ENEA White Paper

The Carrier-Grade Data Plane:

Leveraging the Enea dSPEED Platform to Achieve Highly Available User Plane Processing

Market pressures demand that each new generation of networking and communications system deliver both higher performance and new functionality at a lower price. In many applications, the ability to provide continuous operation, even in the presence of failures (also known as high availability), is a key customer requirement.

To enhance the availability of their systems, equipment makers combine redundant hardware with network supervision, fault management, and shelf management software. This software, collectively known as high availability (HA) middleware, facilitates continuous operation, enhances service quality, and reduces operating cost by making it easier to monitor, repair, configure, and upgrade live systems.

Historically, equipment makers have developed their HA software in house. A growing number, however, are turning to COTS middleware solutions. These out-of-the-box middleware platforms, compliant with industry standards such as The SA Forum specifications, reduce time to market, enhance portability, and enable equipment makers to utilize interoperable, best-in-class products from multiple suppliers.

COTS middleware solutions have traditionally focused on the control plane. Equipment makers, however, require a comprehensive management solution that addresses the DSP-centric user plane (data plane) as well as the control plane. Enea's high-availability Element middleware, when coupled with the Enea® dSPEED Platform™, is the first integrated HA management solution that spans both the data plane and the control plane.

Mobile broadband drives data plane complexity

Soon, mobile broadband users will vastly outnumber wired broadband users. This increase in broadband traffic brings with it a corresponding increase in the number of high-speed data streams that must be processed and managed.
To deliver this high-speed data plane processing in an efficient manner, equipment makers are utilizing multi-core DSPs, which can process a large number of channels

in parallel while sharing resources such as memory and peripherals. By deploying clusters of these devices, known as DSP farms, on a single blade, equipment makers can greatly reduce per-channel cost and power dissipation.

As equipment makers utilize more and more DSPs to bolster data plane processing bandwidth and efficiency, managing those DSPs and maximizing system availability is become more challenging. The greater the number of DSPs, the greater the complexity of restarting individual DSPs and blades. Higher DSP density also makes it more difficult to isolate faults to a single DSP core and maintain continuous operation for the remaining cores on the blade. In addition, the growing complexity of these emerging designs forces new generations of DSPs to handle the load, which in turn causes the network equipment providers to re-implement this platform SW over and over for new cores, IO architectures, boards and other components. This clearly creates a nearly endless cycle of re implementing the DSP software platform over and over again with little reuse benefits for the R&D team.

Enea dSPEED Platform

The Enea dSPEED Platform provides the fine-grain debug, management, and control capabilities needed to ensure continuous operation for data plane blades equipped with high-density DSP farms. The platform makes it easy for equipment makers and service providers to monitor, control, and debug individual DSP cores on multiple DSPs and DSP cores. It also enables them to pin point faults to an individual DSP core and isolate that core without impacting overall blade operation.

The Enea dSPEED Platform, together with Enea Element, provides a system-wide management solution for the control plane and the data plane. Enea Element provides control plane management and overall data plane management for multiple blades at the chassis level. The Enea dSPEED Platform manages DSP farms at the blade level, providing carrier-grade services to remote DSP nodes in the data plane.

The Enea dSPEED Platform essentially extends the capabilities of Enea’s OSE®ck DSP real-time operating system with a complete set of DSP management services. These services include start up, configuration, error handling, monitoring, supervision, event notification, logging, tracing, diagnostics, statistics and post-mortem debugging.

The Enea dSPEED Platform also provides a comprehensive suite of inter-process communications and network services built on Enea’s LINX IPC platform.

