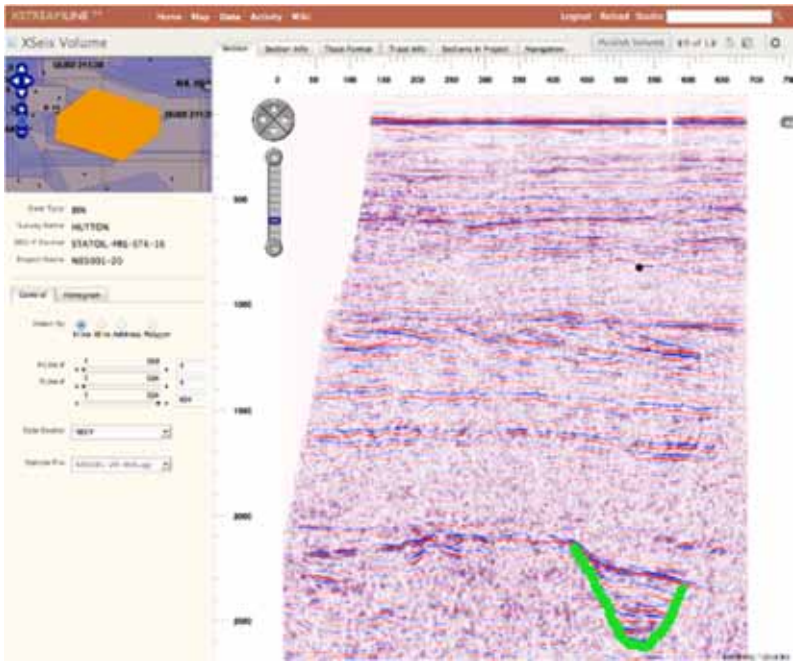
TERADATA | THE BEST DECISION POSSIBLE™

Tackling the *big data* challenges in E&P

Dr Duncan Irving, EMEA Oil and Gas Practice Lead

What if you could…

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?



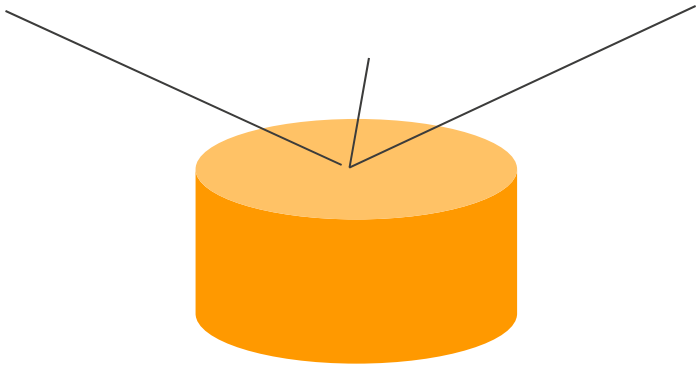
Seismic Volumes

Reservoir Models

Well Sensors

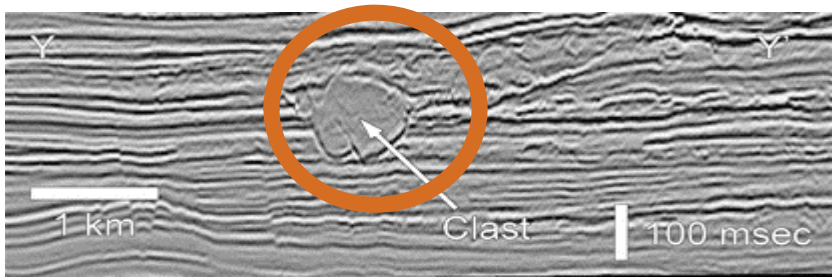
What if you could…

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?

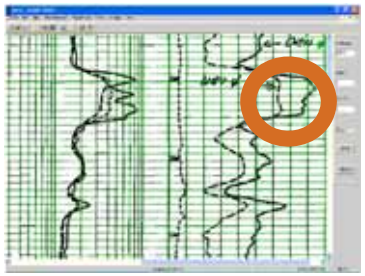


What if you could…

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?



+



+

What if you could…

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?

- guarantee data custodianship based on standards?
- know who knew what, and when?

What is inhibiting this?

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?
- guarantee standards-based data custodianship?
- know who knew what, and when?

- browser capabilities
- network bandwidth
- data access
- data volumes
- compute loads
- poor data model/structures
- 70% of time managing data
- analytical compartmentalisation
- application-centric view
- file/transfer formats rule
- no community ownership
- no granularity
- temporally-enabled data governance

What is out there to help?

- perform all E&P analytical activities through a web browser?
- work collaboratively on a single instance of the data?
- guide the science rather than drive the PC?
- guarantee standards-based data custodianship?
- know who knew what, and when?

- data access
- data
- compute loads
- poor data model/structures
- 70% of time mana
- analytical compart
- application-centric
- file/transfer formats ru
- no community owners
- no granularity
- temporally-enabled data governance

RDBMS

"High Performance XXXX"

"Big Data Solution"

appliances

Data Warehouse

Why can't I make sense of it all?

