WHITE PAPER

June 2013

Fraunhofer ITWM demonstrates GPI 2.0 with Mellanox Connect-IB™ and Intel® Xeon Phi™

In collaboration with:

Fraunhofer ITWM

intel

Over the last decade, specialized heterogeneous hardware designs ranging from Cell over GPGPU to Intel Xeon Phi have become a viable option in High Performance Computing – mostly due to the fact that these heterogeneous architectures allow for a better flops-per-watt ratio than conventional multi-core designs.

However, the corresponding programming models for heterogeneous architectures so far remain limited. Neither offload models (like Cuda from Nvidia, OpenACC or Intel offload directives) nor the native execution on the accelerator (e.g execution on Intel Xeon Phi) is able to provide a single cohesive view of the underlying fragmented heterogeneous memory.

The upcoming new GASPI standard will able to bridge this gap in the sense that GASPI can provide partitioned global address spaces (so called segments), which span across both the memory of the Host and e.g. an Intel Xeon Phi.

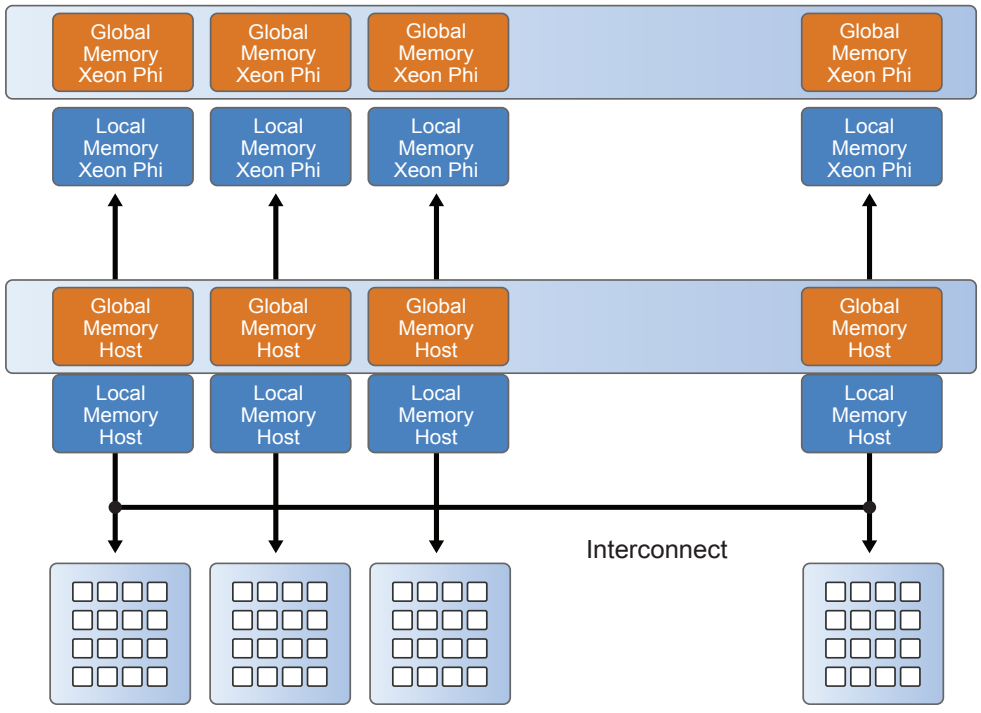

Figure 1. *GASPI Interconnect*

GASPI allows for direct read from/write to of both these segments by means of one-sided, asynchronous, RDMA based communication. Since one-sided communication in GASPI is based on remote completion (so called notifications), GASPI enables its users to implement highly scalable, asynchronous, data-flow based communication patterns for heterogeneous, distributed memory architectures. In GASPI, data may be asynchronously written from the local host segment to the remote Xeon Phi segment, for example. Whenever this data is produced and the data is locally available, then the corresponding notification has been flagged for the respective segment.

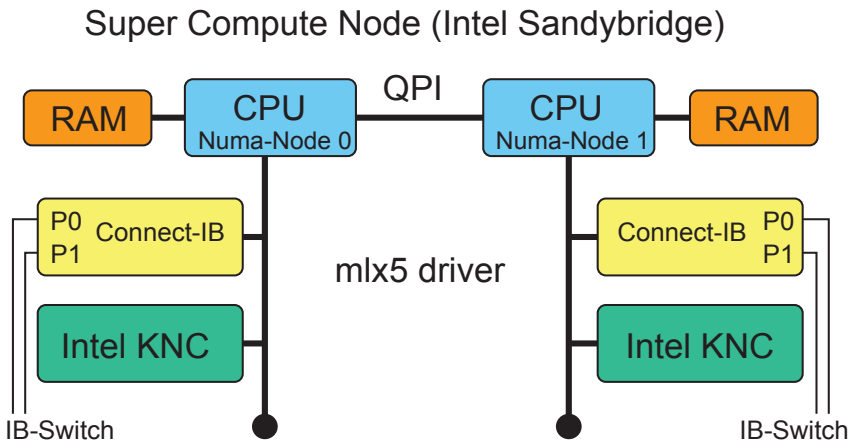

Figure 2. *Diagram shows the functional block diagram of the test system's architecture*

The award winning communication library from Fraunhofer is currently installed on many large scale InfiniBand and Cray-based deployments. Fraunhofer will provide the GPI 2.0 release on June 10th, 2013 which features one of the first implementations based on the latest generation of Mellanox Connect-IB adapters, and will support the GASPI standard. The Connect-IB adapters and drivers dramatically improve data transfer rates and delivers lower latency for all local and remote communication between the respective GASPI segments: From Host to XeonPhi, from Xeon Phi to Xeon Phi, and from Host to Host.

