Open standard middleware enables new classes of HPEC solutions

abaco
S Y S T E M S

Open standard middleware enables new classes of HPEC solutions

The military embedded computing landscape has been transformed from where it was 20 years ago – and that has been almost entirely enabled by the ability of prime contractors, systems integrators and OEMs to leverage the products of COTS manufacturers like Abaco Systems who take leading edge commercial technologies and apply them successfully to the world of military computing.

A look at the commercial landscape today reveals cell phones that are putting vast amounts of location-aware information – and the ability to process that information – directly into the hands of consumers. We see businesses like Amazon and Google creating, managing and leveraging extensive compute resources, based on powerful multi/many core processors from the likes of Intel and NVIDIA, together with interconnect technologies like Ethernet and PCI Express® - turning real-time data into indispensable business information. The Internet of Things has become a deployable reality, with data derived from multiple millions of connected sensors.

Some of these technologies have migrated into the military embedded computing world. Just as cell phones exist on the edge of the network, so now, new generations of small, lightweight, low power, incredibly capable devices are being deployed on the leading edge of the battlefield. The technologies used by companies such as Amazon and Google within their HPC (high performance computing) data centers are being made available to the defense market by Abaco to bring high performance embedded computing (HPEC) to military platforms of all shapes and sizes. Increasingly, more and more sophisticated sensors are being deployed to give strategic and tactical information advantage to warfighters – not to mention their use in maximizing military asset availability and minimizing cost of ownership, and this drives the need for HPEC.

The military is unquestionably deriving enormous benefit from these commercial technologies, because it needs exactly what the commercial world needs. The only differences are that the military needs those technologies to be rugged enough to withstand the rigors of deployment on the battlefield – and it needs those technologies to be supported over the multi-year cycle of the typical program.

That's where Abaco Systems comes in.
This white paper describes the key market drivers, technologies and building blocks that the military is increasingly embracing from the commercial world that provide a clear path to increased operational capabilities within the size, weight and power (SWaP) constraints that are a feature of today's most demanding programs.

Open standard middleware enables new classes of HPEC solutions

Introduction

Knowledge is power. Today, the world is drowning in ever-increasing amounts of data coming at us from many sources. How can we turn this data into knowledge and how can we use this knowledge to improve outcomes across a wide range of potential applications?

The Internet of Things (IoT) is here, resulting in an explosion in the amount of data being produced and distributed around the connected world. Businesses are being built by using Big Data analytics to track trends and predict events before they happen in order to increase productivity and minimise cost and risk across multiple markets. New services are being made possible by networks of data centers, computing clusters, servers and data stores that process enormous data streams with increased efficiency over time. In order to continue to service customer expectations of instant access to information anytime, anywhere, these networks must process more data more effectively while responding to customer requests within acceptable time frames.

Connected world

To do this, knowledge business markets are driving big investments in baseline technologies that support high performance computing platforms, faster networks and high capacity data storage. In turn, these platforms increase our ability to support new and innovative processing, exploitation and dissemination techniques that turn data into deliverable knowledge within shorter time frames. Companies such as Intel®, AMD, NVIDIA®, Micron®, Samsung, AVAGO (PLX), Mellanox®, Broadcom and others are introducing higher performance hardware platforms (processors, fast memory, high speed interconnects, network switches and storage arrays) to enable scale-out of Big Data centers and HPC clusters. In parallel, various standards bodies are creating open software platforms with wide community support to enable increased data

throughput and shorter response times along with application portability across the Cloud, delivering information and services to the connected world.

Technology providers including Apple®, Samsung, Microsoft® and others ship many millions of connected devices - smart phones, tablets, laptops and so on - providing more processing power at the edge of the network. These devices contain hardware technology from companies like ARM®, TI, NVIDIA, Intel, Freescale™ and others that enable whole new service models. In addition to personal connected devices, automobile and commercial vehicle manufacturers are embracing new business models by enabling connected services delivered right into the vehicle.

Intelligent devices at the edge of the network enable a wide range of new services. For example, many of these devices know where they are through the use of GPS. Service providers can also track typical usage trends and behavior patterns. All of this information can help predict what sorts of activities customers engage in and what services consumers are likely to want or need. Such data represents a valuable resource for many commercial service providers or merchandising operators.

