Oracle9*i* and Beyond on the HP-UX 11*i* platform

Sanhita Sarkar
Oracle Corporation
Sanhita.Sarkar@oracle.com

HP WORLD 2003
Solutions and Technology Conference & Expo

Topics for Discussion

- Key developments in Oracle9*i* (Releases 1 and 2)
- A recap of some of the generic performance features in Oracle9*i*
- Configuring Oracle9*i* on HP-UX for optimality
- HP-specific features and enhancements in Oracle9*i*
- Analysis of the performance enhancements – Lab Tests and published Benchmarks
- Activities towards upcoming Oracle releases on HP-UX

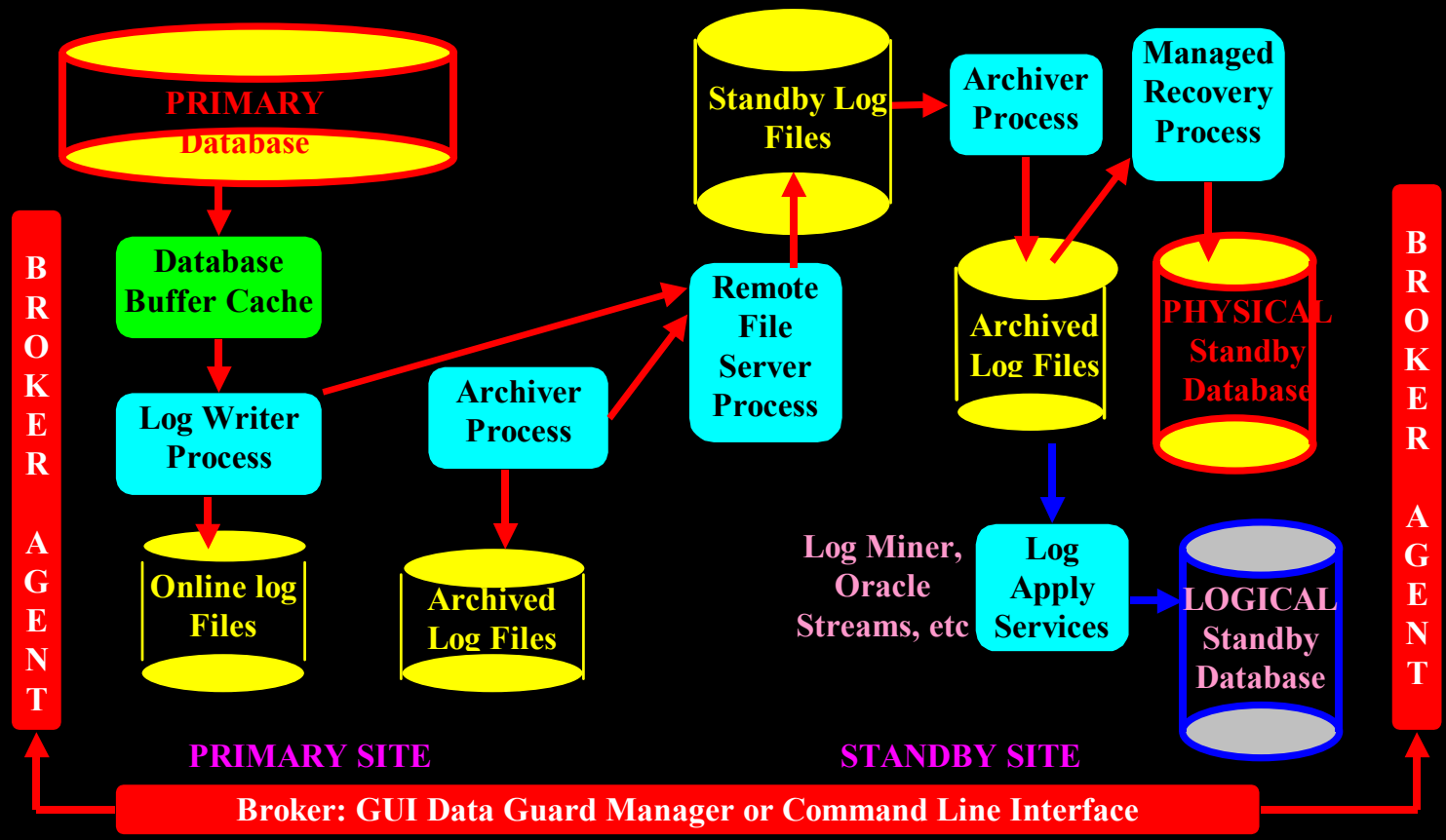
Key developments in Oracle9*i*

- Oracle9*i* Data Guard
- Oracle9*i* Streams
- Oracle9*i* Real Application Clusters (RAC)
- Automated Space and Memory management
- Enhanced Partitioning Techniques
- Integration of OLAP and XML within the Oracle database
- New Performance features

Oracle9*i* Data Guard

- Concept of a Primary and Standby databases
  - Standby : a replica of Primary database to protect against corruption, human errors and disasters, in the event of a failure of the primary database
- Physical and Logical Standby Databases
- Logical Standby concept in Oracle9*i* (9.2)
  - Open for concurrent transactions, may be of a different version of the same OS, different schema and layout, fully SQL maintained, better tolerance to corrupt logs
- Oracle9*i* Data Guard architecture
  - can be a mix of Physical and Logical Standby databases besides the Primary
- Oracle9*i* Data Guard components
  - LGWR, ARCn, Log Transport Services, Remote File Service (RFS) Process, Managed Recovery Process, Log Standby Apply Services and the Data Guard Broker

