The NonStop Advanced Architecture Value Proposition:
HP Integrity NonStop Server Availability Compared to NonStop S-series Server

Bob Kossler
Manager, Product Management
HP
September, 2005
Itanium® processor and the NonStop Advanced Architecture

4-processor system

Upgrade to triplex

* = Logical synchronization units
S-series Processor Block Diagram

- Cache
- MPU
- Memory
  - Memory controller
  - Lockstep checking
    - and
    - ServerNet® Transfer Engine

- To ServerNet X-fabric
- To ServerNet Y-fabric
NonStop Advanced Server

- **Hardware fault-tolerance** – better than S-series
  - recovery from many double faults
  - three NonStop blade complex
  - option for extra ServerNet connections

- **Software fault-tolerance** same as S-series

* = logical synchronization units
## NSAA Duplex and S-Series comparison

### Fault/Event

<table>
<thead>
<tr>
<th>UCME</th>
<th>✓ ✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor failure - transient</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Processor failure - permanent</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Software take over after processor failure</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>LSU failure</td>
<td>✓</td>
</tr>
<tr>
<td>Data integrity</td>
<td>✓</td>
</tr>
<tr>
<td>Data integrity during processor replacement or NSBC failure</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Software fault tolerance</td>
<td>✓</td>
</tr>
<tr>
<td>ServerNet fabric availability</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>I/O adapter availability</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Tolerance for multiple failures</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Load balancing after processor failure</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Fans failure tolerance</td>
<td>✓</td>
</tr>
<tr>
<td>Bulk power supply</td>
<td>✓ ✓</td>
</tr>
</tbody>
</table>

### Fault/Event Results

- NSAA duplex better than S-Series
- NSAA duplex same as S-Series

### Additional Notes
- Commercial data integrity
NSAA Triplex and S-Series comparison

Fault/Event

UCME
Processor failure - transient
Processor failure - permanent
Software take over after processor failure
LSU failure
Data integrity
Data integrity during processor replacement or NSBC failure
Software fault tolerance
ServerNet fabric availability
I/O adapter availability
Tolerance for multiple failures
Load balancing after processor failure
Fans failure tolerance
Bulk power supply

✓ ✓ NSAA duplex better than S-Series
✓ ✓ NSAA duplex same as S-Series
S-series processor fault scenarios and recovery

• Most faults, whether transient or permanent, result in hardware error freeezes (HEFs) and loss of a logical processor that is visible at the application level
  − single microprocessor (e.g., R16K) fault
  − fault in the memory controller or lockstep compare/ServerNet transfer engine
  − memory UCMEs
  − uncorrectable cache errors

• Recovery is done by replacing the PMF if necessary, and then reloading its logical processor
NonStop Advanced Server

- This system's architecture provides many opportunities to
  - Take advantage of hardware redundancy to mask hardware faults
  - Apply recovery techniques and avoid HEFs
- Its bias is to attempt recovery unless the fault is known to be permanent
- In almost all cases, even permanent faults do not result in the loss of a logical processor and resulting application impact
Processor fault recovery

• For most faults, perform a reintegration operation
  – If reintegration fails, delay and try again
  – If reintegration continues to fail, increase the delay between retries
  – Count failures, and dial out for repair if recurrent failures cross defined thresholds
    • Recurrences within a short time interval
    • Lifetime
Reintegration steps

- Isolate affected Processing Elements (microprocessors)
- Copy memory from a “good”, running slice to the target slice
  - Copy operation is done online, similar to a disk REVIVE
  - Impact on the good slice, and hence on applications is minimal
  - Copy time is on the order of 10 minutes (faster on a relatively-idle system)
  - Because of the slice architecture, reintegration must be done for pairs of PEs rather than single PEs
- Restart newly-reintegrated PEs
Processor recovery: if slice repair is required

- Replace the slice
- Reintegrate (4 PEs)

- Reintegration takes place in two steps, with two PEs being reintegrated in each step
  - Reintegration completes for first two PEs within 10 minutes
  - Reintegration completes for other two PEs within 20 minutes
Recovery automation

• All hardware fault recovery actions are carried out automatically by the system, with no operator intervention required

• Actions include:
  − Reintegration
  − Processor reload (if using TFDS)
  − Reintegration link recovery
  − ServerNet fabric recovery
  − PE-to-LSU link recovery
  − LSU recovery
Nonstop Advanced Server – Triplex
Triplex processor fault scenarios

- All single faults (except LSU) are invisible at the application level
  - Single microprocessor (e.g., Itanium-2) fault
  - Memory I/O Bridge fault
  - Memory UCME
  - Uncorrectable cache error
  - Reintegration link fault
  - PE-to-LSU link fault
  - Power/fan fault

- Almost all double faults (again except LSU) are also invisible

- All repairs other than a single LSU can safely be deferred
Triplex processor fault scenarios: PE fault

- All PE faults can be unambiguously identified
  - Many PE faults are detected through self-identifying errors
  - Most other PE faults are detected by timeouts
  - If the faulty PE continues to run, it will eventually be detected by a voting miscompare

- In some cases
  - The offending PE is voted out
  - The logical processor continues to run (in duplex mode)