These paradigms are being extended to the industrial internet and the internet of machines as well. Companies are able to offer extensive health monitoring and asset management services by providing predictive analytics to aero engine, power generation, oil and gas and other industrial customers. Defense and aerospace systems integrators are also benefiting from the use of smart platforms at the edge of the network. Such platforms are capable of delivering game-changing operational reach and effectiveness across both manned and unmanned platforms for land-, air- and sea-based applications.

Open standard middleware enables new classes of HPEC solutions

Enabling knowledge-assisted processing (KAP) at the edge

Today's latest smart phone or tablet packs more compute power than the typical PC of less than a decade ago within a low power portable form factor. As noted previously, these devices know where they are and can send and receive data to and from the network to access information that enables the user to know more about their location as well as any services, environmental conditions (traffic jams, travel or weather information) that may be of interest.

Support

These devices could be referred to as knowledge assisted processing (KAP) platforms since they use on-board processing to provide location-based services by pulling relevant environmental data from a remote data source and making it available to the user in a format they can understand, within a time frame that enables them to take action to support their needs.

Embedded computing platforms can also benefit from KAP paradigms in a variety of ways that are relevant to their operational requirements. For example, sensors on a piece of industrial equipment such as a jet engine can process real-time data - temperatures, pressures, speeds, vibration and fuel consumption, for example - to ensure safe and efficient operation of the equipment. These platforms can compare acquired sensor data to expected values and raise alarms if required. In addition, KAP-enabled systems have the potential to greatly enhance operational effectiveness by dynamically tuning engine performance based on data models derived not just from the on-board real-time data but from many thousands of hours of data collected from many jet engines operating in similar environmental conditions all around the world.

Knowledge

Such optimized operating models are derived from Big Data analytics that look for relationships between very large data sets that encompass a much wider range of variables and knowledge than those collected from a single platform. These models can take into consideration information such as the working life of the engine, maintenance history and population trends across a family of similar platforms, and so on. This knowledge can be used to optimize operational effectiveness of the engine, aircraft and entire fleet by tuning performance parameters to align with business, environmental or safety objectives.

Open standard middleware enables new classes of HPEC solutions

HPEC at the edge

SWaP optimized rugged HPEC platforms

3U VPX HPEC SYSTEM	6U VPX HPEC SYSTEM
>1 TERA FLOPS	4 TERA FLOPS

Diagram 1: Two OpenVPX HPEC systems from Abaco Systems.

As described above, HPC clusters and data center installations are scaling out to provide the Big Data infrastructure needed to support new and evolving business models. In addition, large HPC clusters are designed to support very compute-intensive applications that might include running weather simulation, fluid dynamics, physics or other large scale mathematical models. Data center platforms can be optimized to provide cloud services to multiple users and to handle very large data sets in support of Big Data analytics. Both types of installations use similar open system technologies to provide data processing resources or slices of compute capability to multiple users, either simultaneously or on an exclusive basis.

Highest performance

A typical installation will have multiple racks of Linux servers connected by a high bandwidth scale-out Ethernet network. Each of the server racks will have multiple processor nodes connected via a scale-in fabric which could be Ethernet or InfiniBand®, depending on inter-process communication (IPC) requirements and the topology of the system. For example, InfiniBand is well suited for HPC clusters or compute slices that might include both CPUs and parallel processing accelerators such as graphics processing units (GPUs). Some of the world's highest performance HPC installations harness many thousands of multi-core CPU and many-core GPU slices yielding PetaFLOP performance to accelerate very compute-intensive applications.

Ethernet and InfiniBand fabrics enable very high IPC throughput, very low memory-to-memory latencies together with CPU off-load based on kernel by-pass capability by making use of remote direct memory access (RDMA) software driver stacks. Open Fabrics Enterprise Distribution (OFED) RDMA middleware is one example of a community initiative sponsored by both industry and academia under the Open Fabrics Alliance (openfabrics. org). Another such initiative supports an industry standard IPC middleware called MPI (Message Passing Interface).

These open system platforms provide a consistent application programming interface (API) across multiple installations, enabling application scaling from small to large numbers of processing slices in order to accommodate various job queues within acceptable time frames.

Effectiveness

These same concepts and technologies can now be instantiated on deployed embedded platforms providing TeraFLOP levels of compute performance to expand the reach and effectiveness of a variety of deployed defense and aerospace platforms. The same open middleware APIs used on HPC and data center installations, along with processing slices based on these same Linux® hardware architectures, can be found in the latest rugged form factors offered by companies such as Abaco Systems.