Oracle9*i* Data Guard Architecture



LogMiner Operations

- SQL> EXECUTE DBMS_LOGMNR.START_LOGMNR(OPTIONS => -
  2 DBMS_LOGMNR.DICT_FROM_REDO_LOGS + -
  3 DBMS_LOGMNR.CONTINUOUS_MINE + -
  4 DBMS_LOGMNR.COMMITTED_DATA_ONLY);

- Query V$LOGMNR_CONTENTS view and get the output in a SQL format:
  SQL> SELECT OPERATION, SQL_REDO, SQL_UNDO
  2 FROM V$LOGMNR_CONTENTS
  3 WHERE SEG_OWNER = 'SCOTT' AND SEG_NAME = 'ORDERS' AND
  4 OPERATION = 'DELETE' AND USERNAME = 'RON';

OUTPUT:

OPERATION	SQL_REDO	SQL_UNDO
DELETE	delete from "SCOTT"."ORDERS" where "ORDER_NO" = 2 and "QTY" = 3 and "EXPR_SHIP" = 'Y' and ROWID = 'AAABM8AABAAALm/AAA'	insert into "SCOTT"."ORDERS" ("ORDER_NO", "QTY", "EXPR_SHIP") values(2,3,'Y');
DELETE	delete from "SCOTT"."ORDERS" where "ORDER_NO" = 4 and “ QTY" = 7 and "EXPR_SHIP" = 'Y' and ROWID = 'AAABM8AABAAALm/AAC';	insert into "SCOTT"."ORDERS" ("ORDER_NO","QTY","EXPR_SHIP") values(4,7,'Y');

LogMiner Operations contd...

- Extracting actual data values from redo logs using mine functions
  - Query the V$LOGMNR_CONTENTS view for all updates to scott.emp that increased the sal column to more than twice its original value:

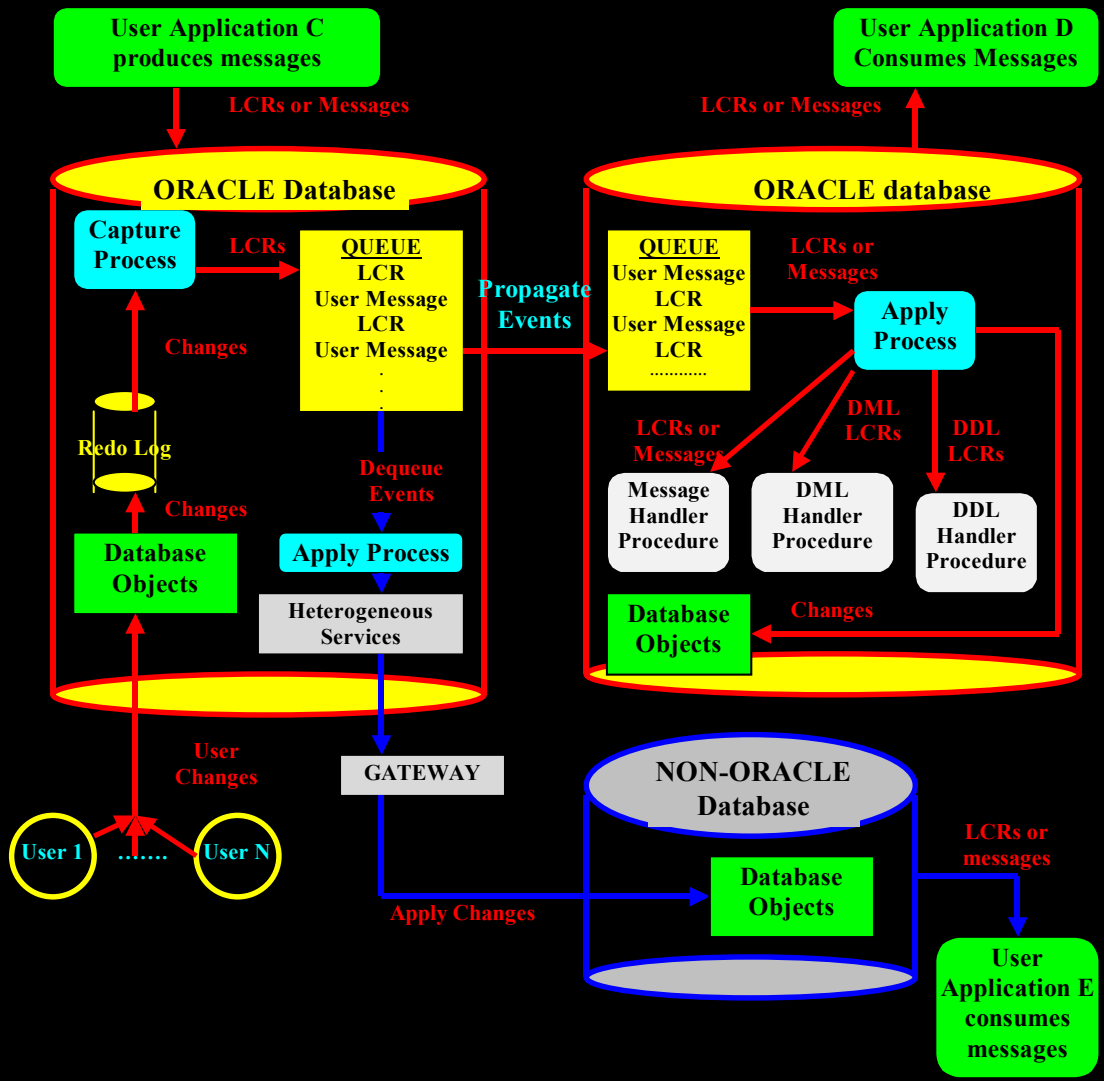
```
SQL> SELECT SQL_REDO FROM V$LOGMNR_CONTENTS
  2   WHERE
  3   SEG_NAME = 'emp' AND
  4    SEG_OWNER = 'SCOTT' AND
  5   OPERATION = 'UPDATE' AND
  6   DBMS_LOGMNR.MINE_VALUE(REDO_VALUE, 'SCOTT.EMP.SAL') >
  7   2*DBMS_LOGMNR.MINE_VALUE(UNDO_VALUE, 'SCOTT.EMP.SAL');
```

