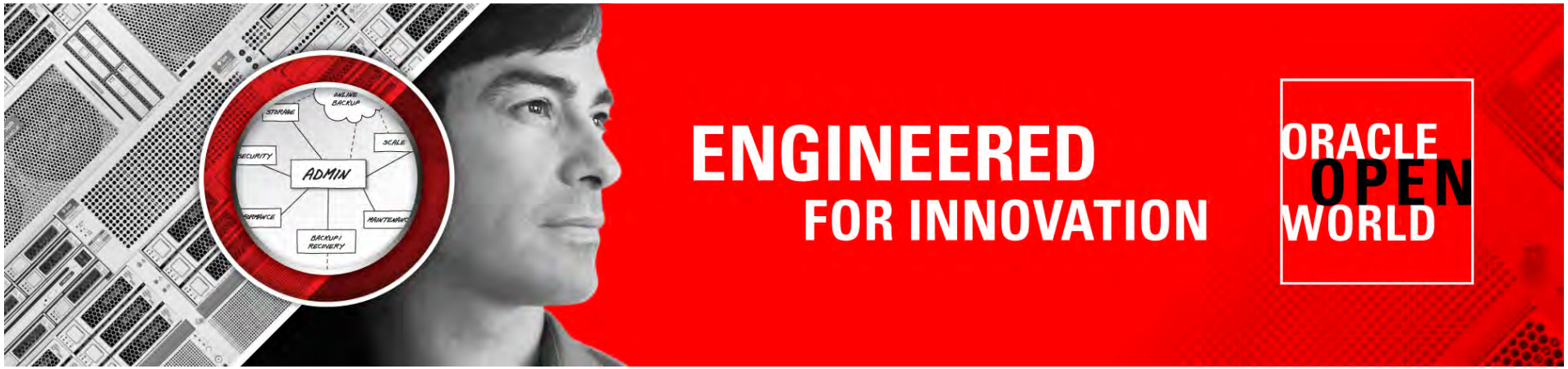
ORACLE®



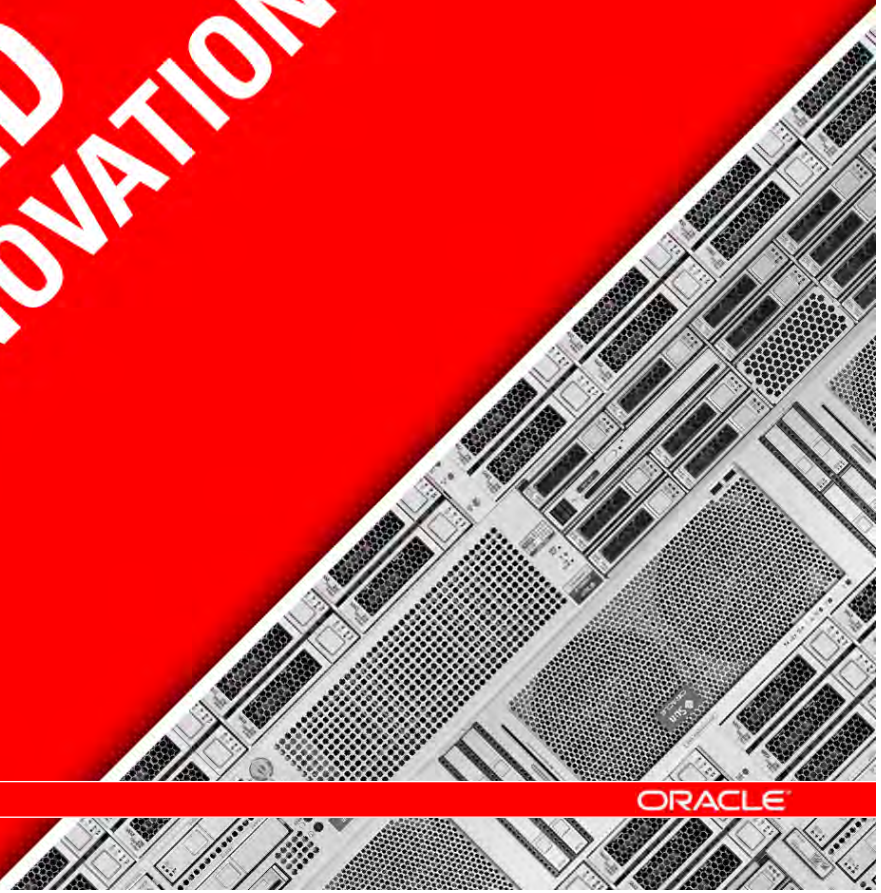
ORACLE®

Get Extreme Performance with Oracle Spatial on Oracle Exadata Database Machine

Jim Steiner Vice President, Server Technologies
Dan Geringer Senior Software Development Manager, Oracle Spatial
Steven Benn Manager, Santos Oil and Gas

ORACLE OPEN WORLD

ENGINEERED FOR INNOVATION



ORACLE

Latin America 2011

December 6–8, 2011

Tokyo 2012

April 4–6, 2012

Oracle OpenWorld Bookstore

- Visit the Oracle OpenWorld Bookstore for a fabulous selection of books on many of the conference topics and more!
- Bookstore located at Moscone West, Level 2
- All Books at 20% Discount

