Oracle9*i* (9.2) on the IA-64 HP-UX 11*i* platform

Sanhita Sarkar
Oracle Corporation
Sanhita.Sarkar@oracle.com

HP WORLD 2003
Solutions and Technology Conference & Expo

Topics for Discussion

- The concept of EPIC, evolution of IA-64 architecture and Itanium based HP-UX 11*i*
- Optimization Techniques for Oracle RDBMS on HP-UX 11.22 and beyond
- Integrated HP-specific enhancements from PA-RISC to Itanium 2 based Oracle9*i* (9.2.0.2) and analysis
- Configuring Oracle9*i* (9.2.0.2) on HP-UX 11.22
- Migration of Oracle9*i* (9.2.0.2) from PA-RISC to Itanium 2 based HP-UX 11*i* systems
- Ultimate advantages for various applications on IPF
- Oracle/HP Performance benchmarks
- Ongoing activities for the upcoming Oracle releases on Itanium 2 (Madison) based HP-UX (11*i* v2) 11.23

EPIC concept, HP-UX11*i* & Oracle9*i*

- Evolution of EPIC concept from RISC architecture – An overview
- The Itanium Processor family (IPF) from Intel and enhanced IA-64 compilers from HP
- The HP-UX 11*i* operating system – version 1.5 (11.20), version 1.6 (11.22) and version 2 (11.23) are based on Itanium, Itanium 2 (McKinley) and Itanium2 (Madison) respectively
- Oracle 9*i* (9.2.0.2 & 9.2.0.3) on Itanium 2 based HP-UX 11.22. Upcoming Oracle releases on Itanium 2 (Madison) based HP-UX 11.23

Optimization Techniques

- Compiler Optimizations for RDBMS performance
- Dynamic Instrumentation with HP Caliper
- Desirable Code scheduling
- Enhanced compile-time tuning capabilities for Itanium-based database applications

Compiler Optimizations

- HP Compiler exploits the features of the IPF architecture in the best possible way
  - Speculation : control and data
  - Predication
  - Pipelining

- **Control Speculation: executes instructions prior to conditions**

```
int a,b;                        ld.s        t1 = [p] ;;
extern int *p;                  add         b = t1,2
extern int global;              cmp.ne.unc  p1,p0 = condition,0 ;;
if(condition) {           (p1)  chk.s       b, L2
a = global;                L1:
b = *p + 2;                ...
}                          L2:  ld          t1 = [p] ;;
                                add         b = t1,2
                                br          L1
```

Compiler Optimizations contd..

- **Data Speculation:** executes load prior to store instructions

```
int a,b;                        ld.a        t1 = [q] ;;
extern int *p;                  add         b = t1,2
extern int *q;                   st          [p] = a
*p = a;                         chk.a        b, L2
b = *q + 2;              L1:
                         ...
                         L2:    ld          t1 = [q] ;;
                                add         b = t1,2
                                br          L1
```

Compiler Optimizations contd..

- **Data Predication:** control dependency to data dependency

```
if (a == 0) {
      x = 5;
} else {
      x = *p;
}
```

Using Branches

```
     cmp.ne.unc p1,p0 = a,0 ;;
(p1) br    L1 ;;
     mov   x = 5
     br    L2 ;;
L1:  ld    x = [p]
L2:
```

Using data predication

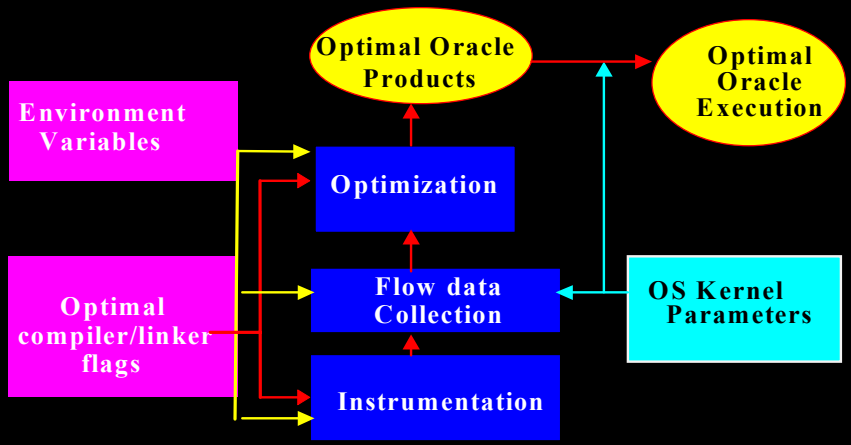
```
     cmp.ne.unc  p1,p2 = a,0 ;;
(p2) mov         x = 5
(p1) ld          x = [p]
```

- High-level compiler optimizations: +O3, +O4, etc
  - Helps inlining, faster numerical codes, faster code for math library functions, etc

