GE
Intelligent Platforms

GPGPU COTS Platforms

High-Performance Computing Solutions

imagination at work

Rugged GPGPU COTS Boards for Military and Aerospace

GPGPU platforms deliver new levels of performance for size, weight and power (SWaP) constrained mission payloads.

The world of high-performance computing is undergoing a revolution, thanks to advances in General Purpose computing on Graphics Processing Units (GPGPU). The idea behind GPGPU is to use a GPU, which typically handles computation for computer graphics only, to perform parallel computation in applications that have traditionally been handled by the CPU.

A multi-GPU platform hosted by one or more CPUs is able to perform heterogeneous computing, harnessing the parallel computing power of the many-core GPUs to provide very large increases in performance with minimal programming complexity.

Additionally, programmers are helped with software development environments such as Compute Unified Device Architecture (CUDA) and OpenCL, which allow them to harness the many-core, parallel processing capabilities of the GPGPU platforms.

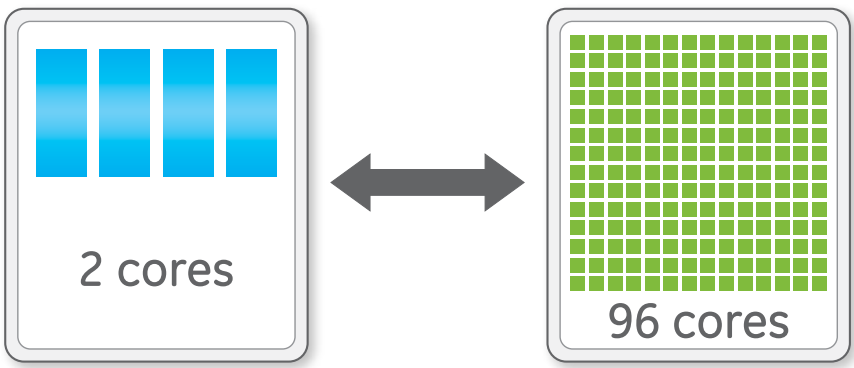
While greatly increasing functional capability, the GPGPU platform also delivers the performance with far less size, weight and power (SWaP). This results in significant savings in cost, risk, and time-to-market.

Lab-proven technologies ruggedized for the rough and tumble world of military applications

Now these benefits are fully available for rugged military and aerospace applications. With the introduction of a full range of GE Intelligent Platforms rugged GPGPU boards and systems, the advantages of GPGPU are no longer confined to controlled environments at universities, research centers and hospitals.

The unique partnership between GE Intelligent Platforms and NVIDIA allows for new product development using NVIDIA GPUs based on the award-winning CUDA architecture, for Military and Aerospace applications.