DigitalGuru
Technical Bookshop

ORACLE

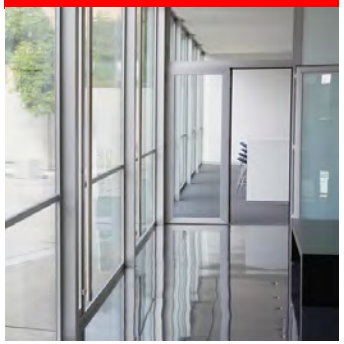
Oracle Products Available Online

Oracle Store

Buy Oracle license and support online today at oracle.com/store

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Program Agenda



- Exadata Database Machine
- Oracle Spatial Architecture
- Santos Oil and Gas Experience
- Best Practices for Oracle Spatial with Exadata
- Questions

ORACLE

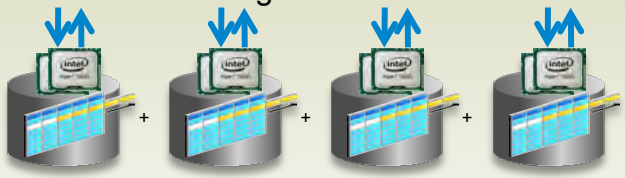
Exadata Database Machine

A complete system - compute, storage, networking

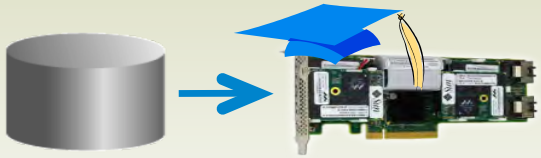
- Database Grid
  - Intel-based database servers
  - Oracle Linux or Solaris 11
  - Oracle Database 11g R2
  - 10 Gig Ethernet (to data center)
- Storage Grid
  - Intel-based storage servers
  - Up to 336 terabytes raw disk
  - 5.3 terabytes Flash storage
  - Exadata Storage Server Software
- InfiniBand Network
  - Internal connectivity (40 Gb/sec)

Key Exadata Innovations

- **Intelligent storage**
  - Smart Scan query offload
  - Scale-out storage

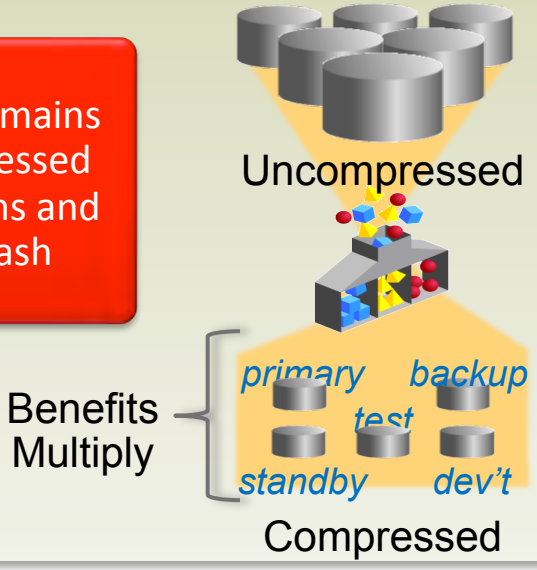


- **Smart Flash Cache**
  - Accelerates random I/O up to 30x
  - Doubles data scan rate



- **Hybrid Columnar Compression**
  - 10x compression for warehouses
  - 15x compression for archives

Data remains compressed for scans and in Flash



 | |

Oracle Spatial Architecture

Benefits on Oracle Exadata Database Machine

ORACLE

Benefits of Oracle Exadata for Spatial Workloads

- Only Oracle Spatial is designed for the Exadata architecture
- Breaks new boundaries for ingesting spatial data
- Data warehouse performance increases of up to 100x faster
- Box and distance queries up to 25x faster

ORACLE

Part of Oracle Database Kernel

- Code resides in Database Address Space
- Index operations performed in Database SGA
- Spatial operators are parallel enabled by default and execute in Database Address Space for maximum scalability
- Seamlessly exploit Database security, compression, partitioning services

Database Compression

- Exadata Hybrid Columnar Compression
  - Reduces storage required for Point data sets
  - Over 10x compression in high archive mode
- SecureFile LOB Compression / GeoRaster DEFLATE
  - Reduces storage required for Raster data sets
  - For Gridded data sets: 5 to 100x compression (data dependent)
  - For DEMs and Images: ~1.5x compression
- GeoRaster JPEG-F Compression
  - 2 - 15x compression lossy

ORACLE

High Performance Spatial Query and Index Operations

- First filter brings approximate results into SGA; exact results performed in memory
- Spatial operators performed in parallel against partitioned and non-partitioned data sets
- Spatial query analysis up to 100x faster
- Millions of spatial objects evaluated in minutes
  - Point in polygon analysis
  - Polygon to polygon analysis
  - Deviation from route
  - Distance covered

High Performance Database Load

- Massive ingestion of Spatial data from sensors, weather data, satellites, and other streams by government environmental and analysis agencies
- Millions of Spatial objects ingested in minutes
  - Weather readings
  - Traffic readings
  - Sensor readings