- "I am old"
- "I am an enterprise architect, not a web programmer!"
- "I am a web programmer, not an enterprise architect!"
- "I see red and twitch involuntarily whenever I hear *Big Data*"
- **I have all of these things and I'm still confused**

Even with all these pieces, no one has yet managed to bring all necessary data to bear on business questions in a useful timeframe

RDBMS

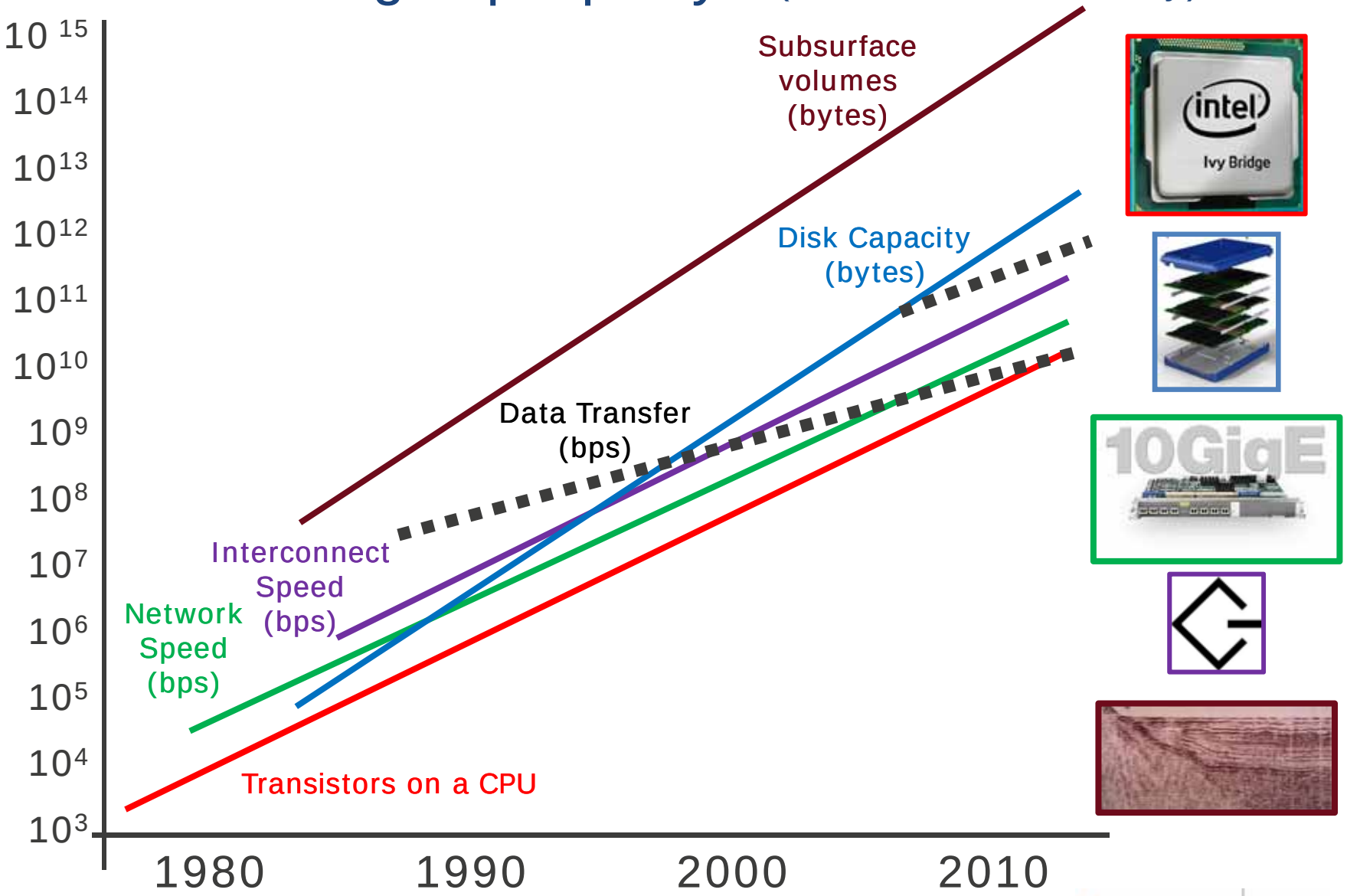
"High Performance XXXX"

"Big Data Solution"

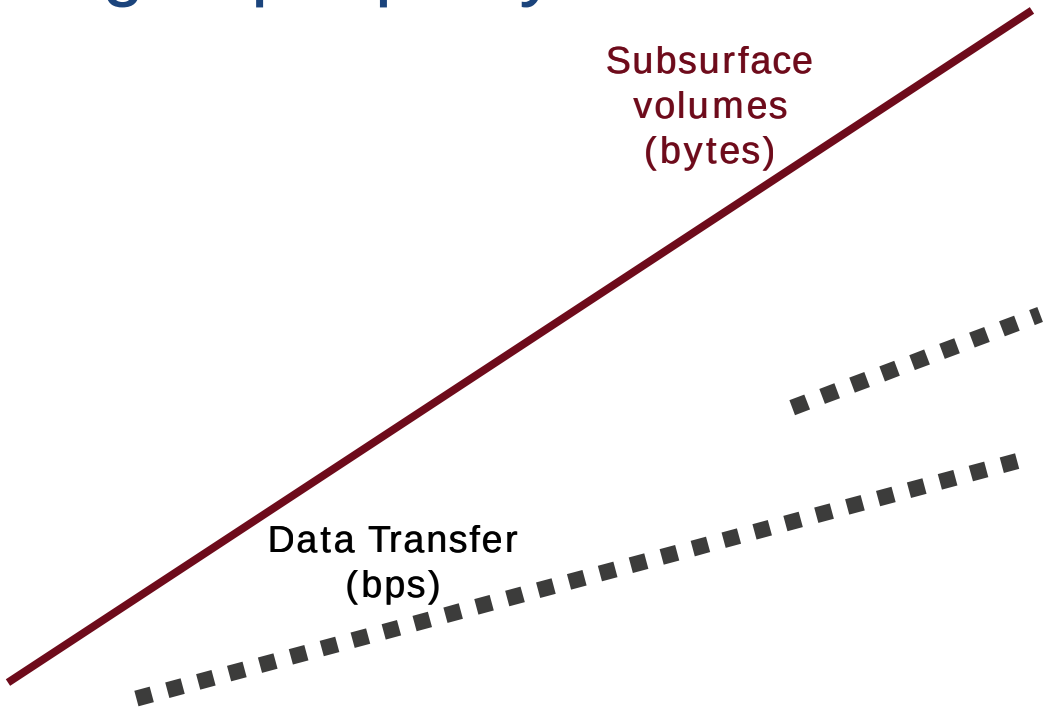
appliances

Data Warehouse

Am I building it properly? (A lesson from history)