PARALLEL COMPUTING

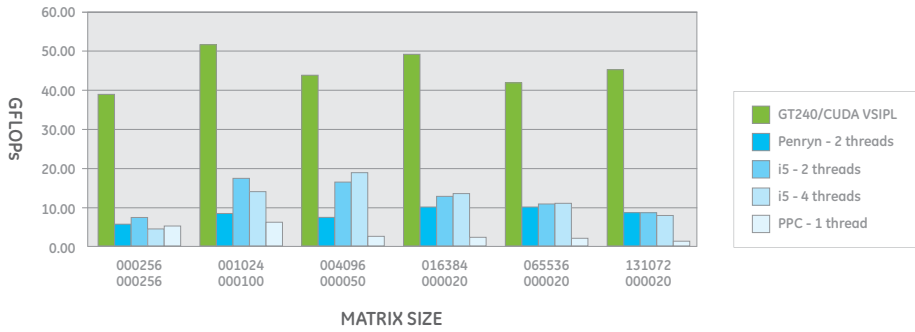


Optimized for throughput computing

Enhanced Performance for Far Less Size, Weight and Power

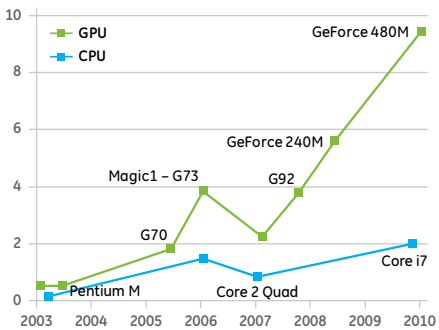
The graph shows GFLOPs performance for the VSIPL multiple FFT operation. It compares performance on different GE Intelligent Platforms embedded platforms; a GPU platform, multi-core Intel Penryn and i5 platforms and an e600 based PowerPC platform. GE's platform-optimized AXISLib DSP library product is used on all platforms. Data is in GPU memory for the GPU case and in-cache for the Intel and PPC cases.

MULTIPLE FFT PERFORMANCE - GPU vs INTEL

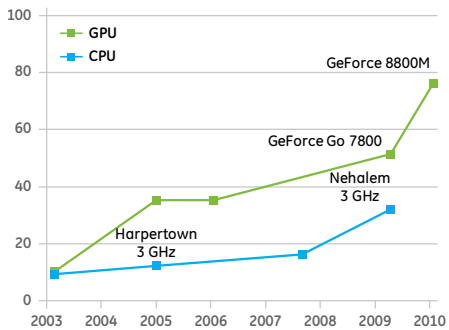


INCREASING PERFORMANCE ADVANTAGE OF GPU

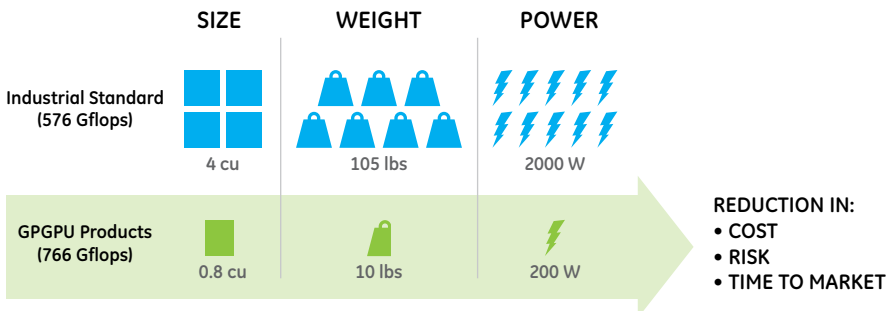
PEAK PERFORMANCE (Gflop/Watt)



PEAK MEMORY BANDWITH (GB/sec)



GE's new GPGPU platforms are particularly well-suited to many of the processing needs of Military and Aerospace applications where size, weight and power are key considerations along with resistance to extended temperature, shock and vibration. GPGPU technology allows system designers to pack more punch into less space and use less power for your applications.



Rugged GPGPU COTS Boards for Military and Aerospace

Radar

One of the biggest challenges for today's radar systems is to provide more capability—range, number of targets, speed, etc. —while meeting ever more stringent SWaP constraints. The extra speed offered by the GPGPU platforms translates directly to more area coverage and more security for the operating team.

One rack containing 72 conventional processors (18 6U boards) and producing a peak capability of 576 GFLOPS can take up 4 cubic feet, weigh over 105 pounds and consume over 2000 watts. GPGPU technology can allow system designers to fit an unprecedented amount of processing power into a very compact package. The use of three 3U VPX boards can yield peak processing power of 766 GFLOPS in less than 0.4 cubic feet.

Development Ease

Increases in performance will be obtained in application areas such as Software Defined Radio, sonar, and medical imaging. But what is less obvious is the change in development strategy offered by GPGPU technology. The only other technology currently offering massively parallel processing capability is Field Programmable Gate Arrays (FPGAs). Although FPGAs provide very high-performance data processing, developing high-performance FPGA cores requires a very specialized skill set built on a hardware engineering background, whereas developing code for GPGPU processors is much more of a software issue. For companies with a background in multi-processor GPP/DSP-based system architecture, the move to GPGPU will be much less disruptive than a move to FPGA processors. The processing power, system size and power consumption are compelling factors, but the addition of programming ease makes such a system tough to match.

