Oracle OpenWorld 2009

Report from a developer's perspective

Giuseppe Lo Presti, IT-DM-DA

DM Technical Meeting, November 17th, 2009

- General remarks

- Highlights from most interesting talks
  - Tom Kyte
  - Steve Feuerstein
  - Jonathan Lewis
  - Bryn Llewellyn
  - ...

- Conclusions

General remarks

- Very large "machinery" for logistics
  - Large number of parallel sessions (~1,800)
  - Huge number of attendees (> 40,000)
  - Huge celebration events...

- However, commercial ~~conference~~ show
  - Plenary talks only marketing oriented
    - And the important message was "we are the best"...
  - Which meant not so much time left for the many interesting technical sessions

- What follows is a summary of the most relevant technical sessions

"Why you should care about Oracle 11g *now*"

- Talk about <u>a</u> new feature – the Function Result Cache – which we could profit from even in 10g.
  - The FRC is a *shared* cache of function results
  - Outperforms recomputing a function's result each time if the function relies on relatively **static** data
- Main message: in writing PL/SQL code it is good practice to:
  - Split data access from application logic
    - To the limit of not having a single SELECT statement into PL/SQL procedures that deal with the business logic!
  - Then optimize each single data access function on its own
    - Eventually caching data in a SQL-table in memory
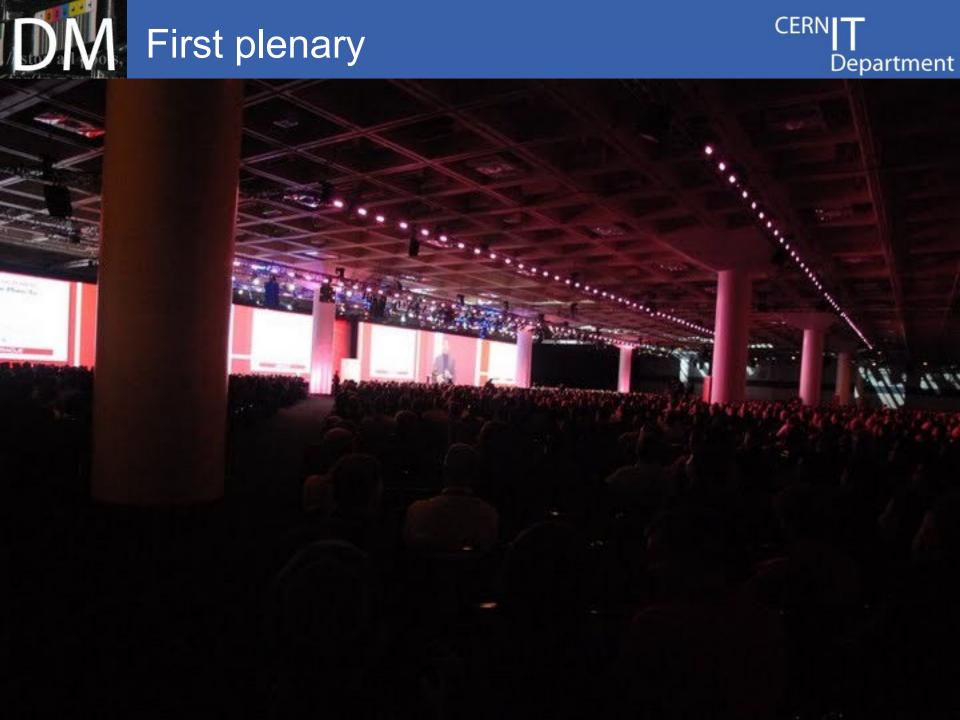   - In the future using the result cache...

SQL Developer 2.1 demo

- New features incorporated in SQL Developer. Highlights:
  - Unit testing / code coverage tests
  - Data viewer includes filters and hiding columns
  - Flashback: a viewer for the undo log
    - Including an Undo SQL statement we could run to restore a previous state
  - dbdoc: ability to generate HTML doc out of a schema
  - Graphical data modeler
  - …
- [Ads] sqldeveloper.oracle.com for more details...

How to get 10x improvements from PL/SQL

- Demo: a "standard" 3-tier application profiling
  - 12ms resp. time @ 83.5 TPS with Oracle 11g
  - 1.2ms resp. time @ 836.1 TPS (!) with TimesTen DB
- Pure in-memory DB
  - Less paths to take into account when retrieving data
  - No latencies with remote I/O
    - But HA support means at least some network overhead...
  - IMDB Cache builds on the same technology: a cache run in front of a standard Oracle 11g DB.
- Still missing
  - Triggers, autonomous transactions
- Just implemented (11gR2)
  - OCI and Pro*C APIs

Telling the schema about our constraints matters

- A nice dev-oriented talk with a number of recommendations:

  - 1. **Datatypes are constraints.**
    E.g. using numbers or strings instead of date types could mislead the optimizer: a BETWEEN 20001231 AND 20010101 clause may lead to FTS!