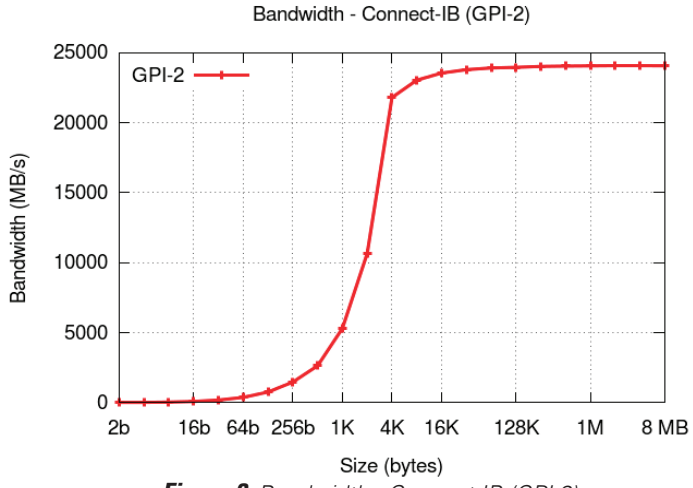

Figure 3. *Bandwidth - Connect-IB (GPI-2)*

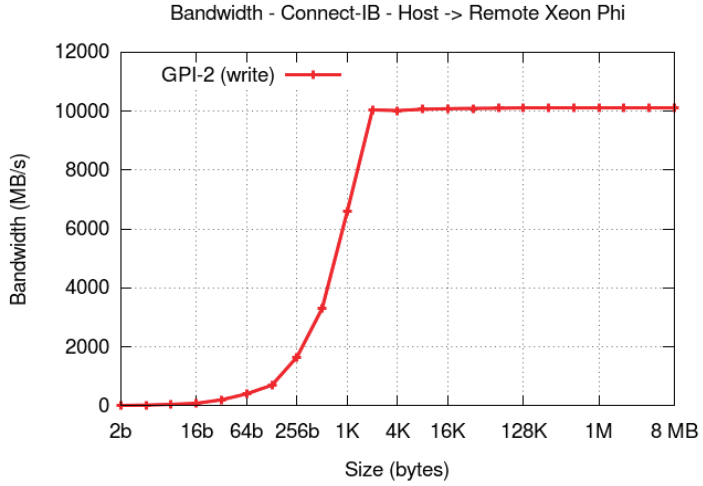

Figure 4. *Bandwidth - Connect-IB - Host -> Remote Xeon Phi*

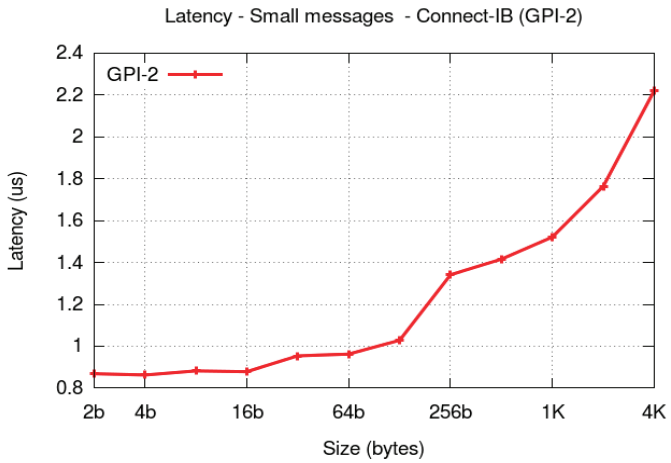

Figure 5. *Latency - Small messages - Connect-IB (GPI-2)*

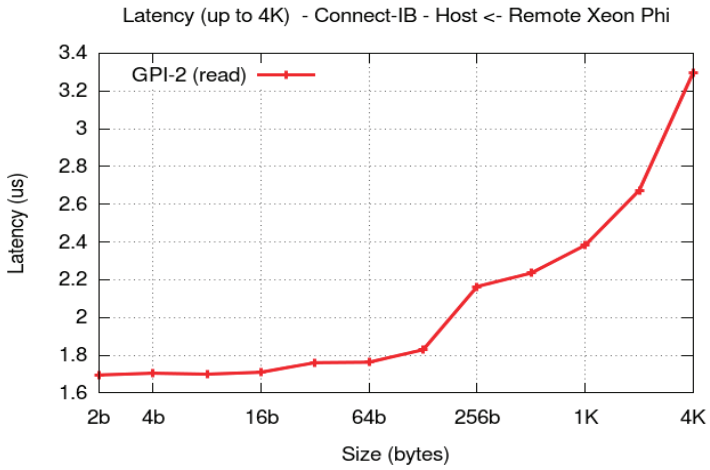

Figure 6. *Latency (small msg sizes) on heterogeneous gaspi segments*

GPI 2.0 demonstrates an implementation of highly scalable dataflow communication patterns for heterogeneous architectures. The results above show the high quality and performance of the new GPI 2.0 software release and its outstanding results for heterogeneous architectures.

Mellanox TECHNOLOGIES
350 Oakmead Parkway, Suite 100, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.mellanox.com