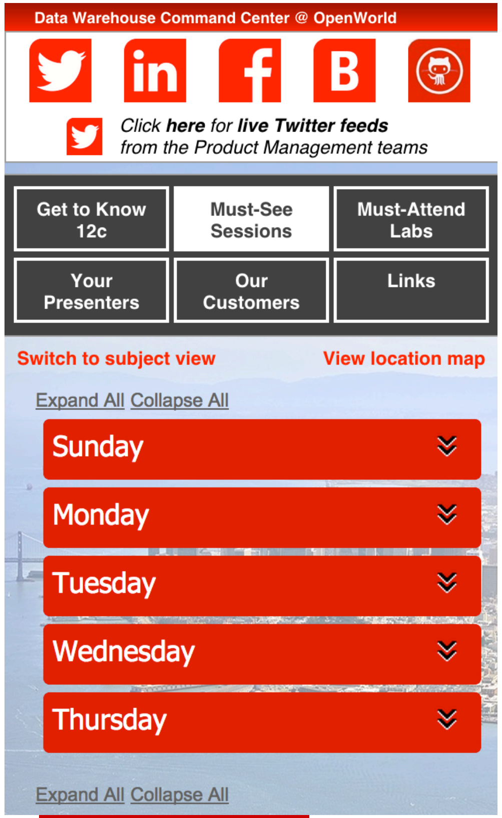
DW & Big Data on your smartphone



Data Warehouse Command Center @ OpenWorld

*Click **here** for **live Twitter feeds** from the Product Management teams*

Smartphone app helping you get the most from this year's OpenWorld

Access to all the most important information

- Presenter profiles
- Must-see sessions
- Must-attend hands-on labs
- Useful links

http://tinyurl.com/kmbsxbu

ORACLE®

Using Analytical SQL to Intelligently Explore Big Data

September 28–
October 2, 2014
San Francisco

ORACLE OPEN WORLD

Joerg Otto, Head of Database Engineering
IDS GmbH - Analysis and Reporting Services

Marty Gubar
Director Product Manager

Keith Laker, Oracle
Senior Principal Product Manager

ORACLE

Safe Harbor Statement

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.

Agenda

1 Analytical SQL – simplicity and sophistication

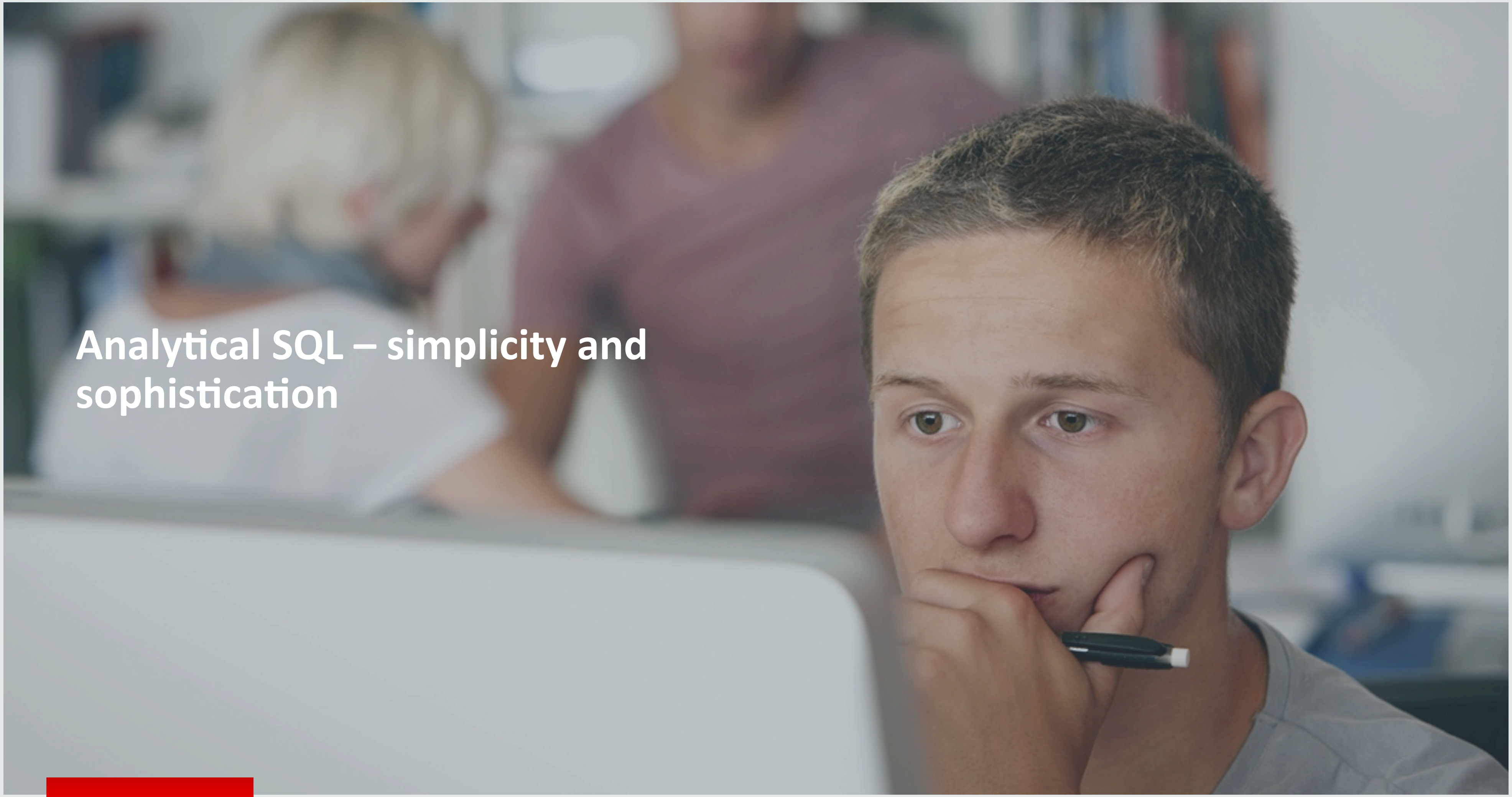
2 Using analytical SQL to solve complex business problems

