Chair for Network Architectures and Services
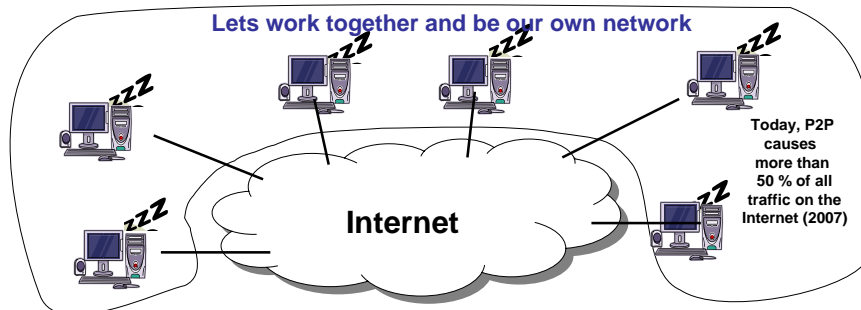Department of Informatics
TU München – Prof. Carle

**Peer-to-Peer Systems
and Security
IN2194**

**Chapter 1
Peer-to-Peer Systems
1.1 Introduction**

Prof. Dr.-Ing. Georg Carle
Dipl.-Inform. Heiko Niedermayer

Overview

- Motivation
- Overlay and Peer-to-Peer
- Graph Theory Basics
- Routing and Searching

Motivation

Lets work together and be our own network

Today, P2P causes more than 50 % of all traffic on the Internet (2007)

Internet

- More and more private users on the Internet
- Powerful private end systems
- Flatrates with always-on users

Why waste these resources?

- A need to provide services independently from commercial or dedicated server providers.

- Application-specific network structures instead of machine location-based addressing.

For instance, friends want their computers to be together. → *Why not their own network - their own overlay network?*

Term: Overlay

"'Overlay networks is a term for networks that run on top of an existing infrastructure but provide certain additional functionality.'" (Source: www.overlay-networks.info)
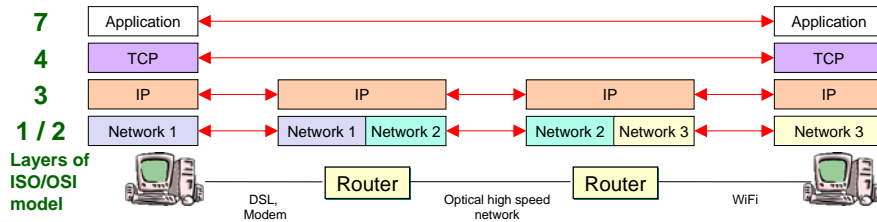
Overlay

Underlay
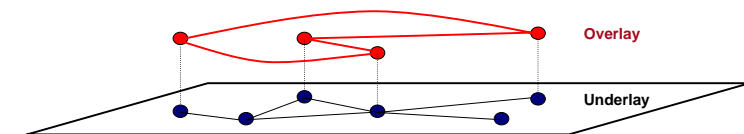
Overlay networks

- add another layer of abstraction in some place in the protocol stack.
- introduce a new structure of who is connected with whom.
- This new structure uses the connectivity of the lower layers for its links. The layer below the overlay is called underlay.
- Overlay networks are usually formed for some kind of reason, usually to provide a desired additional functionality.
  - e.g. send a message to all members of a group (multicast), the overlay represents the group and provides group membership functionality

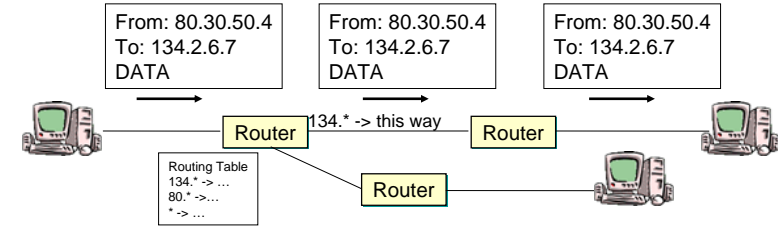## IP – an overlay in the network stack

The most well-known overlay is the Internet Protocol (IP) itself.

7	Application
4	TCP
3	IP
1 / 2	Network 1

Layers of ISO/OSI model

DSL, Modem — Router — Optical high speed network — Router — WiFi

The Internet Protocol (IP)

- abstracts from the technology and physical location present in the lower layers (Medium access, Physical Layer).
- forms a structure that is optimized to provide connectivity between all connected networks and their entities, no matter where they are and what access technology they use.

IP – an overlay in the network stack

From: 80.30.50.4
To: 134.2.6.7
DATA

From: 80.30.50.4
To: 134.2.6.7
DATA

From: 80.30.50.4
To: 134.2.6.7
DATA

Router   134.* -> this way   Router

Routing Table
134.* -> …
80.* ->…
* -> …

Router

The Internet Protocol (IP)

- For an packet (IP datagram) with any arbitrary target any system (router) on the way can efficiently decide where it has to go.
  - **Routing**: Operation to determine optimal paths and structure in a network, usually results in a routing table.
  - **Forwarding**: Operation of the router when it sends a packet to its next hop according to routing table.
- If the target does not exist, the system will recognize this at some point and return an error message (ICMP protocol).

