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

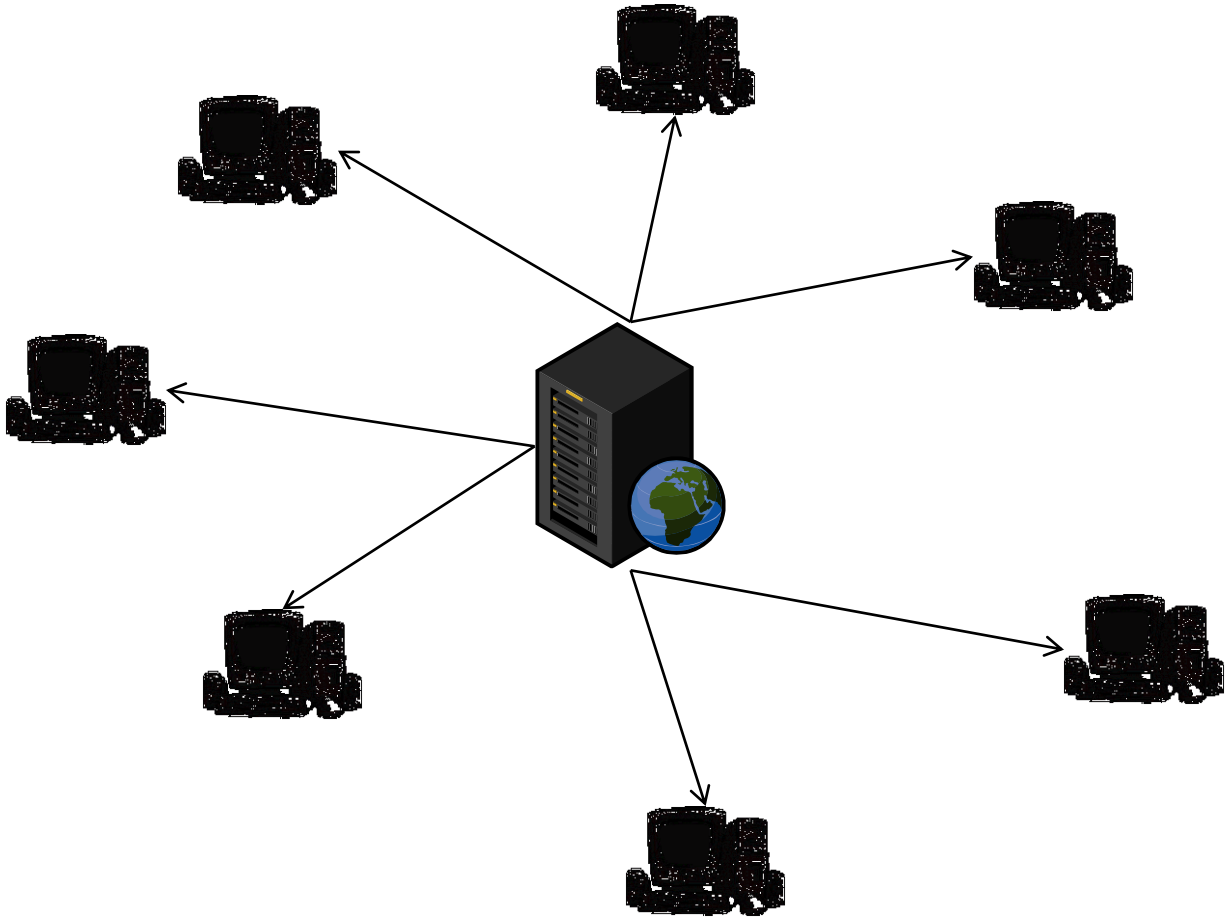
Peer-to-Peer Systems and Security IN2194