Customer Experience

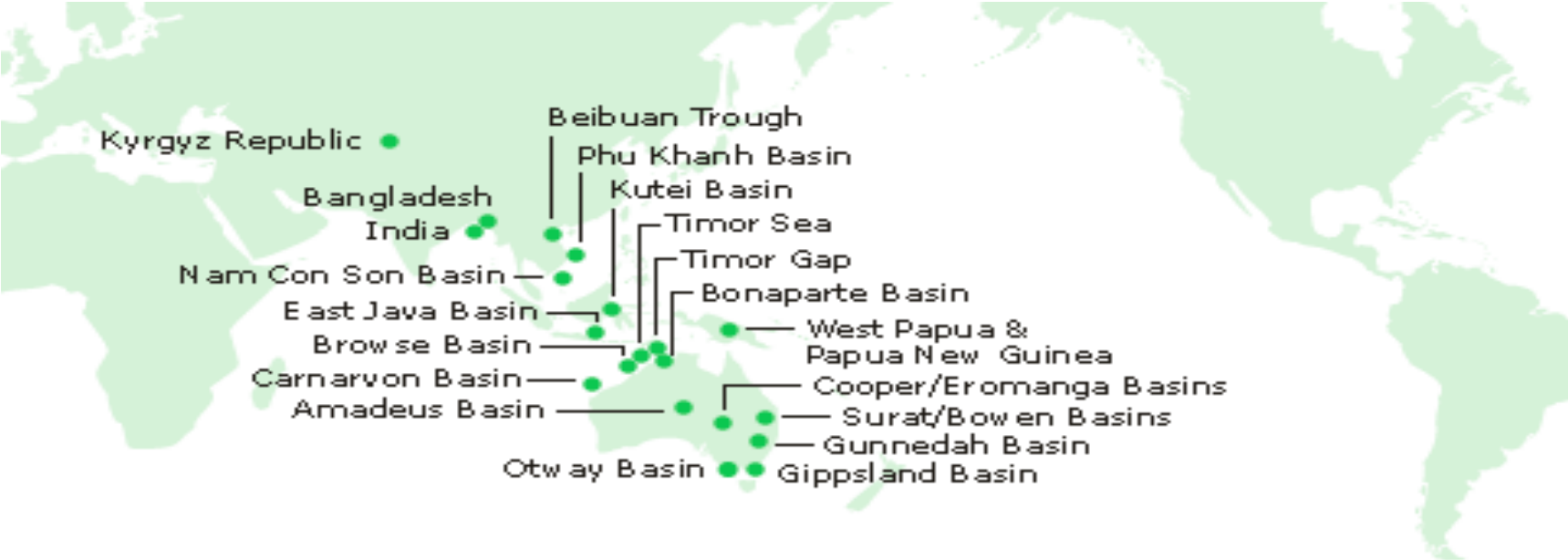
Santos Oil and Gas

ORACLE

About Santos

- Australian energy pioneer since 1954
- Leading gas producers, supplying Australian and Asian customers
- Santos' total production was 49.9 million barrels of oil equivalent
- 146,800 square kilometers exploration portfolio
- Supplies 16% of Australia total Natural Gas Market
- Annual Revenues of $2,181million
- 4500 employees
- Major Focus in Coal Seam Gas/Gladstone LNG ($16 Billion project)

About Santos – Areas of Operation



Current Challenges

- Conflicting Service Levels around Database
- Exploding data consumption
- Performance
  - Need to produce reports on probable reserves to the market. Current report taking 175 Hours to produce the report.

- Business Challenges
  - Huge demand for LNG in growing service area throughout Asia Pacific Region – China, Korea, etc.
  - Major investment in future Coal Seam Gas Business

POC Results

Type of Query or Operation	Type of Data	Time [existing system]	Time DB Machine	How many times faster? See Notes for system being compared to	Notes
Datafile creation	Datafiles	55 minutes	00:10:00	5.5X	Database 5 times greater !
Import DB	RMAN	9 hours	01:30:00	6X	
Allocate performance	Generate	55 minutes.	00:12:46	4X	
Reporting table build	Build	2.5 hours	00:20:14	7.5X	
Test 1	Reporting	6 seconds	00:00:01	6X	
ETL	Load & Translate	32 minutes	00:04:26	7X	
Index build		100 minutes	00:12:02	8X	
Gather stats		36 minutes	00:08:43	4X	
HCC test – query high	Biggest table	52460	840	62X compression	Spatial data
Materialized view creation		60 hours	02:42:00	21X	

Summary Existing System: Existing configuration M5000 8 x 2.150Ghz CPUs 128GB & FusionIO SSD

Summary DB Machine Configuration: Quarter Rack X2-2 (DB Server x 2, Storage Server x 3)

Overall Statement: Customer expectation was exceeded both for speed where Exadata outperformed a SSD configuration and Hybrid Columnar Compression space reduction

ORACLE

Results – In production for Spatial dataset

- 52 x compression of spatial data
- Average 7x increase in query performance
- Best case 90x increase in performance
- Key report drops for determining gas blend data from production information drops from from 16 days to run drops to less than 12 hours – 32x faster
- Estimated ROI on 2.7 Million dollar investment will be less than 6 months

"The decision to go Exadata was simple. The performance benefits we gained were compelling."

ORACLE

Best Practices

Oracle Spatial with Oracle Exadata Database Machine

Daniel Geringer
Senior Software Development Manager
Oracle's Spatial Technologies

ORACLE

What Is Exadata?

What Is the Oracle Exadata Database Machine?

