ORACLE
BIG DATA

ORACLE®

Industrialization of Geospatial Workflows: Building Industrial Strength Cybersystem / IT Base Platforms

Steven Hagan, Vice President, Server Technologies

Industrialization of Workflows

- HENRY FORD
- 100 YEARS AGO
- Industrialized the Manufacturing Workflow
- Popularized the use of the Assembly Line in Manufacturing
- One Machine Driven Platform for the entire car process

United Nation Analysis – September 2013

Initiative on Global GeoSpatial Information Management

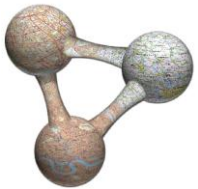
Future Trends

- Technology Trends in Data Creation, Maintenance, and Management
- Reliance on '*big data'* technologies
- The *right* information at the *right time*
- Machine-processable descriptions of data.
- Semantic technologies will play an important role
- Skills and Training: train the individuals is at least five years
- Requirement for enhanced data management systems

The number of actors involved in generating, managing and providing geospatial information has increased significantly in the last ten years and … [this will] continue and … accelerate in the coming five to ten years.

UN Initiative on Global Geospatial Information Management
July 2013

Government

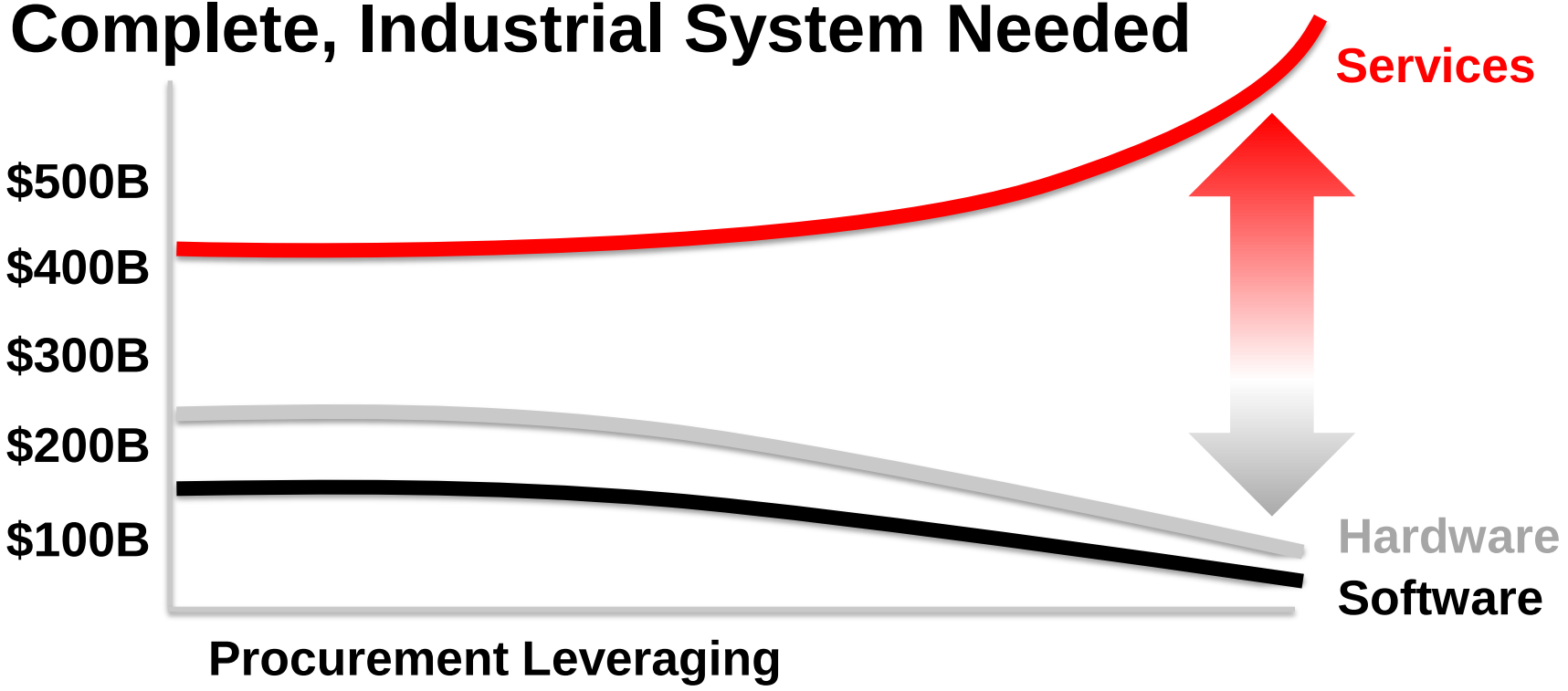


Crowd Sourcing

Sensors/Machines

Business … Future

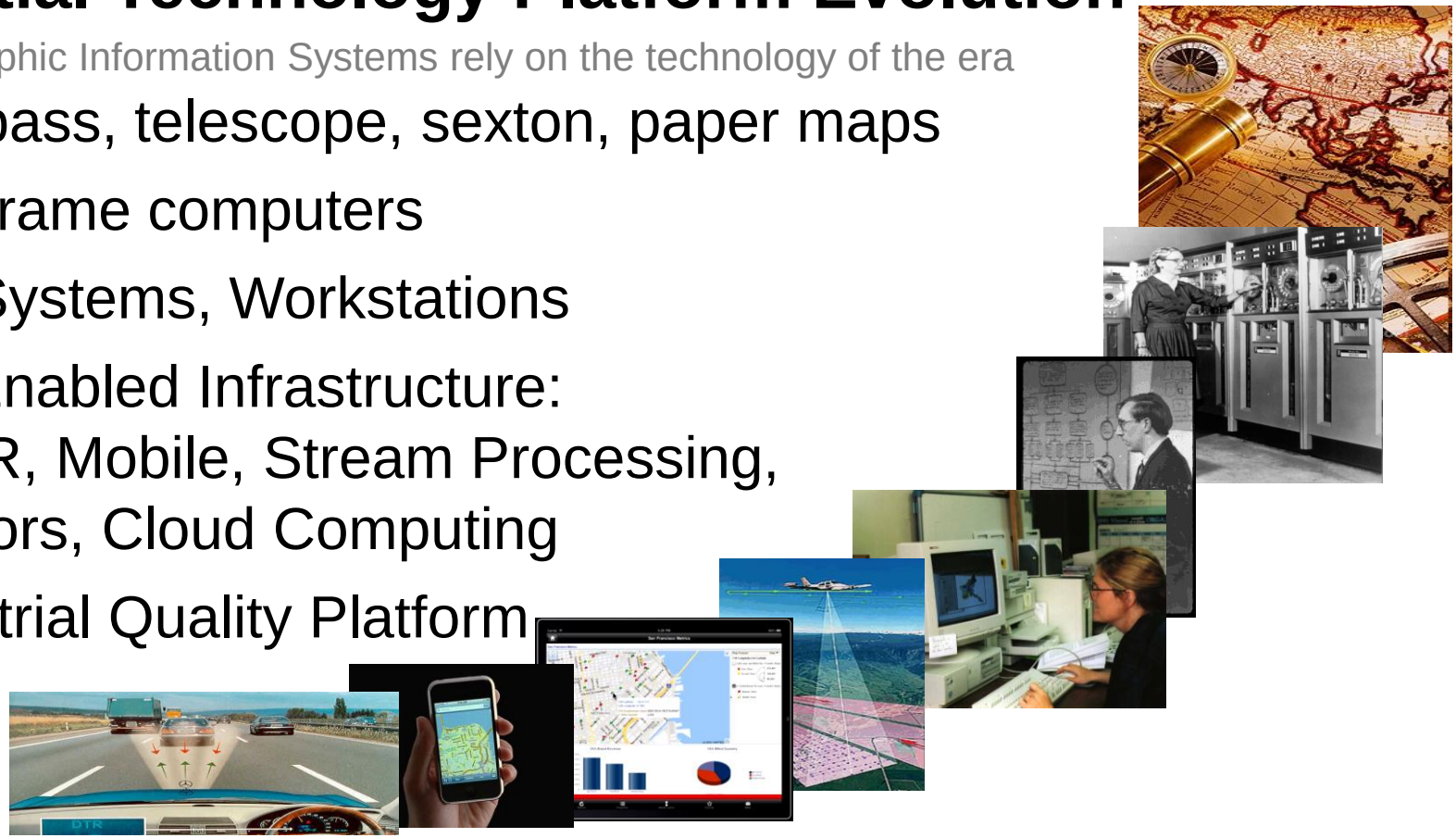
HW/SW Efficiencies: But Labor Costs Growing - Complete, Industrial System Needed



