Algorithms and Programming IV
Basics of Peer-2-Peer Systems (20-1)

Summer Term 2020 | 29.06.2020
Claudia Müller-Birn, Barry Linnert
Recap: Peer-to-Peer Architecture

- Is composed of a large number of peer processes running on separate computers
- All processes have client and server roles: servent
- Patterns of communication between them depends entirely on application requirements
- Need to place and retrieve individual computers is more complex than in client-server architecture
Our Topics Today

- Motivation Peer-to-Peer (P2P) Systems
- Overlay Networks (Unstructured and structured Overlay Networks)
  - Unstructured Overlay Networks
    - Centralized P2P Systems
    - Pure P2P Systems
    - Mixed P2P Systems
  - Structured Overlay
    - DHT-based P2P Systems
Motivation Peer-to-Peer (P2P) Systems

• Represent a paradigm of distributed systems and applications in which data and computational resources are contributed by many hosts on the Internet.

• Enable the sharing of data and resources on a very large scale by eliminating any requirement for separately managed servers and their associated infrastructure.

• Used to provide file sharing, web caching, information distribution and other services, exploiting the resources of tens of thousands of machines across the Internet.
Network Virtualization

- Network virtualization is concerned with the construction of many different virtual networks over an existing network such as the Internet.

- Each virtual network can be designed to support a particular distributed application.
Overlay Networks

An *overlay network* is a virtual network consisting of nodes and virtual links, which sits on top of an underlying network (such as an IP network).

An overlay network offers something that is not otherwise provided:

- a service that is tailored towards the needs of a class of application or a particular higher-level service (e.g., multimedia content distribution);
- more efficient operation in a given networked environment (e.g., routing in an ad hoc network);
- an additional feature (e.g., secure communication).
Advantages and Disadvantages of Overlay Networks

(+) • They enable new network services to be defined without requiring changes to the underlying network.
• They encourage experimentation with network services and the customization of services.
• Multiple overlays can be defined and can coexist.

(-) • Introduce an extra level of indirection
• Add to the complexity of network services
Purpose of Use

Peer-to-peer overlays are used for an efficient and scalable sharing of individual peers’ resources among the participating peers.

Depending on the type of the resources which are shared, the peer-to-peer overlays can be identified as oriented for:

- Data-sharing (data storage and retrieval);
- Bandwidth-sharing (streaming);
- CPU-sharing (distributed computing).

Skype: An example of an overlay network
• A public volunteer computing project hosted by the University of California, Berkeley.

• Analyzing radio signals, searching for signs of extraterrestrial intelligence.

• Distributed computing software that runs either as a screensaver or continuously while a user works, making use of processor time that would otherwise be unused.

• Released to the public on May, 1999 and closed by end of March 2020.
Peer-to-Peer System

- Resources are shared between the peers.
- Resources can be accessed directly from other peers.
- Peer is provider and requester (servent concept).

Unstructured Overlay

Structured Overlay
Characteristics of Unstructured Overlay Network

The **placement** of content (files) is completely unrelated to the overlay topology and the content typically needs to be located.

- Peers & resources have **no special** identifier.
- Each peer is responsible **only** for the resources it submitted.
- The location of the resources is **only** known to submitter.
- New resources can be introduced at any location.

The main **task** in such networks is to search

- Find all peers storing/being in charge of resources fitting some criteria.
- Direct communication when peers have been identified.
Characteristics of Structured Overlay Network

The overlay **topology** is tightly controlled and files (or pointers to them) are placed at precisely specified locations.

- Location of resources **not** only known to submitter
- Each peer may well be responsible for resources **it has not** submitted
- Introduction of new resource(s) at **specific** location, i.e. to give peers and resources (unique) identifiers
- PeerIDs and ObjectIDs (RessourceIDs) should be from the **same** key set
- Each peer is responsible for a specific range of ObjectIDs, i.e., RessourceIDs

The main **task** is to lookup.

- To “route” queries across the overlay network to peers with specific IDs.
## Structured vs. Unstructured P2P Systems

<table>
<thead>
<tr>
<th></th>
<th>Unstructured peer-to-peer</th>
<th>Structured peer-to-peer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>Self-organizing and naturally resilient to node failure.</td>
<td>Guaranteed to locate objects (assuming they exist) and can offer time and complexity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bounds on this operation; relatively low message overhead.</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>Probabilistic and hence can not offer absolute guarantees on locating objects; prone to</td>
<td>Need to maintain often complex overlay structures, which can be difficult and costly</td>
</tr>
<tr>
<td></td>
<td>excessive messaging overhead which can affect scalability.</td>
<td>to achieve, especially in highly dynamic environments.</td>
</tr>
</tbody>
</table>
Peer-to-Peer

Unstructured Overlay
- Centralized P2P
- Pure P2P
- Hybrid P2P

Structured Overlay
- DHT-based