- Oracle SUN hardware uniquely engineered to work together with Oracle database software
- Key features:
  - Database Grid – Up to 128 Intel cores connected by 40 Gb/second InfiniBand fabric, for massive parallel query processing.
  - Raw Disk – Up to 336 TB of uncompressed storage (high performance or high capacity)
  - Memory – Up to 2 TB
  - Exadata Hybrid Columnar Compression (EHCC) – Query and archive modes available. 10x-30x compression.

ORACLE

What Is the Oracle Exadata Database Machine?

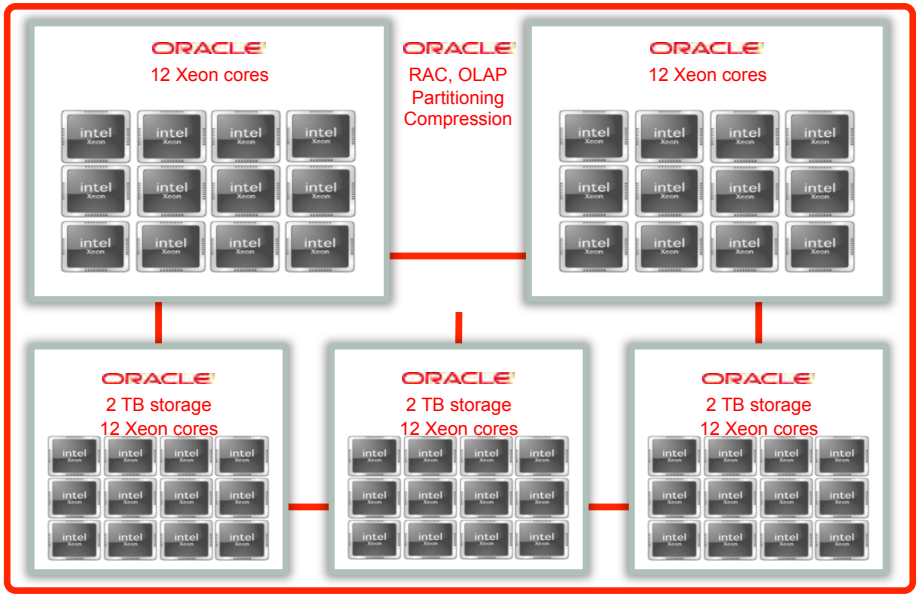
- Key features (continued):
  - Storage Servers – Up to 14 storage servers (168 Intel cores) that can perform massive parallel smart scans. Smart scans offloads SQL predicate filtering to the raw data blocks. Results in much less data transferred, and dramatically improved performance.
  - Storage flash cache – Up to 5.3 TB with I/O resource management

Exadata Database Machine Configurations

- X2-2 - configured as Quarter, Half or Full racks
  - X2-2 Quarter Rack – 24 database cores
  - X2-2 Half Rack – 48 database cores
  - X2-2 Full Rack – 96 database cores
- X2-8 – configured as a full rack
  - 128 database cores

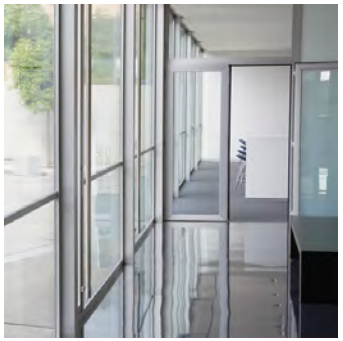
ORACLE

Exadata X2-2 Quarter Rack Diagram



 | |

Presentation Outline



- Best practice strategies discussed throughout the presentation
- Specific examples of:
  - Parallelizing spatial queries against partitioned tables
  - Parallelizing massive spatial computations with Create Table As Select (CTAS):
    - Parallelizing spatial operators
    - SDO_JOIN and parallelizing spatial functions
- Massive spatial ingest with the spatial index enabled.

Oracle Spatial Focused on Parallelizing Spatial Computations

- US Rail
  - Bulk nearest neighbor queries to find closest track, and project reported train positions onto tracks
- Validate home appraisals for a Government Sponsored Enterprise (GSE)
  - Find all the parcels touching parcels to validate appraisals
- Satellite Imagery Provider
  - Find all the useful portions of cloud covered imagery
    - 850,000 strip images
    - 58,000,000 cloud cover geometries

Parallelizing Spatial Queries Against Partitioned Tables

Parallel Spatial Query Against Partitioned Tables

- Partition pruning occurs first.
- If a spatial operator's query window spans multiple partitions, partitions are spatially searched in parallel.
- True for all spatial operators

ORACLE

Parallel Spatial Query Against Partitioned Tables

- Example:
  - A re-insurance company maintains portfolios for hundreds of insurance companies.
  - Which companies will be (or are) affected by the projected path of a hurricane.
  - 36 million rows, 64 partitions, each with about 571,000 rows.
    - 50 seconds serial
    - 1.28 seconds parallel (on a ½ rack Exadata V1 database machine)….. 39 times faster

Spatial Operators Can Parallelize with Create Table As Select (CTAS)

Parallelizing Spatial Operators
US Rail Application

- Requirement
  - GPS locations for each train collected throughout the day
  - Each location has other attributes (time, speed, and more)
  - GPS locations have a degree of error, so they don't always fall on a track.
  - Bulk nearest neighbor queries to find closest track, and project reported train positions onto tracks
- Information used for:
  - Tracking trains
  - Analysis for maintenance, ensure engineers are within parameters, etc…