3 12c – faster and smarter

4 Big Data SQL – analyze all your data

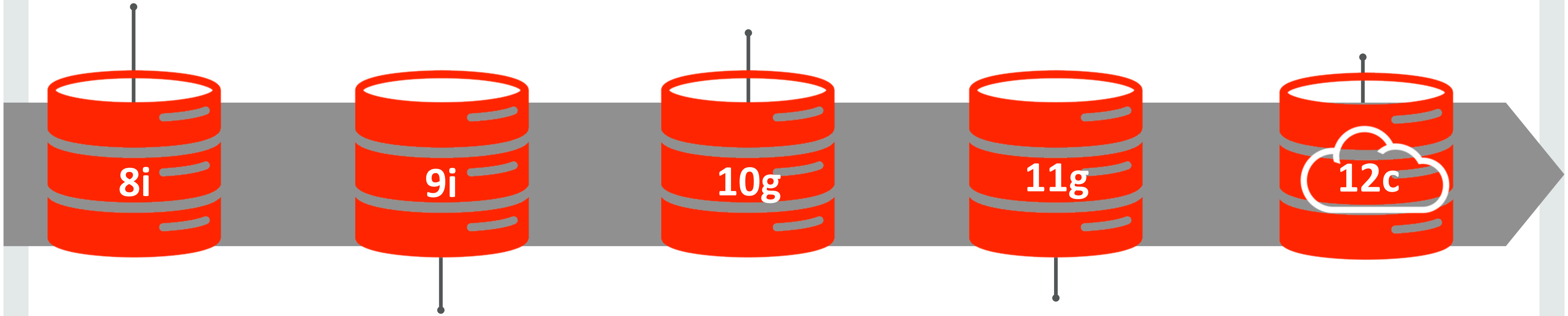
5 Summary

ORACLE



Analytical SQL – simplicity and sophistication

The On-Going Evolution of SQL



- Introduction of “window” functions

- Enhanced window functions (percentile, etc)
- Rollup, grouping sets, cube

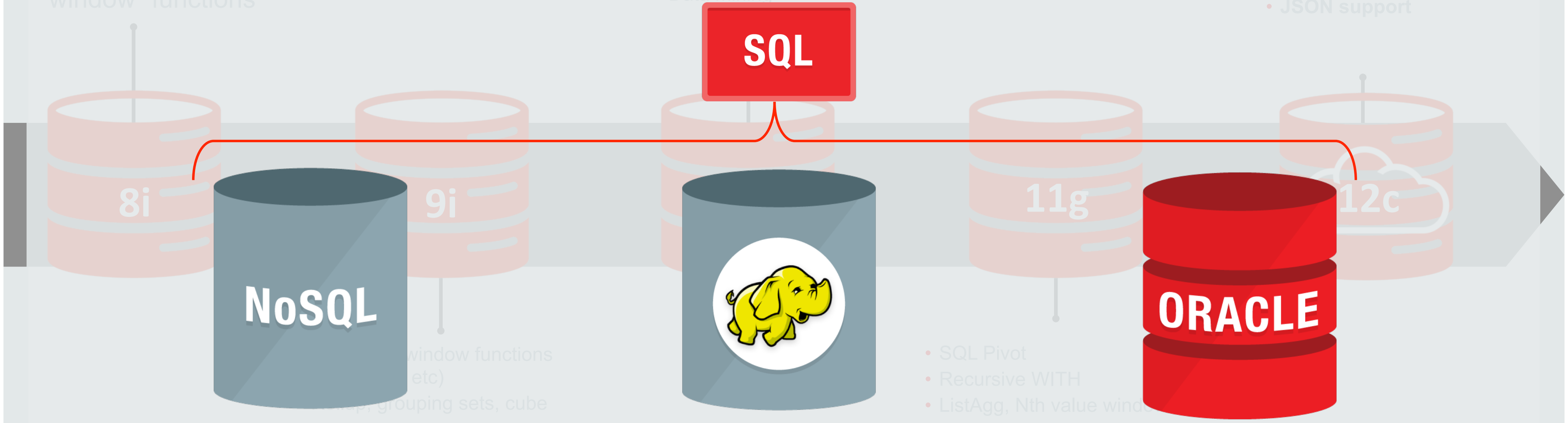
- Statistical functions
- SQL model clause
- Partition Outer Join
- Data mining

- SQL Pivot
- Recursive WITH
- ListAgg, Nth value window

- **Pattern matching**
- Top N clause
- **Approx Count distinct**
- **JSON support**

The On-Going Evolution of SQL

Oracle Big Data SQL



Using analytical SQL to solve complex business problems

Joerg Otto,
Head of Database Engineering
IDS GmbH –
Analysis and Reporting Services

Allianz Group At A Glance

- Insurance, Asset Management and Banking
- Founded 1890, headquarters in Munich / Germany
- 78 Million Customers in more than 70 countries
- > 147.000 Employees in (2013)
- € 110 Billion Revenue (2013)
- € 6,3 Billion Net income (2012)

Analytical Functions – How we use it

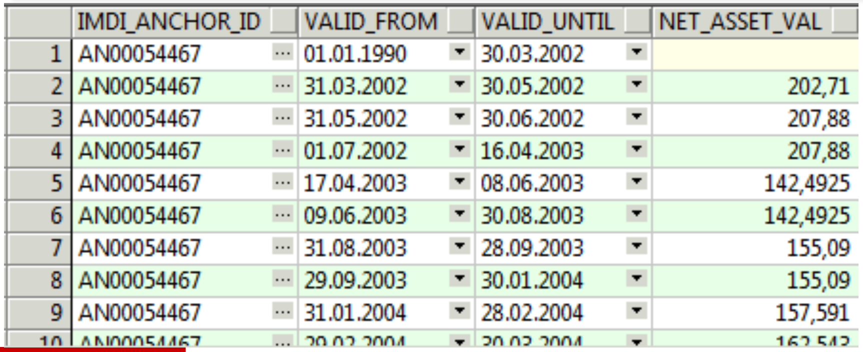
- What is most relevant investment metadata information?
  - Rank
- What is the 3-month moving average of stock price?
  - Moving Window
- What is the percentage growth of 2010 premiums written over 2009?
  - Period-over-period comparisons
- What are January's net income as a percentage of the entire year's?
  - Compare aggregates on different levels
- What are deviations between delivered and market data?
  - with_bucket, standard deviation
- What is the ratio of an e.g. investment's market value in a fund?
  - Share holdings calculations

Analytical Functions – Types

- LAG/LEAD functions
  - Direct inter-row reference using offsets
- Ranking
  - cume_dist, rank, dense_rank, percent_rank, ntile
- Reporting Aggregate
  - sum, avg, min, max, variance, stddev, count, ratio_to_report
- Statistical Aggregates
  - correlation, linear regression, covariance
- Window Aggregate
  - min, max, count, avg, sum, variance, stddev, first_value, last_value

Analytical Functions – LEAD example

```
SELECT imdi.imdi_anchor_id,
       imdi.imdi_valuation_date AS valid_from,
       LEAD(imdi.imdi_valuation_date)
            OVER(PARTITION BY imdi.imdi_anchor_id
                 ORDER BY imdi.imdi_valuation_date) - 1 AS valid_until,
       imdi.imdi_f_asset_val_net AS net_asset_val
  FROM ids.inmd_market_data_inv imdi
 WHERE imdi.imdi_anchor_id = 'AN00054467'
```