Dynamic Instrumentation with HP Caliper

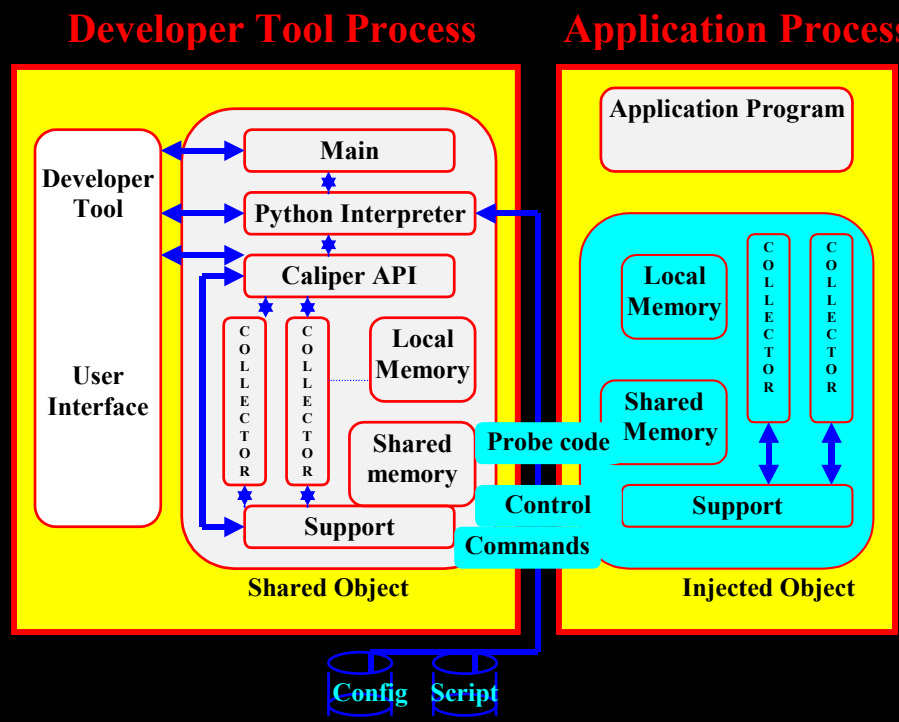
- Caliper-integrated dynamic instrumentation of RDBMS executables
  - Performed at run-time of an application program unlike static PBO-ing (Profile-based Optimization) on PA-RISC
  - Instruments only those parts of the program actually executed
  - Eliminates overhead of a separate instrumentation process
    - It is a two-step PBO process done with **+Oprofile=collect** build followed by the **+Oprofile=use** build.
  - Captures true run-time characteristics of RDBMS applications and provides useful analysis on the application behavior
  - Up to 30% improvement in performance over basic-optimized (+O2) binaries
  - Additional 10% improvement over PBO-ed binaries on PA-RISC

PBO-ing on PA-RISC vs on Itanium 2



Profile Based Optimization on HP-UX (PA-RISC)

Calipering (Dynamic Instrumentation) on HP-UX (Itanium 2)



Application Code Scheduling & compile-time tuning

- HP compiler optimizing scheduler for IPF allows code to be scheduled for Itanium and Itanium 2 based implementations
  - Use +DS{blended| itanium | itanium2 |native} compiler options
  - Application performance depends on the scheduling model
    - Code scheduled for Itanium2
      may be 40% slower on Itanium than the code compiled with +DSblended
      may be 7% faster on Itanium2 than code compiled with +DSblended
  - Compatibility is assured but recompilation is a better option for optimal performance
- Enhanced compile-time tuning capabilities with HP IPF compilers
  - Profiling, scheduling, link mode choices for shared and archived libraries with *–dynamic, -exec, -minshared*
  - Up to 10% improvement in OLTP performance over PA-RISC based RDBMS binaries with proper link mode choices

Integrated enhancements from PA-RISC Oracle9*i* (9.2)

- HP SCHED_NOAGE Process Scheduling Policy
  - overrides *SCHED_TIMESHARE* process scheduling policy
  - processes don't increase or decrease in priority or get preempted - improves latch performance
  - set system privileges RTSCHED and RTPRIO for oracle user
  - set Oracle initialization parameter HPUX_SCHED_NOAGE to specify desired process priority levels
  - Oracle9*i* (9.2.0.2) automatically sets the parameter to a permissible value and continues with the *SCHED_NOAGE* policy with the new priority
  - Use this policy only if benign in terms of system resource sharing
  - Tune the priority level appropriately to avoid CPU hog by Oracle processes causing basic system calls to hang
  - up to 10% enhancement for OLTP applications; improves latch performance

Integrated enhancements contd…

- HP Lightweight Timer
  - Uses lightweight library call gethrtime() to calculate elapsed time when Oracle9*i* initialization parameter timed_statistics=true
  - With Oracle9*i* (9.2.0.2) timed statistics are *automatically* collected by the Oracle database if the dynamic initialization parameter *STATISTICS_LEVEL* is set to *"TYPICAL" (default)* **or** *"ALL"*
  - Allows one to collect run-time statistics at all times while running an Oracle instance. Unlike Oracle9*i* (9.0.1), restarting the instance is not necessary
  - Up to 10% enhancement when timed_statistics=true by using gethrtime()
  - No negative impact with timed_statistics=true by default in Oracle 9.2.0.2

