Gonna talk about

- Gnutella / Gnutella2

  - Short history

  - Network description

  - Protocol features

  - How it works

History

- Justin Frankel and Tom Pepper of Nullsoft in early 2000

- First closed source, later under GPL

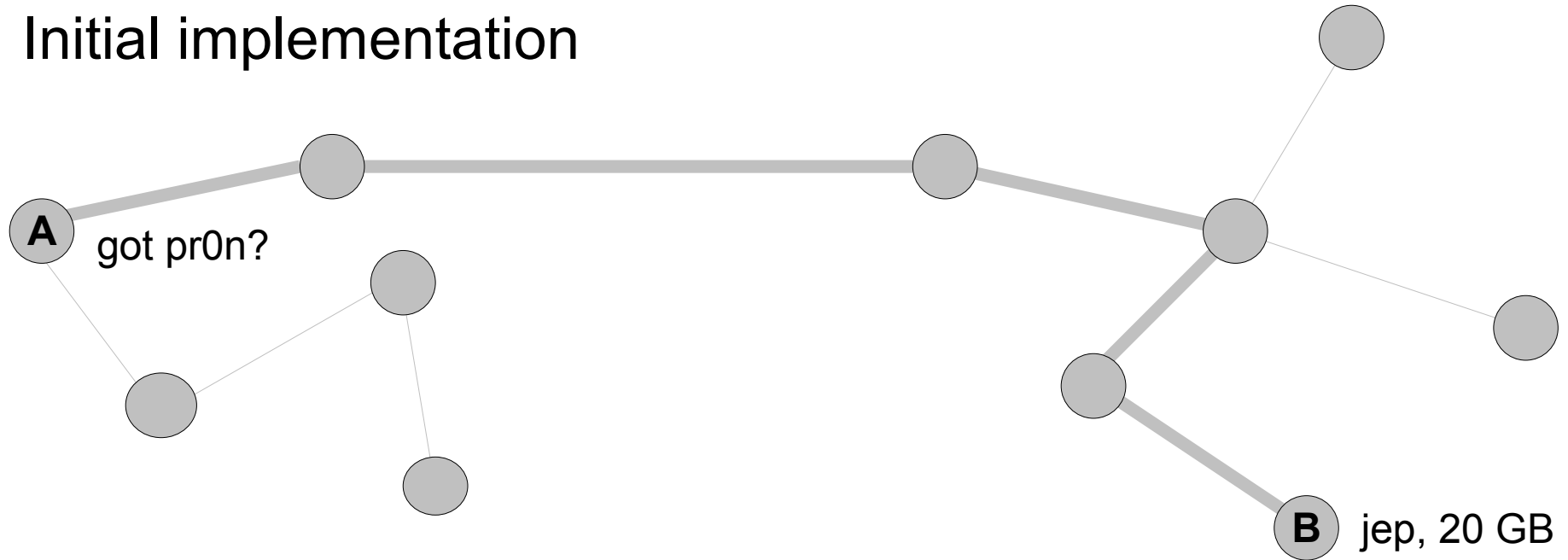- AOL tries to shut down, over legal issues

Network

- Third most popular

- Fully distributed, no sentral servers

- All nodes are equal (1st version)

