SAND2011-4202C

Science Partnership for Extreme-scale Computing
June 2011

Los Alamos National Laboratory
Oak Ridge National Laboratory
Sandia National Laboratories

SPEC builds on previous collaborations

- **The Los Alamos/Sandia Alliance for Computing at the Extreme Scale (ACES)**
  - **ACES is deploying the Cielo Petascale capability platform for ASC**

- **The Oak Ridge/Sandia Institute for advanced Architectures and Algorithms (IAA)**
  - **IAA led to an ASCR CS/math institute, early funding for co-design activities**

- **The Oak Ridge/Los Alamos Hybrid Multicore Consortium (HMC)**

SPEC has been very been very active since its inception

- Initial meeting at SOS14 in March, 2010
- Weekly Tri-lab telecons
- Four way NDAs signed with 7 companies
- MOU signed by laboratory directors – November, 2010
  - Co-directors are Jeff Nichols, Andy White and Sudip Dosanjh
- Numerous meetings with potential industry partners
  - >30 meetings with computer companies (dozens of SPEC-industry telecons as well)
- Defining a SPEC technology roadmap that will advance the HPC ecosystem
- SPEC co-design effort on climate modeling

Elements of SPEC's Strategy

- **Create viable Exascale industry partnerships that advance the HPC ecosystem**
- **Build a broad coalition of support**
- **Identify cross-cutting issues and technologies (e.g., memory, silicon-photonics, programming models, file systems)**
- **Use competition to identify the best technical solutions**
- **Develop mechanisms to enable co-design (includes technical and IP considerations)**

SPEC-industry discussion points include:

- Pre-Exascale systems must be representative of the Exascale systems
  - Programming continuity (i.e., no revolutionary programming change between pre-Exascale and Exascale systems)
- Constraints
  - 1 EF
    - Specify a performance goal for targeted DOE applications (e.g., an average with a minimum)
  - Power must be <20 MW
  - >64 PB of memory (may be multiple levels)
  - Mean time between job interrupts on the order of a day
  - System cost < $200M
  - R&D cost < ??
- Co-design methodology and IP
- Performance portability across different systems through a common programming model and architectural abstraction

A few technical observations from our discussions…

Heterogeneous multicore nodes are in our future

intel

Home › Newsroom › News Stories ›

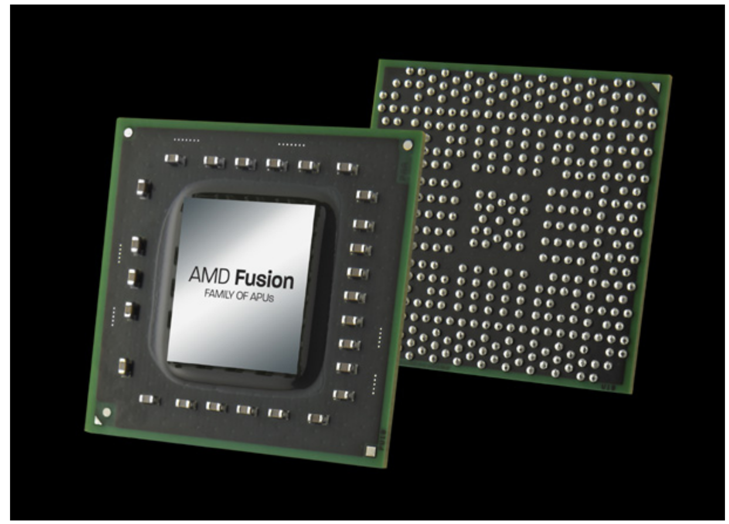
Intel News Release

383

retweet

Intel Unveils New Product Plans for High-Performance Computing

Intel® Many Integrated Core Chips to Extend Intel's Role in Accelerating Science and Discovery