Chapter 1
Peer-to-Peer Systems
1.1 Basics

Dr. Heiko Niedermayer
Christian Grothoff, PhD
Prof. Dr.-Ing. Georg Carle

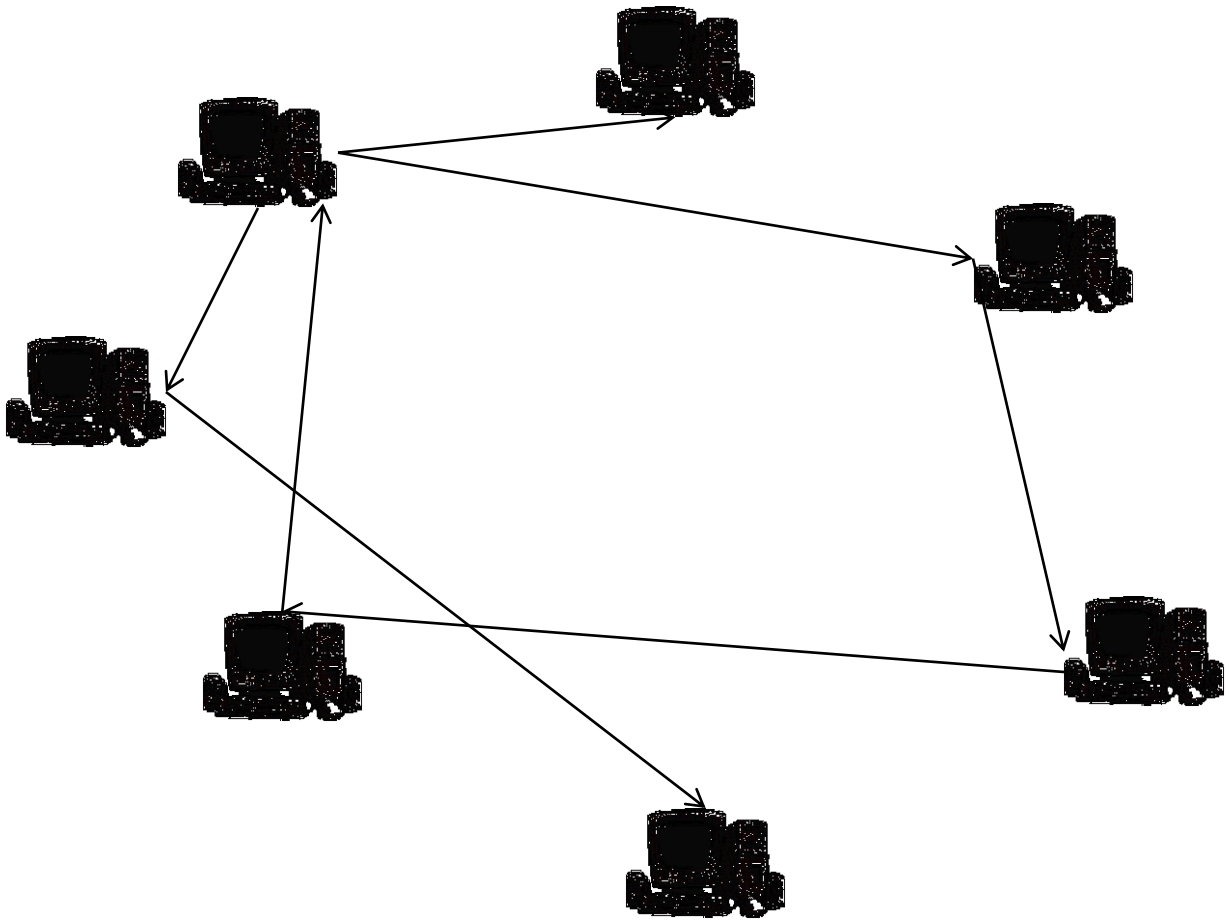
Client / Server

- Server distributes data it owns to clients



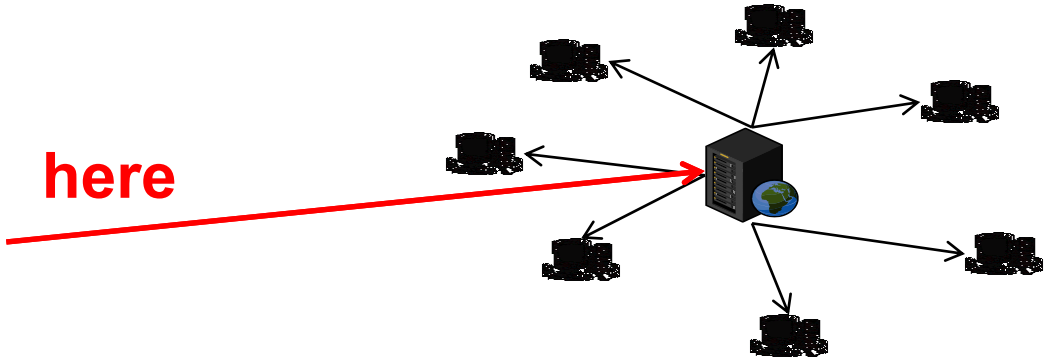
Peer-to-Peer

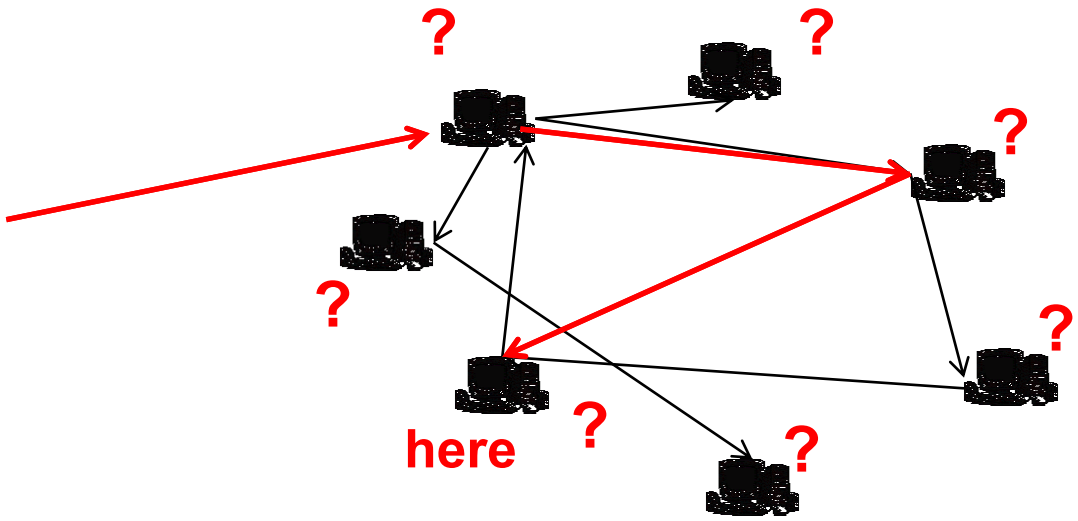
- Peers distribute data they have to other peers



