SMART DATA FAST.™

VOLTDB

REAL-TIME BIG DATA ANALYTICS IN THE IBM SOFTLAYER CLOUD WITH VOLTDB

OUR SPEAKERS

John Hugg
Founding Engineer
VoltDB

Pethuru Raj Chelliah
Infrastructure Architect
IBM SoftLayer

Skylab Vanga
Infrastructure Architect
IBM SoftLayer

THE VOLTDB ARCHITECTURE FOR FAST DATA

✓ Per-machine performance + shared nothing scale out

✓ Millions of SQL statements per second per machine

✓ Up to 500k ACID-Transactional *write* proc calls per sec

✓ Full disk durability and sync cluster replication for HA

✓ Powerful tools for streaming analytics

    ✓ Continuous queries, correlation, windowing, aggregation…

✓ Hadoop ecosystem integration

VoltDB is really different than everything else

THE MODERN OPERATIONAL APPLICATION

Fast Data   =   Ingest   } Streaming Analytics

Analyze

Decide   } Transactions
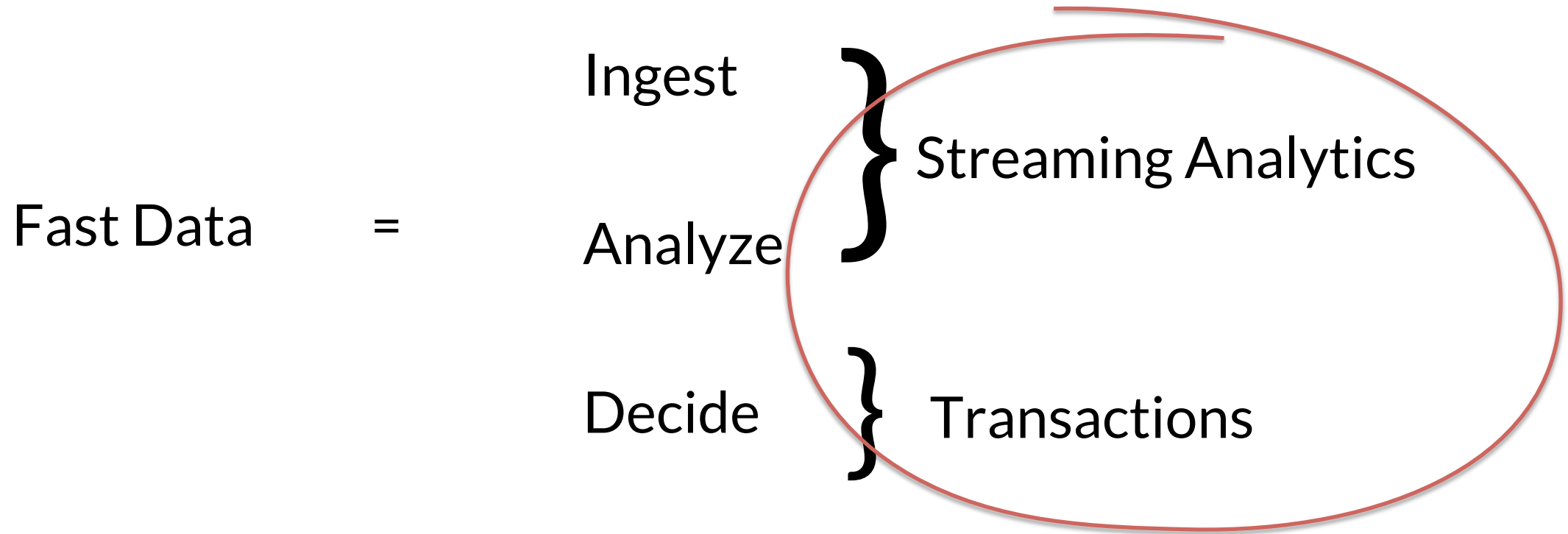
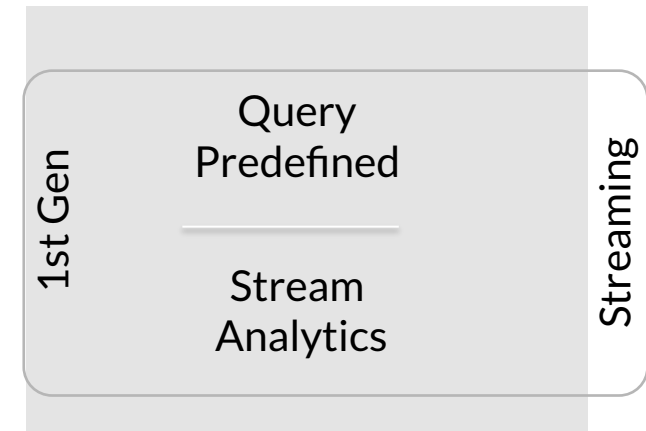# 1ST GENERATION FAST DATA: STREAMING ANALYTICS