Integrated enhancements contd…

- HP asynchronous driver for I/O operations
  - Oracle processes can execute async I/O from shared memory and heap
  - Asynchronous Flag in Oracle shared global area uses non-block polling facility of HP asynchronous driver
  - Allows asynchronous polling for I/O completions and submissions
  - Helps to attain scalability of parallel I/O processes
  - HP asynchronous kernel driver *"asyncdsk"* should be installed (use *sam* utility to check)
  - Device file */dev/async* should be configured with proper ownership and permissions
  - Set system privilege *MLOCK* for Oracle user
  - Set oracle initialization parameter *disk_asynch_io to true*
  - Use raw devices instead of file systems - HP restriction
  - Up to 15% enhancement for DSS applications

Enhancements contd..
Oracle9*i* (9.2.0.3) RAC & HP HMP

- HyperFabric is a cluster interconnect fabric from HP
  - has high speed (a link rate of 4 GB/s over fiber over 200 meters) and excellent scalability (up to 16 nodes via point-to-point connectivity and up to 64 nodes via fabric switches
  - supports HP proprietary *HyperMessaging protocol (HMP)*
  - IPF Version B.11.22.00 called *"HyprFabrc-00"*

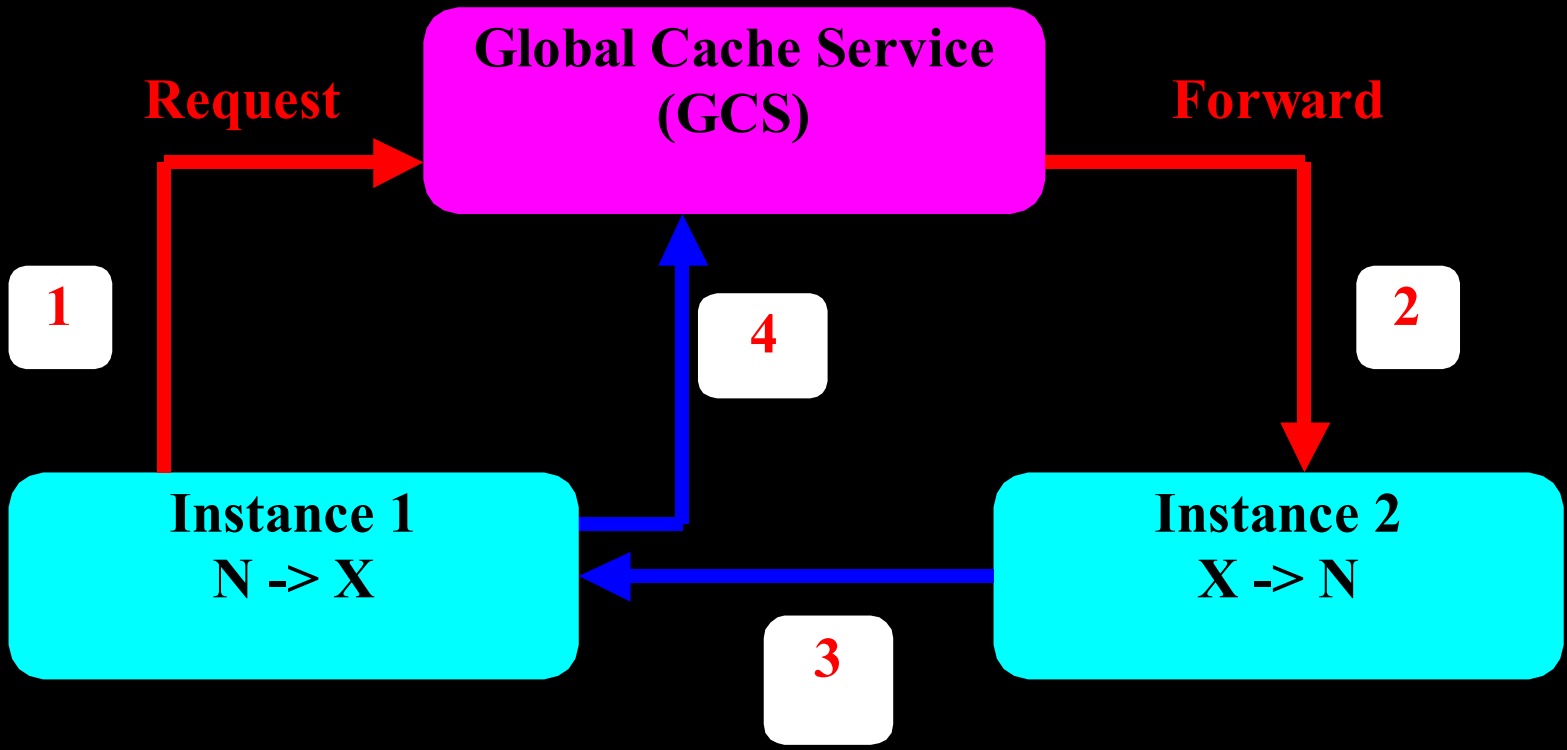
- HMP is the HyperMessaging IPC protocol over HyperFabric
  - provides a true Reliable Datagram model for remote direct memory access
  - allows direct access to memories and caches of other servers in the cluster and provides faster data transfer than UDP or TCP.

