

Support for Metadata Federation in AMGA 2.1

What is AMGA ? (ARDA Metadata Grid Application)

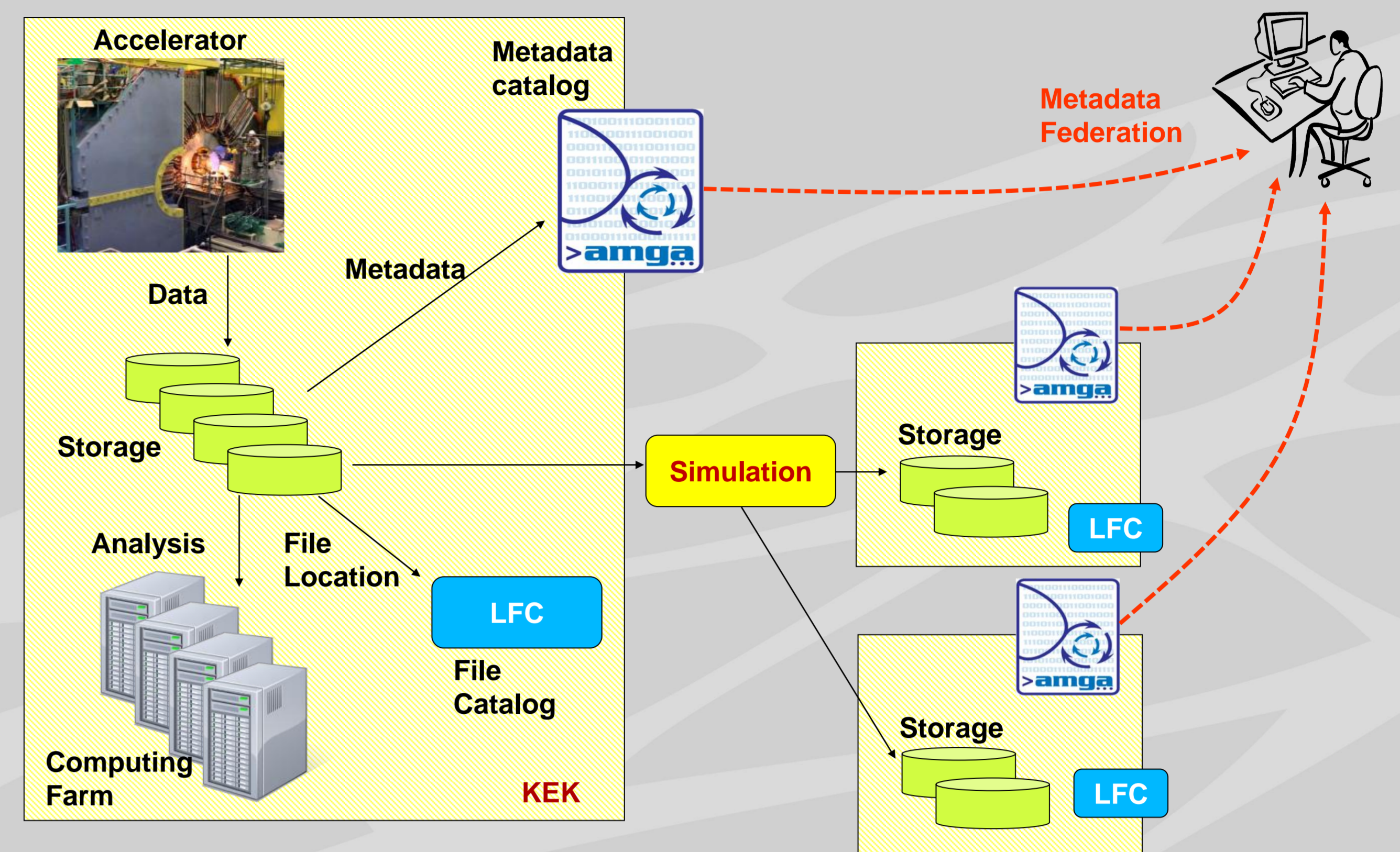
- Metadata Catalogue Service of EMI Middleware
 - Metadata is data about data, On the Grid: information about files
 - But also simplified DB access on the Grid
- Current Release : v 2.1
- More Information : <http://cern.ch/amga>