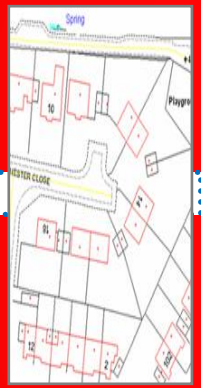
Spatial Technology Platform Evolution

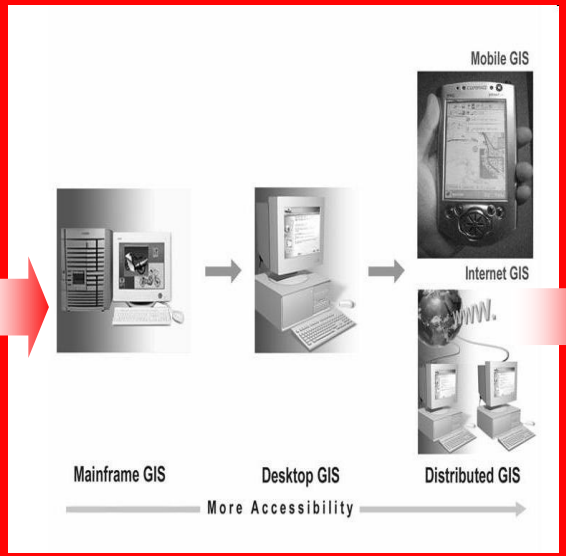
Geographic Information Systems rely on the technology of the era

- Compass, telescope, sexton, paper maps
- Mainframe computers
- GIS Systems, Workstations
- GeoEnabled Infrastructure:
  LiDAR, Mobile, Stream Processing,
  Sensors, Cloud Computing
- Industrial Quality Platform



Industrial Workflows Mandate both Geospatial Technologies and Information Technologies





Mapping

Digital data file

Geographic Information Systems

Spatial Information Technology

Big Data / Spatial in Public Sector

Examples of Different Program Areas

Use Cases

1. Fraud Prevention

2. Maintenance & Utilities

3. Constituent Sentiment

4. Threat Identification

5. Economic Analysis

6. Healthcare

7. Regulatory Compliance, Licensing & Enforcement

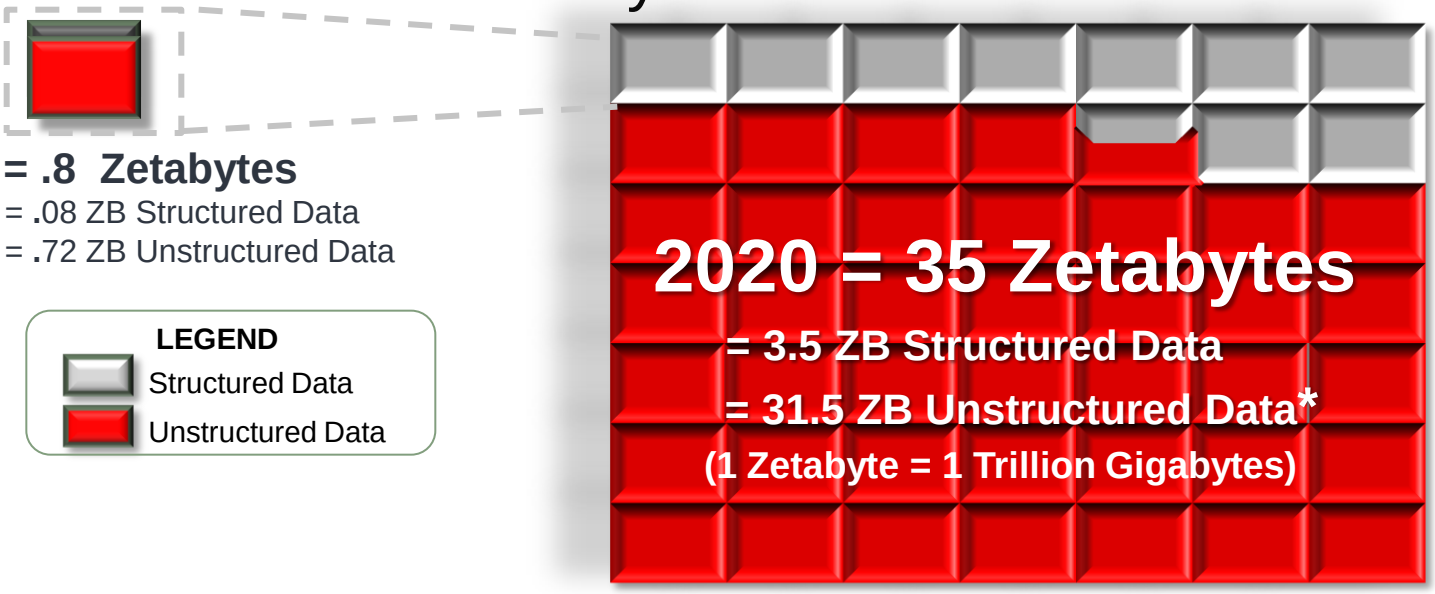
8. Open Government

9. Tax Collections

Global Digital Data Growth: Far Exceeds Storage Mfg

Volume Growing leaps and bounds by 40+% YoY!

