Comparison of Embedded Solutions for Signal and Image Processing: GPU vs. Intel Core i7 vs. MPC8641D

MERCURY WHITE PAPER

By Anne Mascarin, Product Marketing Manager
Scott Thieret, Technology Director, GPU Computing
Mercury Computer Systems

MERCURY COMPUTER SYSTEMS™

INTRODUCTION

Many advanced applications for high-performance embedded computing demand massive amounts of computing power. Real-time, on-platform systems in applications such as persistent surveillance and electronic warfare have huge GFLOPS and/or GB/s requirements, and strict power and size limits. Traditional CPU-based boards simply do not meet computational or size, weight, and power (SWaP) constraints.

GPUs are currently used in high-performance embedded computing applications, where performance metrics are paramount. This paper:

- Compares the performance of the ATI RV770 GPU with two microprocessors, Intel Core i7 and Freescale™ MPC8641D.
- Describes the GPU performance of the Mercury Sensor Stream Computing Platform (a deployable VXS system, which features GPUs) in terms of GFLOPS/watt for streaming FFT and convolution performance.

PERFORMANCE COMPARISONS

The performance of a quad-core (16-FPU) Intel Core i7 processor was measured using Intel's IPP library executing out of L1 cache (versus streaming DRAM to DRAM). The performance of an MPC8641D was measured using the Mercury SAL library and executing out of L1 cache. The performance of an ATI RV770 GPU (800 FPU) was measured using Mercury-optimized FFT routines streaming through the chip (DRAM to DRAM).

Peak theoretical performance varies greatly for the three devices: 94 GFLOPS for the Core i7, 20 GFLOPS for the MPC8641D, and 1200 GFLOPS for the RV770 GPU. (See Figure 1.)

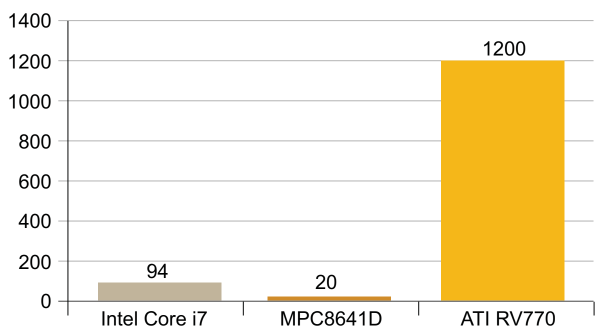

Figure 1. Peak Theoretical Performance

While peak performance is an interesting comparison, the performance of each processing device on relevant algorithms is more relevant. For embedded signal processing – particularly radar and signals intelligence systems – the 1K complex single-precision floating-point FFT is one of the most common benchmarks. This algorithm was measured at 65 GFLOPS on-chip (L1 to L1) using all four cores on the Core i7, 14 GFLOPS on-chip (L1 to L1) using both cores on the MPC8641D, and 305 GFLOPS on the RV770 GPU while streaming (DRAM to DRAM). (See Figure 2.)

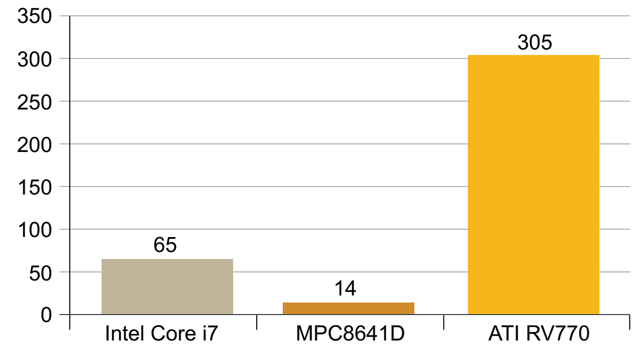

Figure 2. 1K FFT Single-Precision Floating-Point Performance

Performance per watt, a critical measurement for embedded applications, was calculated using the same results. Using 130W as the total chip power, the Core i7 had 0.72 peak theoretical GFLOPS/watt, and the MPC8641D had 0.50 GFLOPS/watt. The RV770 GPU had a much higher 9.23 GFLOPS/watt using peak board power, which included power for the GPU, memory, interfaces, and power conversion. (See Figure 3.)