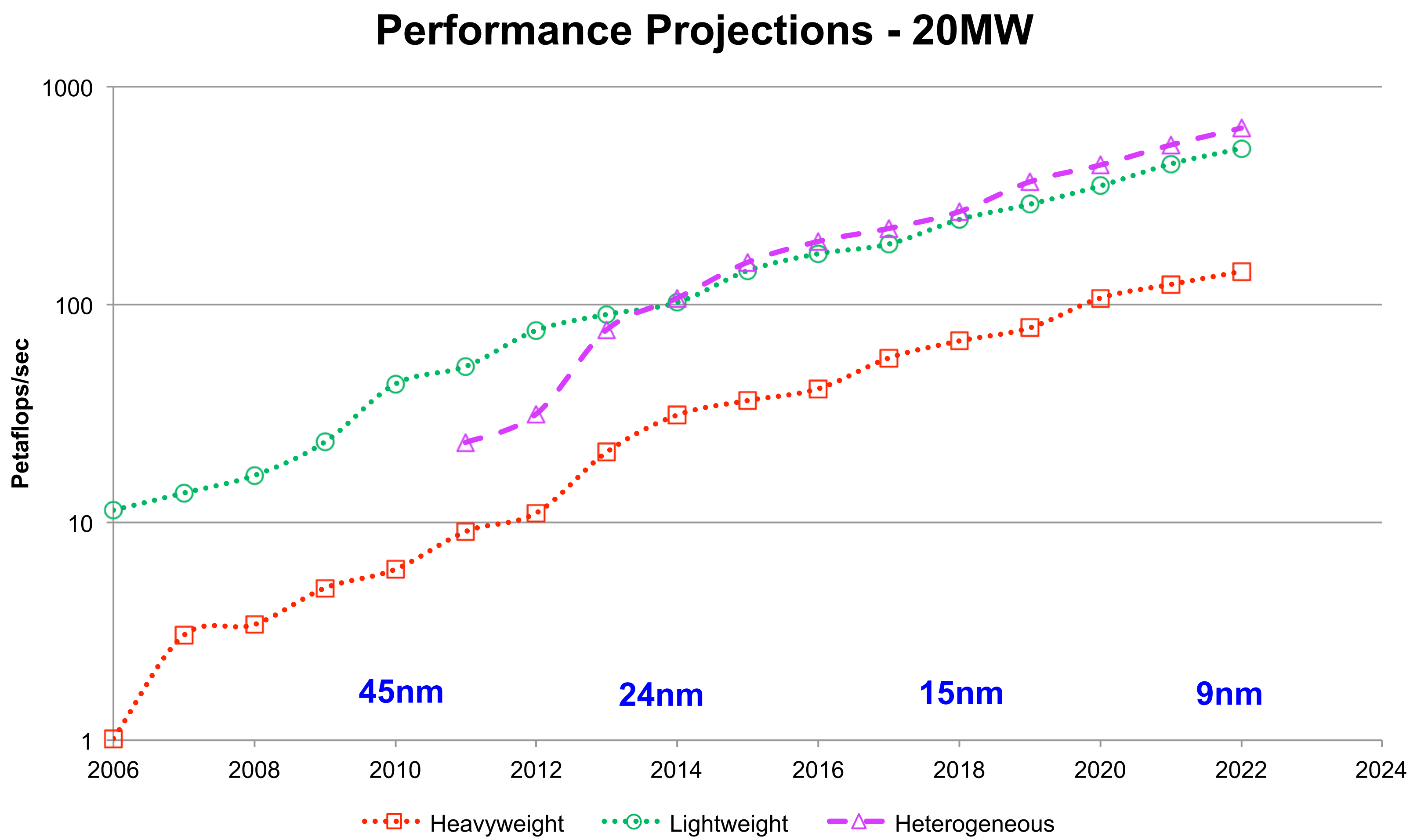
NVIDIA Announces "Project Denver" to Build Custom CPU Cores Based on ARM Architecture, Targeting Personal Computers to Supercomputers

