Dirk Pleiter 30 December 2013

# Integration of Non-Volatile Memory Technologies in HPC Architectures: Challenges and Opportunities

I/O performance of HPC systems is a major challenge on the path towards exascale computers. As of today the computing performance continues to improve at a faster path than the performance of the I/O subsystem. At the same time new application areas, in particular from life sciences, are emerging that challenge this design point even further. Many of these applications have in common that they perform irregular accesses to huge amounts of data. Some of these data-intensive applications have large memory footprints and would benefit from storage class memory to stage large amounts of transient data. Lastly, advances in storage architectures suitable for exascale systems have to take place within a tight power envelope, which mandate major improvements in power and energy efficiency.

Architectural opportunities based on the use of existing non-volatile memory technologies as well as those arising from advances in non-volatile memory (NVM) technologies can play a role in addressing these challenges. The Blue Gene Active Storage (BGAS) architecture is an example of this approach as it realises an active storage concept within a top-class HPC architecture with integrated NAND flash memory cards.

Taking an existing realisation of an active storage concept, the BGAS architecture, which was presented by Blake Fitch (IBM), as a starting point the BoF addressed different questions related to the use of such architectures from scientific applications developer's point of view. What opportunities for new scientific research are enabled by active storage architectures? How can such architectures be used? What are suitable programming models?

In his opening presentation Rick Stevens (ANL) highlighted the opportunities for systems biology. Dealing with a quickly growing volume of genomics data which are processed in increasingly complex workflows, NVM technologies and active storage architectures allow to keep larger amounts of data inside the system and to access or process this data more efficiently. Felix Schürmann analysed the opportunities for large-scale brain simulations where integrated NVM allow to hold neuronal spiking network data for later reuse or to buffer transient data for analysis and visualisation. Chris Carothers (RPI) extended this to workflows in aerodynamics simulations and high-performance data analytics in medicine (breast cancer detection and diagnosis).

Wolfgang Frings (JSC) discussed in his presentation on programming models different approaches on how to use a heterogeneous pair of parallel clusters with different resource characteristics, namely a Blue Gene/Q system and a BGAS system, which are optimized for data intensive and compute intensive tasks, respectively. Active message concepts would, e.g., allow compute nodes to send data to a (loosely coupled) BGAS system for parallel, data-intensive analysis. Vice versa, BGAS nodes could send data to the compute system to accelerate massively-parallel, compute-intensive tasks.

A copy of the presentations can be found here.