- In other cases
  - The slice fails
  - All of its logical processors continue to run (in duplex mode)

- In all cases:
  - The failure is invisible to applications
  - Full checking and fault tolerance continue to be in effect

- Recovery:
  - Reintegrate (either 2 PEs or 4 PEs)
Triplex processor fault scenarios: memory UCME

- Detected by memory ECC, either in normal operation or by the memory sniffer process
- The logical processor continues to run
  - Full checking and fault tolerance continue to be in effect
- The fault is invisible to applications
- Recovery (if detected in normal operation):
  - Reintegrate (2 PEs)
- Recovery (if detected by memory sniffer):
  - Rewrite the affected word
  - If unsuccessful, try to take the page out of service
  - If page can’t be taken out of service cleanly, reintegrate (2 PEs)
Triplex processor fault scenarios: uncorrectable cache error

• Detected by cache ECC
• The logical processor continues to run (in duplex mode)
  – Full checking and fault tolerance continue to be in effect
• The fault is invisible to applications
• Recovery:
  – Reintegrate (2 PEs)
Triplex processor fault scenarios: reintegation link fault

- Reintegration links are continually checked whether reintegration is in progress or not
  - Transient faults during non-reintegration periods are simply counted and dialed out if a threshold is exceeded
  - Transient faults during reintegration cause reintegration to fail and be retried

- The logical processor continues to run (in triplex mode)
  - Full checking and fault tolerance continue to be in effect

- The fault is invisible to applications

- Permanent faults cause dialout, but the system remains fully operational during link down and repair
  - Reintegration operations cannot be performed

- The link is recovered automatically once it has been replaced
Triplex processor fault scenarios: memory I/O bridge fault

- All memory I/O bridge faults are unambiguously detected
  - Most memory I/O bridge faults are detected by timeouts
  - In some cases, the I/O bridge itself will detect the error and notify the appropriate PE(s)
  - All other faults are detected by voting mismatches

- In all cases:
  - The PE(s) is/are voted out
  - Their logical processors continue to run (in duplex mode)
  - Full checking and fault tolerance continue to be in effect
  - The fault is invisible to applications

- Recovery (automatic):
  - Reintegrate (2 or 4 PEs)
Triplex processor fault scenarios: PE-to-LSU link fault

- Transient faults are handled transparently by the link retry logic
- Permanent faults are handled the same as a PE fault
  - Repair is required prior to reintegration
Triplex processor recovery: voting mismatch in the LSU

• For any single fault, it will be easy to identify the offending PE as there will be a two-against-one vote

• There are no known causes of three-way disagreement
  – If one should ever occur, it would be detected and treated as a software error and the logical processor will be halted

• Recovery:
  – Normal case: reintegrate (2 PEs)
  – Three-way disagreement: reload the logical processor
Triplex processor recovery: 
LSU fault

• Most faults are detected by the LSU, as almost all of its components are self-checked
• Its associated PEs learn about the fault either through an interrupt to the PEs or through timeouts
• Recovery:
  – Reload the corresponding logical processor
  – If faults continue to occur, leave the logical processor down and dial out for repair
Nonstop Advanced Server – Duplex

memory reintegration link

CPU 0
CPU 1
CPU 2
CPU 3
memory

slice A
slice B
LSU 0
LSU 1
LSU 2
LSU 3
ServerNet
ServerNet
ServerNet
ServerNet
Duplex processor fault scenarios: summary

• Most faults do not cause a logical processor failure and are invisible at the application level
  – Single microprocessor (e.g., Itanium-2) fault
  – Memory I/O Bridge fault (most cases)
  – Memory UCME
  – Uncorrectable cache error
  – Reintegration link fault
  – PE-to-LSU link fault
  – Power/fan fault

• A voting mismatch not accompanied by a detectable error causes loss of a logical processor

• An LSU fault may cause loss of a logical processor
Duplex processor fault scenarios: PE fault (1 of 2)

- Many PE faults are detected through self-identifying errors

- In some of these cases:
  - The offending PE is voted out
  - The logical processor continues to run (in simplex mode)
  - Full checking and fault tolerance continue to be in effect for the other logical processors associated with that slice

- In other cases:
  - The slice fails
  - All of its logical processors continue to run (in simplex mode)

- In all of the above cases, the failure is invisible to applications

- Recovery:
  - Reintegrate (either 2 PEs or 4 PEs)
Duplex processor fault scenarios: PE fault (2 of 2)

- Remaining fault modes can cause loss of its logical processor if the faulty PE continues to run without obvious error
  - Detection is done through a voting mismatch
- Recovery:
  - Most cases: reintegrate (2 PEs)
  - Remaining cases: reload logical processor
Duplex processor fault scenarios: memory UCME

- Detected by memory ECC, either in normal operation or by the memory sniffer process
- The logical processor continues to run
- The fault is invisible to applications
- Recovery (if detected in normal operation):
  - Reintegrate (2 PEs)
- Recovery (if detected by memory sniffer):
  - Rewrite the affected word
  - If unsuccessful, try to take the page out of service
  - If page can’t be taken out of service cleanly, reintegrate (2 PEs)
Duplex processor fault scenarios: uncorrectable cache error

