ORACLE®

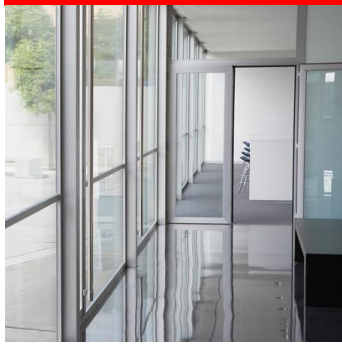
ORACLE SPATIAL

ORACLE

Best Practices for Spatial Analytics in a Data Warehouse

Hans Viehmann
Product Manager EMEA

Agenda

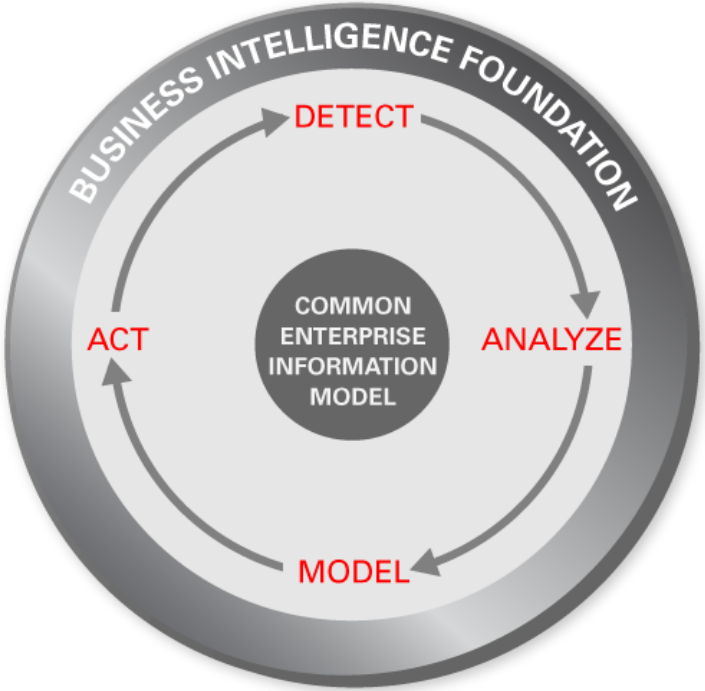


- Introduction
- Integration of Spatial Data into Business Intelligence environments
- Best Practices
  - Geospatial Warehousing on Engineered Systems
  - Geospatial Analytics on Streaming Data
- Summary

Basis of all work – a „Single Source of Truth“

with Spatial Data integrated

- consolidated data repository
  - for reporting, ad hoc queries, analytics, modelling
  - updated at regular intervals
- tools to set strategic goals, track initiatives and detect variances
  - ability to turn insight into action