Where is X?

- Where is my data?



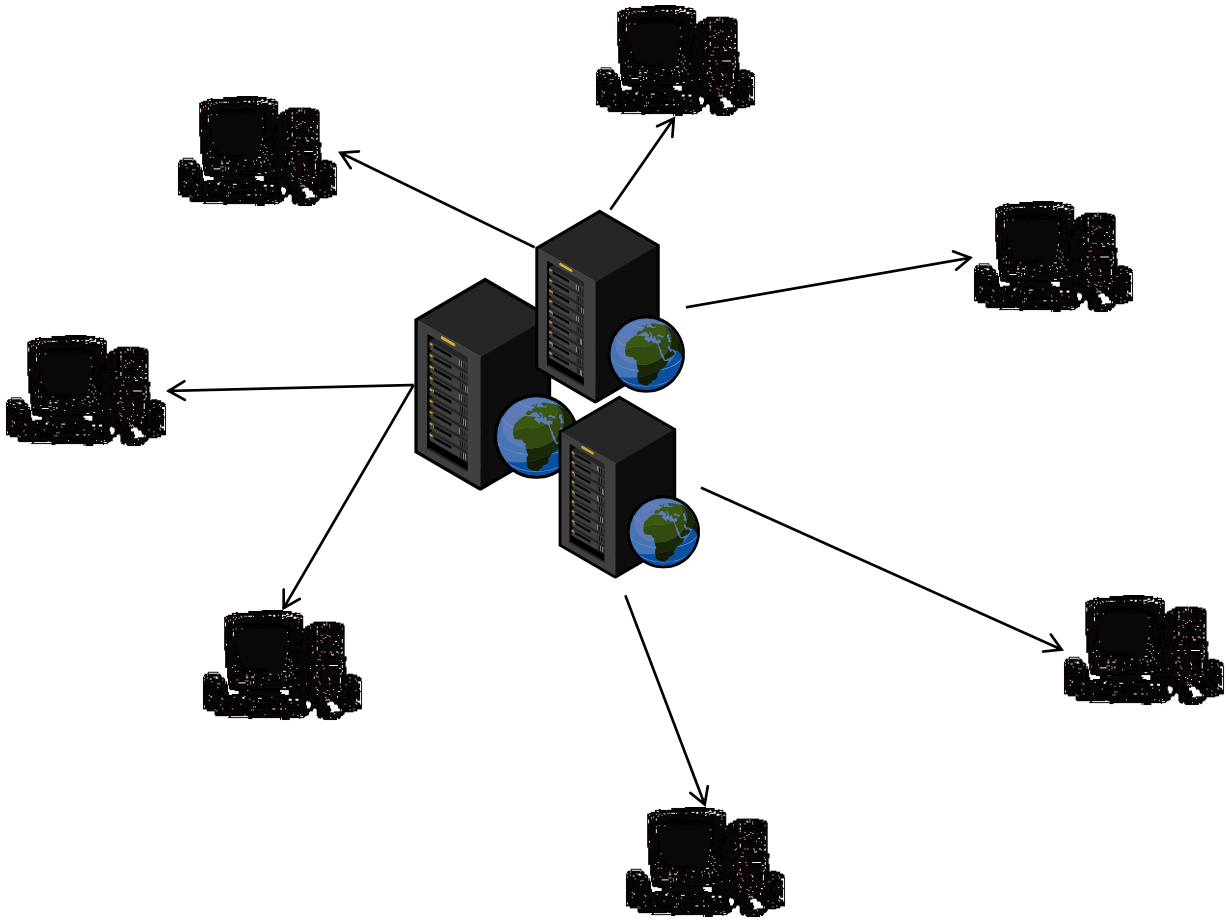


Scalability

- Scalability
  - System can become larger and larger and will still operate good enough.
  - Requires
    - Resources (CPU / Storage) increase with participation (or large enough)
    - Delay / Lookup remains small enough, e.g. log(n) with n number of nodes

