Monitoring BitTorrent swarms

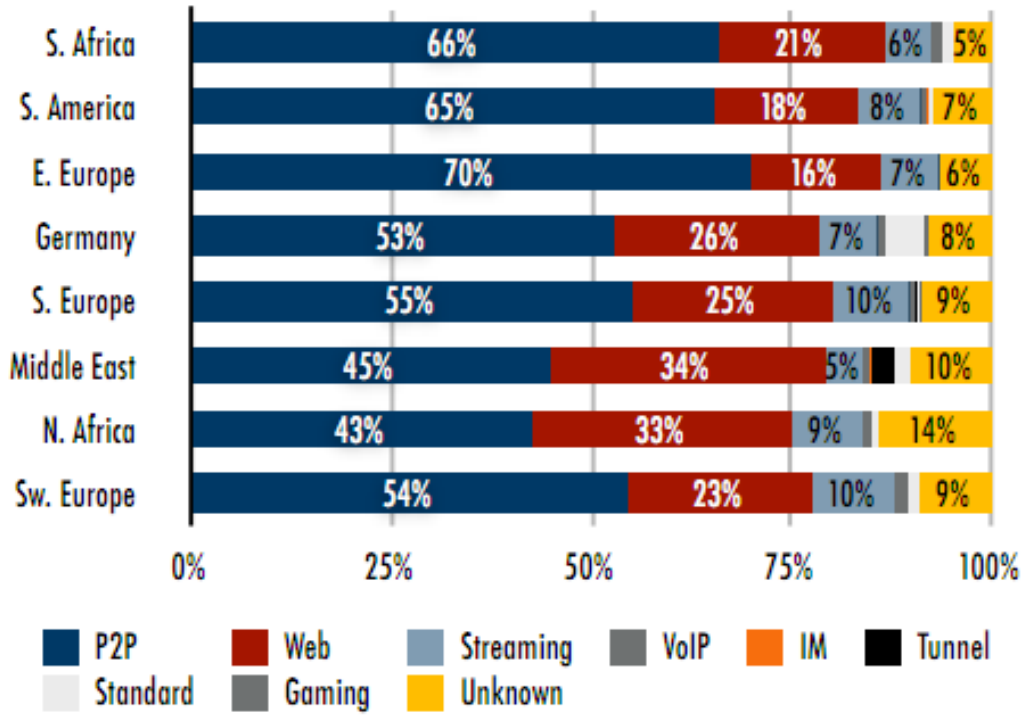
António Homem Ferreira
Student nº 55935
MERC

INESC-ID, Communications Networks and Mobility
Instituto Superior Técnico - Campus TagusPark

Outline

Motivation

Increasing usage of P2P communications
BitTorrent accounts for 60% of worldwide P2P traffic



Motivation

- P2P protocols generate a great amount of traffic
- Affects quality of service of other protocols
- P2P communications results in high inter-ISP traffic
- Inter-ISP traffic is expensive for ISPs

Work Objectives

- Study monitoring real Internet BitTorrent swarms
- Data gathered used to identify patterns in BitTorrent networks:
  - Locality - decrease inter-ISP traffic
  - Content - increase content availability

BitTorrent Protocol

File-sharing protocol

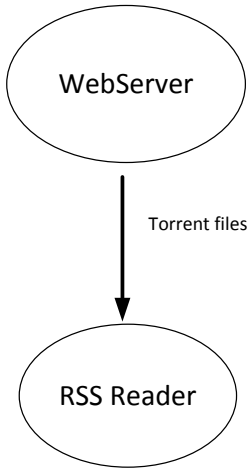
- Peers share files among themselves
- Files are divided into smaller pieces
- Central Server (Tracker) which provides a list of active peers
- Mechanisms that make peers chose to download from the ones with the largest bandwidth

Contributions

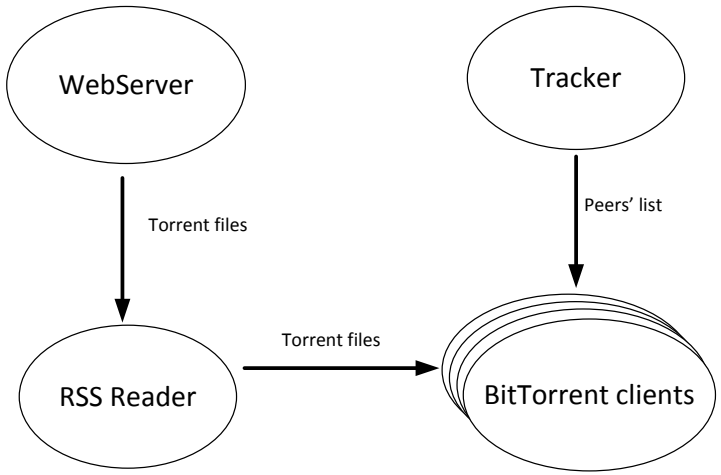
This study differs in:

- Locality studied for two different periods
- Content study that focuses on repeated content
- PSM - an approach to increasing content availability
- Peer and Tracker behaviors that can affect locality

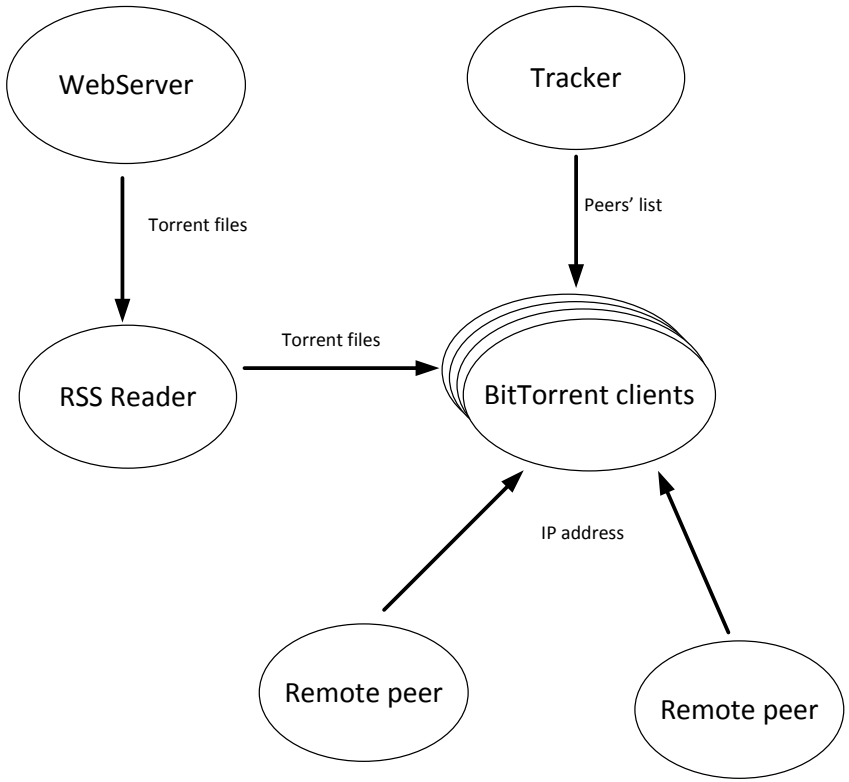
Methodology for gathering data



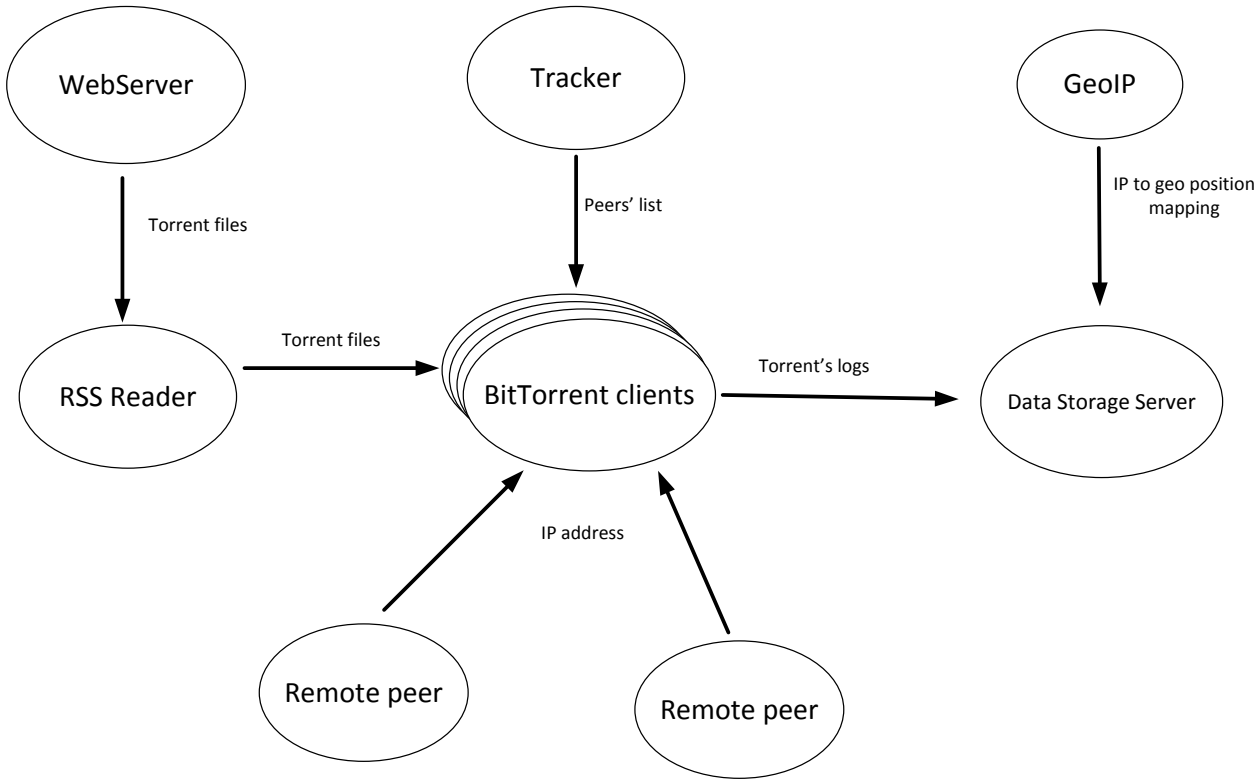
Methodology for gathering data



Methodology for gathering data



Methodology for gathering data



Methodology for analyzing data

The torrent database:

- 3211 torrent files from PirateBay, isohunt and btjunkie
- Torrent files collected from the 25^{th} of April 2011 to the 18^{th} of July 2011