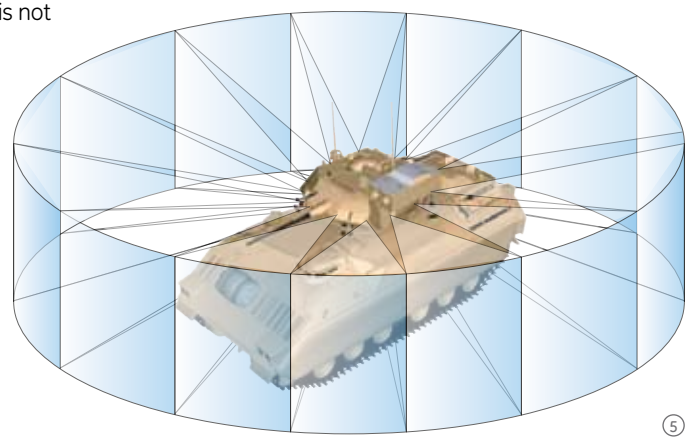
Data Encryption/ Decryption

There are several standards for encryption of data, including Advanced Encryption Standard (AES). AES is the first publicly accessible and open cipher approved by the U.S. National Security Agency for top-secret information, typically requiring 256 bit keys at this level. The time to encrypt a block of data increases linearly with the size of the key. The computation load required to maintain encryption of a real-time data stream can be prohibitive. With the advent of CUDA and the addition of crucial arithmetic, bitwise logical and shift operations as well as the ability to use texture caches to index tables, GPUs are now a viable option to general-purpose processors for data encryption/decryption. Performance gains up to 10x have been demonstrated.

Situational Awareness

Surveillance of large areas has historically been achieved by using an array of sensors connected to a bank of monitors, with separate or multiplexed displays for each video stream. Such arrangements present the operator with a confusing array of disparate video feeds, require a great deal of space, and consume a large amount of power. In a dynamic, real-time scenario, there is also a danger of information overload for an operator attempting to interpret such large volumes of imagery. Interrelationships between sensors is not always obvious, and important contextual visual information can be overlooked. Many such systems rely on the operator for "event" detection, but large volumes of information coupled to the effects of stress and fatigue can significantly reduce operator effectiveness.

Our image processing subsystem overcomes these issues and greatly improves the performance of surveillance assets and their operators. We offer a previously unattainable level of situational awareness to platforms such as armored vehicles, aircraft, remote unmanned platforms and security and surveillance systems.



IED Detection

Improvised Explosive Devices (IEDs) are a major cause of injuries and fatalities among ground troops. A number of techniques for automated detection of IEDs are used, and all of these require processing a high volume of data. The effectiveness of the solution depends on how fast the algorithm can reliably operate on that data. GPGPU technology is proving to be a highly effective solution for such high-throughput computations.

Ground Change Detection relies on real-time image processing, and may be applied to sensors mounted on ground vehicles or UAVs. The system needs to apply image registration and stabilization, and moving object extraction, before comparison with normalized geo-referenced data, all while dealing with lighting and legitimate scene changes. Ground Penetrating Radar (GPR) allows construction of a 3D model of the ground, identifying any suspicious objects or changes from normalized data. GPR can be applied to ground mobile or airborne systems. In addition, behavioral modeling—based on live sensor imagery possibly combined with wide-area surveillance data—can be used to identify hostile intent and potential threats, giving operational forces time to assess the risk and take appropriate defensive actions.

