The AMGA Metadata Service in gLite

B. Koblitz, CERN-IT

with N. Santos

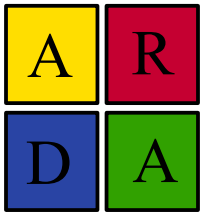
EGEE User Forum, CERN, Switzerland

March 2nd 2006

Overview

- What is AMGA? What is Metadata on the Grid?
- Interface definition
- Implementation of AMGA
- Performance Comparison with File Catalogues
- Replication of Metadata with AMGA

What is AMGA, What Metadata?



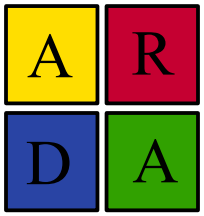
- **AMGA (ARDA Metadata Grid Application) is the Metadata catalogue of gLite 1.5**
- **AMGA started out as ARDA's tool to investigate metadata access on the GRID**
- **Metadata is relationally structured data for grid jobs (lives normally in databases)**
- **AMGA works in 2 modes:**
  - Side-by-Side a File Catalogue (LFC): File Metadata
  - Standalone: General relational data on Grid
- **AMGA has 2 front ends:**
  - SOAP with EGEE standardised interface
  - Text-based TCP streaming protocol (proprietary, documented)
- **http://project-arda-dev.web.cern.ch/project-arda-dev/metadata/**

A Common Interface

- **AMGA implements a common interface designed in close collaboration of gLite-DM and ARDA teams**

  (P. Kunszt, R. Rocha, N. Santos, B.K.)

  contains man ideas from UK GridPP Metadata group, LHCb (Bookkeeping, GANGA), HEP...
- **EGEE standard for metadata access (AMI, FiReMan)**
- **Design Ideas:**
  - Versatile: Performance (HEP), Security (BioMed)
  - Provides simple relational data manipulation interface
  - Modular: Interface for Entry manipulation, schemes, security
  - Few requirements on back end, can be SQL-DB, XML...
  - Allows stateless & statefull implementati
- **Description of WSDL at https://edms.cern.ch/document/573725**



Basic Concepts

- **Schema (directory)**
  - Has hierarchical name and list of attributes `/prod/events`
- **Attributes**
  - Have name and storage type
  - Interface handles all types as strings
- **Entry**
  - Live in a schema, assign values to attributes
- **Query**
  - SELECT ... WHERE ... clause in SQL-like query language

Examples

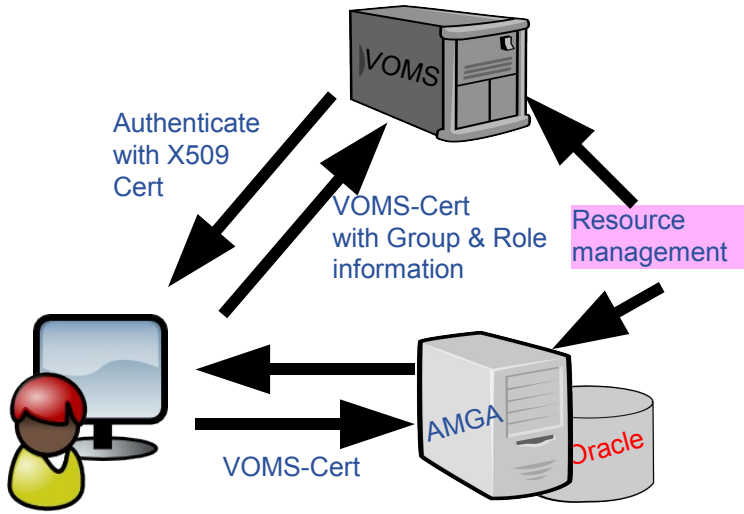
```
createdir /jobs
addattr /jobs jobStatus int
addentry /jobs/job1 jobStatus 0
updateattr /jobs jobStatus 1 jobID>100
selectattr /DLibrary:FileName /DLAudio:Author /DLAudio:Album
        '/DLibrary:FILE=/DLAudio:FILE and like(/DLibrary:FileName, “%.mp3")‘
```

Security Concepts

- **Security very important for BioMed, not for HEP**

Security ↔ Speed

- **AMGA supports:**
  - SSL connections (Optional!)
  - Authentication based on Password, X509 Cert, Grid Proxy
  - Posix-ACLs and Unix permissions for entries and collections
  - Built-in group-management like AFS or via VOMS