NVIDIA Licenses ARM Architecture to Build Next-Generation Processors That Add a CPU to the GPU

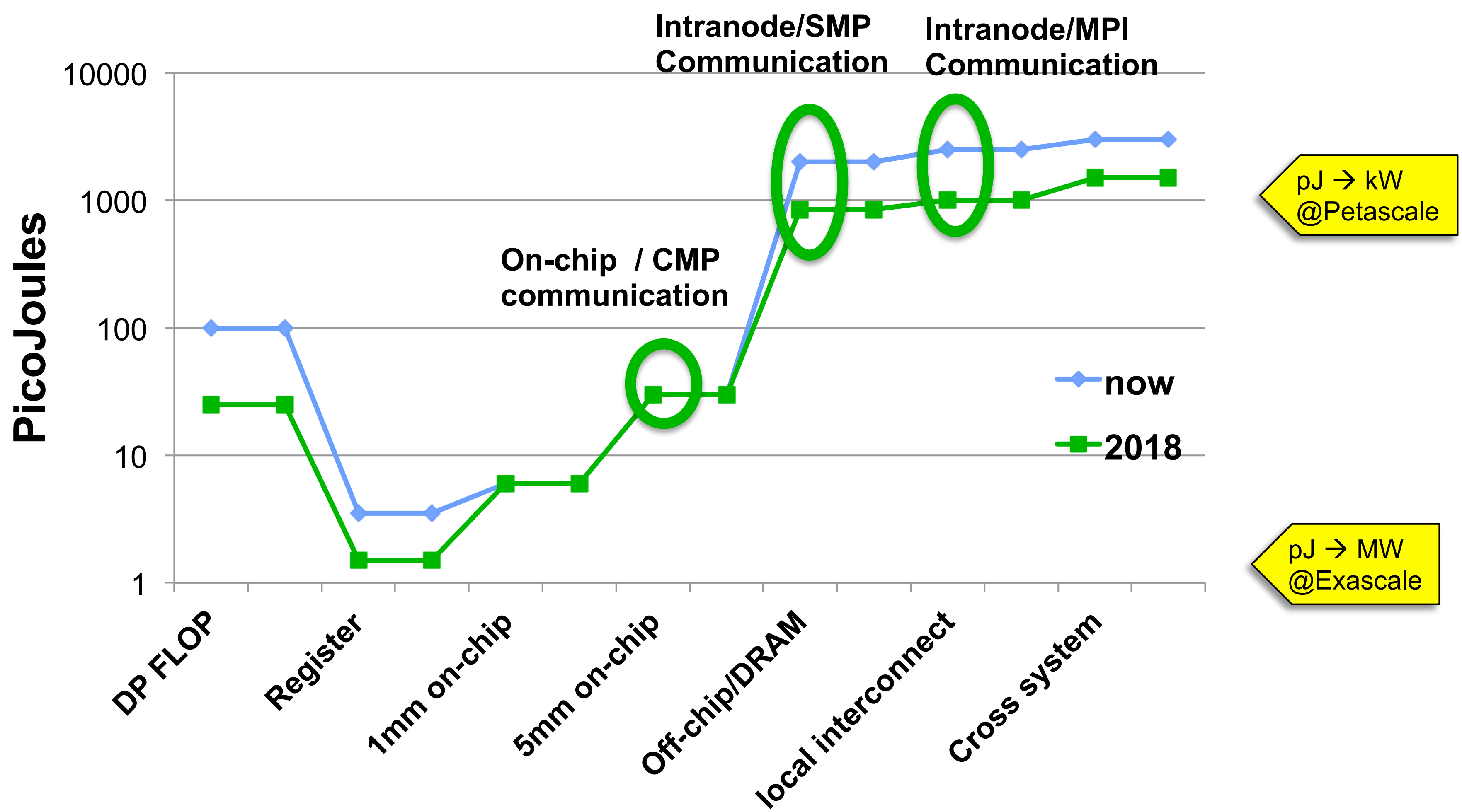
Later this decade
a 10 TF Node might be:

