Performance Testing of NoSQL Applications
Key Big Data Technologies

• Map Reduce
  • Apache Hadoop, Cloudera, Hortonworks, MapR etc.

• NO SQL
  • Cassandra, Mongo DB, Oracle NoSQL, Neo4j etc.

• Messaging queues
  • Kafka, ActiveMQ, RabbitMQ, ZeroMQ etc.

• Search
  • Lucene, Elastic Search, Solr
Agenda

• Introduction to NoSQL
• Performance Testing Challenges
• Performance Testing Approach
• Performance Testing Solutions
Introduction to NoSQL

• Large data volumes
• Complex unstructured data
• Deeper insights
• Storing images, videos
NoSQL Architecture

- NO SQL
  - Cassandra, Mongo, Neo4J etc.
- Fault Tolerant
- Available
- Scalable
- Elastic
- Schema less
Big Data Performance Test Focus Areas

- Data Ingestion
- Data Processing
- Data Persistence
- Reporting & Analytics
Performance Testing Challenges

- Diverse technologies
- Unavailability of tools
- Test scripting
- Test environment
- Limited monitoring solutions
- Lack of diagnostic solutions
Performance Testing Approach

1. Setup Test Environment
2. Identify Workload Characteristics
3. Prepare Test Clients
4. Performance & Benchmark Tests
5. Analyze Results & Tune components
6. Optimum Configuration
7. Failover & Reliability Tests
Performance Testing Solutions

• Performance Test Tools
  • YCSB (Yahoo Cloud Serving Benchmark), SandStorm, JMeter

• Monitoring Tools
  • Nagios, Zabbix, Ganglia, JMX utilities

• Diagnostic Tools (APM)
  • visualVM, AppDynamics, Compuware
YCSB

• Testing client for NoSQL databases

• Performance database operations as per specified workloads

./bin/ycsb load cassandra-10 -P workloads/workloada -P cassandra.props -threads 50 -s > loaddata-cassandra.results
SandStorm

• Clients for NoSQL and message queues
• Design realistic test scenarios
• Monitoring of test environment
Critical Performance Parameters

- Data Storage
- Commit Logs
- Concurrency
- Caching
- JVM parameters
Q&A
Thank You

For more info mbatterywala@impetus.co.in