  - 2. **Length matters.**
    VARCHAR2(n) with n >> what's needed is a significant waste once you multiply by avg # of opened cursors and # of connections

  - 3. **Check constraints.**
    Telling Oracle about only allowed values on a given column makes the Optimizer filter out data that by constraint cannot match some given WHERE clauses. To see this in effect the execution plan includes a FILTER line with an always false condition (typically NULL IS NOT NULL).

- 4. **Not NULL constraints.**
  Typical example: SELECT count(*) translates into a FTS if available indexes are all on nullable columns

- 5. **FK constraints.**
  A not obvious observation: a
    SELECT A.x FROM A, B WHERE <join condition>
  is resolved by the optimizer in:
    SELECT x FROM A WHERE <FK> IS NOT NULL

- 6. **Dimensions.**
  Pure metadata objects which describe how to "roll-up" data, e.g. when date columns need to be aggregated by months, quarters, years.

Hints on Hints

- How to use hints – the advanced way
  - *I was late but most examples are available on his slides and at* jonathanlewis.wordpress.com *(his blog)*

- Message: DON'T use hints (?)
  - Now if you really want – then use two hints per table in your query!
    - What is the number of theoretical execution plans on a 4-table join when each table has 2 indexes? O(800K)!

- Among others...
  - USE_HASH and [NO_]SWAP_HASH_HOIN
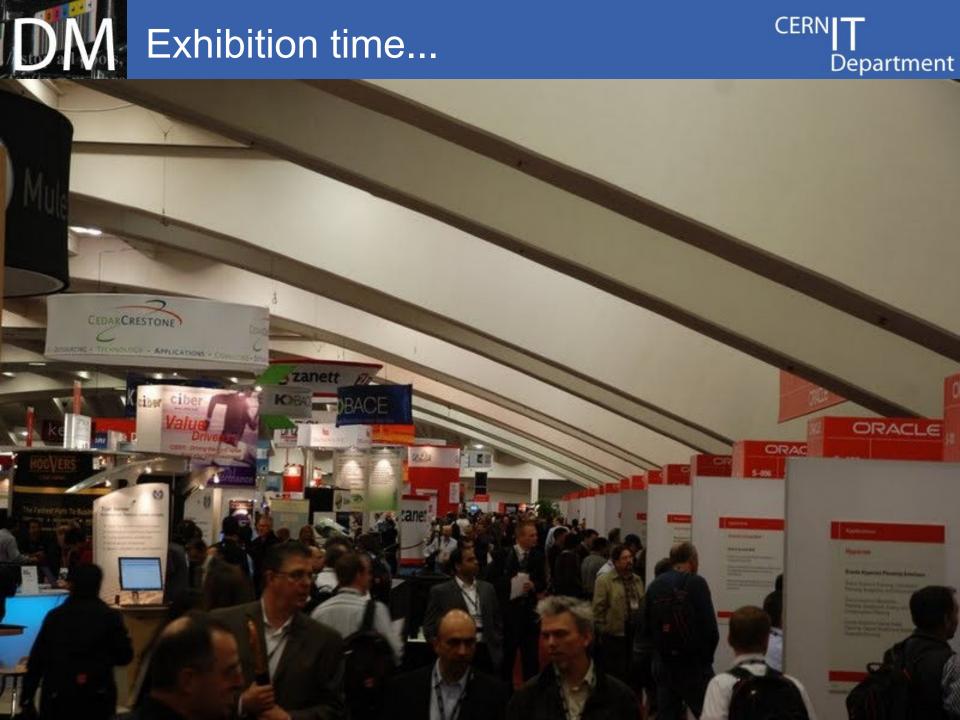 - Oracle 11gR1 has 236 hints...

PL/SQL Hierarchical Performance Profiler

- A new tool made available in 11gR2 to measure:
  - Time spent by subprograms/SQL statements
  - Time spent in each called subprogram
  - Time spent by the callers

- Scenario: pure dev, no production. You run:
  ```
  BEGIN
    DBMS_HProf.start_profiling('<DIR>', 'my_run.trc');
    my_procedure();
    DBMS_HProf.stop_profiling();
  END;
  ```
  - Then, run plshprof on the generated .trc file to get a set of HTML pages with the calling graph of the code
    - *I can give more details if interested...*
  - Note that this is not about profiling a running session with 'real' data, you need to provide everything!

- "Data Intensive" Scientific Applications session
  - Technical contributions by Oracle R&D people
- Maria (CERN)
  - *OK I'll skip that...*
- T.Frazier (LLNL), the National Ignition Facility
  - Event driven architecture: each step from calibration to config to analysis depends on each single event
  - XML is used to glue components together
  - Challenges: end-to-end monitoring
- D. Gawlick (Oracle), Complex Event Processing
  - Proposal for <u>extending</u> SQL to recognize "patterns" as reg-exps, defining rules and events
  - Prototype developed (patient care), looking for another important use case...