- CPU cores -- 10
- GPU
  - Cores – 1000
  - Threads – 100/core
- Fast integrated memory
  - Capacity – 100GB
  - Bandwidth – 1-2 TB/s
- DRAM
  - Capacity – 300 GB
  - Bandwidth – 100 GB/s
- Interconnect
  - ~100 GB/s
- Applications will need to manage locality and parallelism to achieve any reasonable level of performance
- Not clear if mobile devices will require dependability (correctness and reliability)

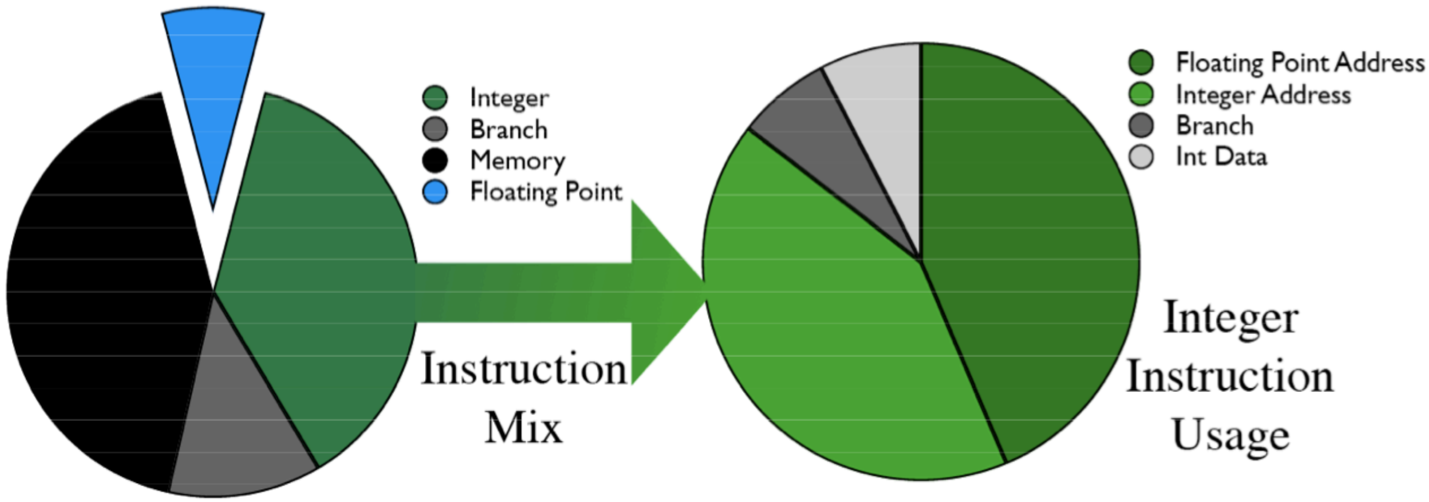
Meeting the 20 MW power goal will be a challenge



We need to reduce the pJs required to move a bit and applications will need to manage locality



