

# Distributed computing peer-to-peer algorithms for collective communication and network tomography

Kiril Dichev

Heterogeneous Computing Laboratory  
School of Computer Science and Informatics  
University College Dublin

May 11, 2012



# Motivation

- In distributed computing, algorithms often deal with heterogeneity and hierarchy of the network (examples: peer-to-peer)
- Can high-performance computing learn from distributed computing when utilizing heterogeneous platforms?
- Use case: BitTorrent protocol

# BitTorrent overview

The BitTorrent protocol is designed to be:

- fair
- dynamic and adaptive
- fault-tolerant
- efficient
- scalable

# BitTorrent vs MPI: Broadcast operation

- We compare BitTorrent and MPI when performing a broadcast operation
- We use only the Ethernet network on different sites of Grid'5000
- On homogeneous clusters, MPI always performs better
- On a hierarchy of clusters and for large enough messages (tens of MB), BitTorrent outperforms modern MPI implementations on the tested setting

# BitTorrent vs MPI: Large-message broadcasts on hierarchical clusters

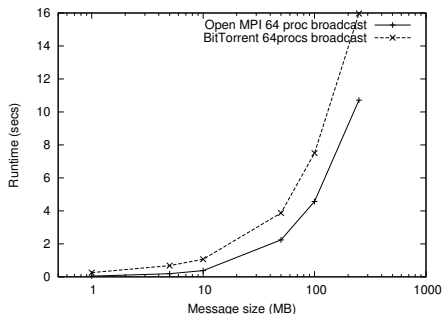


Figure: 16 multi-core nodes on one Bordeaux cluster

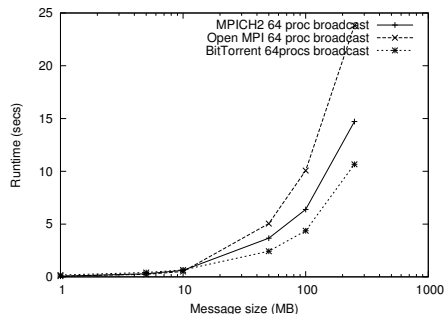


Figure: 64 nodes across 3 clusters on Grenoble

# BitTorrent communication

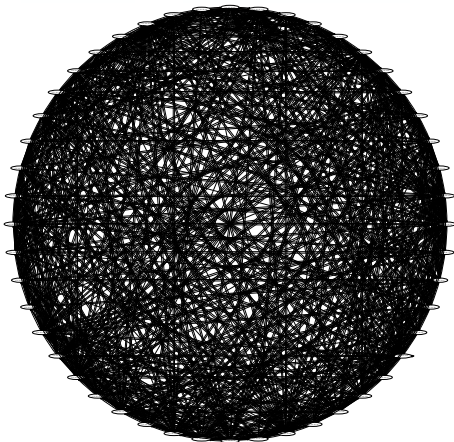


Figure: A runtime-based graph profiling the inter-node communication with BitTorrent. Each directed edge stands for data chunks transferred between two peers. In comparison, the Open MPI broadcast algorithm would build a minimal subgraph spanning all nodes (linear tree algorithm).

# Network tomography with BitTorrent and clustering algorithms <sup>1</sup>

Another interesting application of BitTorrent is network tomography

- Measure the flow of data during BitTorrent broadcasts
- Use clustering algorithms on the measured data flow
- Results are promising - efficient and reliable network tomography during intense communication, including bandwidth bottlenecks

---

<sup>1</sup>This work was done in collaboration with Fergal Reid from Clique Research Cluster

# Example: Physical Clusters and Network Tomography on Bordeaux site

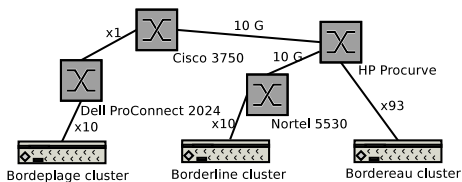


Figure: Physical topology of Bordeaux site

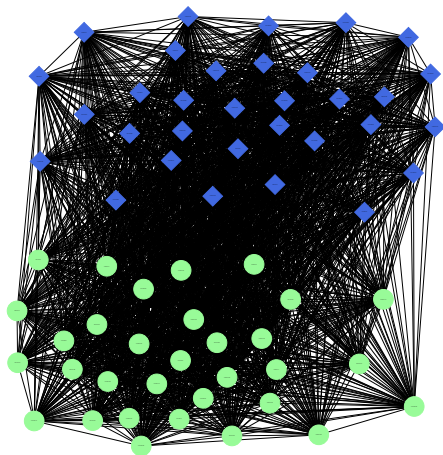


Figure: Clustering generated by network tomography



# Example: Network tomography across four sites

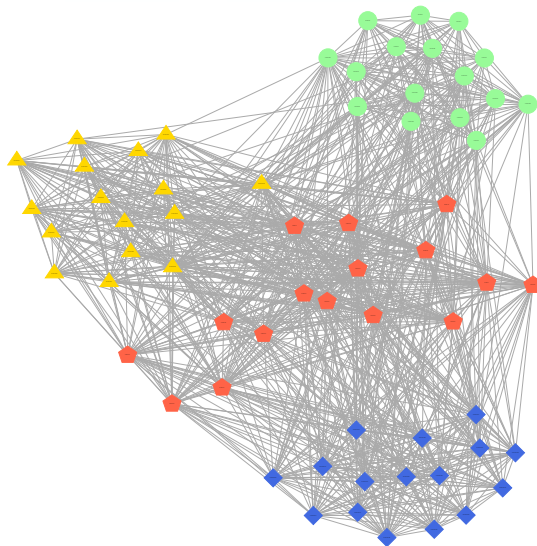


Figure: Experiment on four sites on Grid'5000 - logical clustering according to

Questions?