Integrated Tools accessing Data Warehouse

using Oracle Mapviewer for visualization

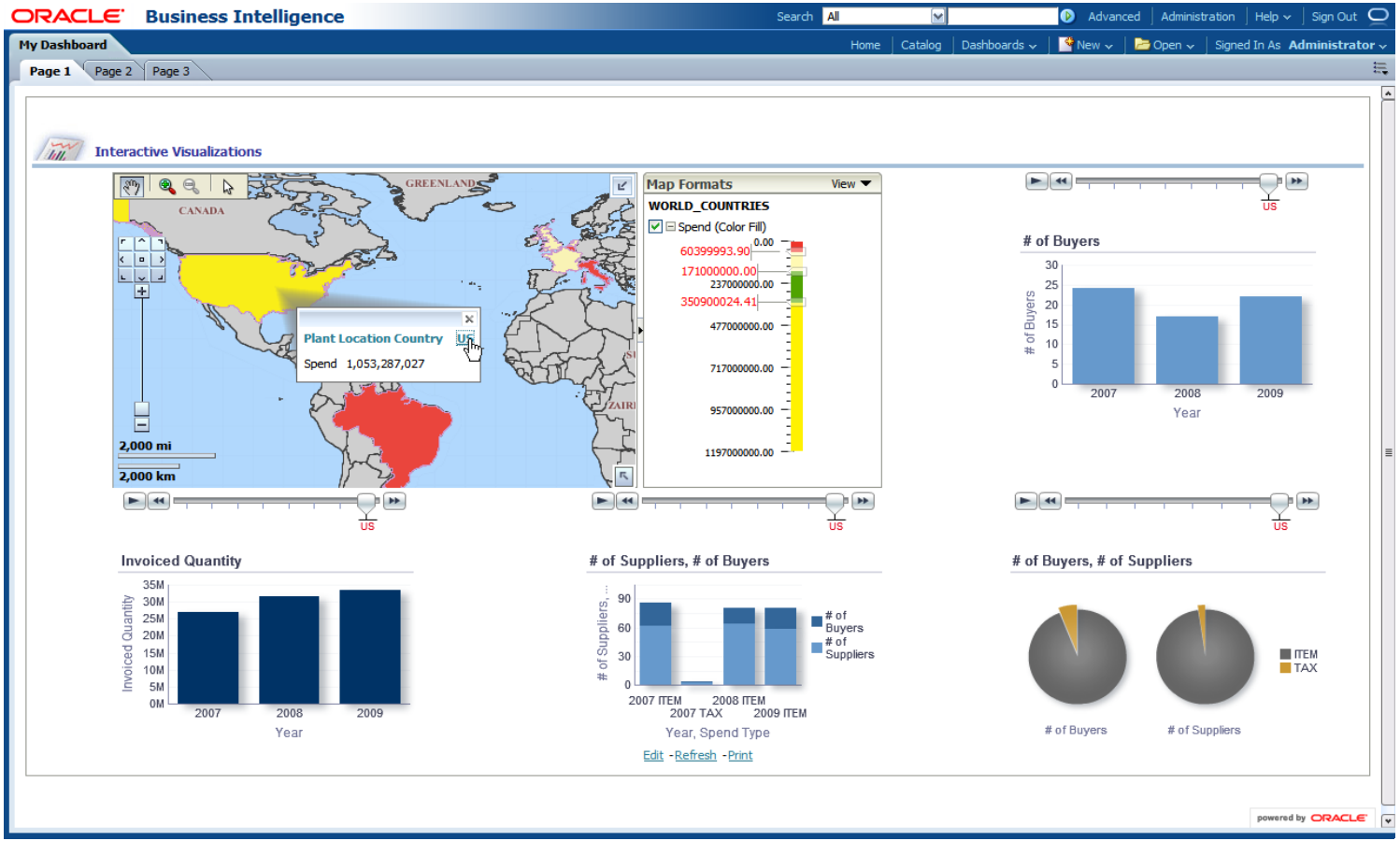
Data Integration | Scorecards | Interactive Dashboards | Reporting & Publishing | Ad-hoc Analysis | Office Integration | Search | Detect & Alert | Collaborate | Mobile | Embedded

Common Enterprise Information Model

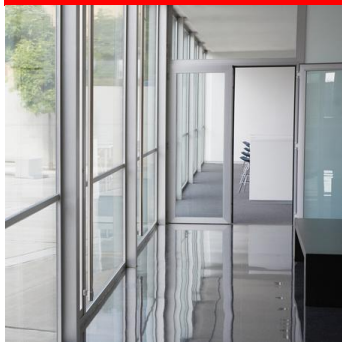
- Common Metadata Foundation across all Data Sources
- Common Security, Access Control, Authorization, Auditing
- Common Request Generation and Optimized Data Access Services
- Common Clustering, Workload Management, & Deployment
- Common Systems & Operational Lifecycle Management

OLTP & ODS Systems | Data Warehouse Data Mart | Exadata | OLAP Sources | Packaged Applications (Oracle. SAP. Others) | Unstructured & Semi-Structured | Excel XML/Office | Business Process

ORACLE



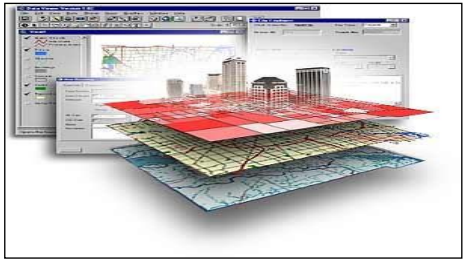
Agenda



- Introduction
- Integration of Spatial Data into Business Intelligence environments
- Best Practices
  - Geospatial Warehousing on Engineered Systems
  - Geospatial Analytics on Streaming Data
- Summary