Bootstrap process

- Very first start

  - Pre-existing address list
  - web caches (GwebCaches)
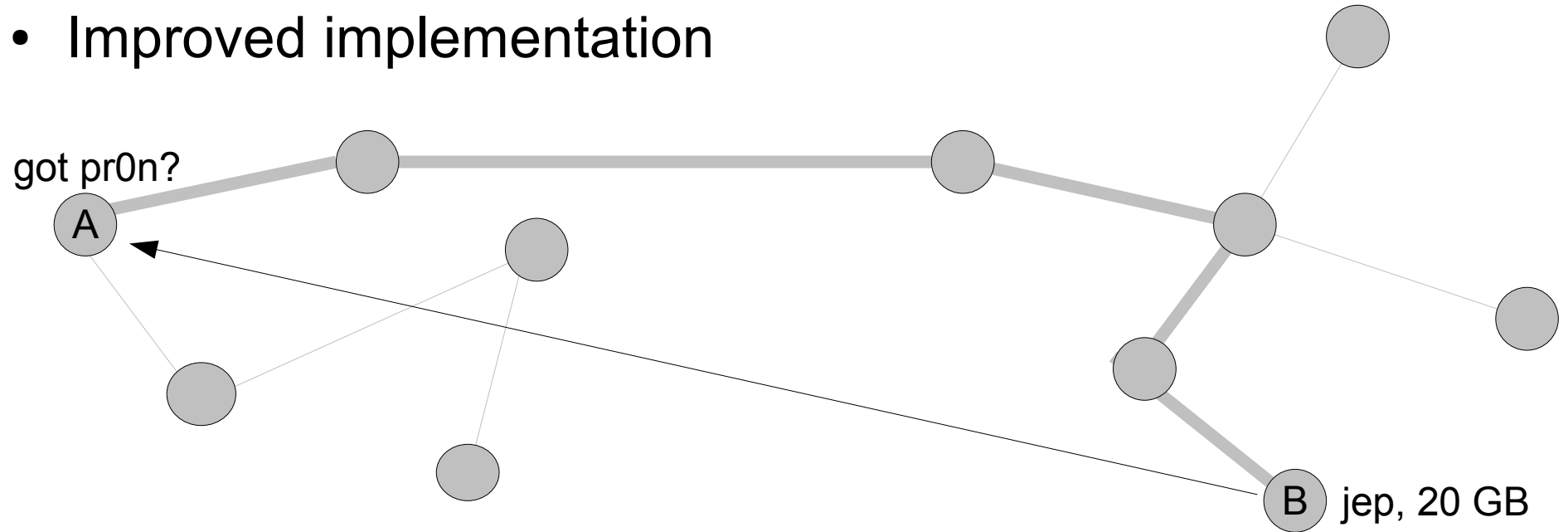 - X-Try header during the handshake
  - IRC

Gnutella search(1)

- Initial implementation

A got pr0n?

B jep, 20 GB

- Inefficient
- Flooding-based protocol
- Expensive (TCP)
- Basically unscalable distributed system
- But – accurate search

Gnutella search(2)

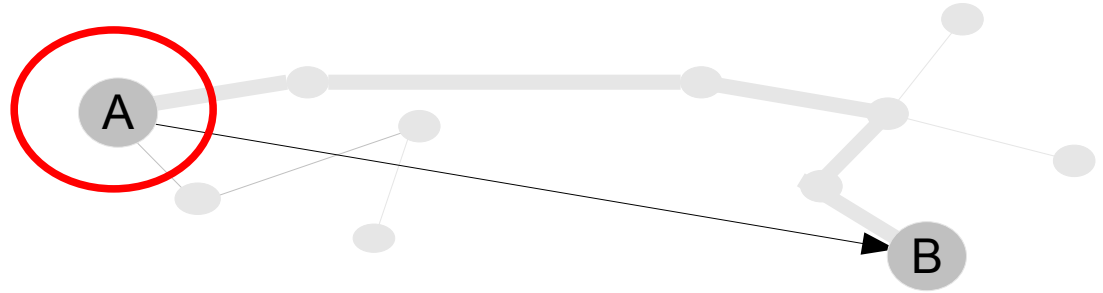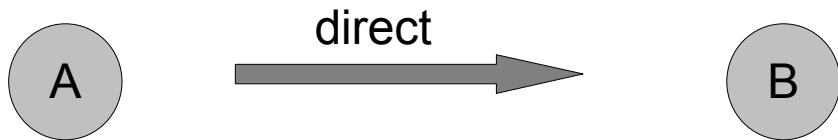- Improved implementation

got pr0n?

