NMP ST8 Dependable Multiprocessor (DM)

Dr. John R. Samson, Jr.
Honeywell Defense & Space Systems
13350 U.S. Highway 19 North
Clearwater, Florida 33764
(727) 539 - 2449
john.r.samson@honeywell.com

High Performance Embedded Computing Workshop (HPEC)
23 – 25 September 2008
Outline

• Introduction
  - Dependable Multiprocessor * technology
    - overview
    - hardware architecture
    - software architecture

• Current Status & Future Plans

• TRL6 Technology Validation

• Summary & Conclusion

* formerly known as the Environmentally-Adaptive Fault-Tolerant Computer (EAFTC);
The Dependable Multiprocessor effort is funded under NASA NMP ST8 contract NMO-710209.

This presentation has not been published elsewhere, and is hereby offered for exclusive publication except that Honeywell reserves the right to reproduce the material in whole or in part for its own use and where Honeywell is so obligated by contract, for whatever use is required thereunder.
DM Technology Advance: Overview

- A high-performance, COTS-based, fault tolerant cluster onboard processing system that can operate in a natural space radiation environment

**NASA Level 1 Requirements (Minimum)**

- High throughput, low power, scalable, & fully programmable >300 MOPS/watt (>100)
- High system availability > 0.995 (>0.95)
- High system reliability for timely and correct delivery of data >0.995 (>0.95)
- Technology independent system software that manages cluster of high performance COTS processing elements
- Technology independent system software that enhances radiation upset tolerance

**Benefits to future users if DM experiment is successful:**

- 10X – 100X more delivered computational throughput in space than currently available
- Enables heretofore unrealizable levels of science data and autonomy processing
- Faster, more efficient applications software development
  - Robust, COTS-derived, fault tolerant cluster processing
  - Port applications directly from laboratory to space environment
    - MPI-based middleware
    - Compatible with standard cluster processing application software including existing parallel processing libraries
- Minimizes non-recurring development time and cost for future missions
- Highly efficient, flexible, and portable SW fault tolerant approach applicable to space and other harsh environments
- DM technology directly portable to future advances in hardware and software technology
Desire - ‘Fly high performance COTS multiprocessors

- To satisfy the long-held desire to put the power of today’s PCs and supercomputers in space, three key issues, SEUs, cooling, & power efficiency, need to be overcome

DM Solution

DM has addressed and solved all three issues

- Single Event Upset (SEU): Radiation induces transient faults in COTS hardware causing erratic performance and confusing COTS software

DM Solution

- robust control of cluster
- enhanced, SW-based, SEU-tolerance

- Cooling: Air flow is generally used to cool high performance COTS multiprocessors, but there is no air in space

DM Solution

- tapped the airborne-conductively-cooled market

- Power Efficiency: COTS only employs power efficiency for compact mobile computing, not for scalable multiprocessing

DM Solution

- tapped the high performance density mobile market
DM Hardware Architecture

- System Controller A
- System Controller B
- Data Processor 1
- Data Processor N
- Instruments
- Co-Processor
- Main Processor
- Memory Volatile & NV
- Bridge/Controller
- High-Speed Network I/O
- N Ports Net & Instr IO
- Network A
- Network B
- S/C Interface A
- S/C Interface B
- Mass Data Storage Unit *
- Custom S/C or Sensor I/O *

* Examples: Other mission-specific functions
DMM Top-Level Software Layers

DMM – Dependable Multiprocessor Middleware

- **System Controller**
  - S/C Interface SW and SOH And Exp. Data Collection
  - Mission Specific Applications
  - Policies Configuration Parameters
  - DMM
  - OS – WindRiver VxWorks 5.4
  - Hardware
    - Honeywell RHSBC
  - Hardware FPGA
  - OS/Hardware Specific
  - cPCI (TCP/IP over cPCI)

- **Data Processor**
  - Application
  - Generic Fault Tolerant Framework
  - DMM
  - OS – WindRiver PNE-LE (CGE)
    - Linux
  - Hardware
    - Extreme 7447A
    - FPGA

- **Application Programming Interface (API)**

- **Scientific Application**

DMM components and agents.
DMM Software Architecture “Stack”

DM System

System Controller

Mission Specific Parameters

DMS, CMS, and AMS

VxWorks OS and Drivers

System Controller

Data Processors

MPI Application Process

DMS, CMS, and AMS

Linux OS and Drivers

Data Processor with FPGA Co-Processor

Data Processor

Application Data Check Points

MDS

Linux OS and Drivers

Data Processor

RTMM

- HA Middleware
- Platform Components
- Application Components
- Mission Specific Components
- Dependable Multiprocessor MW Specific Components

- JM – Job Manager
- JMA – Job Manager Agent
- MM – Mission Manager
- FTM – Fault Tolerance Manager
- FEMPI – Fault Tolerant Embedded Message Passing Interface
- SCIP – Space Craft Interface Message Processor