- The HP Cluster Manager
  - named the ***"ServiceGuard Extension for RAC"*** or **SGeRAC,** a special edition supporting Oracle9*i* RAC on HP 9000 servers.
  - Components: Cluster Manager, Distributed Lock Manager, Package manager, Network Manager
  - monitors health of all nodes, automatic application package failover, data integrity

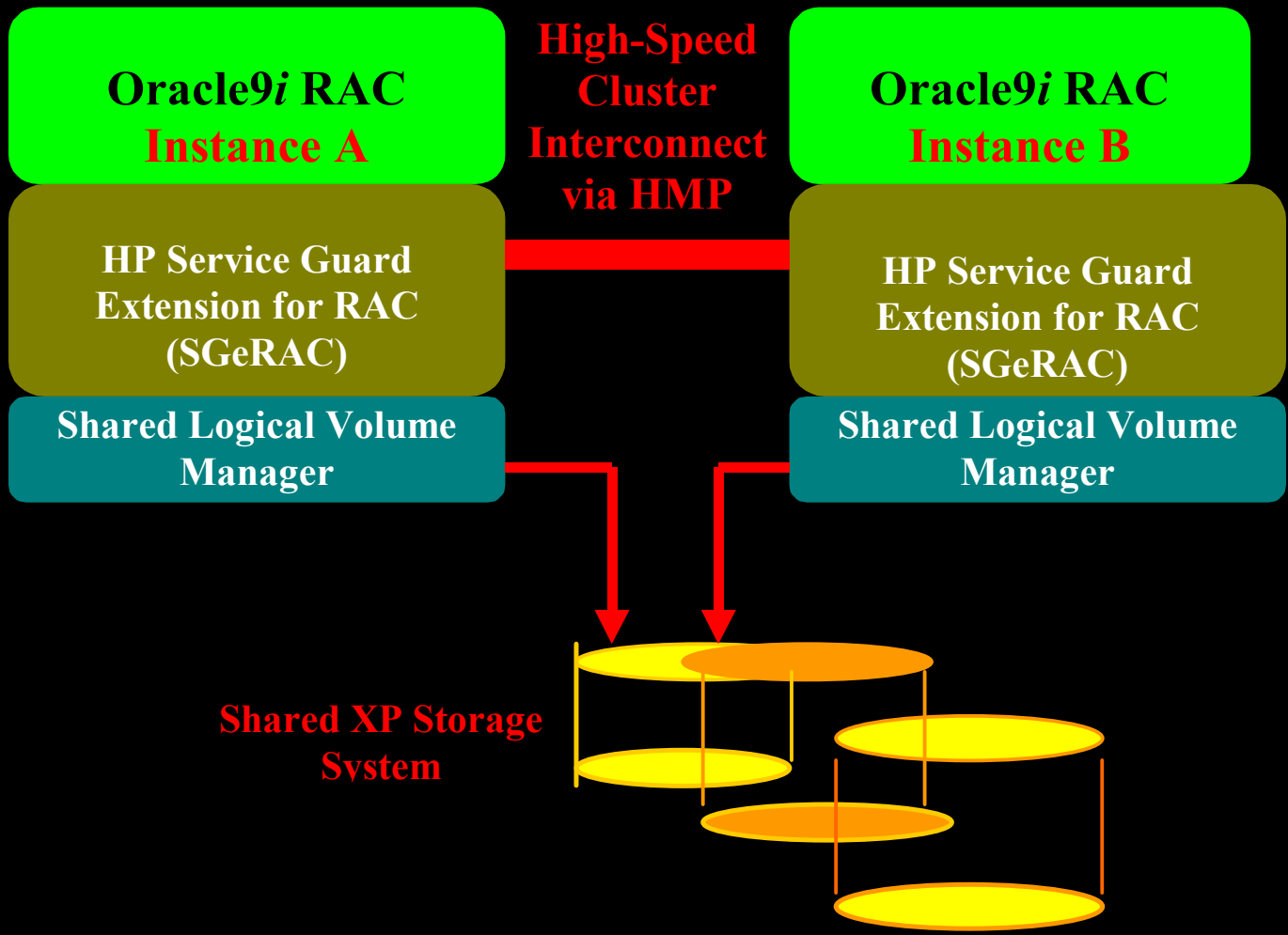
Enhancements contd..
Oracle9*i* (9.2.0.3) RAC & HP HMP