- K. Jernigan (Oracle), Compressed File Systems:

  new file-oriented features with Oracle 11g

  – SecureFile LOB type to handle files 'natively'

  – DBFS to provide direct filesystem interface to data stored into the db (either files or structured data)

    - A fopen() may trigger a SQL procedure, which "materializes" the file's content by e.g. running a SELECT

    - Basically, DBFS is an **updatable fs-oriented view** over the db

  – "Total Recall": full versioning support

    - Supports queries with clauses like
      WHERE file_revision = any time in the past

  – DBFS HSM Store: older files are archived to tape, with a disk staging area for storing recently accessed files

- Clients

  – FUSE on Linux + custom OCI-based CLI

- K. Jernigan (again), Exadata storage
  - Column-based compression, not gzip
    - 10x in warehouse mode, 15-70x in archive mode
  - Hybrid Columnar compression introduced in 11g
    - Single row lookup time comparable to uncompressed data
  - Support for mixed modes on different partitions
    - Older/cold data can be stored in archive mode
- Exadata delivers compression integrated at all levels
  - Including data on RAM!
    - Query processing optimized to take the columnar format into account
- Example: use case from the financial domain
  - Compression ratio: 11x
  - Query performance: 40x faster
  - Load test: Data pump takes 28% more time

Top 10 – no, 11 – new features of Oracle 11g

- Another nice talk about new features for devs

  - 1. **Parallelism**.
    A DBMS_Parallel_Execute package allows e.g. to split the execution of a PL/SQL procedure in chunks by rowid

  - 2. **Analytics** - "The coolest thing to happen in SQL after SELECT".
    Support for returning data of a given aggregate all together

    ```
    SELECT count(*), c,
        listagg(detailField, ',') WITHIN GROUP [ORDER BY detailField]
        FROM MyTable
        GROUP BY c;
    ```

  - 3. **Execute on directories**.

  - 4. **Recursive subquery** factoring. ANSI SQL replacement for CONNECT BY.

Cont'd

- 5. **Improved time travel**.
  We can now query at any time in the past (in the undo logs?) with a costant cost not dependent on the time

- 6. "You've got mail".
  File Watchers to execute code when a file gets in a given location (cf. DBFS)

- 7. Deferred segments creation.

- 8. **Flash Cache.**
  Support for SSDs as 2nd level cache

  - Dirty blocks directly go to disk

- 9. Parallel improved.

- 10 & 11. **Edition-based redefinition**.
  Quoting Tim: "this is THE killer feature of Oracle 11g, that's why it deserves two features!".

Oracle Partitioning in 11g

- New features and how-to use partitioning in 11g

- Highlights

    - "Partly" unusable local indexes are now supported

    - Improved statistics gathering

        - AUTO_SAMPLE_SIZE introduced

        - If a single partition changes, statistics can incrementally be updated without being regenerated for the whole table

- Exadata V1
  - Fastest db for data warehousing
- Exadata V2
  - Fastest db also for OLTP and random I/O
  - Why? because of 400GB RAM + 5TB flash storage + 40 Gbps Infiniband links...
    - => O(1M) random I/O ops/sec
  - Costs: O(2M$), to compare with IBM at O(10M$) - which by the way does not deliver a fault tolerant configuration.
- Cost comparison per TB
  - Oracle is at 5.7K$/TB
  - IBM Teradata or Netezza are at O(20K$/TB)

Plan stability and baselines

- A dev/DBA talk on a rather hot topic...

- So far (10g): stored outlines
  - But they can become suboptimal

- With 11g: baselines
  - The plan is fixed, however the optimizer calculates new plans and uses them IFF better && "accepted"
  - A baseline is a collection of plans for a given SQL statement
  - Plans stay in a baseline even after the SQL is flushed from the buffer cache
    - The SQL Management Base effectively is a metadata repository for all available statements, baselines and SQL profiles
  - Can be enabled at SESSION and SYSTEM level

- Usage and caveats
  - To see baselines:
    DBMS_Xplan.Display_sql_plan_baseline(
    sql_handle=>..., format=>'basic note');
    - the Notes will mention if a plan from a baseline is used
  - To populate baselines, we can load plans from the current cursor cache
  - To manage them we have:
    DBMS_SPM.alter_sqlplan_baseline();
  - And still, if for the very same query we want different plans depending on bind variables' values
    (e.g. on skewed data distributions)...
    - This was only covered during Q&As – and the answer is: enable adaptive cursors (only in 11g)
      - The author wrote an article on it, first hit in google...
        - http://www.oracle.com/technology/pub/articles/oracle-database-11g-top-features/11g-sqlplanmanagement.html

Conclusions

- Overall good experience from a technical standpoint
  - I don't want to comment on the rest...
  - I followed a couple more sessions, which were (too much) more marketing oriented

- (Most) presentations are available at: http://openworld.vportal.net/index.cfm
  - username = 'lopresti', ask me for the password...

- Questions?