YOU Must Make Policy Decisions on What Data to Keep

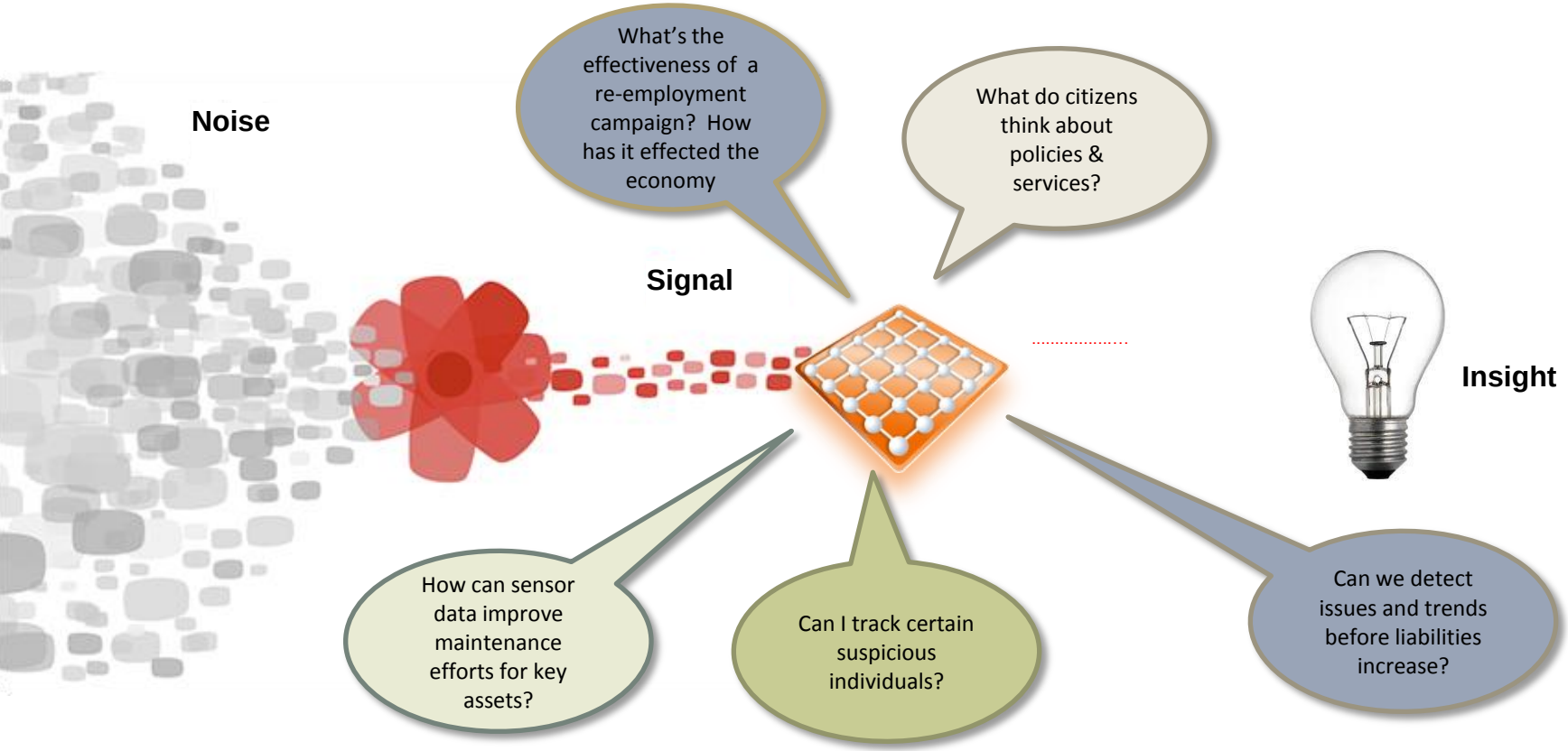


- Chart conservatively assumes a constant 9:1 ratio of unstructured data vs. structured data (based upon IDC's estimate that 90% of all digital data is unstructured).
- Chart does not reflect IDC's projection that unstructured data is currently growing twice as fast as structured data at the rate of 63.7% vs. 32.3% CAGR.

Source: IDC Digital Universe Study, A Digital Universe Decade – Are Your Ready?, 2010


1 NEW DEFINITION IS ADDED ON urban
1,600+ READS ON Scribd
13,000+ HOURS MUSIC STREAMING ON PANDORA
12,000+ NEW ADS POSTED ON craigslist
370,000+ MINUTES VOICE CALLS ON skype
98,000+ TWEETS
20,000+ NEW POSTS ON tumblr.
THE LARGEST SOCIAL READING PUBLISHING COMPANY
New Craigslist Ads
320+ NEW twitter ACCOUNTS
13,000+ iPhone APPLICATIONS DOWNLOADED
100+ NEW Linked in ACCOUNTS
IN 60 SECONDS_
1 associatedcontent NEW ARTICLE IS PUBLISHED
THE WORLD'S LARGEST COMMUNITY CREATED CONTENT!!
QUESTIONS ASKED ON THE INTERNET...
100+ Answers.com
40+ YAHOO! ANSWERS
6,600+ NEW PICTURES ARE UPLOADED ON flickr
600+ NEW VIDEOS
50+ WORDPRESS DOWNLOADS
25+ HOURS TOTAL DURATION
70+ DOMAINS REGISTERED
60+ NEW BLOGS
695,000+ facebook STATUS UPDATES
125+ PLUGIN DOWNLOADS
168 MILLION EMAILS ARE SENT
694,445 SEARCH QUERIES
1,700+ Firefox DOWNLOADS
1,500+ BLOG POSTS
79,364 WALL POSTS
510,040 COMMENTS
Google
Google Search
GO-Globe.com

Detecting Signal in Noise = New Insight



Industrialization of Geospatial Workflows: Drivers:

- BIG DATA GENERATION: – Terabytes, Petabytes, Exabytes, Zettabytes, Yottabytes
  - Sensors, RFID, VIDEO, LIDAR, Raster, 3D, INTERNET OF THINGS
  - SDIs, INSPIRE, Terrain and City Models
  - Social Media, Tagged Data, History / Archive / Version Data
  - Linked Open Data –- Persistent Relationships, Semantics, Ontologies
- BIG but Inexpensive Hardware:
  - CLOUD Platforms – Public and Private
  - More powerful – Clusters of Commodity Servers, Virtualization: = Greener
  - Massively parallel database machines – Software Enablement – e.g. Hadoop
- BIG Software & INDUSTRIAL WORKFLOWS
  - Location Enable All Applications: ERP, CRM, Business Intelligence, Public Sectors
  - REAL TIME Analytics –Biggest value from fastest response – Streams and Events –- Spatially Aware System – no separate GIS
  - CyberSecurity, Encryption, Privacy
  - Support Standards – W3C, OGC, ISO, Wide Range

