SAP HANA on IBM Power Systems

Alfred Freudenberger
North America SAP on Power Systems
Sales Executive
afreude@us.ibm.com
512-659-8059

Blog: SAPonPower.wordpress.com
IBM & SAP HANA

Production Node 3
512GB

Production Node 1
512GB

Shared filesystem - GPFS
HDD
Flash
data01
log01
DB partition 1
- index server - statistic server - SAP HANA
®
- SAP HANA
®
DB
worker node
data01(1)

DB partition 2
HDD
Flash
data02
log02
DB partition 3
Flash
- index server - statistic server
- SAP HANA
®
DB
worker node
data03

Hot Standby
512GB
HDD
Flash
- index server - statistic server
- SAP HANA
®
DB
standby node
HDD replica
primary data
SAP HANA
®
data01(2)
data01(3)
data02(1)
data02(2)
data02(3)
data03(1)
data03(2)
data03(3)
log01(1)
log01(2)
log01(3)
log02(1)
log02(2)
log02(3)
log03(1)
log03(2)
log03(3)

Production Node 2
512GB

- index server - statistic server
- SAP HANA
®
DB
worker node
data02
log02
data01(1)
log01(1)
log03(3)

DB partition 2
HANA
HANA
Node1
Node2
Node3
Node4
NodeN
GPFS
GPFS
GPFS
GPFS
GPFS
IMDB
IMDB
IMDB
IMDB
IMDB
App1
App10
App10
App10
App10
4
5
6
8
9
0

ISL: 2x 10GigE
ISL: 2x 10GigE

SAP HANA on Power
SAP HANA Deployment Options for Different Use Cases

- **Technology Platform**: Customer specific development.
- **Content**: Reports that are re-built on HANA using a BI frontend.
- **Accelerated Apps**: Speed up existing business suite functionality/applications through a side-by-side scenario with HANA.
- **Products on In-Memory**: Existing applications and strategic investment areas that may be disruptive.

Examples:
- e.g. Demographics & Google Maps
- e.g. SAP ERP: Operational reporting
- e.g. SAP ERP: CO-PA
- e.g. SAP BW -7.3 SP5-on HANA, Business Suite
SAP HANA on IBM POWER

- SAP HANA on Power is targeting enterprise customers requiring an SAP HANA-based solution on IBM Power Systems servers.

- IBM intention is not to offer it as an appliance, but in a flexible form combining the HANA license from SAP and IBM Power Systems servers, middleware and services.
SAP HANA on IBM Power – Expected Customer Value

Intended for mission critical 7X24 Enterprise customer operations
- Not an Appliance, running on traditional POWER8 and POWER7+ servers
- Best Reliable, Available, Serviceable (RAS) in the market
- On-Demand Capacity
- Can be integrated into and tailored to a customer’s environment

Throughput per core
- Significant Power SMT throughput advantages versus Intel x86
- Value: fewer cores, fewer footprints, and lower operating costs

Virtualization out of the box
- Lower virtualization layer overhead on multi-threaded HANA workloads
- Value: fewer cores, fewer footprints, and lower operating costs

Price performance
Intelligence Moved into Memory
- Scheduling logic, caching structures
- Energy Mgmt, RAS decision point
  - Formerly on Processor
  - Moved to Memory Buffer

Processor Interface
- 9.6 GB/s high speed interface
- More robust RAS
- “On-the-fly” lane isolation/repair

Performance Value
- End-to-end fastpath and data retry (latency)
- Cache → latency/bandwidth, partial updates
- Cache → write scheduling, prefetch, energy
## Reliability Features

<table>
<thead>
<tr>
<th>RAS Feature</th>
<th>POWER7, POWER8</th>
<th>x86</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Application/Partition RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Live Partition Mobility</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Live Application Mobility</td>
<td>Yes</td>
<td>Yes, support issues</td>
</tr>
<tr>
<td>Partition Availability priority</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>System RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS independent First Failure Data Capture</td>
<td>Yes</td>
<td>EX – MCA Recovery</td>
</tr>
<tr>
<td>Memory Keys (including OS exploitation)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Processor RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processor Instruction Retry</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Alternate Processor Recovery</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dynamic Processor Deallocation</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dynamic Processor Sparing</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Memory RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chipkill™</td>
<td>Yes</td>
<td>Yes, some vendors</td>
</tr>
<tr>
<td>Survives Double Memory Failures</td>
<td>Yes</td>
<td>Yes, optional</td>
</tr>
<tr>
<td>Selective Memory Mirroring</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Redundant Memory</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>I/O RAS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extended Error Handling</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>I/O Adapter Isolation (PI-Bus and TCEs)</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