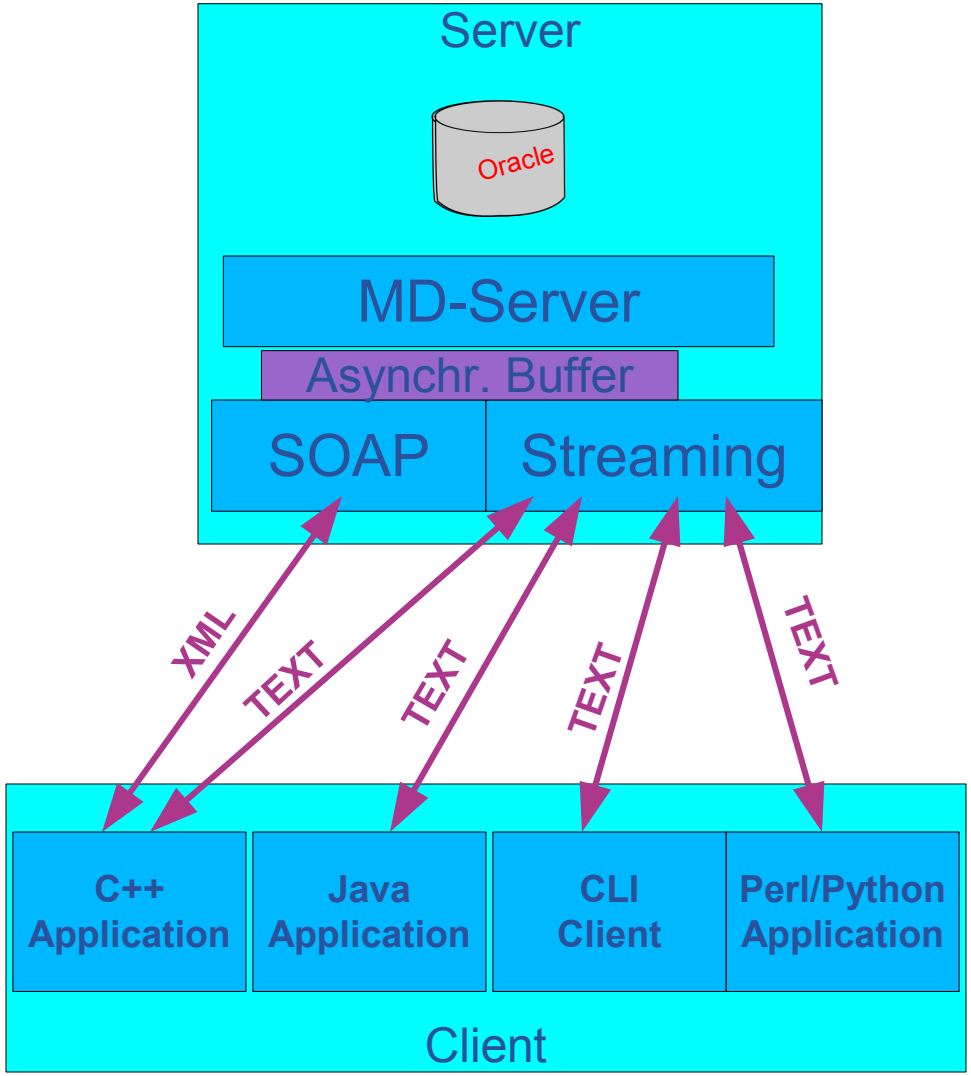


- **Currently no security on attribute basis**
  - AMGA allows to create views: Safer, faster, similar to RDBMS