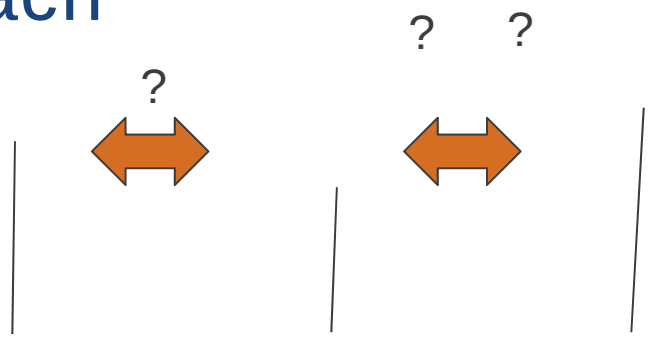


Am I building it properly?



Questioning power is not keeping up with the potential value of the answer.

Application-centric approach

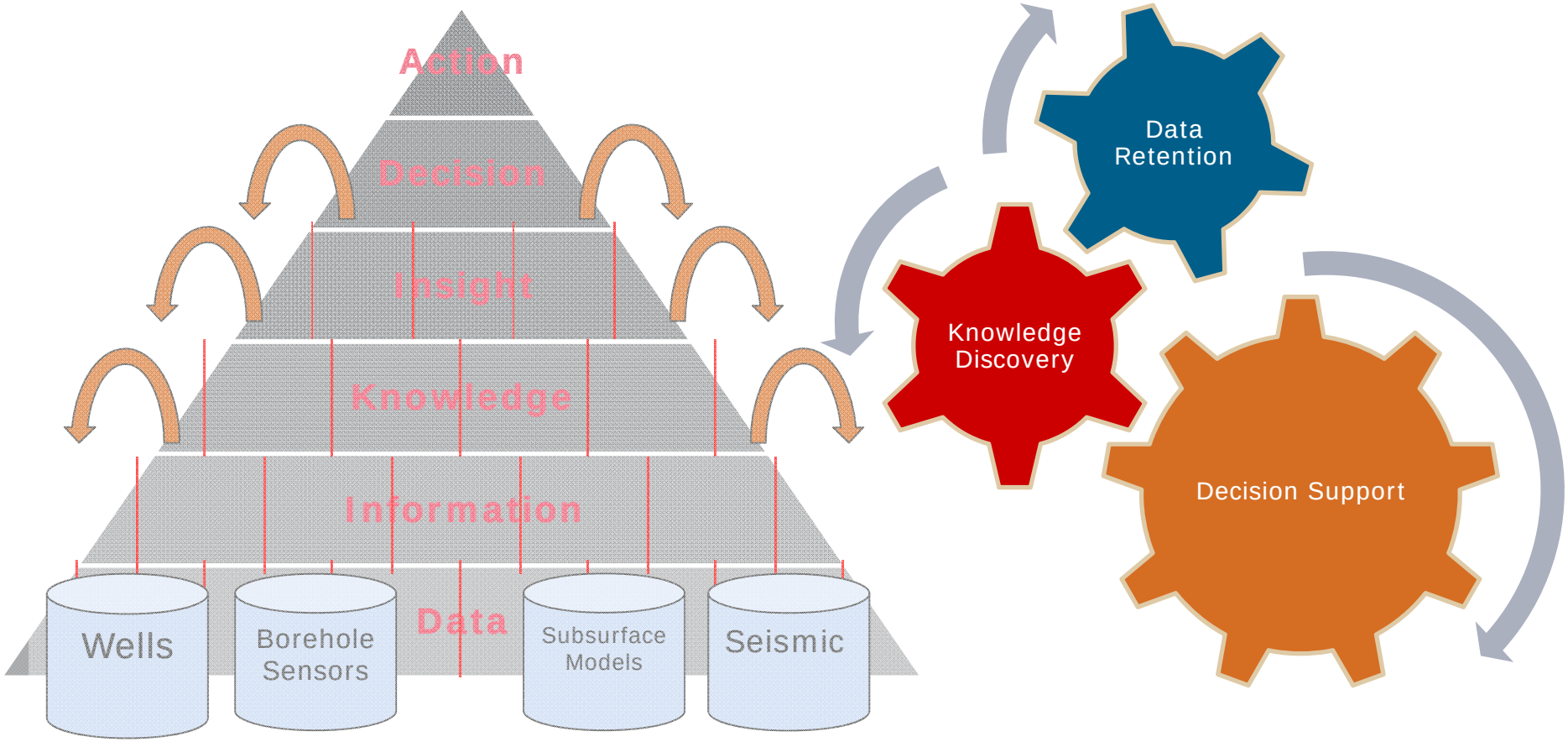


The quest to integrate decision making led to:

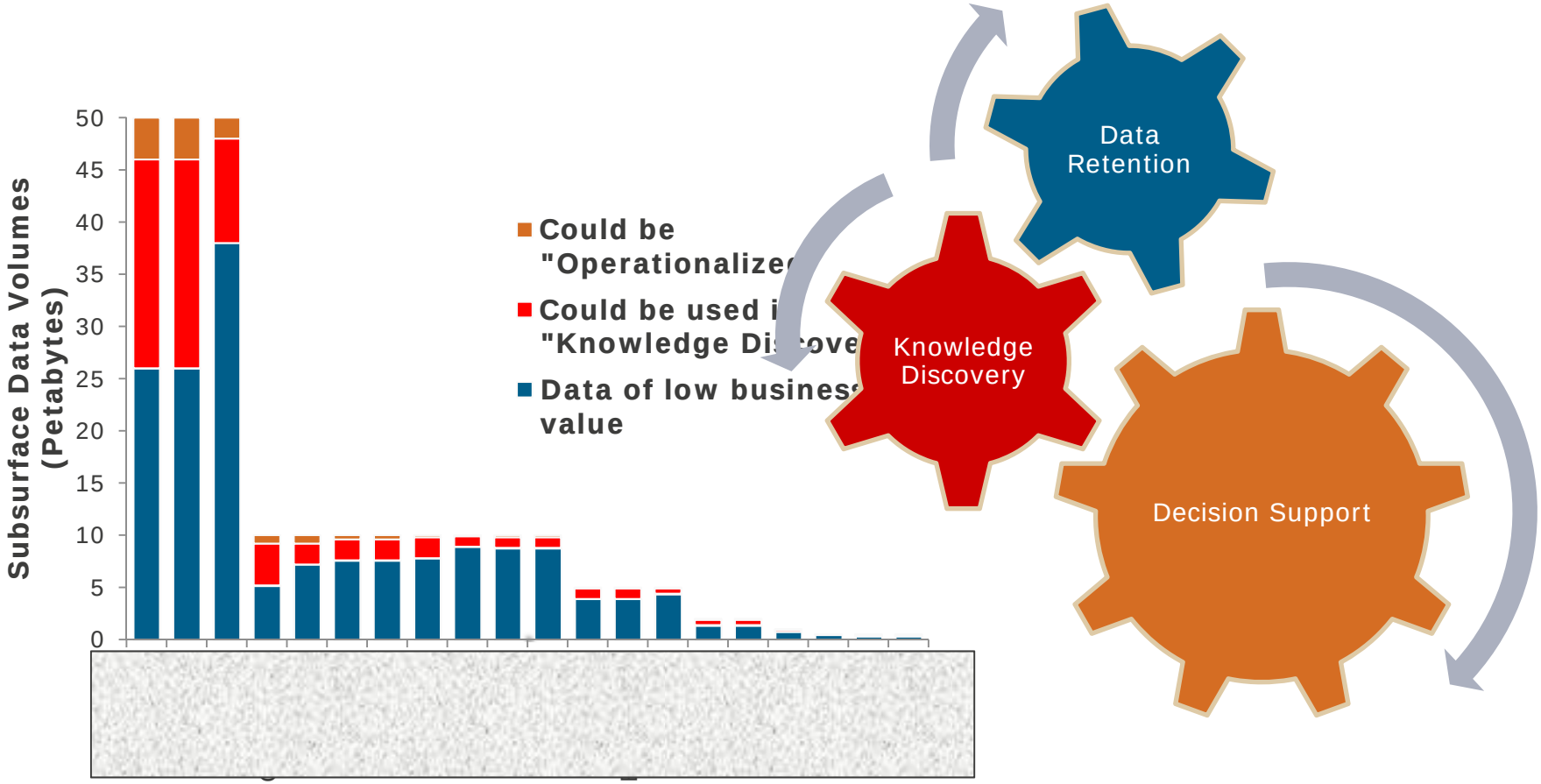
- Application Service Provision
- Collaborative Visualisation
- Virtualized workstations
- Proprietary file formats

- No long-term stewardship of:
  - File formats
  - Decision milestones
  - Data dependencies
  - Information management