- AS – Application Services
- MDS – Mass Data Storage
- CMS – Cluster Management Services
- AMS – Availability Management Services
- DMS – Distributed Messaging Services
- RTMM – Radiation Tolerant Mass Memory
### Examples: User-Selectable Fault Tolerance Modes

<table>
<thead>
<tr>
<th>Fault Tolerance Option</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMR Spatial Replication Services</td>
<td>Multi-node HW SCP and Multi-node HW TMR</td>
</tr>
<tr>
<td>NMR Temporal Replication Services</td>
<td>Multiple execution SW SCP and Multiple Execution SW TMR in same node with protected voting</td>
</tr>
<tr>
<td>ABFT</td>
<td>Existing or user-defined algorithm; can either detector detect or detect and correct data errors with less overhead than NMR solution</td>
</tr>
<tr>
<td>ABFT with partial Replication Services</td>
<td>Optimal mix of ABFT to handle data errors and Replication Services for critical control flow functions</td>
</tr>
<tr>
<td>Check-pointing Roll Back</td>
<td>User can specify one or more check-points within the application, including the ability to roll all the way back to the original</td>
</tr>
<tr>
<td>Roll forward</td>
<td>As defined by user</td>
</tr>
<tr>
<td>Soft Node Reset</td>
<td>DM system supports soft node reset</td>
</tr>
<tr>
<td>Hard Node Reset</td>
<td>DM system supports hard node reset</td>
</tr>
<tr>
<td>Fast kernel OS reload</td>
<td>Future DM system will support faster OS re-load for faster recovery</td>
</tr>
<tr>
<td>Partial re-load of System Controller/Bridge Chip configuration and control registers</td>
<td>Faster recovery that complete re-load of all registers in the device</td>
</tr>
<tr>
<td>Complete System re-boot</td>
<td>System can be designed with defined interaction with the S/C; TBD missing heartbeats will cause the S/C to cycle power</td>
</tr>
</tbody>
</table>
DM Technology Readiness & Experiment Development Status and Future Plans

**Key:**
- Complete

**Technology Validation**
- TRL4
  - Technology Concept Demonstration
- TRL5
  - Technology in Relevant Environment
- TRL6
  - Technology Demonstration in a Relevant Environment

**Preliminary Design Review**
- TRL7
  - Flight Experiment

**Critical Design Review**
- TRL7
  - Technology Validation

**Radiation Testing**
- Critical Component Survivability & Preliminary Rates
- Complete Component & System-Level Beam Tests

**Dates & Milestones**
- 5/30/05: TRL4 Technology Validation
- 5/17/06: TRL5 Technology Validation
- 10/27/06: NASA ST8 Project Confirmation Review
- 9/08 & 1/09: TRL6 Technology Validation
- 5/31/06: Preliminary Design Review
- 6/27/07: Final Experiment HW & SW Design & Analysis
- 5/06, 4/07, & 5/07: Preliminary Radiation Testing
- 5/08, 7/08, 8/08, & 10/08: Final Radiation Testing
- 5/06, 4/07, & 5/07: Critical Component Survivability & Preliminary Rates
- 5/08, 7/08, 8/08, & 10/08: Complete Component & System-Level Beam Tests
- 9/08 & 1/09: TRL6 Technology Validation
- Launch 11/09: Mission 1/10 – 6/10

**Notes**
- Per direction from NASA Headquarters 8/3/07; The ST8 project ends with TRL6 Validation; Preliminary TRL6 demonstration 9/15/08; Final TRL6 demonstration 1/10/09
DM TRL6 Testbed System

System Controller:
- Wind River OS
  - VxWorks 5.4
- Honeywell Ganymede SBC (PPC 603e)

Data Processor:
- Wind River OS
  - PNE-LE 4.0 (CGE) Linux
  - Extreme 6031
  - PPC 7447a with AltiVec co-processor

Memory Card:
- Aitech S990

Networks:
- cPCI
- Ethernet: 100Mb/s

SCIP – S/C Interface Process
DM TRL6 (Phase C/D) Flight Testbed

Custom Commercial Open cPCI Chassis

System Controller (flight RHSBC)

Backplane Ethernet Extender Cards

Flight-like COTS DP nodes

Flight-like Mass Memory Module
Automated SWIFI (SW Implemented Fault Injection) Tests:

KEY:
- DP - COTS Data Processor
- NFTAPE – Network Fault Tolerance And Performance Evaluation tool
System-Level Proton Beam Tests:

Key:
- DP - COTS Data Processor

Diagram:
- S/C Emulator
- Radiation Source
- Borax Shield
- Aperture for Radiation Beam
- Proton Beam
- DP Board on Reversed cPCI Backplane
- Ethernet Switch
- TRL6 Test Bed
- System Controller
- DP Boards
- cPCI
Dependable Multiprocessor Experiment Payload on the ST8 “NMP Carrier” Spacecraft