- Examples:
  Spark Streaming, Storm, Kinesis, CEP (Streambase, Esper, etc…), et al

- Technical:
  - Lacks integrated "state" for transaction processing (operational)
 - Complex programming model
  - No ability to do ad hoc queries

- Functional:
  - 1st Gen. only offers streaming analytics on windows of events
  - Separate database required for any meaningful work
  - Proprietary interface is inconsistent with the rest of the data pipeline
  - Does not support applications requirement for interaction

1st Gen

Query
Predefined

Stream
Analytics

Streaming

2ND GENERATION FAST DATA:
COMBINE STREAMING ANALYSIS WITH THE OPERATIONAL STORE

- Operational workloads want to understand data as it changes
- Stream processing work needs state to enable richer analytics

- Operations and Stream Processing are set to converge
  - Convergence is necessary to use streaming analysis in real-time
  - Automated application interactions are informed by data
  - Brings the application into the "data analytics" world

- Streaming Analytics alone is *passive*, Fast Data is *interactive*

2nd Gen	Support Operational Work _____ Ad hoc	VoltDB
1st Gen	Query Predefined _____ Stream Analytics	Streaming
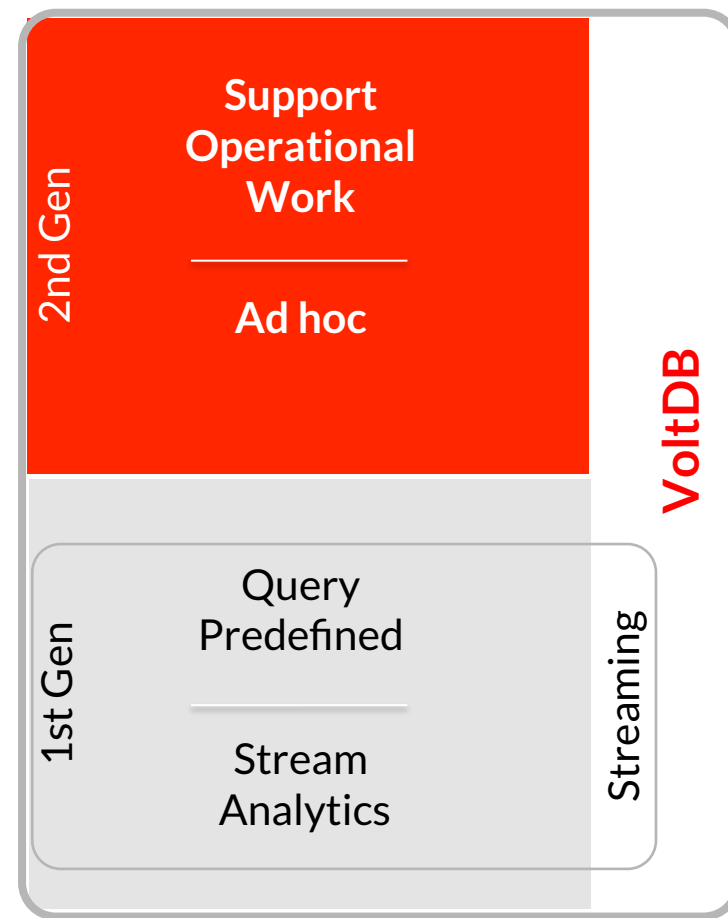
FAST DATA PIPELINE