Analytical Compartments v Data Flow

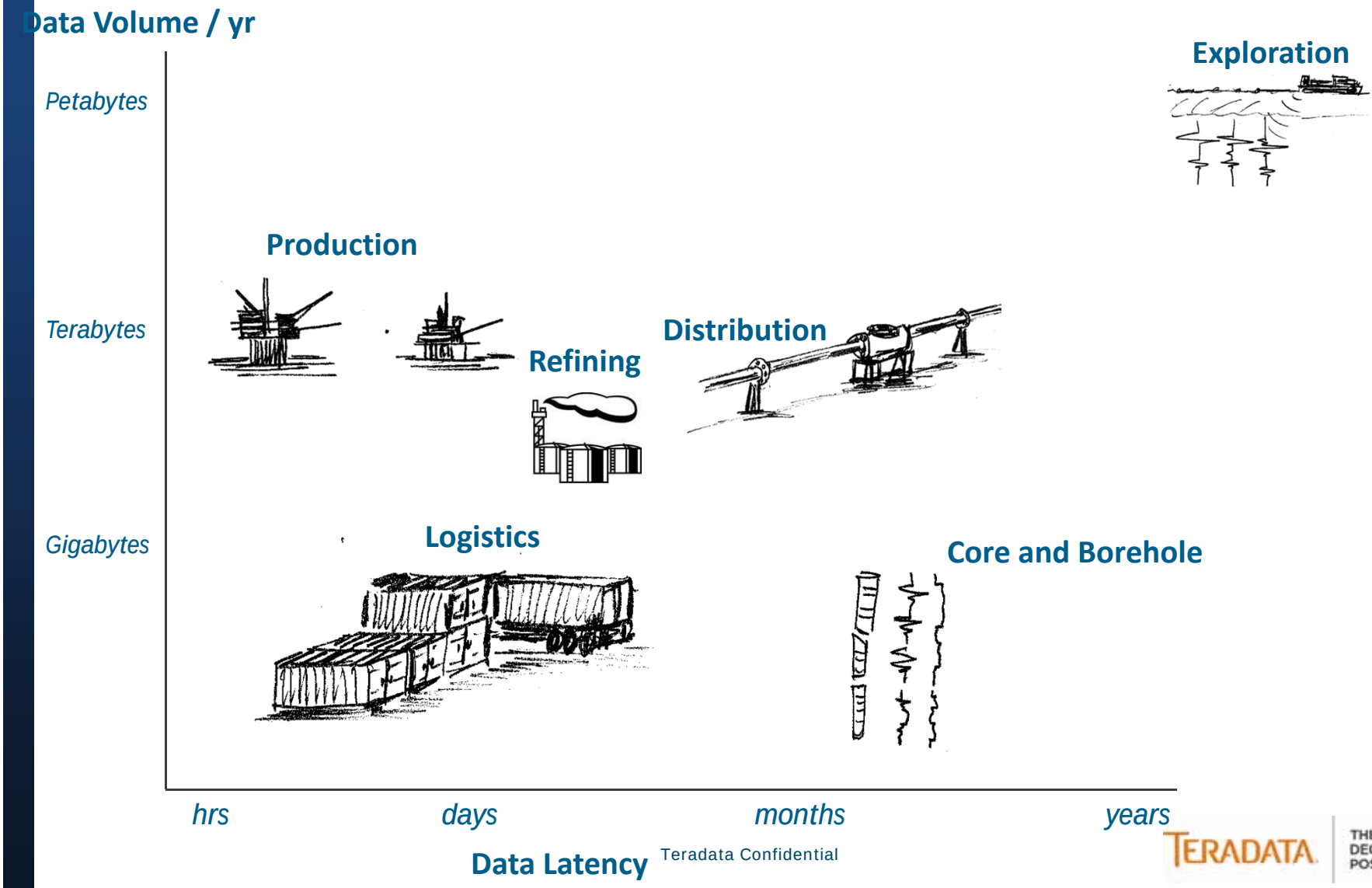


E&P Data Usage Modes



Caveat: These numbers are half-baked estimates.
Error is ~50%

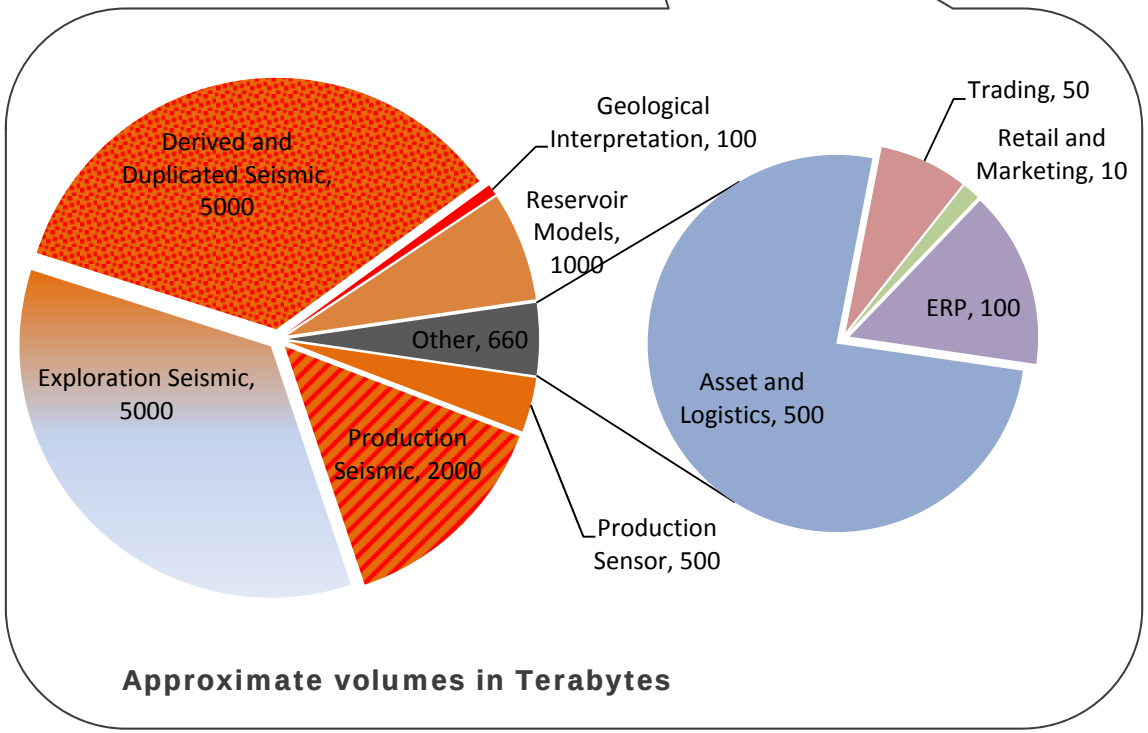
The E&P analytical landscape



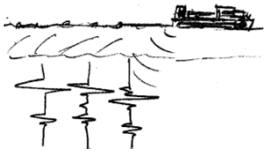
The E&P analytical landscape

Many other industries are curing their "big data" problems:

- Analytical Integration
- Massive data volumes
- Mixed workloads
- Query Concurrency