	IMDI_ANCHOR_ID	VALID_FROM	VALID_UNTIL	NET_ASSET_VAL
1	AN00054467	01.01.1990	30.03.2002	
2	AN00054467	31.03.2002	30.05.2002	202,71
3	AN00054467	31.05.2002	30.06.2002	207,88
4	AN00054467	01.07.2002	16.04.2003	207,88
5	AN00054467	17.04.2003	08.06.2003	142,4925
6	AN00054467	09.06.2003	30.08.2003	142,4925
7	AN00054467	31.08.2003	28.09.2003	155,09
8	AN00054467	29.09.2003	30.01.2004	155,09
9	AN00054467	31.01.2004	28.02.2004	157,591
10	AN00054467	29.02.2004	30.03.2004	162,543

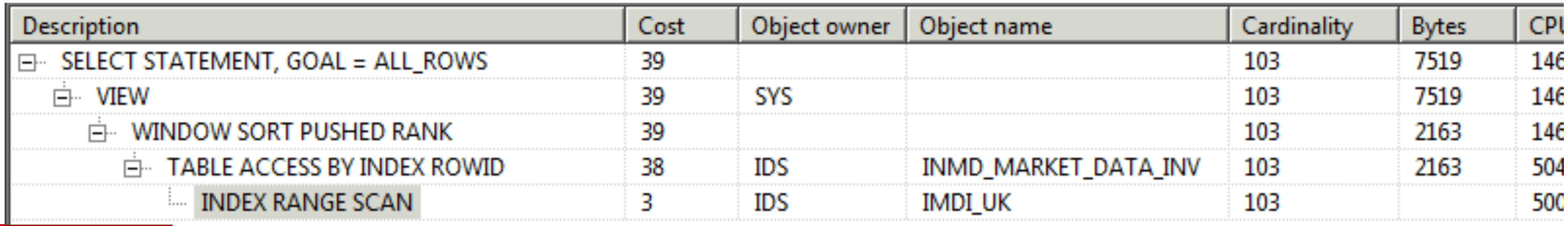
Example:
- delivers a time series for validity date

Analytical Functions – RANK

- Prioritization of rows and rows in data sets (in window functions)
  - Used to build waterfalls for views

```
SELECT *
  FROM (SELECT imdi.imdi_anchor_id,
               imdi.imdi_valuation_date,
               RANK() OVER(PARTITION BY imdi.imdi_anchor_id
                           ORDER BY imdi.imdi_valuation_date DESC) AS rk,
               imdi.imdi_f_asset_val_net
          FROM ids.inmd_market_data_inv imdi
         WHERE imdi.imdi_anchor_id = 'AN00054467'
        )
 WHERE rk = 1;
```

Example:
delivers latest net asset value of an investment



Description	Cost	Object owner	Object name	Cardinality	Bytes	CPU
SELECT STATEMENT, GOAL = ALL_ROWS	39			103	7519	146
VIEW	39	SYS		103	7519	146
WINDOW SORT PUSHED RANK	39			103	2163	146
TABLE ACCESS BY INDEX ROWID	38	IDS	INMD_MARKET_DATA_INV	103	2163	504
INDEX RANGE SCAN	3	IDS	IMDI_UK	103		500

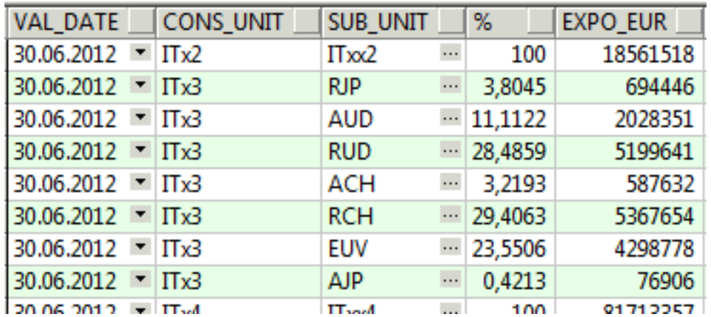
Analytical Functions – RATIO_TO_REPORT

- Computes the ratio of a value to the sum of a data set of values
  - Used to calculate breakdowns and distributions of investment classes

```
WITH
summe AS (
SELECT valid_until, cons_unit,
       SUM(exposure_eur) exp_eur,
       SUM(exposure_por) exp_por),
  FROM vo_a_positions_aggregated
 WHERE …
 GROUP BY …)
SELECT e.valid_until,
       e.cons_unit,
       e.scur_code,
       e.val,
       100 * SUM(e.exposure_eur) / s.exp_eur
       100 * SUM(e.exposure_por) / s.exp_por
  FROM vo_a_positions_aggregated e,
       summe s
 WHERE e.valid_until = s.valid_until
   AND e.cons_unit = s.cons_unit
 GROUP BY …
```

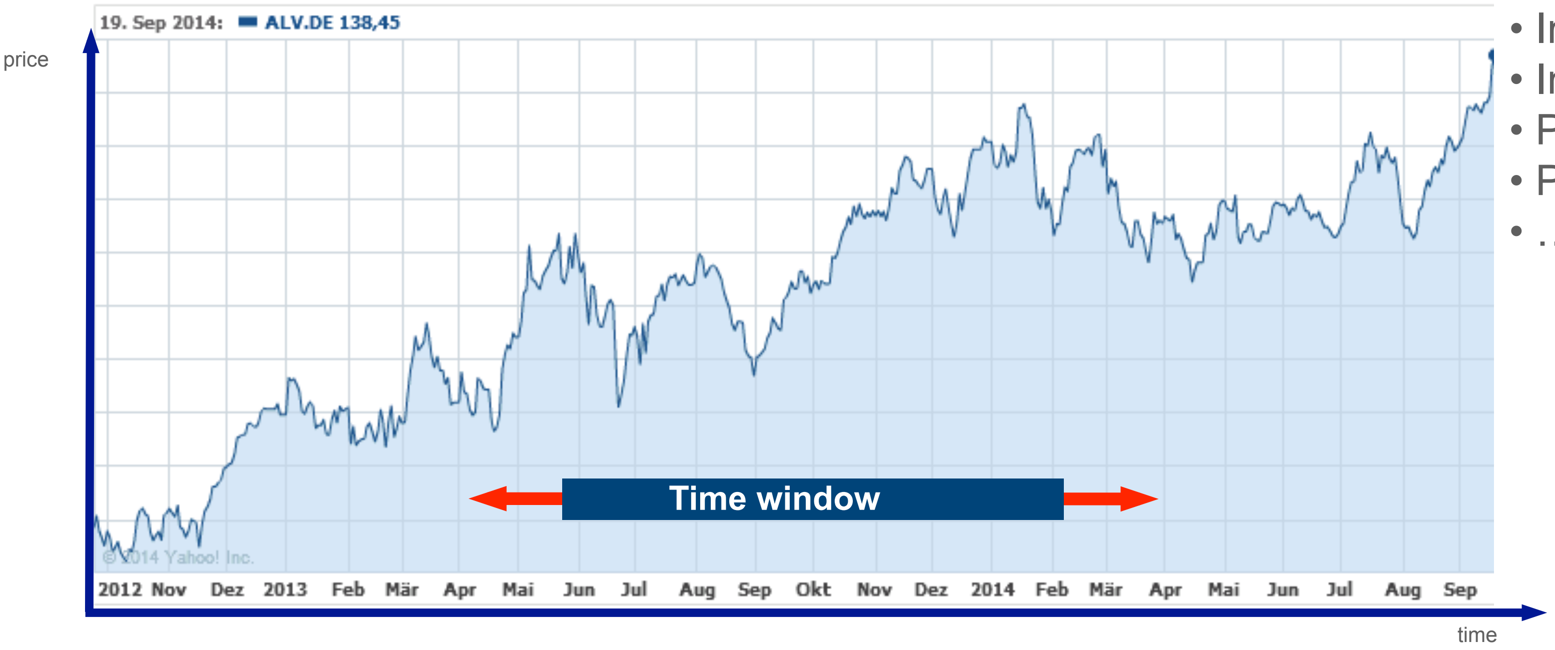
```
SELECT val_date, cons_unit , sub_unit
       RATIO_TO_REPORT(SUM(expo_eur))
       OVER (PARTITION by val_date, cons_unit) as ”%”,
       SUM(expo_eur) expo_eur,
FROM vo_a_positions_aggregated
WHERE …
GROUP BY …
```

