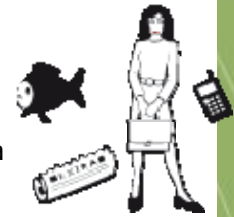


GridTorrent Transfer Mechanism

Motivation and Challenges

In terms of the GREDIA project, vast amounts of data are being manipulated. The platform should be able to efficiently transfer multimedia content, produced by geographically dispersed sources and to cope with high request rates – to the point of flash crowd. GridFTP, the transfer mechanism currently deployed in Grids, heavily relies on centralized mechanisms, which constitute performance bottlenecks and single points of failure. GridFTP does not make optimal use of all bandwidth resources in cases where the same data must be transferred to multiple sites, nor does it scale to a large number of concurrent requests. Therefore, Gredia is in need of a data transfer mechanism that effectively deals with large and concurrent file uploads and downloads, even when numerous requests rely on a single data source, maximizing bandwidth utilization.



GridTorrent Solution

GridTorrent is an implementation of the popular BitTorrent protocol designed to interface and integrate with well-defined and deployed Data Grid components and protocols (e.g. GridFTP, RLS). Just like BitTorrent, GridTorrent is based on P2P techniques, that allow clients to download files from multiple sources while uploading them to other users at the same time, rather than obtaining them from a central server. By dividing files into fragments, GridTorrent can combine the best out of the two protocols:

- It exploits BitTorrent's peer and fragment selection, thus providing an optimized data transfer service.
- It takes advantage of the striped version of GridFTP protocol, as it has the ability to directly communicate with GridFTP servers, thus being backwards compatible.

Innovation

GridTorrent focuses on real-time optimization of data transfers on the Grid, fully supporting the induced security mechanisms. Based on collaborative sharing, GridTorrent allows for low latency and maximum bandwidth utilization, even under extreme load and flash crowd conditions. It allows transfers from multiple sites to multiple clients and maximizes performance by piece exchange among the participants. A very important characteristic of the proposed architecture is that it is designed to interface and exploit well-defined and deployed Grid components and protocols, thus being completely backwards compatible and readily deployable.



Business Impact

Data intensive scientific applications widely benefit from the capabilities offered by Grid Computing. However, over the last years, Grid Computing has attracted the business world as well, serving as a solution for business applications manipulating large amounts of data. Such business applications could greatly benefit from a transfer mechanism like GridTorrent, especially when they face the danger of heavy workload conditions, which can cause the collapsing of conventional GridFTP servers.



Interoperability

The GridTorrent implementation is entirely written in Java. The GridTorrent client has bindings with Globus Toolkit 4 libraries and exploits the GridFTP client API, the Replica Location Service API and the Grid Security Infrastructure API. These bindings enrich our prototype with the abilities to use existing grid infrastructure, such as data stored in GridFTP servers, metadata stored in Globus RLS and x509 certificates that are already issued to users and services for authentication, authorization, integrity protection and confidentiality. Thus it is completely interoperable with existing Grid solutions.



Partners Involved



Contact

Institute of Communication And Computer Systems

<http://www.iccs.gr/>

Dimitrios Tsoumakos