Exploration

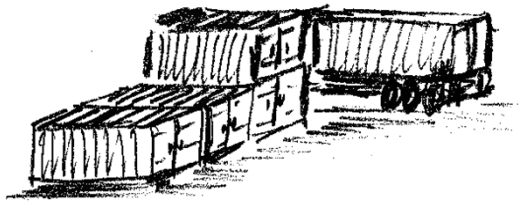


Production

Refining

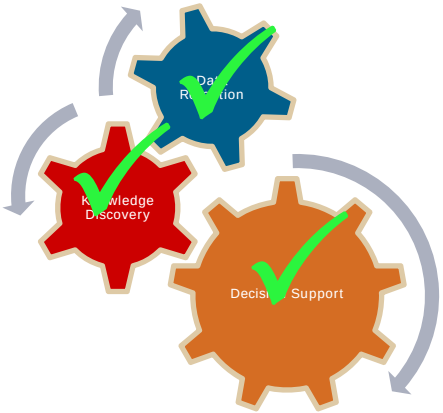
Distribution

Logistics

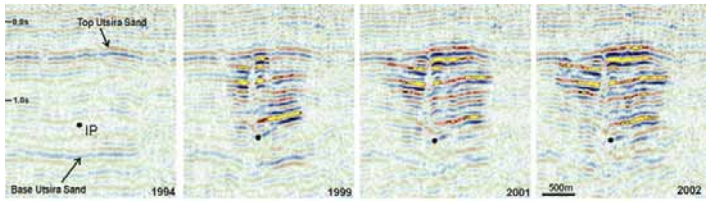


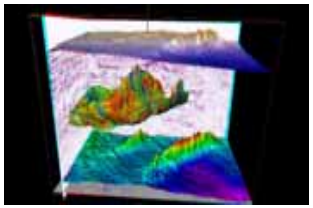
Integrated Operations Analytics

Wellfield Data Warehousing in a supermajor



How to do a rapid comparison between “new” and everything ever seen – and make rapid decisions on it.





Carry out $100M+ seismic survey every three years to re-image producing reservoir

Spend 2-3 months reprocessing data and reincorporate into workflow

Hope geoscientists and engineers can control reservoir flows at the weekly scale based on imaging from the year+ timeframe

Current approach

Install seafloor seismic imaging array and stimulate with in-reservoir tectonic events, and supply-vessel based airguns

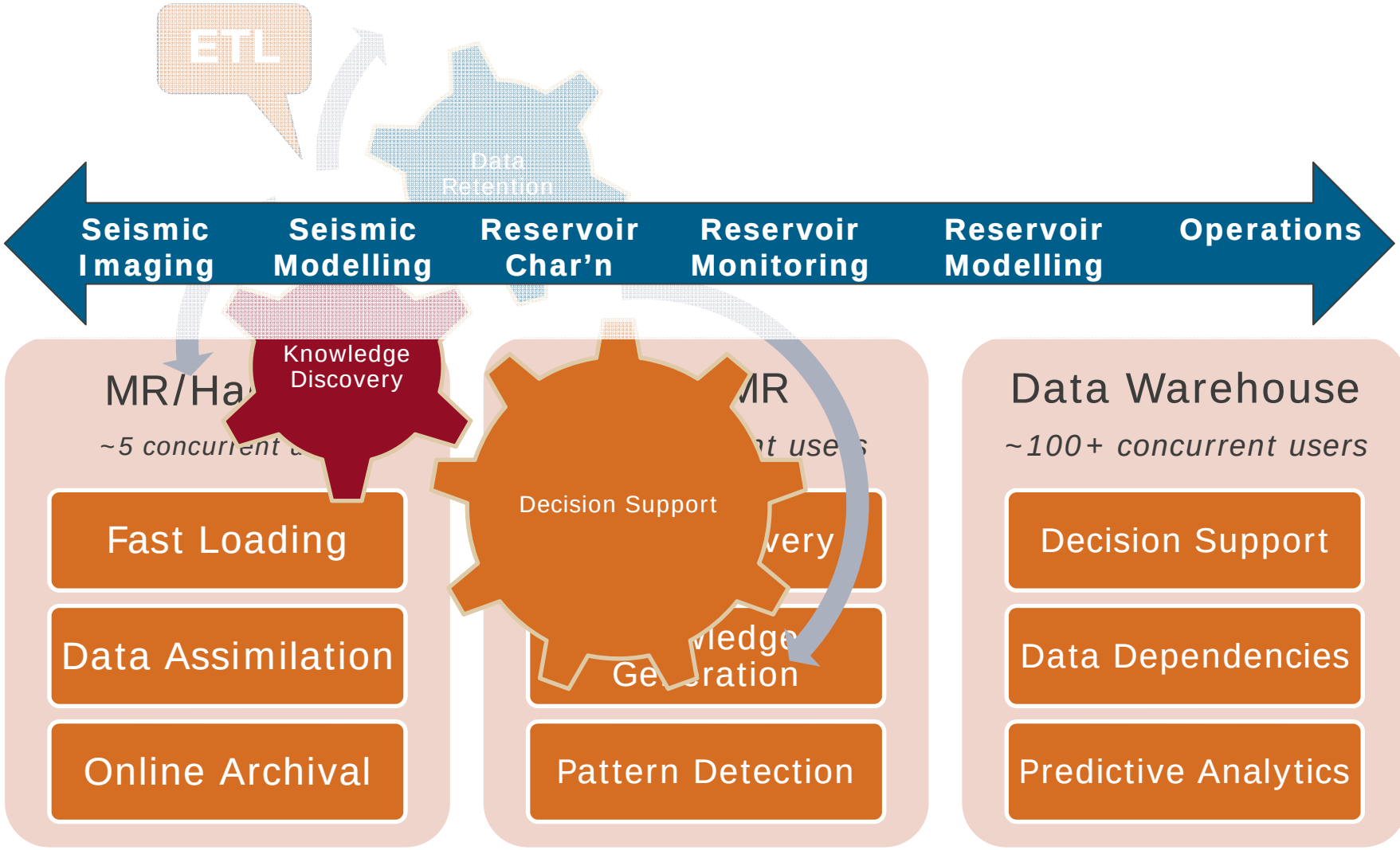
Spend 2-3 days reprocessing data and reincorporate into workflow

Allow geoscientists and engineers to respond to HSE and production issues to see how the reservoir is evolving in a right-time manner!