Whereas HPC clusters can be measured in thousands of compute slices consuming hundreds or thousands of kilowatts within a 100,000 square foot installation, a typical deployable HPEC platform might occupy a small number of cubic feet with a power budget of one or two Kilowatts or less. Diagram 1 shows two examples of rugged, scalable HPEC systems from Abaco.

Open standard middleware enables new classes of HPEC solutions

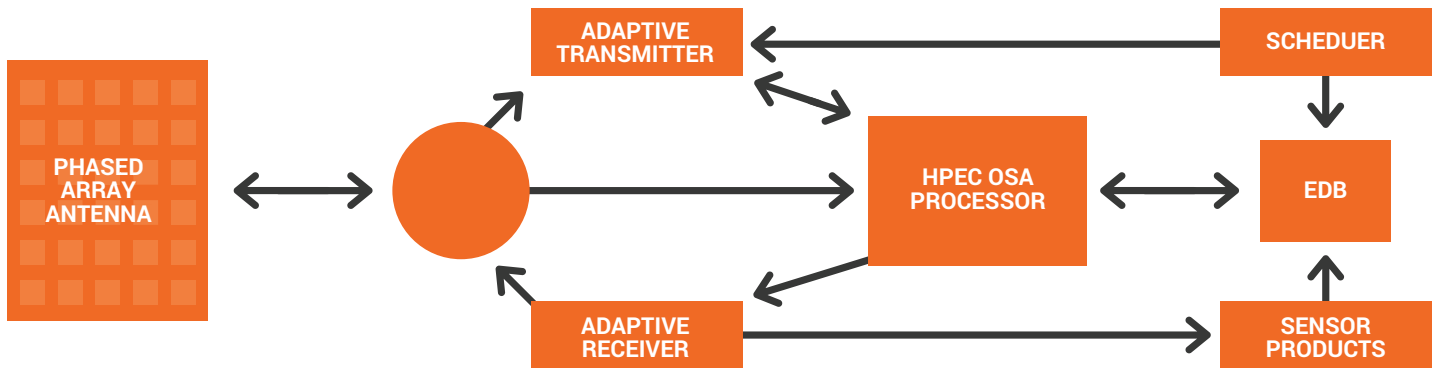## *Application portability and scalability*

Diagram 2: Block diagram of knowledge assisted cognitive radar processor

Such open architecture HPEC systems can run the same applications developed on HPC clusters. The open APIs and middlewares provide application portability while the compute-, fabric- and storage hardware modules can be scaled from few to many slices to provide the best SWaP profile to address various deployed system mission objectives.

Such general purpose HPEC platforms can be configured to provide high data throughput to cater for continuous data streams from high resolution sensors such as radar or radio antennae, sonar acoustic heads or multi-spectral camera arrays. In addition to real-time sensor- and image processing, these platforms can record and retrieve relevant data from on-board storage or from a network resource in order to tailor the application to specific operational environments while maintaining the ability to draw on other data sources to adapt to new operational environments within a single sortie.

Multi-mode capabilities
The ability to consolidate multiple applications onto a single reconfigurable HPEC system enables a whole range of multi-mode capabilities, sometimes combining the functionality of more than one legacy processor onto a single multi-role platform. Such strategies can reduce the amount of discrete single mode systems on older platforms and replace them with fewer multi-purpose systems with more capability to address a variety of operational requirements while greatly reducing the weight and power requirements of the overall vehicle.

Such HPEC systems can also facilitate the development of cognitive sensor processing systems that are able use knowledge-assisted processing techniques to optimize sensor inputs

and outputs and thereby maximize mission capabilities and effectiveness. They do this by analyzing both acquired real-time mission data and archived data that enables the mission data to be placed in a wider knowledge-based context.Consider a multi-mode cognitive radar platform (diagram 2) that knows where it is and understands the environment it might be working in. Such systems could draw on a combination of real-time, real world sensor inputs as well as operating models and archived data such as multispectral maps stored in on-board environmental data bases (EDB). It could also make use of networked sensor inputs from other assets as well as pre-planned operational rules of engagement to greatly increase its ability to find weak targets in complex environments while minimizing the ability of others to detect its presence within the theater of operations.

Reducing latency
Such systems could use all of these data resources to predict or look ahead to where the vehicle will be within the next few seconds and adjust both transmit and receive modes to tune the radar antenna to best effect by comparing archived data maps to its current position in advance. This strategy would reduce processing latency on the real-world signals of interest by predicting expected returns from unwanted background features and mapping them out of the sensor processing chain.