## Examples for Overlay Networks

Virtual Private Networks (VPNs)

- used to connect local networks that are located in different places using other networks, e.g. the Internet or phone lines.

10.3.2.11   Site-to-Site   10.3.2.34

Network A   Internet   Network B

End-to-Site

Provider   10.3.2.67
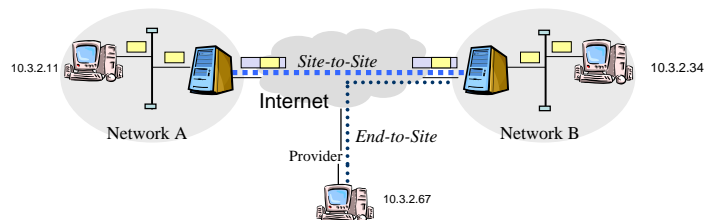
Domain Name System (DNS)

- used to map names of resources to IP adresses
- DNS servers internally form an overlay
  - Hierarchy with a group of root servers to more and more specialized servers for subdomains.

Term: Peer-to-Peer
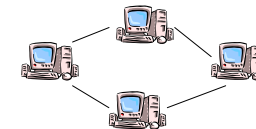
Peer-to-Peer systems

- Distributed systems that consist of equals (peers) with no predefined distinction between client and server and no dedicated servers or central authority.

Characteristics

- Peer-to-Peer networks are decentralized and take advantage of resources at the edge of the Internet, in particular the computers of users.
- End systems do not primarily serve the purpose of the Peer-to-Peer system.
  → Their resources must not be exhausted by the Peer-to-Peer network
- Computers are not always-on.
  → Environment is less stable and more dynamic than in the traditional client-server case.

Peer-to-Peer or not Peer-to-Peer

Auctions / Ebay
- Peer-to-Peer
  - Money and goods exchange (nothing to do with the network)
- Not Peer-to-Peer
 - The platform itself (Auctions, Accounts, Information transfer) and its Information Management

Skype
- Peer-to-Peer
  - Lookup, User Interaction, Data Exchange
- Not Peer-to-Peer
  - Login, Account Management

Many Peer-to-Peer systems are not purely Peer-to-Peer.


Overlay vs. Peer-to-Peer

- Overlay and not Peer-to-Peer ?
  - Virtual Private Networks with VPN servers.

VPN Server

- Peer-to-Peer system and no Overlay ?
  - Usually Peer-to-Peer systems create a new structure (overlay) on top of an underlying network.
    - No perfect examples without overlay.
  - Possible examples
    - Peers in an ad-hoc or sensor network may not add a new structure with new identities.
    - Peers in a LAN playing a P2P game use IP.
    - Students in a lecture organize where they sit, etc. However, again no new adressing or communication structure.
    - …
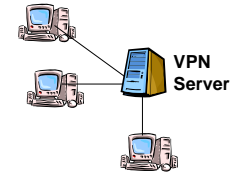
Spontaneous meeting, no central coordination
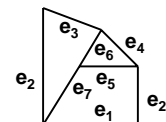

Some terms from Graph Theory

- Graph G=(V,E)

- Vertex set $V = \{v_1, v_2, \ldots v_n\}$
  - We usally say **nodes.**
  - $n = |V|$

- Edge set $E = \{e_1, e_2, \ldots e_m\}$
  - We usually say **links.**
  - $m = |E|$
 - Can have attributes like distance, etc.
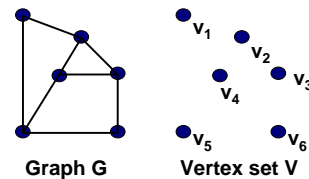

Graph G   **Vertex set V**

Edge set E

Some terms from Graph Theory

- Distance d(i,j)
  - Shortest path between nodes v_i and v_j

- Diameter D of G
  - Longest distance in graph G

- Degree
  - Node degree = number of edges adjacent to node
  - Degree of a graph = max. node degree

- A graph is <u>connected</u> if there is a path from any node in the graph to any other node in the graph.