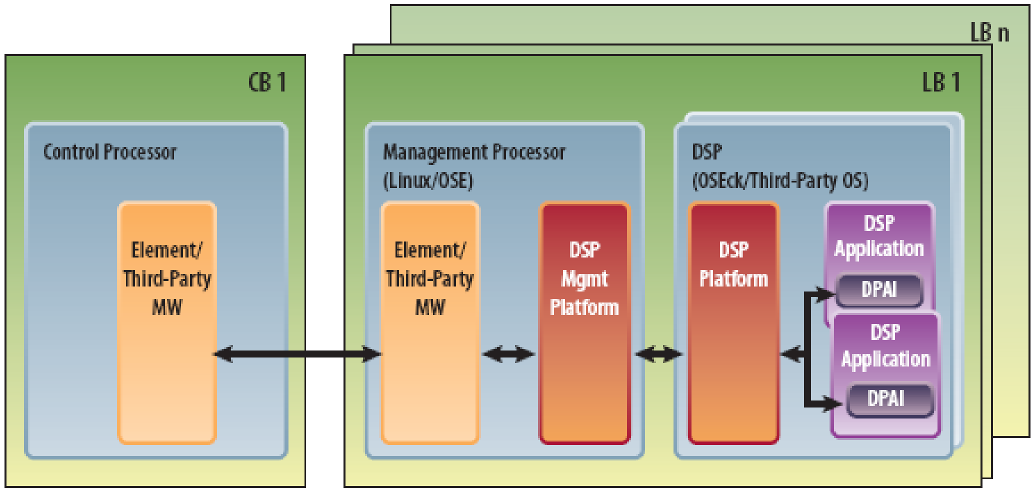

Figure 1 The Enea dSPEED Platform Overview

The management component of the Enea dSPEED Platform, known as the DSP Management Platform (DMP), typically resides on a dedicated management processor. However, it may also be deployed on the control processor. This is becoming more commonplace as equipment makers strive to reduce the number of line blades by integrating control and management on a single device. The target component of the Enea dSPEED Platform, known as the DSP Platform (DP), runs on Enea's OSEck real-time operating system, which resides on the DSPs under management. Enea's LINX services handle inter-process communications (IPC) between the DMP host and DP target.

Fault detection

Given The Enea dSPEED Platform's breadth of features and capabilities, the question then is – how does dSPEED meet the needs of developers working to meet the requirement for more robust systems supplying carrier grade services to demanding customers?

Highly available systems must be able to detect when a node, process or communication fault occurs. They must also be capable of notifying interested parties. In the control plane, the management software periodically monitors each node's "heartbeat" to determine whether it is alive or has failed. In the data plane at the DSP core level, the management software must be able to identify failures far more quickly in order to prevent DSP core failures from affecting other DSP cores, either on the same DSP or other DSPs.

Error handling
Centralized error handling is critical for responding immediately to fatal errors, bringing the DSP core to a safe state without affecting other cores, and providing error information to the host management software (DMP user). The DSP-resident DP can detect three different types of errors: Processor errors (through the OSEck error handler) such as memory access exceptions and illegal instructions; Kernel errors (OSEck software error), such as a bad system call parameter or memory corruption; and user/application errors reported as fatal or non fatal through the DP Application Interface (DPAI)

Supervision (heartbeat)
The Enea dSPEED Platform provides three different supervision (i.e., heartbeat) mechanisms. Only the LINX mechanism is required, but platform level supervision is highly recommended.

- A basic link level supervision mechanism for quickly detecting failing switches or communication peripherals.
- A platform level supervision mechanism (implemented as ping pong signals over LINX) for detecting overloaded DSPs and crashed DSPs that are running out of control. The interval and timeout is configurable. DSP applications can subscribe to supervision signals so that the DMP is notified if the application does not respond.
- A DSP watchdog that can be used as an additional supervision mechanism. A watchdog timeout triggers an NMI to invoke the appropriate dSPEED error handler.

Monitoring
Another Enea dSPEED Platform service that can be used to detect faulty behavior is monitoring. The DMP user can monitor DSP cores, CPU and memory load and set thresholds that trigger an alarm event to the DMP user. The OSEck system information API is used by the DSP to retrieve the needed information.

Isolation and containment

When a fault is detected, it is critical that the faulty DSP core be quickly put into a safe state so that it does not affect other cores in that DSP (or in other DSPs).

Fatal error handler
When a fatal error occurs, the DP locks all interrupts to prevent faulty users or interrupt processes from taking control.

An interrupt or IPMI (ATCA/uTCA) call is typically used to notify the host that an error has occurred. The error information (CoreID, error code, file, line, and error message string) is provided through the chosen communication media, such as sRIO or Ethernet.

Core NMI/Reset

Through the DMP Application Interface (DMAI), the user can take control over a crashed DSP (e.g. failing supervision) by using NMI, reset lines, or IPMI calls (if the hardware permits). When an NMI is triggered, the DSP enters the fatal error handler and generates error information to the host. Other cores on the same DSP can then finish the current task prior to a DSP restart.

Restart

The dSPEED platform provide the management services that enable DSPs and/or the management processor to be restarted after a fault or upgrade.

Boot loading

Currently, the platform supports two efficient DSP boot loading mechanisms: sRIO and Ethernet. To minimize downtime, the boot sequence is optimized such that it does not perform time consuming self tests (unless ordered to do so by the user)

Host restart

If a fault occurs on the management processor (DSP host), that fault is typically handled by the control plane middleware, such as Enea Element. If the host is restarted, the DMP can detect running DSP cores at startup, reestablish the LINX connection, and resume supervision without causing any faults on the DSP cores.

Debugging faults

The ability to obtain comprehensive, accurate debug information about what causes faults is essential for constantly improving DSP system availability. The ability to detect hardware problems through diagnostics also helps limit the number of restarts needed before faults can be detected and the faulty DSP core is taken offline -- while the rest of the DSP system continue to run.