Data in Motion

Data at Rest

Customer-Facing

- Personalization
- Customer experience

Streaming
Analytics
- Filtering
- Windowing
- Aggregation
- Enrichment
- Correlations

+

Operational
Interaction/
Transactions
- Context-aware
- Personal
- Real-time

Export

Batch/Iterative
Analytics
- Statistical correlations
- Multi-dimensional analysis
- Predictive analytics

Operations-Facing

- Network optimization
- API monitoring
- Sensors

VoltDB

Real-time Analytics by VoltDB on IBM SoftLayer Cloud

Agenda

- The IT Trends

- About VoltDB

- The Reference Architectures

- A Simple Demo

9

innovation that matters

The Top Disruptive Transformations

- **Digitization & Deeper Connectivity leads to massive amounts of Data**

- **Data Analytics technologies and tools enables data to information and to knowledge**

- **Knowledge-filled up Services and Applications for the Envisioned Smarter Planet**

innovation that matters

The Prominent Implications

- Trillions of Digitized Objects through digitization and edge technologies

- Billions of connected devices through deeper connectivity and service-enablement of physical, mechanical, electrical, electronics and IT systems

- Millions of software services (device-specific as well as agnostic)

- Zettabytes of multi-structured data

- Newer Analytical Capabilities – prognostic, predictive, prescriptive and personalized insights

- Cognitive Systems

The Prominent Challenges

- Hyper-converged Data Analytics Platforms

- Analytics-aware IT Infrastructures

innovation that matters

What to do with the bulging data?

- Indulge in analytics to extract insights

- Implement Insights-driven Applications

innovation that matters

Why Data Analytics is vital?

Data analytics is a process for squeezing out actionable insights out of data towards tactical and strategic decision-enablement for individuals as well as institutions

Insights

1. Enabling business Process Innovation
2. Exploring fresh avenues for fresh revenues
3. Elevating business efficiency and value
4. Enhancing business operations, retaining customer loyalty, attaining newer customers,
5. Expounding outside-in thinking in product designs

innovation that matters

VoltDB for Real-time Data Analytics

- VoltDB is an in-memory scale-out database built to power the next generation of streaming data applications.

- VoltDB's modern architecture provides fast data ingestion and export with massive scalability, real-time analytics, and data enrichment.

- VoltDB is a relational database system (RDBMS) for high-throughput, operational applications requiring:

  - Orders of magnitude better performance than a conventional DBMS

  - SQL as the native data language

  - ACID transactions to ensure data consistency and integrity

  - Built-in high availability (database fault tolerance)

  - Database replication (for disaster recovery)

