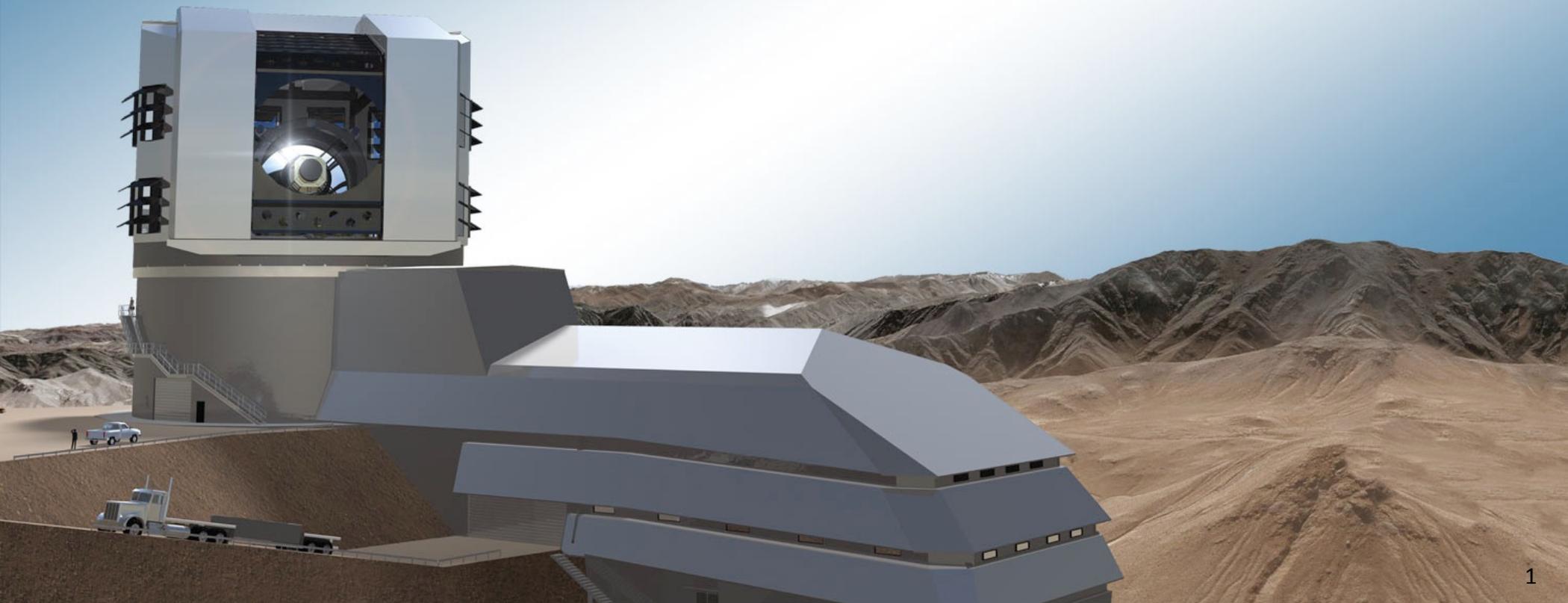


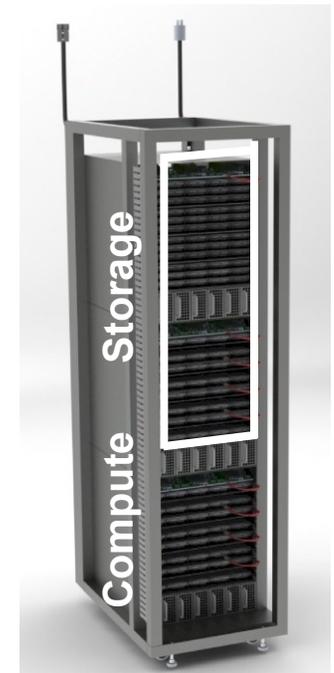
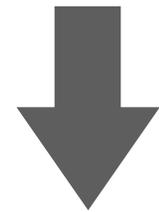
Thoughts on Next-Generation Computing and LSST DESC

