T3E Resiliency Enhancements

Dean Elling

Software Engineer

SGI

A Brief History

PE Resiliency

- **Initial releases of UNICOS/mk**
  - system panicked
  - processes hung
  - system would have to be rebooted

A Brief History

PE Resiliency

- **UNICOS/mk matures**
  - failed PE was isolated
  - processes were cleanly terminated
  - application PE region was partitioned
  - command PE remained unusable

A Brief History

PE Resiliency

- **UNICOS/mk 2.0.3**
    - SWS Warmboot of software panicked PE
    - failed PE was cleanly integrated back in to the running system

T3E Resiliency Enhancements

UNICOS/mk 2.0.5 Features

- Mainframe Warmboot
- Dynamic PE Renumbering

Mainframe Warmboot

Goal

The goal was to improve the warmboot process by performing the warmboot entirely on the Cray–T3E mainframe.

Mainframe Warmboot

Overview

- **Target the PE initialization diagnostic for a specific PE**
- **Load and execute the targeted diagnostic**
- **Load mkpal**
- **Load the UNICOS/mk archive**
- **Raise reset**

Mainframe Warmboot

sgi

System Impact

- hdw_boot.uv, mkpal.cray–t3e and the UNICOS/mk archive must reside on local disk *(/dumps/current)*

- new */etc/warmboot* system administrator command

Mainframe Warmboot

Command

warmboot [–a archive] [–b bootpal] [–d dir] [–f] [–m mkpal] –l lpe [–y]

–a archive   Specifies the directory and filename of the UNICOS/mk archive.

–b bootpal   Specifies the directory and filename of the hdw_boot.uv binary file.

–d dir   Specifies the directory containing the UNICOS/mk archive, bootpal and mkpal files. The *a, b* and *m* options will override the *d* option. The default of dir is */dumps/current.*

–f   Force the warmboot without any attempts to halt the PE.

–l lpe   Identifies logical PE to be warmbooted. (Required)

–m mkpal   Specifies the directory and filename of the mkpal binary file.

–y   Answer 'y' (yes) to all prompts.

Mainframe Warmboot

Comparison

- ## SWS Warmboot
    - **Establish GRING proxy connection**
    - **Load diagnostic across proxy and execute**
    - **Load UNICOS/mk archive across proxy**
    - **Load mkpal across proxy**
    - **Load configuration parameters across proxy**
    - **Raise Reset**

```
cyclone–sws 2.0.4$ time t3epeboot –p 0x1ff
real  1m13.98s
user 0m12.25s
sys   0m8.53s
```

Mainframe Warmboot

sgi

Example

- **Cyclone (SN6302) a 544 PE System**

cyclone# time /etc/warmboot –l 0x1ff

Warmbooting LPE 0x1ff

	seconds	clocks
elapsed	6.50377	487783077
user	0.00733	549600
sys	0.74290	55717500

cyclone#

Mainframe Warmboot

sgi

Warmboot Caveats

- **Software panicked PEs**

- **Transient hardware errors**
  - transient memory errors
  - for more information on which hardware errors Warmboot is generally safe to use contact SGI customer service

- **What about hardware failed PEs?**

Dynamic PE Renumbering

Goal

The goal was to improve system MTTI by avoiding a cold boot in order to recover the application or command space after a hard PE failure.

Dynamic PE Renumbering

Overview

- **Stop the scheduling of processes on the affected PE(s)**
- **Migrate processes running on the affected PE(s)**
- **Halt the affected PE(s)**
- **Swap entries in the hardware route table stored on the R-chip (R_NET_LUT)**
- **Swap special routes (MK_SROUTES_TABLE)**
- **Update the Configuration Server and GRM and then warmboot the affected PE(s)**

Dynamic PE Renumbering

sgi

System Impact

- **Routing performance degradation**
  - **logical PEs would no longer be physical neighbors**
- **System boot files must reside on local disk**
  - **hdw_boot.uv, mkpal.cray–t3e, and the UNICOS/mk archive must reside on local disk for Mainframe Warmboot of the affected PEs**
- **One–for–one or four–for–four PE swaps**
  - **four–for–four PE swaps would be required on T3Es with a non–zero *lut_mode* (Cray–T3E's with more than 256 PEs)**
- **New *etc/renumber* system administrator command**

Dynamic PE Renumbering

sgi

Expectations

- **A renumber may require the halting of additional PEs**
- **PEs on a board with an I/O connection cannot be renumbered**
  - **This only applies to four–for–four PE swaps**
- **Processes/applications may be lost on the affected PEs**
- **After a renumber, cannot warmboot PEs from the SWS**
  - **Mainframe Warmboot must be used (*/etc/warmboot*)**
  - **Recommend the use of Mainframe Warmboot only**
- **Sites will be expected to reserve PEs for replacing failed PEs**

Dynamic PE Renumbering

sgi

Replacement PEs

- **Command PEs with no system critical daemons running on them**
  - PEs with a hard label set via */etc/grmgr* and daemon binaries with a label set via */bin/setlabel*
- **PEs which were not booted during initial boot of the mainframe**
- **How many replacement PEs should be reserved?**
  - Cray–T3E's *lut_mode* determines how many PEs must be swapped by a renumber operation
  - site's PE failure history
  - time between maintenance activities to replace failed PEs

Dynamic PE Renumbering

Command

renumber [–a archive] [–b bootpal] [–d dir] –f lpe [–m mkpal] [–n] [–p] –r lpe

–a archive   Specifies the directory and filename of the UNICOS/mk archive.

–b bootpal   Specifies the directory and filename of the hdw_boot.uv binary file.

–d dir   Specifies the directory containing the UNICOS/mk archive, bootpal and mkpal files. The *a, b* and *m* options will override the *d* option.

–f lpe   Identifies the failed LPE. (Required)

–m mkpal   Specifies the directory and filename of the mkpal binary file.

–n   After renumbering, do NOT warmboot the PEs which neighbor the failed PE. This only applies to Cray–T3E's running with a non–zero *lut_mode*.

–p   List the processes that would be affected by the renumbering of the specified PEs. The actual renumber is not performed.

–r lpe   Identifies the replacement LPE. (Required)

Dynamic PE Renumbering

sgi

Example

- Hard PE failure identified
- Administrator selects PE to be swapped for the failed PE
- Administrator executes the renumber command to swap PEs
- System runs with routing performance degradation
- At the next cold boot, physical PE renumbering can be done via *t3ems* on the SWS

T3E Resiliency Enhancements

Conclusion

Mainframe Warmboot and Dynamic PE Renumbering are a continuation of efforts in establishing UNICOS/mk as the leader in overall system resiliency.

Mainframe Warmboot Dynamic PE Renumbering

sgi

More Information

- **UNICOS/mk General Administration Guide, 004–2601–002**
- ***warmboot* (8) man page**
- ***renumber* (8) man page**