innovation that matters

The Big Picture

innovation that matters
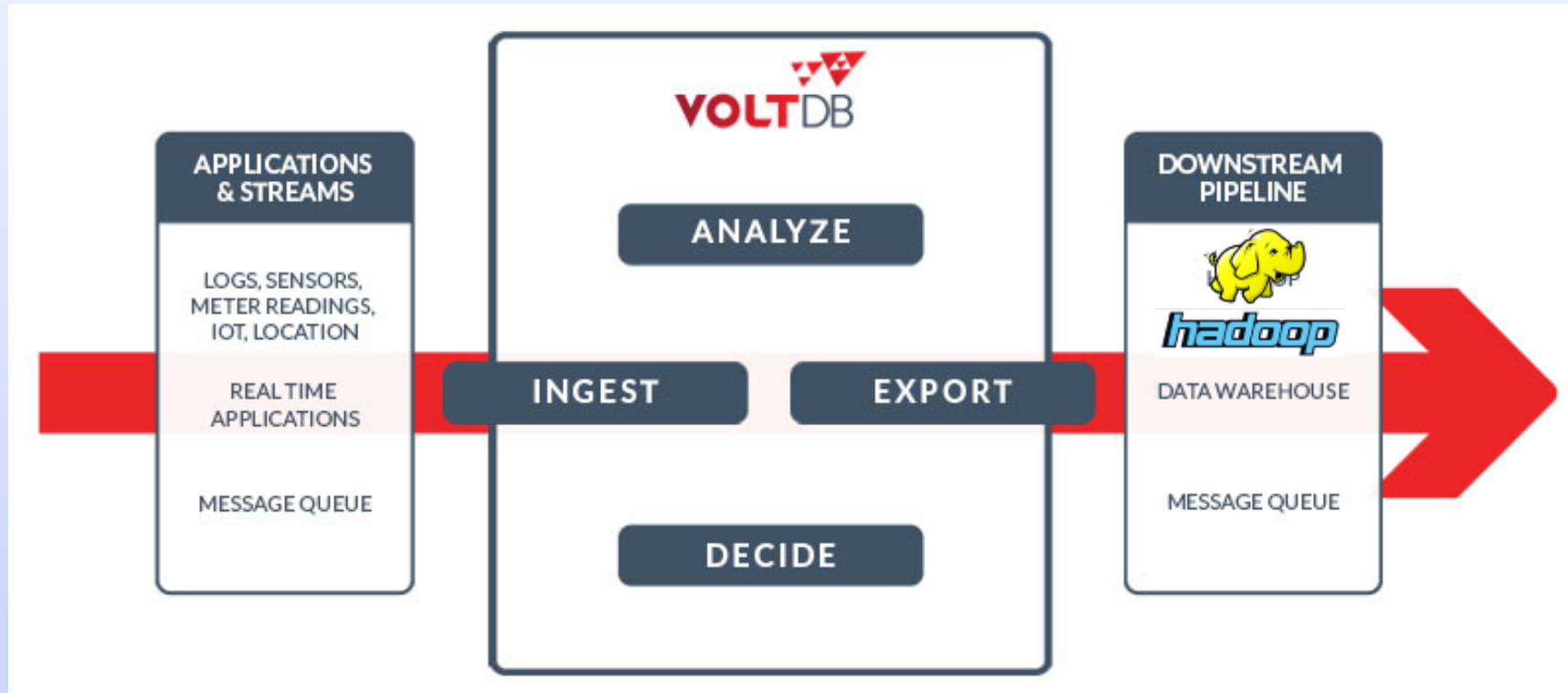
How VoltDB eliminates the RDBMS Problems for real-time analytics

- Data and associated processing are partitioned together and distributed across the CPU cores ("partitions") of a shared-nothing hardware cluster.

- Data is held in main memory, eliminating the need for buffer management.

- Transactions execute sequentially in memory, eliminating the need for locking and latching.

- Synchronous multi-master replication provides built-in high availability.

- Command logging provides a high performance replacement for "write-ahead" data logging.

