Dynamic File Bundling for Large-scale Content Distribution

Song Zhang	University of Calgary
Niklas Carlsson	**Linköping University**
Derek Eager	University of Saskatchewan
Zongpeng Li	University of Calgary
Anirban Mahanti	NICTA

October 26, 2012

LiU expanding reality

Motivation

- Use of Internet for content delivery is massive …
- Content delivery can consume significant resources
  - How to make scalable and efficient?
- Consider content provider wanting to server catalogue of many contents

Motivation

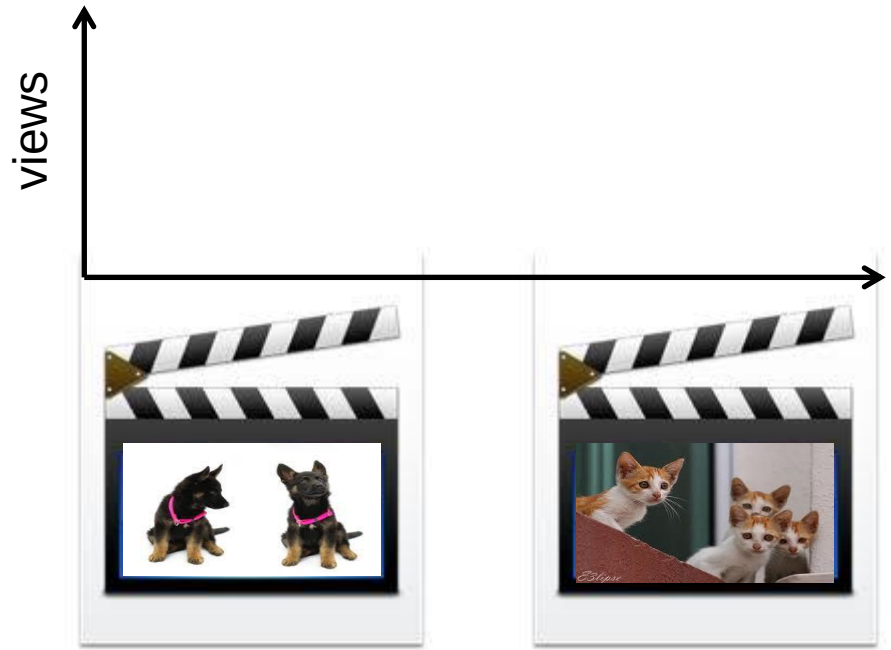
Motivation

- Not all videos will reach the same popularity and have the same impact

Motivation

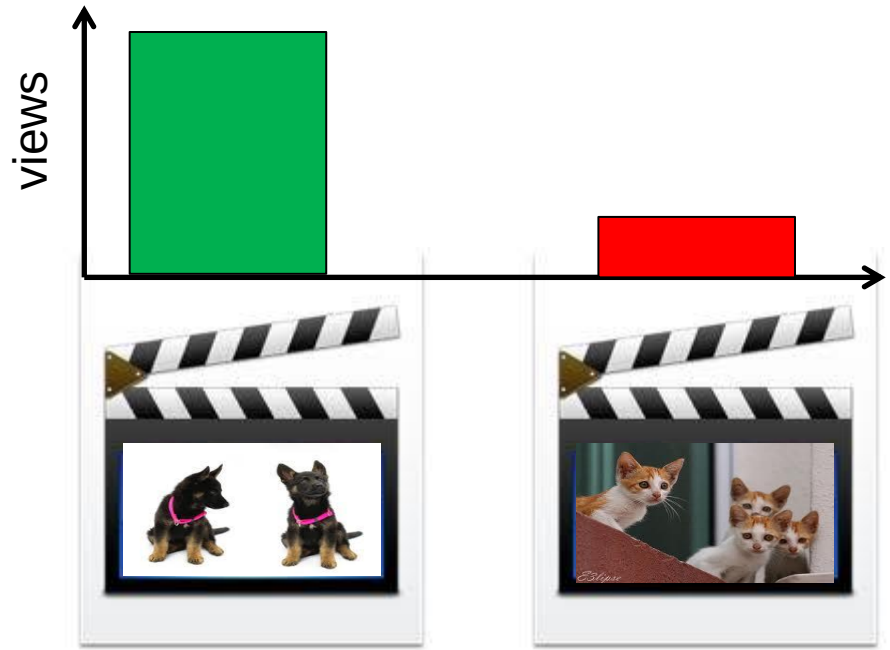
- Not all videos will reach the same popularity and have the same impact

Motivation



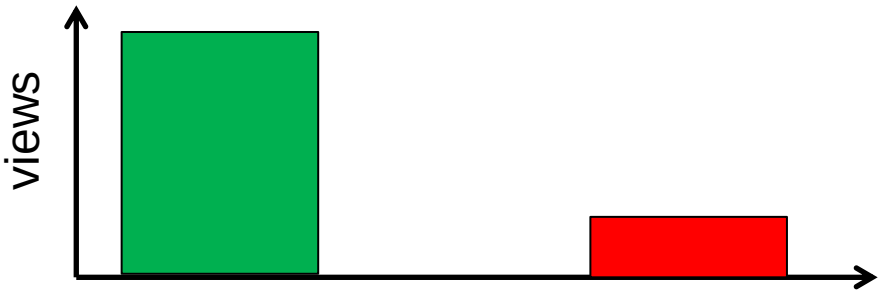
- Not all videos will reach the same popularity and have the same impact

Motivation

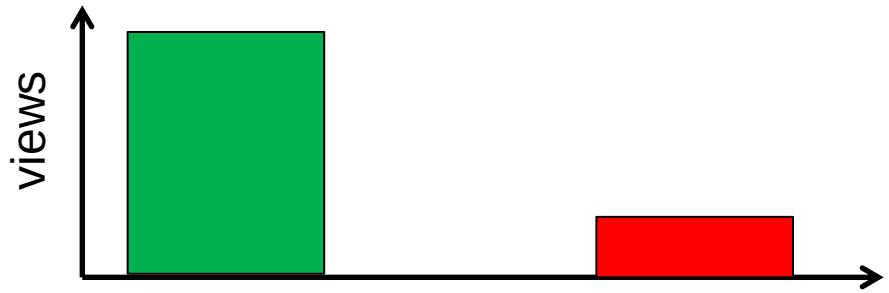


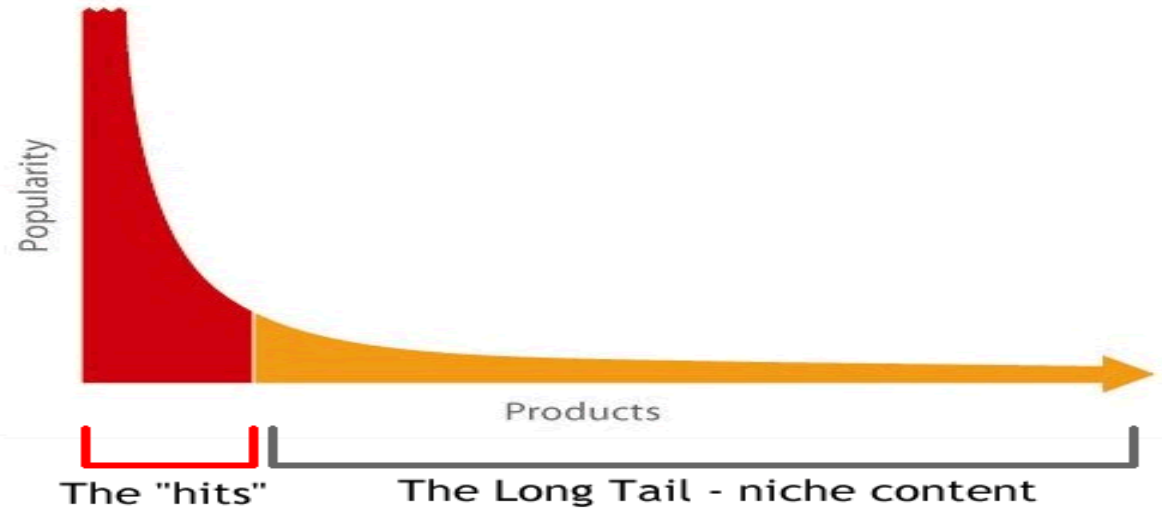
- Not all videos will reach the same popularity and have the same impact