The application can use the HPEC processor cluster to handle both real-time signal processing streams as well as non-pipelined workloads that might use archived information to minimize environmental interference or tune the transmitter in order to minimize the illumination of clutter while adapting receive filters to look for certain return wave forms or frequencies that might be expected from targets of interest.

Open standard middleware enables new classes of HPEC solutions

Architecture considerations

MEMOERY
ACCESS
CLOCK
CYCLES

=2 — CPU REGISTERS

1 — LEVEL1 CACHE

1*N — LEVEL2 CACHE

LEVEL3 CACHE

ON-BOARD SDRAM

LOCAL SSD

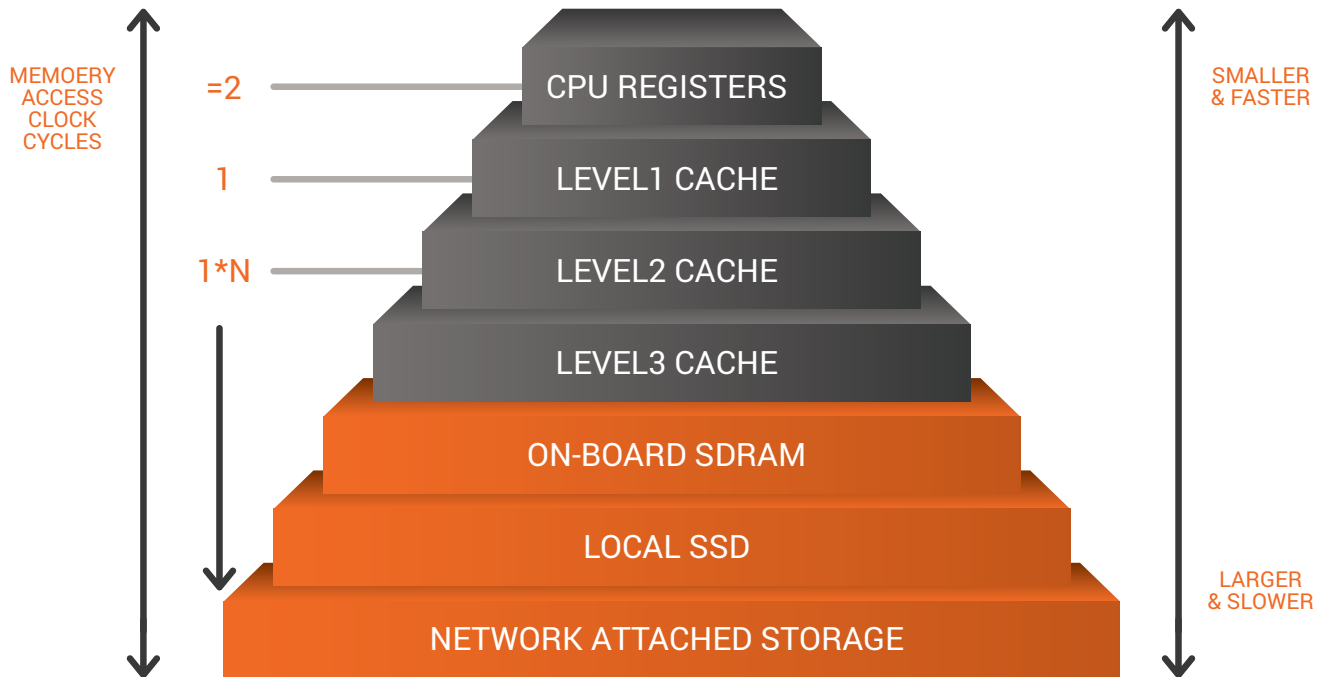NETWORK ATTACHED STORAGE

SMALLER
& FASTER

LARGER
& SLOWER

Diagram 3: CPU memory access cycles.

One of the key factors that should be addressed by system architects when designing such systems is the ability of the processor to access, process and react to relevant mission data within actionable time frames. These time frames could be measured in terms of seconds or milliseconds depending on the relative speed of the platform in relation to any targets or threats. Designers must therefore consider the locality of the data and the speed at which it can be accessed and processed together with any inherent system latencies in order to adapt their application to best effect.

High speed
Modern multi-core processors incorporate high speed pipelined data buses with multiple levels of on-chip memory caches and multiple high bandwidth memory controllers as well as high speed storage and network interfaces.