Datacenter

- Scale using a large number of servers, today large enough

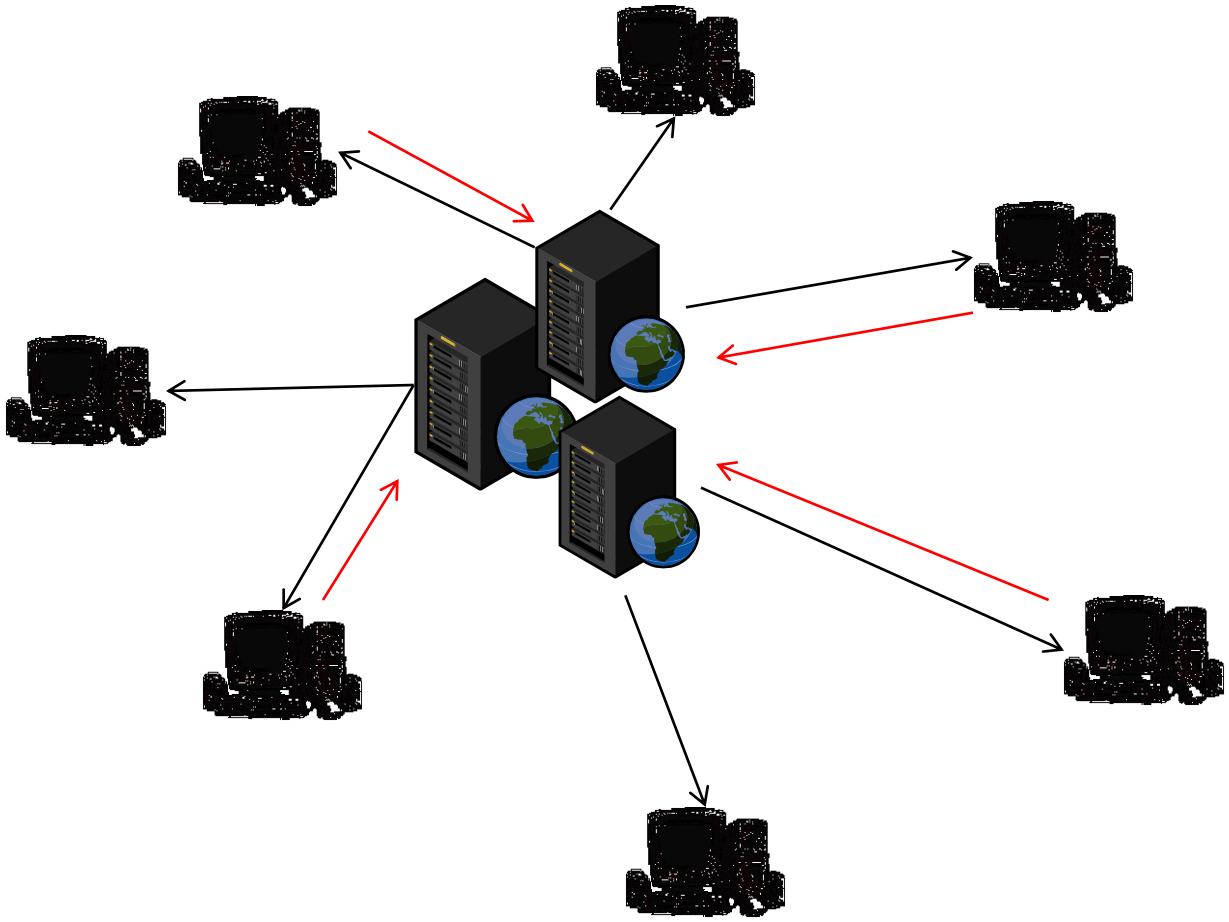


User-Generated Content

- E.g. Facebook, Youtube

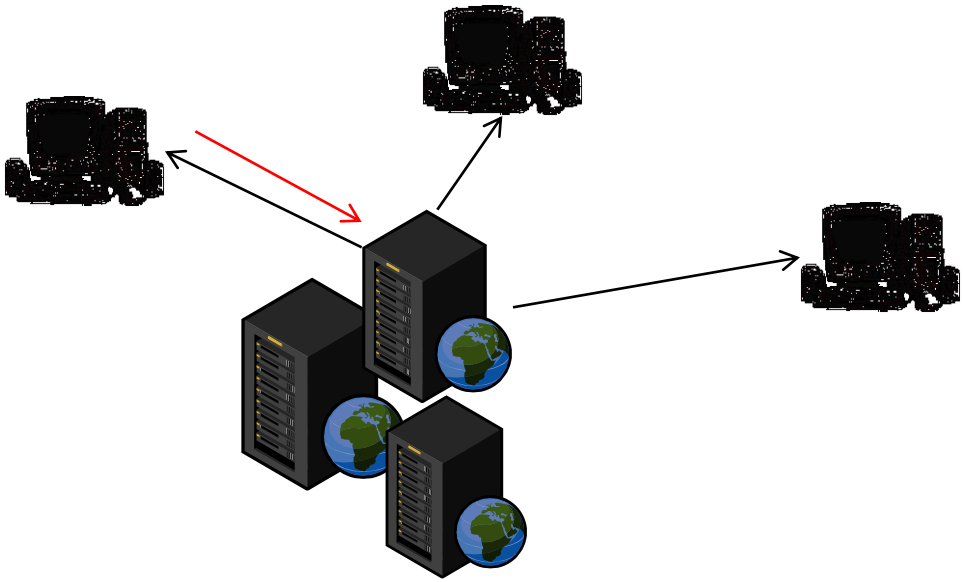
Generate content („Peer“)

Distribute content („Server“)



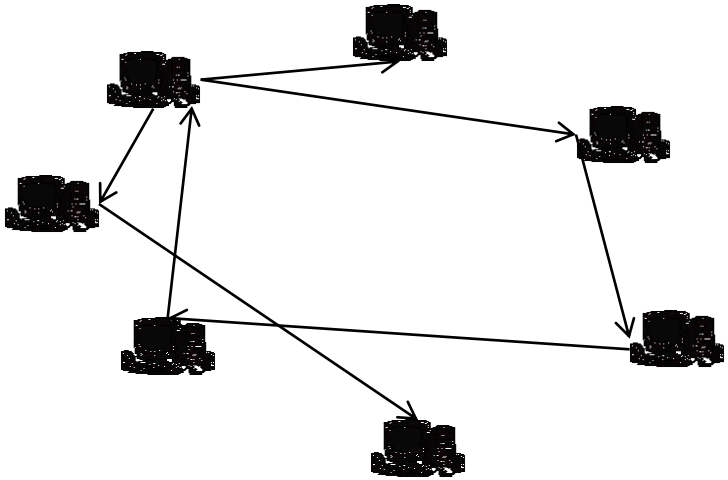
Cloud Computing

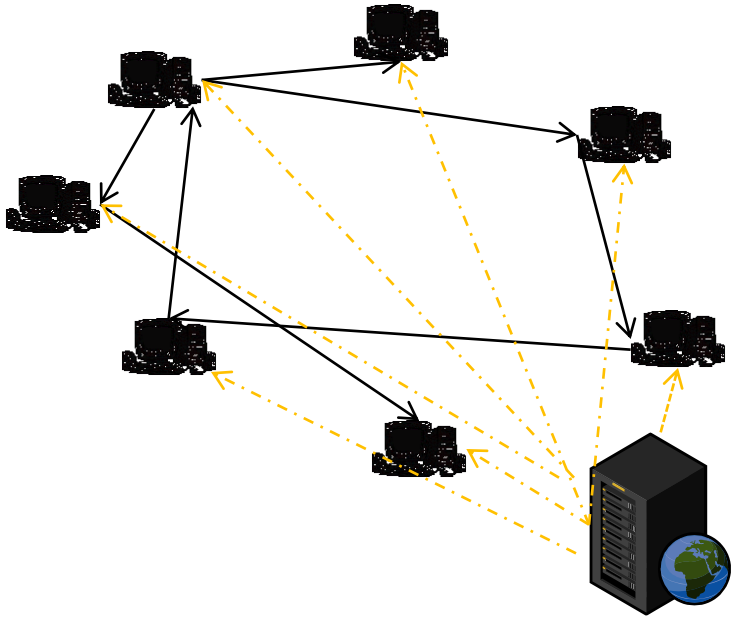
- Client / Peer generates resources for itself
- Depending on application, server may serve them to others as well



Peer-to-Peer vs Distributed Computing

- Pure P2P assumes no single administrative control or federation, no realiable operation, … no central planer or optimizer in the background

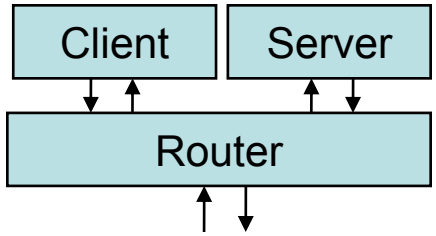




Terms: Peer, Peer-to-Peer

Peer

- A host in the network, usually at the edge of a network, in a private home.
- A peer operates as client and server in parallel.
  - Peers are meant to be equal, no pre-defined distiction into client, server or other roles.
  - In reality, not all peers are equal.