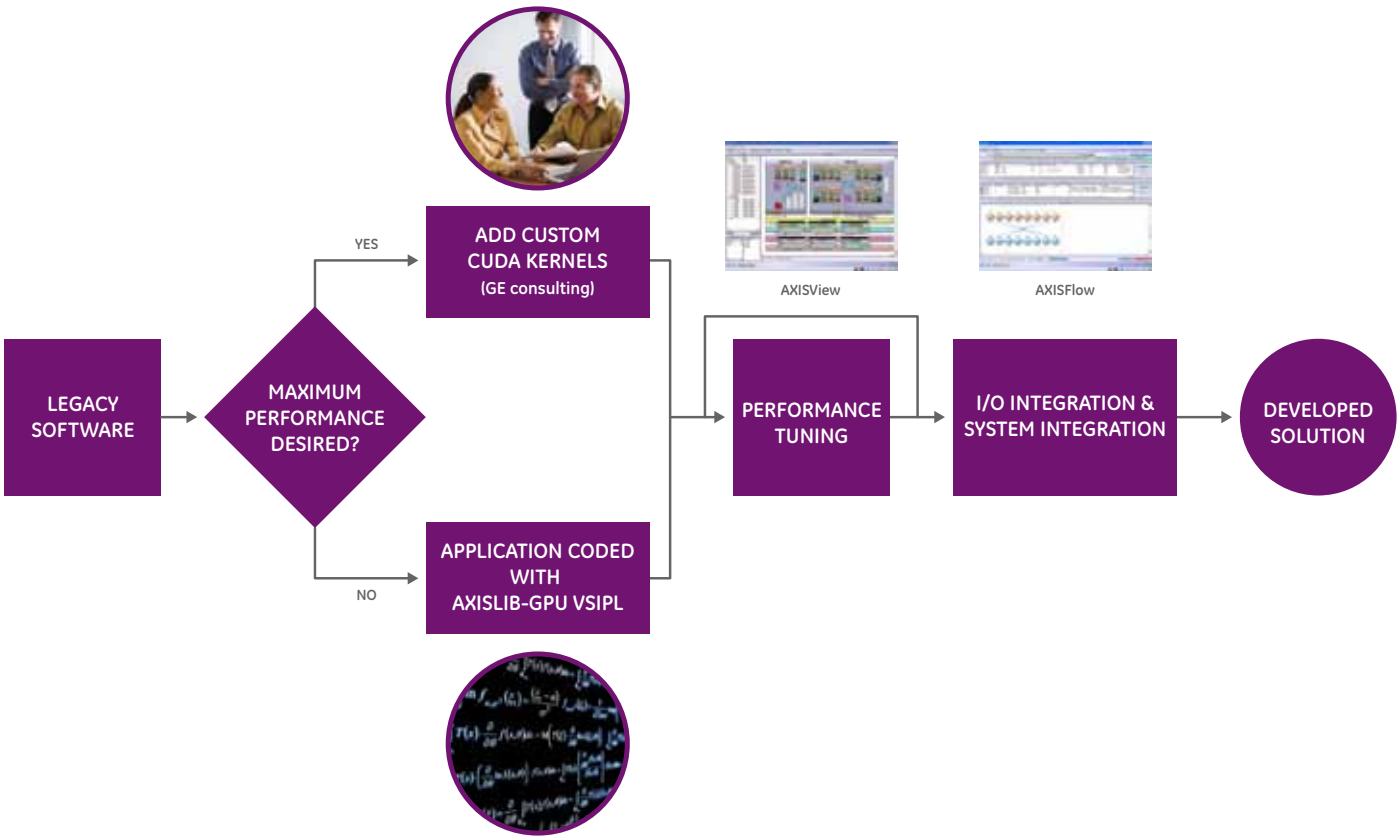
Target Tracking

GPGPU-based automatic video trackers and image processors are at the heart of target tracking systems where they provide the highest performance solutions in the smallest, fully ruggedized hardware packages. Target detection and target acquisition processes identify objects within an area of the video image display that meets the user-defined target criteria. A range of detection algorithms are built into the system to meet situational requirements. When one or more targets have been detected, the tracking system can enter automatic or manual tracking mode.

Automatic target acquisition may be prioritized by using several different factors, such as target nearest to the boresight or the largest target. If a system is in autotrack mode, the video tracker automatically tracks the selected target and can control almost any type of pan-and-tilt or gimbal system to track the target. GPGPU-based automatic video trackers feature a wide range of proven high-performance detection and tracking algorithms that can be tailored to the operational scenario, including centroid, phase correlation and edge detection. Algorithms may be combined to meet particularly demanding tracking scenarios.