Post-mortem debugging

The Enea dSPEED Platform's post-mortem debug facilities provide a wealth of information that can help programmers identify the errors that cause a program or processor crash.

Following a crash, the platform can perform a system dump that provides static DP Information such as a text log, trace log DMP user-defined memory sections, and DP user-defined memory sections.
The platform can also perform an OSEck core dump that provides information about error messages, memory usage, processes, signals, saved registers and stack content.

Diagnostics

The diagnostic tests verify that OSEck and important hardware functions are working correctly. The Enea dSPEED Platform provide three levels of diagnostics: offline diagnostics, self test, and extended self test with non-destructive test. The offline diagnostics service enables destructive tests to be uploaded to the DSP and checked for results. This service is typically used for comprehensive tests such as checking every bit of memory.

Self test is a short test of the vital functions needed for successful booting, such as memory, the DSP core, and peripherals used for communications (e.g. Ethernet or sRIO).The self test is useful to avoid crashes after boot-up in a faulty DSP environment. A longer version of the self test, which can be executed during runtime, includes a full suite of non-destructive tests for functions such as memory, peripherals, clock, timer and the CPU/DSP core.

Redundancy

The Enea dSPEED Platform's DSP management services can be used to implement a variety of redundancy policies. The platform itself does not take any actions on faults, but it does provide the DMP user with comprehensive information when a fault occurs. The user can then make the appropriate decision based on knowledge about the application behavior. For example, if a shared execution image is used for all DSPs on a board, the management services can be used to place one DSP core in standby. This enables the DSP to assume the active role should one of the active cores fail (N+1 redundancy).

Checkpointing

One of the most important capabilities of any high-availability management platform is the ability to periodically checkpoint or save the state of an application process. Because it has limited application knowledge, the Enea dSPEED Platform leaves the responsibility for saving specific application state to the application itself. This function is often done via a checkpointing service provided by the system middleware.

As stated previously, the Enea dSPEED Platform is integrated with Enea Element – a complete framework for building high performance, high availability applications – which includes a high-performance, scalable checkpointing service.

Together, they form a complete system wide management and supervision layer with all the tools necessary to develop highly available control and data-processing applications. Although the checkpointing service is provided by Element, the Enea dSPEED Platform provides complimentary services for the equally important task of logging all operations that occur before and after a checkpoint.

Text and trace log

The Enea dSPEED Platform provides both a text log and a trace log. Both logs can be read out after a crash using the post-mortem debugging service. The text log records formatable text strings of software events defined in the source code to a circular DSP local log of configurable size. The DMP user can monitor and read the DSP log. The text log includes two configurable filters. One controls what enters the DSP local log. The other, set by the DMP user, controls what if anything is monitored.

The trace log is a binary log for storing more frequent events such as OS system calls. It also uses a circular DSP local log of configurable size, provides a filter, and can be read by the DMP user. DSP applications can also log their own events.

Future Improvements

Enea plans a number of significant improvements for both the Enea dSPEED Platform and the OSEck RTOS that are geared toward increasing system availability.

Handle failing communications at the link level. E.g. switching from sRIO to Ethernet should the sRIO switch between the DMP and DP fail.

Memory protection. Protect the RTOS and the DSP platform memories from being overwritten by applications.

Core protection. Prevent individual DSP cores from writing to the memory of other DSP cores or memory.

Core restart. Ability to restart individual cores on a multi-core DSP (instead of restarting the entire DSP).

Monitoring framework. Allows the DP user to publish application specific information that can be monitored by the DMP user.

Conclusion

Highly available, high-performance real-time processing is critical to delivering the converged multimedia broadband services that consumers demand. From wireless base stations to advanced media gateways, network equipment designers are increasingly relying on clusters of multi-core DSPs to deliver the scalable performance needed for these compute-intensive applications.

The Enea *dSPEED* Platform provides the fine-grain monitoring, control, debug and fault management capabilities TEMs need to get to market quickly and affordably with equipment that delivers both high performance and high availability.

For more information please visit www.enea.com WP 45 32008

Corporate Headquarters
P.O. Box 1033
Skalholtsgatan 9
SE-164 21 Kista, Sweden
Phone: +46 (0)8 507 140 00
Email: info@enea.se
Web : www.enea.com

US Headquarters
25 South Arizona Place
Chandler, AZ 85225
Phone: 480-753-9200
Toll-free: 866-844-7867
Email: info@enea.com
Web: www.enea.com

Asian Headquarters
1-4-2 Kanda
Ogawa-machi, Chiyoda-ku
Tokyo, Japan 101-0052
Phone: +81 3 5207 6167
E-mail: osesales_jp@enea.se
Web: www.enea.com

ENEA

Embedded for Leaders