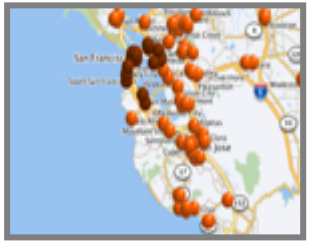
Strategic Vision for Geospatial Data

Provide the Platform



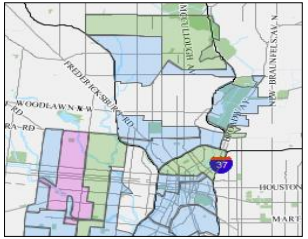
- ubiquitous Spatial services in IT infrastructure
  - significantly lower costs
  - simplified application development
  - integrated operational systems
- enable Spatially-enabled solutions to focus on business context, not infrastructure services
  - Database and Application Server manage deployment infrastructure
    - Security
    - Scalability
    - Standards Compliance
    - Load Balancing
    - Failover
    - HW/SW Dependencies
  - support multiple application models with common data model
    - Java
    - GML
    - Future …
    - .Net
    - Web Services/SOA

Natively manage all spatial content



Points -- Locations of Interest

Lines -- Streets, Pipelines, cables, rail, etc.

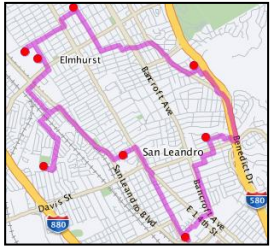


Polygons -- Admin boundaries, sales territories, high risk zones

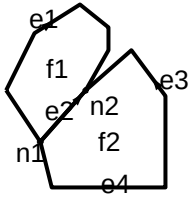
Web Services (OGC)