Example:
- delivers latest exposure distribution of subfonds



VAL_DATE	CONS_UNIT	SUB_UNIT	%	EXPO_EUR
30.06.2012	ITx2	ITxx2	100	18561518
30.06.2012	ITx3	RJP	3,8045	694446
30.06.2012	ITx3	AUD	11,1122	2028351
30.06.2012	ITx3	RUD	28,4859	5199641
30.06.2012	ITx3	ACH	3,2193	587632
30.06.2012	ITx3	RCH	29,4063	5367654
30.06.2012	ITx3	EUV	23,5506	4298778
30.06.2012	ITx3	AJP	0,4213	76906
30.06.2012	ITx4	ITxx4	100	81713357

Analytical Functions – Moving time window example



- Impairments
- Inv. Performance
- Portf. Hierarchies
- Price Series
- …

```
SELECT AVG(last_price) over
           (PARTITION BY inv_id, exchange, vendor ORDER BY val_date
            RANGE BETWEEN INTERVAL '3' MONTH PRECEDING AND INTERVAL '1' DAY PRECEDING) avg_price_month…
```

ORACLE

Investment Data Services

Analytical Functions – Aggregates example

ROLLUP

SPOK ID	Attr. Type	Marketvalue Sum
4547	A	-2.434.690
4547	C	-13.638.163
4547	F	78.300.073
4547	S	2.165.730.474
4547	X	-943.969
4547		2.227.013.724
4551	A	-1.971.228
4551	C	32.572.552
4551	S	714.799.433
4551	X	-24.653
4551		745.376.105
		2.972.389.829

CUBE

Attr. Type	4551	4547	Total
A	-1.971.228	-2.434.690	-4.405.918
C	32.572.552	-13.638.163	18.934.389
F		78.300.073	78.300.073
S	714.799.433	2.165.730.474	2.880.529.907
X	-24.653	-943.969	-968.622
Total	745.376.105	2.227.013.724	2.972.389.829

- used to classify/aggregate data for types of investments (eg. Stocks, fund, derivatives etc.) / “Excel” in the database

- used to aggregate data for portfolio hierarchies

12c – faster and smarter

- SQL Pattern Matching
- Approximate Count Distinct

Pattern Recognition In Sequences of Rows

SQL Pattern Matching” - Concepts

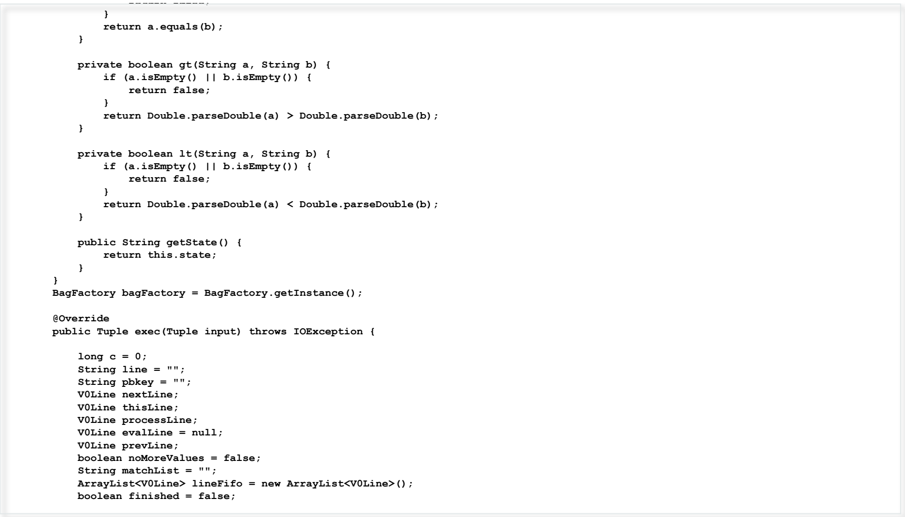
- Recognize patterns in sequences of events using SQL
  - Sequence is a stream of rows
  - Event equals a row in a stream

- New SQL construct MATCH_RECOGNIZE
  - Logically partition and order the data
  - Pattern defined with regular expressions via variables
    - Regular expression matched against a sequence of rows (*forwards/backwards*)
    - Each pattern variable is defined using conditions on rows and aggregates

Pattern Matching With Oracle SQL

Simplifies Development and Application Code: SQL vs. Java

Searching for double bottom (w-style) patterns in stock market data



```
            }
            return a.equals(b);
        }

        private boolean gt(String a, String b) {
            if (a.isEmpty() || b.isEmpty()) {
                return false;
            }
            return Double.parseDouble(a) > Double.parseDouble(b);
        }

        private boolean lt(String a, String b) {
            if (a.isEmpty() || b.isEmpty()) {
                return false;
            }
            return Double.parseDouble(a) < Double.parseDouble(b);
        }

        public String getState() {
            return this.state;
        }
    }
    BagFactory bagFactory = BagFactory.getInstance();

    @Override
    public Tuple exec(Tuple input) throws IOException {

        long c = 0;
        String line = "";
        String pbkey = "";
        V0Line nextLine;
        V0Line thisLine;
        V0Line processLine;
        V0Line evalLine = null;
        V0Line prevLine;
        boolean noMoreValues = false;
        String matchList = "";
        ArrayList<V0Line> lineFifo = new ArrayList<V0Line>();
        boolean finished = false;
```

