel.

- The contribution of the management plan and key personnel to successful and timely completion of the work.

- If work is subcontracted, how much is to small businesses, the extent to which the assignment of work scope appropriate, and to which the prospective lower-tier subcontractor is qualified to do that work.

7.6 Price of Proposed Research and Development

The Evaluation Team will assess the following price-related factors:

- Reasonableness of the proposed work package prices in a competitive environment;

- Proposed price compared to the perceived value;

- Price tradeoffs and options embodied in the Offeror's proposal;

- Financial considerations, such as price versus value.

7.7 Gains over Industry Roadmaps

The most meritorious improvements will make substantial gains over industry roadmaps/trends and substantiate a convincing path to achieving the extreme-scale technology characteristics required by the DOE.

The list of characteristics includes improvements to:

- Application performance and/or time to solution

- Power and/or energy

- Resilience

- User productivity

- Scaling

- Total cost of ownership

- Total delivered value for exascale workflows

ATTACHMENT A: PathForward R&D Examples

A   Areas of Interest

The areas below apply to supercomputers that will be deployed by DOE Office of Science and NNSA. Each Offeror may have different areas of focus based on the goal of providing the strongest possible Exascale system proposal. **Again, we wish the Offeror NOT to respond directly to this list of interest areas – rather the Offeror should self-identify and prioritize the areas that would provide the most impact.** The list of areas in this appendix provides examples of the scope of responses, and is by no means intended to restrict or to change priorities for the Offeror's response. The following are examples of objectives and technologies that could be considered in PathForward R&D proposals. Some of the items below may only apply to certain architectures, and some may be mutually exclusive. Furthermore, this list of topics is not exhaustive. **Offerors need not propose R&D for all of these topics, and may propose alternative topics within these areas of interest.** However *all* proposals must describe the proposed architecture of a conceptual exascale system and how their proposed set of highest priority R&D work packages would impact that system.

A.1   Overall System Architecture

The Offeror shall describe their vision for conceptual exascale system architecture and how their proposed prioritized set of R&D work packages fit into that vision. If the proposal only includes certain technology components and is meant to complement another vendor's conceptual system architecture, enough detail must be provided to clearly illustrate how the proposed R&D would have impact. Examples of system design R&D efforts that may close or accelerate a gap between the Offeror's current technology roadmap and the needs for ECP:

- Designs that simplify changing or upgrading specific node capabilities (e.g., processors, memory, coprocessors) or that enable node substitution in the face of faults that may degrade or kill nodes

- Mechanisms to increase flexibility in resource utilization such as ways to share memory capacity across nodes

- Mechanisms to mitigate the tension between production system use, which primarily entails large jobs, and software development for the system, which involves non-computational tasks such as compilation and short jobs for testing and debugging

- Designs that facilitate compiling for a mix of heterogeneous nodes

- Mechanisms to support isolation and flexibility in resource association, e.g., partitioning resources among jobs

- Techniques to support efficient scheduling of diverse resource types

- Scalable, adaptive, and unobtrusive monitoring, with real-time analysis of platform state

- Real-time autonomic platform management under production workloads, gracefully handling unplanned events without requiring immediate human intervention

A.2   Memory Architecture

Memory capacity has continued to increase, but memory latency and bandwidth have not kept pace. In order to provide the improved memory performance to the multi-core processors or compute accelerators, vendors are exploring stacked memory solutions, but to address capacity requirements, new memory sub-systems will likely have to use multi-level architectures. In addition, these memory architectures are likely to incorporate emerging non-volatile memory technologies. The deep memory hierarchies will require science applications to be rethought and refactored to maximize data locality, and to minimize data movement and input/output (I/O) in order to accomplish the mission needs. However, the potential performance improvement must be balanced against increased programming complexity relative to prioritizing overall mission needs.

Memory sub-system designs are desired for improved memory performance that can avoid the programming complexity of multi-level architectures.

A.3   Node Designs

Research into component and/or node architectures should include planning for how an exascale system can be built with that technology. Therefore, proposals for component architecture research should include milestones that call for the Offeror to make contact with one or more potential system integration teams (either within the Offeror's company or externally) and establish the feasibility of building an exascale system from the proposed technologies.

Mechanisms for producing a highly optimized node that has a tight coupling of components and are highly optimized for HPC are desired. Solutions should describe how they would achieve this goal. Examples of node design R&D include:

- Development of mechanisms to understand and influence the trade-offs between power, resilience and performance, statically (when the node is designed), or statically (when the node is booted), and dynamically (at runtime)

- Integration of standard building blocks into a balanced node architecture for HPC, and alternate configurations of component building blocks for node architectures that may be targeted for high performance data analytics

- Heterogeneous node designs with a mix of large and small cores and support for dynamic (runtime) configuration

A.4 Energy Utilization

The target requirement is a system that achieves high performance on a broad range of DOE applications while minimizing energy use. Solutions should target 20 MW (peak) at system scale while achieving application performance and reliability requirements. Examples of node or system design energy R&D include:

- Designs that improve power efficiency

- Techniques for measurement, runtime control and application control of power utilization

