

Attic: A Case Study for Distributing Data in BOINC Projects

AbdelHamid Elwaer, Andrew Harrison, Ian Kelley and Ian Taylor

Paper: [Attic: A Case Study for Distributing Data in BOINC Projects](#)

Web: <http://www.atticfs.org/>

by Diego

Problem Statement

The BOINC architecture for data distribution is not scalable and reliable because it is based on a centralized server, or fixed set of pre-configured servers, each with a copy of a project's input files.

Besides, this centrally administered approach derives in a complex management due to the servers have to be manually installed and configured for use.

Proposal

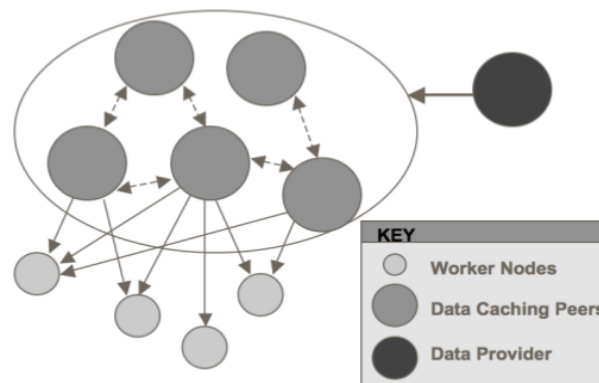
This paper introduces Attic, a software project that creates a secure data distribution network applicable to projects like BOINC that employs P2P data sharing practices to share data across the data distribution layer. This alternative provides a more lightweight and dynamic environment for distributing data.

Properties:

- It has been designed with volunteer computing in mind, and does not require all participants to share data, or share equally, therefore differing it from other types of P2P file-sharing systems.
- Secure decentralized approach and BitTorrent-like file swarming techniques to serve data and manage load
- It allows client nodes to forgo "tit-for-tat" algorithms to receive data.
- It uses HTTP for the transfer data transfer layer (multiple concurrent downloads)
- It can be integrated with BOINC projects

Four main elements:

- A data serving application that replicates data on the network.
- Data Centers (data cacher) that cache data, providing the distributed data source overlay.
- A Look up service that keeps track of which Data Centers have individual data items.
- Client applications that download



data from Data Centers on the network.

Hypotheses

Improve the throughput using the Attic storage system over BOINC projects.

Experiments

Comparison between the Attic approach and the current centralized BOINC data distribution. To do this, they conducted three experiments where Attic was integrated with BOINC and showed the Attic's performance using different parameters, namely: chunk size and the number of data centers.

Parameter for the experiments:

- 19 networked Linux machines from their laboratory.
 - 18 machines used to run combinations of clients and data centers.
Hardware: 2.8GHz P\$ processor and 2GB RAM.
 - 1 machine was used to run either the Attic data look-up server.
Hardware: 2.0GHz Pentium Dual Core processor and 3GB RAM.
- Machines are connected through a LAN network and.
 - The speeds of the ports were configured. A subset of machines with 10Mbps and others with 100Mbps

Results

The experiments' results show that careful choice of chunk size, by analyzing a server's uploading capabilities and file size, can result in large throughput gains for attic when compared to the current BOINC data distribution paths.

P.D.: the paper focus on a brief overview on their approach. The results are relevant in that they test the Attic architecture and load distribution that takes place within the applied scale (no the real world).

A point of failure: the Lookup Server.

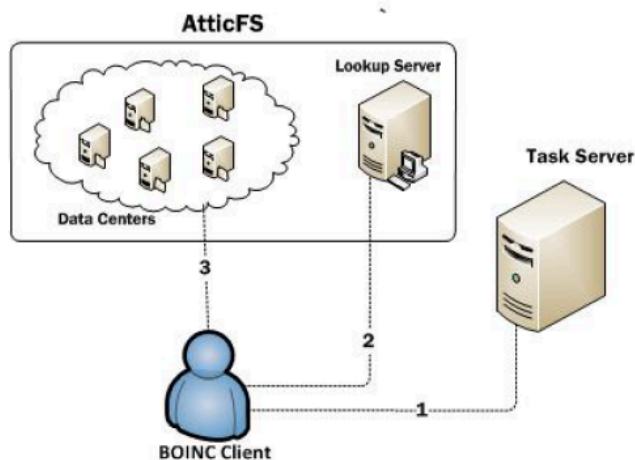


Figure 2. Integrating Attic in BOINC Project