- PirateBay's one hundred most popular torrent files on the 25^{th} of April
- isohunt's twenty most popular torrent files on the 24^{th} of May for each major category (audio, video, and tv)

Methodology for analyzing data

- Different data analysis criteria for each study:
  - Locality study
  - Content study
- Monitoring stopped for all swarms that dropped below 30 peers and 0 seeders for a period of a week
- Monitoring nodes not part of the results.

Results

- Locality Analysis
- Content Analysis
- Peer and tracker behavior

Locality Analysis

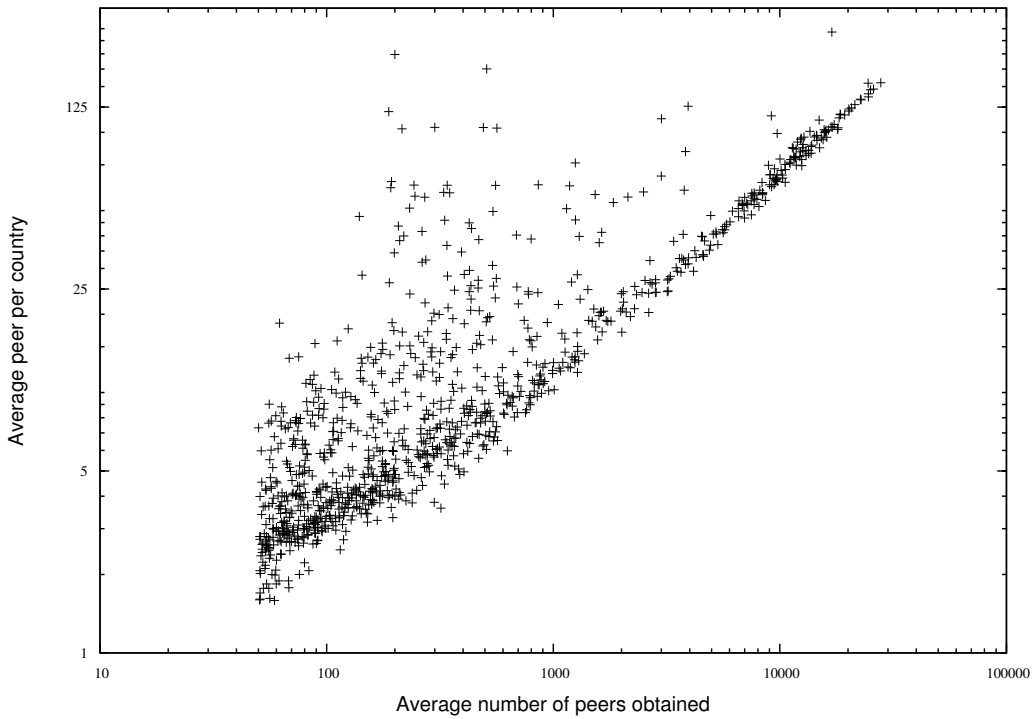


Figure: Average number of peers obtained and average number of peers per country.