Geocoding
Routing

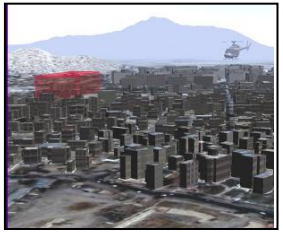
Satellite Imagery



Network Models -- Connectivity Analysis



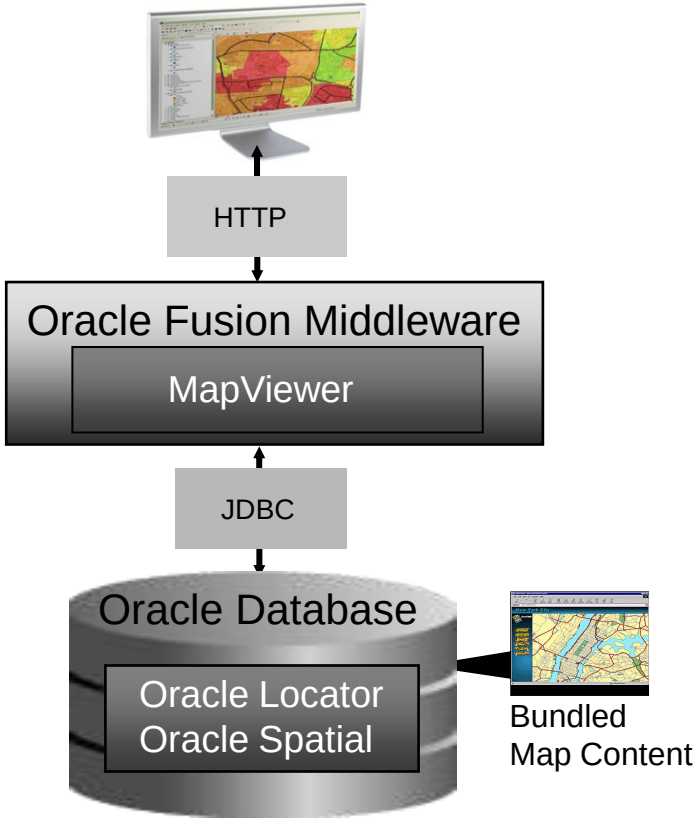
Topologies



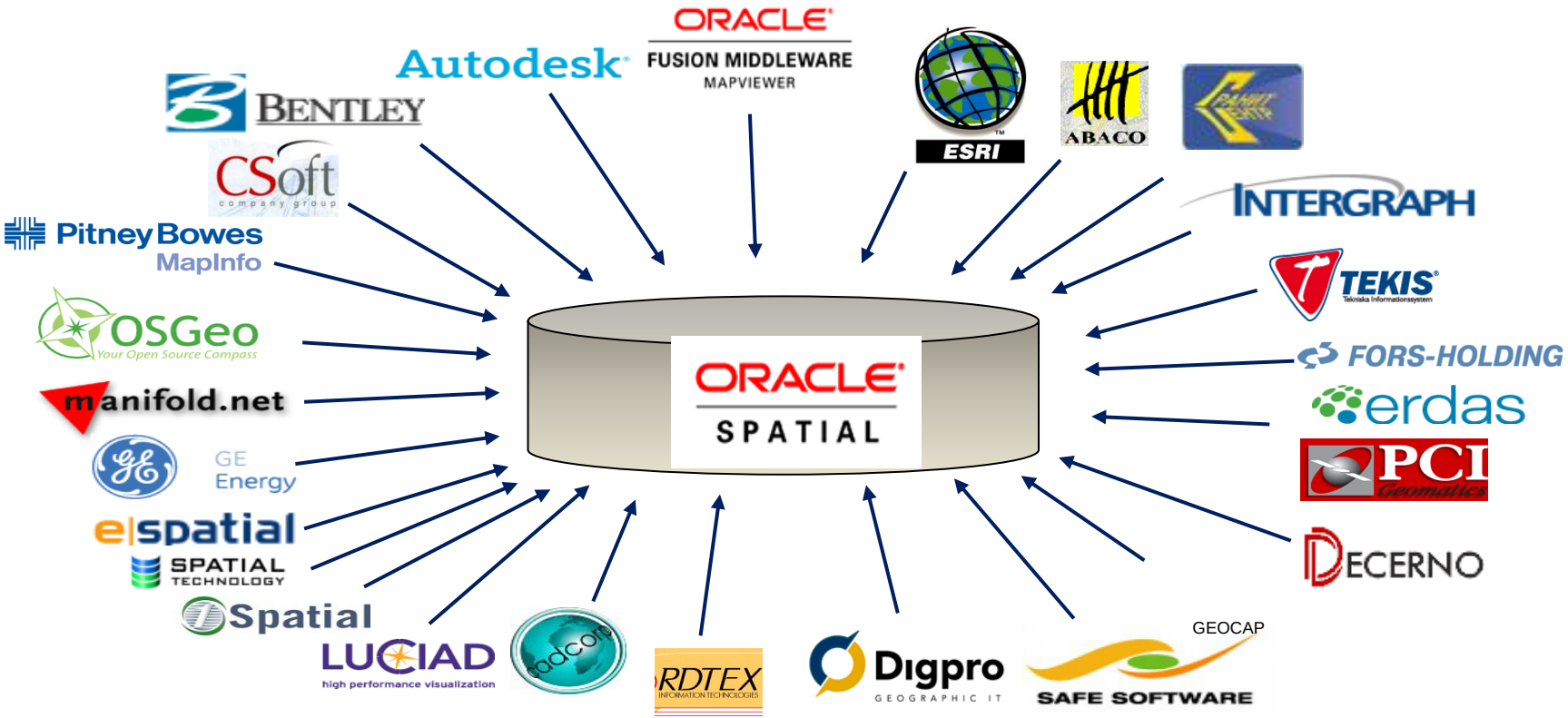
3D and LiDAR

Oracle Spatial Technologies

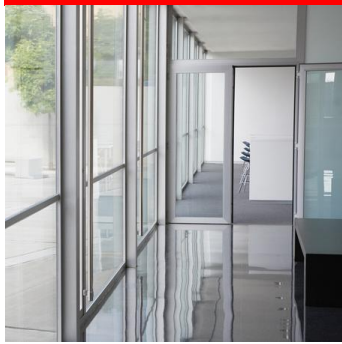
- **Oracle Locator**: Feature of Oracle Database XE, SE, EE
- **Oracle Spatial**: Priced option to Oracle Database EE
- **MapViewer**: Java application and map rendering feature of Oracle Fusion Middleware
- **Workspace Manager**: Long transactions feature of Oracle Database SE, EE
- **Bundled Map Content**: Major roads, administrative boundaries (city, county, state, country) - worldwide coverage eg. from Navteq