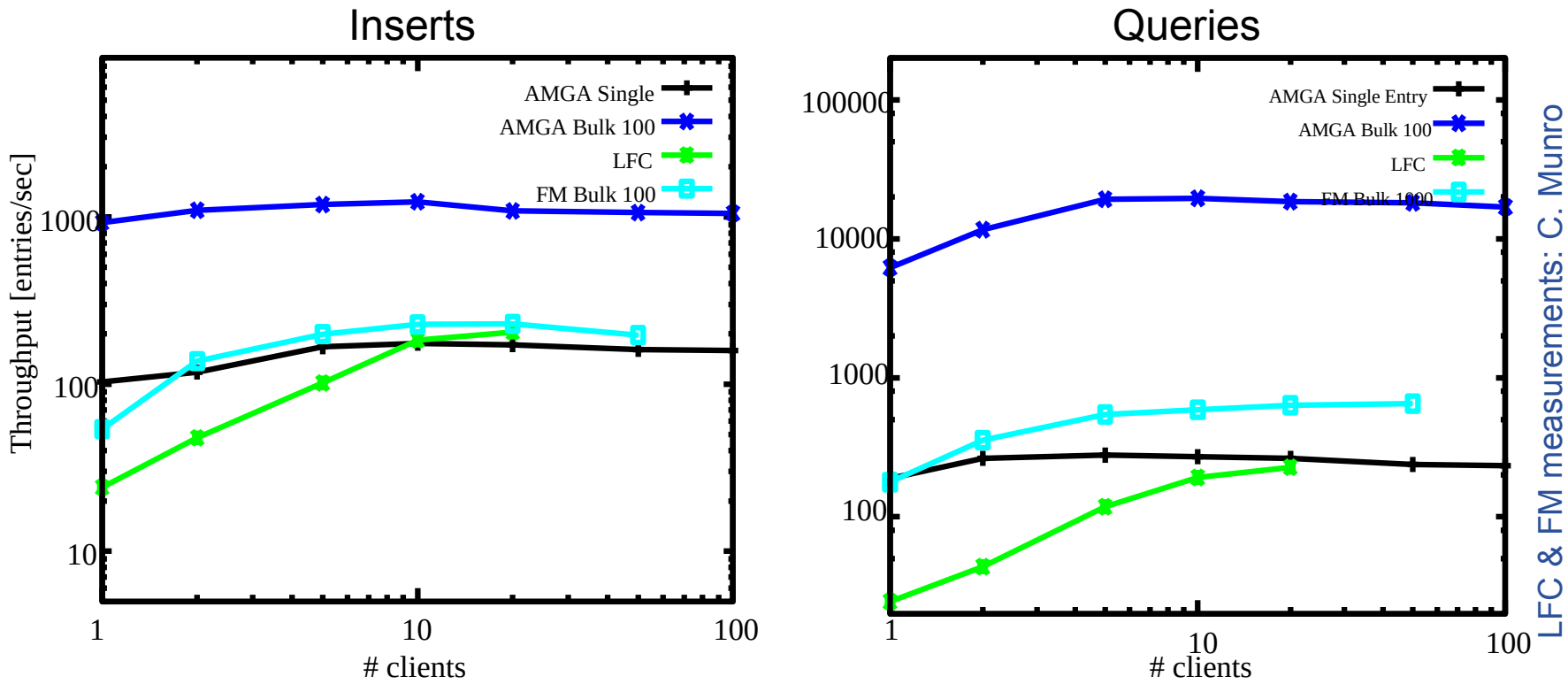
Security tested by GILDA and BioMed teams: Very positive feedback for built-in group management & ACLs

AMGA Implementation

- **AMGA Implementation:**
  - SOAP and Text frontends
  - Streamed Bulk Operations
  - Supports single calls, sessions & connections
  - SSL security with grid certs (negociated by client)
  - Own User & Group management + VOMS
  - PostgreSQL, Oracle, MySQL, SQLite backends
  - Works alongside LFC
  - C++, Java, Perl, Python clients



LAN Performance



- **Protocol comparison with LFC and FiReMan catalogues:**
  - Authentication with X509 Certs, SSL connections
  - LFN/GUID pairs inserted, query for GUID of LFN, Oracle DB
  - AMGA scales very well up to 100 concurrent client (DB Limit)
  - Streamed bulk inserts/queries are very fast!

AMGA in Use

- **AMGA in (pre-)production within several projects:**
- **LHCb and ATLAS: GANGA**
  - Highly dynamic relational data for tracking jobs
  - AMGA used for communication between user clients
- **LHCb Logging and Bookkeeping**
  - 15 Million log entries of jobs
- **EGEE BioMed applications**
  - Highly secure access to medical images metadata
  - Many different security roles, uses VOMS
- **Generic applications:**
  - Metadata for EGEE-GILDA, gLibrary
  - UNOSAT project: Satellite images metadata, GIS