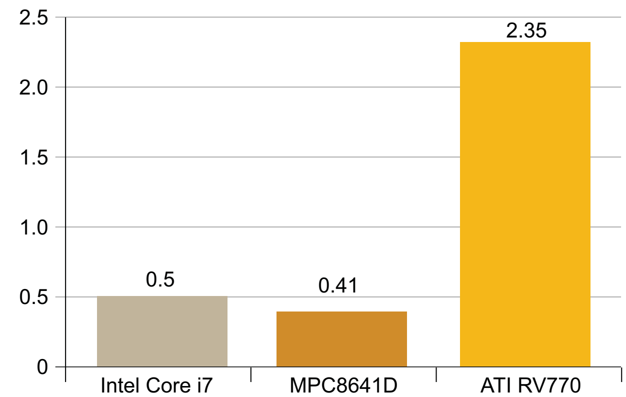

Figure 3. 1K Single-Precision Floating-Point FFT GFLOPS/Watt Performance

Similarly, computing the performance per watt for the FFT resulted in 0.41 GFLOPS/watt for the MPC 8641D, 0.5 1K single-precision FFT GFLOPS/watt for the Core i7, and 2.35 single-precision FFT GFLOPS/watt for the RV770 GPU. This is an almost 5x improvement in performance over the Core i7, and almost 6x performance improvement over the MPC8641D, even though the Core i7 is sourcing and sinking data from/to L1 cache and the GPU is streaming from/to off-chip DRAM. (See Figure 4.)

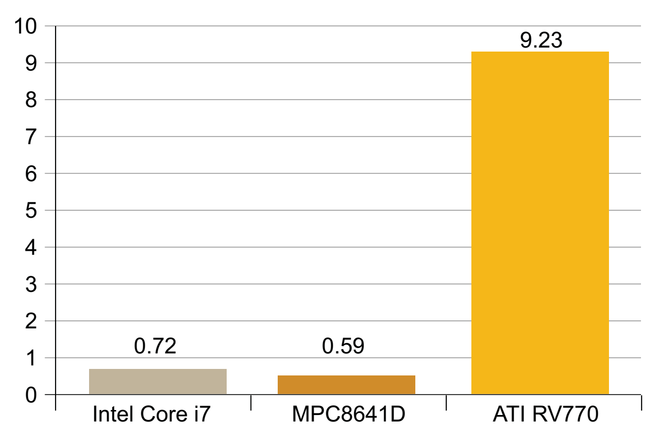

Figure 4. Peak Theoretical GFLOPS/Watt

PERFORMANCE DIFFERENTIAL BETWEEN CPUS AND GPUS

Why do such large performance differences between the Core i7, the MPC8641D, and the RV770 GPU exist? The answer lies in the basic structure and function of the processors. CPUs feature large caches and branch-prediction logic for decision-based code and more complex algorithms. They excel at flow control and disposition of data, for which they were principally architected.

GPUs, on the other hand, are massively parallel array processors with limited branching performance. They operate on large amounts of data simultaneously, because they have been architected to maximize arithmetic performance for graphical operations. They are compute- and memory bandwidth-intensive machines, containing small amounts of cache, which are optimized for large dataset throughput with computational kernels. This property makes GPUs very successful in compute-intensive signal and image processing systems.

GPUS IN DEPLOYABLE SYSTEM PERFORMANCE

The preceding data was generated using workstation cards in a benign environment. GPU performance was also measured in the Sensor Stream Computing Platform (SSCP), a 6U VXS-based system that harnesses the processing power of GPUs for high-performance, data-parallel computing in rugged environments. The SSCP (which has been shipping since October 2008) can be configured with either one or two VXS-GSC5200 boards. Each of these boards can be configured with either one or two GPU-based MXM modules (currently ATI 4870M or NVIDIA QUADRO 3600M). An air-cooled (as opposed to conduction-cooled) VX6-200 Dual Dual-Core Xeon VXS Single-Board Computer is included in the SSCP for I/O and control. The peak theoretical performance of the SSCP system is shown in Figure 5.