- Oracle9*i* Cache Fusion
  - provides an expanded database cache for queries and updates
  - reduced disk I/O synchronization which overall speeds up database operations.
  - enhanced coordination by the Global Cache Service (GCS) with the HP Cluster Manager (CM) e.g., the Service Guard Extension for RAC, leads to advantages like resource affinity, data integrity, application transparency and fault tolerance
- RAC Guard I and RAC Guard II
  - are RAC components and provide an efficient high availability solution
- RAC and HMP integration on IPF HP-UX 11.22 in Oracle 9*i* (9.2.0.3)
- Performance Analysis
  - Successful Stress Tests, Recovery and Destructive Tests
  - On a 16-node HP cluster configuration, a 20% improvement in Oracle9*i* RAC OLTP throughput and response time using HMP over UDP

Cache Fusion: requesting a changed block for modification



Enhancements contd..
Oracle9*i* (9.2.0.3) RAC & HP HMP



Configuring Oracle9*i* (9.2) on HP-UX 11.22

Pre-installation requirements

- OS patches - accuracy a must
  - Quality Pack (patch bundle) plus additional patches
  - 9.2.0.3 RAC needs HyperFabric B.11.22.00 and SGeRAC
  - Refer to Oracle 9.2.0.2 & 9.2.0.3 Release Notes

- Granting Privileges to the ***dba*** group
  - MLOCK, RTSCHED and RTPRIO

  Execute: *# setprivgrp dba MLOCK RTSCHED RTPRIO*

  Edit file */etc/privgroup: dba MLOCK RTSCHED RTPRIO*

- Set the HP-UX kernel parameters
  - shmmax, maxdsiz_64bit, max_async_ports, etc
- Configure asynchronous I/O
  - HP asynchronous kernel driver *(asyncdsk)* and device file /dev/async
  - raw devices for I/O

Configuring Oracle9*i* (9.2) on HP-UX 11.22 contd..

Post-Installation Recommendations

- Large Memory Allocations and Oracle9*i* (9.2) tuning on HP-UX
  - **Oracle initialization parameter** cursor_space_for_time = true
    - Yields better SQL performance by persistent private SQL areas but leads to an increased cursor memory
  - **Large virtual memory default page size "*L*"**
    - Large pages yields better application performance
    - May lead processes to indicate large memory allocations followed by "out of memory" error, for memory-constrained applications
  - **Tuning memory-constrained applications**
    - Maintain cursor_space_for_time = true
    - Decrease page size as follows:

      /usr/bin/chatr +pd <new size> $ORACLE_HOME/bin/oracle

Configuring Oracle9*i* (9.2) on HP-UX 11.22 contd..

- Revalidate natively compiled Java objects for a migrated 9.2.0.2 DB on IPF (special case)

```
SQL> create or replace java system;
```

- Natively compiled PL/SQL for performance
  - Configure Oracle9*i* (9.2) database for native C code compilation of PL/SQL
  - Oracle9*i* uses $ORACLE_HOME/plsql/spnc_makefile.mk and the C compiler/linker and utilities
    - Transforms PL/SQL statements to generate C code
    - Compiles and links the resulting C code into shared libraries which are loaded at run time of PL/SQL
  - Requires system parameters e.g., plsql_native_make_file_name, etc
  - Set plsql_compiler_flags to *NATIVE* at system level to compile whole database as native
  - 30% performance gains for compute-intensive DSS applications
  - Experiment nativity first on a test database to look for any substantial gains
  - See Release Notes for Oracle9*i* (9.2.0.2) to enable PL/SQL nativity

Migrating Oracle DB from PA–RISC to Itanium HP-UX system

- Compatibility of PA-RISC based HP-UX with Itanium-based HP-UX is a great advantage for migration
- Conditions for migrating Oracle RDBMS from PA to IA-based HP-UX
  - Oracle version on PA-RISC must be 8.1.7.x or 9.2.0.2
  - Oracle version 8.1.7.x on PA-RISC should be upgraded to 9.2.0.2 before or immediately after migration. Otherwise, database can't be opened
- Steps for migration:
  1. Make sure the Oracle version is 81.7.x or 9.2.0.2. Upgrade now if necessary
  2. Prepare the Itanium system with HP-UX 11.22 (v1.6) and associated patch upgrades as required
     - If it will run RAC, install SGeRAC and do other cluster configurations
     - Configure the raw logical volumes or file systems as required
     - Make sure Oracle9*i* (9.2.0.2) is installed. Check for *dba* privileges

Migration contd..