innovation that matters

The Data Architecture for Fast and Big Data

innovation that matters
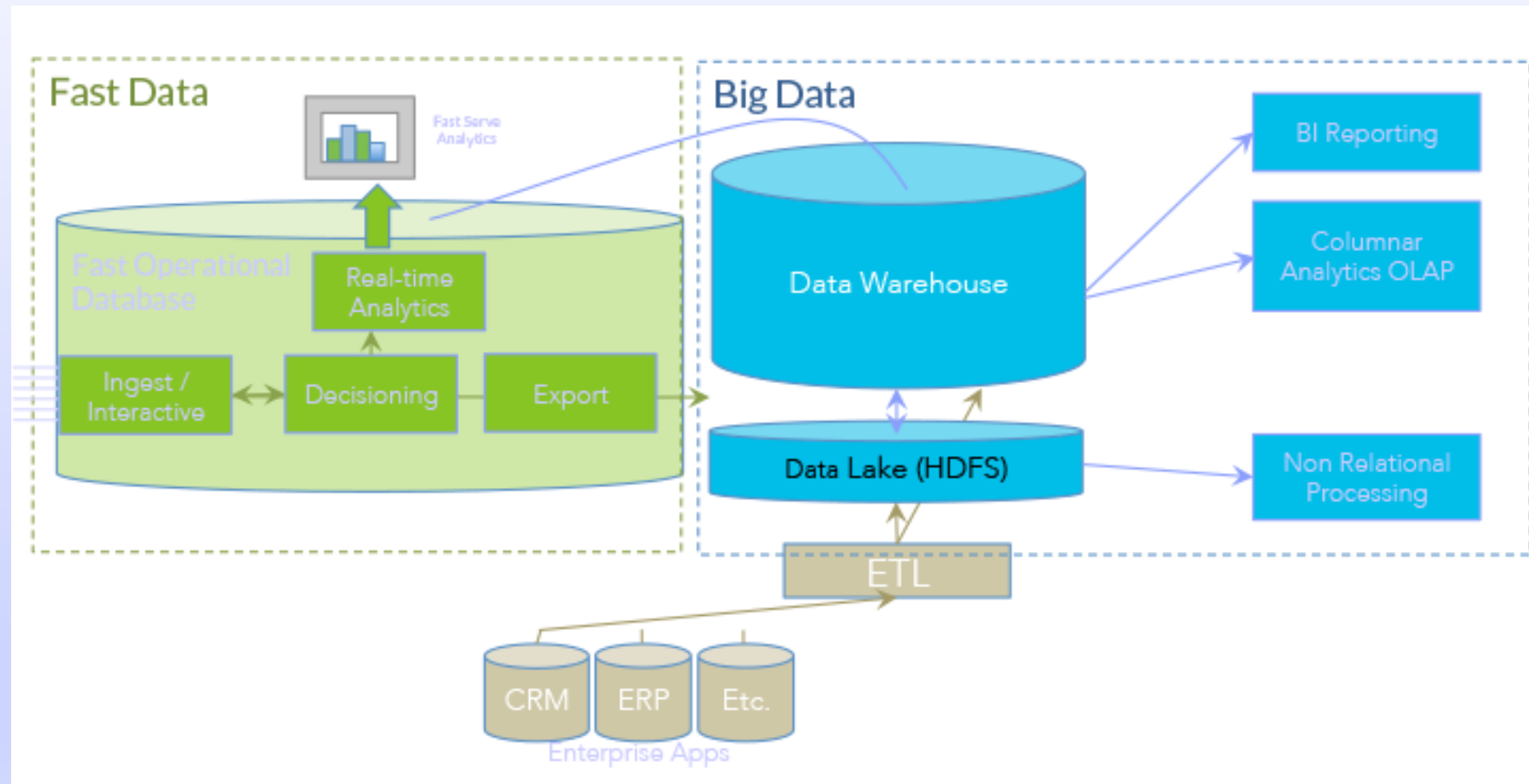
VoltDB on IBM SoftLayer Cloud

- Ideally suited to streaming applications in the cloud.

- In addition to providing extremely fast ingest, VoltDB exports data at high speed to long-term analytics stores.

innovation that matters

The VoltDB export function automatically queues data to the connector. The export client then uses a series of poll and acknowledgement requests to transactionally exchange data between VoltDB and Netezza systems, ensuring at least one delivery of data to the destination system.