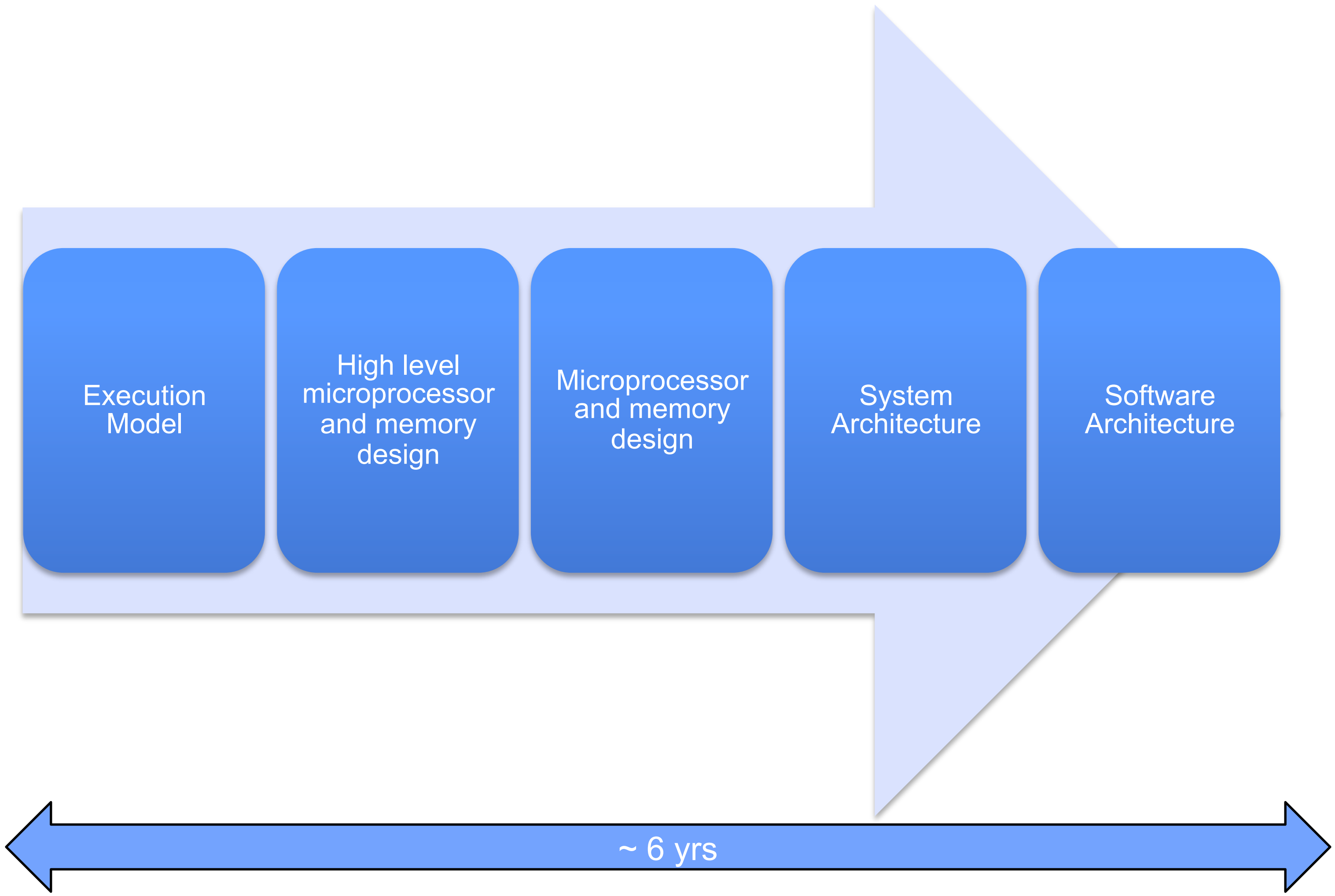
Memory dominated now FLOPS can be overprovisioned