- Steps for migration (Contd…)

3. Copy over Oracle configuration files (init.ora, control files, listener.ora, tnsnames.ora, etc) from PA-RISC to Itanium system

   `rcp $OH-PA/dbs/*.ora  ipf-node:$OH-IA/dbs`

   Copy database files onto a tape or by using *"dd"* as the case is

   OR Move the database from PA to IA system

   - Unmount (***umount***) the filesystems (if using) and deactivate (***vgchange***) volume groups on PA
   - Export the logical volume groups and associated logical volumes

     `vgexport -p -m mapfile -s  /dev/vgORA`
   - Reconfigure the disk arrays as needed on the IA system
   - Import the volume groups per the info in the mapfile

     `vgimport -m mapfile -s  /dev/vgORA`
   - Ensure correct permissions to the volume groups, activate (***vgchange***) the volume groups (extra steps to permit sharing in RAC case) and ***mount*** the activated volume groups (for filesystems)

Migration contd…

- Steps for migration (Contd…)

4. Edit the database files (init.ora, listener.ora, etc) for the IA environment
5. While using raw volumes, check for device driver *asyncdsk* in kernel, configure */dev/async* file, use *disk_asynch_io=true*
6. If Oracle8*i* version 8.1.7.x, then upgrade to Oracle9*i* release 2 (9.2.0.2) now and go to step 9
7. Login as oracle user and startup the database with/without RAC option as the case is
8. Revalidate natively compiled Java objects as user *SYSTEM*

```
SQL> create or replace java system;
```

9. Do a clean shutdown and restart database

Migration contd..

- Performance of a migrated Oracle9*i* (9.2.0.2) database
  - No noticeable performance benefit in recreating the PA-based database natively on the Itanium-based system
  - Recompiling applications “natively” is a benefit but it does not apply to data migrated, as no changes in database structures or block layout takes place during migration
  - Lab tests show no noticeable performance penalty on a migrated database
  - A database migrated from a PA-RISC HP-UX system has been used to run the published Oracle Applications Standard Benchmark 11.5.6 on a 4-way Itanium HP Server rx5670

The ultimate advantages for various applications

- Ease of compatibility on migration from PA-RISC
  - Existing 32-bit and 64-bit applications can immediately run on Itanium 2 based HP-UX 11.22 without recoding or recompiling
  - An advantage for corporate developers with large inventories of existing applications
- Responsiveness of Web-based Oracle9*i* (9.2.0.2) applications on HP-UX 11.22
  - Enhanced cache spec of Itanium 2 – larger L2, better support for branch instructions and load/store operations. Designed for smooth OLTP operations during peak periods
  - A 100% increase in Oracle9*i* (9.2.0.2) OLTP performance on a 4-way HP-UX 11.22 server compared to that on HP-UX 11.20 (Itanium) server
  - OLAP engine within the Oracle9*i* (9.2) DB -> reduced information cycle time -> fast query response time for data mining

The ultimate advantages contd..

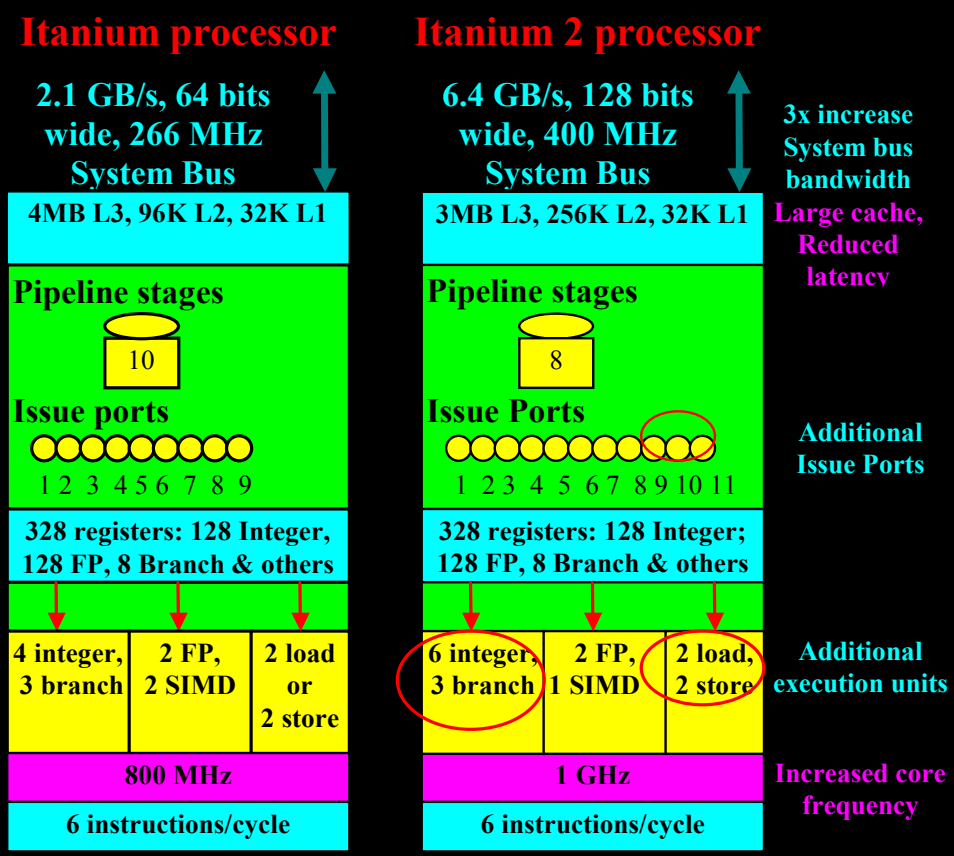
- Customization capabilities
  - Floating point computational performance of Itanium 2, rich data type support of Oracle9*i* (9.2) -> native support of XML data type – SQL operations on XML data, etc. LOBs, heterogeneous Oracle9*i* (9.2) Streams
  - Personalized web information service, simulations, data analysis for varying user profiles – an advantage for growing e-business environments

