SGI and Sun

1996

2011

Google

facebook
Graph Algorithms

Depth-first search

Breadth-first search
Social Media

Which is anonymized and aggregated by PatientsLikeMe... to reveal real-world insights for industry partners.

Health information...
We’re a very tiny circle in the middle of this big universe. So it’s more likely interesting advances will come from outside the company than within the company.

George A. Scangos, CEO, Biogen Idec
Life Science industry trends
Challenges current IT solutions and strategies

Economic Drivers
- Patent expiries
- Fewer blockbusters
- Healthcare reform

Commercial Models
- Integrated Supply Chain
- Customer Relationships
- New Partnership Models

New Discovery Models
- External Collaboration
- New Computational Science Capabilities

Data Flood
- Archiving
- Aggregating / Analyzing
- Retention & Risk
- Intelligent Search

Technology Trends
- Enterprise apps
- Mobility
- Cloud
- Engineered systems

IT Investment Strategy
Research Trends and Challenges

• Research is More Expensive
• Output/$ is down
• Researchers want more Collaboration
• Data-Driven Discovery of Structured and Unstructured Data
• Budgets under stress
IT Complexity Imposes a “Tax” on the Business

- Slows Product Introduction
- Inhibits Research Productivity
- Stifles Business Processes
- Inhibits M&A Synergies

© 2011 Oracle Corporation
## Oracle Healthcare and Life Sciences Vision and Business Strategy

### Vision

Provide industry-leading applications and technology to help our customers grow revenues, increase efficiency, enhance research productivity and improve the quality of healthcare.

### Business Strategy

- Provide the most functionally rich and cost competitive enterprise software and technology solutions to the Healthcare & Life Sciences Industries
- Provide end-to-end solutions
- Provide integrated solutions that decrease costs of business operations
- Provide high quality support to our customers and partners
Oracle’s Strategy

• Deliver Complete, Open, Integrated Business Solutions to customers for their industry

• Provide flexibility and choice by integrating Oracle Software on an Open Standards Based Architecture

• Transform the data center with integrated Systems and Hardware that provide breakthrough performance and unprecedented lower cost of ownership

• Enable customers to deploy Cloud Computing strategies through a full-range of innovative technologies
Oracle Solution

Compute, Storage, Network Building Blocks

Optimized Systems and Solutions

Engineered Systems

Silo
Consolidated
Optimized
Cloud

© 2011 Oracle Corporation
From Data Deluge
to
Data-Driven Discovery
Research Data Lifecycle

- **INGEST**
  - Create
  - Gather

- **REDUCE/COMBINE**

- **ARCHIVE**

- **PUBLISH/CMPARE**

- **ANALYZE**

  - External Repositories
  - Survey
  - Instruments
  - Pubs
Oracle’s R&D
Data Management Platform

Database and Development
As a Service
Framework for Exadata to support data-driven application development

Data Cloud Preservation
The Leading Tape Archive and Preservation Filesystem

Database as File System
Common File Services, Data Provenance, DICOM Image Management

Analytics Engine Acceleration
Semantic Web, Oracle Data Mining Text Mining

SECURITY   SCALE   SKILLS
R&D Successes
Scientific Question begins the data
What if I need new data?
How do I ensure the data is clean?
How do I secure it?
How do I share it?
How do I scale it?
SPREADSHEET ARE NOT THE ANSWER

Hear Directly from DaVita and Purdue
http://www.events vc.com/support/asurvey/
 aebb9d44-02c6-4ce9-a325-6e8cb2337405
<table>
<thead>
<tr>
<th>Step Description</th>
<th>AVG PERF IMPROVEMENT FACTOR</th>
<th>CURRENT PROD AVG TIME per TEST CASE</th>
<th>EXADATA with TUNING AVG TIME per TEST CASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.1 Login (3 Steps)</td>
<td>1.7 x</td>
<td>77.7</td>
<td>46.4</td>
</tr>
<tr>
<td>6.2.2 Navigation and Save (28 Steps)</td>
<td>1.3 x</td>
<td>6.5</td>
<td>5.1</td>
</tr>
<tr>
<td>6.2.3 Text Search (27 Steps)</td>
<td>4.3 x</td>
<td>99.8</td>
<td>23.2</td>
</tr>
<tr>
<td>6.2.4 Structure Search (43 Steps)</td>
<td>5.3 x</td>
<td>74.2</td>
<td>14.0</td>
</tr>
<tr>
<td>6.2.5 Structure Search Combined with Text Search (15 Steps)</td>
<td>5.8 x</td>
<td>232.8*</td>
<td>40.0</td>
</tr>
</tbody>
</table>
Migrating from HP-UX platform to the Exadata Platform
- Estimates ranged from 10 – 14 days of downtime
- Utilized GoldenGate technology to migrate
  - Bi-directional transactional replication between the databases
  - Minimize risk with phased, rolling upgrade
  - Minimal downtime – a few hours to switchover and perform housekeeping tasks

