FOSS Large Data Storage Solution

Gary Jaffe, Priscilla Ramirez, Nurdeen Salami, Rachel Saltzman, Jacob Wilson
Mentors: Michael Ritchson, Shaun Smith

AERO Institute
Palmdale, CA

National Aeronautics and Space Administration,
Armstrong Flight Research Center,
Edwards ,CA

AERO Test Bed Team (ATBT)
"Sandbox"

Introduction

NASA projects require a reliable approach to store large volumes of data. Accordingly, it is crucial to adopt a secure and accessible database that enables fast, reliable, and flexible use. The current database implemented at NASA bears costly license fees with undesirable speed and flexibility. The purpose of the "Sand Box" or the "AERO Test Bed Team" (ATBT) Development team project is to test a new database, Cassandra, for performance capabilities over traditional databases.

System Requirements

- Store vast amounts of Compact Fiber Optic Sensing System (C-FOSS) data in multiple file formats (.csv, .log, .xml, and .jpg)
- Create an intuitive front-end user interface that allows quick searching and viewing of C-FOSS data
- Use benchmarking tests to verify the speed, flexibility, and reliability of data stored in the Cassandra database

Methodology

While reading FOSS .csv file, create and upload data to Database

Invoke a Web Service to handle Cassandra FOSS specific queries from client

Query data from Cassandra Database and return in XML format to client

Parse XML data on client

Display modified information through the AERO Portal

Software Architecture

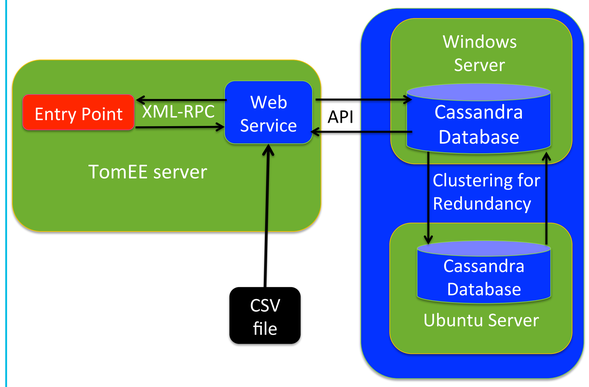


*Diagram of loosely coupled CASSANDRA architecture

Agile Development Team

Business Analyst: Priscilla Ramirez

- Gather ATBT System requirements
- Maintain software compliance and documentation per DPR 7150
- Act as liaison between customer and software team

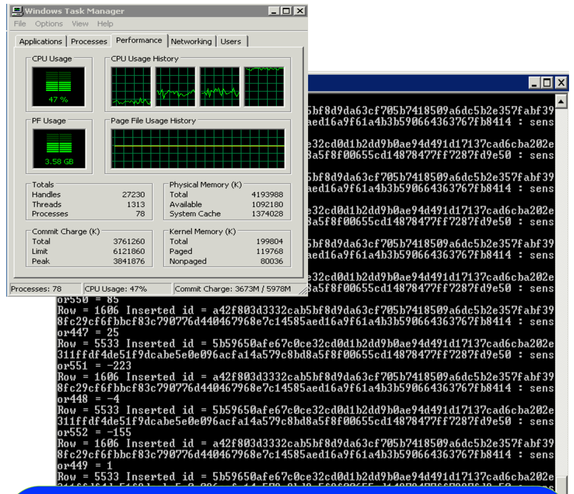
Tester: Jacob Wilson

- Generate Requirements Traceability Matrix (RTM) and test plan
- Explore test automation tools
- Validate data integrity

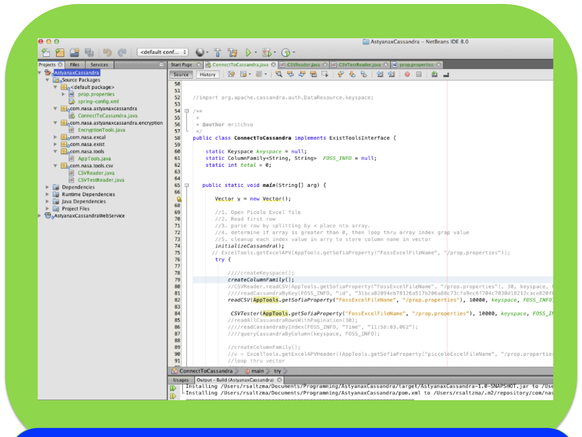
Developer: Gary Jaffe & Rachel Saltzman

- Test Cassandra; architecture, APIs, and multi-cluster system functionality
- Develop Web applications and services in an integrated development environment
- Generate benchmark system tests and reports
- **System Administrator:** Nurdeen Salami
- Build server test rack
- Identify and install server software
- Identify and install benchmark system test and reporting applications

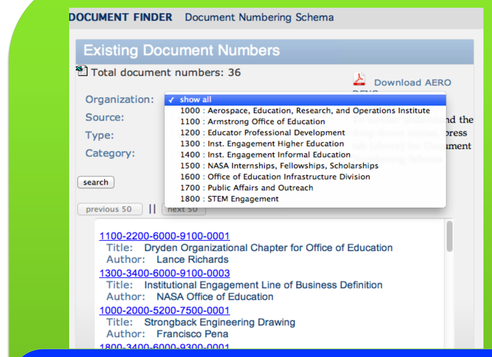
Analysis



Cassandra demonstrated its ability to handle large data files recorded by the FOSS team. The image above illustrates the input of a file containing 14,000 rows and 575 columns of data.



This is sample code that reads, validates, and uploads files to the Cassandra database.



The finished product will resemble the existing method for querying text documents seen above on the AERO Portal.

CONCLUSION

The application of the CASSANDRA database allows for low footprint data storage and retrieval. Validation tests demonstrate its ability as a reliable data storing method. CASSANDRA is an acceptable data storage solution, though an Application Program Interface API may still need to be implemented to finalize the project.

ACKNOWLEDGMENTS

- Michael Ritchson, NASA AFRC
  Software Architect
  Michael.g.ritchson@nasa.gov
- Shaun Smith, AERO Institute
  Mission Support Specialist
  Shaun.smith@nasa.gov