Motivation



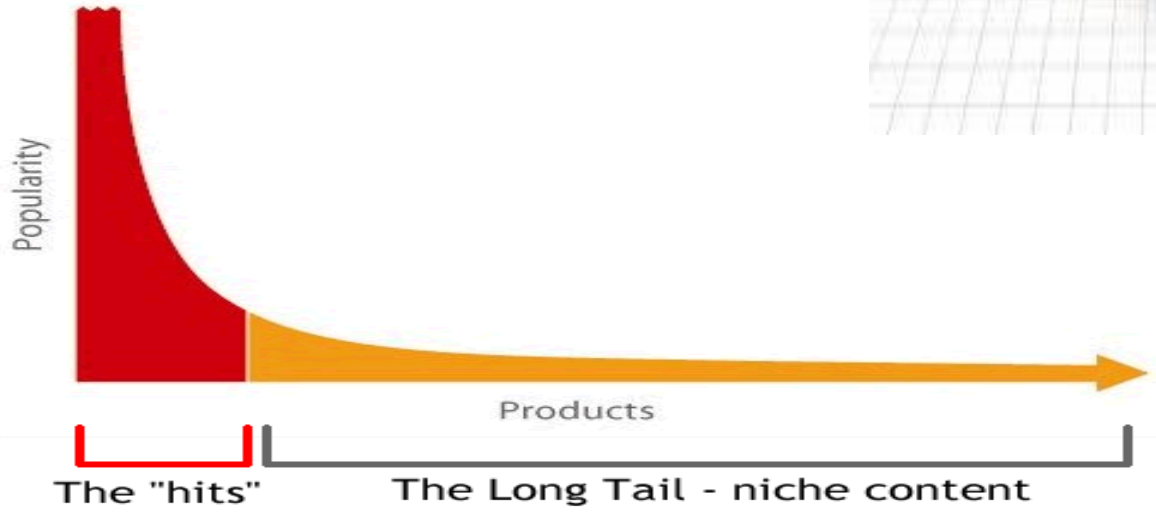
Motivation


views


Popularity
Products
The "hits"
The Long Tail - niche content

Motivation

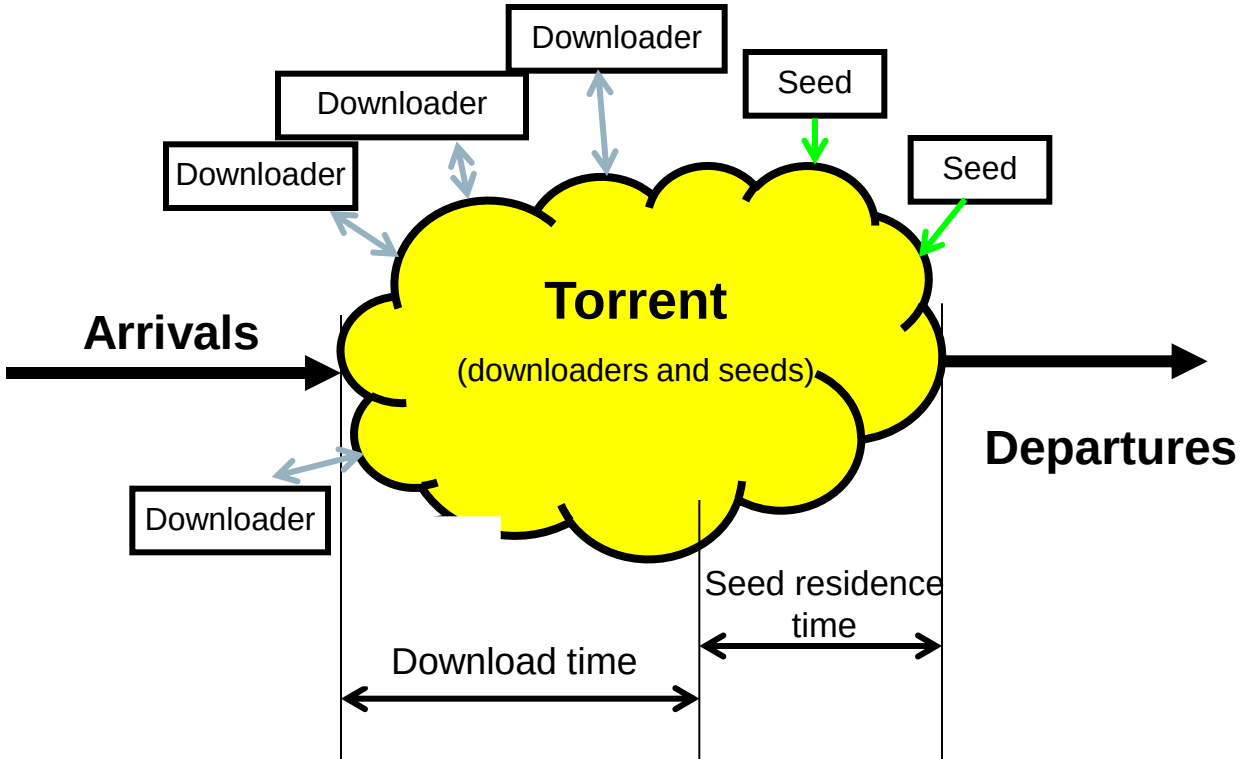
- Cost-efficient solution must scale with regards to both:
  - Request rate
  - Number of availible contents



Background: BitTorrent

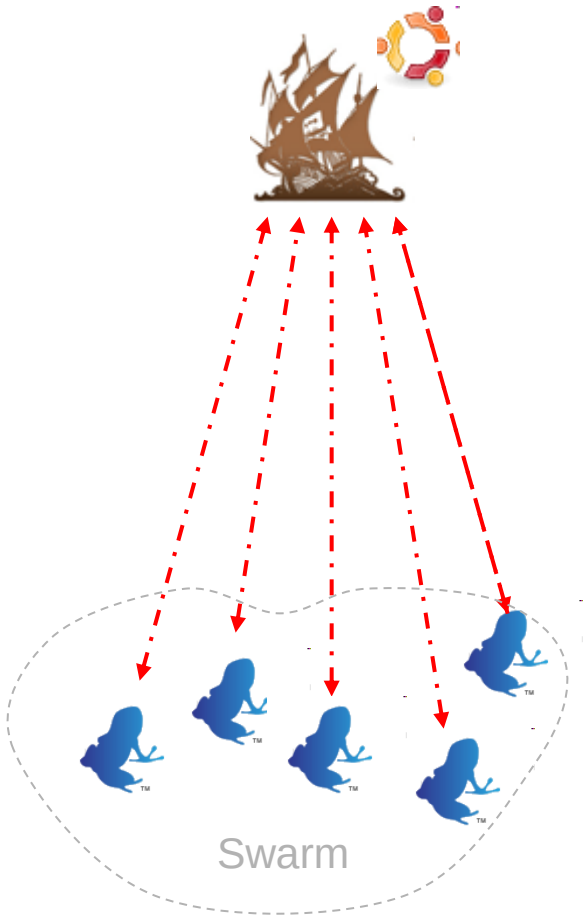
Single file download

- File split into many smaller chunks
- Downloaded from both seeds and downloaders
- Distribution paths are dynamically determined
  - Based on data availability