See the following URLs for addition details:
- [http://www-03.ibm.com/systems/migratetobm/systems/power/virtualization.html](http://www-03.ibm.com/systems/migratetobm/systems/power/virtualization.html)
Advanced RAS: MCA Recovery

- Allows recovery from otherwise fatal system errors

IvyBridge-EX recovery is dependent on Windows, Linux or Hypervisor to handle the problem
POC Results: Effects of simultaneous (hardware) multi-threading
## Comparisons to Intel Systems

<table>
<thead>
<tr>
<th></th>
<th>Intel Sandy Bridge EP E5-26xx</th>
<th>Intel Ivy Bridge EP E5-26xx v2</th>
<th>Intel Ivy Bridge EX E7-88xx v2</th>
<th>POWER7+</th>
<th>POWER8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock rate range</td>
<td>1.8 - 3.6 GHz</td>
<td>1.7 - 3.7 GHz</td>
<td>1.9 - 3.4 GHz</td>
<td>3.1 - 4.4 GHz</td>
<td>3.0 - 4.1 GHz</td>
</tr>
<tr>
<td>SMT options</td>
<td>1, 2*</td>
<td>1, 2*</td>
<td>1, 2*</td>
<td>1, 2, 4</td>
<td>1, 2, 4, 8</td>
</tr>
<tr>
<td>HW threads/socket</td>
<td>16</td>
<td>24</td>
<td>30</td>
<td>32</td>
<td>96</td>
</tr>
<tr>
<td>L1 cache/core</td>
<td>32 KB</td>
<td>32 KB*</td>
<td>32 KB*</td>
<td>32 KB</td>
<td>64 KB</td>
</tr>
<tr>
<td>L2 cache/core</td>
<td>256 KB</td>
<td>256 KB</td>
<td>256 KB</td>
<td>256 KB</td>
<td>512 KB</td>
</tr>
<tr>
<td>L3 cache/core</td>
<td>2.5 MB</td>
<td>2.5 MB</td>
<td>2.5 MB</td>
<td>10 MB</td>
<td>8 MB</td>
</tr>
<tr>
<td>L4 cache/socket</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>128 MB</td>
</tr>
<tr>
<td>Memory Bandwidth</td>
<td>31.4 - 51.2 GB/s</td>
<td>42.6 - 59.7 GB/s</td>
<td>68 - 85** GB/s</td>
<td>100 - 180 GB/sec</td>
<td>230 - 410 GB/sec</td>
</tr>
</tbody>
</table>

* Intel calls this Hyper-Threading Technology (No HT and with HT)
*32KB running in “Non-RAS mode” Only 16KB in RAS mode
**85GB running in “Non-RAS mode” = dual-device error NOT supported
## Performance Comparison – POWER8 vs. x86 E5
IBM POWER8 core and system performance is 2x the x86 Xeon E5-2697 v2 core performance

- Published Benchmarks –
- ALL data is PUBLISHED or SUBMITTED (SPECjbb2013 on the POWER824)

### Performance Comparison Table

<table>
<thead>
<tr>
<th></th>
<th>x86 “Ivy Bridge”</th>
<th>IBM POWER S824</th>
<th>POWER8 vs. x86 Core Performance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intel Xeon E5-2697 v2 (except where noted)</td>
<td>POWER8 @ 3.5 GHz</td>
<td></td>
</tr>
<tr>
<td># Cores</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>SAP 2-Tier</td>
<td>10253</td>
<td>21212</td>
<td>2.1</td>
</tr>
<tr>
<td>SPECint_rate2006</td>
<td>1020</td>
<td>1750</td>
<td>1.7</td>
</tr>
<tr>
<td>SPECfp_rate2006</td>
<td>734</td>
<td>1370</td>
<td>1.9</td>
</tr>
<tr>
<td>SPECjbb2013 (max-jOPS)</td>
<td>63079</td>
<td>167958</td>
<td>2.7</td>
</tr>
<tr>
<td>SPECjEnterprise2010</td>
<td>11260</td>
<td>22543</td>
<td>2.0</td>
</tr>
<tr>
<td>Oracle eBS 12.1.3 Payroll</td>
<td>1017639</td>
<td>1090909 (12-core)</td>
<td>2.1</td>
</tr>
<tr>
<td>Siebel CRM Release 8.1.1.4</td>
<td>10000 (16-core E5-2690)</td>
<td>50000 (6-core)</td>
<td>13.3</td>
</tr>
</tbody>
</table>