Salman Habib

June 2014, LSST DESC Philadelphia Meeting



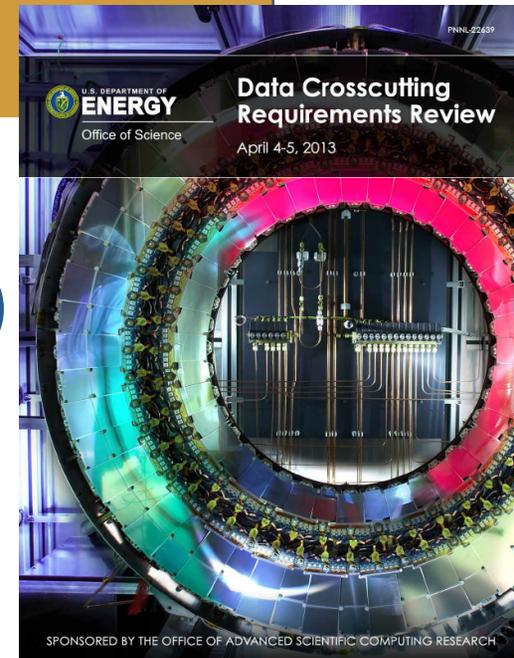
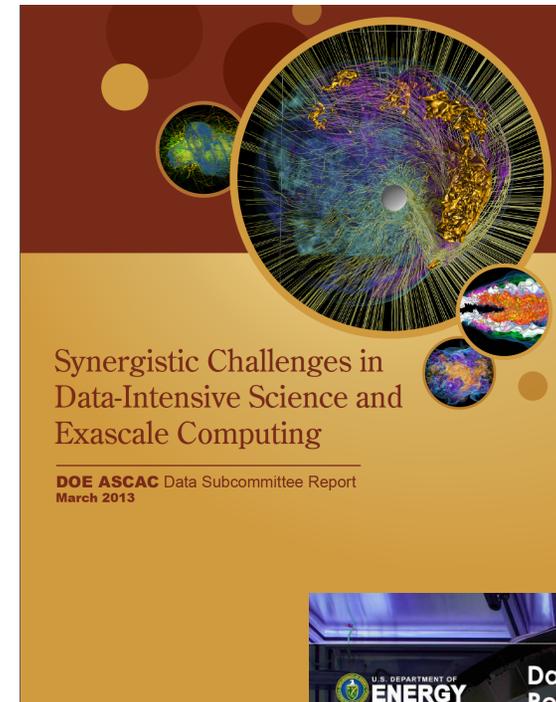
- **LSST DESC Focus is on Level II/III Analysis**
 - Estimates of initial computing needs unclear, range from 150-350 TFlops; initial storage estimates ~1 PB
- **Based on this we would want (at least) the #1 machine in the Top 500 in 2006**
- **In 2022 there may be $O(1000-10000)$ such machines in the US alone!**
- **~PB storage is a trivial requirement (already is)**
- **The estimates above are likely quite wrong, the actual figures will be much larger —**
- **The Simulations WG (+ associated analytics) alone will completely swamp these estimates**
- **Where will the computing come from? And how usable will it be?**



300 TFlops/10PB,
10kW in 2020

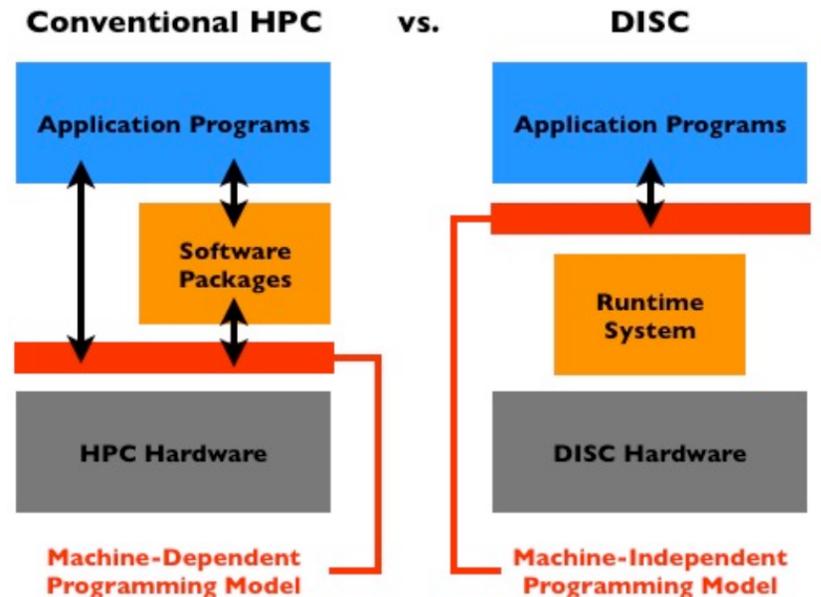
Different Flavors of Computing

- HPC**
 - **High Performance Computing ('PDEs')**
 - ▶ Parallel systems with a fast network
 - ▶ Designed to run tightly coupled jobs
 - ▶ High performance parallel file system
 - ▶ Batch processing
- DISC**
 - **Data-Intensive Computing ('Analytics')**
 - ▶ Parallel systems with balanced I/O
 - ▶ Designed for data analytics
 - ▶ System level storage model
 - ▶ Interactive processing
- HTC**
 - **High Throughput Computing ('Events'/'Workflows')**
 - ▶ Distributed systems with 'slow' networks
 - ▶ Designed to run loosely coupled jobs
 - ▶ System level/Distributed data model
 - ▶ Batch processing