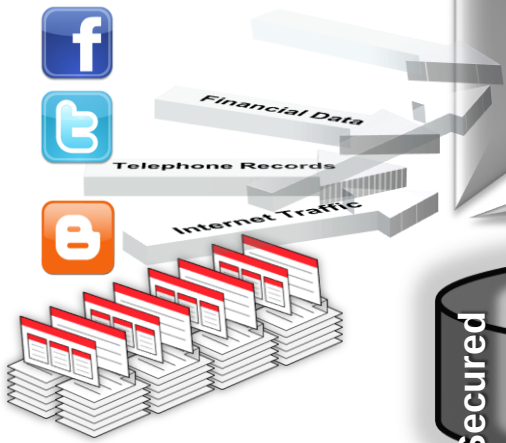
Big, Fast, Aggregation & Workflows = SHAREABLE, REPURPOSABLE LOCATION DATA

- DATA / INFORMATION –
  - Too much to store it all
  - Arriving too fast for humans to process – Need Automated Workflows
    - Must use Real Time Filtering and Analytics
    - This is the Big Data / Hadoop filtering & CEP - Complex Event Processing
    - Set policies on what to keep, what to expunge
  - Must share data among your many Organizations, enabling Aggregation
    - Geography/Mapping/Location, Health Care, Statistics, Commerce, Taxation
  - SHARING requires Interoperability and Semantics / Ontologies / LOD
- This is a DATA MANAGEMENT TOPIC, not a GIS issue.

Your Employee Workflow for Analysis

External Data Sources

Transactional &
Operational Systems
Contents Repository
Databases
Web resources
Blogs, Mails, news



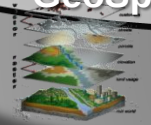
Real-time Data Streams

Search, Presentation, Report, Visualization, Query

Enterprise Data Management Infrastructure

Secured

GeoSpatial



Historical Records

POIs
Demographics
Customer Data
Call Records

Documents

Automatic Responses and Publishing

SMS

Console Alerts

EV Grid Management

Workflow Initiation

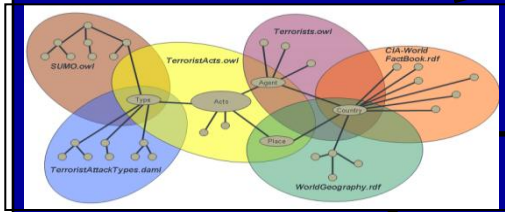
Real-time Dashboards

Ontology-driven Geospatial Workflows - Connect Actionable Knowledge

National Mapping Private Cloud

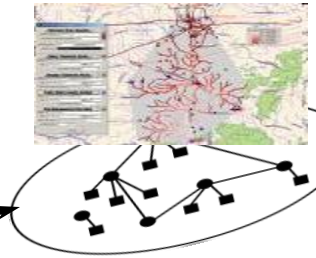
- Simple Features
- GeoRaster
- Topology
- Networks
- Gazetteers
- …

RDF & OWL Metadata

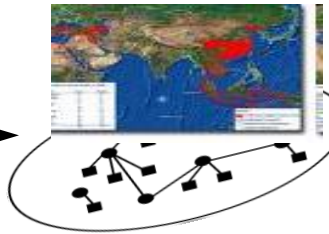


- Data Integration
- National Map schemas
- Geographic names
- Temporal
- Naïve Geography
- …

Application Ontologies



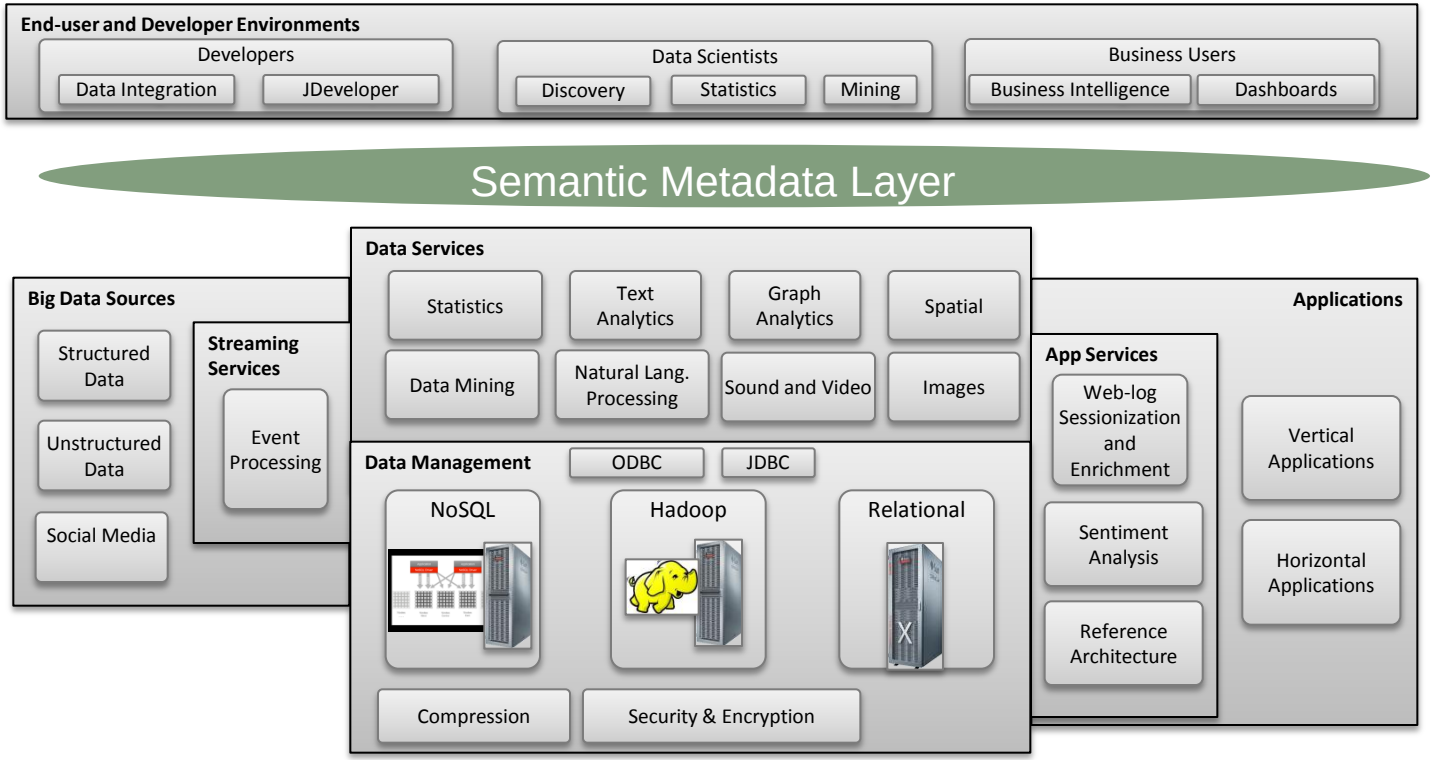
Environmental Monitoring



Famine Relief

Disaster Response

Workflow Supporting Breadth of Enterprise Data



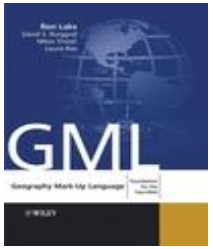
Buzzwords For Apps & Workflows using Graph Technology: What terms to look for:

- Semantic Web
- W3C RDF/OWL/SPARQL
- Graph Data Management
- Social Network Analysis (SNA)
- Knowledge Discovery
- Knowledge Mining
- Big Data
- Taxonomy/Terminology Mgmt
- Faceted Search
- Inferencing / Reasoning
- Property Graphs
- Sentiment Analysis
- Text Mining
- NoSQL Database

You Get Order And Efficiency Using Standards

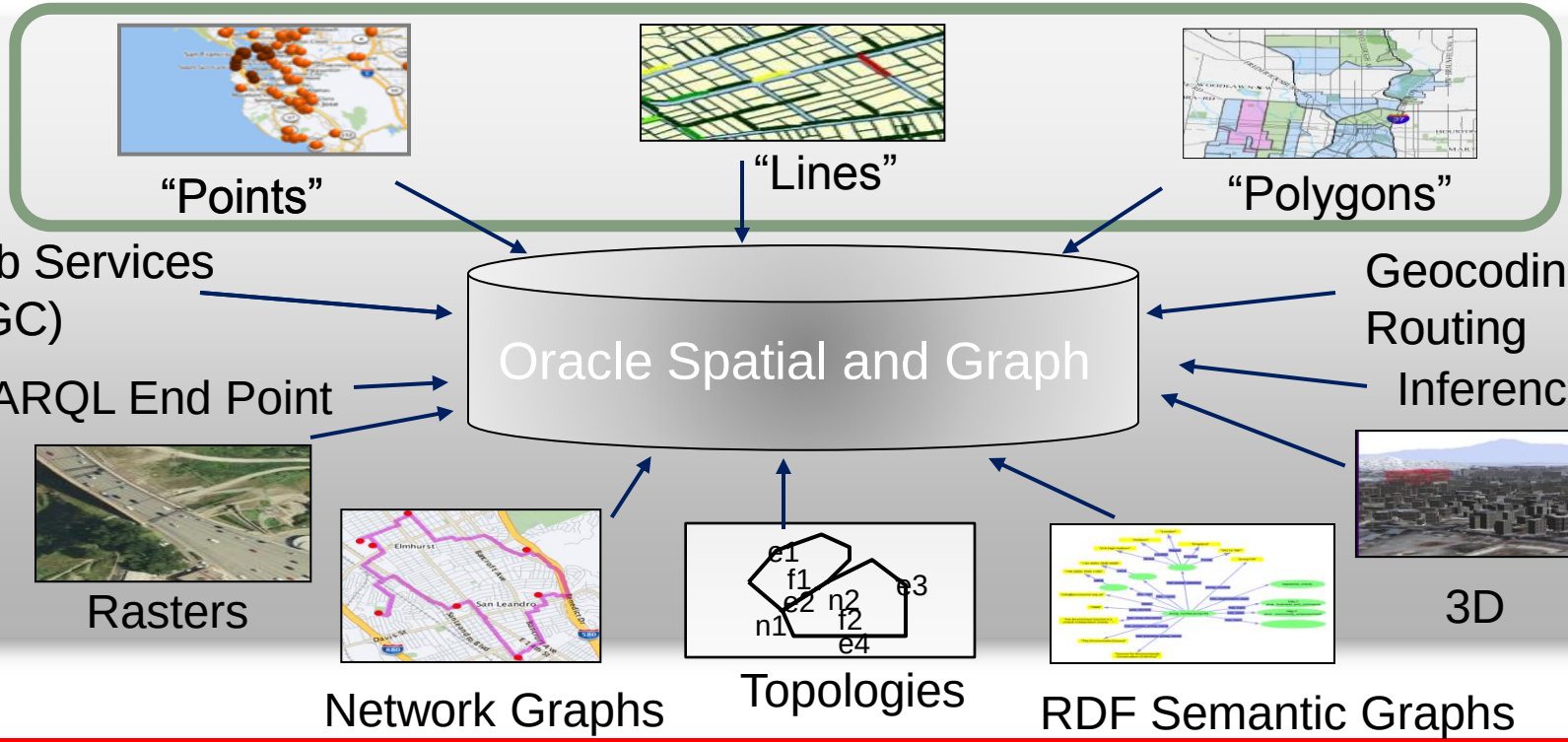
"We intend to complete development for a new suite of tools for developing the next generation of applications. And there are several interesting things with the next generation of tools, but perhaps the single most interesting thing about them is that for the first time a major application company is going to commit to an absolute standards-based development environment."
– Larry Ellison

- **ISO**
  - **TC 211**
  - **TC 204**
- **Open Geospatial Consortium**
  - **Simple Features**
  - **GML**
  - **Web Services**
- **De-facto Standards**
  - **SHP, MGE, DXF, KML**
- **Professional Standards**
  - **ISPRS, FIG, WMO**
- **Java, .NET, Flash**
- **TAGGED METADATA – agree on tags**