- System-wide and site-wide power management methodologies

- Advances that support high-performance, power-efficient processor integration with memory, optics, and networking

- Techniques to reduce cooling energy requirements

A.5 Resilience and Reliability

The target ECP resilience targets will require processor designs that lead to a mean time to application failure requiring user or administrator action of six (6) days or greater in an exascale system, as determined by estimates of system component FIT rates and application recovery rates. The overhead to handle automatic fault recovery should not reduce application performance by more than half. Examples of system resilience design R&D include:

- Designs that improve the resiliency or reliability, for example, improved fault detection, containment, correction, and response time

- Methods that enable dynamic adaptation to a constantly changing system

- Leveraging of hardware/software resilience synergies to improve overall time to solution

- Techniques to improve fault detection accuracy (e.g., fewer undetected errors) and root cause analysis or to reduce their cost and time to repair/recovery

- Framework for representing hardware and system software dependencies, for interpretation of failure modes and autonomic reasoning about remaining recovery paths

- Integration of data management RAS features into the overall system-wide RAS capability

A.6 Data Movement

The performance of applications depends upon many factors such as message injection rates and contention. Offerors should describe improvements to the rates of data movement through all layers of the data hierarchy that optimize application performance.

Solutions should address how best to balance these memory systems in terms of bandwidth and capacity within a node to optimize for application performance and programmer productivity at minimum cost. The Offeror should describe in detail how this will be accomplished.

Solutions should discuss how new functionality enabled by tight integration will contribute to increased communication efficiency. Examples of data movement design R&D include:

- Designs that increase memory capacity without sacrificing bandwidth or latency performance.

- Designs that allow extremely low-latency multi-hop messages

- Improvements to the performance and energy efficiency of messaging, remote memory access, and collective operations

- Analysis of the optical/copper tradeoffs to improve data movement across the system

- Reliable low-energy, long-distance data movement

- Efficient data movement for computation and also across levels of the storage hierarchy

- Mechanisms to avoid contention and to provide QoS guarantees (bandwidth, latency, reliability, etc.)

- Advances to enable efficient latency hiding techniques

A.7   Concurrency

Examples of node and system design R&D to increase concurrency efficiency include:

- Advances that improve the scalability of processor designs as the number of processing units per chip increase

- Advances that improve the inherent scaling and concurrency limits in applications

- Advances that improve the efficiency of process or thread creation and their management

- Advances that reduce the synchronization and activation time of large numbers of on-chip threads or across heterogeneous devices

- Advances to assist in identification of active performance constraints within the system, such as latency or throughput limited sections, memory and network bottlenecks

A.8   Programmability and Usability

While the key challenges in many fields have been effectively articulated in the workshop reports referenced in Section 1, the development of models, algorithms and software is dependent on how hardware and system software designers respond to the challenges of building an Exascale system and vice versa. DOE's aim is to optimize many possible trade-offs in the space of both applications and architectures effectively to ensure that the resulting exascale platforms lead to improved scientific output and greater machine efficiency.

Describe novel features of the hardware that allow applications to use the proposed architecture more efficiently. Highlight descriptions of how the proposed solutions increase performance without increasing programmer effort or software revalidation.

The areas that contribute to scientific delivery that will be considered include but are not limited to:

- Application execution performance

- Application/algorithm scaling

- Managing application load imbalance

- Power/energy consumption

- Perceived application/system reliability

- System density and efficient cooling

Proposals that seek to provide large improvements to scientific delivery with higher application impact as well as more modest increases at lower application impact will be considered to allow Offerors to balance risk and opportunities across a range of work packages. Areas that may be considered include:

- Advances that significantly improve the performance and energy efficiency of arithmetic patterns common to DOE applications but are not well supported by today's processors, for example, short vector operations such as processing in vector registers

- Advances that allow efficient computation on irregular data structures (for example, compressed sparse matrices and graphs)

- Research to determine the most effective option(s) for cache and memory coherency policies; configurable coherency policies and configurable coherence or NUMA domains may be options; coherency policies might also be a power management tool

- Research on efficient mapping of multiple levels of application parallelism to node architecture parallelism

- Advances in software and hardware that allow a user or runtime system to measure and to understand node activities and to adjust implementation choices dynamically

- Advances that enable a programmer to understand and to reason about optimally programming the node, and that expose the right architectural details for consideration; development of a target independent programming system

- Advances that minimize the impact of hardware complexity on software and reduce the cost and time required for revalidating large software systems.

- Execution models that enable the programmer to perceive the system as a unified and naturally parallel computer system, not merely as a collection of microprocessors and an interconnection network

- Programming and execution models that provide for runtime support of the coexistence of threading among all the supported languages (C, C++, Fortran, etc.) within the application and any supporting libraries

A.9 Data and Workflow Management

Describe the data management capabilities of their conceptual system architecture, including workflow scheduling and integration with system-level RAS framework where applicable.

Draft

- Development of integrated models for use of in-system nonvolatile storage (which could be located anywhere within the system architecture) including integration with programming models, abstractions to exploit locality awareness and security models

- Methods to integrate scientific workflows into system resource managers, including abstractions to describe data management resource requirements (storage, bandwidth, etc.) in a non-system dependent way.