250+ Lines of Java UDF

```
SELECT first_x, last_z
FROM ticker
MATCH_RECOGNIZE (
       PARTITION BY stock_id ORDER BY timestamp
       MEASURES FIRST(x.time) AS first_x,
                LAST(z.time) AS last_z
       ONE ROW PER MATCH
       PATTERN (X+ Y+ W+ Z+)
       DEFINE X AS (price < PREV(price)),
              Y AS (price > PREV(price)),
              W AS (price < PREV(price)),
              Z AS (price > PREV(price)  AND
                   z.time - FIRST(x.time) <= 7 ))
```

12 Lines of SQL

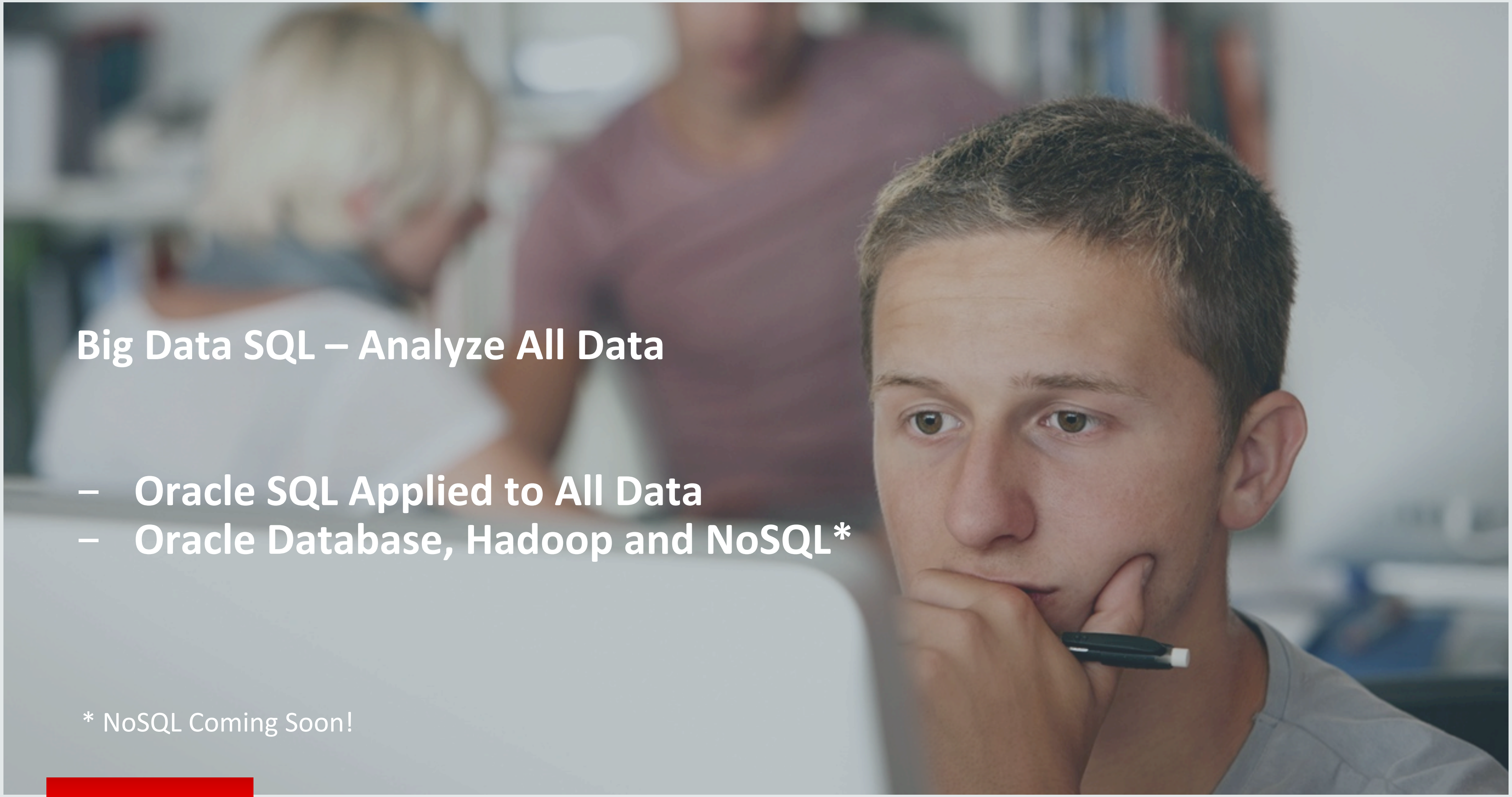
20x less code

Getting An *Approximate* Answer From a Query

When "good enough" is good enough

- Business Problems
  - Not every query requires a completely accurate result
    - Trending, data discovery, social analysis
  - Exact processing of large data sets not economical
  - For interactive analysis, sufficient accuracy of a query not known at start
    - Accuracy dependent on result because of "Think time" before next analysis step
- Solutions
  - Provide "approximate result" capabilities in SQL
    - Guided by user-controllable intentions: maximum runtime, maximum accuracy, number of iterations
  - Framework allows data analysis accuracy to be refined progressively

ORACLE



Big Data SQL – Analyze All Data

- Oracle SQL Applied to All Data
- Oracle Database, Hadoop and NoSQL*

* NoSQL Coming Soon!

Oracle Big Data SQL

Oracle SQL on all your data.

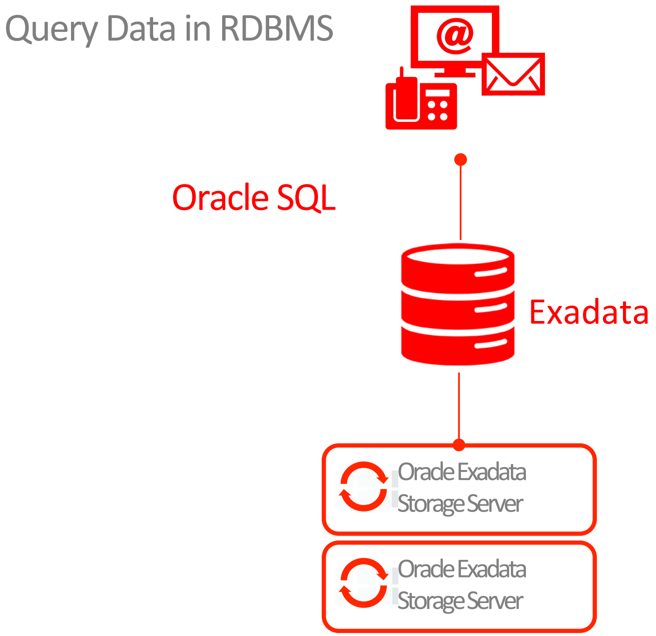
Oracle SQL on Hadoop and beyond

- With a **Smart Scan** service inspired by Exadata
- With **native** SQL operators
- With the **security** of Oracle Database