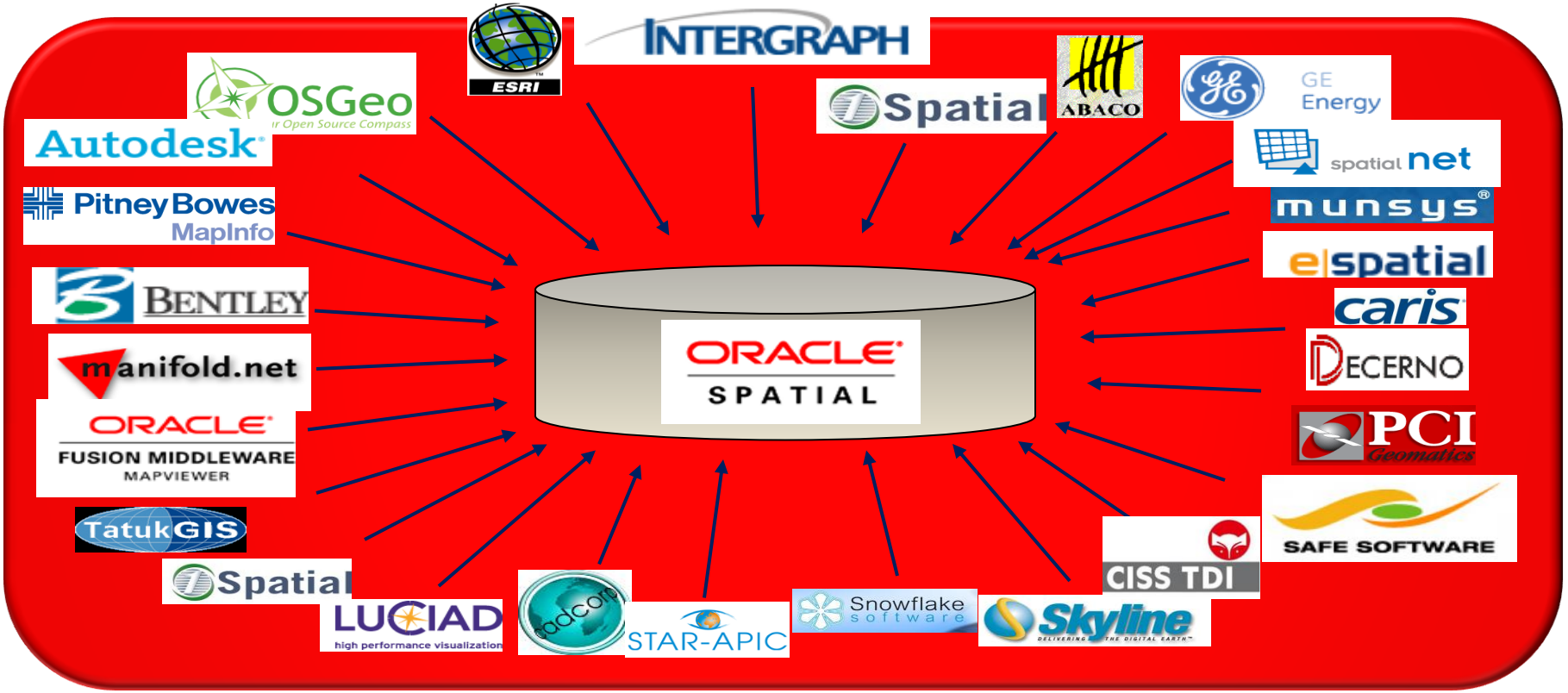


SQL3/MM Spatial

Oracle Spatial and Graph option



Open and Interoperable

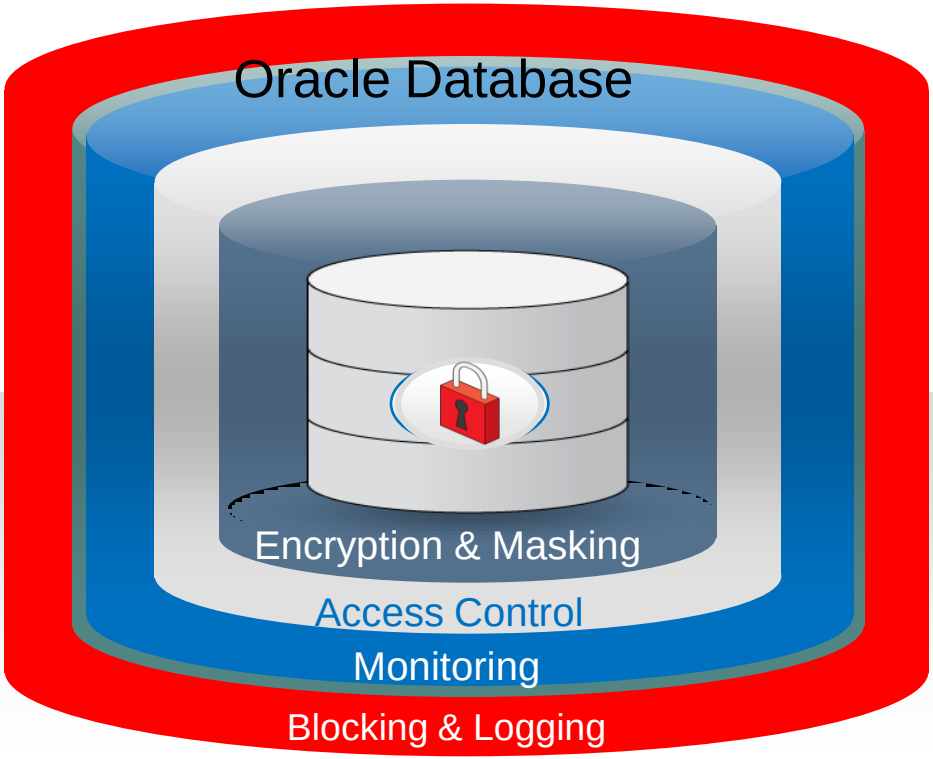


SIMPLIFY WORKFLOWS – AUTOMATE IT

- SIMPLIFY BASIC PARTS OF WORKFLOWS
- YOUR PLATFORM MUST AUTOMATE STANDARD IT ACTIONS

Connecting: CYBERSECURITY is Major Challenge

Information Security and Privacy



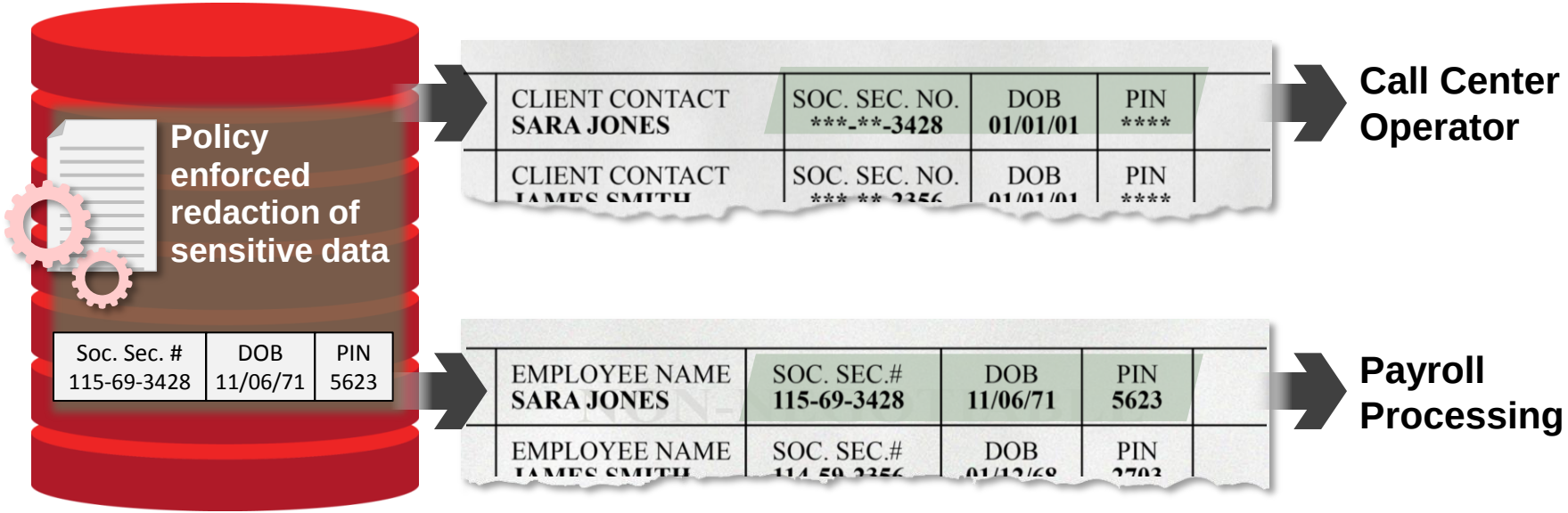
Monitoring
- Configuration Management
- Audit Vault
- Total Recall

