Contents

• Overlay networks and intro to networking
• Unstructured networks
Overlay Networks

An overlay network is a network that is built on top of an existing network.

The overlay therefore relies on the so called underlay network for basic networking functions, namely routing and forwarding.

Today, most overlay networks are built in the application layer on top of the TCP/IP networking suite.
Overlay Networks

Overlay technologies can be used to overcome some of the limitations of the underlay, and at the same time offering new routing and forwarding features without changing the routers.

The nodes in an overlay network are connected via logical links that can span many physical links. A link between two overlay nodes may take several hops in the underlying network.
Requirements for Overlay Networks

1. Support the execution of one or more distributed applications by providing infrastructure for them.

2. Participate and support high-level routing and forwarding tasks. The overlay is expected to provide data forwarding capabilities that are different from those that are part of the basic Internet.

3. Deployed across the Internet in such a way that third parties can participate in the organization and operation of the overlay network.
Protocol Stack

• Layers are part of a network architecture
  – Provide services for layers above
  – Hiding the complexity of the current layer
• Multiple layers are needed in order to reduce complexity
  – Separation of network functions
  – Distribution of complexity
  – OSI, TCP/IP
• Protocols are building blocks of a network design
  – Can exist independently of layering
Background

• What is network architecture?
• Layered architecture
• Network structure
• Network evolution
• Motivation for overlay networks
Network architecture

• A set of principles and basic mechanisms that guide network engineering
  – Physical links
  – Communication protocols
    • Format of messages
    • The way in messages are exchanged
    • Elements of the protocol stack
• Where is the state?
Naming, Addressing, and Routing

NAMING

How to identify and name a node? Even if its address changes.

unicast: to a specific node
broadcast: to all nodes
multicast: to a subset of nodes
anycast: to any one in some subset (IPv6)

ADDRESSING

Where is the node located?

ROUTING

How to route information to the node’s address?
TCP/IP Network Stack

Application Layer

Transport Layer (TCP/UDP)

Networking Layer (IP)

Underlying network (link layer, physical)
Structure of the Internet

End-to-End Principle: application logic and state at the edge, core network is simple

Network maintains routing tables and forwards packets based on IP prefixes

Various protocols for intra-domain routing (OSPF, …)

Hierarchical structure: Autonomous Systems with tiers and peering links, economics of inter-domain routing

Border Gateway Protocol (BGP) for routing between domains

New requirements for IP and network: mobility, multihoming, security, multipath operation, …
The Hourglass

- Diverse applications
- Divergence
- Transport layer (TCP/IP)
- Convergence
- Diverse physical layers

Overlays!
Questions

Q: Where is the routing state in overlay networks?
A: Must be at the edge of the network. Easy to bootstrap, because network is not modified.

Q: What is the benefit of this from the network point of view?
A: Can overcome limitations of the current network and introduce new services.

Q: What is limitation from the network point of view?
A: Must take network structure into account (economics). Traversal solutions can be complex.
Evolution of the Network

Cisco reports trends and estimates in "Cisco VNI: Forecast and Methodology, 2011-2016"

Annual global IP traffic will pass the zettabyte ($10^{21}$) threshold by the end of 2016.

The number of devices connected to IP networks will be nearly three times as high as the global population in 2016.

Video delivery has become one of the recent services on the Web.

Estimates of P2P share of network traffic range from 20% to 40% (used to be 40-70%).
Cisco forecast: Global Video Highlights

Globally, Internet video traffic will be **55 percent** of all consumer Internet traffic in 2016, up from 51 percent in 2011.

The sum of all forms of video (TV, video on demand, Internet, and P2P) will continue to be approximately 86 percent of global consumer traffic by 2016.

Internet video to TV doubled in 2011. Internet video to TV will be 12 percent of consumer Internet video traffic in 2016.

**Video-on-demand** traffic will triple by 2016. The amount of VoD traffic in 2016 will be equivalent to 4 billion DVDs per month.

**High-definition** video-on-demand surpassed standard-definition VoD by the end of 2011. By 2016, high-definition Internet video will comprise 79 percent of VoD.
With video growth, Internet traffic is evolving from a relatively steady stream of traffic (characteristic of P2P) to a more dynamic traffic pattern.
Dynamic Patterns

Busy-hour traffic is growing more rapidly than average traffic. Busy-hour traffic will increase nearly fivefold by 2016, while average traffic will increase nearly fourfold.

Busy-hour Internet traffic will reach 720 Tbps in 2016, the equivalent of 600 million people streaming a high-definition video continuously.
Network Growth

The traffic is expected to increase some 29% each year (2011-2016)

Much of this increase comes from the delivery of video data

P2P traffic has become a smaller component of Internet traffic in terms of its current share

Video is being delivered by a set of protocols, typically coordinated by overlay solutions and CDN solutions

We will cover these in during the course
CDNs

Content Delivery Networks (CDNs) are examples of overlay networks that cache and store content and allow efficient and less costly way to distribute data in massive scale.

CDNs typically do not require changes to end-systems and they are not peer-to-peer solutions from the viewpoint of the end clients.
Challenges for Overlay Networks

The Real World. In practice, the typical underlay protocol, IP, does not provide universal end-to-end connectivity due to the ubiquitous nature of firewalls and Network Address Translation (NAT) devices.

Management and administration. Practical deployment requires that the overlay network has a management interface.

Overhead. An overlay network typically consists of a heterogeneous body of devices across the Internet. It is clear that the overlay network cannot be as efficient as the dedicated routers in processing packets and messages. Moreover, the overlay network may not have adequate information about the Internet topology to properly optimize routing processes.
Layering

- **Routing algorithms**: Dynamics of content demand and supply
- **Overlay Networks**: Dynamics of nodes (churn)
- **Network protocols**: Dynamics of the network, mobility, inter-domain routing

**Typical applications of overlay networks** include:

- Content search and file transfer.
- Distributed directories with efficient lookups.
- Content routing over the Internet.
- Publish/subscribe.
- Distributed storage systems.
- Substrates for multi-player games.