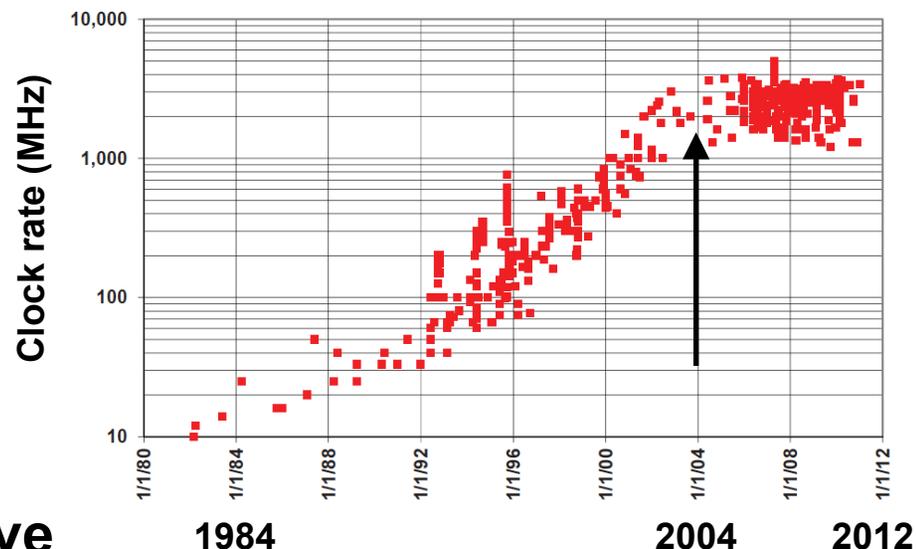
What Does it Mean? Why Should Anyone Care?

- **‘Big’ systems will be different soon**
 - CPU/GPU, ManyCore, etc.
 - Should one continue to ignore this?
- **Need for a consistent system model**
 - Path for workstation to supercomputer scale-up (‘one-step’)
 - Scale-up lockout: Ignorance is not bliss!
- **HPC and DISC (Current)**
 - Architectures very different
 - Usage models very different
- **Can HPC and DISC converge?**
 - Convergence at the hardware level
 - Configurational flexibility
 - If (some form of) convergence is possible, we should all care!

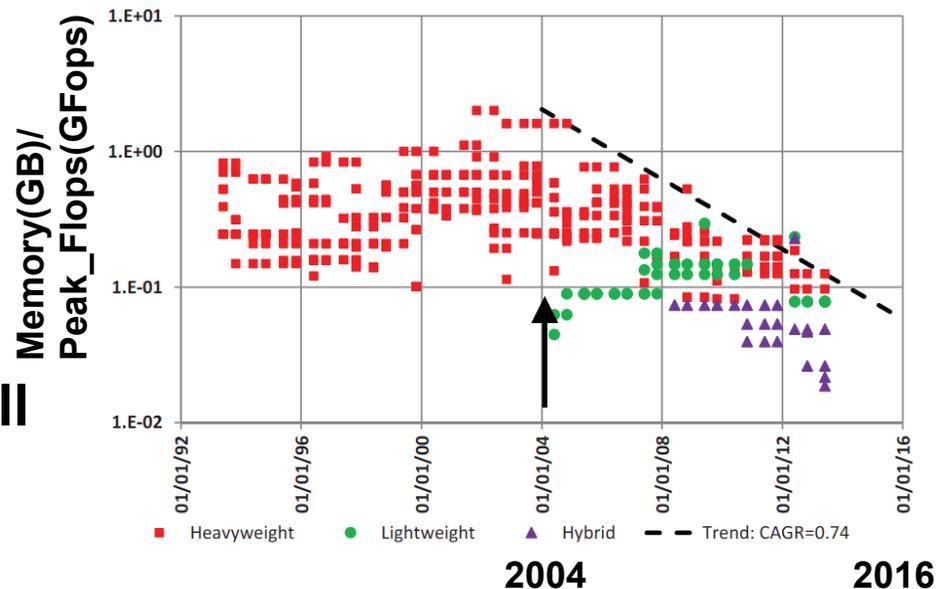


After Bryant (2007)

- **Power is the main constraint**
 - ▶ 30X performance gain by 2020
 - ▶ ~10-20MW per large system
 - ▶ power/socket roughly const.
- **Only way out: more cores**
 - ▶ Several design choices
 - ▶ None good from scientist's perspective
- **Micro-architecture gains sacrificed**
 - ▶ Accelerate specific tasks
 - ▶ Restrict memory access structure (SIMD/SIMT)
- **Machine balance sacrifice**
 - ▶ Memory/Flops; comm BW/Flops — all go in the wrong direction
 - ▶ (Low-level) code must be refactored