Locality Analysis

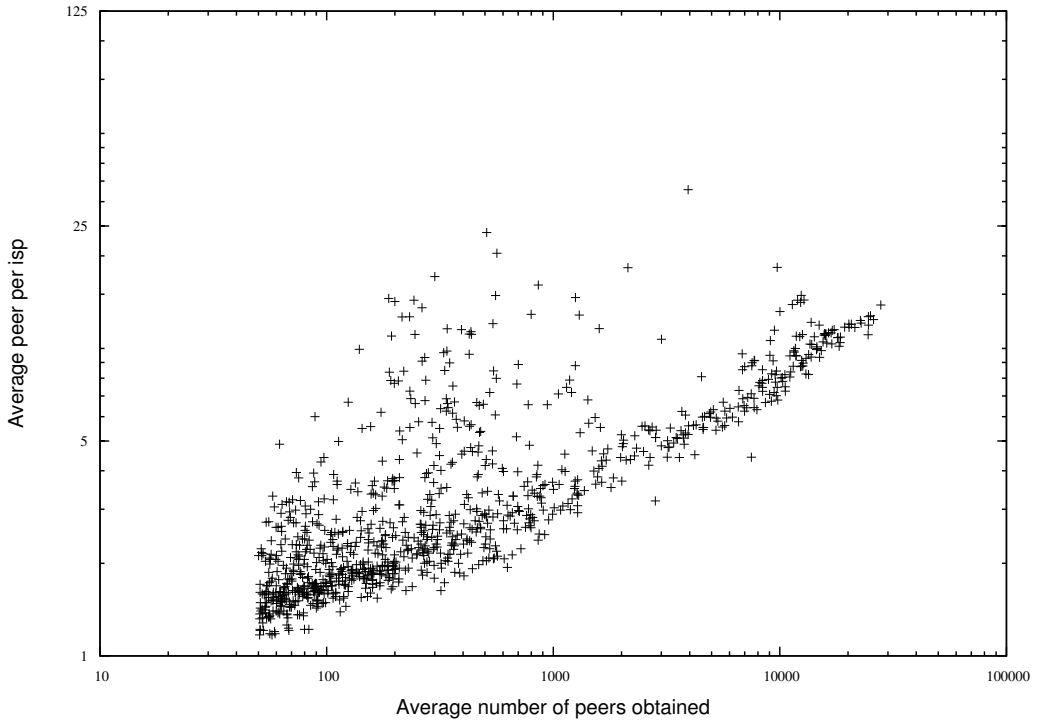


Figure: Average number of peers obtained and average number of peers per ISP.