ORACLE

Parallel and CTAS With Spatial Operators

- Spatial operators can parallelize with CTAS <u>when multiple candidates feed the second argument</u>

- For example, a GPS records 45,158,800 train positions. For each train position:
  - Find the closest track to the train (with SDO_NN)
  - Then calculate the position on the track closest to the train

- See SQL on next slide…

Linear Scalability to process 45,158,800 Train Positions (One SQL statement)

- X2-2 Half Rack, 34.75 hours serially vs. 41.1 minutes parallel
- X2-2 Half Rack (48 cores) over 47x faster … Linear scalability
- X2-2 Full Rack (96 cores) about 94x faster… X2-8 (128 cores) even faster

```
ALTER SESSION FORCE PARALLEL ddl PARALLEL 72;
ALTER SESSION FORCE PARALLEL query PARALLEL 72;

CREATE TABLE results NOLOGGING PARALLEL
AS SELECT /*+ ordered  index (b tracks_spatial_idx) */
                a.locomotive_id,
                sdo_lrs.find_measure (b.track_geom, a.locomotive_pos) measure
   FROM locomotives a,
          tracks b
   WHERE sdo_nn (b.track_geom, a.locomotive_pos, 'sdo_num_res=1') = 'TRUE';
```

Summary - Rail Application Results - Performance

- Results compare serial and parallel “Exadata” runs
- Serial runs on other hardware are slower than Exadata
- On Exadata:
  - X2-2 Half Rack - 34.75 hours serially vs. 41.1 minutes in parallel
  - X2-2 Half Rack (48 database cores) - 47x faster
  - X2-2 Full Rack (96 database cores) about 94x faster
  - X2-8 (128 cores) even faster

ORACLE

Summary- Rail Application Results - Compression

- EHCC (Exadata Hybrid Columnar Compression)
  - Can compress point geometries
  - 12.03 GB vs 1.125 GB EHCC Archive High Compression
  - 10.69X compression

ORACLE

Parallel and CTAS With Spatial Operators

- Works with all spatial operators:
  - SDO_ANYINTERACT
  - SDO_INSIDE
  - SDO_TOUCH
  - SDO_WITHIN_DISTANCE
  - SDO_NN
  - Etc…

ORACLE

Parallelizing Spatial Functions

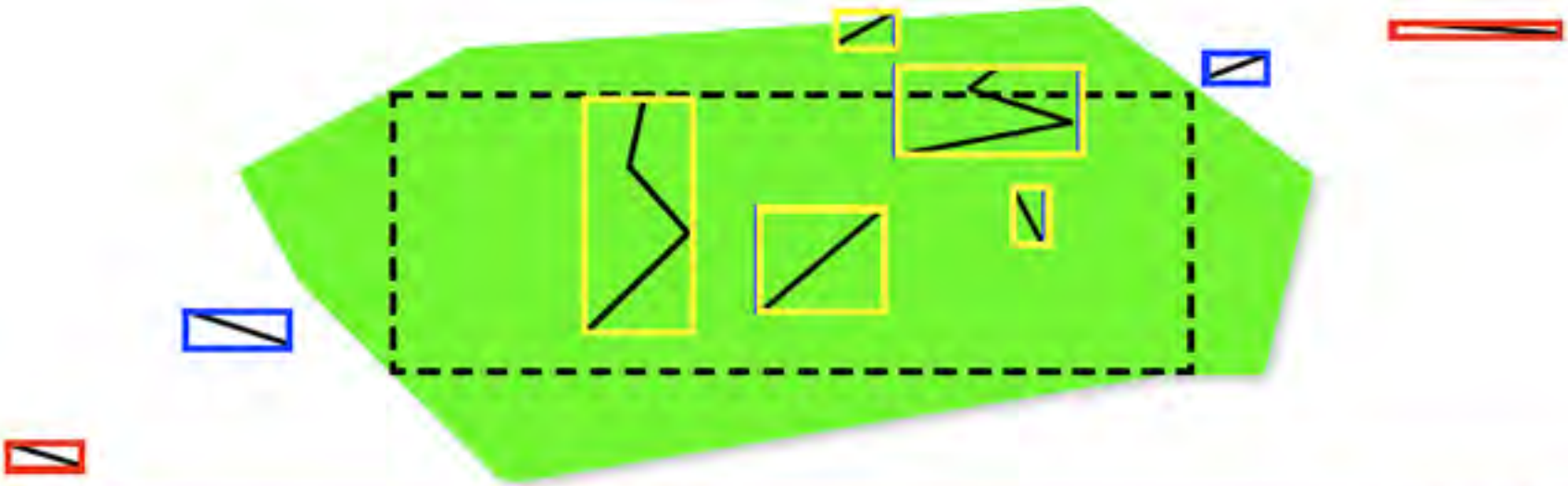
- Validate home appraisals for a Government Sponsored Enterprise (GSE)
  - Find all the parcels touching parcels to validate appraisals
- Satellite Imagery Provider
  - Find all the useful portions of cloud covered imagery
    - 850,000 strip images
    - 58,000,000 cloud cover geometries

ORACLE

Spatial Optimizations – Behind the Scenes

Which of the millions of roads in the U.S. have some interaction with this county?

- Primary filter compares geometry approximations
- Interior optimizations are applied to candidate set.
- Geometry comparisons are done only where required.



 | |

Spatial Operators

When possible

make the query window a polygon.