Access Control
- Database Vault
- Label Security

Encryption & Masking
- Advanced Security
- Secure Backup
- Data Masking

Redacting Sensitive Data

Mask Application Data Dynamically

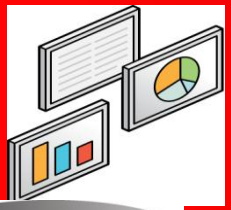


ILM: Hot/Cold Data Classification

Enhanced Insight into Data Usage: “heat map”

Recently inserted, actively updated

Infrequently updated, Frequently Queried

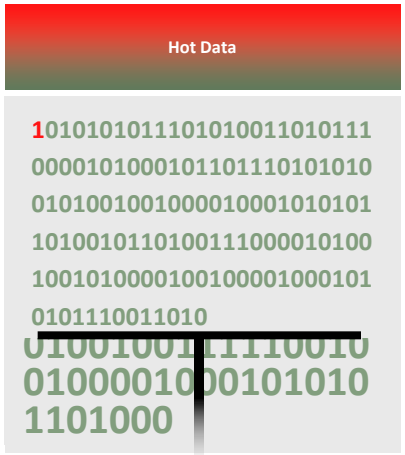


Retained for long term analytics and compliance with corporate policies and regulations

Block and Segment level statistics on last Read and last Update

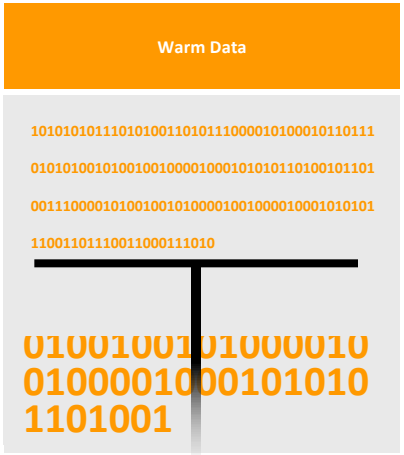
Data Compression

Reduce storage footprint, read compressed data faster



3X

Advanced Row Compression



10X

Columnar Query Compression

15X

Columnar Archive Compression

Why Build A Hadoop Appliance? Or DBMS?

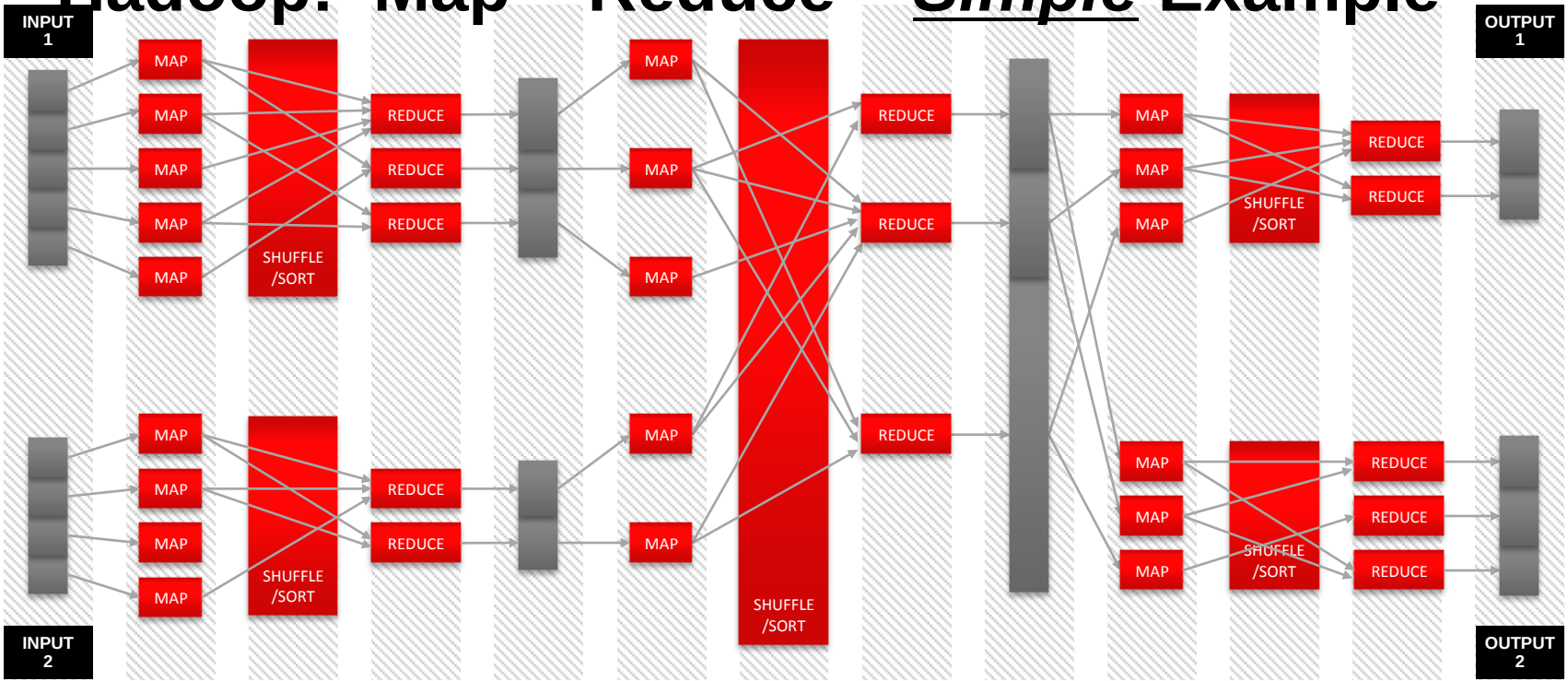
UN-GGIM: “train the individuals is at least five years”