1) SAP results are based on the two-tier SAP SD standard application benchmark running SAP enhancement package 5 for the SAP ERP 6.0 application. Results valid as of June 2, 2014. Source: [http://www.sap.com/benchmark](http://www.sap.com/benchmark)
2) SPECcpu2006 results are submitted as of 4/22/2014. For more information go to [http://www.specbench.org/cpu2006/results/](http://www.specbench.org/cpu2006/results/)
4) SPECjEnterprise2010 results are valid as of 4/22/2014. For more information go to [http://www.specbench.org/jEnterprise2010/results/](http://www.specbench.org/jEnterprise2010/results/)
**HANA Appliance**

**Fast Implementation Support fully provided by SAP**

- SAP HANA Server
- Application
- Database
- Operation System
- Virtualization
- Server
- Network
- Storage

**Appliance delivery approach**
- Solution validation done by SAP and partner
- Preconfigured hardware set-up
- Preinstalled software

**TDI – Tailored Datacenter Integration**

**More Flexibility Save IT budget and existing investment**

- SAP HANA Server
- Virtualization
- Server
- Network
- Storage
- Enterprise Storage

**SAP HANA tailored data center integration**
- Installation needs to be done by customer
- Customer aligns with the hardware partner on individual support mode
Conceptual solution with HANA on Power

- SAP HANA®
- High Availability
  - SAP HANA System Replication, SUSE HA Ext, RH HA Plugin, Symantec HA, Tivoli SA
- File System
  - GPFS®, XFS
- OS
  - Linux Enterprise Server
  - Priority Support for SAP applications
- Server Hardware
  - Any POWER7+ or POWER8
  - Standalone or shared/PVM
- Storage Hardware
  - Customer choice

Power Landscape example

- ECC
- BW HA
- ECC App
- CRM
- BW QA
- CRM QA
- BW App

Additional notes:
1) Not currently supported with Linux on Power. Contact RedHat for product plans.
2) GPFS is currently supported for Linux on Power by IBM. Contact SAP for plans for certification with HANA on Power.
Memory and I/O Intensive Operations

- All columns at initialization/HA
- Selected columns at initialization/HA
- Others on demand

Load time dependent on throughput of IOPS

Response time of on-demand columns dependent on latency/throughput of IOPS

Selected columns at initialization/HA
Others on demand

Delta Merge Process

Extremely high memory demand

I/O latency & throughput dependency

I/O latency dependency

SAP HANA database

Savepoint
COMMIT

Data Backups
Log Backups

Data Area (disk)
Log Area (disk)

© 2014 IBM Corporation
HWCCT – Hardware configuration check tool

Determine if system meets KPI requirements

- Landscape test
  - OS config validity
  - Consistency of landscape based on reference architecture
- File system throughput/latency
- Network throughput for multinode configurations
  - 9.5 GBits for single stream
  - 9.0 GBits for duplex stream