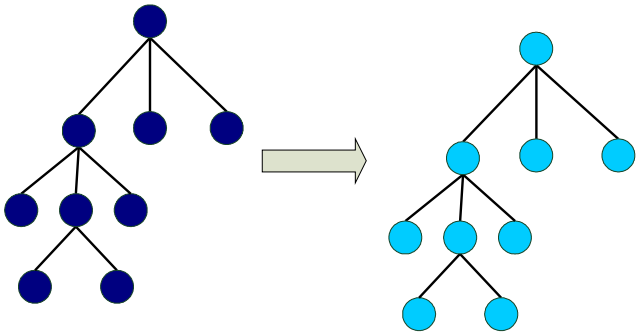
Metadata Replication

- **Replication of metadata on Grid necessary:**
  - Scalability: more connections, smaller latencies
  - Reliability: Avoid single point of failures
  - Administration: Less DBA time, central management possible
- **Replication and Federation is built into AMGA**
  - Replication of Metadata commands allows replication across different databases
  - Setup of replication can be defined by VO managers
  - Asynchronous shipping of logs from master to slaves
  - Single master per directory/directory hierarchy
    - ➔ To improve write performance on master use more powerful solution for database back-end instead!

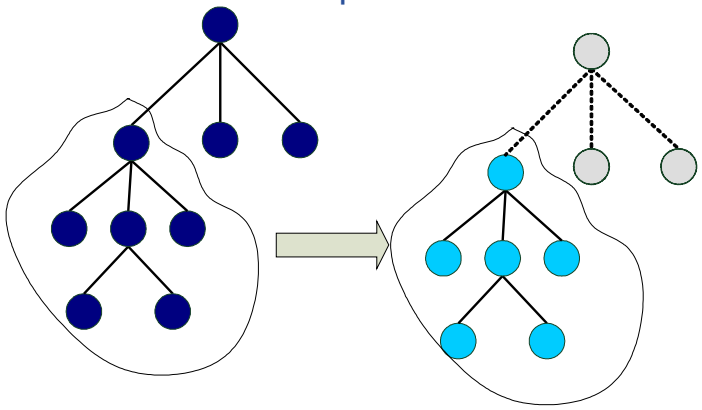
Replication & Federation Modes

- **AMGA replication makes use of hierarchical concept:**

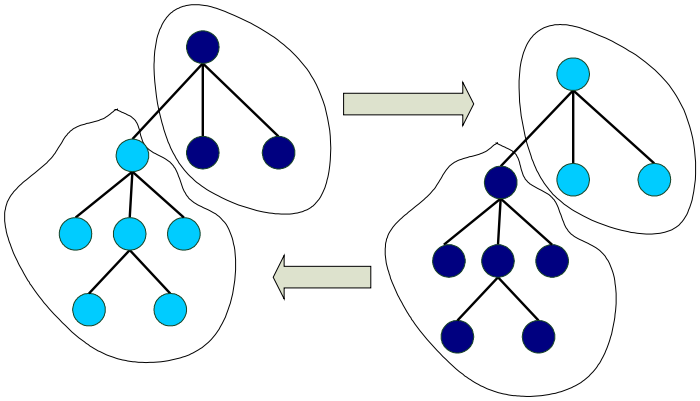
Full replication



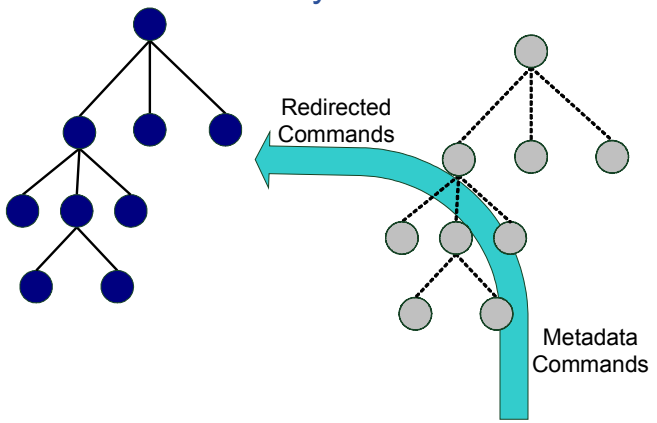
Partial replication



Federation



Proxy



Replication Use Cases

AMGA's different Replication/Federation modes support very different Use Cases:

- **High Energy Physics:**
  - Centrally organized production + several 1000 users (readers)
  - ➔ Replicate read-only copies of entire metadata or tree parts to tiers 1, 2 & 3 (laptop)
- **Biomedical Applications**
  - Federation and Proxying of metadata allows to leave highly confidential data in the places of origin (hospitals)
    - ➔ No copies of data in other database backends!
- **Initial implementation currently tested**
  - UnoSat
  - BioMed: Medical data manager
  - Soon: LHCb metadata