- A graph is <u>k-connected</u> if any k-1 nodes can be removed without causing the resulting subgraph to become disconnected.
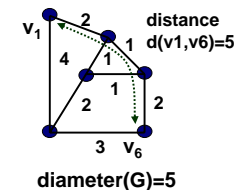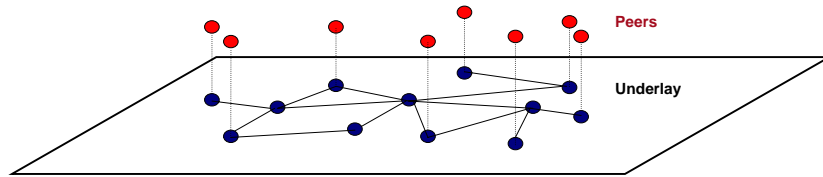

distance d(v1,v6)=5
diameter(G)=5
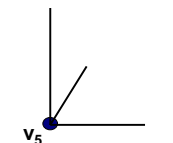
degree (v_5) = 3

Peer-to-Peer network
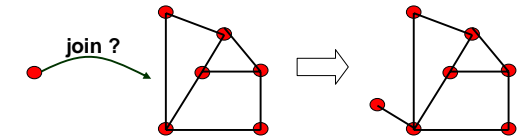

Peers

Underlay

Underlay
- Provides connectivity between all peers in the Peer-to-Peer network (overlay).

Peers $V = \{v_1, v_2, \ldots v_n\}$
- Peers are the nodes of the graph G.
- Peers may have a name (identities are usually necessary).
- The set of edges E needs to be created by the Peer-to-Peer algorithms.
  - The graph needs to be connected.
  - The structure should be good for the purpose of the Peer-to-Peer system.
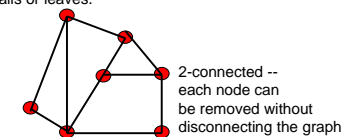
P2P network is not static – Peers join and leave

- Node joins
  - Needs to be added to the network
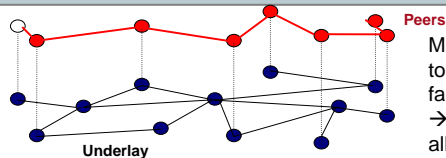 - Usually via some node in the network already known (rendezvous point, list/cache of nodes)

join ?

- Node leaves
  - Important to keep the graph connected
  - Better not rely on a single node that could leave anytime

leave ?

disconnected
when this node fails or leaves.
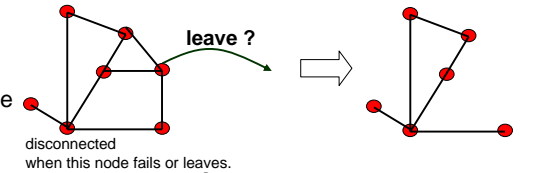
- How to organize such a network?
  - e.g. k-connected graph

2-connected --
each node can
be removed without
disconnecting the graph

Application Requirements

Peers

Multicast from the white game server to its peers. Is this a good graph for fast delivery?
→ No, a balanced tree allows O(logn) diameter.

Underlay

Application
- Peer-to-Peer networks are usually created for an application or application scenario.
  - Filesharing
  - File Distribution
  - Instant Messaging and Voice-over-IP
  - Multicast
  - Peer-to-Peer Video Streaming
  - Anonymous communication and services
  - …
- The application is the purpose of the Peer-to-Peer network.
- The application and its requirements determine if a given graph is a good or a bad choice.

Operational aspects

- Find someone to
  - get something
  - use a service
  - interact
  - interact for a cooperative service or goal
  - maintain network
- Find something (item, data, information, etc.) to
  - get it
  - set it
- Interact with other nodes to cooperatively
  - provide a service
  - share resources
  - run an algorithm
- …

Routing and Searching

„How can we **find** something or someone?"

Routing

- Routing = algorithm / protocol to find a path.
- Given a destination send a message to the target on a direct way.
- Needs
  - Names for peers or items
  - Means to determine and know the way

A, B, C

D, E, F

to C

to C
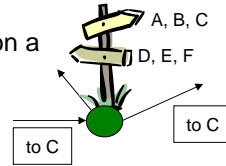
Searching

- Look for someone or something.
  - Location of destination or set of destinations is unknown.
  - Breadth-First / Depth-First Searching, Flooding, etc.
- Usually used when no criteria exists that allows to determine a best direct way.
  - The requested information can be anywhere.

Opera Expert?

No.

Opera Expert?

No. Maybe D,E,H.

...
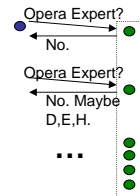
Routing and Searching – Types of Queries

Types of Queries

- Exact
- Fuzzy Queries
  - Find somthing similar according to a metric.
- Range Queries (Bereichsabfrage)
  - Find everything in given Interval, e.g. [c,pfau] or (2,4]
- String Queries
  - substring, startsWith, endsWith, …
- Complex queries
  Find peers / items where for a node-specific and query-specific function $f_{this\text{-}query}$ holds
  $$f_{this\text{-}query}(\texttt{candidate}) > threshold_{this\text{-}query}$$
- SQL / Database queries

„Marienplatz, Munich"

„somewhere in Munich"

„everything on the way from Garching to Munich Marienplatz"

„All restaurants near Marienplatz with good food according to <u>my</u> taste."

Conclusion and Outlook

Conclusion

- What is a Peer-to-Peer system?
  - Not necessarily pure P2P.
- Overlay and Underlay
- Graph Theory Basics
- Routing and Searching