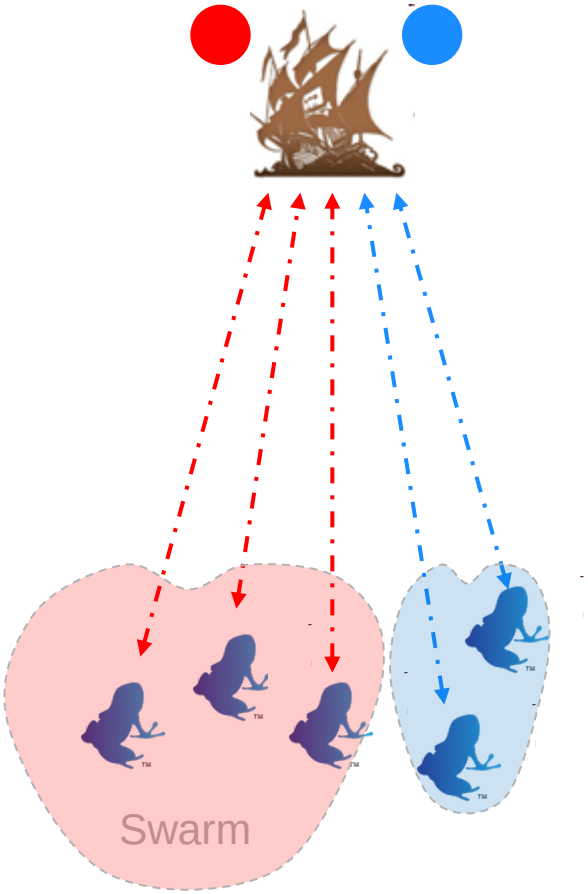
Background: BitTorrent

- Torrent file
  - "announce-list" URLs
- Trackers
  - Register torrent file
  - Maintain state information
- Peers
  - Obtain torrent file
  - Announce
  - Report status
  - Peer exchange (PEX)

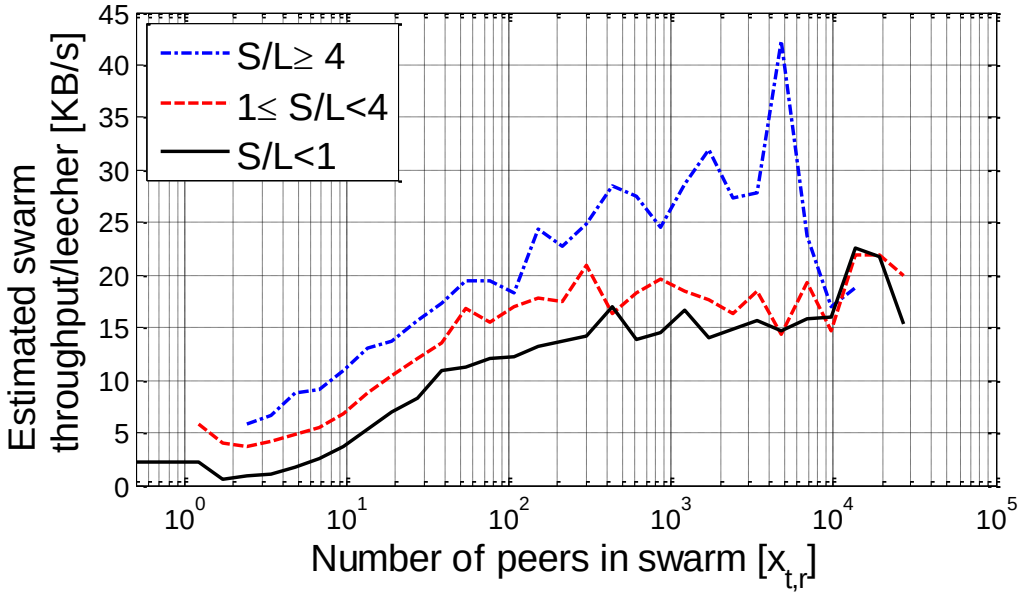


Background: BitTorrent

- Torrent file
  - "announce-list" URLs
- Trackers
  - Register torrent file
  - Maintain state information
- Peers
  - Obtain torrent file
  - Announce
  - Report status
  - Peer exchange (PEX)