The SDO_JOIN Operation

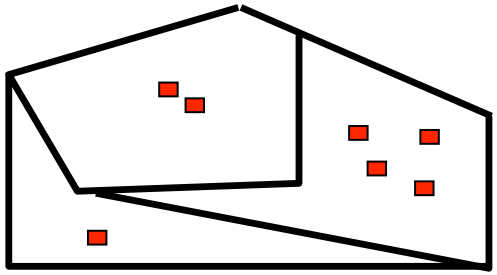
SDO_JOIN – Spatial Cross Product

- Effective way to compare all geometries in one layer to all geometries in another (or most to most)
- Leverages spatial index for both spatial layers
- Can be orders of magnitude faster

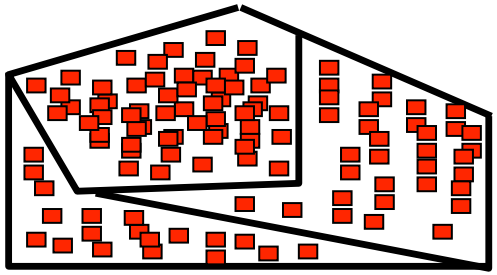
```
SELECT /*+ ordered */ b.risk_zone_id,
                      c.parcel_id
FROM TABLE ( SDO_JOIN ( 'RISK_ZONES', 'GEOM',  ' PARCELS', 'GEOM',
                        'mask=anyinteract')                          ) a,
        risk_zones b,
        parcels c
WHERE a.rowid1 = b.rowid
  AND a.rowid2 = c.rowid;
```

SDO_JOIN – When is it Most Effective?

- If one of the layers is a polygon layer:



- When not many geometries are associated with each polygon, SDO_JOIN may be much more effective



- When many geometries are associated with each polygon, SDO_ANYINTERACT may be more effective
- SDO_ANYINTERACT performs interior optimization

SDO_JOIN – Strategy For Parallelization

- First SDO_JOIN, primary filter only

```
CREATE TABLE result1 NOLOGGING AS
    SELECT a.rowid1 AS risk_zones_rowid,
           a.rowid2 AS parcels_rowid
    FROM TABLE ( SDO_JOIN ('RISK_ZONES', 'GEOM',
                           ' PARCELS', 'GEOM') );
```

SDO_JOIN – Strategy For Parallelization

- Then parallelize spatial function
- For this example, call sdo_geom.relate on each pair.

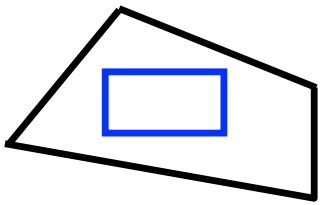
```
ALTER SESSION FORCE PARALLEL ddl PARALLEL 72;
ALTER SESSION FORCE PARALLEL query PARALLEL 72;

CREATE TABLE result2 NOLOGGING PARALLEL AS
SELECT /*+ ordered use_nl (a,b) use_nl (a,c) */
          sdo_geom.relate (b.geom, 'DETERMINE', c.geom, .05) relation,
          b.risk_zone_id,  c.parcel_id
FROM result1 a , risk_zones b,  parcels c
WHERE a.risk_zones_rowid = b.rowid
   AND a.parcels_rowid       = c.rowid;
```

Summary of Strategy – First SDO_JOIN, then SDO_GEOM.RELATE in Parallel

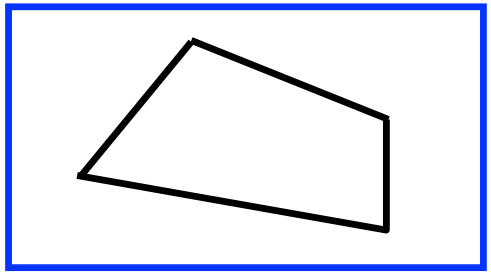
- SDO_JOIN –
  - Primary filter only, very effectively utilizes the spatial index of high risk zones and parcel polygons
  - Returns rowid pairs that likely intersect
- SDO_GEOM.RELATE with DETERMINE mask, in parallel:
  - Given two polygons, returns their relationship
  - Intersection is expensive
  - **Intersection is not always necessary**

OVERLAPBDYINTERSECT	2943738
COVERS	64
TOUCH	106
CONTAINS	1502223
DISJOINT	624854
COVEREDBY	6
INSIDE	661298
OVERLAPBDYDISJOINT	366

Further Optimizations



- High risk zone (in black) CONTAINS the parcel polygon



- High risk zone (in black) is INSIDE the parcel polygon

Strategy – First `SDO_JOIN`, then `SDO_GEOM.RELATE` in Parallel

- Satellite Imagery Provider
  - Find all the useful portions of cloud covered imagery
    - 850,000 strip images
    - 58,000,000 cloud cover geometries
    - 45 days to 4 days (with hardware constraints)
    - Exadata POC scheduled

Strategy – First `SDO_JOIN`, then `SDO_GEOM.RELATE` in Parallel

- Validate home appraisals for a Government Sponsored Enterprise (GSE)
  - Find all the parcels touching parcels to validate appraisals
  - Processed 2,018,429 parcels serially in 38.25 minutes
  - Exadata ½ RAC (48 cores) – 45x faster - 50 seconds
  - X2-2 Full RAC (96 cores) about 90x faster
  - X2-8 (128 cores) even faster
- Satellite imagery requirement from previous slide
  - 45 days to 4 days (non-Exadata with hardware constraints)
  - From 4 days to hours or minutes on Exadata

ORACLE

Massive Spatial Ingest With Spatial Index Enabled

High Ingest Rate – Spatial Index Enabled Strategy

- For OLTP applications that need to insert into tables with the spatial index enabled
- Eliminate spatial index contention. Not always necessary, but really helps if ingest rate is very high
- To eliminate spatial index contention, partition table by time and process id (composite partition key)
  - Assign process id's to Java pool connections
  - Each connection only writes to the partition with the same process id
- Partitioning is transparent to the SQL developer.
- Queries are written against 1 table, Oracle manages which partitions to search.