innovation that matters

A Simple Demo on IBM SoftLayer Cloud

VoltDB installation

innovation that matters

Application main page

innovation that matters
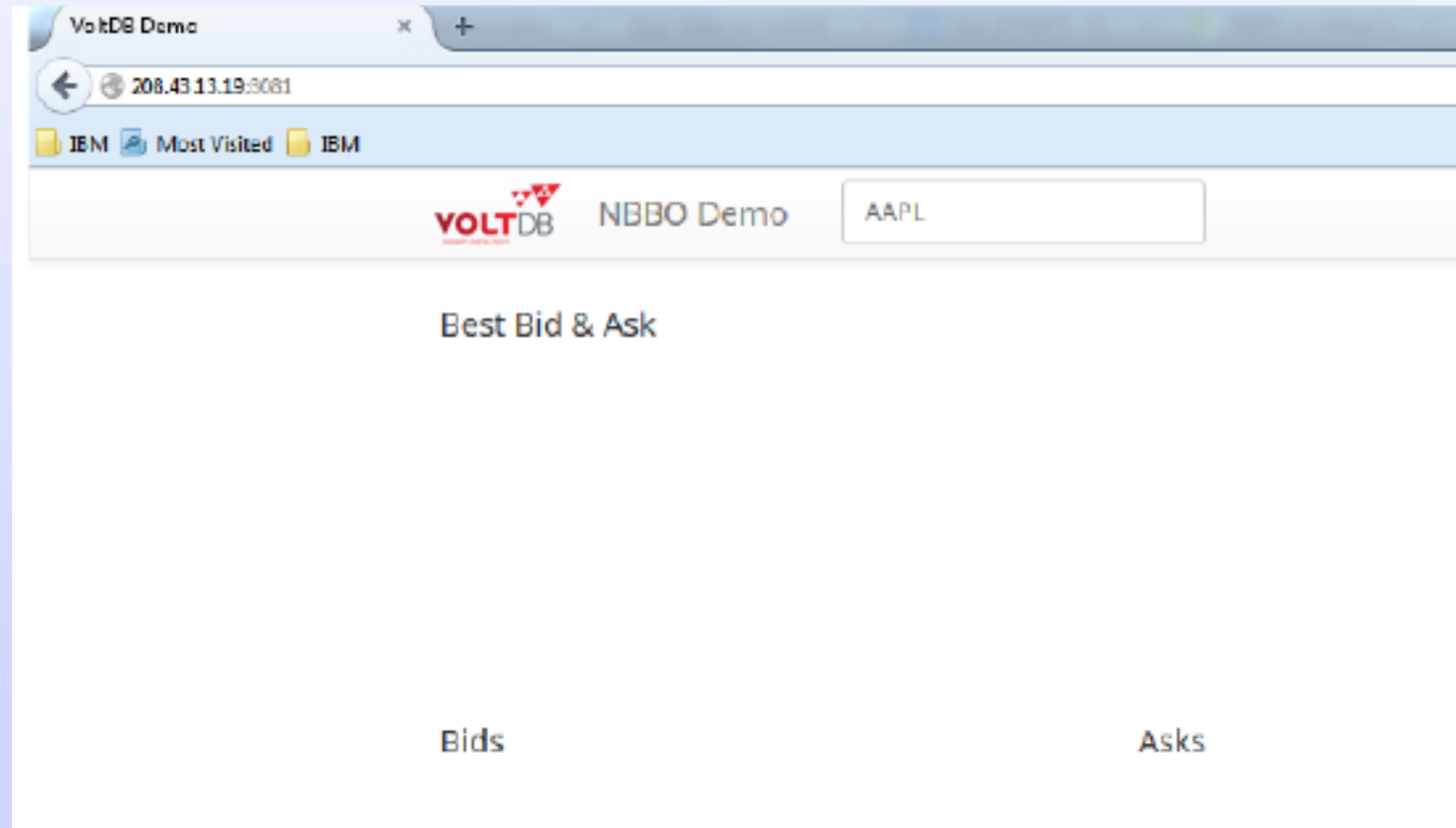
Connecting VoltDB Server