GPGPU COTS Platforms

GE Intelligent Platforms GPGPU-empowered platforms can be easily implemented to either adapt to your legacy applications or to accommodate your new applications.



Application Development Made Simple By Software

AXISLib-GPU provides the industry standard Vector, Signal and Image Processing Library (VSIPL) for NVIDIA-based GPU platforms. The VSIPL API model fits perfectly with GPGPU-enabling DSP application developers to quickly realize significant application speed-ups without the need to learn GPU programming techniques.

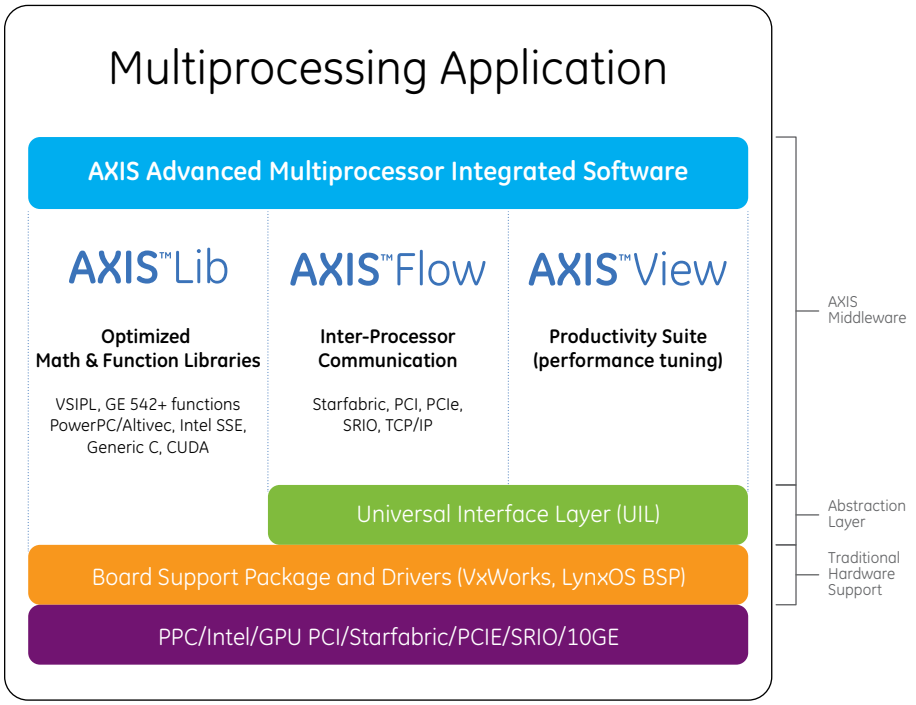
AXISLib-GPU supports the development and speeds the deployment of high-performance DSP and multiprocessing applications on GE's NVIDIA CUDA-enabled GPGPU platforms such as the IPN250, NPN240 and GRA111. Typical applications include radar, sonar, image processing, signals intelligence and intelligence, surveillance, reconnaissance (ISR). It is a set of signal and vector processing libraries providing over 500 high-performance digital signal processing and vector mathematical functions optimized for NVIDIA's many-core GPUs and created to help developers maximize system and application performance.

The AXIS Multiprocessing software suite facilitates the development of complex applications over multiple clusters of GPU platforms.

AXISFlow provides a communications API for multi-threaded/multi-core/multiprocessor communications.

AXISView provides a set of GUI tools enabling system visualization, application instrumentation, debug and monitoring.

INTEGRATED SOFTWARE MODULES



A view of 2 x IPN250s and an NPN240 system using the AXISView graphics tool

CUDA exploits the massively parallel characteristics of NVIDIA's ubiquitous silicon. CUDA is taught in universities worldwide and used in many R&D labs, so a large number of programmers are available and there is a wealth of web-based resources.