<table>
<thead>
<tr>
<th>Volume</th>
<th>Block Sizes</th>
<th>Test File Size *</th>
<th>KPIs</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Initial Write (MB/s)</td>
<td>Overwrite (MB/s)</td>
<td>Read (MB/s)</td>
<td>Latency (μs)</td>
<td></td>
</tr>
<tr>
<td>Log</td>
<td>4K</td>
<td>5G</td>
<td>n.a.</td>
<td>30</td>
<td>n.a.</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16K</td>
<td>16G</td>
<td>n.a.</td>
<td>120</td>
<td>n.a.</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1M</td>
<td>16G</td>
<td>n.a.</td>
<td>250</td>
<td>500</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>Data</td>
<td>4K</td>
<td>5G</td>
<td>10</td>
<td>20</td>
<td>80</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16K</td>
<td>16G</td>
<td>40</td>
<td>100</td>
<td>200</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64K</td>
<td>16G</td>
<td>100</td>
<td>150</td>
<td>250</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1M</td>
<td>16G</td>
<td>150</td>
<td>200</td>
<td>300</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16M</td>
<td>16G</td>
<td>200</td>
<td>250</td>
<td>400</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>64M</td>
<td>16G</td>
<td>200</td>
<td>250</td>
<td>400</td>
<td>n.a.</td>
<td></td>
</tr>
</tbody>
</table>
Extreme Performance - enable business to unleash the power of performance, scale, and insight to drive services and products market faster

MicroLatency™ - facilitate the data path that will accelerate critical applications and achieve a true market based competitive advantage

Macro Efficiency - driven by consolidation of hardware and software, deployment speed, efficient use of IT staff as well as power and cooling savings

Enterprise Reliability - durable and reliable designs that use enterprise class flash and patented data protection technology
### IBM FlashSystem 840 performance

<table>
<thead>
<tr>
<th>Performance criteria¹²</th>
<th>Maximum capacity (12 flash modules)</th>
<th>Middle capacity (8 flash modules)</th>
<th>Minimum capacity (4 flash modules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% Read IOPS</td>
<td>1.1 M</td>
<td>1.1 M</td>
<td>1.0 M</td>
</tr>
<tr>
<td>100% Write IOPS</td>
<td>600 K</td>
<td>400 K</td>
<td>225 K</td>
</tr>
<tr>
<td>70/30 IOPS</td>
<td>750 K</td>
<td>500 K</td>
<td>225 K</td>
</tr>
<tr>
<td>100% large block sequential read</td>
<td>8 GBps</td>
<td>8 GBps</td>
<td>4 GBps</td>
</tr>
<tr>
<td>100% large block sequential write</td>
<td>4 GBps</td>
<td>2.5 GBps</td>
<td>1 GBps</td>
</tr>
<tr>
<td>Read latency</td>
<td>135 µs</td>
<td>135 µs</td>
<td>135 µs</td>
</tr>
<tr>
<td>Write latency</td>
<td>90 µs</td>
<td>90 µs</td>
<td>90 µs</td>
</tr>
</tbody>
</table>

¹ Includes all features and performance enhancements
² Capacity is measured in GB for large block sequential reads and writes
Coherent Accelerator Processor Interface (CAPI) Overview

Typical I/O Model Flow

DD Call → Copy or Pin Source Data → MMIO Notify Accelerator → Acceleration → Poll / Int Completion → Copy or Unpin Result Data → Ret. From DD Completion

Flow with a Coherent Model

Shared Mem, Notify Accelerator → Acceleration → Shared Memory Completion

Advantages of Coherent Attachment Over I/O Attachment

- Virtual Addressing & Data Caching (significant latency reduction)
- Easier, Natural Programming Model (avoid application restructuring)
- Enables Apps Not Possible on I/O (Pointer chasing, shared mem semaphores, ...)

© 2014 IBM Corporation
Possible Example: CAPI Attached Flash Optimization

Eliminates 97% of instruction path
Saves 10 Cores per 1M IOPs

Application
Read/Write Syscall
FileSystem
strategy()
iodone()
LVM
strategy()
Disk & Adapter DD
Pin buffers, Translate, Map DMA, Start I/O
Interrupt, unmap, unpin, iodone scheduling

20K Instructions

< 500 Instructions

Attach flash memory to POWER8 via CAPI coherent Attach

Application
Posix Async I/O Style API
aio_read() aio_write()

User Library
Shared Memory Work Queue
Enabling the Art of the Possible on POWER8

CAPI enables I/O devices to operate on memory in the same way that general purpose processors can operate on memory.