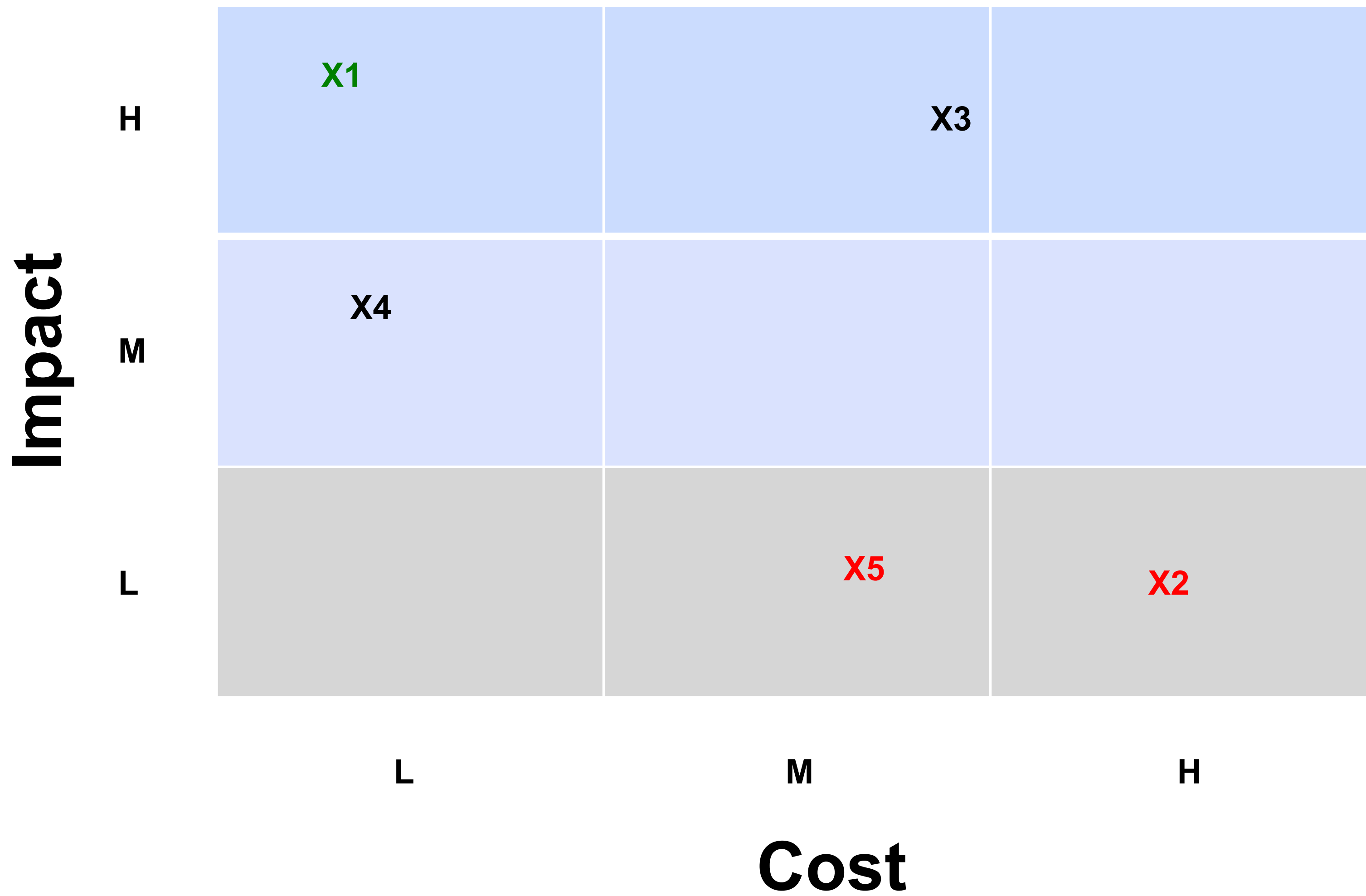
- Most of DOE's Applications (e.g., climate, fusion, shock physics, …) spend most of their instructions accessing memory or doing integer computations, not floating point
- Additionally, most integer computations are computing memory Addresses
- Advanced development efforts are focused on accelerating memory subsystem performance for both scientific and informatics applications

