In-Memory Computing Patterns for High Volume, Real-Time Applications

Narendra Paruchuri  
American Airlines

Murali Ande  
American Airlines
Outline

• Who We are
• Our Use Cases
• Our Journey
• Evaluation Criteria
• Our Journey with In-Memory data models
• Architecture Patterns
• Advantages
• Being mindful
Who we are

Oneworld alliance
14,250 flights, 1,000 destinations, 150 countries

Complex aircraft turn activities before on time departures

500,000 daily customers
6,700 flights
350 destinations
50 countries

Multiple Hubs
CLT, ORD, DFW, LAX, MIA, JFK, LGA, PHL, PHX, DCA

Safety and Regulations

Complex weather situations, reroutes, off-scheduled operations, travel plan changes, last minute
Our Use Cases

- Customer connections and Service recovery
  - 12 Hubs.
  - Connections – delays, tight and missed connections.
  - Rebooking.
  - Expedite Immigration.
  - Taking care of AAdvantage customers and their bags.

- Aircraft Turn Management
  - Aircraft turn events
  - Customer boarding
  - Bags loading
  - Crew check-ins.
  - Customer check-in
  - Fueling
  - Cabin Cleaning
  - Pro actively finding flight delays
  - Alerts and notifications
While systems are modified for key value pair keeping in view the trends in latest technologies, our business operations still require us to perform joins to correlate and coalesce data to facilitate business decisions.
Evaluation Criteria

- ACID compliance with native persistence and third party persistence integration
- Readability and Maintainability of System
- Support High Volume Transactions with simultaneous updates to multiple attributes (12K / Min – Writes, 25K / Min – Reads)
- Replication and Distribution support with Multi – DC
- Streaming connectors like Kafka
- Generate events on data updates
- Docker and Kubernetes support
Our Journey with In-Memory Data Models

Ehcache
Infinispan

Key Value Model

Pain Points:
- Most of the solutions does not offer Joins
- Lot of code to correlate and Coalesce data
- Slow response times.

Ehcache
Infinispan

Key Value Model

Pain Points:
- Pre-defined Object Structures
- Reduced Flexibility
- Longer Deployment Cycles

Object Model

Gigaspaces
Our Journey with In-Memory Data Models

Ehcache
Infinispan

Key Value Model

Gigaspaces

Neo4J

Pain Points:
- Computations and Aggregations

Object Model

Pain Points:
- Pre-defined Object Structures
- Reduced Flexibility
- Longer Deployment Cycles

Node1
Node2
Node3
Our Journey with In-Memory Data Models

- **Ehcache**
  - Infinispan
  - **Key Value Model**

- **Neo4J**
  - Pain Points:
    - Computations and Aggregations

- **Gigaspaces**
  - **Object Model**

- **Cassandra**
  - **Columnar like Model**
  - Pain Points
    - Atomicity – multi tables
    - Consistency
Our Journey with In-Memory Data Models

- Ehcache
- Infinispan

Key Value Model

- Neo4J

Graph

- VoltDB
- Apache Ignite

Relational Model

- Cassandra

Columnar like Model

Object Model

- Gigaspaces

Pain Points
- Atomicity – multi tables
- Consistency

Our business operations still require us to perform joins to correlate and coalesce data for the business.
While there may be several options, we have chosen to implement Ignite for our use cases.
Apache Ignite (Open Source) – Active – Passive topology

- Use for Non-Critical systems that need real-time data aggregation.
Apache Ignite (Open Source) – Active – Passive topology

- Topology aware Feeder Client Node(S)
- Client Listener Node(S)
- Web Application
- Global Traffic Manager
- Websocket
- Real-time Application Data Store ACID compliance with Strong Consistency
- Native Persistence
- SAN Storage
- WAN Replication
- SAN Storage
- Active Data Center A
- Passive Data Center B
- • Use for Non-Critical systems that need real-time data aggregation.
Apache Ignite and Cassandra Active-Passive topology

- **Global Traffic Manager**
  - **Active Data Center A**
  - **Passive Data Center B**
  - **Cassandra Cluster**
  - **WAN Replication**

- **Client Listener Node(S)**
  - **Websocket**
  - **Web Application**

- **Topology aware Feeder Client Node(S)**

- **Real-time Application Data Store**
  - **ACID compliance with Strong Consistency.**
  - **Cassandra Connector**

- **Cassandra Cluster**

- **Combines OLTP & OLAP**

- **Use for Non-Critical systems that need real-time data aggregation.**
Advantages of using Ignite along with Cassandra

- Bandwidth and Response times improvement (order of magnitude improvement).
- Improved Availability than standalone Cassandra system, as Ignite offers sophisticated clustering support, such as detecting and remediating split brain conditions.
- Horizontal and Vertically scalable.
- More efficient, as Ignite can use all the memory available on a node, and not only JVM memory.
- ANSI-99 SQL and ACID Transaction Guarantees (Improved Consistency)
- You can run other analytics off of Cassandra
- Support for jdbc and odbc make it easier to integrate with existing tech, such as Hibernate and Spring Data.
- No data remodeling required for existing Cassandra deployment, as Apache Ignite can read from it as well as relational databases.
Apache Ignite (GridGain) – Active – Active topology

- Use for Critical and Vital systems with low RTO and RPO.

Real-time Application Data Store ACID compliance with Strong Consistency

Global Traffic Manager

- WAN Replication

Client Listener Node(S)

Topology aware Feeder Client Node(S)

Websocket

Web Application

SAN Storage

Native Persistence

Active Data Center A

Active Data Center B
Being mindful

- Not a magic bullet
- Evaluate use case
- Consider future data growth
- Take advantage of all available technologies
- Lack of standards
- Not plug and play
- Vendor challenges
- Governance
Why not traditional RDBMS database

- Database will work but at what cost?
- Complexity of Architecture with Multi DC and Multi Cloud providers.
- Tool consistency
- Scalability Vertical vs Horizontal
- No collocated data processing
- Leveraging auto scaling using Containers and Kubernetes is not available.

Cloud Scaling Breaks Domain Driven Design
Advantages

• Extremely fast response times
• Highly scalable
• Improved complex processing on events.
• Improved real time analytics
• Reduced cost of operation
• Inexpensive with improvements in memory technology
QUESTIONS?
In-Memory Computing Patterns for High Volume, Real-Time Applications

Narendra Paruchuri
American Airlines

Murali Ande
American Airlines