**ST8 Orbit:**
- Sun-synchronous
- 955 km x 460 km @ 98.2° inclination

**Software**
- Multi-layered System SW
  - OS, DMM, APIs, FT algorithms
- SEU-Tolerance
  - Detection
  - Autonomous, transparent recovery
- Applications
  - 2DFFT, LUD, Matrix Multiply, FFTW SAR, HSI
- Multi-processing
  - Parallelism, redundancy
  - Combinable FT modes

**Flight Hardware**
- Dimensions
  - 10.6 x 12.2 x 24.0 in.
  - (26.9 x 30.9 x 45.7 cm)
- Weight (Mass)
  - ~ 61.05 lbs
  - (27.8 kg)
- Power
  - ~ 121 W (max)

The ST8 DM Experiment Payload is a stand-alone, self-contained, bolt-on system.
DM Markov Models

Data Flow Diagram for DM Markov Models

- **SWIFI Tests** (slide 13) → Log files
- **Post-Processing** (slide 16) → Filter Raw Data
- **Canonical Fault / Error Syndrome Matrix** (slide 20) → Component area → Circuit knowledge → Error syndrome rates (Candidate App)
- **App Profile Database** (slide 32) → Test App
- **Comp-level Rad. Tests** (slide 5) → Relevant environment
- **Space Rad. Tool** (slide 7) → Comp-level fault rates
- **Arch. Mapping Function** (slide 10) → Error syndrome rates (Candidate App)
- **Scaling and Reduction Function** (slide 36) → Error syndrome rates (Candidate App)
- **Error Syndrome-to-Markov Variable Mapping** (slide 38) → Markov models (slide 40)
- **Reliability**
- **Availability**
- **Performance**

* The Space Rad tool will be bypassed when validating model against system-level radiation tests.
DM Technology - Platform Independence

- DM technology has already been ported successfully to a number of platforms with heterogeneous HW and SW elements
  - Pegasus II with Freescale 7447a 1.0GHz processor with AltiVec vector processor with existing DM TRL5 Testbed
  - 35-Node Dual 2.4GHz Intel Xeon processors with 533MHz front-side bus and hyper-threading (Kappa Cluster)
  - 10-Node Dual Motorola G4 7455 @ 1.42 GHz, with AltiVec vector processor (Sigma Cluster) with FPGA acceleration
  - DM flight experiment 7447a COTS processing boards with DM TRL5 Testbed
  - DM TRL6 flight system testbed with 7447a COTS processing boards, with AltiVec
    -- > 300 MOPS/watt for HSI application (> 287 MOPS/watt including System Controller power)
  - State-of-the-art PA Semiconductor dual core processor
    -- demonstrated high performance working under DM DMM umbrella
    -- > 1077 MOPS/watt for HSI application
DM Technology - Ease of Use

• Successfully ported four (4) real applications to DM testbeds
  - HSI *
    - eminently scalable MPI application
    - ~ 14 hours to port application to DM system with DMM, hybrid ABFT, and in-line replication
    - ~ 4 hours to implement auto-correlation function in FPGA
  - SAR *
    - eminently scalable MPI application
    - ~ 15 hours to port application to DM system with DMM, hybrid ABFT, in-line replication, check-pointing
  - CRBLASTER (cosmic ray elimination application) **
    - eminently scalable MPI application
    - ~ 11 hours to port application to DM system with DMM, hybrid ABFT, and in-line replication
    - scalability demonstrated ~ 1 minute per configuration
  - QLWFP2C (cosmic ray elimination application) **
    - fully-distributed MPI application
    - ~ 4 hours port application to DM system with DMM
    - scalability demonstrated ~ 1 minute per configuration
  - NASA GSFC Synthetic Neural System (SNS) application for autonomous docking *
    - ~ 51 hours to port application to DM system with DMM (includes time required to find a FORTRAN compiler to work with DM)

* Port performed by Adam Jacobs, doctoral student at the University of Florida and member of ST8 DM team
** Port performed by Dr. Ken Mighell, NOAO, Kitt Peak Observatory, independent 3rd party user/application developer with minimal knowledge of fault tolerance techniques, per TRL6 requirement
Summary & Conclusion

- Flying high performance COTS in space is a long-held desire/goal
  - Space Touchstone - (DARPA/NRL)
  - Remote Exploration and Experimentation (REE) - (NASA/JPL)
  - Improved Space Architecture Concept (ISAC) - (USAF)

- NMP ST8 DM project is bringing this desire/goal closer to reality

- DM TRL6 Technology Validation Demonstration 9/15 & 9/16/08
  - system-level radiation tests validated DM operation in a radiation environment
  - demonstrated high performance, high Reliability, high Availability and ease of use

- DM technology is applicable to wide range of missions
  - science and autonomy missions
  - landers/rovers
  - CEV docking computer
  - MKV
  - UAVs (Unattended Airborne Vehicles)
  - UUVs (Unattended or Un-tethered Undersea Vehicles)
  - ORS (Operationally Responsive Space)
  - Stratolites
  - ground-based systems & rad hard space applications