innovation that matters

IBM/MTA Confidential | Cloud POC

innovation that matters

IBM
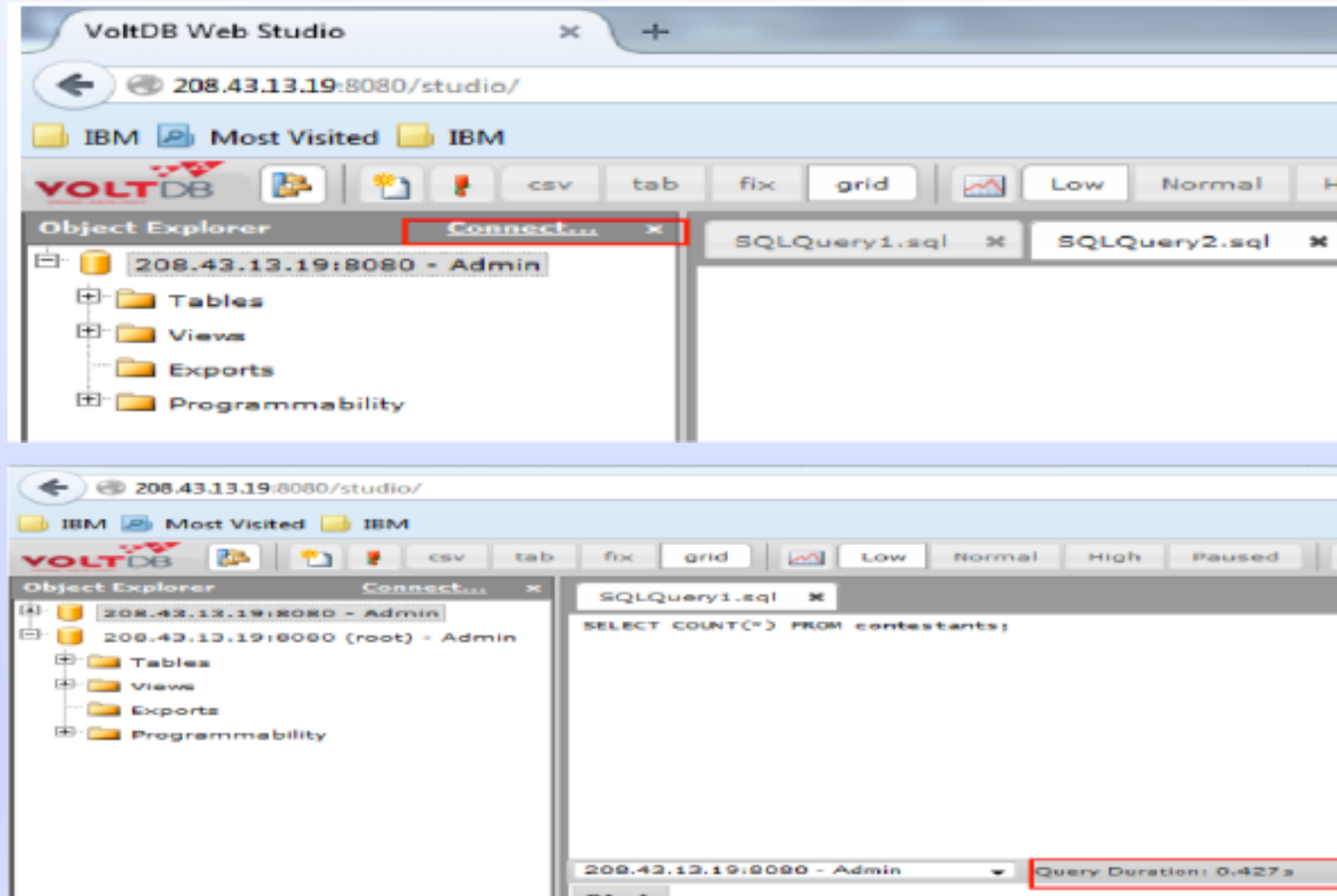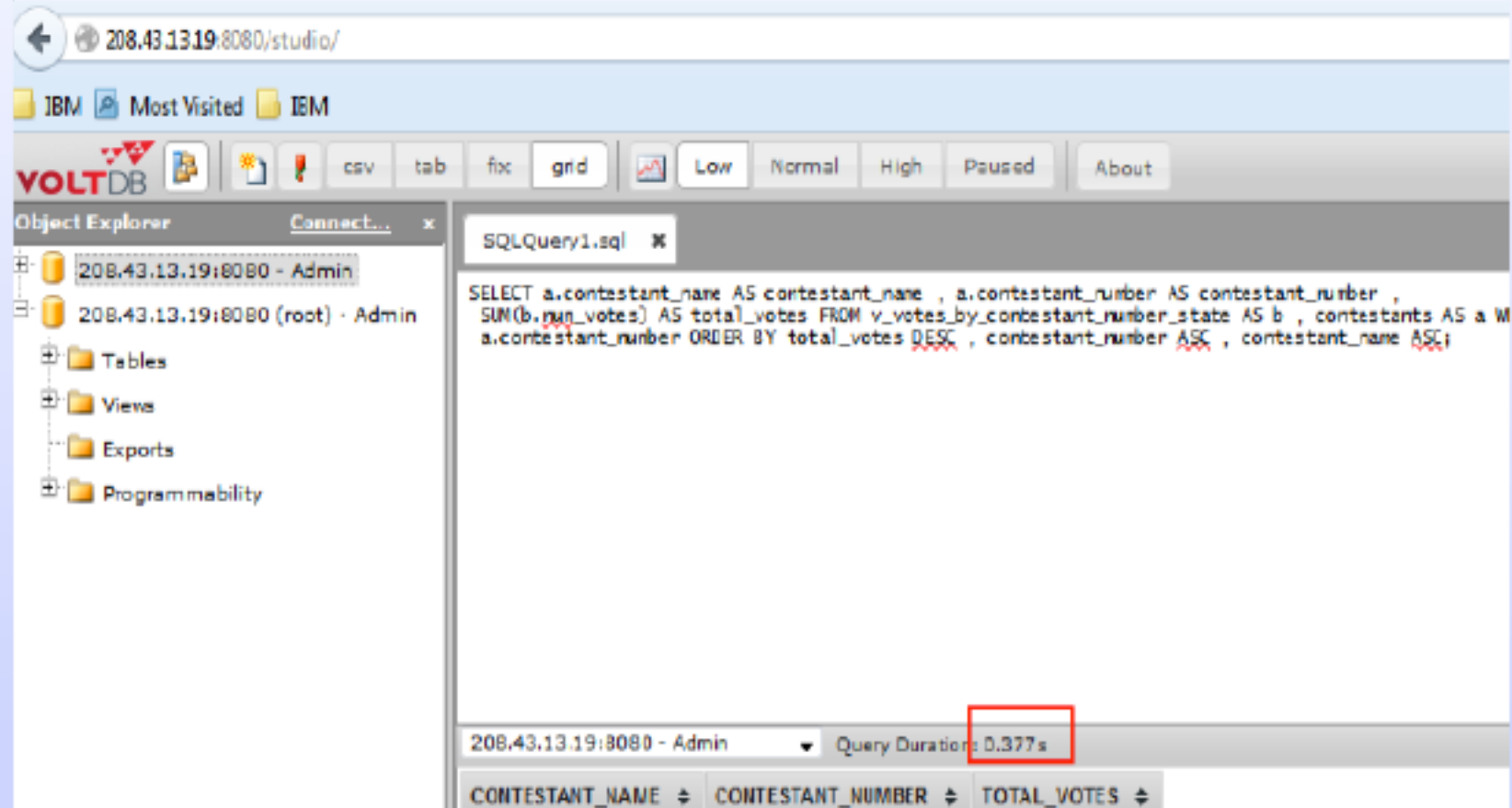
Thanks for listening

innovation that matters

*

QUESTIONS?

- Use the chat window to type in your questions
- Try VoltDB yourself:

  - ➢ Free trial of the Enterprise Edition:
    - www.voltdb.com/Download

  - ➢ Open source version is available on github.com

-
    Email us your question: askanengineer@voltdb.com