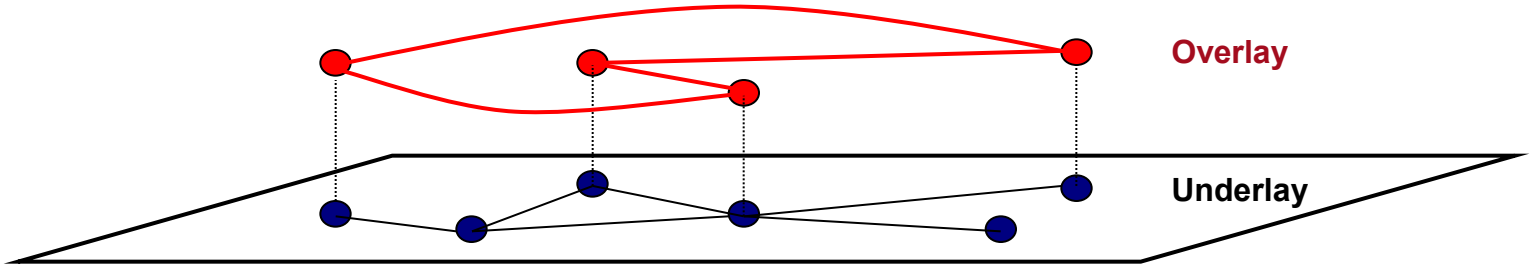


Peer-to-Peer

- Paradigm to provide functionality by a set of peers.
- Control is shifted from a server controlled by a single administration to a set of peers, controlled by their users.

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)

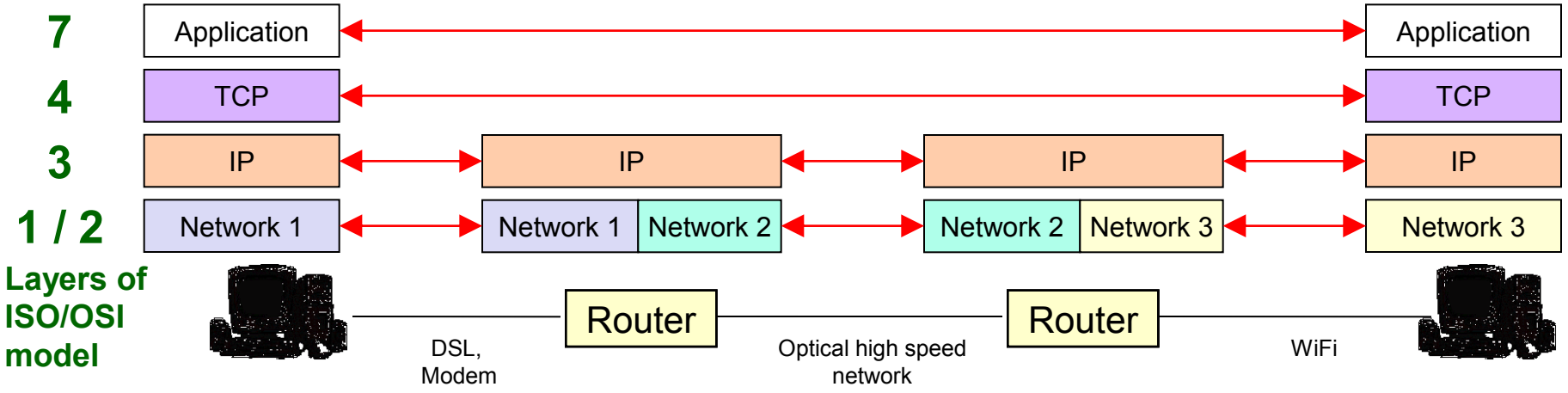


Overlay networks

- 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

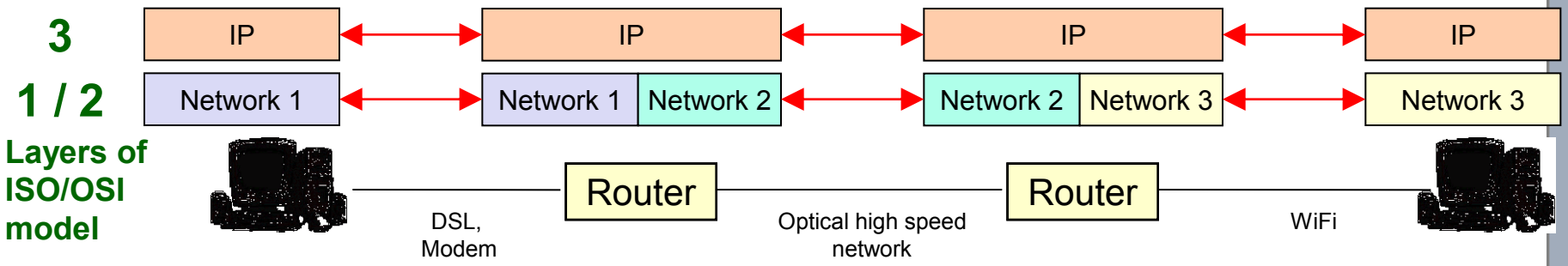
The most well-known overlay is the Internet Protocol (IP) itself.