ORACLE

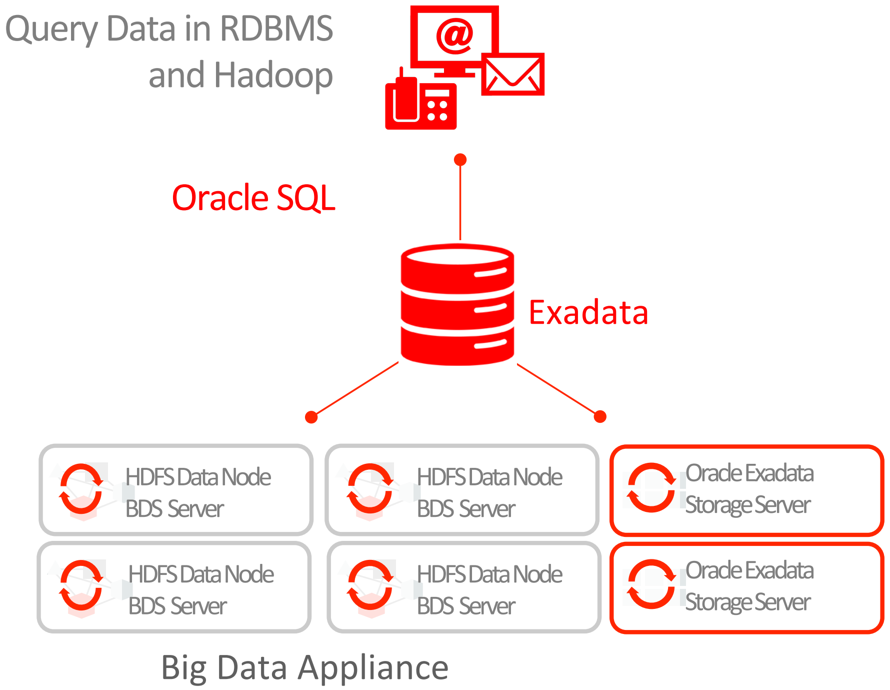
Intelligent Query Optimization

Exadata: Applies SmartScan Close to the Data



Intelligent Query Optimization

Exadata & Big Data SQL: Applies SmartScan Close to All Data



Fast

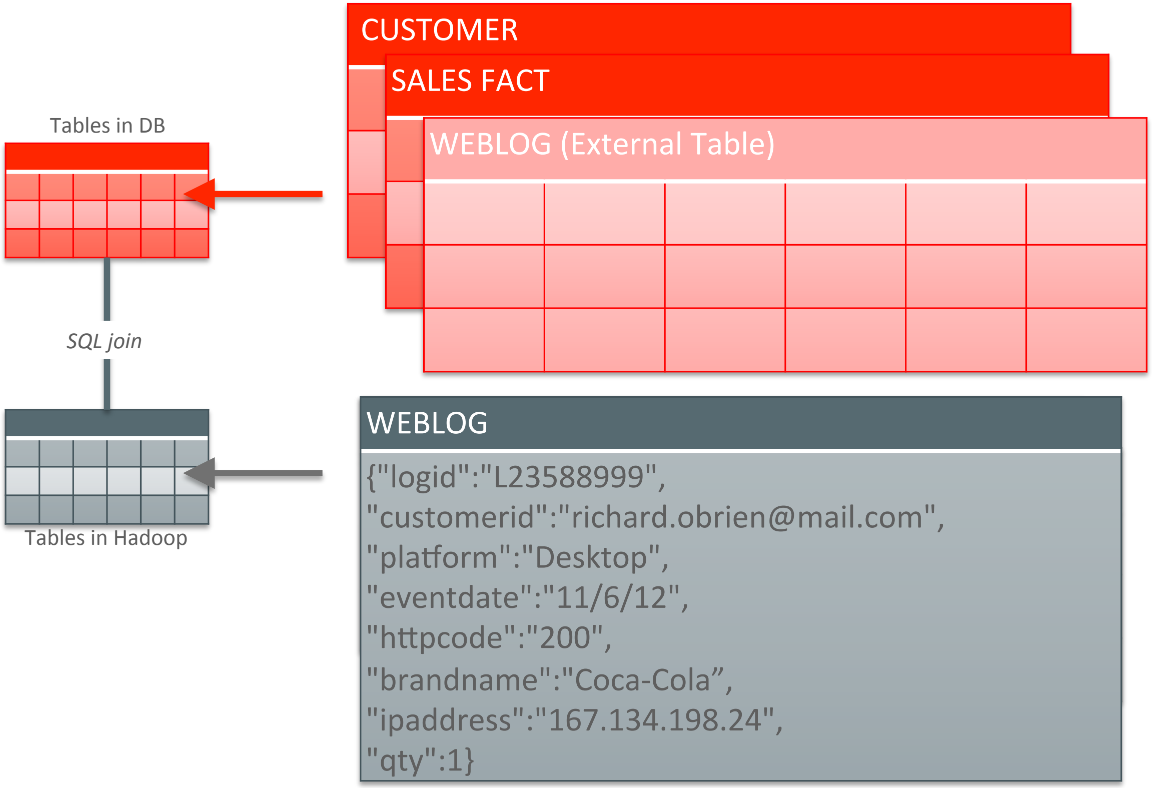
Massive Parallelism

Filtered Locally

Minimized Data Movement

ORACLE

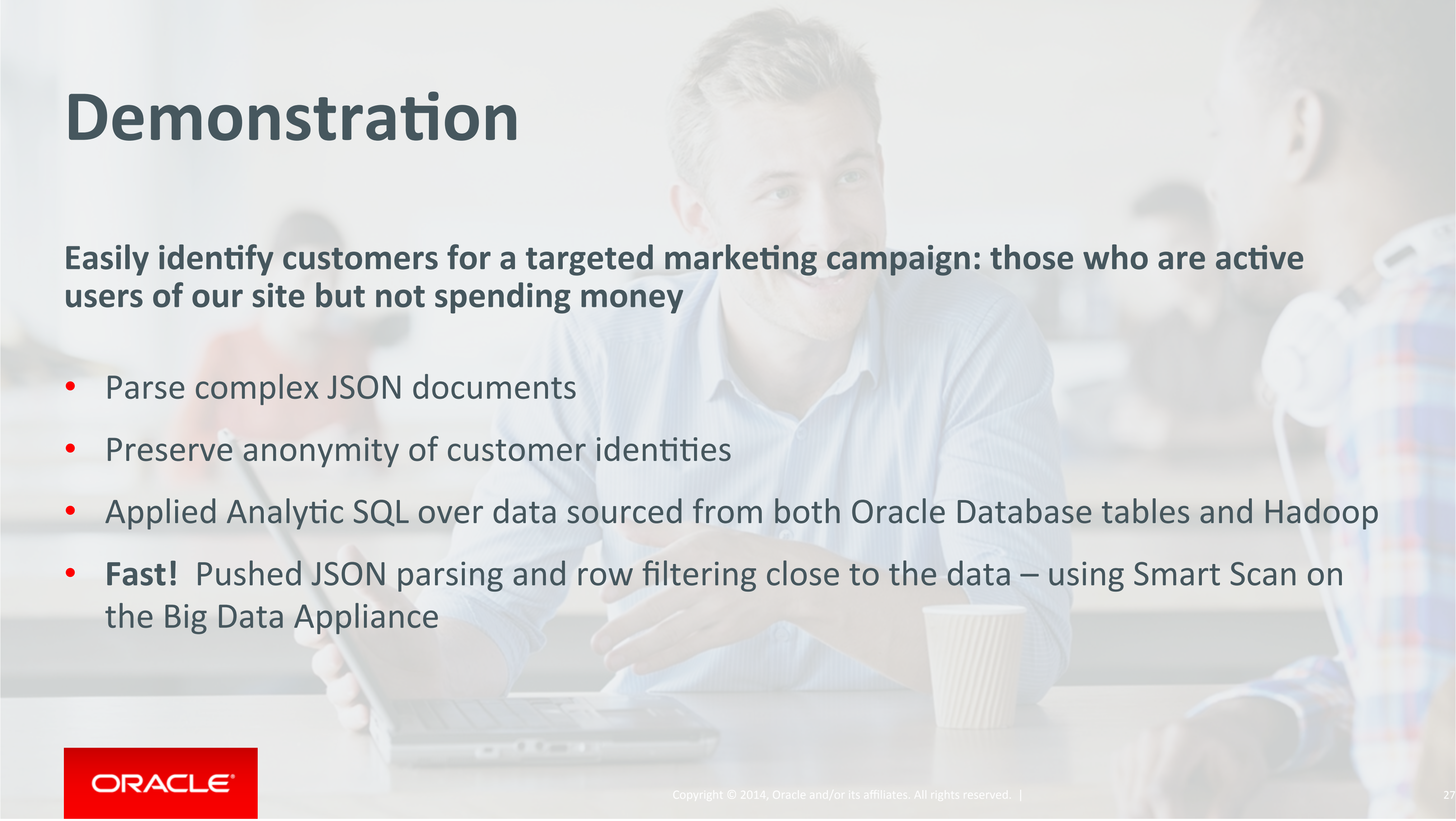
Oracle Big Data SQL Demonstration



{"logid":"L23588999",
"customerid":"richard.obrien@mail.com",
"platform":"Desktop",
"eventdate":"11/6/12",
"httpcode":"200",
"brandname":"Coca-Cola”,
"ipaddress":"167.134.198.24",
"qty":1}

