# What is P2P?

P2P (peer-to-peer, person-to-person) file-sharing protocol refers to protocols for the dissemination and exchange of information using a special program to search for other users on the network, in which all users (peers) are equal participants, in contrast to the client-server model of information exchange.

File sharing technology is a groundbreaking innovation that allows individuals and companies to seamlessly access, transfer and distribute multimedia files such as music, videos, documents and games.

The peculiarity of this decentralized approach is its ability to provide faster downloads while reducing the load on the server compared to traditional file transfer.

It has revolutionized the way we share digital content, allowing users to transfer data from one device to another without relying on centralized servers or external storage devices.

For some protocols (hybrid P2P, for example, Napster or BitTorrent), you need to connect to the server to find out information about files or about users from whom you can take a certain file. The data exchange itself is carried out between network users, which is the main feature of P2P.

In such networks, all computers (network participants) are equal. In particular, everyone can perform both the functions of the server where the files are stored and the functions of the client downloading them. According to a similar model, what we used to call "torrents" works.

Programs (clients) that support this technology use a special peer-to-peer network protocol (BitTorrent), one of its features is that each network participant who downloaded any fragment of the transmitted file to his device can distribute it to other participants.

Several factors have contributed to the widespread and simplified peer-to-peer file sharing. These include an increase in Internet bandwidth, widespread digitization of physical media, and the increasing capabilities of home personal computers. Users can transfer one or more files from one computer to another over the Internet through various file transfer systems and other file sharing networks.

# Peer-to-peer network device

There are a number of machines in the network, and each can communicate with any of the others. Each of these machines can send requests to other machines to provide any resources within this network and, thus, act as a client. Being a server, each machine must be able to process requests from other machines on the network, to send what was requested. P2P file sharing works by creating a network between devices, allowing each node to upload files from other nodes or upload files to them without relying on a central server.

Each machine must also perform some auxiliary and administrative functions (for example, store a list of other known "neighbors" machines and keep it up-to-date).

Any member of this network does not guarantee its presence on a permanent basis. It can appear and disappear at any time. But when a certain critical network size is reached, there comes a moment when there are many servers with the same functions in the network at the same time.

An example of such a network: I2P, Gnutella2.

P2P Architecture:

1. Centralized Directory
2. Query Flooding
3. Exploiting Heterogeneity
4. Centralized Directory

A centralized Directory is somewhat similar to client-server architecture in the sense that it maintains a huge central server to provide directory service. All the peers inform this central server of their IP address and the files they are making available for sharing. The server queries the peers at regular intervals to make sure if the peers are still connected or not. So basically, this server maintains a huge database regarding which file is present at which IP addresses. The first system which made use of this method was Napster, for Mp3 distribution.

Working:

* Now whenever a requesting peer comes in, it sends its query to the server.
* Since the server has all the information of its peers, so it returns the IP addresses of all the peers having the requested file to the peer.
* Now the file transfer takes place between these two peers.

1. Query Flooding

Unlike the centralized approach, this method makes use of distributed systems. In this, the peers are supposed to be connected to an overlay network. It means if a connection/path exists from one peer to another, it is a part of this overlay network. In this overlay network, peers are called nodes, and the connection between peers is called an edge between the nodes, thus resulting in a graph-like structure. Gnutella was the first decentralized peer-to-peer network.

Working:

* Now when one peer requests for some file, this request is sent to all its neighboring nodes i.e. to all nodes connected to this node. If those nodes don’t have the required file, they pass on the query to their neighbors and so on. This is called query flooding.
* When the peer with the requested file is found (referred to as query hit), the query flooding stops and it sends back the file name and file size to the client, thus following the reverse path.
* If there are multiple query hits, the client selects from one of these peers.

Gnutella: Gnutella represents a new wave of P2P applications providing distributed discovery and sharing of resources across the Internet. Gnutella is distinguished by its support for anonymity and its decentralized architecture. A Gnutella network consists of a dynamically changing set of peers connected using TCP/IP.

This method also has some disadvantages, the query has to be sent to all the neighboring peers unless a match is found. This increases traffic in the network.

1. Exploiting Heterogeneity

This P2P architecture makes use of both the above-discussed systems. It resembles a distributed system like Gnutella because there is no central server for query processing. But unlike Gnutella, it does not treat all its peers equally. The peers with higher bandwidth and network connectivity are at a higher priority and are called group leaders/supernodes. The rest of the peers are assigned to these supernodes. These supernodes are interconnected and the peers under these supernodes inform their respective leaders about their connectivity, IP address, and the files available for sharing.

KaZaA technology is such an example that makes use of Napster and Gnutella. Thus, the individual group leaders along with their child peers form a Napster-like structure. These group leaders then interconnect among themselves to resemble a Gnutella-like structure.

Working:

* This structure can process the queries in two ways.
* The first one is that the supernodes could contact other supernodes and merge their databases with their database. Thus, this supernode now has information about a large number of peers.
* Another approach is that when a query comes in, it is forwarded to the neighboring super nodes until a match is found, just like in Gnutella. Thus query flooding exists but with limited scope as each supernode has many child peers. Hence, such a system exploits the heterogeneity of the peers by designating some of them as group leaders/supernodes and others as their child peers

# Types of peer-to-peer (P2P) file sharing

There are two types of P2P file sharing: centralized and decentralized, each with its own advantages and risks.

P2P file sharing can be divided into two main types: centralized and decentralized peer-to-peer networks.

1. Centralized P2P networks

* + A centralized P2P network has a central server that manages the network and stores information about files that are shared.
  + Users connect to this central server to search for files or upload their own files.
  + Napster and Limewire are examples of centralized P2P networks.

2. Decentralized P2P networks

* + There is no central server in a decentralized P2P network. Instead, each device on the network acts both as a client and as a server.
  + This means that each device can download files from other devices, as well as upload files to them.
  + Examples of decentralized P2P networks are BitTorrent and Kazaa.

It is important to note that both types of P2P networks have their advantages and risks. Centralized networks are generally easier to use, but may be more vulnerable to lawsuits due to their dependence on a central authority. Decentralized networks provide greater anonymity, but can also carry a greater risk of malware or copyright infringement. Thus, users need to understand the differences between these two types of P2P networks before engaging in file sharing.