A

B jep, 20 GB

- Direct UDP connect
- Less traffic on the wire
- Use of distributed hash tables

Gnutella download(1)

Perfect scenario

A ——— direct ———→ B

Firewalled

A ——— direct ———→ | B

Gnutella download(2)

A

B

"Push" request

A

give me the file X

ok, here's a connection

woot!

B

firewall
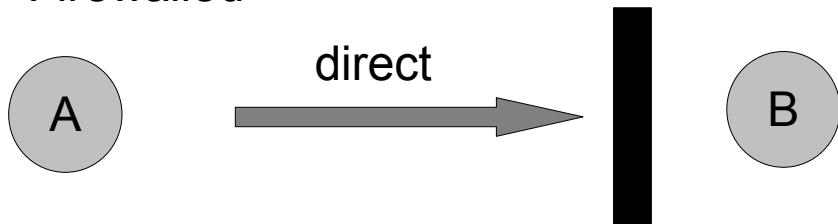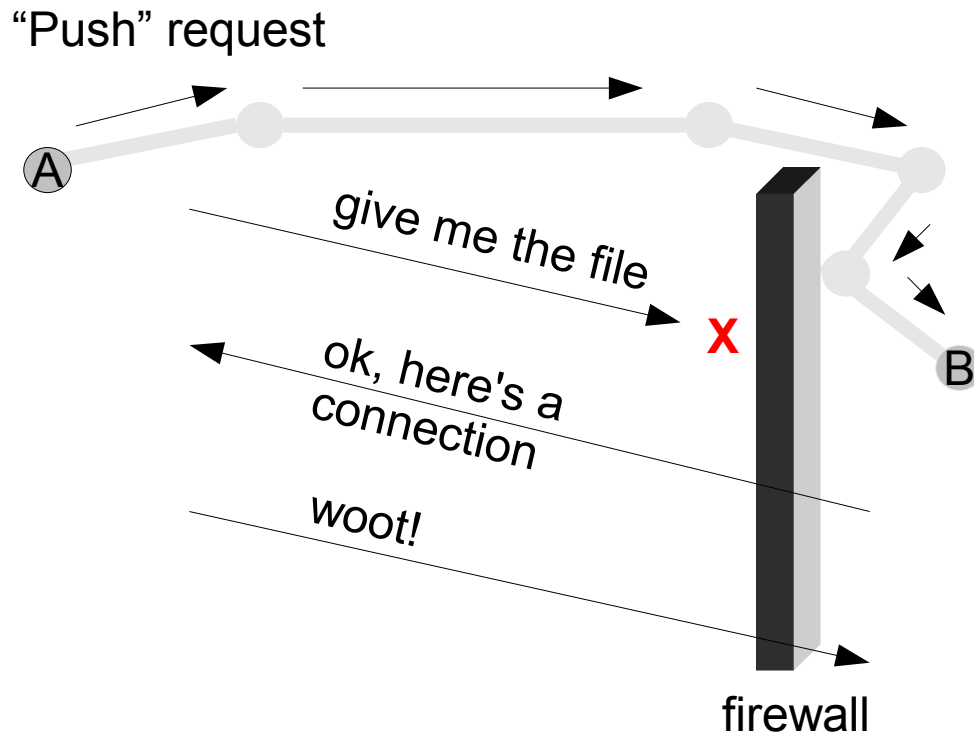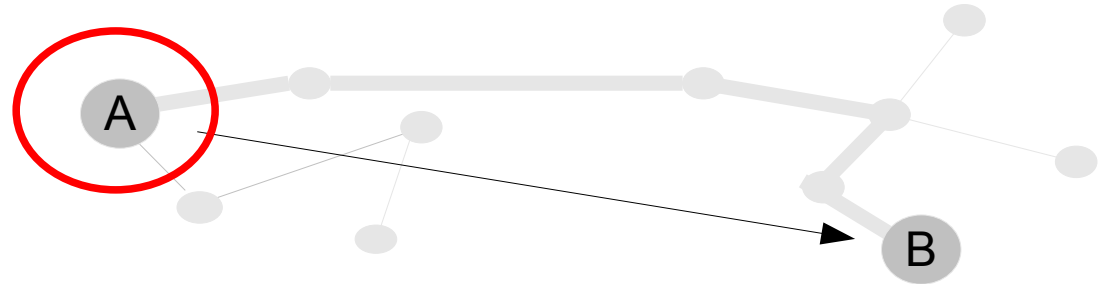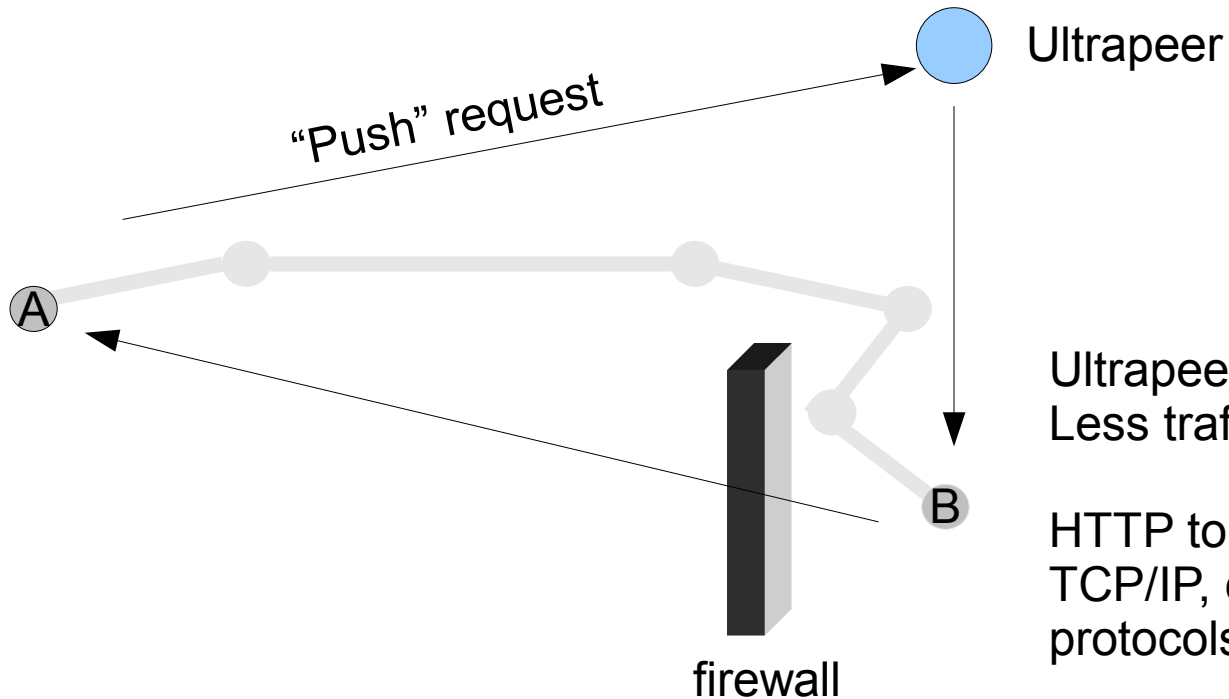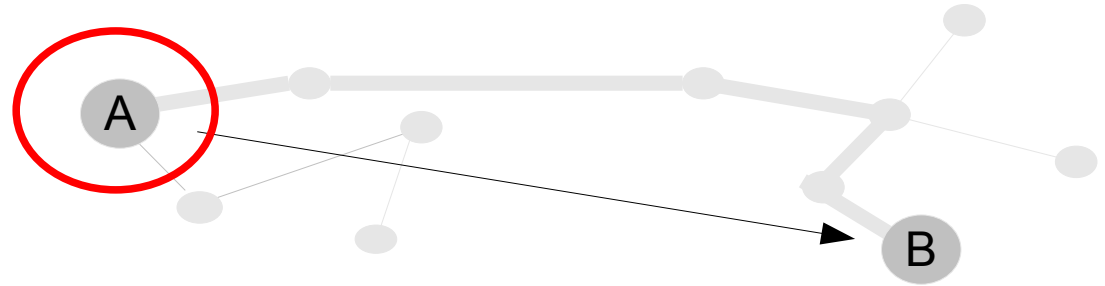
Not reliable method, as nodes on the "push" request route might brake – nodes come and go.
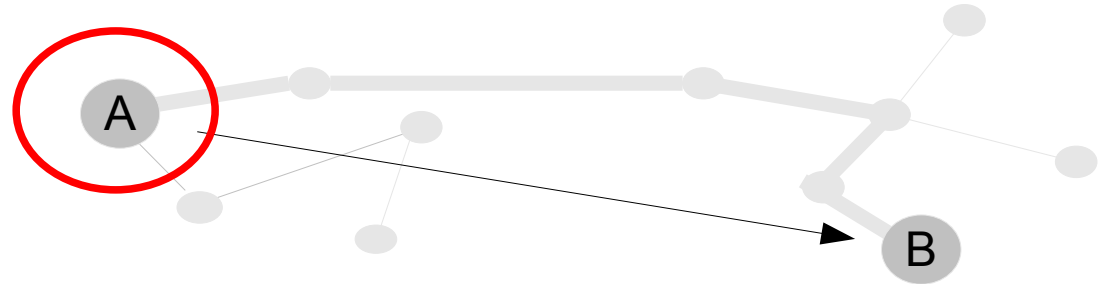
Gnutella download(3)