Open and interoperable



Agenda



- Introduction
- Integration of Spatial Data into Business Intelligence environments
- Best Practices
  - Geospatial Warehousing on Engineered Systems
  - Geospatial Analytics on Streaming Data
- Summary

Oracle Strategy

Complete Stack

- Best-of-breed
- Open
- Vertical Integration
- Extreme Performance
- Engineered Systems

Complete Customer Choice

- On-premise
- Private Cloud
- Public Cloud
- Hybrid Cloud

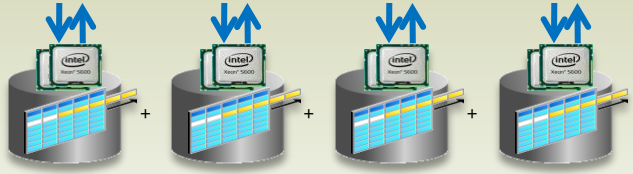
Exadata Database Machine

An engineered system - compute, storage, networking

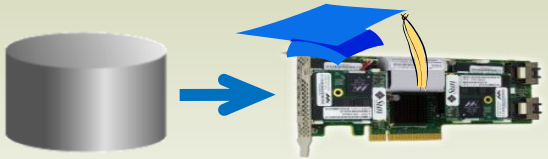
- Database Grid
  - Intel-based database servers, up to 128 cores
  - Oracle Linux or Solaris 11
  - Oracle Database 11g R2
  - 10 Gig Ethernet (to data center)
- Storage Grid
  - Intel-based storage servers, up to 168 cores
  - Up to 336 terabytes raw disk
  - 5.3 terabytes Flash storage
  - Exadata Storage Server Software
- InfiniBand Network
  - Internal connectivity (40 Gb/sec)

Key Exadata Innovations

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

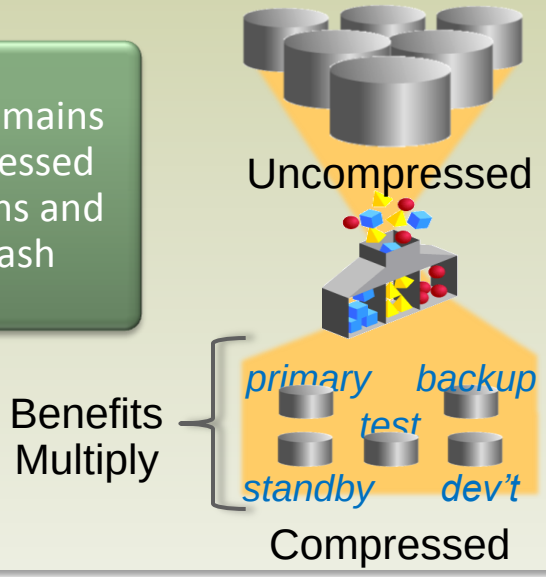


- **Smart Flash Cache**
  - Accelerates random I/O up to 30x
  - Doubles data scan rate



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

Data remains compressed for scans and in Flash



Oracle Exadata for Spatial Workloads

EXTREME PERFORMANCE

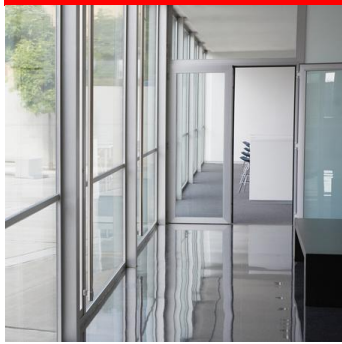
- Oracle Spatial is particularly suitable for the Exadata architecture
  - CPU intensive geospatial calculations
  - excellent use of parallelism
  - load distribution between server nodes and storage heads
  - index operations benefiting from large memory
- Breaks new boundaries for ingesting spatial data
  - using the fast I/O and high network bandwidth
- same Architecture for Transaction Processing, Data Warehousing, Spatial Data Infrastructure, …

