

# Distributed Systems: Project Proposal

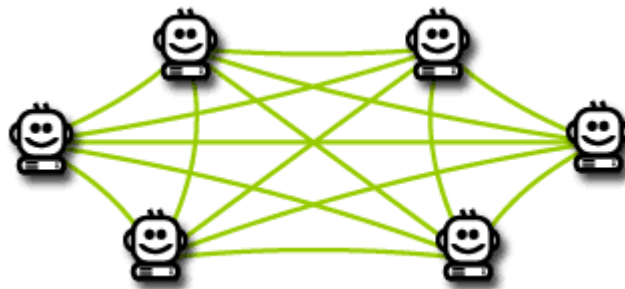
## ***Topic: Access and Billing in p2p environment***

Group Name: PHOENIX

Gergana Krumova  
Mariana Marin  
Alexander Stasiv  
Lazar Adzigogov

### **Challenge (Identified problem)**

Today there is a large explosion of p2p-based systems on internet. One representative is KaZaa, world's number one sharing software application, which enables users in p2p systems to search for and download music, documents, images, play lists, software and videos. One of the existing problems is how to manage secure access and bill the engaging parties.



P2P system: *Picture taken from KaZaa Web site*

For example: Any node (user A) using KaZaa system can search and find the requested file (music, document), that exist at another user B. Now user A can establish direct link for downloading the requested file from user B, but what KaZaa does not offer is billing of downloading a file, user A is not charged. Therefore, the challenge is to make p2p system, where billing can be accomplished. In addition, while downloading the file, KaZaa does not offer any confidentiality, protecting the content from eavesdropping. Therefore the second challenge of the new p2p system will be to provide confidentiality (need for access control and encryption mechanisms), only user A that paid for the downloaded file can be able to see its contexts, nobody else, even if the file during the downloading process is eavesdropped.

## Recommended Solution (See the conceptual drawing on the next page)

The following solution gives the overall idea and the usage of the system in practice. The proposed name for the project is **ABC (Access and Billing Control)**.

Our p2p billing system will be used by two types of users: Providers and Consumers. A provider is a user that wants to share (sell) information (music, documents, video, movies, etc) for a given price and for a given speed of downloading. A consumer is a user that searches for particular information and if he/she is willing to buy the found document, he/she pays for it. A user can be a provider/consumer at the same time. A user becomes provider by downloading application called ABC node, which enables the user to specify which documents he/she wants to sell for specified price. This module updates this information to a specific server called CBL (Certificates, Billing, Locator), where practically the user accounts is created with the information/prices that he offers. The downloaded module also provides confidentiality (encryption mechanism), when a particular consumer wants to download specific document from the provider. A user connects to the system as a consumer in the same way, and after successful authentication (signing in process was successful, here he puts the amount of money he has), the ABC offers to the consumer some kind of search engine in order to help the user to find the requested general information. If the consumer found what he was looking for and is willing to pay for it, it establish secured authenticated channel with the provider and downloads the document. All the information about the users sharing information and completed transaction will be stored possibly on MySQL server, since it is free of charge.

The system will be managed by special users (Administrators), which will take care for the chain of p2p CBL servers that hold the user's information's (configuring and maintenance). There will be special application (GUI) for this purpose.

# ***Detected Problems in Peer- To – Peer networks regarding access control and billing***

## **1. Searching Information**

P2P networks as communication model connect together a large number of different computers, pooling their resources. They provide an efficient way of storing and accessing large amounts of data, as demonstrated by the popularity of music file sharing networks such as KaZaa. One big drawback is that they are de-centralized, i.e. there is no central server that keeps tracks of all the information in the system. This means searching for information on these networks is slow and cannot guarantee to find information.

Proposed solution:

Our distributed systems will use middleware, specific dedicated nodes of the network will play role as storage facilities for keeping index of content of each node (user) that are sharing/selling general information.

## **2. High Availability**

Services such as sharing/ billing/selling must be available during the whole day, in order to preserve the customers of any distributed selling. If one point of service is available, its failure will be the death of the entire system. Therefore, some kind of mechanism must be provided in order to ensure high availability.

Proposed solution:

Providing high availability can be achieved by data replication. Our distributed system will have subsystem called “Replication Control”, which will expect dedicated node failures and deal with it by making multiple copies of the data, ensuring that end user will not see any node failures and will continue to use the system like nothing happened. Failure of one or more dedicated nodes will be transparent to the end user. Initially we are considering using active replication protocol. Type of database is MySQL, and the update will be made manually by the ‘Replication control’, not automatically by the database system.

The problem of interrupted downloading of information is solved by segmented downloading. When the user will connect again, the downloading will proceed.

## **3. Security**

The distributed systems for sharing information raises the prospect of it being very easy to publish information anonymously, for example, pirated music and video. P2P networks can also allow an user to download and use copyrighted material in a way that violates intellectual property laws, and to share files in a manner that violates an organizations security policies. The notion of authentication, authorization, ownership arises. In addition, there are other threats that exist, such as viruses, worms, malicious programs, that spread very easily on file-sharing distributed systems and can cause malfunctioning of end user system, when they download some document.

Proposed Solution:

Our distributed system will provide authentication, meaning that the user who will download the file, will be assured that he is downloading from the right person, and the user who is giving the information, is assured that he allows copying of document to authenticated person. Initially we are planning to use MAC.

In addition, encryption mechanism through authenticated channel will be used while the document is transported over the network, therefore protection against eavesdropping is ensured. Nobody else but the downloading person can be able to decrypt the document.

## 4. Management and Maintenance

Distributed Systems by definition are harder to maintain, since they are consisted of more than one node. There is a need to do diagnose and locate problems very easily.

Proposed Solution:

Special kind of users, Administrators will be responsible for managing the distributed systems. The subsystems, 'Replication control' , and 'Fail-safe control', will be responsible for detecting failures, and reporting to the Administrators, the type of failure occurred and the measures taken, the failures was fixed, or human interaction is required. For the administrators, specific software module will be implemented.

## 5. Scalability

By default one of the requirements of every distributed system is to be scalable. The performance of the system should not drop, if the number of the users increases and geographically the system will expand its distribution.

Proposed Solution:

The system design is scalable. As number of the users' increases, the subsystem 'Replication control' will take care for load-balancing, replication copies, caching if necessary, pre-fetching.

## 6. Priority

Since there will be many users in the distributed system, the Internet will be loaded with heavy traffic, not only by downloading, but also with control messaging. Should the distributed system support priority among users, in case where one user makes a lot of transactions/spends a lot of money, therefore downloading more documents, than the users that download only one document? Does the user with higher downloads have higher priority then other users?

Proposed Solution:

For now, we do not have some initial solution regarding this problem.

Reference:

(1) NewScientist.Com