- Reliability, High Availability and Data Security
  - Efficient error handling model of IPF ensures non-disruption of applications – makes environment reliable; Oracle9*i* RAC Guard provides automatic failover and recovery
  - Oracle9*i* RAC technology, HP HMP and HyperFabric ensure HA
  - Efficient RSA computations (encryption) by Itanium – pipelined Multiply/Add instructions, register width, floating point registers.

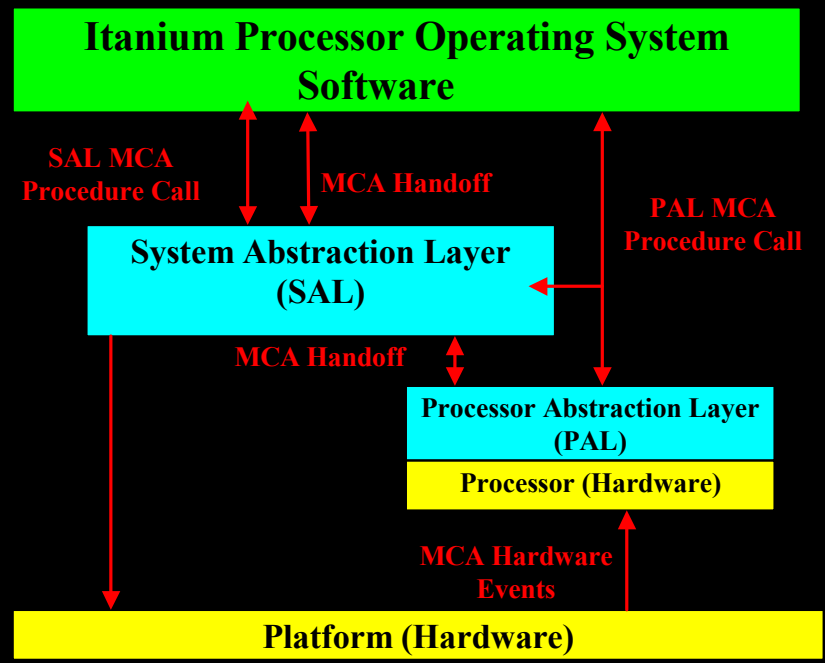
    Database-level security: Oracle9*i* (9.2) -> Label Security, Advanced Encryption security

The ultimate advantages contd..