Great Performance but How Do We Migrate? Exadata + Golden Gate
Oracle’s End-to-End Content Solution
Aligning the Content Lifecycle with Storage

Create
Documents Images Video Emails
Users generate content and upload to the system

Manage
Oracle Enterprise Content Management
ECM provides publishing, approval workflows, and application integration

Preserve
Oracle Database 11g Storage Archive Manager
Secure movement of files between tiers of storage, maintains integrity and sustains data over time

Store
Sun Flash Disk Tape
Secure, reliable and cost-effective access throughout lifecycle
Integrated Content & Storage Solution
Pre-configured and tested to reduce cost and lower risk

• Certified Hardware, Software, Services Complete
  • Includes Content Management Suite for Enterprise Content Lifecycle Management
  • Reference Architecture Certified and Optimized for Enterprise Content Management
  • Services for setup, installation and implementation for four key focus areas

• Most Cost-Effective
  – Matches cost of storage to content access requirements to minimize TCO
  – Blends disk for faster access and tape for less active access
  – Simplified management and operational infrastructure

• Fastest Deployment with Low Risk
  – Pre-tested and architected solution maximizes supportability and minimizes deployment risk
  – Enforce the highest levels of security through proven technologies certified by Oracle
  – ECM Suite reduces silo’s, consolidates redundant services and drives greater productivity

• Highly Scalable
  – Seamlessly scale within each tier of storage
  – Transparently move data between tiers based on policy
Develop In-DB Analytical Pipelines

Life & Health Sciences data

Functional Genomic Databases

Proteomics Database

Pharmacological databases

Deductive Analysis

Answer complex questions about the relationships in genomic, clinical and pharmacological data

Inductive Analysis

Finding relationships for classification, class discovery and prediction

© 2011 Oracle Corporation
Prioritization of Biological Targets for Drug Discovery

- Better prioritization of possible drug targets, based on
  - integrating large diversity of data,
    - electronic patient records, chemical structures, biological sequences, images, biological pathways, scientific papers,…
  - usage public and internal ontologies

- A search engine (“Target Assessment Tool”) is built to assist researchers
Oracle’s Analytic Foundation
In-Database Data Mining

**Results**
- Faster time for “Data” to “Insights”
- Lower TCO—Eliminates
  - Data Movement
  - Data Duplication
- Maintains Security

Model “Scoring”
Data remains in the Database
Embedded data preparation

- Cutting edge machine learning algorithms inside the SQL kernel of Database
- SQL—Most powerful language for data preparation and transformation

Data remains in the Database

**Traditional Analytics**
- Data Import
  - Data Mining Model “Scoring”
- Data Preparation and Transformation
  - Data Mining Model Building
- Data Prep & Transformation
- Data Extraction

**Oracle Data Mining**
- Model “Scoring”
  - Embedded Data Prep
  - Model Building
  - Data Preparation

**Savings**
- Hours, Days or Weeks
- Secs, Mins or Hours

© 2011 Oracle Corporation
11g Statistics & SQL Analytics (Free)

- Ranking functions
  - rank, dense_rank, cume_dist, percent_rank, ntile

- Window Aggregate functions
  (moving and cumulative)
  - Avg, sum, min, max, count, variance, stddev, first_value, last_value

- LAG/LEAD functions
  - Direct inter-row reference using offsets

- Reporting Aggregate functions
  - Sum, avg, min, max, variance, stddev, count, ratio_to_report

- Statistical Aggregates
  - Correlation, linear regression family, covariance

- Linear regression
  - Fitting of an ordinary-least-squares regression line to a set of number pairs.
  - Frequently combined with the COVAR_POP, COVAR_SAMP, and CORR functions

Descriptive Statistics
- DBMS_STAT_FUNCS: summarizes numerical columns of a table and returns count, min, max, range, mean, median, stats_mode, variance, standard deviation, quantile values, +/- n sigma values, top/bottom 5 values

Correlations
- Pearson’s correlation coefficients, Spearman's and Kendall's (both nonparametric).