Each of these can be used to store and retrieve relevant data sets of different sizes at varying speeds (latencies). Diagram 3 provides a framework that could be used when considering how to maximize processing efficiencies and minimize latencies in meeting timing requirements.

HPEC commercial off-the-shelf (COTS) processing modules take advantage of these multi-core CPUs, many-core GPUs as well as InfiniBand and Ethernet switch fabric modules for inter-process communication and system control. Data I/O is supported with high speed PCIe™ expansion plane connectivity and/or 10 Gigabit Ethernet. The OpenVPX standard is supported by a wide community of board- and system vendors through the VME International Trade Association (VITA.com).

Cost-effective
These modules are available from multiple vendors and can be scaled from small 3U VPX multi-board systems to much larger 16- or 18-slot 6U VPX platforms to meet different size, weight, power and cost (SWaP-C) requirements. This ecosystem provides a clear migration path to a fully rugged deployable system architecture that can interface directly to host applications developed on commercial HPC servers, thus providing a cost-effective means to deploy advanced, knowledge-assisted sensor processing solutions.

Open standard middleware enables new classes of HPEC solutions

SWaP optimization for HPEC platforms

Optimizing an application for real-time performance can be a time-consuming task. However, there are application development tools that can reduce the time and effort required to take full advantage of the SWaP-constrained platform in a very efficient manner.

Such tools allow the visualization of the application through multiple windows such that the developer can quickly understand how the application is mapped to the available hardware resources without having to develop low level code.

One such tool from Abaco provides a user-friendly graphical user interface, a high performance IPC middleware library as well as more than 600 digital signal processing (DSP) and math function libraries to build and run advanced sensor processing applications. Diagram 4 shows a typical multi-threaded pipelined DSP application scaled across multiple compute nodes using the InfiniBand data plane fabric.

Diagram 4: Task mapping across a 6U OpenVPX multi-board, multi-node CPU/GPU system using InfiniBand and Ethernet.

Developers can maximize performance per watt on such systems by partitioning various processing tasks across the available resources to best effect. For example, certain signal processing loops can be greatly accelerated by ensuring data is available in on-chip memory when needed. This can sometimes be achieved by splitting data streams into parallel processing pipe lines. Other tasks may require all-to-all data movement, in which case high speed inter-process communication fabrics such as InfiniBand

can offer an effective way to share data and processing loads across groups of CPUs and GPUs within the system.

Application development tools such as Abaco's AXIS enable developers to quickly try different approaches in order to achieve timing requirements across the system. Such tools can be used on commercial servers as well as on deployable OpenVPX systems in order to scale the application to best effect, taking maximum advantage of the available platform resources while also ensuring application portability.

Diagram 5 shows a typical signal processing application running on a small foot print HPEC cluster using Gen 3 PCIe fabric for inter-process communication. This application was developed on a 6U OpenVPX Linux cluster with an InfiniBand data plane and re-hosted on a small multi-board 3U OpenVPX platform running VxWorks with a peer-to-peer (P2P) PCIe data plane.

The same application code is running on both the 6U Linux InfiniBand cluster and the 3U platform, providing application portability and re-use across platforms to cater for a variety of SWaP-C requirements.

Diagram 5: Signal processing application mapped to a 3U OpenVPX Intel 4th Generation Core-i7 HPEC cluster running VxWorks and using gen 3 PCIe P2P data plane fabric.

Open standard middleware enables new classes of HPEC solutions

Conclusion

System integrators are now able to develop and deploy advanced knowledge-assisted embedded platforms using open system architectures to achieve very high performance in order to expand mission capabilities across a range of sensor processing applications. As noted, HPEC platforms can be adapted to use real-time sensor data in combination with archived intelligence to adapt operation of a multi-mode radar system to cater for changing and varied operational environments with the potential to service multiple missions.

Such concepts can be applied to image-, video-, sound- and electromagnetic spectrums to provide a true multi-spectral cognitive sensor processing capability over time.

Open system architectures and open software middleware offerings with industry standard APIs afford application scaling and portability from commercial server and HPC clusters to deployable, rugged HPEC OpenVPX platforms enabling multi-generation technology insertion road maps to support expanded operational needs now and new multi-role processing platforms into the future.

As we transition from processors with one to four cores to higher core counts – twelve to sixteen is now commonplace - and to 40G Ethernet interconnects on 6U and 10G Ethernet on 3U, this ability to scale up and down without compromising agnosticism of architectures and vendors is key to efficient development.

WE INNOVATE. WE DELIVER. YOU SUCCEED.