Securely analyze customer behavior and sales transactions for targeted marketing

Customer behavioral data from weblogs stored as JSON documents

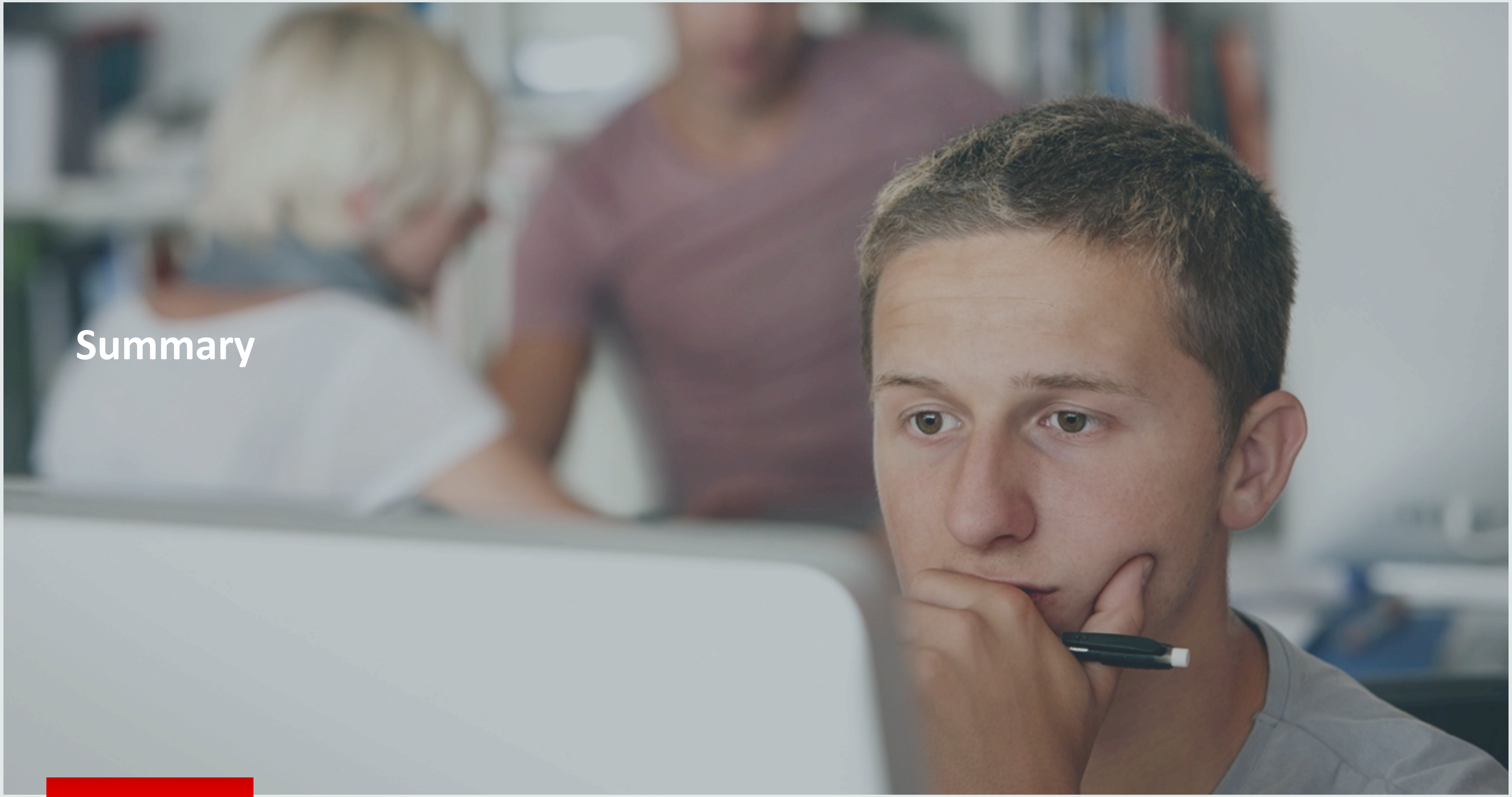


Demonstration

Easily identify customers for a targeted marketing campaign: those who are active users of our site but not spending money

- Parse complex JSON documents
- Preserve anonymity of customer identities
- Applied Analytic SQL over data sourced from both Oracle Database tables and Hadoop
- **Fast!** Pushed JSON parsing and row filtering close to the data – using Smart Scan on the Big Data Appliance