- Log apply services (LogMiner) convert the data from the redo logs into SQL statements and then executes these SQL statements on the logical standby database
- Because the logical standby database remains open, tables that are maintained can be used simultaneously for other tasks such as reporting and queries. Helps detecting any errors at application level and execute the necessary undo operations

Oracle9*i* Streams

- A new information sharing feature between databases
  - Propagation of data, transactions and events in a data stream from one database to another; Heterogeneous data sharing
- Stages of information transfer
  - **Capturing Events**
    - Database changes and application generated events are put into a queue
    - Implicit Capture by the server Capture process from redo logs -> Logical records (LCRs) put into a queue
    - Explicit Capture of user messages or LCRs directly put into a queue
  - **Staging and Propagating Events**
    - LCR staging provides a holding area for security, auditing and tracking LCR data between queues for propagation within the same database or heterogeneous ones
  - **Applying Events**
    - An Oracle background Apply process dequeues events from a queue
    - Applies each event to a database object or passes the event as a parameter to a user-defined procedure - message handlers, DML handlers and DDL handlers.

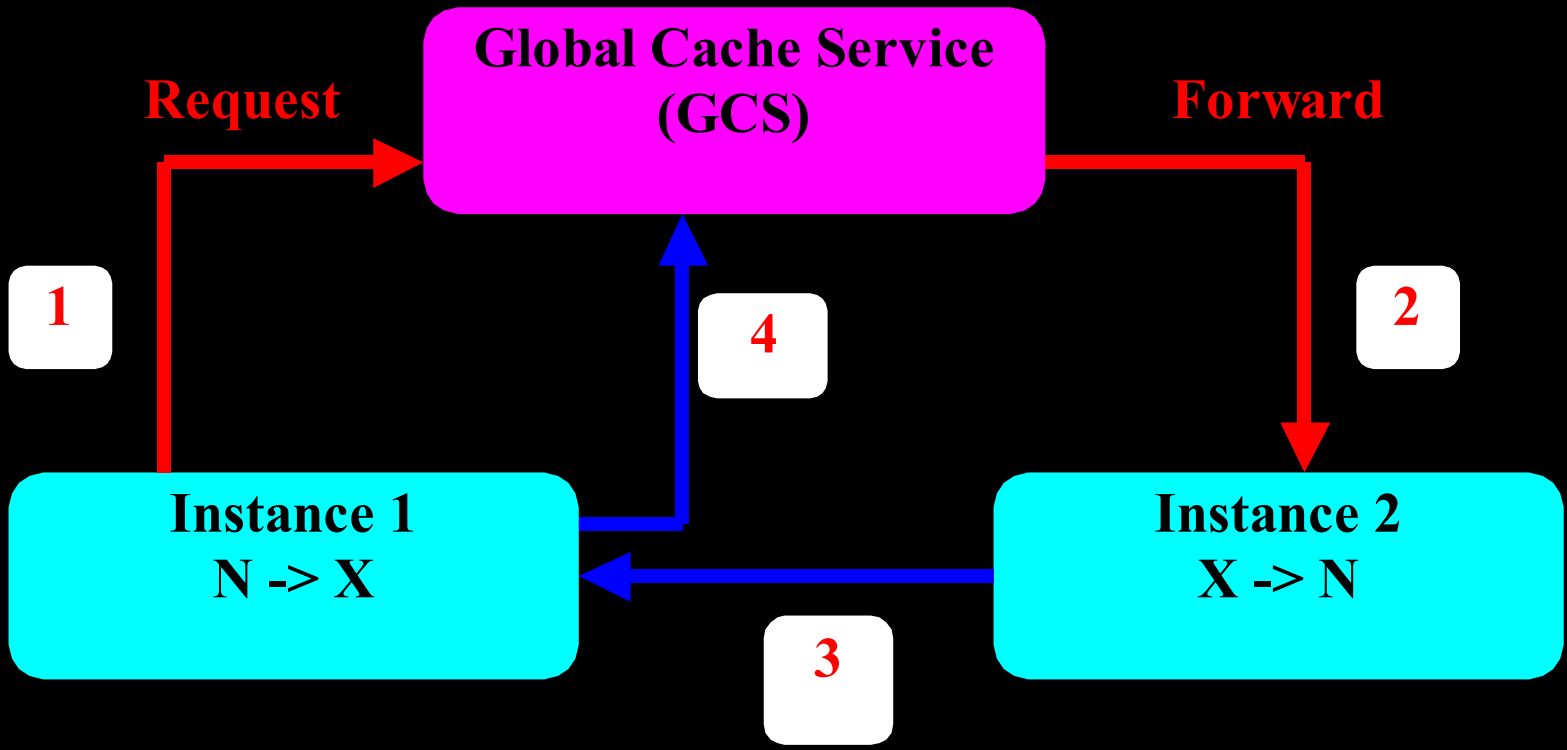
Streams: Heterogeneous data sharing



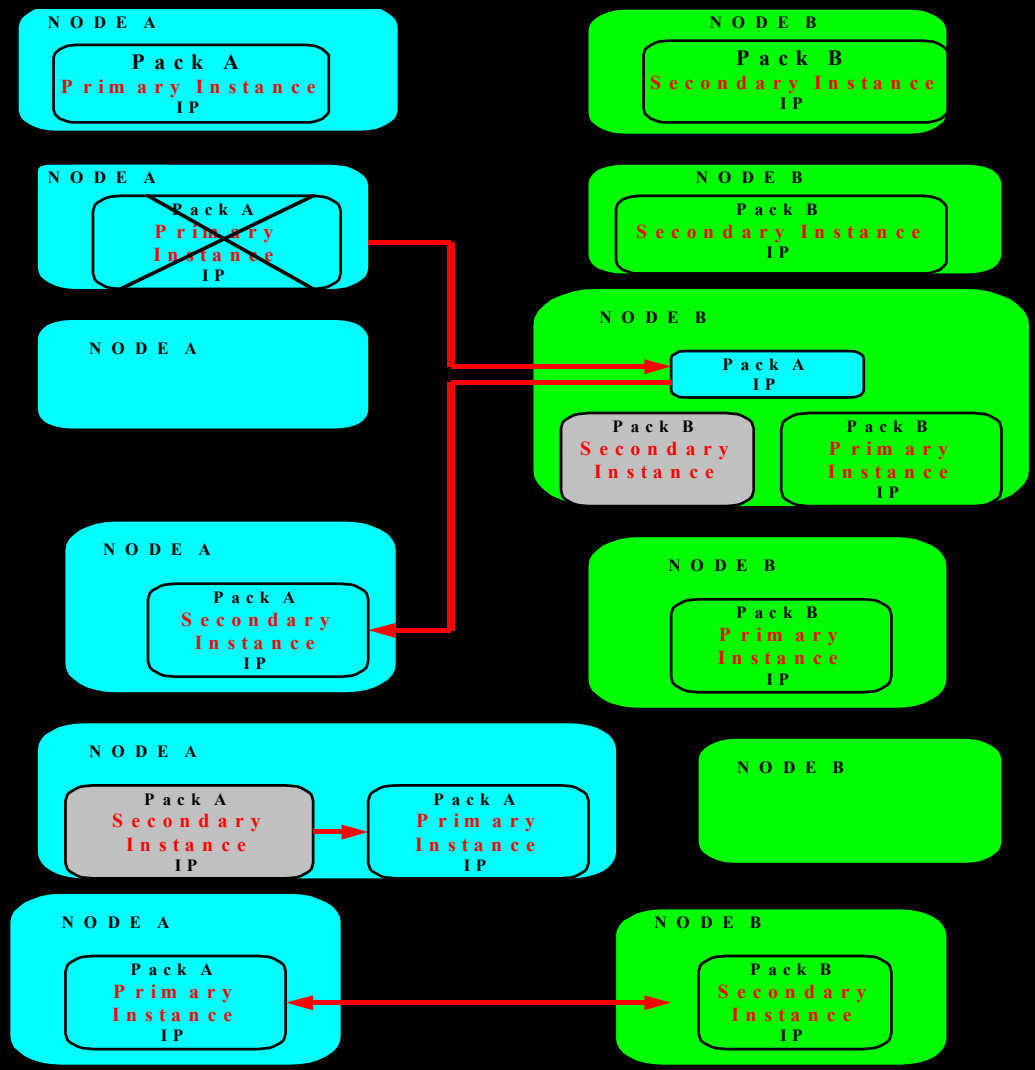
Oracle9*i* Real Application Clusters (RAC)

- A cluster software architecture providing a single view of the Oracle database across the cluster server nodes
- Cache Fusion
  - a diskless, cache-to-cache transfer mechanism of read-consistent data blocks between instances. Uses Global cache, Cluster Manager and fast IPC to ensure cache coherency.
- Allows application transparency, fault tolerance, resource affinity, data integrity
- RAC Guard I and II components for high availability
- RAC Guard I
  - Automatic failover and recovery using RAC Guard Packs
  - Enforced Primary/Secondary Configurations; a read-only Secondary
  - No delay in reestablishing connection after a failover; bounded recovery time
- RAC Guard II
  - Extension to N-node fully active configuration
  - Uses service names for ease of workload transfer and connection failover