Time to Build

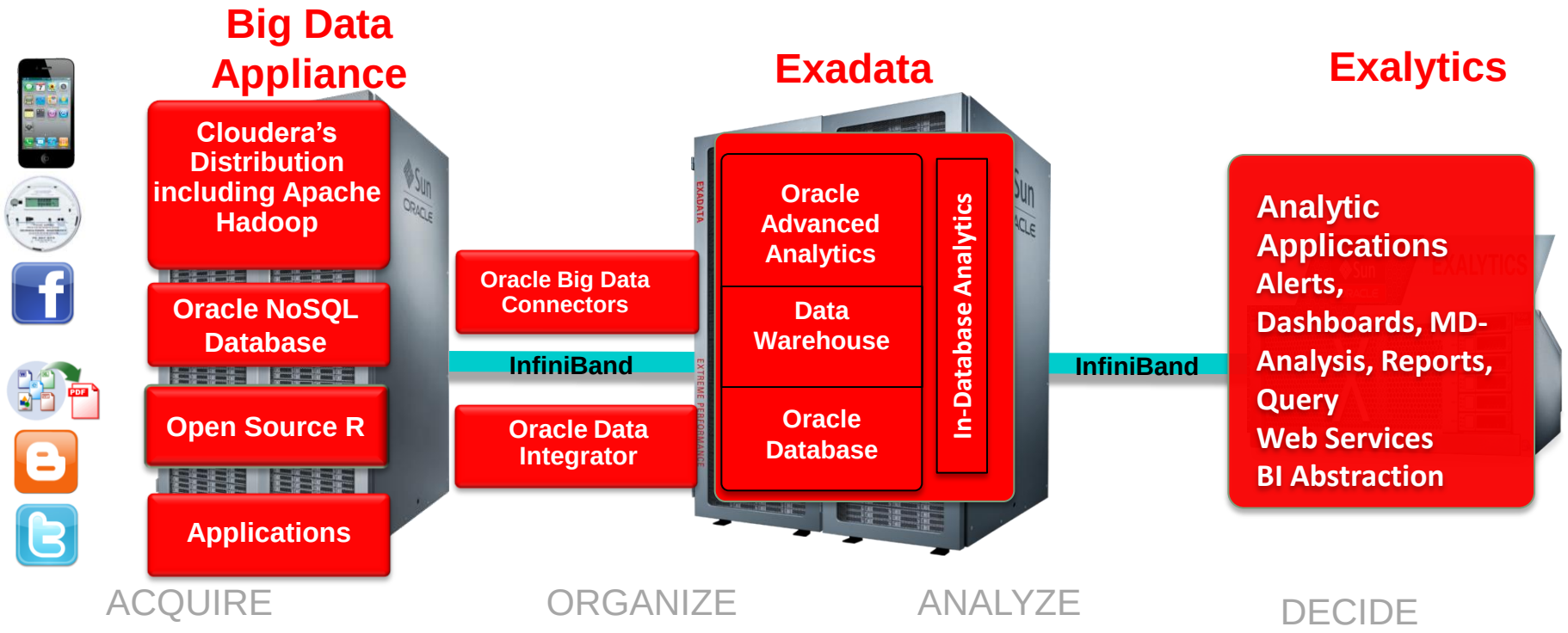
Optimizations

Maintenance

Big Data: Batch-Oriented Processing using Hadoop: Map – Reduce – *Simple* Example

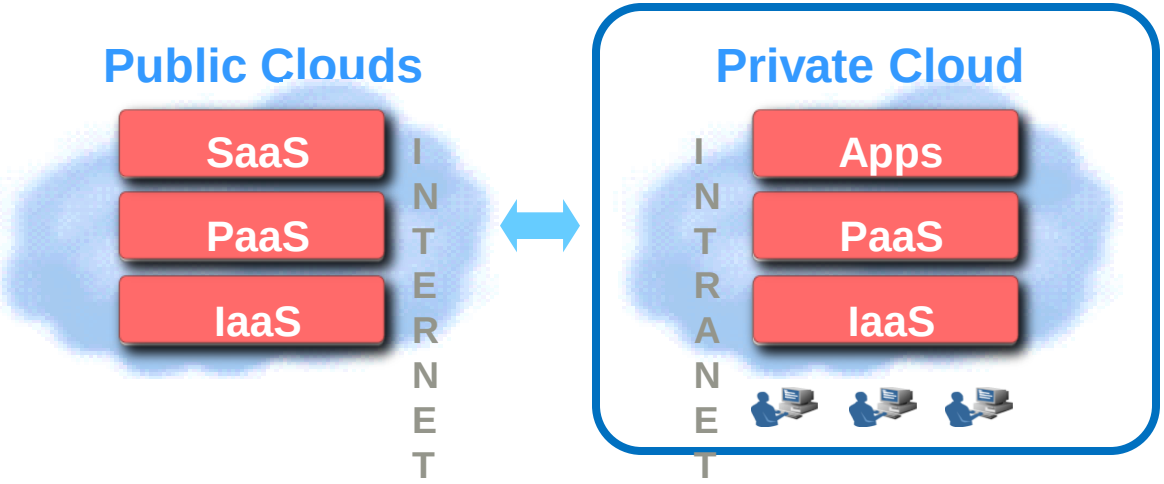


Oracle Industrial Geospatial Workflow Platform



Public Clouds and Private Clouds

- Used by multiple tenants on a shared basis
- Hosted and managed by cloud service provider



- Exclusively used by a single organization
- Controlled and managed by in-house IT

Trade-offs

Lower *upfront* costs	Lower *total* costs
Outsourced management	Greater control over security, compliance, QoS
OpEx	CapEx & OpEx

Oracle Technology Supplies both Public and Private clouds

Oracle's Spatial Stack

Built-in Geospatial features

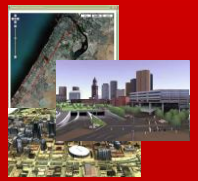
- Spatial analytics in Applications and tools
- Fusion Middleware MapViewer, Event Processing
- Oracle Spatial and Graph database
- Bundled and Cloud-based Map Content
- Designed for Exadata

Industrialization of Geospatial Workflows: Best Success Requires Complete Platforms

Big & Fast Data

Volunteered Geographic Information

Sensors Streaming Data



Geo-referenced Video, 3D, LiDAR

Simplified Spatial IT

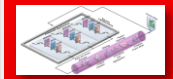
Support for Open Standards

Spatial Database, Application Server, BI, tools

Support by Leading Partner solutions

Spatially-enabled Engineered Systems

Deep Analytics



Real-time Spatial Event Processing

Dense Visualization

Spatial Analysis

On Premise, On Cloud, Shared Services

Shared GeoSpatial Services
Location Aware Everything

Cloud Computing

+

Oracle Engineered Systems