Cross Tabs
- Enhanced with % statistics: chi squared, phi coefficient, Cramer's V, contingency coefficient, Cohen's kappa

Hypothesis Testing
- Student t-test, F-test, Binomial test, Wilcoxon Signed Ranks test, Chi-square, Mann Whitney test, Kolmogorov-Smirnov test, One-way ANOVA

Distribution Fitting
- Kolmogorov-Smirnov Test, Anderson-Darling Test, Chi-Squared Test, Normal, Uniform, Weibull, Exponential
Oracle Exadata Database Machine

- **Fastest** for data warehousing & OLTP
- **Best** data warehousing & OLTP cost/performance
- Best for consolidation
  - Only database machine that runs and scales *all* workloads
- 100% redundant hardware & scalable on demand
Exadata Database Machine X2-2 Full Rack
Pre-Configured for Extreme Performance

- 8 x64 Dual-processor Database Servers (Sun Fire X4170 M2)
  - 96 cores (12 per server)
  - 768 GB memory (96GB per server)
  - 10 GigE connectivity to Data Center
    - 16 x 10GbE ports (2 per server)
- 14 Exadata Storage Servers X2-2
  - All with High Performance 600GB SAS disks
  OR
  - All with High Capacity 2 TB SAS disks
- 3 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Cisco Ethernet switch
- Keyboard, Video, Mouse (KVM) hardware
- Redundant Power Distributions Units (PDUs)
Exadata Database Machine X2-8 Full Rack
Extreme Performance for Consolidation, Large OLTP and DW

- 2 x64 Eight-processor Database servers (Sun Fire 4800)
  - High Core, High Memory Database Servers
  - 128 CPU cores (64 per server)
  - 2 TB (1 TB per server)
  - 10 GigE connectivity to Data Center
    - 16 x 10GbE ports (8 per server)
- 14 Exadata Storage Servers X2-2
  - All with High Performance 600GB SAS disks
  OR
  - All with High Capacity 2 TB SAS disks
- 3 Sun Datacenter InfiniBand Switch 36
  - 36-port Managed QDR (40Gb/s) switch
- 1 “Admin” Cisco Ethernet switch
- Redundant Power Distributions Units (PDUs)

Data Weighted
Exadata Delivers Extreme Consolidation
Create a “Database-as-a-Service” platform

- **Large Memory**
  - Many databases can be consolidated

- **Extreme Performance**
  - OLTP, DW, data mining, batch, reporting, loading, backups, files in the database
  - Encryption, compression

- **Workload Isolation**
  - Manage SLAs via Quality of Service (QoS)
  - CPU and I/O resource management
  - Instance caging – Isolation

Shrink data center costs, increase system utilization and promote data integration

© 2011 Oracle Corporation
Enabling Research and Discovery
Addressing the Research Data Lifecycle

- Productivity of PI and IT
- Acceleration of Analytics
- Collaboration
- Preservation
Oracle and the Marketplace

• Teradata
• Greenplum
• Netezza
Oracle’s R&D
Data Management Platform

- **Database and Development As a Service**: Framework for Exadata to support data-driven application development
- **Data Cloud Preservation**: The Leading Tape Archive and Preservation Filesystem
- **Database as File System**: Common File Services, Data Provenance, DICOM Image Management
- **Analytics Engine Acceleration**: Semantic Web, Oracle Data Mining Text Mining

**SECURITY**  **SCALE**  **SKILLS**
Hardware and Software

Engineered to Work Together

- Innovation delivered faster
- Better performance, reliability, security
- Shorter deployment times
- Easier to manage and upgrade
- Lower cost of ownership
- Reduced change management risk
- One-stop support
Oracle’s Optimized Solutions
Fast To Deploy, Optimized For The Enterprise

- Up to 3x Faster Than Any Consolidated Platform
  - Siebel CRM

- Fastest Payroll Processing By 52%
  - PeopleSoft HCM

- 4x Faster e-Commerce Query Transactions
  - WebLogic Suite

- Up to 10x Faster Than Intel with 3x less CPU Overhead
  - WebCenter

- Bulletproof Non-stop Business Services for High Availability
  - E-Business Suite

- Fastest DB for ERP, OLAP, DSS, Payroll, PeopleSoft Campus
  - Database

© 2011 Oracle Corporation