ORACLE

High Ingest Rates – Eliminate Contention

- Example is hourly, but could be quarterly (every 3 months)
- Time portion of composite key is a virtual column

```
CREATE TABLE composite_example
  ( t              timestamp
  , process_id     number
  , geom           sdo_geometry
  , hour_partition as (substr(t,1,12)) )
  PARTITION BY RANGE ( hour_partition, process_id)
   (PARTITION DAY1_H5_1  VALUES LESS THAN ('30-NOV-10 05',2),
    PARTITION DAY1_H5_2  VALUES LESS THAN ('30-NOV-10 05',3),
    PARTITION DAY1_H5_3  VALUES LESS THAN ('30-NOV-10 05',4),
    PARTITION DAY1_H6_1  VALUES LESS THAN ('30-NOV-10 06',2),
    PARTITION DAY1_H6_2  VALUES LESS THAN ('30-NOV-10 06',3),
    PARTITION DAY1_H6_3  VALUES LESS THAN ('30-NOV-10 06',4),
    PARTITION REST   VALUES LESS THAN (MAXVALUE,MAXVALUE));
```

What We Tested

Inserting With Spatial Index Enabled (Implemented These Recommendations)

- For fastest ingest, each partition written to by only one writer.
- Batch inserts are critical to performance:
  - FORALL bulk inserts (if PL/SQL) … benchmark does this
  - JDBC Update Batching (if Java)
- Set sdo_dml_batch_size = 15000 (this really helps)
  - Benchmark commits every 15000 inserts
- Increase SGA and Log file sizes
  - May not be necessary, if resources are not saturated
  - For benchmark, SGA set to 32 Gb, and each log file to 16 Gb.

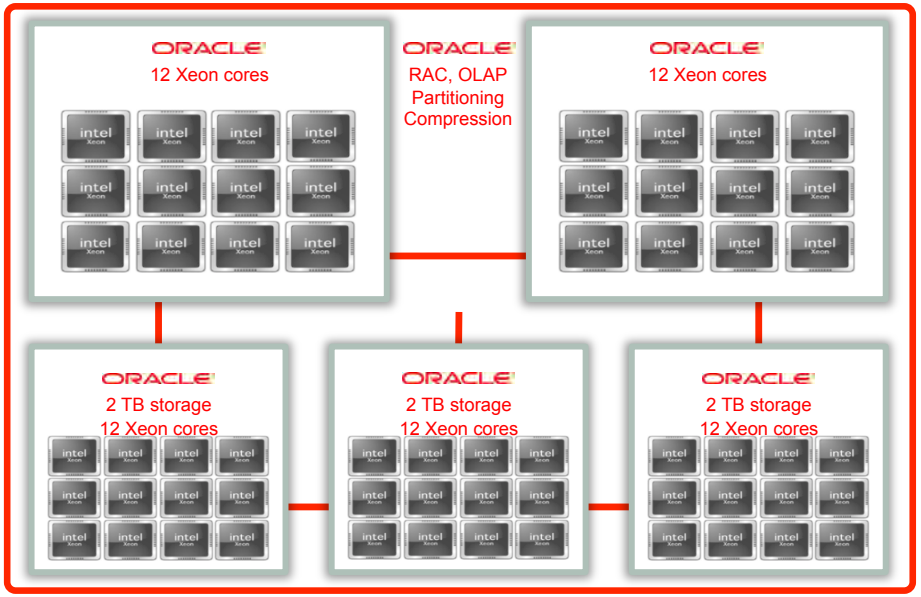
ORACLE

How the Benchmark Works

- Data generated dynamically in PL/SQL
- You control the number of parallel processes to run
- All processes are clones of each other
- Each process gets passed a process id
- Each process insets into an exclusive table (simulating an exclusive partition), for example:
  - Process 1 inserts into tracks_1 (spatial index enabled)
  - Process 2 inserts into tracks_2 (spatial index enabled)
  - Etc..
  - All tables start empty with a spatial index enabled
- You can specify on which node in the RAC to run the process, or use the scan listener to decide.