Cache Fusion: requesting a changed block for modification



RAC Guard I Operation



Other Notable Features in Oracle9*i*

- Automated space and memory management
  - Locally managed SYSTEM tablespace
  - EXTENT MANAGEMENT LOCAL clause in CREATE DATABASE
  - Built-in advisories for diagnostic tasks related to memory tuning
- Composite Range-List Partitioning
  - Partition by range value, with sub-partitions by list value
  - Better customer data maintenance for geographical regions by month
- Integration of OLAP and XML inside the Oracle9*i* database
  - Oracle XML DB provides XMLType datatype support to store and manipulate XML data, execute SQL operations (queries, OLAP) on XML data and vice versa, format SQL query results into XML, etc
  - Oracle OLAP option provides access to both relational and multidimensional data stored in Oracle data files through SQL or an OLAP API. OLAP engine runs within the Oracle kernel and uses OLAP DML for analytic calculations

Notable Features in Oracle9*i:* *Performance features*

- **Dynamic sampling**
  - Provides better quality of performance stats. Use *optimizer_dynamic_sampling* between 0 and 10
- **Performance Tuning Intelligent Advisories**
  - Shared Pool Advisory; Mean-Time-To-Recover (MTTR) Advisory; PGA Aggregate Target Advisory (tuning pga_aggregate_target)
- **Data Segment Compression**
  - Can be applied to heap-organized (partitioned, not index-organized) tables; eliminates column value repetition within individual blocks; compresses several partitions or complete partitioned table
  - Reduces disk and memory usage; enhances read-only operations

```
ALTER TABLE sales MOVE PARTITION sales_q1_1998 TABLESPACE
                                        ts_arch_q1_1998 COMPRESS;
ALTER TABLE sales MODIFY PARTITION sales_q1_1998 REBUILD
                                        UNUSABLE  LOCAL INDEXES;
ALTER TABLE sales MERGE PARTITIONS sales_q1_1998, sales_q2_1998
                    INTO PARTITION sales_1_1998 TABLESPACE ts_arch_1_1998
                                        COMPRESS UPDATE GLOBAL INDEXES;
```

Notable Features in Oracle9*i*
Performance features Contd….

- First_Rows_n optimization
  - cost based approach to return first n rows (n=1,10,100,1000)
- Cursor_Sharing parameter
  - Setting to SIMILAR instead of the default EXACT
- Index monitoring
  - Identifying unused indexes
- System stats estimates by optimizer
- New optimizer hints
  - NL_AJ, NL_SJ, CURSOR_SHARING_EXACT, FACT/NO_FACT
- Fast_Start_MTTR_Target parameter
  - expected mean time to recover
- Dynamic memory management
  - PGA & SGA : PGA_AGGREGATE_TARGET, SGA_MAX_SIZE

Configuring Oracle9*i* on HP-UX

Pre-installation requirements

- OS Patches - accuracy a must
  - Quality Pack plus additional patches
  - Refer to Oracle9*i* (9.2) Installation Guide and 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
  - Make sure the HP asynchronous kernel driver *(asyncdsk)* is installed and you are using raw devices for I/O
  - Configure asynchronous I/O device file /dev/async

Configuring Oracle9*i* contd..

Post-Installation Recommendations

- Large Memory Allocations and Oracle9*i* tuning on HP-UX
  - Setting **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* contd..

Post-Installation Recommendations Contd…

- 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
  - Require 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) to enable PL/SQL nativity

HP-specific enhancements in Oracle9*i*

- 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 HPUX_SCHED_NOAGE parameter to specify desired process priority levels
  - Oracle9*i* (9.2) automatically sets the parameter to a permissible value and continues with the SCHED_NOAGE policy using the new priority.
  - Up to 10% enhancement for OLTP applications, improves latch performance
  - 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

HP-specific enhancements contd…

- HP Lightweight timer
  - Uses lightweight library call *gethrtime()* to calculate elapsed time when oracle initialization parameter timed_statistics=true
  - Replaces the earlier implementation using the heavyweight gettimeofday()
  - 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 by using *gethrtime()* when *timed_statistics=true*
  - No negative impact with *timed_statistics* = *true* by default in Oracle9*i* (9.2)

HP-specific 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 async driver
- HP asynchronous kernel driver *"asyncdsk"* should be installed (use *sam* utility to check)
- Configure async device file */dev/async* 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
- Allows asynchronous polling for I/O completions followed by submissions scalability of parallel I/O processes
- Up to 15% enhancement for DSS applications