Summary



Key Benefits of Analyzing Big Data with SQL

Highlights

Key benefits provided by Oracle's analytical SQL for Big Data:

1. Increased performance
2. Enhanced developer productivity
3. Improved manageability
4. Minimized learning effort
5. Investment protection (through ANSI SQL compliance)

Where to get more information

- SQL Analytics Home Page on OTN
- Big Data Home Page on OTN
- Oracle Learning Library

- Big Data, Data Warehouse and SQL Analytics blog

Follow us on social media

in B

ORACLE

DW and Big Data Demo Booths

Come and visit us in the Moscone South Exhibition Hall

	Regular Hours	Dedicated Hours
Monday	9:45 a.m.–6:00 p.m.	9:45 a.m.–10:15 a.m.
Tuesday	10:00 a.m.–6:00 p.m.	10:00 a.m.–10:45 a.m.
Wednesday	9:30 a.m.–3:30 p.m.	9:30 a.m.–10:15 a.m. 2:45 p.m.–3:30 p.m.

ORACLE

DW and Big Data Sessions @OOW2014

Monday

Time	Session Title	Location
2:45pm	*Oracle Big Data: Strategy and Roadmap*	*South 104*
4:00pm	*Using Analytical SQL to Intelligently Explore Big Data*	*North 131*
5:15pm	Oracle Big Data Appliance: Deep Dive and Roadmap for Customers and Partners	South 104

Tuesday

Time	Session Title	Location
10:45pm	Data Warehousing and Big Data Customer Panel	South 302
12:00pm	Top Five Things to Know About Oracle Database In-Memory	South 104
6:00pm	Meet the Experts - Oracle’s Big Data Management System	South 303

DW and Big Data Sessions @OOW2014

Wednesday

Time	Session Title	Location
10:15pm	Big Data & Predictive Analytics: Fiserv Data Mining Case Study	South 301
12:45pm	If You Think Partitioning Is Only for Performance, Think Again	South 103
3:45pm	Oracle Big Data SQL: Deep Dive (SQL over Relational, NoSQL, and Hadoop)	South 103
4:45pm	Parallel Execution and Resource Management in Concurrent Environments	North 131

Thursday

Time	Session Title	Location
12:00pm	Oracle Database In-Memory Meets Oracle Optimizer	South 104

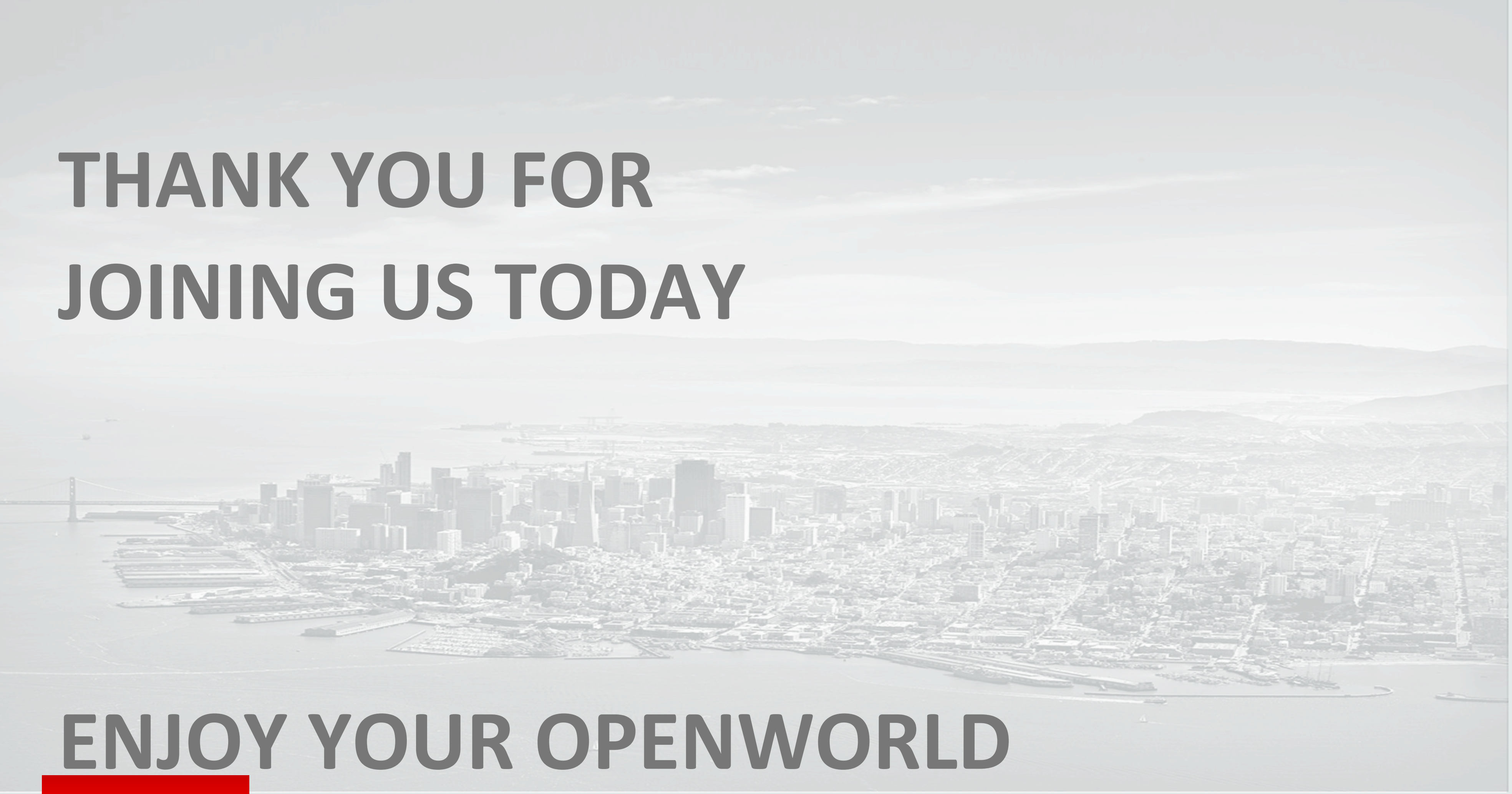
ORACLE

DW and Big Data Hands-on Lab @OOW2014

Oracle Big Data SQL: Unified SQL Analysis Across the Big Data Platform

Date	Time	Location
Monday	11:45am - 12:45pm	Hotel Nikko - Peninsula
Tuesday	3:45pm – 4:45pm	Hotel Nikko - Peninsula
Wednesday	1:15pm – 2:15pm	Hotel Nikko - Peninsula
Thursday	11:30am – 12:30pm	Hotel Nikko - Peninsula

ORACLE



THANK YOU FOR JOINING US TODAY

ENJOY YOUR OPENWORLD

Hardware and Software

Engineered to Work Together

ORACLE®