ORACLE

Exadata X2-2 Quarter Rack Diagram



 | |

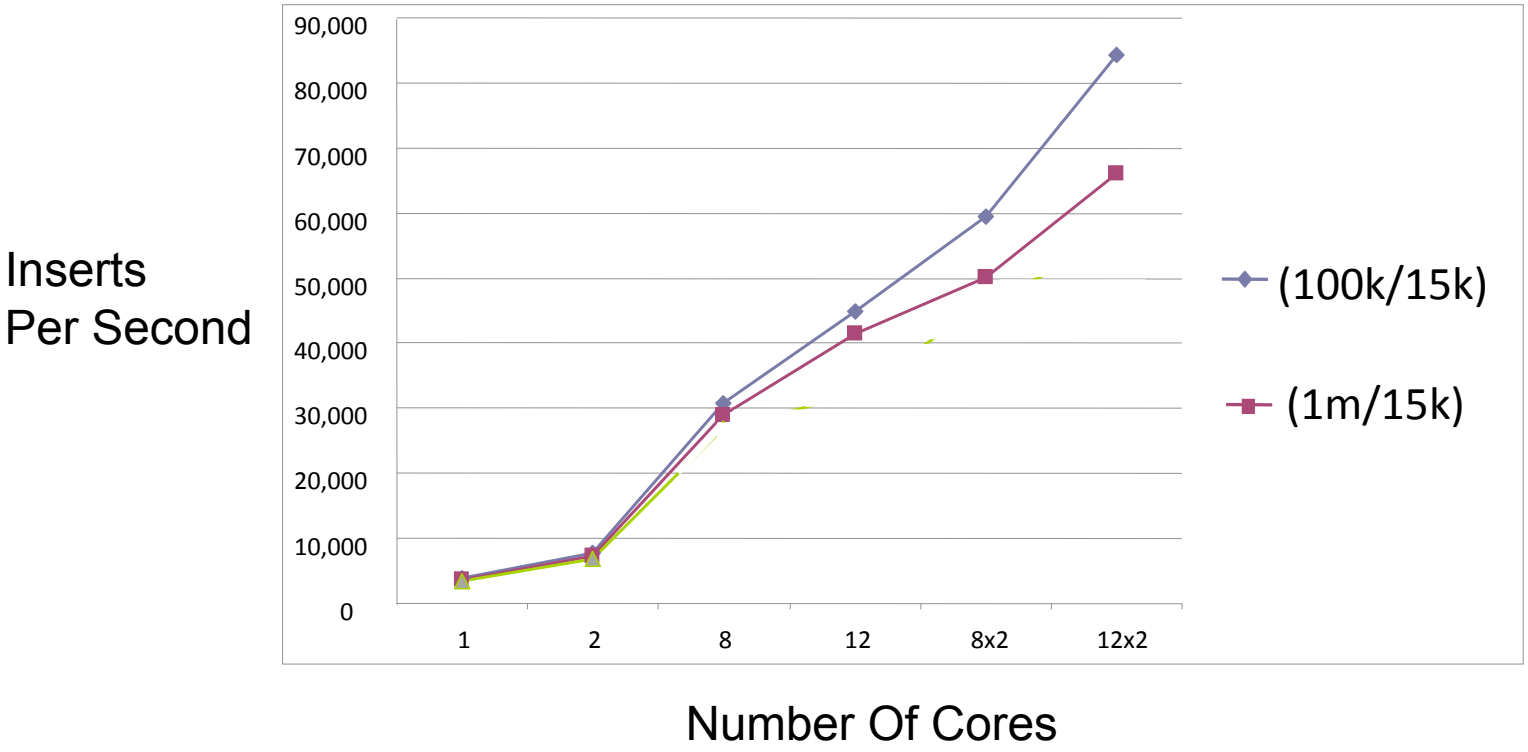
Exadata ¼ RAC Results - Spatial Index Enabled

- Parallel Degree - how many processes are running in parallel
- Each process writes to an exclusive table.
- 1m/15k means each process inserts 1,000,000 rows and commits every 15000 (with sdo_dml_batch_size=15000)
- 8x2 means 16 processes, 8 running on node 1 and 8 on node 2

Parallel Degree	Inserts/Sec (100k/15k)	Inserts/Sec (1m/15k)
1	3,901.92	3,701.10
2	7,646.74	7,352.68
8	30,757.95	28,870.05
12	44,843.68	41,557.92
8x2	59,547.36	50,237.07
12x2	84,309.38	66,123.57

ORACLE

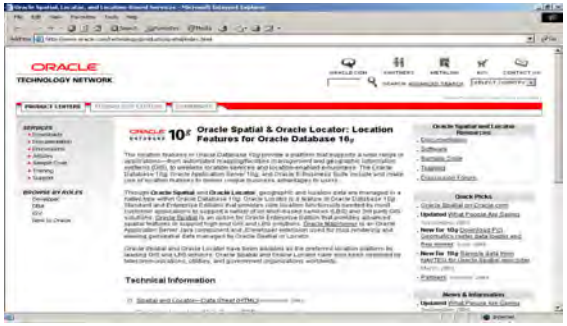
Exadata ¼ RAC Results

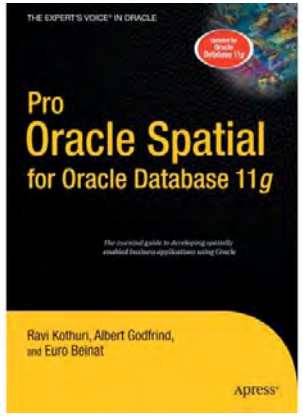


Find out more...

oracle.com/database/spatial.html

oracle.com/technology/products/spatial





oracle.com/technology/products/spatial/htdocs/pro_oracle_spatial.html

Q&A

ORACLE

Hardware and Software

ORACLE

Engineered to Work Together

ORACLE®