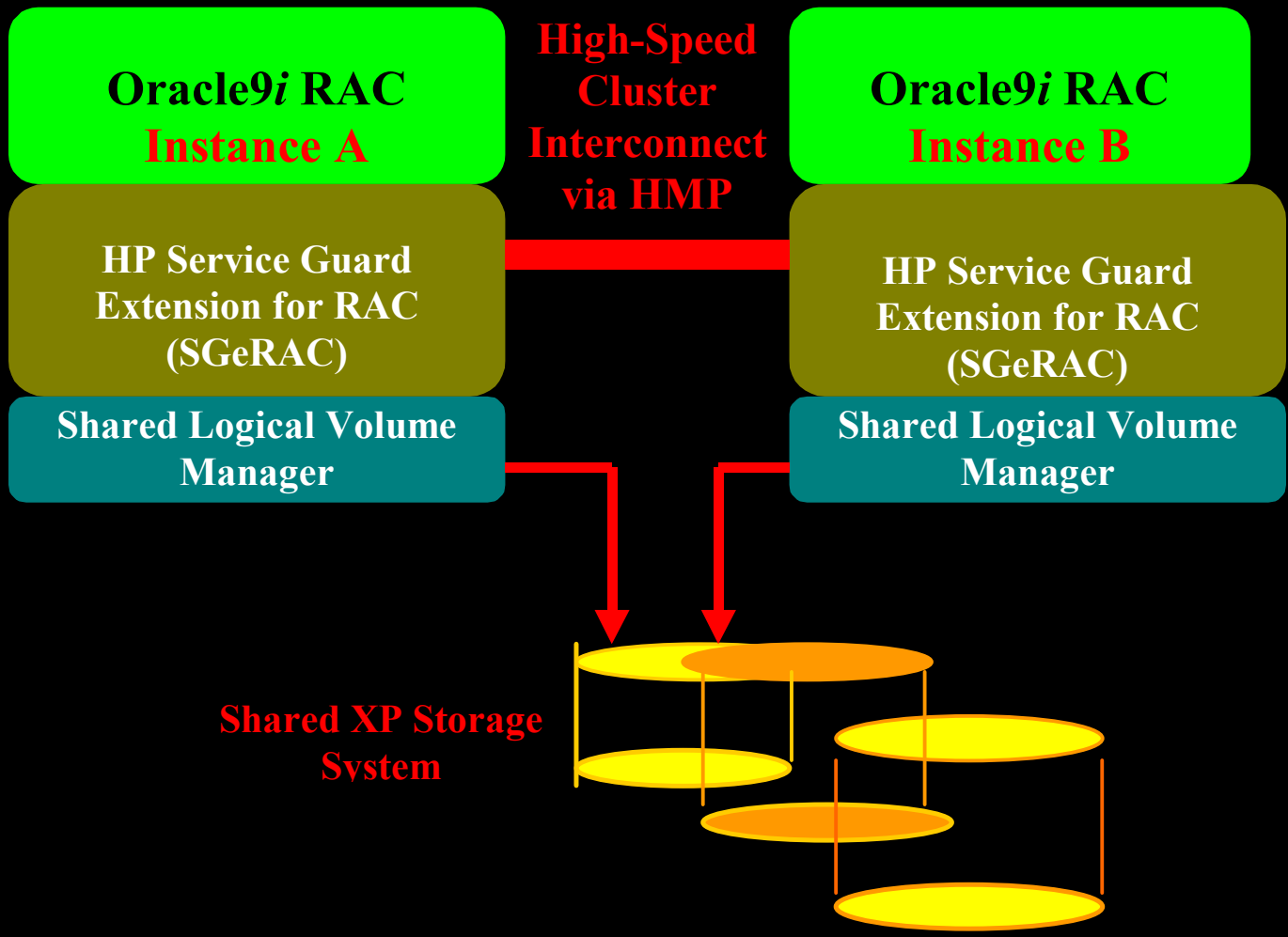
Enhancements contd..
Oracle9*i* RAC & HP HMP

- HyperFabric II is a cluster interconnect fabric from HP
  - It has higher 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)
- 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
  - called 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* 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 HP SGeRAC leads to advantages like resource affinity, data integrity, application transparency and fault tolerance
- RAC Guard I and RAC Guard II
  - provide an efficient high availability solution
- Successful Stress Tests, Recovery & Destructive Tests on integration
- Performance
  - Scalability of 1.95 from a two-node to a four-node HP rp5470 cluster
  - On a 16-node HP cluster configuration, a 20% improvement in Oracle9*i* RAC OLTP throughput and response time using HMP over UDP

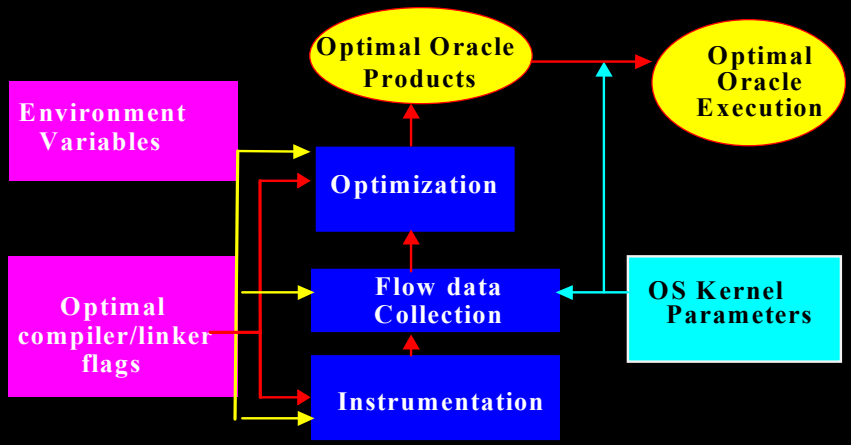
Enhancements contd..
Oracle9*i* RAC & HP HMP



Oracle9*i* (9.2) on HP-UX 11.22 (Itanium 2)