USE CASE EXAMPLE : BELLE II

- 50 times of data and metadata than Belle
- Metadata on users' simulated data may be stored at separated AMGA servers to reduce complexity and boost performance
- Federation mechanism may be used to integrate all metadata distributed at multiple sites

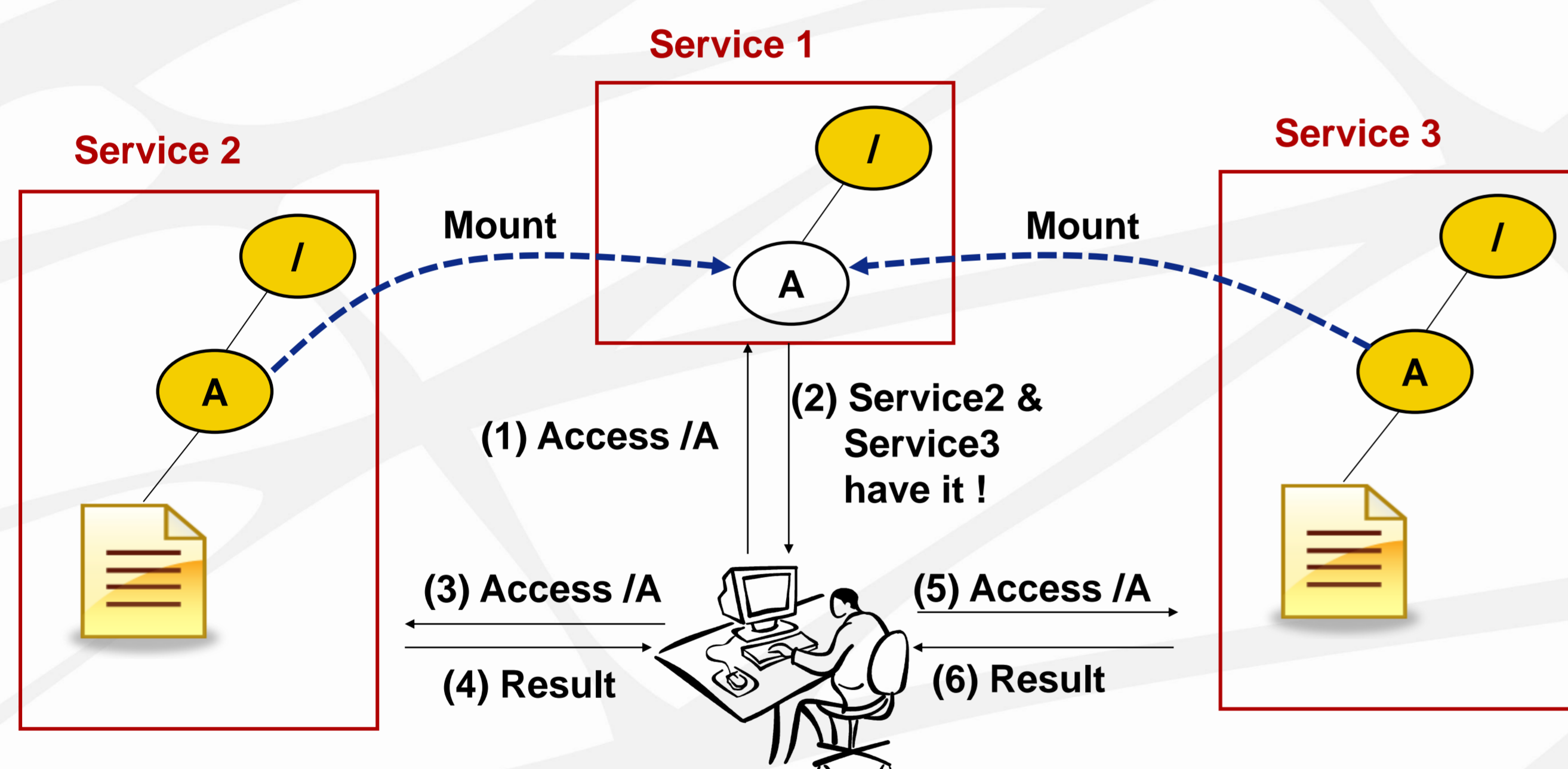
Metadata Federation

- A Mechanism to integrate distributed metadata
 - Provide a user with a virtualized view on metadata as if one metadata server has all data which are actually distributed at multiple sites.
- Two types : Server Side, Client Side
- Parallel Federation Support
- Policy based Fault Handling