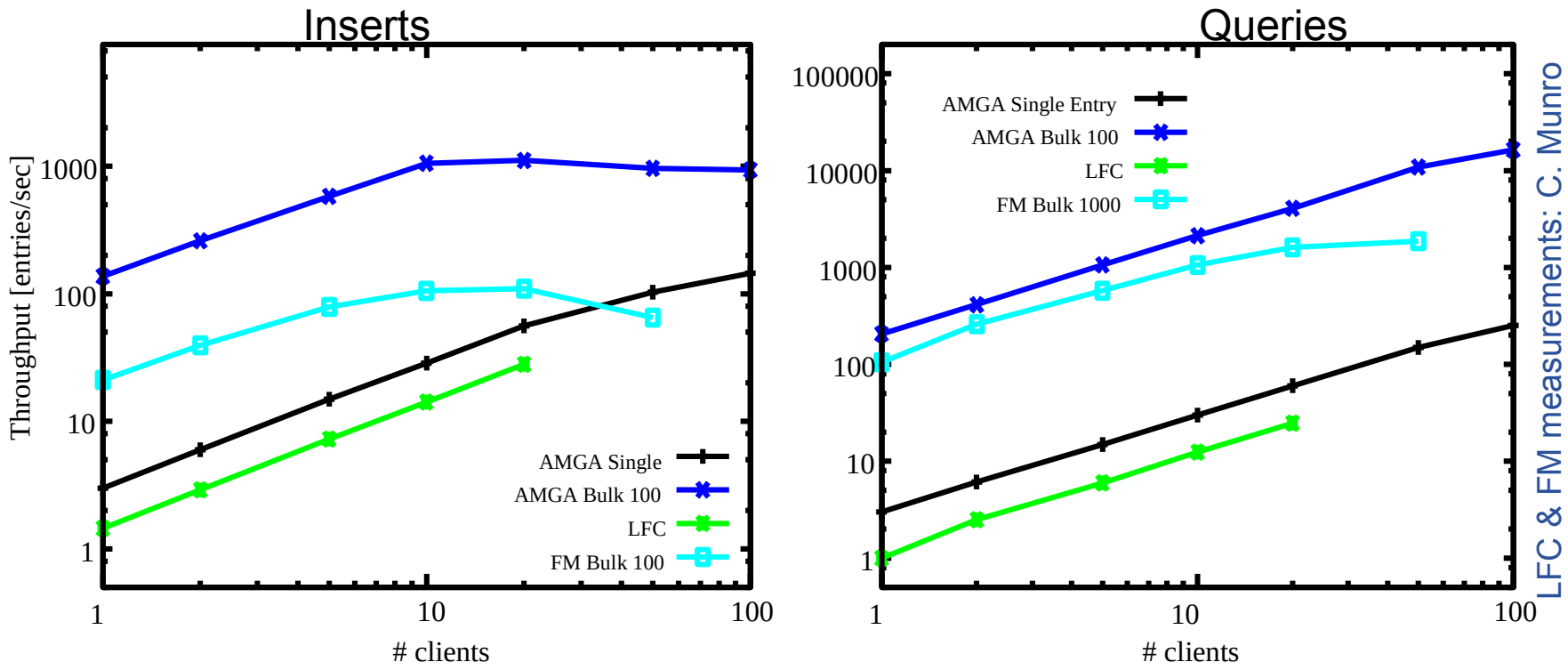
Summary

- **ARDA metadata server AMGA now in gLite 1.5**
- **Implements most of EGEE Metadata Interface**
- **AMGA has seen heavy performance/stability testing**
  - AMGA's protocol throughput 10 times higher than LFC's or FiReMan's due to combination of sessions and bulk streaming
- **AMGA includes very flexible (and optional) security features**
- **AMGA currently evaluated / in preproduction:**

  LHCb(GANGA, bookkeeping), MDM, gLibrary, UnoSat...
- **Replication features now available in AMGA (beta)**
- **AMGA Web Site**

