OpenVMS V8.4-1H1 performance

Performance improvements on HP Integrity bl8x0c–i4 and rx2800-i4 servers

Colin Butcher
XDelta Limited

www.xdelta.co.uk
+44 117 904 8209
XDelta – who we are

- Independent consultants since 1996:
  - UK based with international reach
  - Over 30 years experience with OpenVMS
- We design and implement solutions:
  - Mission critical systems
  - Cross-sector experience
  - Engineering background
- Gartner (2009):
  - Identified XDelta as one of few companies world-wide capable of OpenVMS migration related projects
HP Integrity -i4 servers - highlights

- “Poulson” 2.53GHz 8 core processor with shared L3 cache
- Around 30% per core greater throughput
- Reduced NUMA effects for same core count
- Better memory latency and bandwidth
- Improved floating point and integer performance
- bl870c-i4 (32 cores) about 1.3x better than bl890c-i2
HP Integrity -i4 servers – hardware

- bl860c-i4:
  single width, 16 cores, 384GB, 4x 10GigE, 3x mezz, 1c2d SAS

- bl870c-i4:
  double width, 32 cores, 768GB, 8x 10GigE, 6x mezz, 2c4d SAS

- bl890c-i4:
  quad width, 64 cores, 1.5TB, 16x 10GigE, 12x mezz, 4c8d SAS
  - OpenVMS currently supports a maximum of 32 cores
  - OpenVMS also supports nPARs

- rx2800-i4:
  2U rack, 16 cores, 384GB, 4x 1GigE, 6x PCIe, 1c8d SAS
Migrating from Alpha to Integrity

- Multi-core processors, NUMA, hyperthreading
- 10GigE network
- 8GigFC SAN
- Blade chassis connectivity for bl8x0c-i4
- EVA to 3Par storage migration
- bl870c-i4 and bl890c-i4: good for GS1280 migration
Migration to Integrity and beyond

Platform type / scale

Urgency

Small IA64 (eg: rx2660, rx2800-i2)

Mid-range IA64 (eg: BL870c-i2)

Large IA64 (eg: BL890c-i2)

Small VAX (eg: microVAX)

Mid-range VAX (eg: VAX 4000)

Large VAX (eg: VAX 7000)

Small Alpha (eg: DS25)

Mid-range Alpha (eg: ES45)

Large Alpha (eg: GS1280)

HP VM
Server hardware differences (-i2 to –i4)

- Higher clock rate
- Double the core count (8 cores)
- Greater memory capacity
- Reduced memory latency
- Shared on-chip cache
- 10GigE LoM (LAN on Motherboard) – LAN only, not FCoE

- Still use 8GigFC mezzanine cards
Chassis hardware – c7000 / c3000

- Virtual Connect (GigE, 1/10GigE, 8GigFC)
- Flex10
- LAN side of FlexFabric

- 10GigE chassis based switching
- 10GigE passthrough
- 1GigE passthrough

- 8GigFC chassis switching
- 4GigFC passthrough
Infrastructure hardware

- 3Par storage arrays at 8GigFC
- SSD devices for local storage and 3Par storage arrays
- 8GigFC SAN – HP / Brocade switches
- 10GigE networking – HP Procurve, Cisco
OpenVMS V8.4-1H1 on –i4 servers

• Complete build of base system from sources

• -i4 hardware support (32 cores supported, threads off)

• New LoM driver

• VSI branding
CPU architecture - Intel 9500 – “Poulson”
System architecture – rx2800-i4
Blade architecture – bl8x0c-i4
QPI fabric – bl870c-i4 and bl890c-i4
High core count

• CPU 00 is the primary CPU – try to reduce its workload

• Fastpath CPU selection – be aware of physical layout

• CPU choice for dedicated lock manager

• CPU choice for TCPIP packet processing engine

• Consider physical layout - RADs and NUMA
Hyperthreading

- Hyperthreading is extremely workload dependent
- In general the OpenVMS scheduler does a better job
- Enable / disable hyperthreads and reboot
- “CPU” count will appear to double when enabled
  Note: OpenVMS currently supports a maximum of 32 “CPUs”
Memory architecture – bl890c-i4
NUMA (non-uniform memory access)

- OpenVMS uses large shared memory regions:
  - XFC (50% available memory by default)
  - RMS global buffers
  - Global sections (especially database caches)
  - Memory disc driver (MD devices)

- Useful starting point for OpenVMS is “mostly UMA”
Preliminary Performance Results i2 vs. i4

- The following slides contain preliminary data on performance differences between selected i2 and i4 servers running OpenVMS E8.4-1H1.