- Detected by cache ECC
- The logical processor continues to run (in simplex mode)
- The fault is invisible to applications
- Recovery:
  - Reintegrate (2 PEs)
Duplex processor fault scenarios: reintegration link fault

- Reintegration links are continually checked whether reintegration is in progress or not
  - Transient faults during non-reintegration periods are simply counted and dialed out if a threshold is exceeded
  - Transient faults during reintegration cause reintegration to fail and be retried

- The logical processor continues to run (in triplex mode)
  - Full checking and fault tolerance continue to be in effect

- The fault is invisible to applications

- Permanent faults cause dialout, but the system remains fully operational during link down and repair
  - Reintegration operations cannot be performed

- The link is recovered automatically once it has been replaced
Duplex processor fault scenarios: memory I/O bridge fault

- Most memory I/O bridge faults are detected by timeouts
  - In some cases, the I/O bridge itself will detect the error and notify the appropriate PE
  - The PE(s) is/are voted out
  - Logical processors continue to run (in simplex mode)
  - The fault is invisible to applications

- Recovery:
  - Most cases: reintegrate (4 PEs)
Duplex processor fault scenarios: PE-to-LSU link fault

- Transient faults are handled transparently by the link retry logic
- Permanent faults are handled the same as a PE fault
  - Repair is required prior to reintegration
Duplex processor recovery: voting mismatch in the LSU

- If a voting mismatch occurs and one of the PEs does not respond in a timely manner, it is assumed to be the faulty component.
- If a voting mismatch occurs while one of the PEs is on probation and the offending PE cannot be identified:
  - The PE on probation is assumed to be the faulty component.
  - Recovery is the same as for a PE fault.
Duplex processor recovery: voting mismatch in the LSU

• If there is no error indication on either slice and neither PE is on probation, it is not possible to determine which of the two PEs has sent faulty data
• The logical processor is halted
• Recovery:
  − Reload the logical processor
  − If voting mismatches continue to occur, replace one slice, reintegrate, and reload the logical processor
  − If they still continue to occur, replace the other slice, reintegrate, and reload the logical processor
Duplex processor recovery: LSU fault

- The LSU is self-checked and identifies its own faults
- Its associated PEs learn about the fault either through an interrupt to the PEs or through timeouts

Recovery:
- Reload the LSU’s logical processor
- If faults continue to occur, leave the logical processor down and dial out for repair
## Comparison of Fundamental Features

<table>
<thead>
<tr>
<th></th>
<th>S-series Design</th>
<th>Triplex</th>
<th>Duplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Integrity</td>
<td>Complete Checking and Data protection</td>
<td>Complete Checking and Data protection</td>
<td>Complete Checking and Data protection except during repair</td>
</tr>
<tr>
<td>Hardware Single Fault</td>
<td>No Single Points of Failure</td>
<td>No Single Points of Failure</td>
<td>One known class of (very rare) failures can cause system outage.</td>
</tr>
<tr>
<td>Tolerance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardware Multiple Fault</td>
<td>Not Designed into architecture but</td>
<td>Designed into architecture. Many</td>
<td>Microprocessor and memory covered for at least two faults. Other components may not be.</td>
</tr>
<tr>
<td>Tolerance</td>
<td>certain cases can be survived</td>
<td>double faults transparent.</td>
<td></td>
</tr>
<tr>
<td>Memory Faults</td>
<td>UCME causes CPU Freeze</td>
<td>UCMEs are transparent and repairs can be deferred.</td>
<td>UCMEs are transparent and repairs can be deferred.</td>
</tr>
</tbody>
</table>
## Comparison of Fundamental Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>S-series Design</th>
<th>Triplex</th>
<th>Duplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repairs</td>
<td>Safe to defer repair for 4 hours</td>
<td>Safe to defer repair for days. Still have two checked processors.</td>
<td>Same as S-series but slightly greater exposure to silent data integrity errors.</td>
</tr>
<tr>
<td>System Availability</td>
<td></td>
<td>Significant reduction in unplanned outages due to software halts triggered by Hardware Faults.</td>
<td>Comparable to S-series.</td>
</tr>
<tr>
<td>Application Availability</td>
<td>Single processor halt may cause outages for non-FT applications.</td>
<td>Fewer single processor halts due to hardware results in fewer non-FT application outages</td>
<td>Fewer single processor halts due to hardware results in fewer non-FT application outages</td>
</tr>
<tr>
<td>Future Processor Upgrades</td>
<td>Online</td>
<td>Must be done offline, at least initially. Increases planned downtime by a small factor.</td>
<td>Must be done offline, at least initially. Increases planned downtime by a small factor.</td>
</tr>
<tr>
<td>Software Fault Tolerance</td>
<td>No Changes.</td>
<td></td>
<td>No changes.</td>
</tr>
</tbody>
</table>
Conclusion

• Duplex configurations provide comparable processor availability to S-series systems
• Triplex configurations provide much better processor availability than S-series systems
• Both duplex and triplex provide full checking and data integrity
  – Exception: duplex during slice repair and some phases of recovery