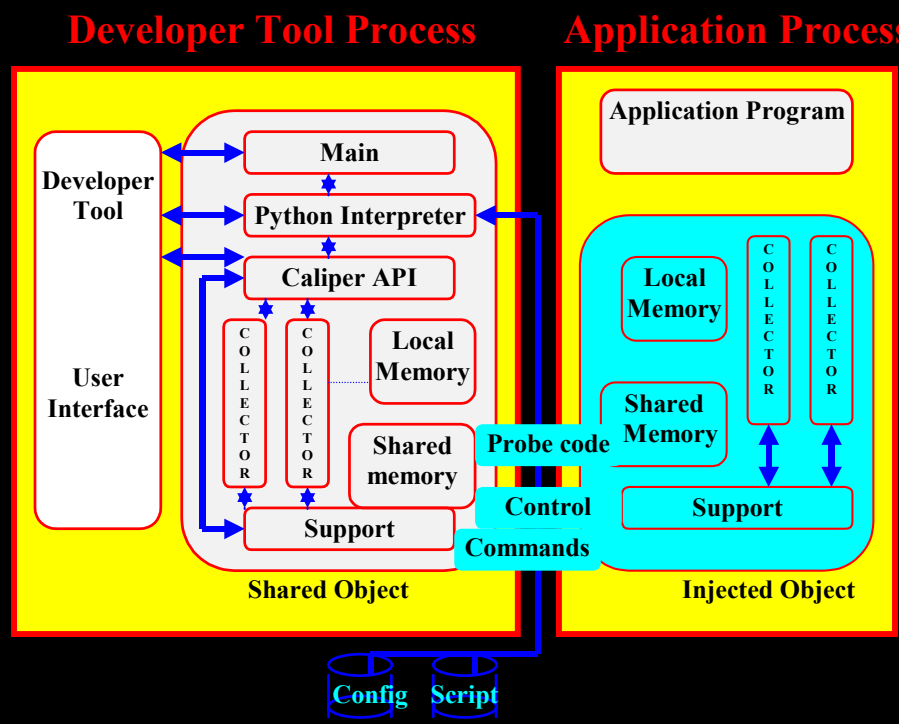
- EPIC concept -> IPF architecture – Itanium, Itanium 2
- HP Compiler Optimizations on IPF
  - Predication, Pipelining and Speculation
  - Better inlining of math functions
  - HP compiler optimizing scheduler for IPF allows code to be scheduled for Itanium and Itanium 2 based implementations
- Caliper-integrated dynamic instrumentation for Oracle RDBMS
  - Performed at run-time of an application program unlike static PBO-ing (Profile-based Optimization) on PA-RISC
  - Captures true run-time characteristics of RDBMS applications and provides useful analysis on the application behavior
- Enhanced compile-time tuning capabilities with HP IPF compilers
  - Profiling, scheduling, link mode choices for shared and archived libraries with *–dynamic, -exec, -minshared*

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)



Advantages for Oracle applications on IPF HP-UX 11*i*

- 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.
  - Lab tests show a 100% increase in Oracle9*i* (9.2) OLTP performance on a 4-way rx5670 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.

Advantages contd..

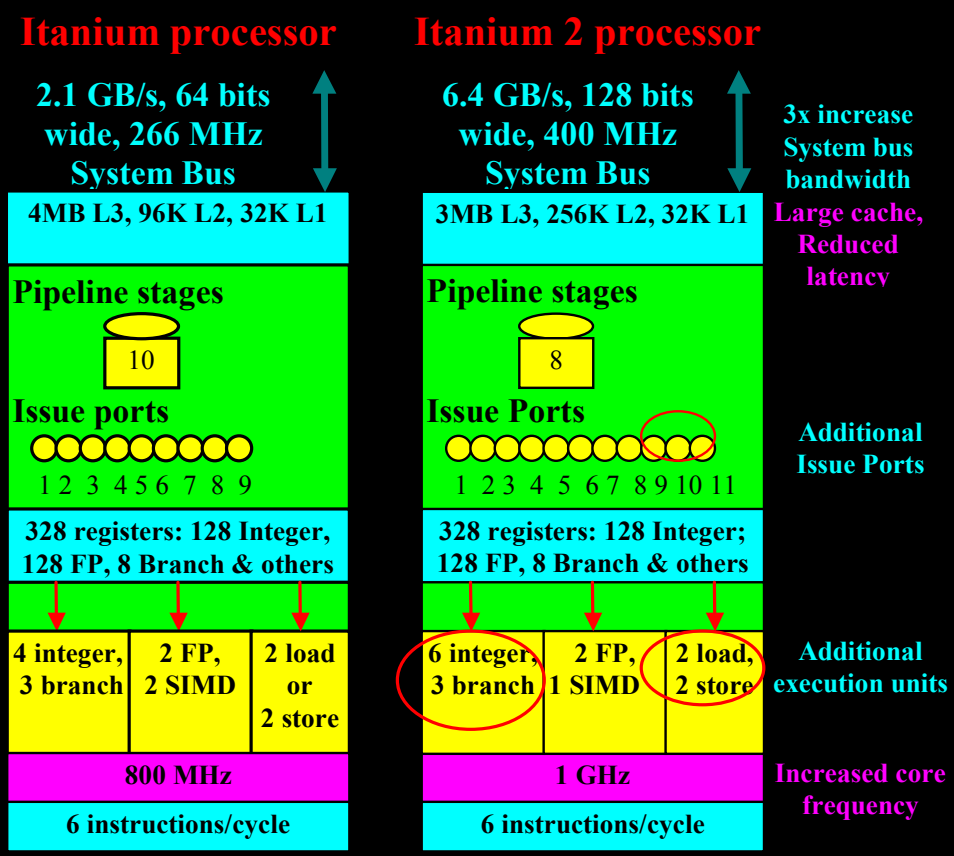
- Customization capabilities
  - Floating point computational performance of Itanium 2, rich data type support of Oracle9*i* (9.2) -> native support of XML data type – SQL operations on XML data, etc; LOBs, heterogeneous Streams
  - Personalized web information service, simulations, data analysis for 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 by Itanium – pipelined Multiply-Add instructions, register width, floating point registers