- **4X** Smaller footprint
- **8X** Improvement
- **10X** Bandwidth Increase & **7X** Reduced Latency
- **12X** Less space & Energy

- POWER8 Data Clusters
- Java Acceleration on GPU
- RDMA & I/O
- CAPI Flash
HANA HA/DR Options

Storage subsystem replication

Sync or Async  
Write order dependent  
Asymmetric

Sync  
Write order dependent  
Symmetric

GPFS replication

Sync or Async  
Write order dependent  
Symmetric

HANA System Replication

Sync or Async  
Write order independent  
Hardware and vendor agnostic  
Symmetric  
Support NZDT upgrades

© 2014 IBM Corporation
Business Goals

- Improve analytics
- Decrease IT costs
- Increase performance
- Reduce time to close books
- IT to become invisible

Technology Enablers

- In-memory Columnar DB
- Converged infrastructure
- Flash Storage
- Cloud
- Security compliance
- Memory compression
- Fault tolerance
- HA, DR

Real Time Enterprise

In-memory Columnar DB
Converged infrastructure
Flash Storage
Cloud
Security compliance
Memory compression
Fault tolerance
HA, DR

IBM FlashSystem

ERP, BW, SCM, PLM, CRM, SRM on separate conventional DB instances
## SAP HANA on Power Services Draft Plan

### STG Led

1. **Setup** (PoC / Installation / upgrades / updates)
   - STG or BP sales led, contract via STG Lab Services
   - Sales order 'service units' based on sizing guidelines
   - Service offerings catalog based on System x services
   - STG passes lead to TSS for 2 + 3

### TSS Led

2. **Support** (defect/non-defect, Single Point of Contact)
   - TSS led where TSS operates via annual contract
3. **Proactive services** (annual health checks...)
   - TSS led where TSS operates via annual contract
   - TSS may utilize STG Lab Services

### Services Planning Workgroup

<table>
<thead>
<tr>
<th>IBM Global SAP Alliance:</th>
<th>Jim Dilley, Uwe Boettcher</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTS TSS Business Development:</td>
<td>Klaus Neumann</td>
</tr>
<tr>
<td>STG Lab Services:</td>
<td>Eric Aquaronne, Ian Jarman, Virginie Cohen, Michael Barenys</td>
</tr>
</tbody>
</table>

### STG Lab Services

- Planning & configuration
- Installation
- Link with SAP systems
- Backup & recovery
- High availability
- Coaching & education
- Migrate, upgrade

### IBM Lab Services:

- Donal O’Connell,
  - Lead consultant Europe
- Kurt Koehle,
  - Lead consultant NA
How to get from here to there

Not an Appliance:
Instead, an SAP and IBM Enterprise offering focused on supporting Enterprise customers
SAP sized for SAP applications and SAP HANA database requirements
Data migration through SLT

Ordered and fulfilled through normal IBM direct and BP order processing channels
Minimum configuration certified with SAP
Above minimum supported as validated by Hardware Configuration Check Tool (HWCCT)
Recommended IBM configurations (S, M, L, XL)

Supporting P7+ and P8 systems running SLES 11 SP3 and RHEL 7 and above

Implementation thru IBM Services, SI Services and Partners:
Plan, install, configure thru STG Lab Services (Service Units per Catalog)
Support, manage, operate, monitor, SpoC thru GTS/TSS
Business implementation services thru GBS, SI’s, etc.
Migration Services thru GBS Migration Factory, SI’s, SAP Services, etc.
Summary

HANA on Power

- Delivers the reliability necessary to support Suite on HANA
  - “Bet your business”
- Outstanding performance for analytics
- I/O and memory bandwidth to support transaction processing
- Virtualization to continue the journey of server consolidation
  - Not send you back to the 1990s with “islands of automation”
  - Or non-prod with a different stack than prod
• Join me in the **Expert Networking Lounge** on **Wednesday October 22 at 11:30 in Lounge 4**

• Stop by the IBM booth. IBM experts will be on hand to share IBM’s extensive portfolio of breakthrough technologies, innovative solutions and comprehensive services for SAP clients.
  
  • Ask us about cloud computing, enterprise mobility, IBM solutions for the SAP HANA platform, and social tools for business and commercial engagement.
  • See demos and user cases from the IBM’s Client Center: Lab for SAP Solutions.

• Find customer case studies and more on the **IBM-SAP Alliance website**: www.ibm-sap.com

Blog: SAPonPower.wordpress.com