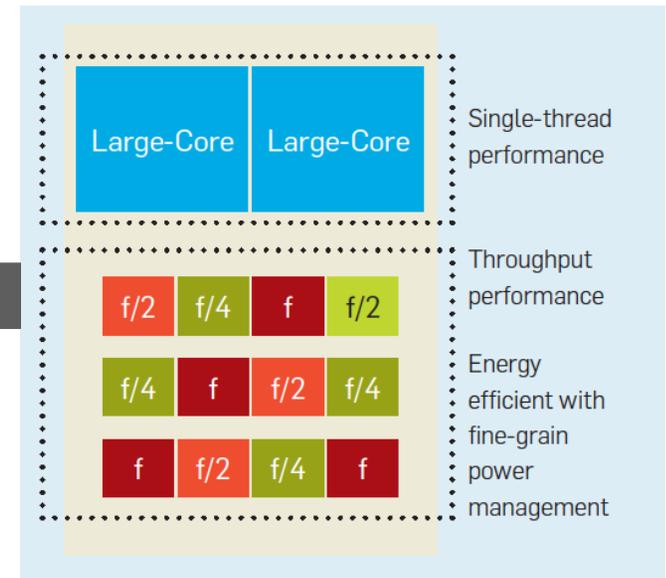
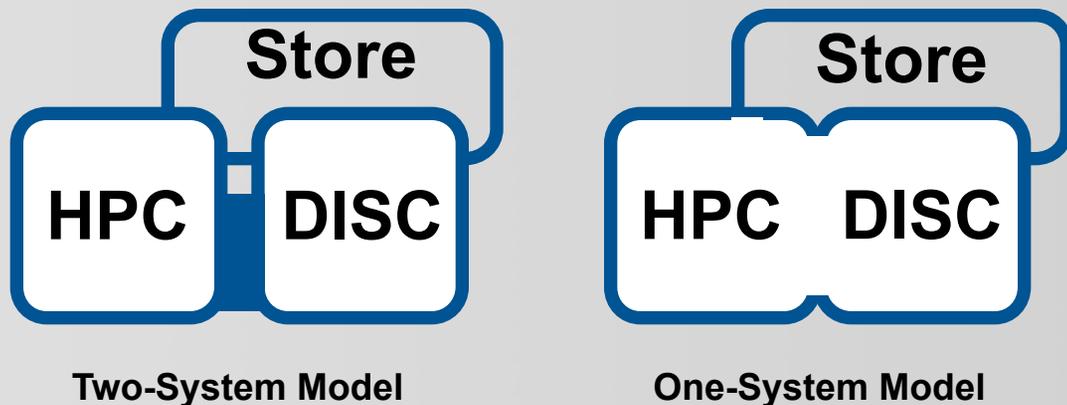


Kogge and Resnick (2013)



Possible Future Directions

- **The future of HPC is not 'HPC'!**
 - ▶ HPC systems were meant to be balanced under certain metrics
 - ▶ These range from ~ 0.1 to ~ 0.001 on the same system and will get worse
 - ▶ A question of \$\$, not technology
- **Data centers are important drivers**
 - ▶ Power is the essential constraint
 - ▶ Trends towards software defined infrastructure



Future heterogeneous manycore system, Borkar and Chien (2011)

“— challenges involved in meeting the increasing computational needs of the field and suggested steps to take **full advantage of cost-effective computing solutions**. The present practice is to handle much of the computing within individual projects. **Rapidly evolving computer architectures and increasing data volumes require effective crosscutting solutions** that are being developed in other science disciplines and in industry. Mechanisms are needed for the continued maintenance and development of major software frameworks and tools for particle physics and long-term data and software preservation, as well as **investments to exploit next-generation hardware and computing models**. Close collaboration of national laboratories and universities across the research areas will be needed to take advantage of industrial developments and to avoid duplication.”

- **DOE HEP Response**

- ▶ **Forum for Computational Excellence**
- ▶ **Initial activity is data gathering**
- ▶ **Software/Tools/Systems**
- ▶ **Need your input!**
- ▶ **Website coming soon**
- ▶ **Contacts:**
habib@anl.gov,
roser@fnal.gov

- **DOE ASCR moving into data space**

- ▶ Focus area demos (for SC14)
- ▶ EXDAC (EXtreme Data Analysis for Cosmology)

- **NERSC**

- ▶ DISC(ish) clusters
- ▶ HPC to DISC transition (Edison)
- ▶ LSST allocation can go in new directions (as with DES)

- **Argonne**

- ▶ Magellan cloud prototype

- **Community response**

- ▶ Multiple projects: DES, DESI, LSST DESC, WFIRST-AFTA, —
- ▶ Use cases, application kernels