Locality Analysis - Regional content

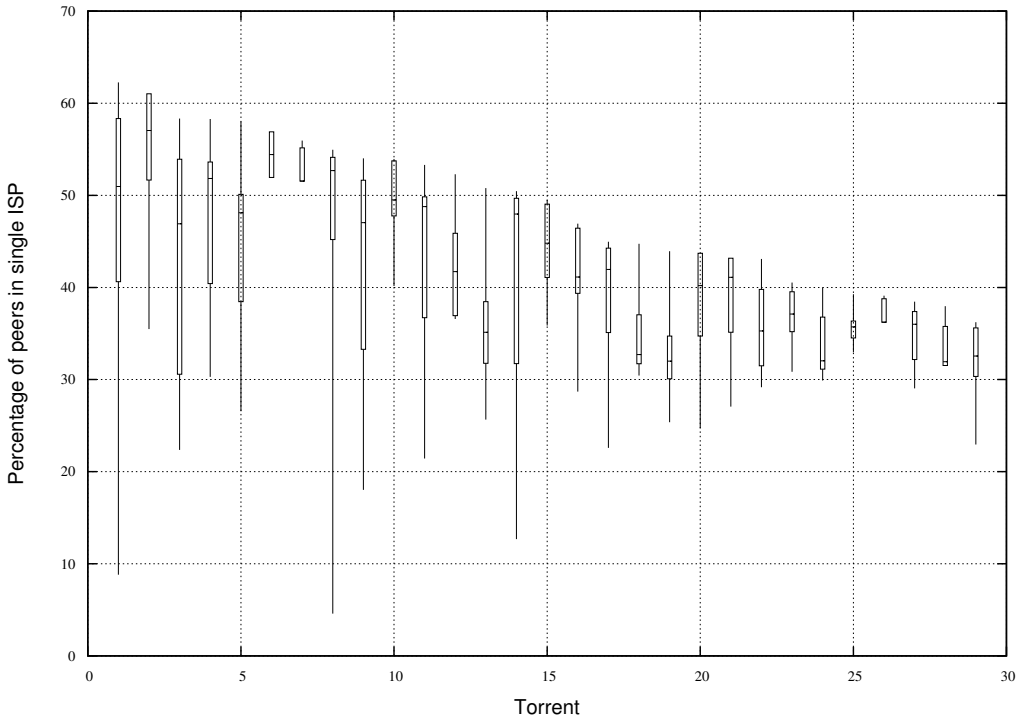


Figure: Regional torrents with 30% of all peers belonging to the same ISP, at least 75% of the times.

Content Analysis

- Content pollution
- Content repetition

Content Analysis - Content pollution

- Files with most parts the same
- Files did not represent any real content
- From the 3211 files collected, 221 were found to be polluted content

Content Analysis - Content repetition

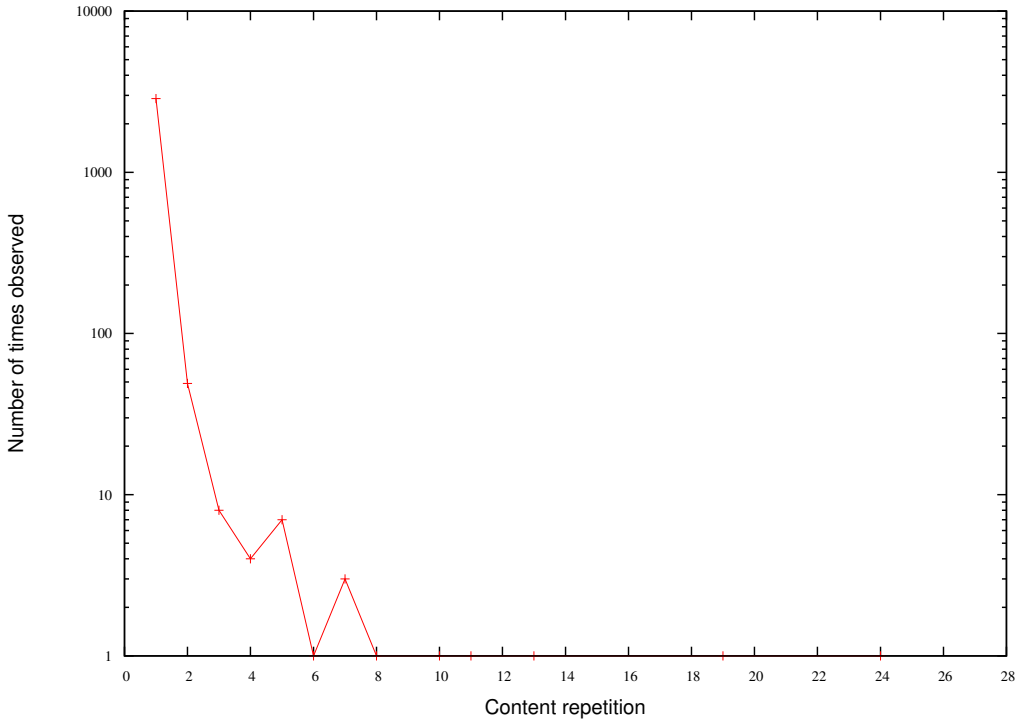


Figure: Histogram of the content repetition frequency.