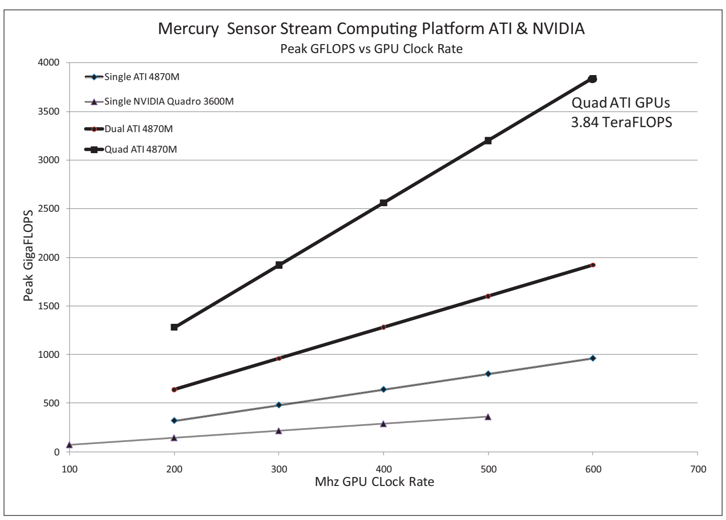

Figure 5. Peak GFLOPS vs. GPU Clock Rate in an SSCP System

A crucial feature of the SSCP is the ability to "tune" the power signature of the GPUs. This feature is particularly useful for on-platform applications, where peak algorithm performance per watt is extremely important. In the previous chart, the peak GFLOPS of the various SSCP configurations (single, dual, and quad ATI 4870M, and single NVIDIA QUADRO 3600M) increase linearly with an increase in GPU clock rate. This is just one of the "knobs" the SSCP allows users to turn to optimize the performance per watt on a deployed system-by-system basis.

SSCP performance was also measured in terms of GFLOPS per chassis watt. Because the GFLOPS performance of the SSCP increases linearly with an increase in clock speed (Figure 5), dividing the GFLOPS performance for a particular algorithm (1K complex single-precision FFT and fast convolution in this example) by the chassis power dissipated by a particular SSCP configuration yields a new metric, the GFLOPS per chassis watt. The power is full "draw-from-the-wall" watts; this includes fans, power supply inefficiencies, and the x86 host processors, as well as the power for the GPU(s). Figure 6 shows the SSCP performance in terms of GFLOPS/chassis watt for various SSCP configurations.

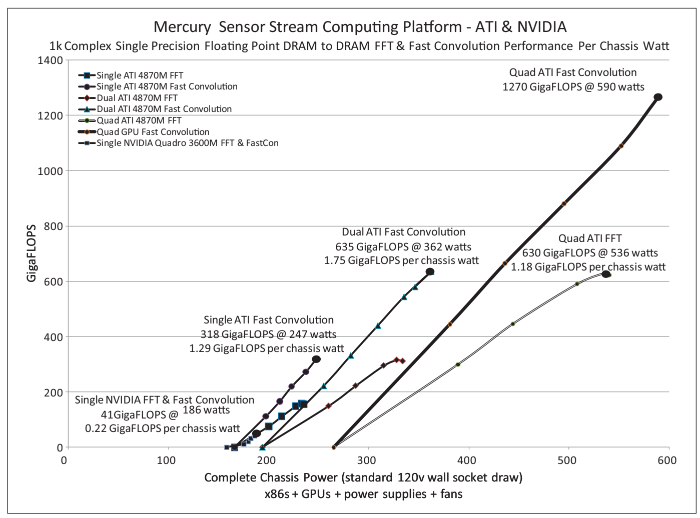

Figure 6. 1k Complex Single-Precision Floating-Point DRAM to DRAM FFT and Fast Convolution Performance per Chassis Watt

SIGNIFICANTLY HIGHER GPU PERFORMANCE

The ATI RV770 GPU significantly outperforms both the Intel Core i7 and the MCP8641D processors in terms of GFLOPS and GFLOPS/watt for peak and FFT/fast convolution metrics. Even more dramatic (two orders of magnitude) performance improvements of the RV770 are seen when running Mercury image processing libraries over quad-core Xeon® processors running IPP in areas such as image formation and analysis for persistent surveillance. Mercury's hardware design permits rapid technical insertion for our GPU-based products. Because of its outstanding performance-per-watt metrics, a system similar to the Sensor Stream Computing Platform will be flown on-platform in a major UAV program in 2010.

2362.00E-0410-WP-openvpx-advant

MERCURY COMPUTER SYSTEMS™

Corporate Headquarters
201 Riverneck Road
Chelmsford, MA 01824-2820 USA
+1 (978) 967-1401 • +1 (866) 627-6951
Fax +1 (978) 256-3599
www.mc.com

Europe
Mercury Computer Systems, Ltd.
Unit 1 - Easter Park, Benyon Road • Silchester, Reading • RG7 2PQ UNITED KINGDOM
+ 44 0 1189 702050 • Fax + 44 0 1189 702321

Asia
Nihon Mercury Computer Systems K.K.
No. 2 Gotanda Fujikoshi Bldg. 4F • 5-23-1 Higashi Gotanda • Shinagawa-ku, Tokyo 141-0022 JAPAN
+81 3 3473 0140 • Fax +81 3 3473 0141

Challenges Drive Innovation®