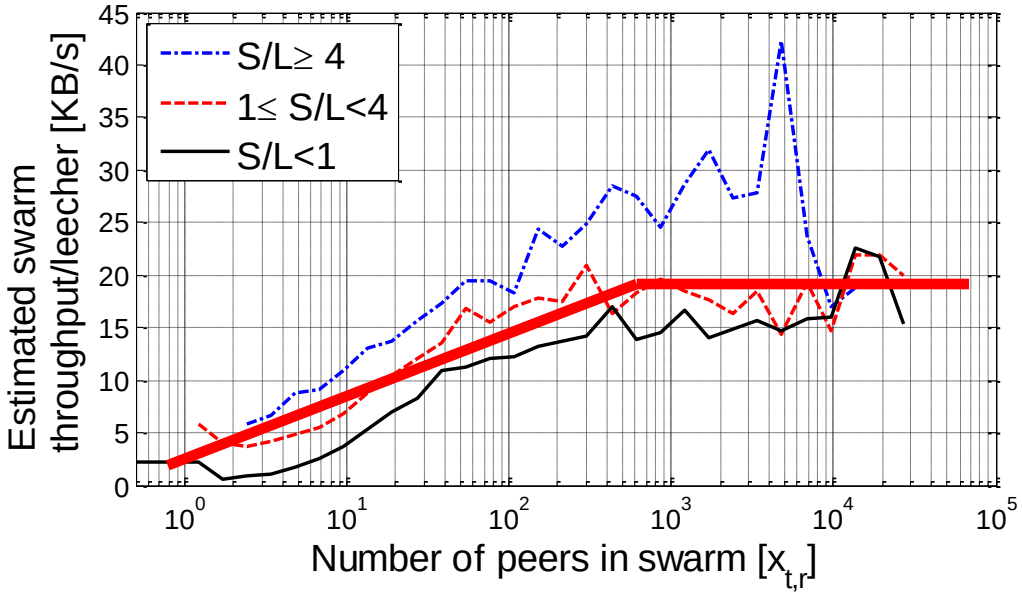
Example 1: Download rate in BitTorrent



- Download rate

Analysis and results presented in IPTPS 2009 and IEEE/ACM ToN

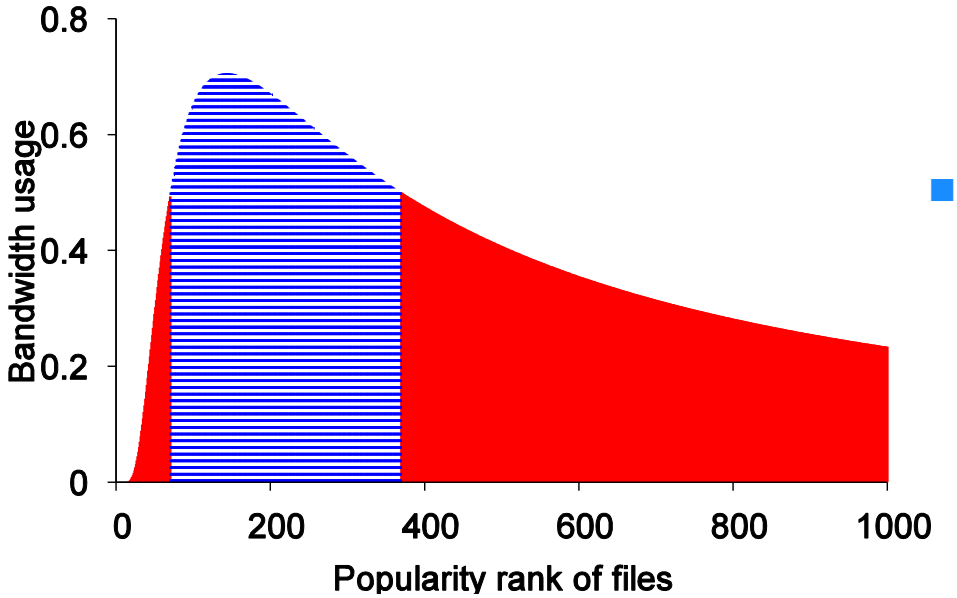
Example 1: Download rate in BitTorrent

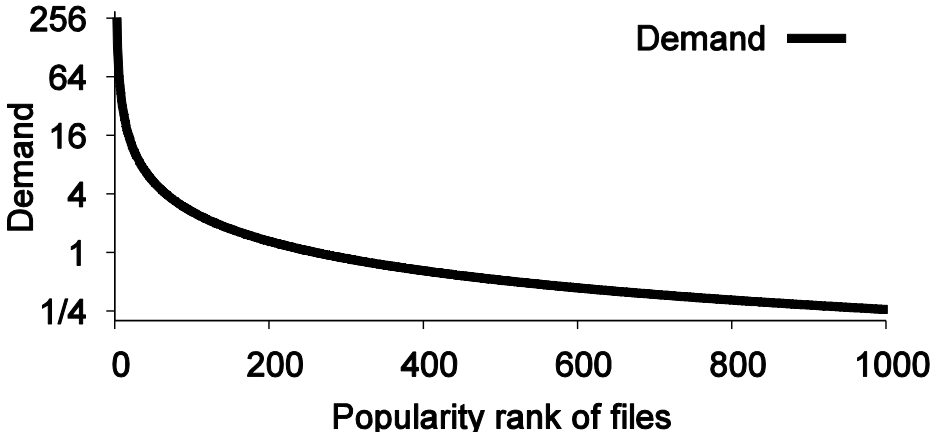


- Download rate

Analysis and results presented in IPTPS 2009 and IEEE/ACM ToN

Example 2: Peer-assisted system with service guarantees



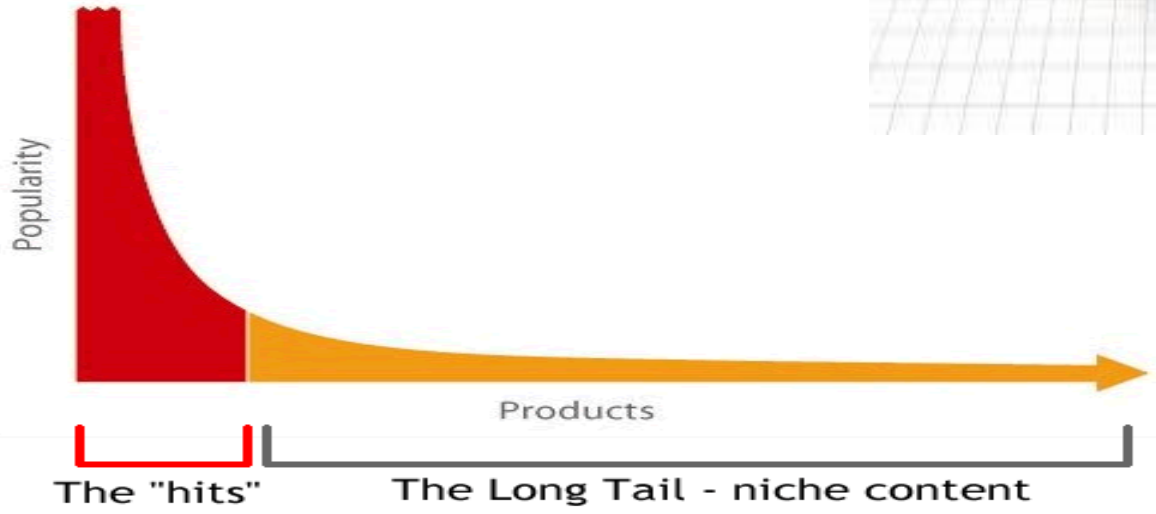


- Server b/w usage

Analysis and results presented in IEEE P2P 2012

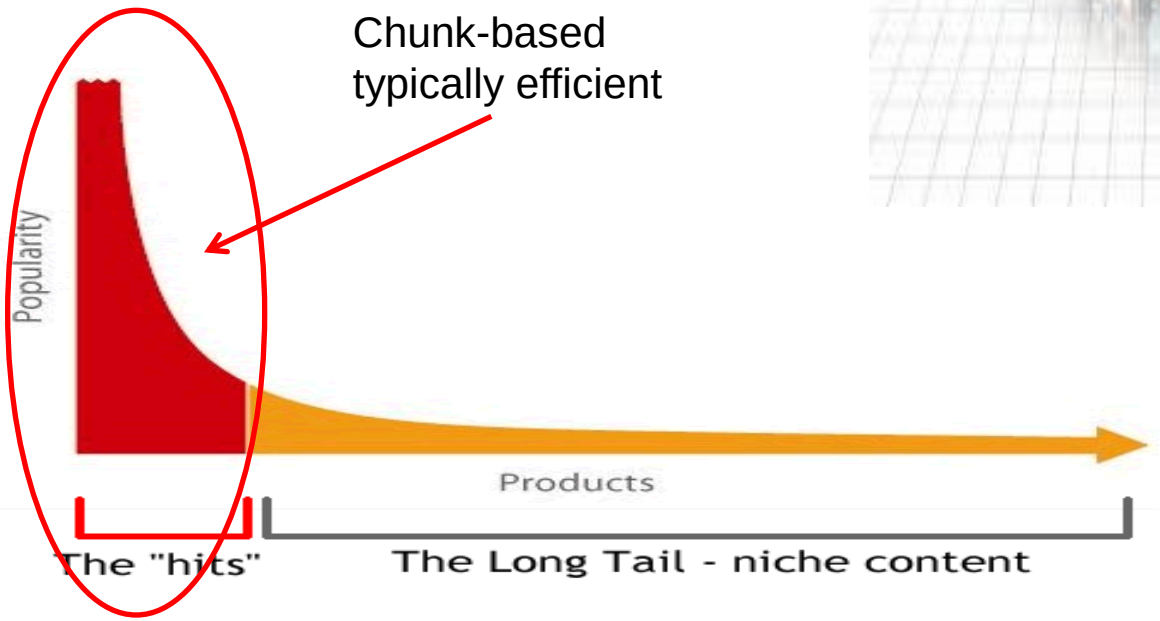
Motivation

- Cost-efficient solution must scale with regards to both:
  - Request rate
  - Number of availible contents



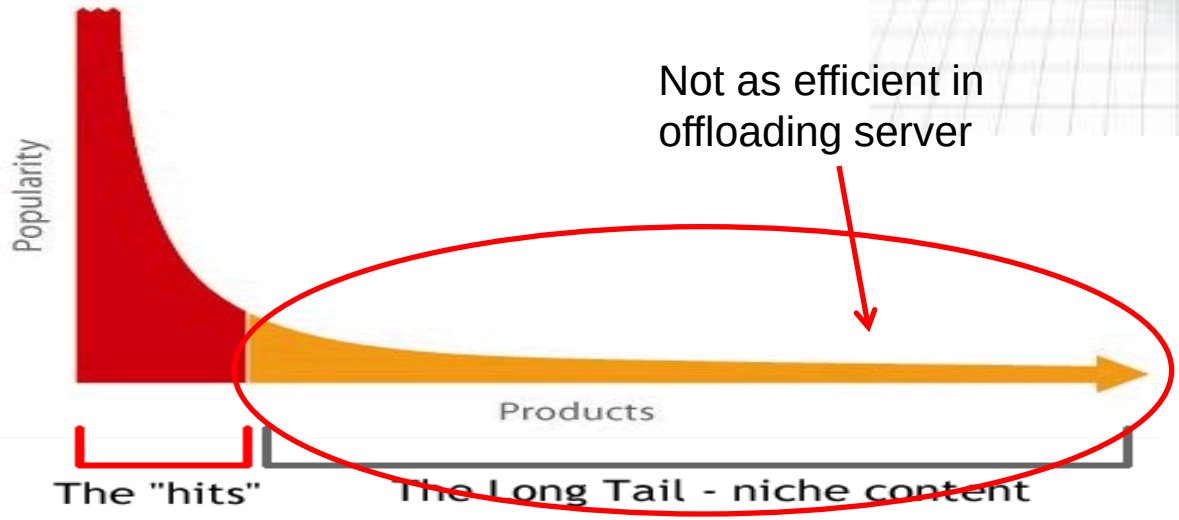
Motivation

- Cost-efficient solution must scale with regards to both:
  - Request rate
  - Number of availible contents



