“Opportunities and Challenges for Peta-Architectures”

Dr Pawel Gepner – HPC Platform Architecture Specialist, Intel Corporation

Abstract:

There is a great deal of enthusiasm for making large scale systems and use of PFLOPS size computing. Systems on such scale are more powerful than most of today's HPC installations. After systems are designed and purchased, and before they can be used at PFLOPS scale, they must be installed in a facility, integrated with the existing infrastructure and environment, tested and then deployed for use. Unless system testing and integration is done effectively, there are risks that large scale systems will never reach their full potential. PFLOPS systems’ capabilities depend on advances in device technology, architectural structures, and parallel algorithms. We will discuss the field of PFLOPS computing in terms of the evolution of new architectures and readiness to deliver effective performance. There are a number of difficult technical problems that need to be solved.

One of the challenges is conquering latency: coupled with the need to achieve one PFLOPS aggregate sustained performance, this class of the system will require enormous system concurrency as 100000 processor cores need to be installed. Concurrency of this scale is not trivial and managing latency and extreme concurrency drives much of the development that need to be done. Building systems of such class requires specific questions to be answered:

- Can we produce a usable PFLOPS system using commercial, off-the-shelf (COTS) hardware components?
- Is a hybrid hardware technology approach, such as GPGPUs, superior to a COTS-based design and how effectively it can be used?
- How can the power consumption for such a system be reduced to acceptable levels and implemented in existing data center?
- What hardware services are needed to manage latency and what kind of instrumentations are necessary for fail-over scenario?
- What is the best design for data storage and I/O subsystem on such a class of structure?

On the applications level we will see even more challenges to fully utilize this scale of system. We see opportunity and potential but can we find a lot of applications that are ready to exhibit hundreds of thousands of concurrent threads today? Many others questions appear when we planning PFLOPS system not all of them can be answered but at least are valuable to ask:

- What is the best way to implement a choice of applications on proposed system designs?
- What will be the memory and I/O requirements of future applications?
- What exotic new applications might be enabled by PFLOPS systems

These types of questions need to be asked before we make decision and start implementation of a PFLOPS system. Of course we can build PFLOPS system today but the challenge is to do it smart.

Paweł Gepner is Platform Architecture Specialist for Europe Middle East & Africa Operation focused on High Performance Computing.

Paweł Gepner in his role as Platform Architecture Specialist is to ensure customers of server and HPC products receive world-class technical service and support, technology training and other marketing services.

Paweł Gepner has joined Intel in 1996 as Field Application Engineer for Central and Eastern Europe. In 1999 he has been appointed to a current position.

In 2001, Paweł Gepner became EMEA Architecture Specialist focused on HPC area. He led couple of server development projects including first Fault Tolerance Systems based on IA-32 from Stratus Technology. He was responsible for driving Pentium III server project at IBM R&D Center in Greenock. He also led the team that developed Bull’s Itanium 2 system. He was also involved in Itanium 2 projects at Siemens AG and Eriksson.

Paweł Gepner led the development team for the first teraflop computing projects in Poland and first Itanium 2 teraflop installations in EMEA. He was driving many of the HPC projects in EMEA including TASK, SKODA, VW, CERN, Airbus and many others.

In addition to the Platform Architecture Specialist role he is also Intel Corporation spoke person responsible for communication with the press reg. technical and technology aspect of Intel’s products and technology.

Paweł Gepner is graduated in Computer Science and he holds master's and Ph.D. degrees from Warsaw Technical University, Poland

Paweł Gepner has written 25 technical papers on Computer Science and technology. He is also a board member and technology advisor for many international scientific and commercial HPC projects.