Ideal approach

Shortened Timeframe

Best of Breed Big Data Architecture



MapReduce/Hadoop

- **Is Not** a database
  - HDFS
  - No schema, indexes, optimizer
  - No high availability, security
- Not high performance
- Not a data warehouse
  - No integrated data, no history
  - Severe data skew
- Not mature technology
  - Early open source
  - A few ISV tools integrating
  - Single points of failure
- Not a cloud technology, per se

Uses

- Seismic Processing
- Trace Sorting
- 1D/2D filtering and transformation
- Online Seismic Archiving
- Repurposing WITSML and other sensor feeds

It is a great place to “get to know” your data.

SQL-MapReduce

- **Is** a database
  - MPP database
  - schema, indexes, optimizer
  - security
- **Is** high performance
- Not a data warehouse
  - No integrated data but it can talk to a DW via SQL
- Not mature technology
  - A few ISV tools integrating
  - Single points of failure

Can enable AaaS

Uses:

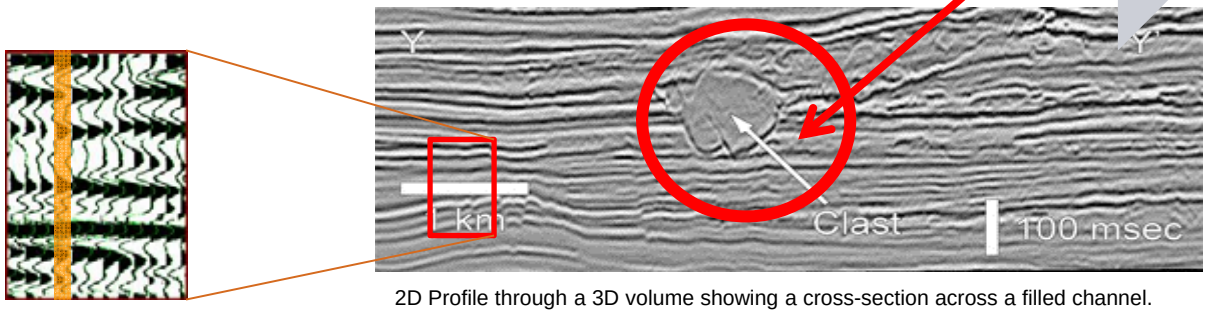
- Pattern matching
- Feature extraction
- "Spot the difference"
- Statistical investigations:
  - Clustering
  - Likelihoods
  - Associations
- "Fail fast" hypothesis reduction:
  - Seismic Modelling
  - Reservoir Modelling frameworks

How do I describe a pattern in SQL-MR?

Simple partitioning in-trace (vertical) analytics and adjacent trace analytics

Broader pattern matching – how do I state spatial relationships, meso-scale textures? easy with SQL-MR!

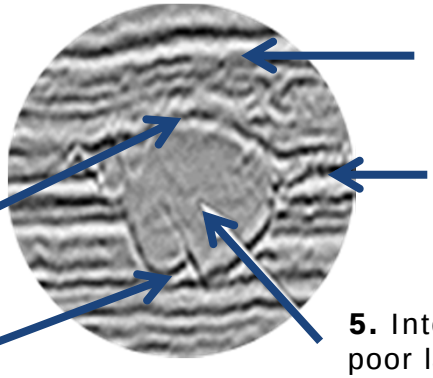
How do I find everything that "looks like" a channel in my 10 Tb 3D (inc. pre-stack) image volume?



2D Profile through a 3D volume showing a cross-section across a filled channel.

The Power of SQL-MR

Java MR functions perform the standard textural descriptions (reflectivity, variance, etc) and **SQL** asks the questions – above, inside, below.