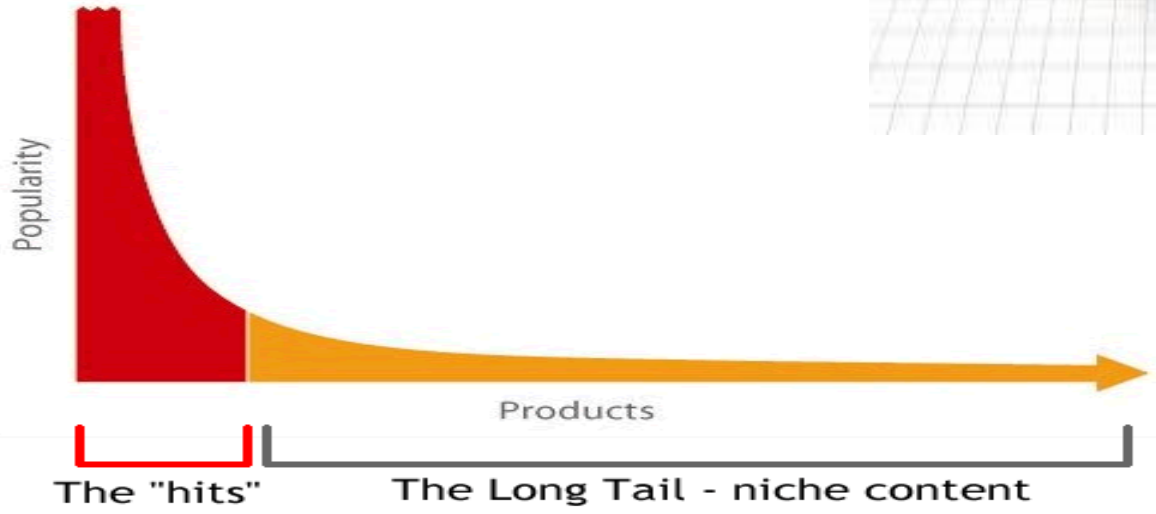
Motivation

- Cost-efficient solution must scale with regards to both:
  - Request rate
  - Number of availible contents



Motivation

- Cost-efficient solution must scale with regards to both:
  - Request rate
  - Number of availible contents



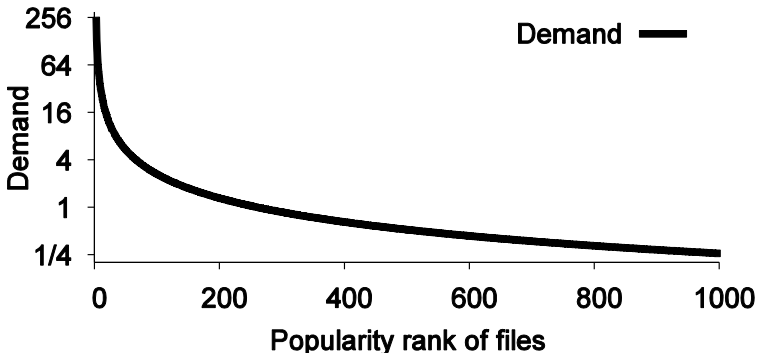
Bundling (or torrent inflation)

Using some of the available upload bandwidth from currently downloading peers to “inflate” torrents for files that would otherwise require substantial server b/w

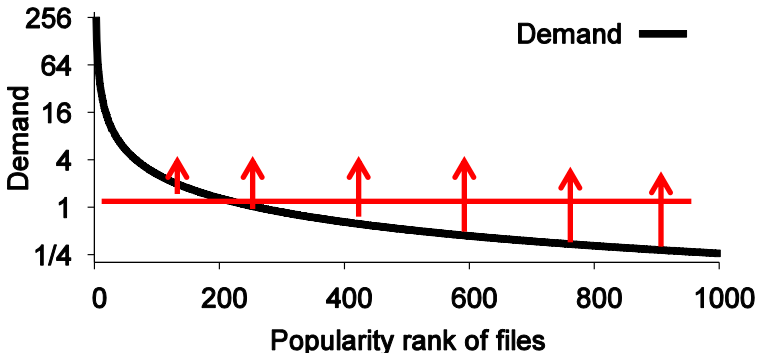
How to pick which files to bundle/inflate?

- Static bundling
  - Bundle at content creation
- Dynamic bundling
  - Adapt which files to bundle based on current popularities and conditions