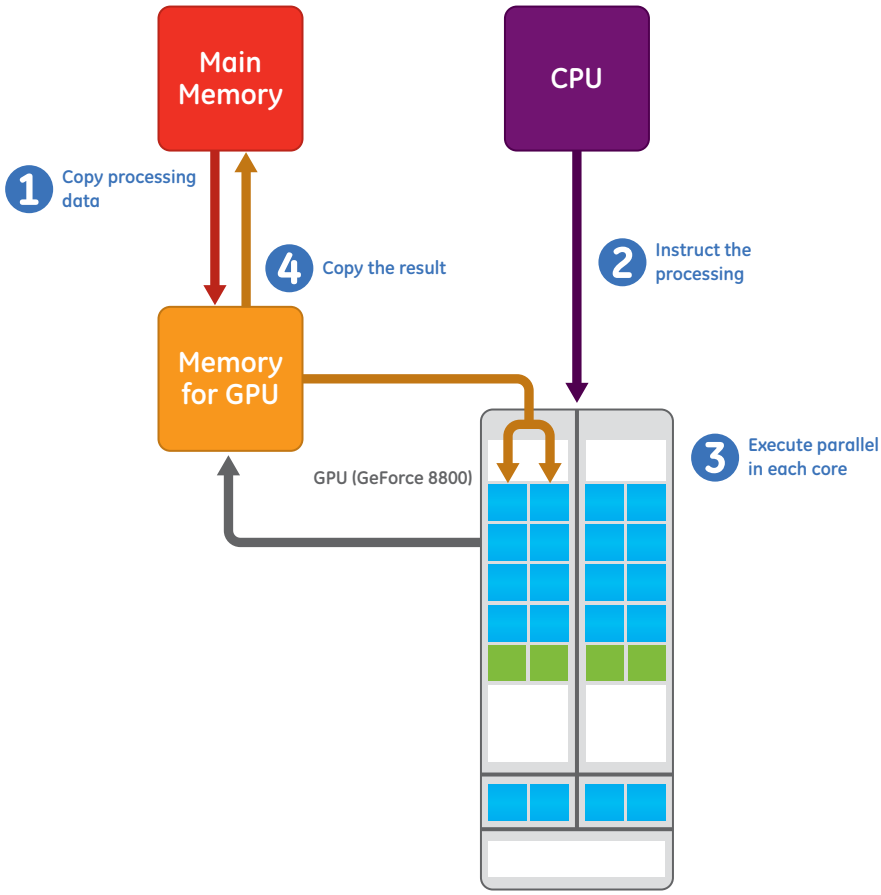
CUDA software development tools:
- NVIDIA C Compiler for parallel GPU code
- CUDA Debugger
- CUDA Visual Profiler
- SDK with best-practice guides
- Parallel Nsight®

Advanced libraries that include:
- NVIDIA Performance Primitives (image and video)
- Basic Linear Algebra Subprograms
- FFT
- VSIPL

GE GPGPU products also support Open Computer Language (OpenCL), the first open language for writing programs that execute across CPUs, GPUs, and other processors. It includes a language for writing kernels, defines APIs, and provides parallel computing using task-based and data-based parallelism.

Open Graphics Library (OpenGL) is a standard specification defining a cross-language, cross-platform API for writing applications that produce 2D and 3D computer graphics. This is used in the graphics output processes.

PROCESSING FLOW ON CUDA



C for CUDA extends C by allowing the programmer to define C functions, called kernels, that when called are executed N times in parallel by N different CUDA threads, as opposed to only once like regular C functions.

A kernel is defined using the __global__ declaration specifier and the number of CUDA threads for each call is specified using a new <<<...>>> syntax:

```
__global__ void vecAdd(float* A, float* B, float* C)
{
        int i = threadIdx.x;
        C[i] = A[i] + B[i];
}
int main()
{
        // Kernel invocation
        vecAdd<<<1, N>>>(A, B, C);
}
```

COTS GPGPU Boards and Systems

GRA111

NVIDIA GT 240 CUDA Capable Graphics Processer

The OpenVPX-compatible GRA111 is the second generation 3U VPX graphics board, employing the state-of-the-art NVIDIA GT 240 GPU to bring desktop performance to the rugged Military and Aerospace market. In addition to meeting increased demand for graphics rendering performance, the GRA111 is the first rugged implementation of a CUDA capable GPU.

