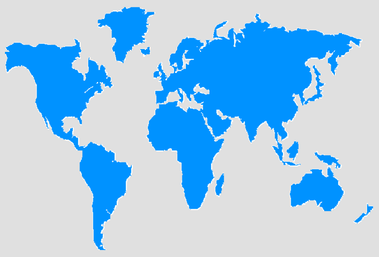
StreamGlobe: P2P Stream Sharing



Processing and Sharing Data Streams in Grid-Based P2P Infrastructures

Richard Kuntschke, Tobias Scholl, Bernhard Stegmaier, Alfons Kemper, and Angelika Reiser
Lehrstuhl Informatik III: Datenbanksysteme
Technische Universität München – Fakultät für Informatik

Network Basics:

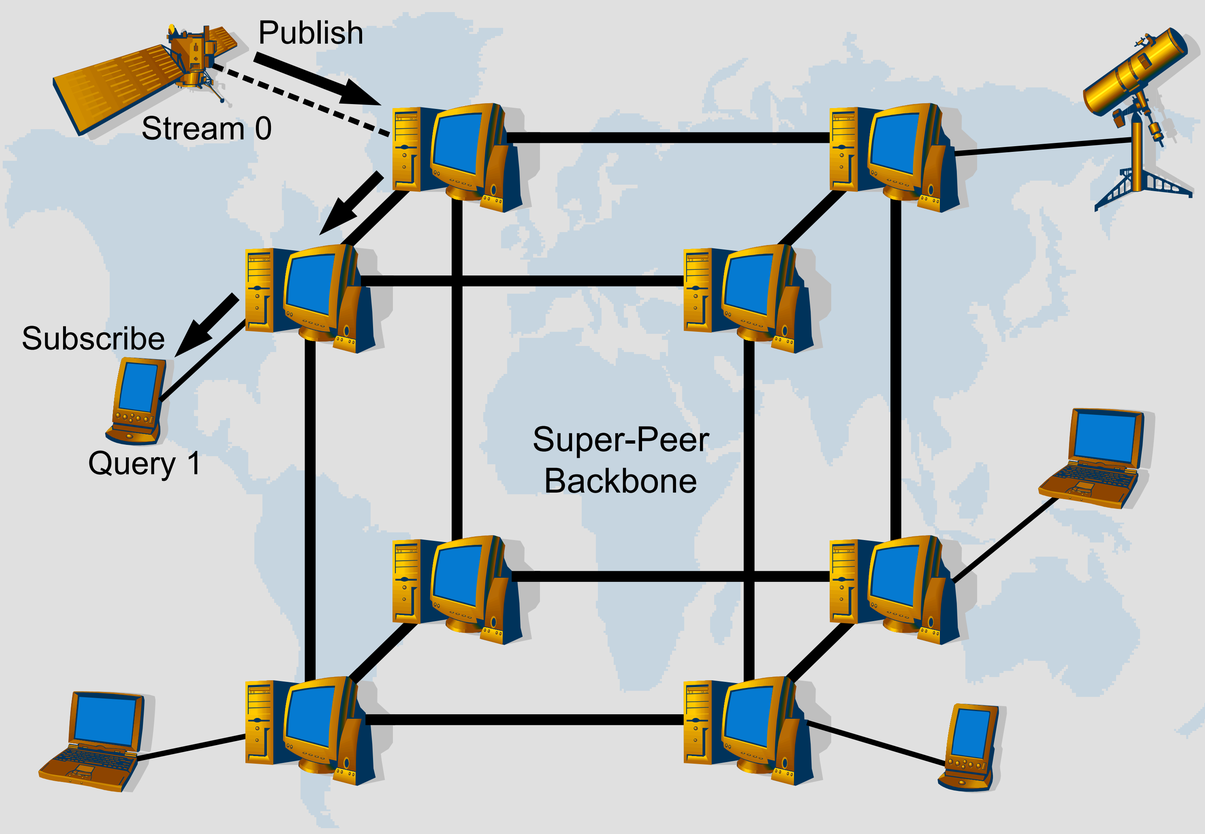
- Grid-based P2P network
- Super-Peer Backbone
- Super-Peers: Powerful stationary servers
- Thin-Peers: Less powerful, possibly mobile peers/sensors

Deficiencies of traditional approach:

- Redundant transmission of data streams
- Redundant execution of stream transforming operators
- Transmission of unnecessary data