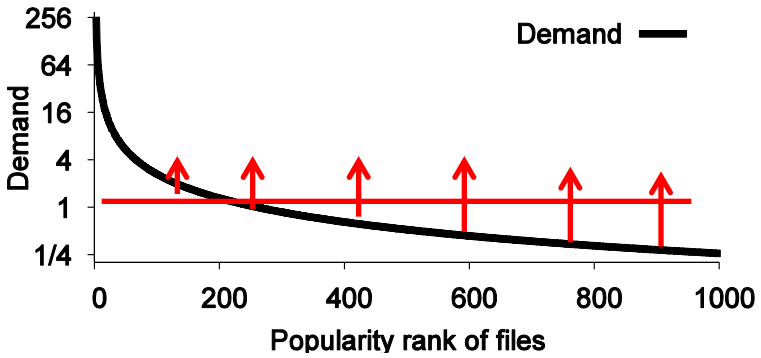
Dynamic bundling policies

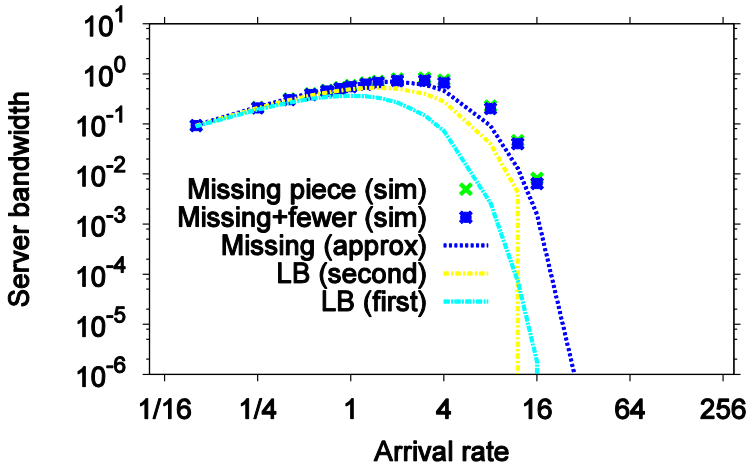


Dynamic bundling policies

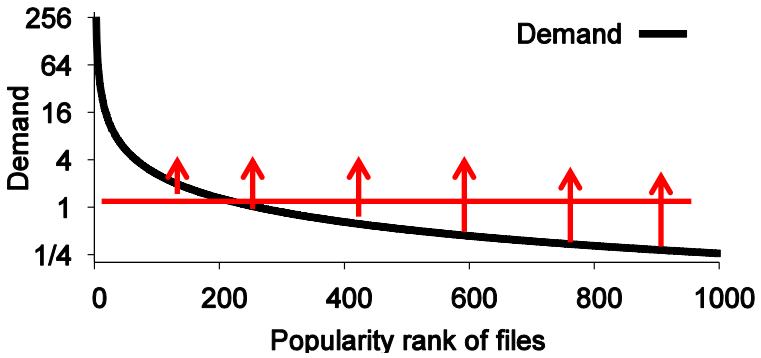


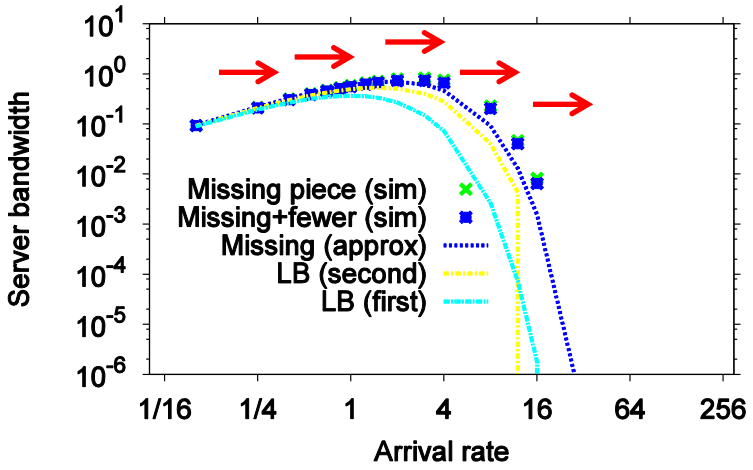
Dynamic bundling policies



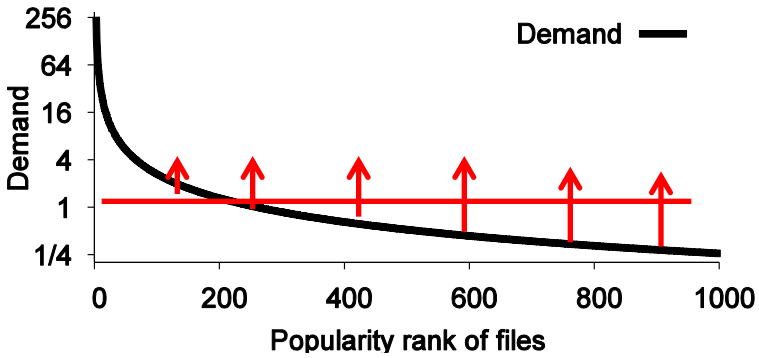


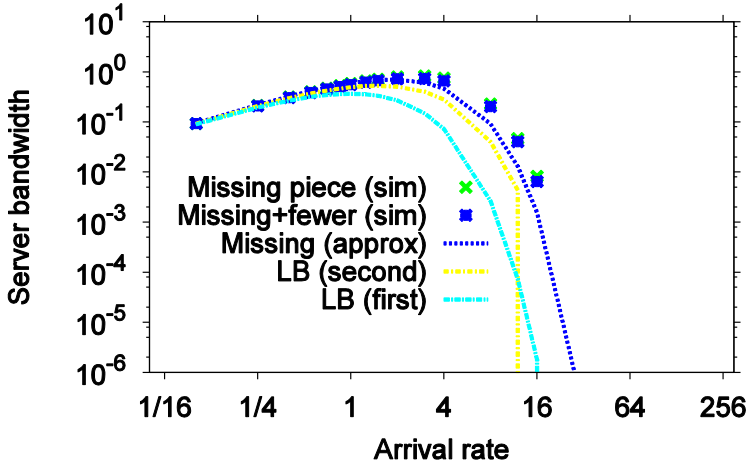
Dynamic bundling policies

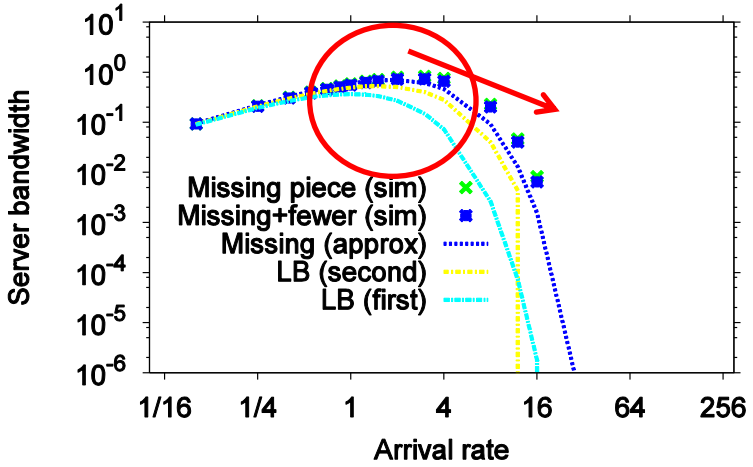




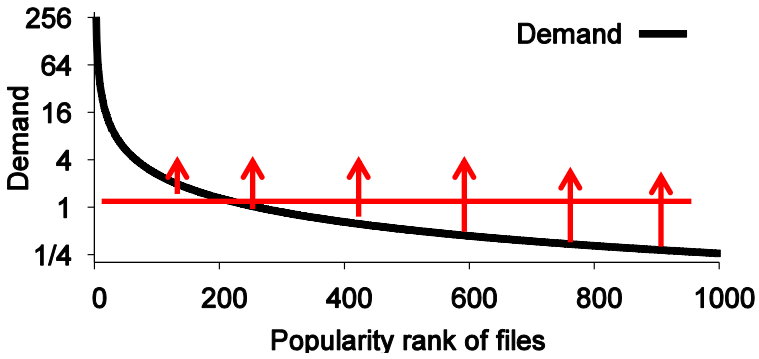
Dynamic bundling policies

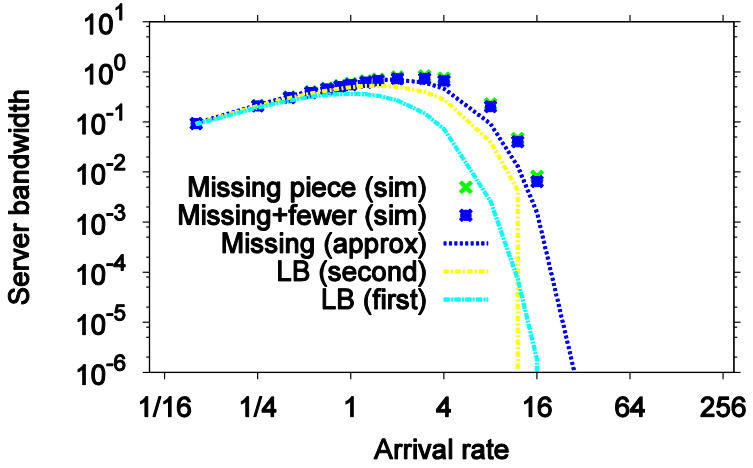


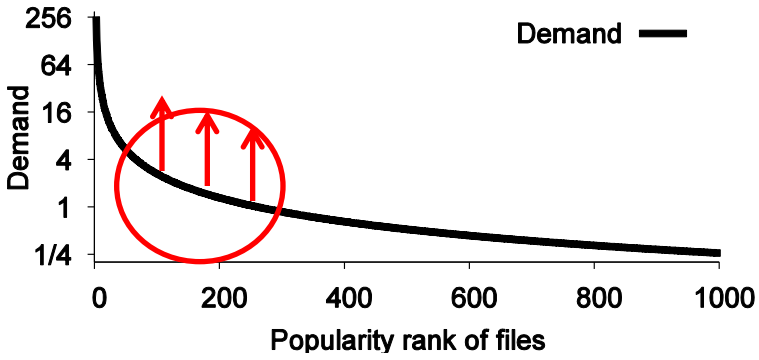


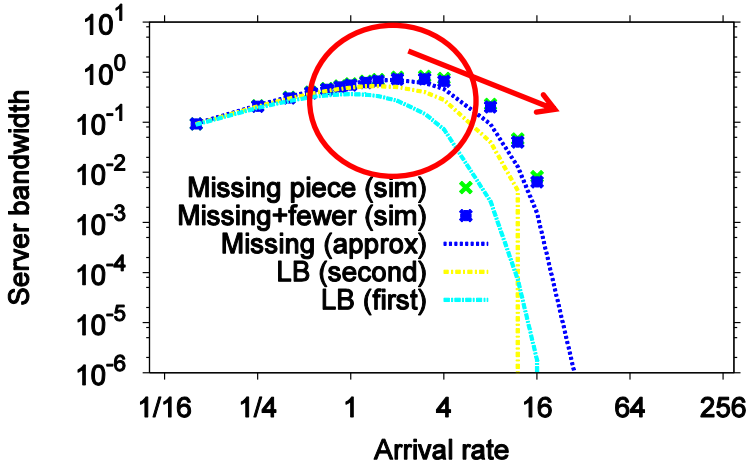


Dynamic bundling policies

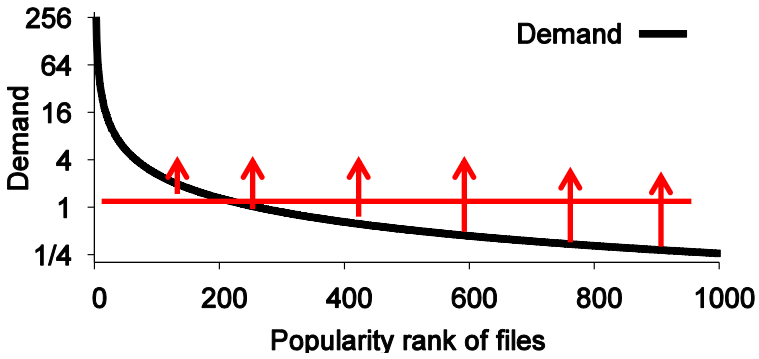