- The data was generated from VSI-written programs used to measure certain aspects of system performance.
- The results shown here should not be used as a general characterization of overall system performance or as an indication of how any specific application may perform.
i2 vs. i4 Memory Bandwidth

![i2 vs. i4 Memory Bandwidth Chart](chart.png)
i2 vs. i4 Memory Latency

- BL860c i2 (1.47ghz): 237 ns
- BL860c i4 (2.4ghz): 196 ns
- RX2800 i2 (1.73ghz): 243 ns
- RX2800 i4 (2.67ghz): 108 ns
i2 vs. i4 Floating Point Performance

<table>
<thead>
<tr>
<th>CPU</th>
<th>Frequency</th>
<th>Relative Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL860c i2</td>
<td>1.47ghz</td>
<td>1.00</td>
</tr>
<tr>
<td>BL860c i4</td>
<td>2.4ghz</td>
<td>1.33</td>
</tr>
<tr>
<td>RX2800 i2</td>
<td>1.73ghz</td>
<td>1.14</td>
</tr>
<tr>
<td>RX2800 i4</td>
<td>2.67ghz</td>
<td>1.57</td>
</tr>
</tbody>
</table>
i2 vs. i4 Integer Performance

<table>
<thead>
<tr>
<th>Platform</th>
<th>Model</th>
<th>Frequency</th>
<th>Relative Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL860c i2</td>
<td></td>
<td>1.47ghz</td>
<td>1.00</td>
</tr>
<tr>
<td>BL860c i4</td>
<td></td>
<td>2.4ghz</td>
<td>1.26</td>
</tr>
<tr>
<td>RX2800 i2</td>
<td></td>
<td>1.73ghz</td>
<td>1.17</td>
</tr>
<tr>
<td>RX2800 i4</td>
<td></td>
<td>2.67ghz</td>
<td>1.49</td>
</tr>
</tbody>
</table>
Performance engineering – use T4

• Avoid guesswork - run T4 all the time

• Without good data you cannot do good performance work

• A faster machine just waits more quickly
• Don’t make it go faster, stop it going slower
• The fastest IO is the IO you don’t do
• The fastest code is the code you don’t execute
Summary - VMS V8.4-1H1 on –i4 servers

- Disable devices you don’t use
  SYSMAN IO SET EXCLUDE=(EWC,EWD,…)

- Experiment with memory interleave setting
- Use memory reservations

- Fastpath settings for device types
- Dedicated CPU for TCPIP + LCKMGR

- Experiment with hyperthreading
OpenVMS Rolling Roadmap

**OpenVMS V8.4-1H1**
*Architecture: Itanium*

**Itanium® Processor 9500 series**

**HP Integrity System Support**
- rx2800 i4
- HP Integrity Server Blades
  - BL860c i4
  - BL870c i4
  - BL890c i4
- i2 versions of the above
- Blades FlexFabric LAN support

**Software**
- Improved performance over i2
- Availability Manager – update to 64-bit desktop

**OpenVMS V9.0**
*Architecture: Common*

**New File System**
- Eliminate 2TB volume size limit
- Improved performance

**Architecture: Integrity**
Additional servers & I/O, depending on feedback

**Architecture: X86-64**
- Selected HP servers
- OpenVMS as a VM guest
- Binary Translator
- Updated Language Standards
  - C
  - C++
  - FORTRAN

Q2 2015

**Architecture: Itanium**

**HP Integrity system support**
- V8.4-1H1 supported servers and more - such as rx2660, rx3600, rx6600, ...
- More network and storage devices
- Kittson-based systems (when available)

**Software**
- Improved performance, reliability
- New TCP/IP stack
- Support 64 cores (threads off)
- Enhanced GNV/CRTL for open source porting and development
- JAVA 1.8
- Updated Open Source Kits
  - Apache, gSOAP, Samba
  - SSL, Kerberos
  - and more

Q4 2015

Q3 2016

Q3 2017

Two releases are planned between V8.4-1H1 and V9.0. The order in which work makes it into these releases will be determined by readiness, HW availability, and customer feedback.
OpenVMS V8.4-1H1 performance

Performance improvements on HP Integrity bl8x0c–i4 and rx2800-i4 servers

Colin Butcher
XDelta Limited

www.xdelta.co.uk
+44 117 904 8209