

# Performance Analysis and Optimization of AMGA for the WISDOM environment

Wednesday, February 13, 2008 11:35 AM (20 minutes)

In the WISDOM environment, thousands of job agents distributed on the Grid may have access to an AMGA server simultaneously (1) to take docking tasks out of the AMGA server to execute on the machine that they are sitting, (2) to get the related ligand and target information, and (3) to store the docking results. The docking tasks take about 10 to 30 minutes to finish depending on the machine that they run and the docking configuration.

We have carried out some performance analysis on the current AMGA implementation. Due to the overhead required to handle GSI/SSL connection on the Grid, it showed about 350% poorer throughput compared with a direct DB access. In the current version of WISDOM, AMGA is used as a placeholder for a task distribution table where docking tasks are stored and maintained. We have found a serious performance degrade due to the overhead caused by the need to lock the whole table to prevent different agents from taking the same task.

## 3. Impact

First, in order to address the SSL/GSI-related performance issue, we have proposed a load-balanced multiple server and a DB connection pool technique in AMGA. Our preliminary test results demonstrate a linear performance improvement in proportion to the number of AMGA servers.

Secondly, to deal with the performance degrading problem associated with the locking of the whole table, we modified the AMGA source code and added a new API that allows the two separate AMGA APIs, SELECT and UPDATE needed to take a task, to be invoked at once. Our preliminary tests show that the new API allows about 50 tasks to be retrieved per second in contrast with one task per second being retrieved using the two separate SELECT and UPDATE API calls.

## 4. Conclusions / Future plans

We addressed performance issues on the use of AMGA in the WISDOM environment and presented some new techniques to drastically improve the performance of AMGA. The techniques are expected to be integrated in the new release of WISDOM environment, being deployed in the EGEE biomed VO infrastructure for the next WISDOM data challenge.

## Provide a set of generic keywords that define your contribution (e.g. Data Management, Workflows, High Energy Physics)

WISDOM, AMGA, metadata catalog, performance measurement

## 1. Short overview

AMGA is a gLite-metadata catalogue service designed to offer access to metadata for files stored on the Grid. We evaluated AMGA to analyze whether it is suitable for the WISDOM environment, where thousands of jobs access it simultaneously to get metadata describing docking results and the status of jobs. In this work, we address performance issues on AMGA and propose new techniques to improve AMGA performance in the WISDOM environment.

**Primary author:** Mr AHN, Sunil (KISTI)

**Co-authors:** Dr KOBLITZ, Birger (CERN); Mr KIM, Namgyu (KISTI); Mr LEE, Seehoon (KISTI); Dr HWANG, Soonwook (KISTI); Dr BRETON, Vincent (CNRS)

**Presenter:** Mr AHN, Sunil (KISTI)

**Session Classification:** Life Sciences

**Track Classification:** Existing or Prospective Grid Services