⇨ Increased network traffic
⇨ Increased peer load

Publish & Subscribe in a P2P Network



StreamGlobe Basics:

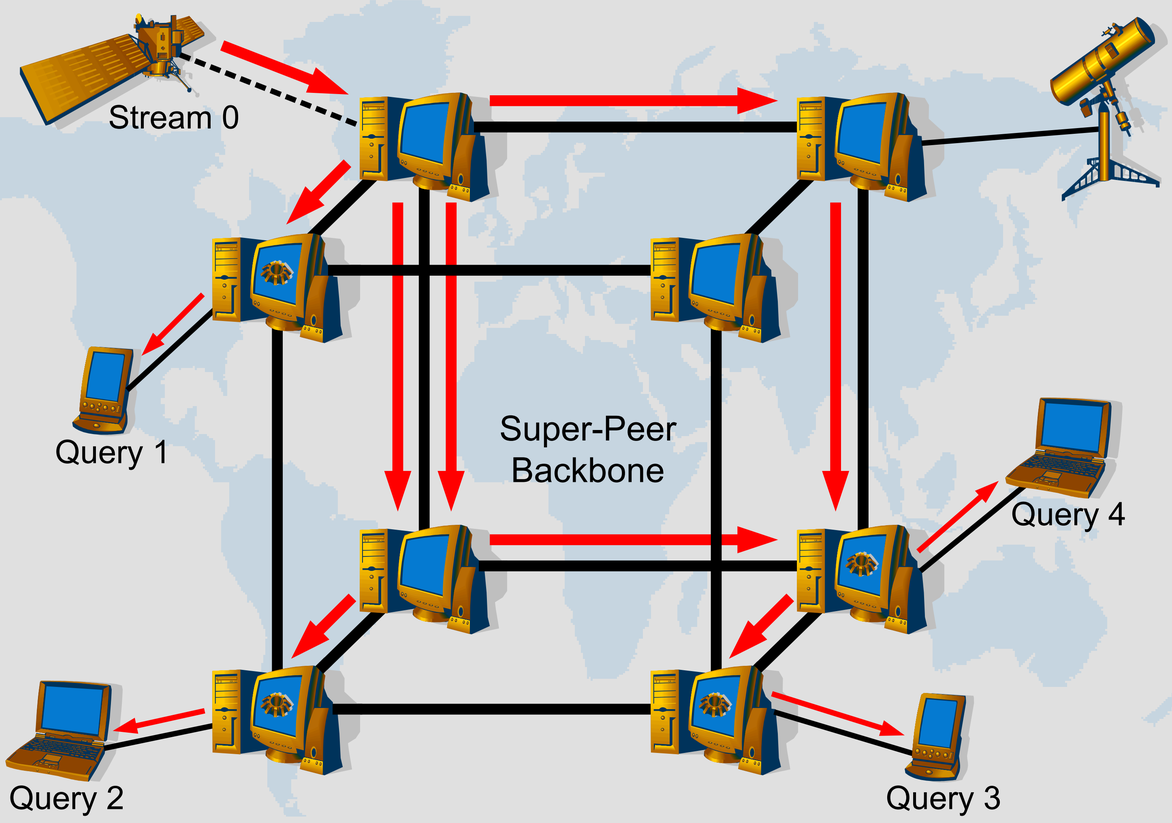
- StreamGlobe: Data Stream Management System (DSMS)
- Super-Peers process and route data streams
- Thin-Peers publish and subscribe to data streams

Benefits of StreamGlobe approach:

