

FOREWORD

Jim Held PhD.

Intel Fellow
Director, Tera-scale Computing
Research
Intel Labs
Intel Corporation

“Systems with teraFLOPS of parallel performance that can process tera-bytes of data.”

“Visual-computing applications require tera-scale performance in order to simulate worlds based on complex physical models.”

The Intel® Tera-scale Computing Research Program is Intel’s overarching effort to shape the future of Intel processors and platforms, in order to accelerate the shift from frequency to parallelism for performance improvement. Intel researchers worldwide are already working on R&D projects to address the hardware and software challenges of building and programming systems with teraFLOPS of parallel performance that can process tera-bytes of data. This level of performance will enable exciting new and emerging applications, but will also require addressing challenges in everything from program architecture to circuit technologies. This issue of the Intel Technology Journal includes results from a range of research that walks down the ‘stack’ from application design to circuits.

Emerging visual-computing applications require tera-scale performance in order to simulate worlds based on complex physical models. They use rich user interfaces with video recognition and 3D graphics synthesis, and they are highly parallel. How can we build them? Architecting designs for such applications that fully exploit their inherent parallelism is a major software engineering challenge. As with most kinds of architecture, new programs will be based on a combination of preexisting patterns and an exploitation of application frameworks that support them. Tim Mattson and Kurt Koetzer describe their work to find the parallel patterns that are needed for concurrent software in their article entitled “A Design Pattern Language for Engineering (Parallel) Software.”

The non-deterministic nature of concurrent execution has made debugging one of the toughest parts of delivering a parallel program. Gilles Pokam and his colleagues, in their article “Hardware/Software Approaches for Deterministic Multi-processor Replay of Concurrent Programs” describe their work on hardware and software to support debugging by recording and replaying execution in order to allow analysis and discovery of the subtle timing errors that come with the many possible executions of parallelism.

Future tera-scale platforms may be heterogeneous with a mixture of types of compute elements. Our August 2007 issue of the Intel Technology Journal included articles that described support for mixed-ISA co-processing. In “Programming Model for Heterogeneous Intel® x86 Platforms” in this issue, Bratin Saha’s and his colleagues describe work in IA-ISA to provide support for shared memory with a mixture of cache coherence models.

Mani Azimi and his colleagues' article "Flexible and Adaptive On-Chip Interconnect for Tera-scale Architectures," describes research into on-die network fabric, and they show our evolution from an analysis of the challenges and alternatives to the development of the protocols to exploit the potential of a network on chip. Effective use of a mesh network will require sophisticated support to provide the routing and configuration management for fairness, load balancing, and congestion management.

Perhaps the largest platform hardware challenge for tera-scale computing is the longstanding one of access to memory to match the tremendous compute density of many cores on a die. Moreover, an effective solution must also meet the declining cost and power consumption targets of the mainstream market segments. Dave Dunning and his colleagues, in the article "Tera-scale Memory Challenges and Solutions" outline the problems and our research agenda in this critical area.

The continuing challenge for the core of tera-scale platforms is how to continue to increase energy efficiency. As process technology advances continue to give us more transistors, we can add more cores, but unless we improve their efficiency, we won't be able to use them. Ram Krishnamurthy's team continues to make progress in improving the energy efficiency of computations with designs for ALUs that exploit near-threshold voltage circuits and extremely fine-grained power management. Their work is described in an article "Ultra-low Voltage Technologies for Energy-efficient Special-purpose Hardware Accelerators."

Finally, for some research questions there is no substitute for a silicon implementation: therefore, we built the Tera-scale Research Processor to explore a tile-based design methodology as well as to understand the performance and power efficiency that is possible with intensive floating-point engines and an on-die network. In our final article "Lessons Learned from the 80-core Tera-scale Research Processor," Saurabh Dighe and his colleagues review these results and discuss what conclusions we draw from them. They summarize what we learned from many experiments with this chip. We recently announced our second-generation many-core research prototype, the Single-chip Cloud Computer, which builds on this work.

I hope you find these articles informative, and the future they are part of creating, as exciting as we do at Intel Labs. We look forward to continuing work with academia and the industry to meet the challenges of mainstream parallel computing.

"Effective use of a mesh network will require sophisticated support to provide the routing and configuration management for fairness, load balancing, and congestion management."

"For some research questions there is no substitute for a silicon implementation."