ORACLE

Use Cases from Real-world Experiences

1. Re-Insurance Company
   - Risk management in a hurricane scenario
2. US Rail
   - Bulk nearest neighbor queries to find closest track, and project reported train positions onto tracks
3. Validate home appraisals for a Government Sponsored Enterprise (GSE)
   - Find all the parcels touching parcels to validate appraisals
4. Satellite Imagery Provider
   - Find all the useful portions of cloud covered imagery

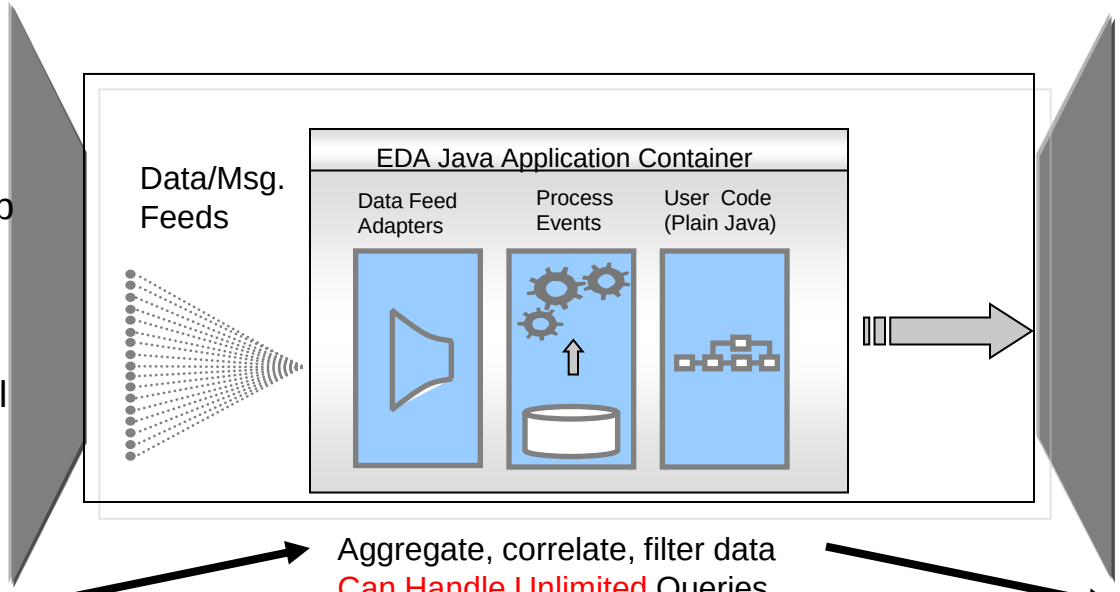
Agenda



- Introduction
- Integration of Spatial Data into Business Intelligence environments
- Best Practices
  - Geospatial Warehousing on Engineered Systems
  - Geospatial Analytics on Streaming Data
- Summary

Streaming Event-Driven Architecture (EDA)

- Streams/Relations
  - From any source: data streams, web services, Java, Database
- Adapters
  - Translate external events/data into java objects for processing



- Processors
  - Set of queries applied to the streams
- Listeners
  - Handle triggers raised by the processors
- Events
  - Implemented as JavaBean or Map

Aggregate, correlate, filter data
Can Handle Unlimited Queries

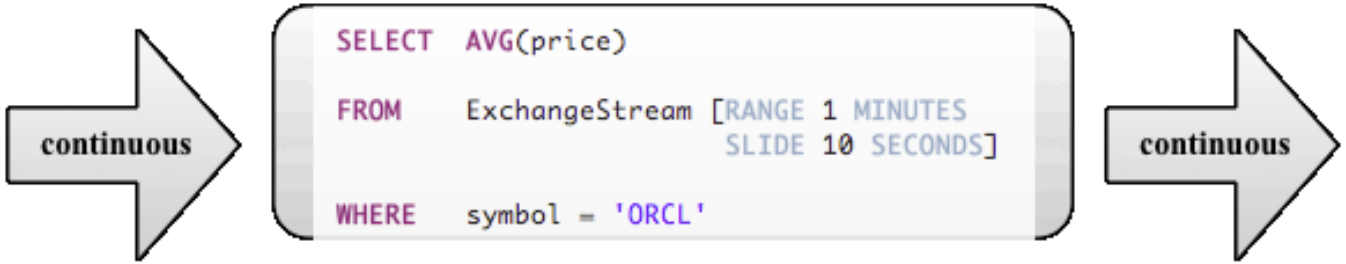
Incoming Data Streams
1,000,000+ Messages/Sec

Latency
Microsecond Latency (Avg.)