1. Capping boundary: Strong, broadly convex reflection (**nPath**)

2. Basal boundary: Very strong, grossly concave reflection with local minima (**nPath**)

3. Overburden facies: laterally moderately continuous with high amplitude stratigraphy (**nPath**)

4. Incised facies: laterally continuous (low variance) and high amplitude stratigraphy (**nPath**)

5. Internal facies: moderate variance, poor lateral continuity and steeply-dipping beds (**statistical description**)

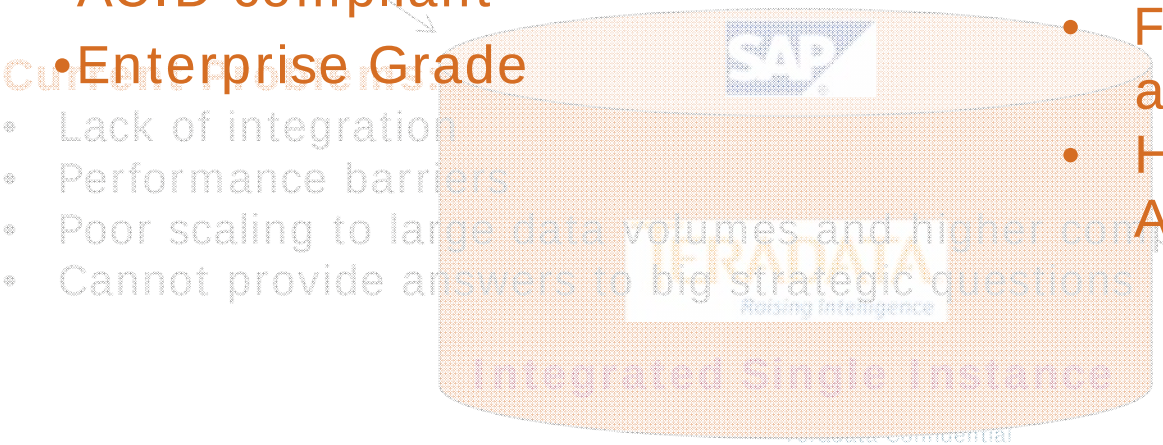
The Active Data Warehouse in E&P

Operationalised DSS

- Master Store for all data
- data is stored in a manner that allows it to be useful

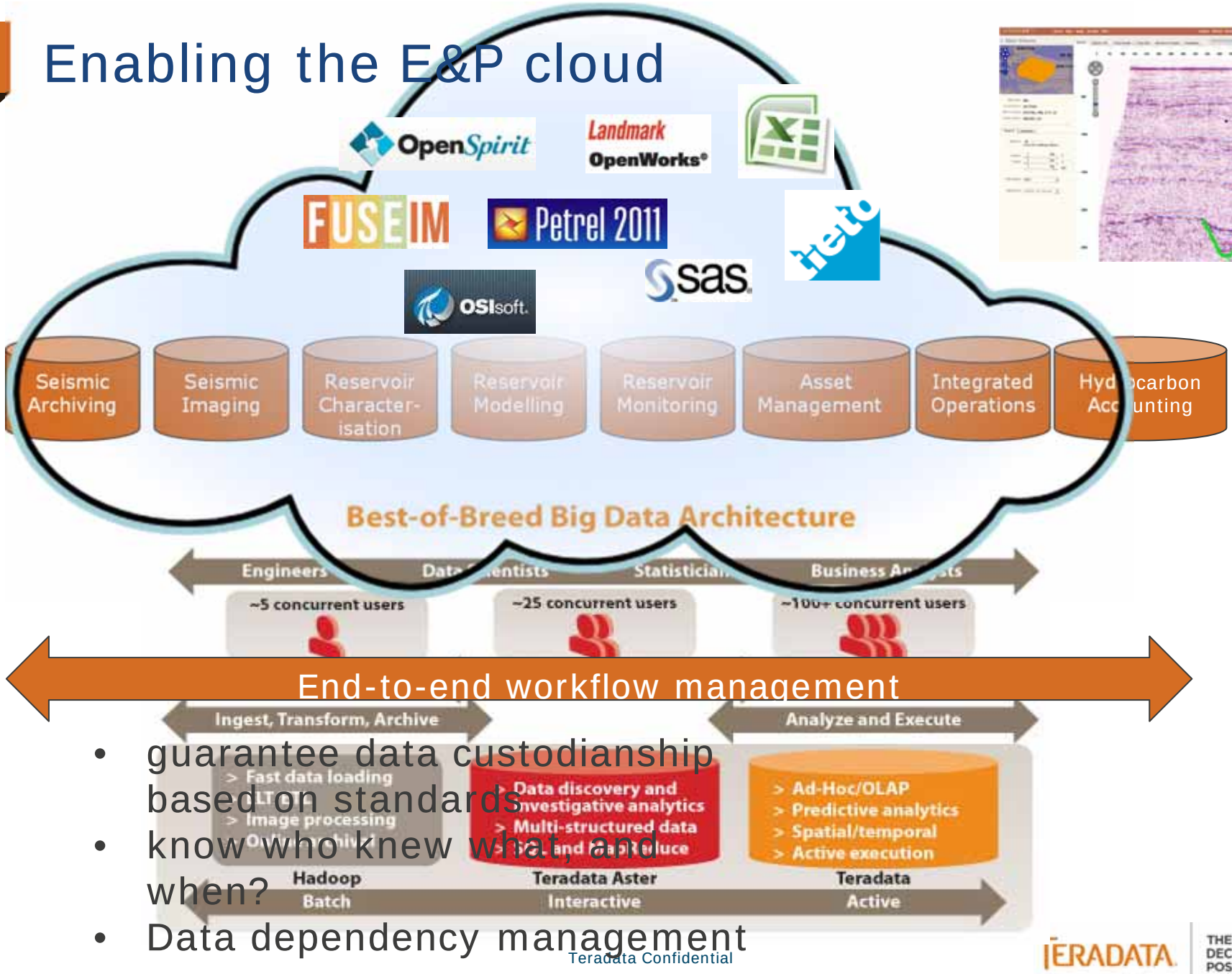
Features

- Fully-integrated
- Highly relational
- Fully-decomposed
- ACID compliant
- Enterprise Grade



- Establish workflows to integrated operations
- Integrate Subsurface insights with Asset and Maintenance actions
- Link production activities via Logical Data Models to other areas of the business
  - Permanent Seismic Monitoring
  - Fracturation Models and regimes
  - Hydrocarbon Accounting

Enabling the E&P cloud



- guarantee data custodianship based on standards
- know who knew what, and when?
- Data dependency management

THE BEST DECISION POSSIBLE™

TERADATA

TERADATA ASTER

aprimo
A Teradata Company