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



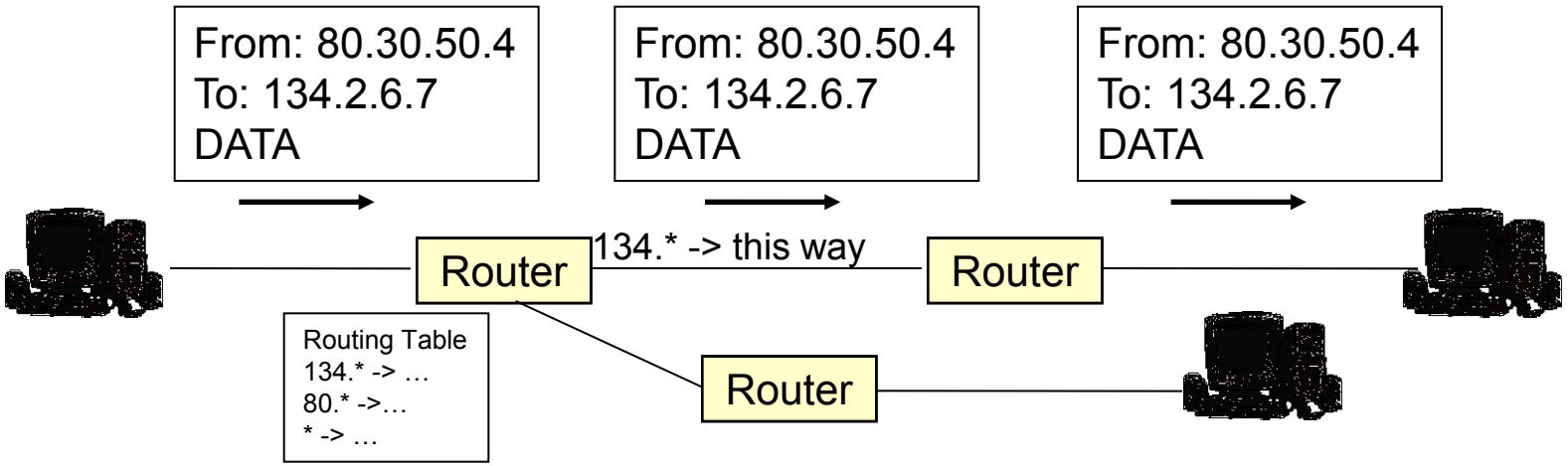
- The problem to route a packet to its destination is usually separated into the *routing* and the *forwarding*.

Routing

- Operation to determine optimal paths and structure in a network.
- Method: Routers run a routing protocol and a shortest path algorithm to determine the best paths to all reachable destinations in the network.
- Result: Routing table with entries (destination, next hop towards the destination) The next hop is a neighbor that can be reached with the local layer 2 protocol (e.g. ethernet).

destination	next hop
131.*.*.*	115.3.1.2
80.12.*.*	114.3.1.2
…	…

IP – an overlay in the network stack



Forwarding

- Operation of the router when it sends a packet to its next hop according to routing table.
- Situation: a packet arrives at a router.
- Response: router checks its routing table for entries that match with the destination given in the packet.
  - Use the best next hop specified in the corresponding entry of the routing table.
  - If the target does not exist, some router on the path will recognize this and return an error message (ICMP protocol).
- Necessary: For any packet (IP datagram) with any arbitrary target any system (router) on the way can efficiently decide where it has to go.

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

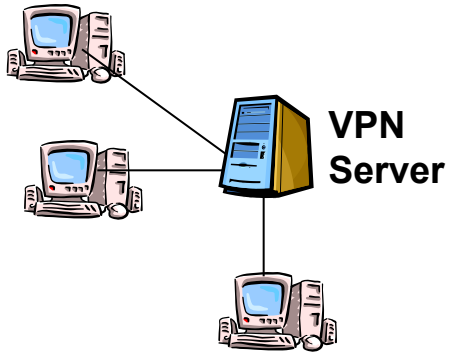
Skype

- 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.