Resulting Data fed to vast business opportunities with Java language

Continuous Query Language (CQL)

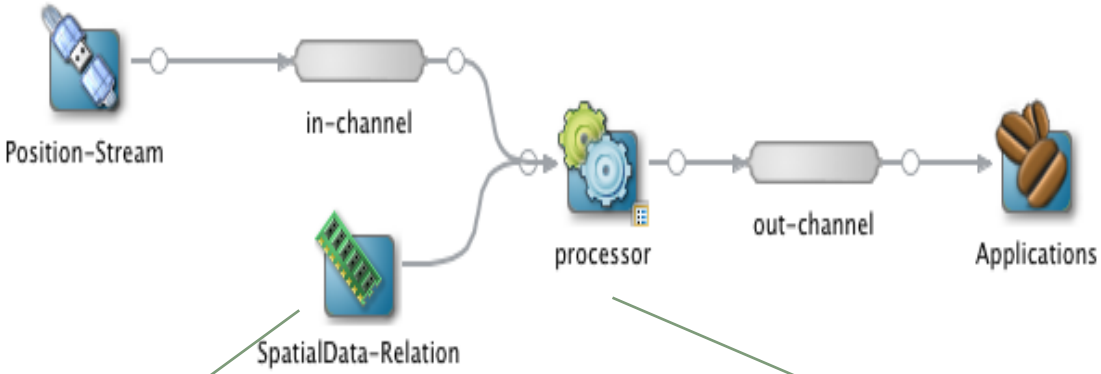
- Stream-Relation algebra
  - Continuous semantic
  - Driven by event, driven by time
  - Allows us to leverage SQL principles/implementation