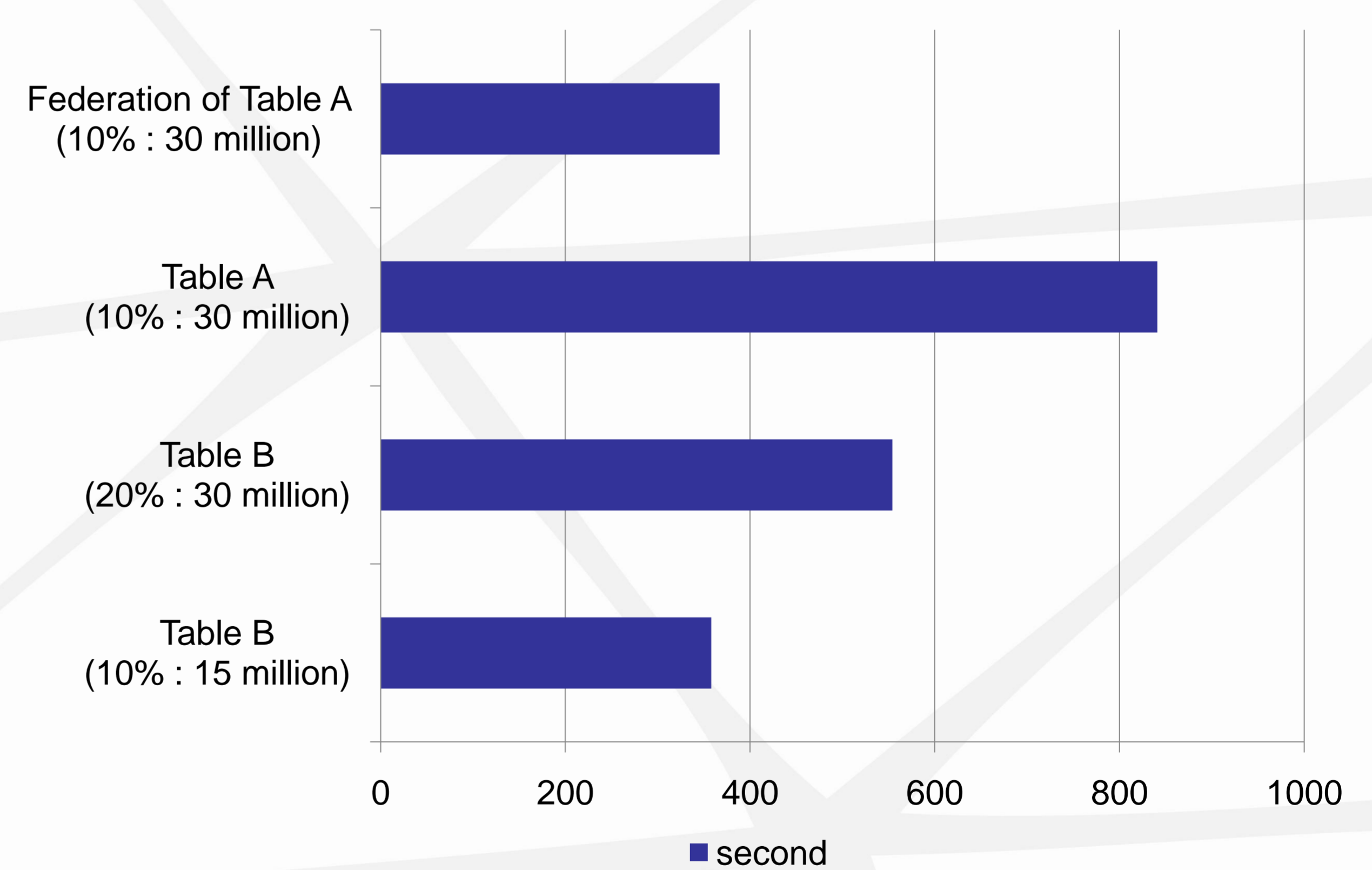
Client-Side Federation

- Low Overhead at Server Side
- Only works with C++ APIs Currently



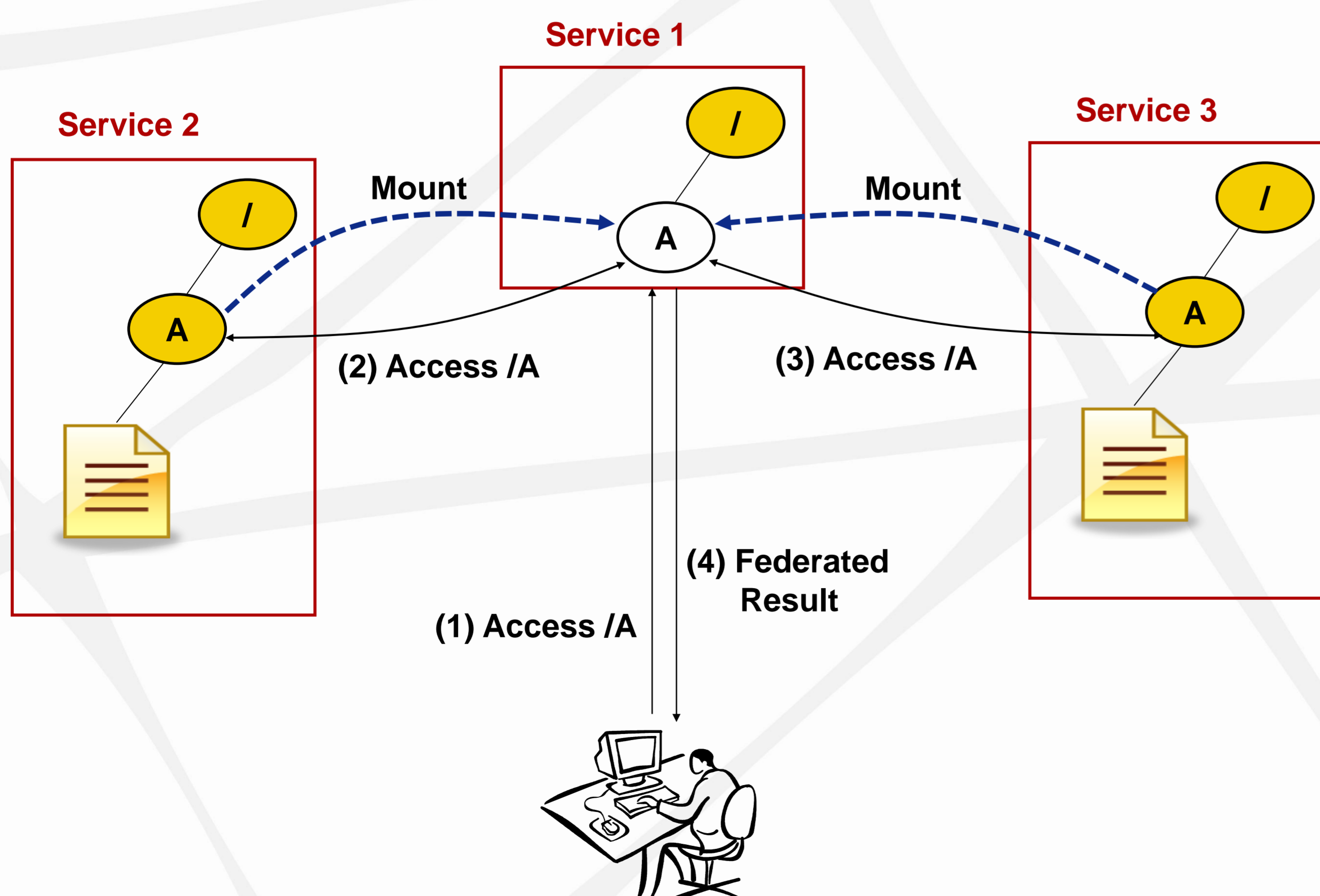
Preliminary Performance Evaluation of Federation

- Server 1
 - Table A : 300 million rows
 - Table B : 150 million rows
 - Each table has 7 columns, Values are random in a certain range
- Server 2
 - Table B : 150 million rows
- Using 2 Federated tables retrieves metadata more than twice faster than using 1 large table



Server-Side Federation

- High Overhead on Server Side
- Works with all the APIs; C++, Java, PHP, Python, Perl



Current Limitations on AMGA Federation

- No Schema Heterogeneity Support
- No Distributed Join Support
- No Transaction Supported
- No Relative Path Support