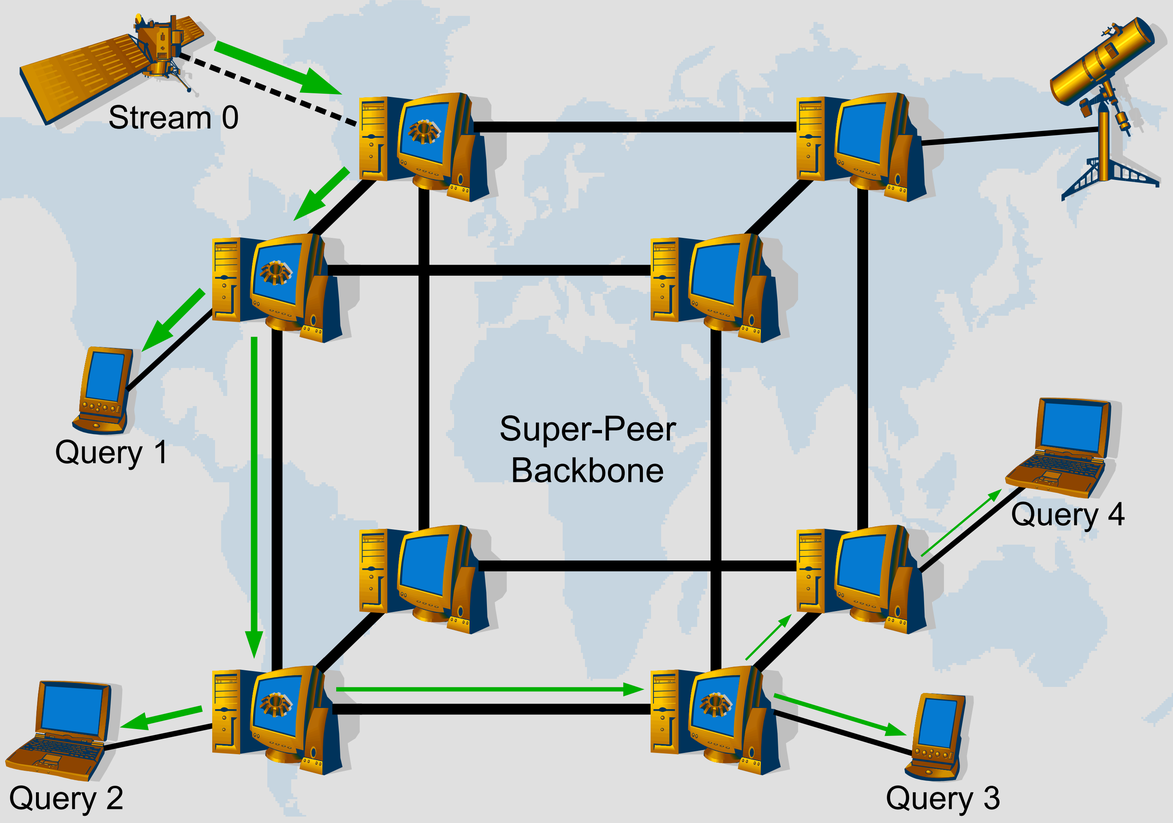
- Stream sharing avoids redundant stream transmission
- Sharing computational results avoids redundant computation
- Early filtering and aggregation avoid unnecessary transmission

⇨ Reduced network traffic
⇨ Reduced peer load

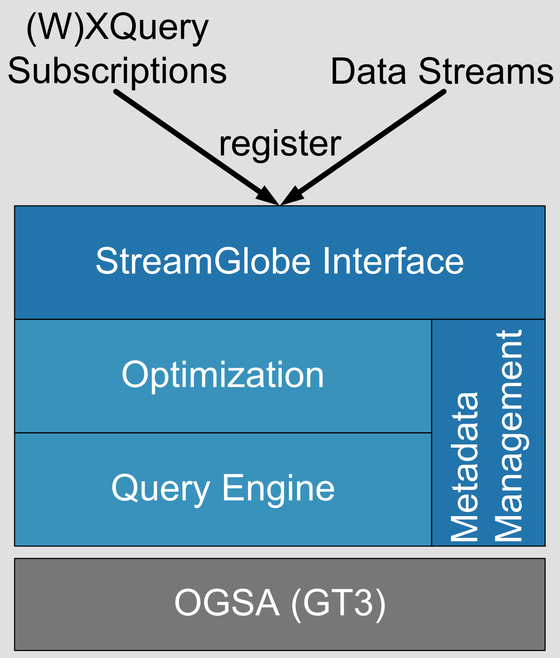
Traditional Approach without Stream Sharing



StreamGlobe Approach using Stream Sharing



Peer Architecture Overview



- Built on top of *Open Grid Services Architecture*
- Layers implemented as collaborating grid services
- Availability of services according to capabilities of peers
- *FluX* query engine for subscription evaluation

Optimization using Data Stream Sharing

Optimization Techniques:

- In-network query processing
  - Distribute query processing operators in the network (query shipping)
  - Early filtering and aggregation
- Multi-query optimization
  - Share processing results (i.e., data streams)

Optimizer:

- Cost-based optimizer
- Various cost functions possible

Optimization Benefits:

- Reduced network traffic
- Reduced computational load on peers
- Load balancing
- Increased flexibility
- Parallelization
- Reduced latency

http://www-db.in.tum.de/research/projects/StreamGlobe