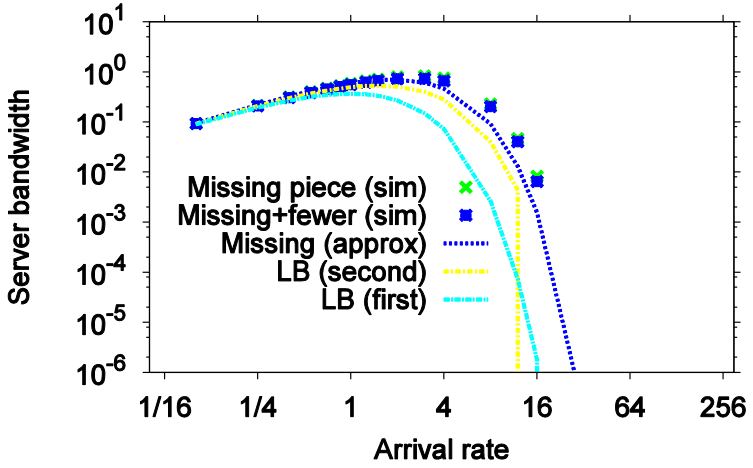


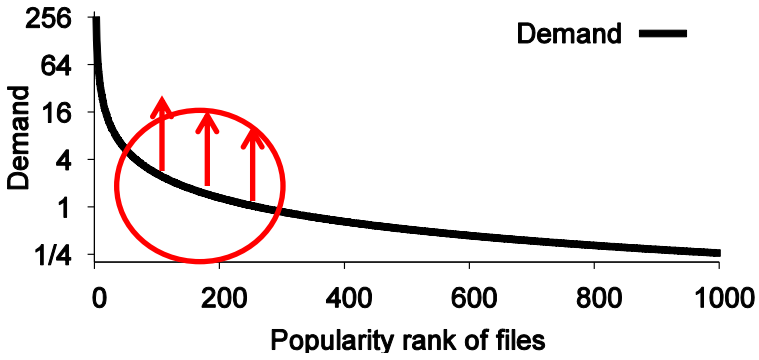


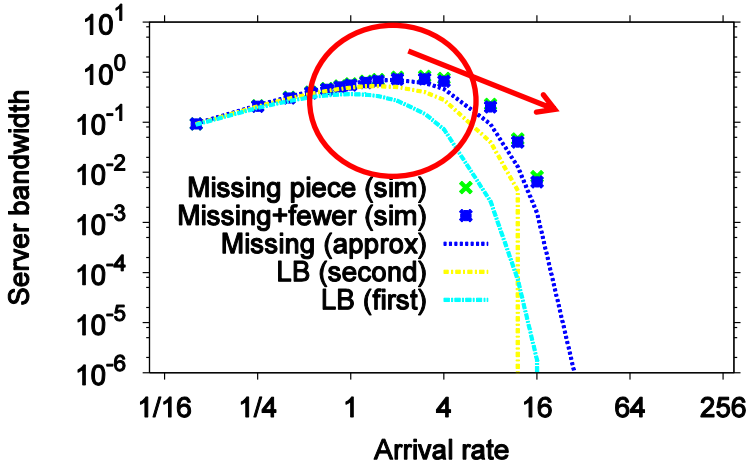


Dynamic bundling policies

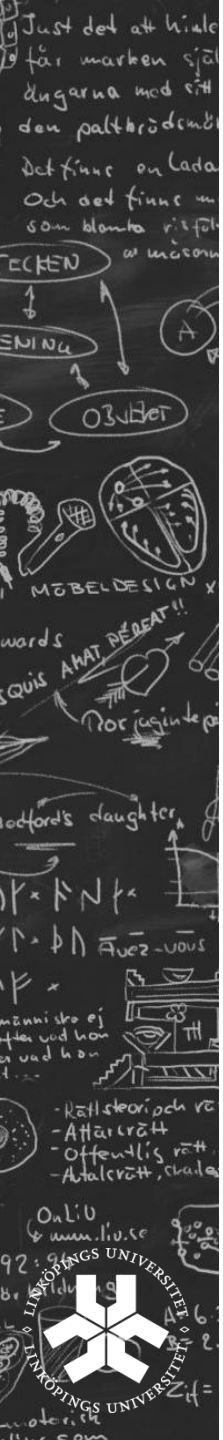








Inflate lukewarm files (e..g, IEEE P2P 2012; Networking 2010)

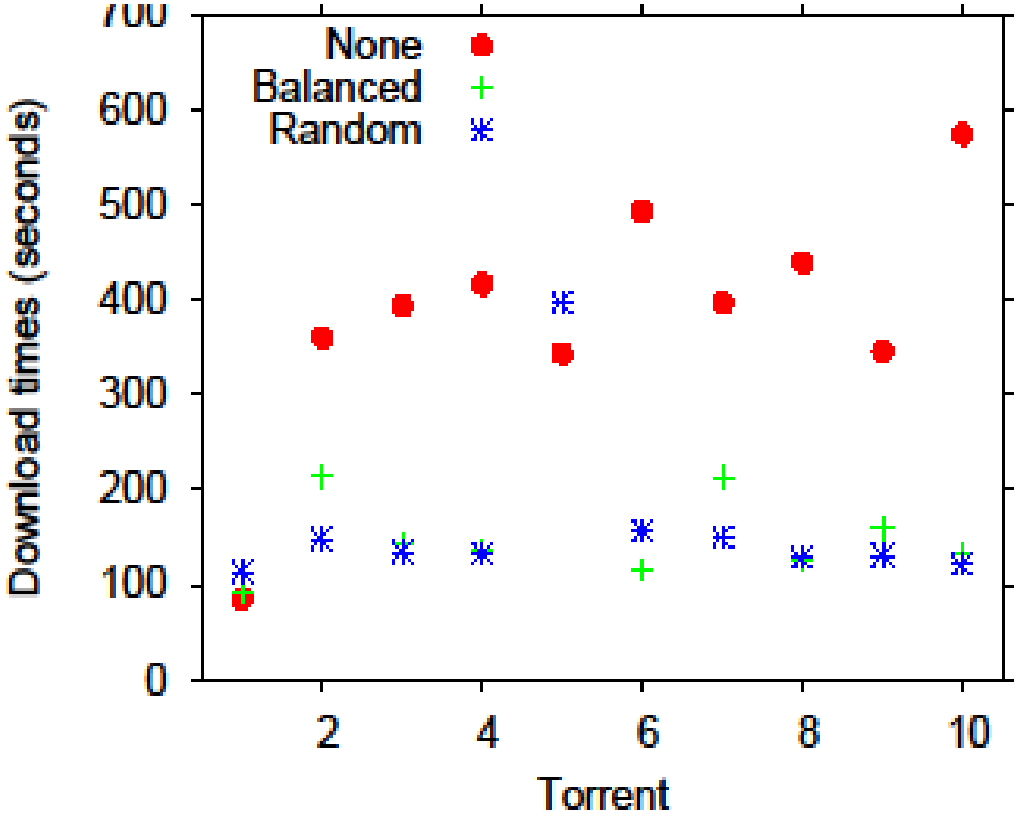


Proof of concept implementation

- Instead of looking at policy choices ...
- This paper presents the design and proof of concept implementation of a dynamic bundling system
  - Want to minimize changes to the BitTorrent protocol
  - Introduce concept of “super bundle” and “individual bundles”
  - Performance evaluation **using simple baseline policies** on PlanetLab

