Bit Torrent

Friday, 12 June 2020 12:44 PM

Key points-

1) Protocol for peer to peer file sharing. A bitTorrent client is an application that uses this protocol.
2) follows a hybrid architecture instead of P2P.
3) Data is downloaded and uploaded directly to and by peers.
4) Allow easy exchange of large files.
5) A bittorrent client requests files from multiple clients in parallel.
6) Small chunks of data is called pieces.
7) If a client downloads a piece successfully, bittorrent tells other clients that ok, this client has this piece of data downloaded. So they can download it too.
8) These collection of collaborating clients are called "Swarms"

Torrent file.

→ Client joins a swarm by downloading a (.torrent) file that tells the following info.

1) Gives info about the file like how big it is, the size of it's pieces & how it starts interacting other clients
2) Gives detail about a Tracker (server that tracks who is participating in the swarm)

→ When client joins a swarm, it req a list of clients from the tracker, and starts communication with these client over TCP (initially as a "Leecher")

→ When the size of swarm increase we can use something like trackerless torrents that use DHT (distributed hash table).

1) It's a distributed coordination mechanism
2) Information on swarm accross many nodes.

! → What exactly bit-torrent does?

① It breaks up file into N pieces

→ for throughput pieces are large : 256KB - 1MB

→ for latency, broken into subpieces.

→ this is to ensure that TCP stream transferring the file is long lived enough that it's congestion window can grow to reasonable size.

② Piece also ensures Integrity

→ A torrent contains SHA-1 hashes of one piece
↳ cryptographic hash func^n

STORY TIME!!

In 2006, HBO released a new series, "ROME". Sound torrents were there with very large swarms. But many people found that clients were not able to download it. It was observed that very very fast peers provided pieces with incorrect hashes, & when client downloaded & found hash is wrong they re-requested it. It seemed to be like a network error to them but there's a theory that HBO was regulating these peers.

↳ That's why now-a-days clients can blacklist peers.

Underhood Mechanism

Peers exchange what pieces they have.

Peers download the rarest piece among peers first.

If any piece is unavailable in all peers no one can download it.

If very few peers have the piece, it becomes bottleneck for the download.

This is called "Rarest first policy"

TFT policy -

→ Tit for Tat

Top P peers

Choking

Most fun on initial checkout

You send data to peers, who send you data.

The peers who contribute can download faster, this creates an incentive to send pieces to peers.

Use your uplink capacity to be divided equally into 'p' top.

NOTE → BIT-TYRANT

↳ Increases throughput by 70%

↳ You preferentially gain access for download.

↳ It ensures that each peer uses uplink capacity exactly equal to just what they need. So there is room to add more peers to p top.