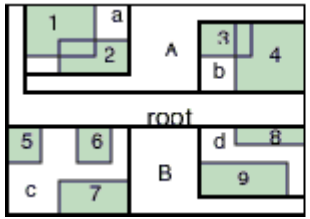


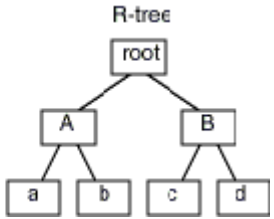
Window of events

Rate of output

Real-time processing using spatial data

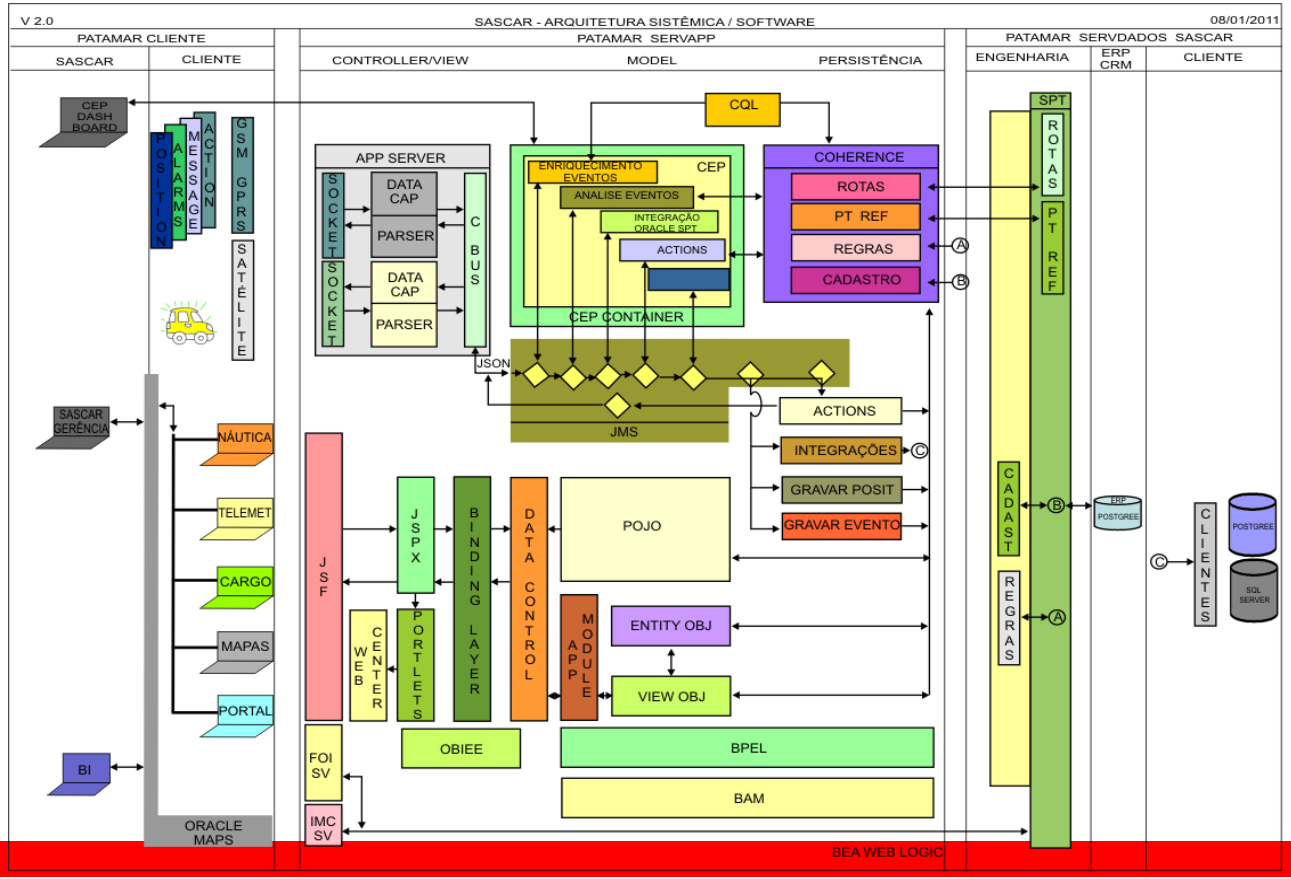




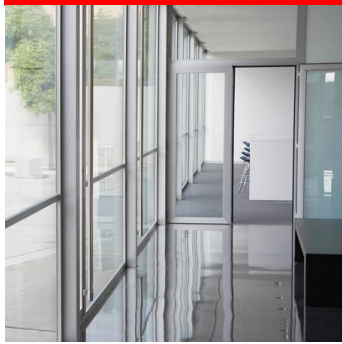



```
SELECT *
FROM in-channel[now] as S,
  SpatialData-Relation as R
WHERE  contains@spatial(R.geometry, S.pos, 0d)
  = true
```

Implementation example – Sascar (fleet management)



Agenda



- Introduction
- Integration of Spatial Data into Business Intelligence environments
- Best Practices
  - Geospatial Warehousing on Engineered Systems
  - Geospatial Analytics on Streaming Data
- Summary

Summary – best practices

- for integration of spatial data with enterprise data use a standards-compliant spatial database
  - look out for data integration tools, process integration capabilities and seamless BI platform besides map display
- if you need significant processing power and need to save cost, consider Exadata
- if you need spatial analytics on streaming data, consider CEP (complex event processing) tools
- presentation on Cloud Computing and „Big Data“ in Cloud Computing session Friday, 9:00am

Hardware and Software

ORACLE

Engineered to Work Together

“Oracle continues its remarkably consistent strategy of including spatial capabilities within its database and application deployment platform. … Oracle offers a wide range of spatial capabilities across its product line that provide foundation-level SIM capabilities for enterprise applications – including industry-specific applications.
Oracle maintains a dominant position as the spatial data repository for medium-sized and large spatial systems. …”

David Sonnen and Dan Vesset

IDC Analyze the Future

Source: IDC #224740. Worldwide Spatial Information Management 2010-2014 Forecast and 2009 Vendor Shares

ORACLE

ORACLE®