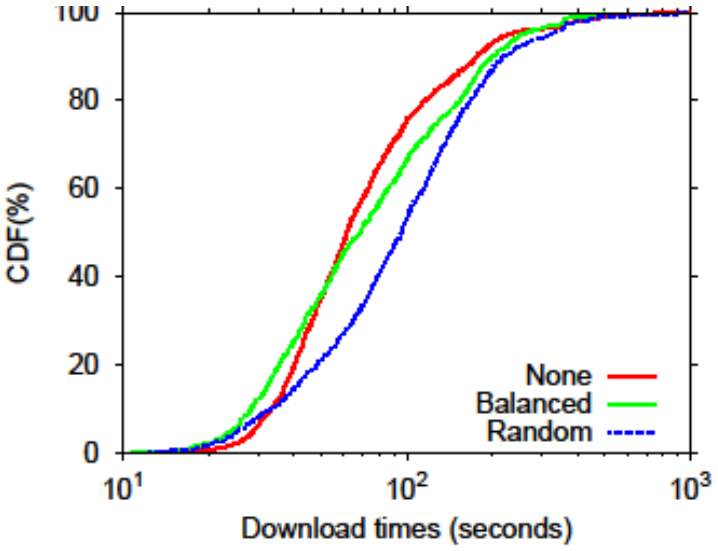
- Four parts
  - Tracker communication (bundle negotiation event and additional information)
  - Bundling selection component at tracker
  - Piece selection rule (as each peer have individual preferences and responsibilities)
  - Disk writing algorithm

Example results: Steady-state

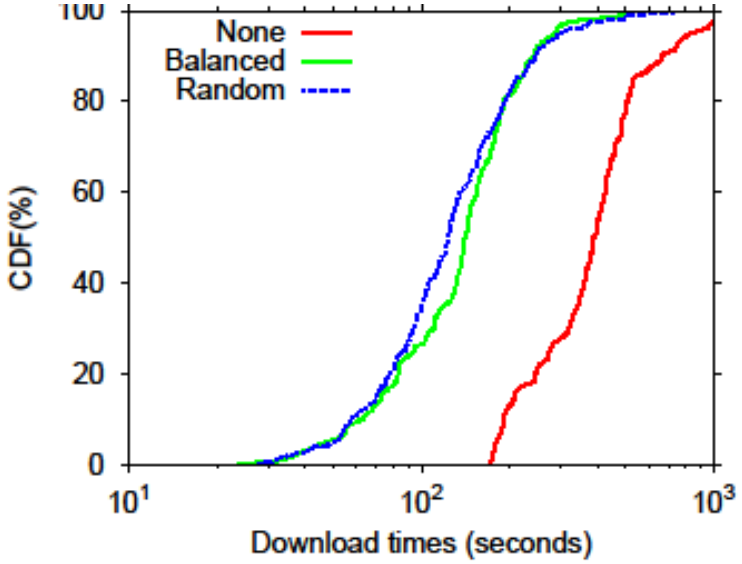


- Average download times
  - Again, naïve baseline policies
  - Still performance improvements ...
    - Big improvements for less popular files, at small cost at popular file

Example results: Steady-state



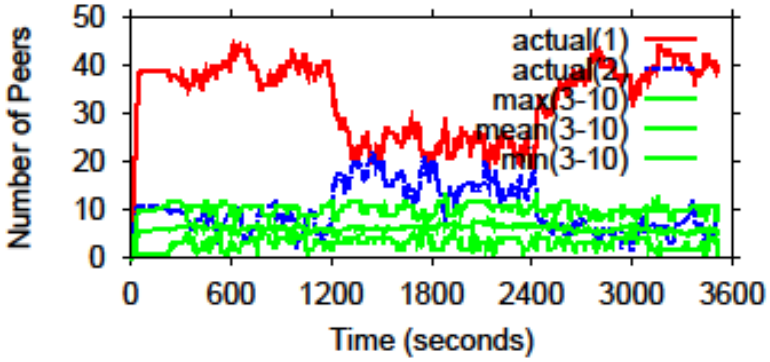
(a) Popular file

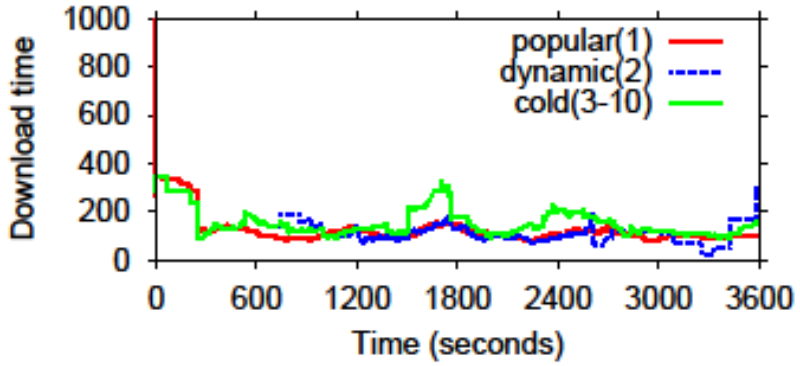


(b) Less popular file

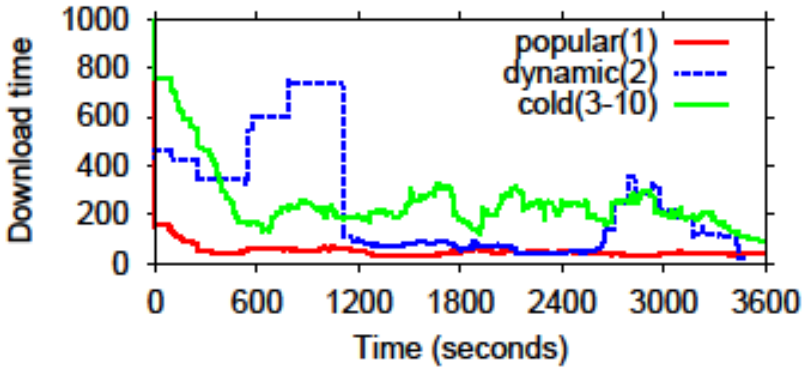
- Download time distributions
  - Again, naïve baseline policies
  - Still performance improvements ...
    - Big improvements for less popular files, at small cost at popular file

Example results: Dynamic scenario





- Smaller download times
- Adapt to current conditions



Contributions

- Presented the design of a dynamic bundling system, including a proof of concept implementation
  - Allow flexible adaptive policies to be implemented and tested with relatively small changes to the BitTorrent protocol
- Some example results are presented for naïve baseline policies
- Future work include
  - Design of a fully dynamic bundling system
  - PlanetLab testing of more advanced policies (as proposed/evaluated in other works; e.g., IFIP Networking 2010 and IEEE P2P 2012)

Thank you!

- Song Zhang — University of Calgary
- **Niklas Carlsson** — **Linköping University**
- Derek Eager — University of Saskatchewan
- Zongpeng Li — University of Calgary
- Anirban Mahanti — NICTA