Content Analysis - Content repetition

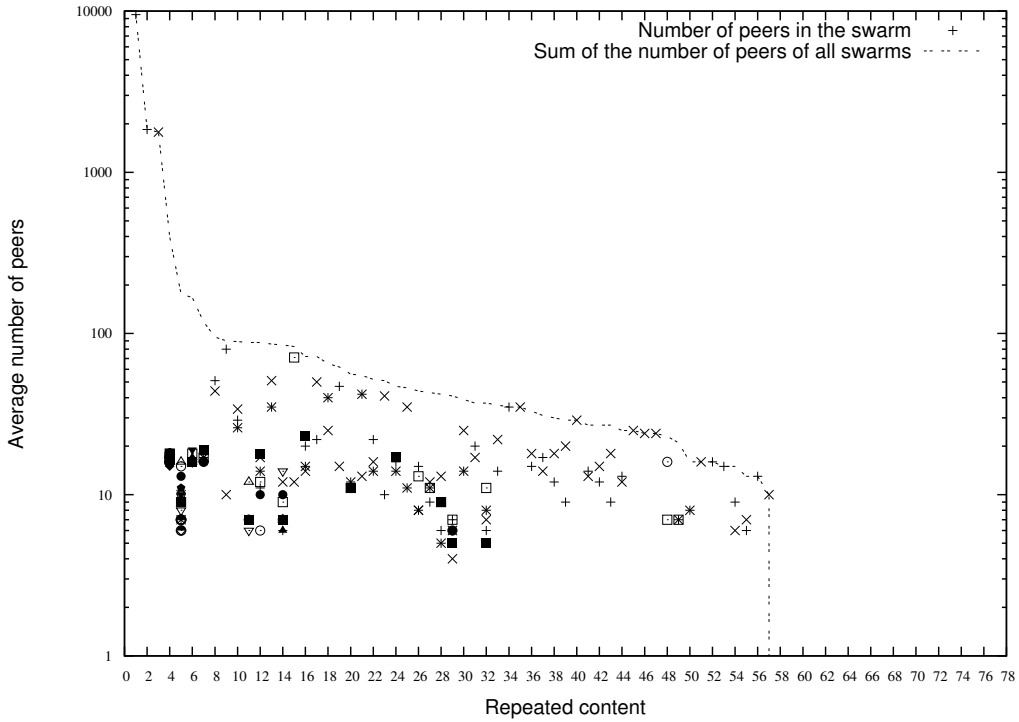


Figure: Average number of peers per content.

Content Analysis - Content repetition

	Torrent 1	Torrent 2	Torrent 3	Total
Num. Countries	87	68	38	101
Avg. Peers/Country	6.20	5.25	2.61	9.85
Num. ISPs	238	178	54	342
Avg. Peers/ISP	2.26	2.01	1.83	2.91

Table: Torrent aggregation benefits.

PSM

- Peers query for swarms that share similar content to the one they are currently downloading
- Peers participate in multiple swarms, isolated from each other
- Its databases are populated through the different queries received

PSM

- Service outside BitTorrent
- Requires no modifications to the BitTorrent protocol
- BitTorrent client application requires only an extension/add-on to make use of the service

PSM

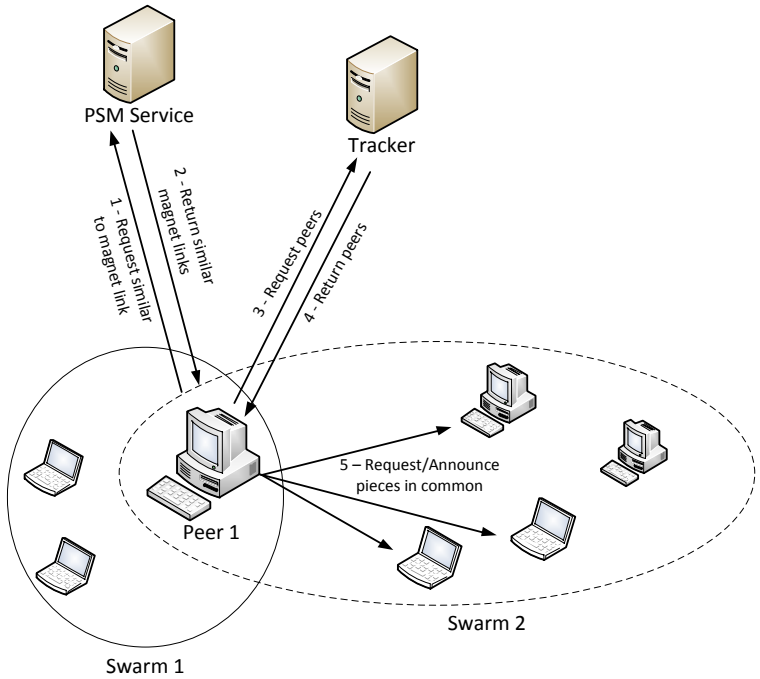


Figure: PSM workflow.

Conclusions

- High locality properties in both large and small size swarms
- Many patterns observed:
  - Content shared
  - Peer behavior
- Findings suggest that BitTorrent traffic can be restrained and inter-ISP traffic decreased

Publications

This paper was submitted to be presented at INFOCOM'2012 and we are now waiting for its evaluation.

* António Homem Ferreira, Ricardo Lopes Pereira and Fernando M. Silva. Partial Swarm Merger: Increasing BitTorrent content availability.