Paper 262-2009

High Performance Grid and Cluster Computing Industry Trends

Dana Rafiee

ABSTRACT/INTRODUCTION

As computing power requirements and the amount of data increase, organizations continue to be faced with the challenge of satisfying business needs on the data within business usable timeframes. In the past, there was only one choice, process on a server. If more processing was needed, it was time to get a larger server. The same was true for databases. They typically use a traditional storage mechanism. The primary answer to each traditional method is to buy bigger, more expensive server hardware and disk storage. Imagine the potential for a redesign that offers distributed computing capacity and distributed storage. This presentation will discuss these industry trends along with innovative designs created by Destiny Corporation that address the combination of issues, applied to various client sites.

BUSINESS PROBLEM

Case Study: Washington University, St. Louis, MO

This organization was challenged with long running times of jobs in their current environment. They were running applications that executed for up to two weeks each. They were performing Genetic Simulations (Monte Carlo) and statistical bootstrapping. The goal was to reduce overall execution time from two weeks to a few hours. The new design included the following:

- Platform Computing Software
- SAS® Grid Computing Software
- Enterprise Miner for Grid Computing
- 90 Dual Core Blades
- 32 bit Red Hat Linux
- 16 GB RAM/Blade
- Shared SAN Storage
- 790+ Node environment
- LDAP Security

Case Study: Deloitte, Global Audit

It was extremely challenging for Deloitte IT to keep up with the success of their global audit practice. In addition, the Sarbanes-Oxley act requires audits of publicly traded firms. As with all auditing firms, Microsoft Excel™ and SAS act as primary tools for many auditors. However, when the complexity and amount of data outgrow the capability of the tool and the local environment, a new design must be considered. The requirement was to be able to maintain a parameter driven Excel based front end, while being able to process the back end on computing power and storage with complete flexibility and minimal limitations.

The challenge was to load balance the server usage with no user intervention and tie back the results to the Excel Desktop without changing the user's interface.

The design employed the following

- Platform Computing Software

- SAS Grid Computing Software

- SAS/ITRM Software

- (30) Dual Core Windows 2003 Servers

- (16) 64 bit AIX partitions

- 16 GB RAM per partition

- 4 processors/partition

- Shared SAN Storage

- LDAP secured

- Above replicated across Dev/Test/QA/Prod/BCP

The results for nine months of Performance Statistics are as follows:

- Many thousands of jobs executed

- No failures

- No downtime

- Scalability achieved allowing Deloitte to increase and effectively serve its customer base.

Case Study: Fortune 1000 Global Marketing Firm

This organization was challenged by execution on legacy environments where jobs would execute for over 11 hours on a single, large capacity mainframe computer. The goal was to reduce execution time and migrate to an environment that would be more cost effective. The designed employed In-Database Processing and Advanced Grid capabilities. These new capabilities allow for movement of processing into the grid based database environment that automatically distributes storage and the data along with offering local processing. Some of the capabilities include the in database ability for

- C++ Code

- Java Code

- Where Clause subsetting

- Join Redistributions

- Sorting

- Scoring

- More

The results for processing today are as follows:

- Production jobs processing time has been reduced from 11 hours to 11 minutes.

During the live presentation, we will show and discuss several architectural designs, including the ones discussed in this paper. For more information, please contact the author.

CONCLUSION

In conclusion, Destiny Corporation has tracked trends in Business and Information Technology since 1987. We have heard our customer's issues from both a Business and an IT perspective. Living and working with large data has been a growing challenge for our customers as data acquisition continues yielding masses of stored data available to the organization for analysis. The storage and analysis mechanisms must adapt. Our analysis has shown that High Performance Computing is an excellent answer to the challenge of today's large data warehouse business needs.

CONTACT INFORMATION

Dana Rafiee
Destiny Corporation
2075 Silas Deane Highway
Rocky Hill, CT 06067 USA
Phone: 860-721-1684
Fax: 860-721-9784
E-mail: drafiee@destinycorp.com
Web: www.destinycorp.com

All brand and product names are trademarks of their respective companies.