The GRA111 supports the 16-lane PCI Express implementation, providing the maximum available communication bandwidth to a CPU such as GE Intelligent Platforms SBC341. The PCI Express link will automatically adapt to the active number of lanes available, and so will work with single board computers in 4- and 8-lane configurations.

SBC341

VPXcel3, Penryn

The SBC341 is a 3U VPX Single Board Computer based on the Intel Core 2 Duo (Penryn) processor and is part of the VPXcel3 product family. This processor is 45nm micro architecture delivering superior performance and energy-efficiency.

This single board is optimized for use in VPX systems with multiple PCI Express port options including an x16, x4, and x1. The x16 is designed to provide a dedicated communication path to a graphics processing unit, such as the GRA110. The x4 and x1 PCI Express ports allows high-speed communication to other single board computers and I/O cards in order to build complex systems.

IPN250 and NPN240

Intel + NVIDIA

These OpenVPX CPU+GPU combination boards feature NVIDIA CUDA-capable GT240 96-core GPUs, where the IPN250 uses 1 NVIDIA GPU and the NPN240 uses 2 NVIDIA GPUS. The GPUs enable the boards to be 15 to 20 times more energy efficient than CPU boards. Multiple boards can be linked to single or multiple hosts to create multi-node CUDA GPU clusters capable of thousands of GFLOPS.

Main Features

- PCIe Gen 2 fabric interfaces
- Multiple capture and output formats supported
- OpenVPX / REDI - 6 Ruggedization build levels

SBC324 and SBC624

2nd Generation Intel Core i7

The VPXcel3 SBC324 is a 3U VPX Single Board Computer built around the 2nd Generation Intel Core i7 processor with fully integrated graphics and memory controller in one device. This 32 nm processor incorporates Advanced Vector Extensions (AVX) technology that provides tremendous signal processing capability in a 3U form factor. Coupled with the QM67 chipset, the rugged OpenVPX SBC324 is available with dual- or quad-core processing at up to 2.5 GHz.

The VPXcel6 SBC624 Single Board Computer is also based on the 2nd Generation Intel Core i7 processor and QM67 chipset. The SBC624, which offers a range of air- and conduction-cooled build levels, provides an unmatched level of I/O bandwidth for both onboard and offboard functions. Onboard XMC mezzanine expansion sites provide enhanced system flexibility.

Getting Started

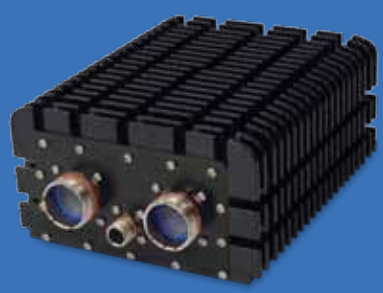
The CUDA Starter Kit is intended to allow engineers a quick and easy start up with a CUDA-based system, either for evaluation and benchmarking or to develop a specific application. Containing all the hardware that is needed to begin development, customers can immediately begin loading their chosen operating system on the delivered equipment.

NVIDIA drivers and CUDA SDK are delivered on CD for Wind River Real Time Linux®, Concurrent Red Hawk real time Linux, Red Hat, SuSE, Windows® XP, and Windows Vista. For latest versions and technical support, please refer to www.nvidia.com

SC-S-CUDA3USK1
3U GPGPU Development Chassis

MAGIC1
Pre-configured Rugged 3U GPGPU Development Chassis Solution Set Contains:

- SBC341 Intel CPU
- GRA111 NVIDIA GPU
- 64GB solid state SATA drive



SC-S-CUDA6USK1
6U OpenVPX Development Chassis

Supports multiple fabric topologies within a single chassis without having to define the system backplane requirements in advance of application development, which greatly reduces risk and time to solution for your applications

GE Intelligent Platforms Contact Information

Americas: **1 800 433 2682** or **1 434 978 5100**
Global regional phone numbers are listed by location on our web site at **www.ge-ip.com/contact**

defense.ge-ip.com/gpgpu

06.11 10M GFA-1790