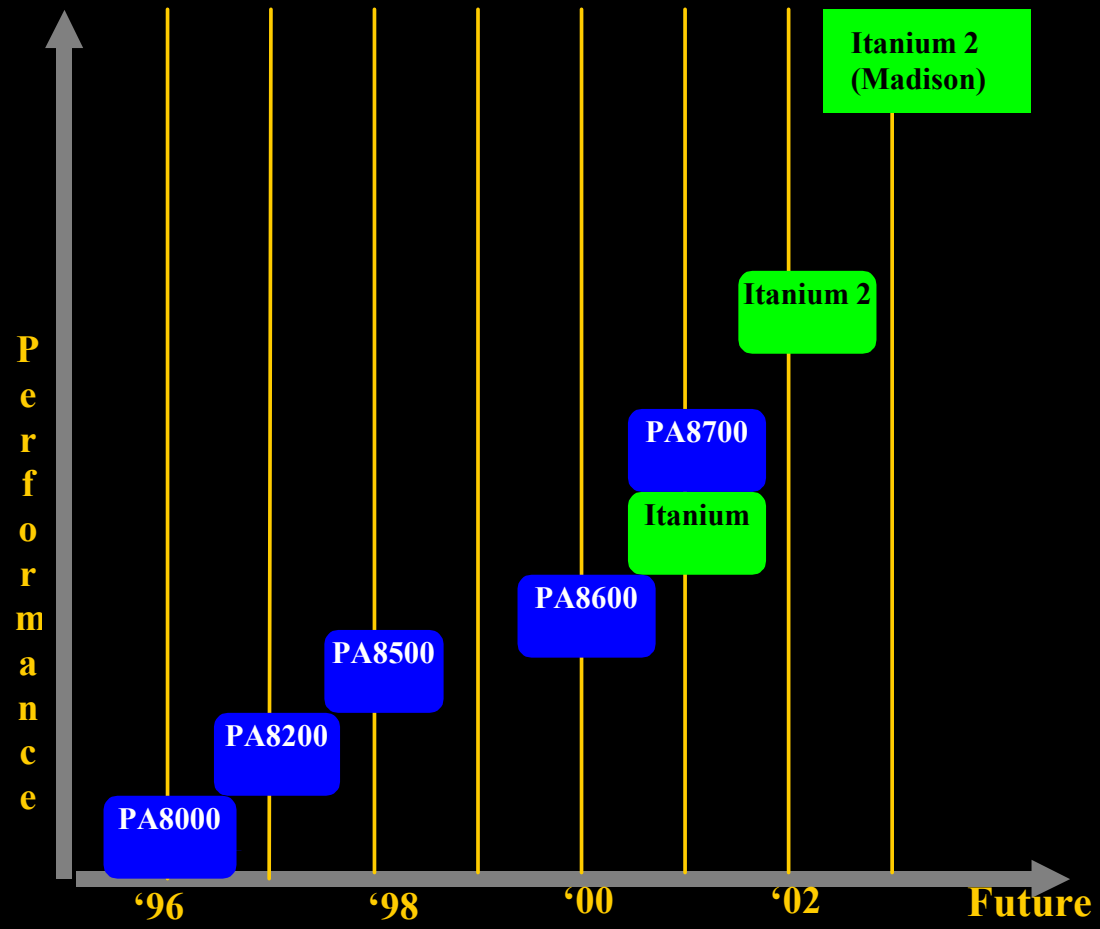
Itanium vs Itanium 2 Chip Architecture



Error Handling by PAL and SAL on Itanium 2



The ultimate advantage contd..
Itanium 2 performance



Performance of PA-RISC Processor Family versus Itanium Processor Family

Recent HP/Oracle IA-64 Benchmarks

- **Oracle Applications Standard Benchmark 11.5.6 on Split Configuration**
  - Oracle9*i* (9.2) single-node result on a 4-way rx5670 server (Itanium 2: Madison, 1.5GHz, 6MB L3 cache) running HP-UX 11.23 (Database Tier). Application Tier on PA-RISC based HP-UX 11*i* (11.11)
    - 6440 users with an average response time of 0.6 secs
  - Oracle9*i* (9.2) single-node result on a 4-way rx5670 server (Itanium 2, 1GHz, 3MB L3 cache) running HP-UX 11.22 (Database Tier). Application Tier on PA-RISC based HP-UX 11*i* (11.11)
    - 4200 users with an average response time of 1.51 secs
- **Two-tier SAP Standard Applications SD (Sales & Distribution) Benchmark**
  - Oracle9*i* (9.2) on a 4-way HP rx5670 server (Itanium 2: Madison, 1.5 GHz, 6MB L3 cache) running HP-UX 11.23
  - 860 SAP SD benchmark users with 1.97 seconds average dialog response time

Recent HP/Oracle IA-64 Benchmarks Contd…

- **OLTP benchmark (TPC-C)**
  - Oracle Database 10G (upcoming Oracle release)
    - On a 64-way Itanium 2 (Madison: 1.5GHz, 6MB L3) Integrity Superdome (HP-UX 11*i* v2)
      - Performance metric - 824K tpmc; price/performance - $8.28/tpmc
    - On a 4-way Itanium 2 (Madison: 1.5GHz, 6MB L3) Integrity rx5670 (HP-UX 11*i* v2)
      - Performance metric - 131K tpmc; price/performance - $7.25/tpmc

Future Oracle releases on IA-64 HP-UX 11*i*

- Investigate & integrate new features of HP Integrity Server Family running HP-UX 11*i* v2 (11.23)
  - Numa-awareness & cell-local memory, dynamic expandability, multi-OS flexibility, high performance features
- Make efficient use of the new HP C Compiler optimizations for Oracle RDBMS on HP-UX 11.23 (Madison)
  - Optimizer flags, new prefetching techniques, enhanced dynamic profiling, etc
- Investigate automated storage management at database level with respect to conventional HP LVMs
- Enhanced asynchronous I/O interface to support file systems
- Selective use of threaded libraries for Oracle client applications
- Investigate performance of new Oracle features on HP-UX 11.23

References

- Paper “Oracle9*i* (9.2) on the IA-64 HP-UX 11*i* Platform”
  - HP World 2003 Conference Proceedings, Atlanta, GA
- Paper “A Fast Track to Oracle9*i* Release 2 on HP-UX 11*i* “
  - HP World 2002 Conference Proceedings, Los Angeles, CA
- Release Notes for Oracle9*i* (9.2.0.2) and Oracle9*i* (9.2.0.3)
- HP Online Documentation

  http://www.hp.com/dspp
- Oracle Online Documentation

  http://otn.oracle.com
- Intel Online Documentation

  Developer Services at http://www.intel.com

Acknowledgements

- Hewlett Packard: Development Alliance Lab
- Hewlett Packard: Performance Benchmarking Group
- Oracle: Product Management and Alliance
- Oracle: Product Line Engineering (Release)
- Oracle: Release documentation Group
- Oracle: Server Technologies Division
- Oracle: Platform Technologies Division

Interex, Encompass and HP bring you a powerful new HP World.

interex
shared knowledge • shared power

encompass
AN HP USER GROUP

hp
invent