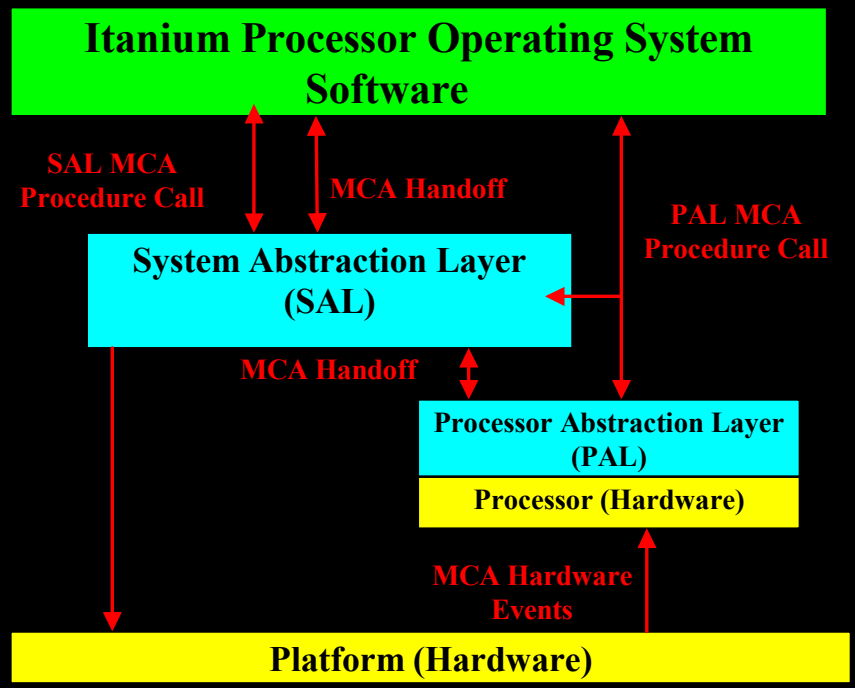
  Database-level security: Oracle9*i* (9.2) -> Label Security, Advanced Encryption security

Advantages contd..

Itanium vs Itanium 2 Chip Architecture



Error Handling by PAL and SAL on Itanium 2



Recent HP/Oracle Benchmarks

- **OLTP benchmark (TPC-C)**
  - Oracle9*i* (9.2)
    - On a 64-way PA-8700 (875 MHz) Superdome (HP-UX 11*i*)
      - Performance metric - 423K tpmc; price/performance - $15.64/tpmc
  - 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 - 824 K tpmc; price/performance - $8.28/tpmc
    - On a 64-way PA-8700 (875 MHz) Superdome (HP-UX 11*i*)
      - Performance metric - 541K tpmc; price/performance - $10.69/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

- **DSS benchmark (TPC-H)**
  - Oracle9*i* (9.2)
    - On a 64-way PA-8700 (875 MHz) Superdome (HP-UX 11*i*)
      - Performance metric - 27K QphH@3000GB; price/performance - $222/QphH@3000GB
      - Performance metric - 25K QphH@1000GB; price/performance - $213/QphH@1000GB

Recent HP/Oracle Benchmarks contd…

- **Oracle Applications Standard Benchmark**

  Oracle Applications Standard Benchmark 11.5.3

  - Oracle9*i* RAC database using HP's Hyper Messaging protocol (HMP) over HyperFabric II
  - 2,296 users on a 2-node rp5470 cluster to 4,368 users on a 4-node rp5470 cluster yields scalability of 1.95

  Oracle Applications Standard Benchmark 11.5.6 (Split Configuration)

  - Oracle9*i* (9.2) single-node result on a 4-way rx5670 server (Itanium2: Madison,1.5GHz, 6MB L3) running HP-UX 11.23 (Application Tier on PA-RISC based HP-UX 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 (Itanium 2, 1GHz, 3MB L3) running HP-UX 11.22 (Application Tier on PA-RISC based HP-UX 11.11)
    - 4200 users with an average response time of 1.51 secs

- **Two-tier SAP Standard Applications SD Benchmark**
  - Oracle9*i* (9.2) on a 4-way HP rx5670 server (Itanium2: 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

Future Oracle Releases on 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
- Continue enhancing Oracle RDBMS on PA-RISC based HP-UX 11*i*

References

- Paper “A Fast Track to Oracle9*i* Release 2 on HP-UX 11*i* “
  - HP World 2002 Conference Proceedings, Los Angeles, CA
- Paper “Oracle9*i* (9.2) on the IA-64 HP-UX 11*i* Platform”
  - HP World 2003 Conference Proceedings, Atlanta, GA
- Release Notes, Installation Guide, Administrator’s Reference for Oracle9*i* (9.2)
- 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