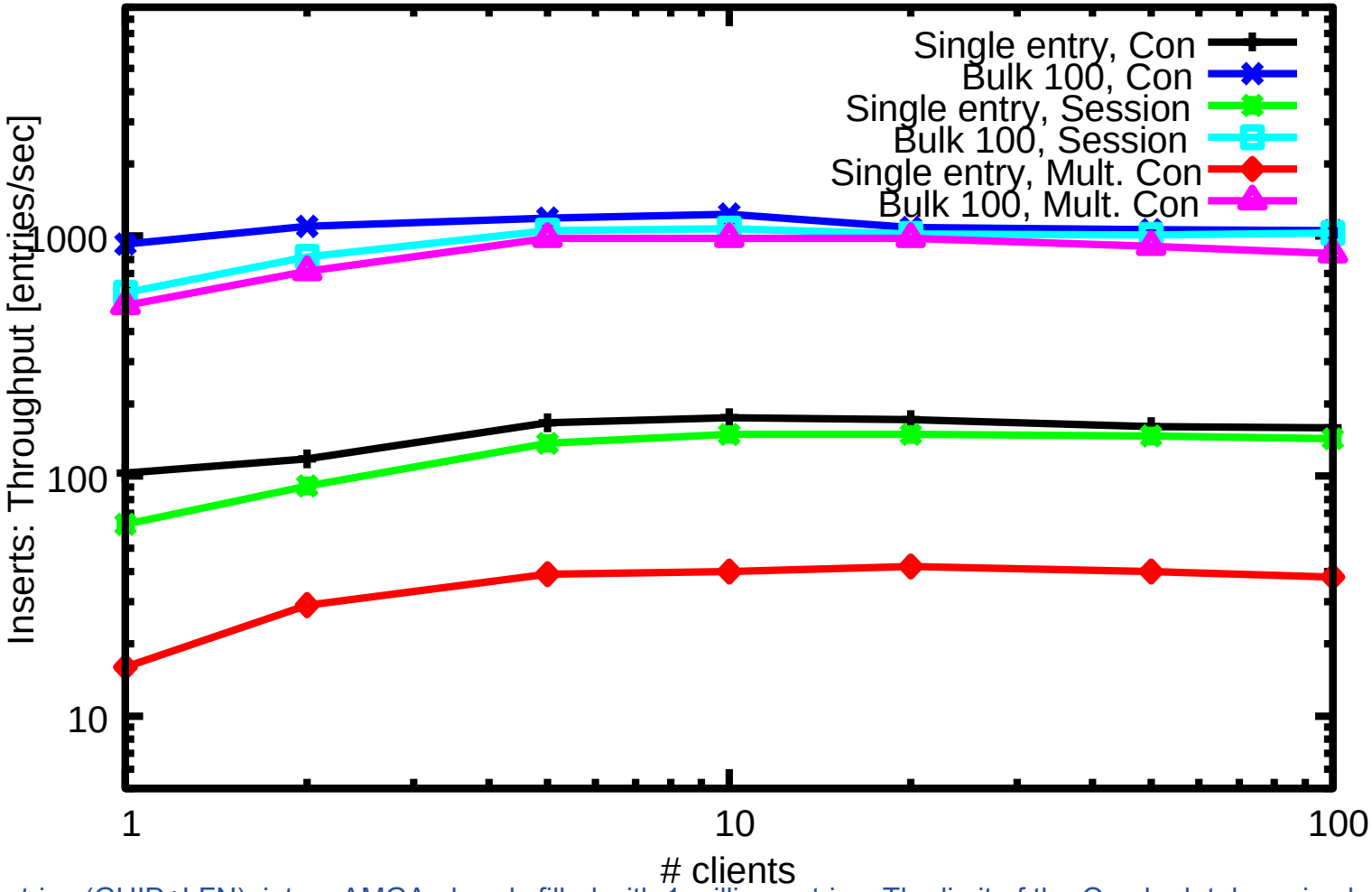
  http://project-arda-dev.web.cern.ch/project-arda-dev/metadata/

WAN Performance



- **Comparison with FC protocols, connection from Taiwan:**
  - 300ms latency dominates performance
    - Reduce round-trips with sessions or holding connections
  - (Streamed) bulk operations vital for WAN performance