Science Partnership for Extreme-scale Computing

It is urgent to begin soon for co-design to have an impact



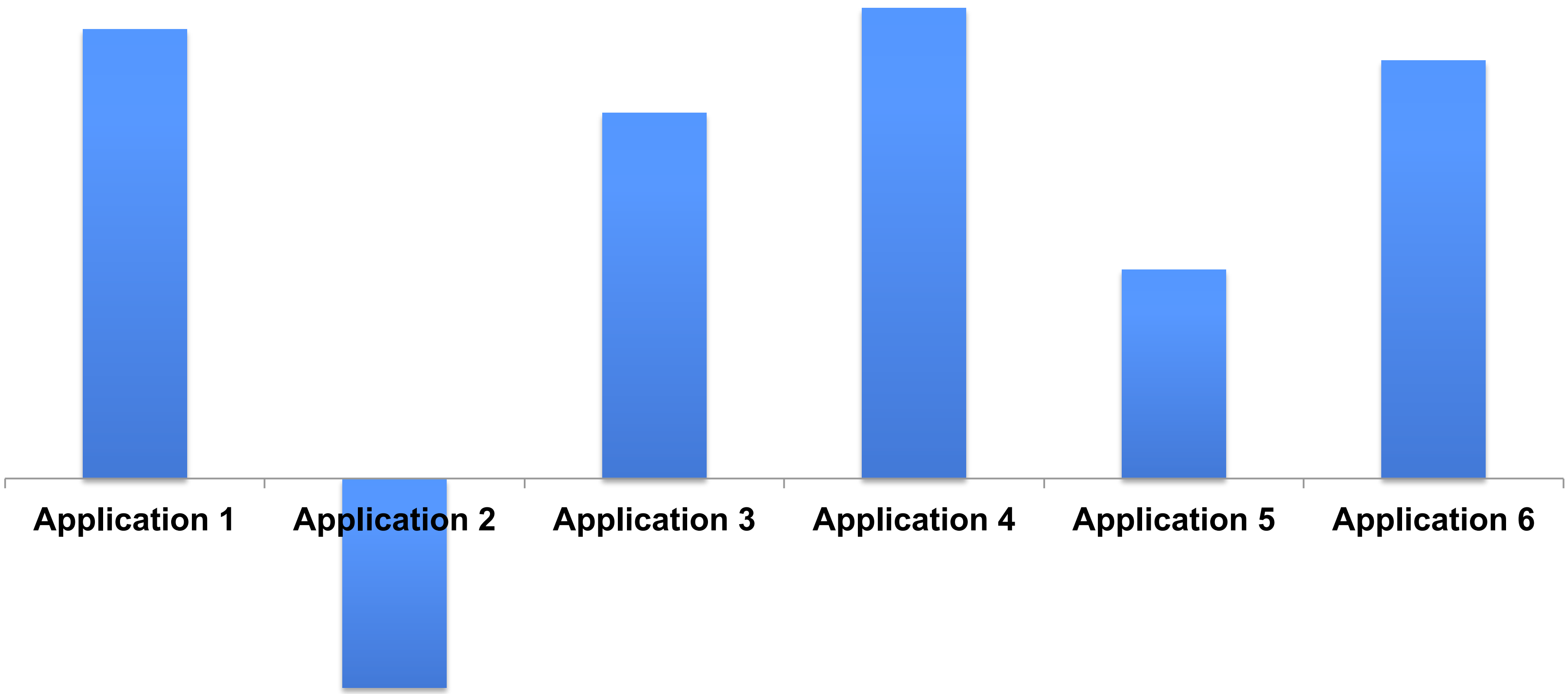
We will need data to make decisions at key points in the design process



Determine the benefit of X_n architectural choices that have a given cost (Si area, energy, R&D)

Science Partnership for Extreme-scale Computing

SPEC will work with industry partners, codesign teams and HQ to make decisions



Is there a weighting? A minimum?

We have the potential to influence many elements of an Exascale system

- Elements we might influence
  - Cores/node, threads/core, scheduling width/thread
  - Memory capacity and bandwidth
  - Logic in memory subsystem (improve effective bandwidth)
  - Interconnect performance
  - Dependability
- However, we must understand and leverage industry roadmaps