

Case Study



Adopting Performance-driven Development* to build a near real-time eXtreme transaction processing (XTP) and analytics platform

Abstract

An Information Service firm partnered Infosys to build an almost real-time market information storage and analytics platform to devise trading strategies. Based on live data, this platform had extremely ambitious performance requirements. It was required to support 4000 concurrent users, 5 GB of input data/minute, 6 million transactions/minute and an information retrieval latency of less than 2 seconds. Infosys' performance-driven development approach combined with the powerful Grid technology helped the client's platform not only successfully achieve but go beyond the stringent performance targets.

* Performance-driven development is a methodology used by Infosys to ensure a proactive and holistic focus on Performance & Scalability throughout the Software Development Life Cycle (SDLC).

Background

The client provides news, information and related products to professionals working in financial service firms, media entities, corporate markets and consumers around the world. The offered products and services help these professionals to arrive at sound business decisions and improve productivity.

The client proposed to build an innovative, cutting edge technology platform to serve as a data source for devising trading strategies, computing analytics and formulating algorithms for market analysis. Besides these criteria, they also wanted the platform to fulfill regulatory requirements. The objective of the platform was to provide ready-to-consume information to their consumers. Its other necessity was to possess the capability to capture, store and query market data from as far as 15 years back to almost real-time.

The client faced a number of hurdles in building the required platform, the most significant of these being:

1. Defining a scalable architecture to satisfy the platform's ambitious requirement. The stringent Service Level Agreements (SLA) challenged the technical limits of many of the technologies used to create such a platform
2. Designing the platform to achieve extremely low latency in information retrieval (response times less than 2 sec) and high throughput
3. The platform was required to support a very large number of simultaneous users growing exponentially year after year at 100%, while ensuring high availability of 99.999%
4. Handling huge data volumes which were created by the requirement of making available online up to 15 years worth of data, with an expected growth rate of 60% per year
5. The expected 100% growth in the total user base year after year
6. Catering to performance requirements that were in excess of industry standards then:
 - a. A peak transaction processing requirement of 6 million transactions/minute as opposed to the then TPC-C record of 4 million transactions/minute
 - b. A scalability requirement of 60% growth every year that exceeded Moore's law

Infosys' Solution

The client collaborated with Infosys to build a platform to address these extremely challenging requirements. As part of the engagement, Infosys used a proactive and holistic performance-driven development approach to amplify the performance and scalability of the application. The approach involved an agile software development model, the key activities of which included:

- Validating non-functional requirements and conducting a feasibility analysis
- Characterizing and modeling the business workload
- Defining the solution architecture and validating performance and scalability via a proof-of-concept
- Arriving at performance baselines for various architecture and technology components individually, as well as a composite application, as part of the build cycle
- Capacity planning
- Modeling the system via performance benchmarking for scalability analysis
- Analyzing and optimizing code response time as a part of daily unit testing cycles
- Testing and tuning system performance



- A data acquisition layer built using bespoke Java code to assimilate data from multiple data sources
- An application grid for data processing and analytics computations using an enterprise distributed caching product — named GigaSpaces — employed for higher throughput and response times
- A data storage grid built on Oracle 10g Real Application Clusters for catering to the heavy volume and mixed workload for read and write functions
- Exploiting the finer capabilities of Oracle 10g - Grid Console for grid nodes administration, performance monitoring, load balancing and nodes accounting
- Application presentation and data distribution framework built using Oracle's streaming web services integrated on top of Oracle 10g J2EE Application Server Cluster

Benefits to the client

- The client successfully built a platform catering to the near real-time information processing, persistence and distribution requirements of its customers
- The platform achieved and exceeded every performance target specified in the requirements
- The cost to the client was optimized by:
 - Utilizing low cost Sun T2000 Nodes
 - Bespoke Java code employed instead of a licensed ETL product for data acquisition
 - Utilizing grid and virtualization capabilities



- Achieved 4.2 million transactions/minute with a peak of 6 million transactions/minute
- Handled high data volumes – 5 GB of data/minute
- Devised an information lifecycle management strategy for 30 TB of data
- Achieved information retrieval response times of 500 milliseconds with 4000 concurrent users against the required target of 2 seconds



- Homogeneous Grid comprising of SUN T2000 Server Nodes (1 CPU with 8 cores and 32 GB RAM) powered by SUN N1 Grid Engine
- SUN StorEdge 9990 Storage Array
- Application Grid using multithreaded Java code
- Oracle 10g RAC Grid
- Oracle 10g J2EE Application servers
- GigaSpaces Enterprise Distributed Cache as the Data Grid Solution



For more information, contact askus@infosys.com

About Infosys

Many of the world's most successful organizations rely on Infosys to deliver measurable business value. Infosys provides business consulting, technology, engineering and outsourcing services to help clients in over 30 countries build tomorrow's enterprise.

For more information about Infosys (NASDAQ:INFY), visit www.infosys.com.