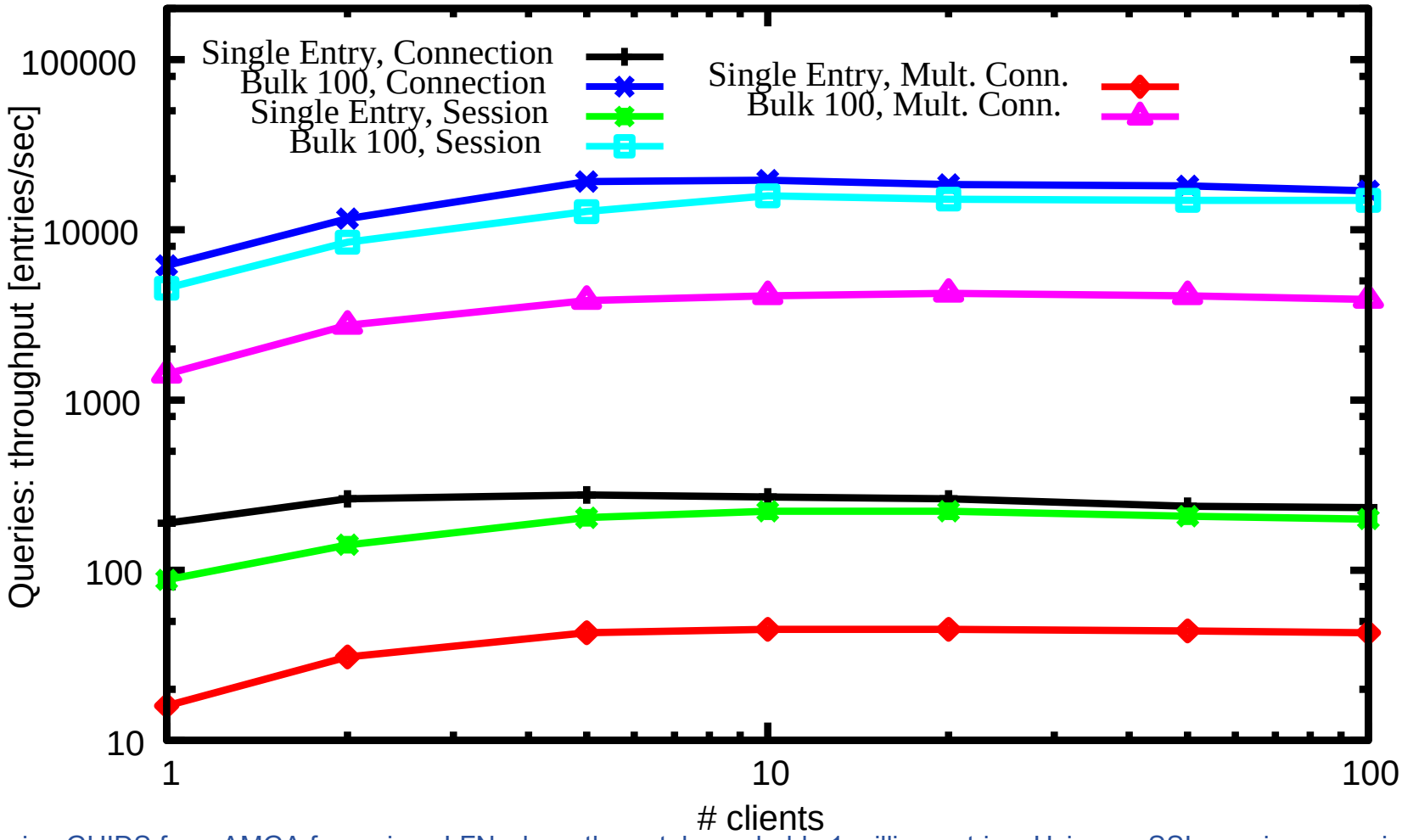
AMGA Performance: Inserts



Inserting entries (GUID+LFN) into a AMGA already filled with 1 million entries. The limit of the Oracle database is about 1000 entries/sec which can be reached with bulk operations of 100 entries in all of the 3 connection modes of AMGA: Multiple individually authenticated TCP connections, a single SSL session spanning multiple TCP connections and a single TCP connection. The SSL session is nearly as fast as the single connection.

All measurements with an Oracle backend.

AMGA Performance: Queries



Retrieving GUIDS from AMGA for a given LFN where the catalogue holds 1 million entries. Using an SSL session spanning several TCP connections is about 10 times faster than having several TCP connections where the client authenticates in each. A single connection is only sightly faster than a session. A bulk operation with 100 entries is able to increase the speed over the single entry operation in all cases by about another factor of 10.