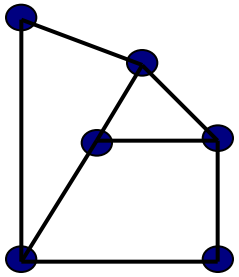


- Peer-to-Peer system and no Overlay ?
  - Usually Peer-to-Peer systems create a new structure (overlay) on top of an underlying network.
    - No perfect examples without overlay.
  - Possible examples
    - Peers in an ad-hoc or sensor network may not add a new structure with new identities.
    - 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.
    - Friends voting what to do…
    - …

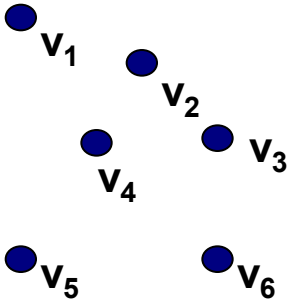
Spontaneous meeting, no central coordination

Some terms from Graph Theory

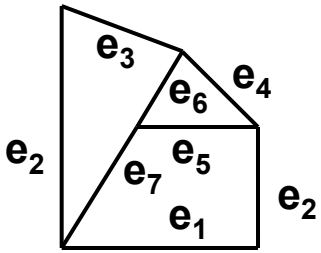
- Graph G=(V,E)
- Vertex set V = {v_1, v_2, …v_n}
  - We usally say **nodes.**
  - n = |V|
- Edge set E = {e_1, e_2, …e_m}
  - We usually say **links.**
  - m = |E|
  - Can have attributes like distance, etc.



Graph G



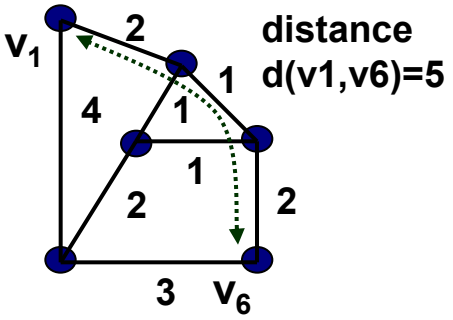
Vertex set V



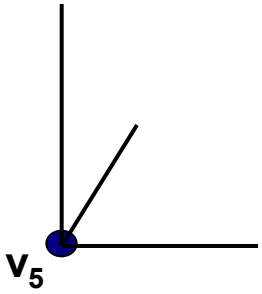
Edge set E

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 connected if there is a path from any node in the graph to any other node in the graph.
- A graph is k-connected if any k-1 nodes can be removed without causing the resulting subgraph to become disconnected.

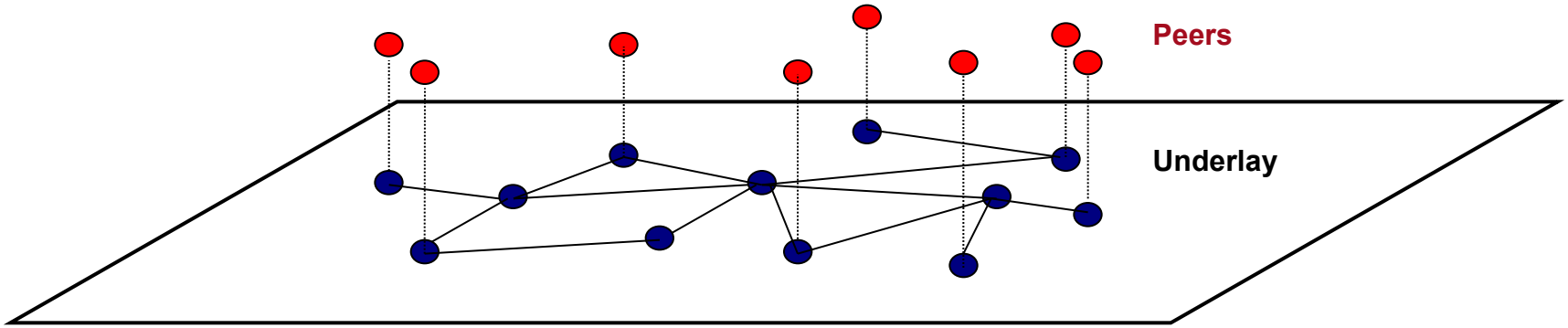


diameter(G)=5



degree (v_5) = 3

Peers and 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.

Complexity / Performance

- Let us consider a graph with n nodes.
- Complexity metrics (complexity to achieve a certain performance)
  - Number of links m = |E| in graph
  - Maximum/Average State of a node
    - Degree of a node, Routing table size, Non-graph-related information
  - Number of Responsibities (e.g. number of files provided to network)
  - Traffic routed through a node / Number of Requests
  - Maintenance Traffic / Overhead (amout of traffic vs maintenance traffic)
  - …
- Performance metrics
  - Diameter or average distance in Graph
  - Robustness / Connectedness of a graph
  - Performance / Quality of application-specific operations
  - Fair distribution of work
  - ….