A

B
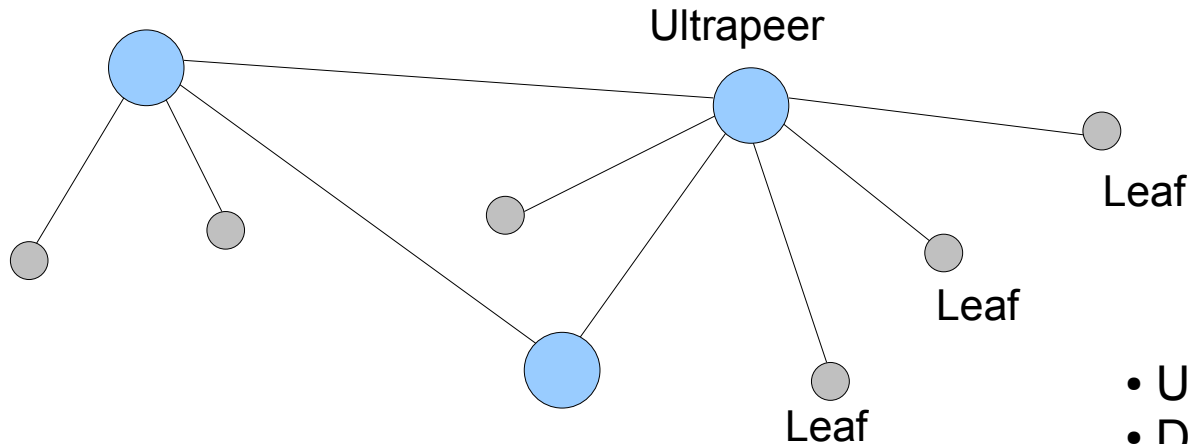
Ultrapeer

"Push" request

A

firewall

B

Ultrapeer-leef conn. are more stable
Less traffic on the network

HTTP to download files – ne need
TCP/IP, can work on top of other
protocols.

Gnutella network improvements

New network structure

Ultrapeer

Leaf

Leaf

Leaf

Leaf

- Ultrapeers responsible for routing
- Deeper search
- QRP (Query Routing Protocol)
- DQ (Dynamic Querying)
- Less traffic

Gnutella 2

- Design and Protocol features

  - Leaves and hubs (connect hundreds of leaves)
  - UDP with reliability flag
  - Search using walk system – connect one hub at a time
  - Extensible binary packet format
  - Meta data for searching
  - Compression