The two important categories for overlay networks are [28]:

- Unstructured networks that do not assume a particular way of distributing content to nodes and building links between nodes. Example systems include the classical P2P systems such as Gnutella and Freenet as well as gossip algorithms discussed later in this chapter.

- Structured networks that make more assumptions regarding the links between nodes and how content is placed on the nodes. Example systems are the DHT algorithms that will be presented next.
Overlay Nodes

The overlay nodes are expected to support the network in the following ways:

1. Providing infrastructure for the overlay network and supporting the execution of distributed applications.

2. Participating in the overlay network and perform routing and forwarding of messages and packets.

3. Deployed over the Internet in such a way that they are reachable by third parties.
Example Overlay: RON

- Consider a native link failure in CE
- Only overlay link AE is affected.
- The native path AE is rerouted over F (ACE → ACFDE)

![Diagram showing overlay routing]

Figure 3.1 Example of resilient overlay routing.

However, the overlay network must address concerns with node dynamics (churn) and security. On top of the overlay network, various pub/sub routing algorithms and systems can be developed. One example is the Scribe pub/sub system implemented on top of the Pastry DHT. The solutions need to consider the dynamics of content demand and supply. Therefore issues pertaining to dynamic operation need to be investigated at each layer.

3.2 Usage

Overlay networks can be used in different operating environments. The two typical environments are the clusters and the Internet (wide-area). Clusters can typically be assumed to be reliable, secure, and administered. In addition, they are not prone to network partitions or unpredictable operating conditions. Wide-area environments, on the other hand, lack these qualities. They are unreliable, unpredictable, insecure, prone to network partitions, and do not have centralized administration. DHTs balance hop count with the size of the routing tables, network diameter, and the ability to cope with changes.

Typical requirements for wide-area overlays include:

- Decentralized operation, in which the nodes co-operate and form the system without central coordination.
- Scalability to millions of nodes.
- Fault tolerance and support for churn.
Example Overlay: Session Initiation Protocol (SIP)

1. REGISTER
   Registrar at test.com

2. Upload location
   Location Server

3. SIP Request
   Proxy

4. Where is Bob?

5. Answer: at test.com

6. SIP request

Figure 2.13 SIP entities.

2.7.2 Method Types

There are six method types for request messages in the SIP specification. There are also additional methods defined in other RFCs that extend the core SIP standard.

- ACK that confirms that a final response has been received in a transaction.
- BYE that is sent by a User Agent Client to indicate to the server that it wishes to terminate the call.
- CANCEL that is used to cancel a pending request.
- INVITE that indicates the user or service is being invited to a session. The body of this message typically includes a description of the session with the Session Description Protocol (SDP).
- OPTIONS that is used to query a server about its capabilities.
- REGISTER that is used by a client to register an address with a SIP server.

2.7.3 Establishing a Session

The session establishment process begins with an INVITE message. This message is sent by a calling user to a callee user and the INVITE message invites the callee user to participate in a session. The message includes information about the session, for example the type of session and its properties. Typically, the Session Description Protocol (SDP) is used to describe sessions.
Network Invariants and Metrics

The correctness and performance of a routing algorithm can be analyzed using a number of metrics. Typically it is expected that a routing algorithm satisfies certain invariant properties that must be satisfied at all times. The two key properties are safety and liveness. The former states that undesired effects do not occur, in other words the algorithm works correctly, and the latter states that the algorithm continues to work correctly, for example avoids deadlocks and loops. These properties can typically be proven for a given routing algorithm under certain assumptions. Important metrics: shortest path, routing table size, path stretch, forwarding load, churn.
<table>
<thead>
<tr>
<th>Trend</th>
<th>Challenges</th>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2P</td>
<td>Growth in traffic, upstream bottlenecks</td>
<td>P2P caching, locality-awareness</td>
</tr>
<tr>
<td>Internet Broadcast</td>
<td>Flash crowds</td>
<td>P2P content distribution, multicast technologies</td>
</tr>
<tr>
<td>Internet Video-on-Demand</td>
<td>Growth in traffic, especially metropolitan area and core</td>
<td>Content Delivery Networks (CDNs), increasing network capacity, compression</td>
</tr>
<tr>
<td>Commercial Video-on-Demand</td>
<td>Growth in traffic in the metropolitan area network</td>
<td>CDNs, increasing network capacity, compression</td>
</tr>
<tr>
<td>High-definition content</td>
<td>Access network IPTV bottleneck, growth in VoD traffic volume in the</td>
<td>CDNs, increasing network capacity, compression</td>
</tr>
<tr>
<td></td>
<td>metropolitan area network</td>
<td></td>
</tr>
</tbody>
</table>
Overlay Applications

Typical applications of overlay networks include:

• Content search and file transfer.
• Distributed directories with efficient lookups.
• Content routing over the Internet including voice and video.
• Publish/subscribe and notification.
• Distributed storage systems.
• Multi-player games.
Terminology

- Peer-to-peer (P2P)
  - Different from client-server model
  - Each peer has both client/server features
- Overlay networks
  - Routing systems that run on top of another network, such as the Internet.
- Distributed Hash Tables (DHT)
  - An algorithm for creating efficient distributed hash tables (lookup structures)
  - Used to implement overlay networks
- Typical features of P2P / overlays
  - Scalability, resilience, high availability, and they tolerate frequent peer connections and disconnections
Peer-to-peer in more detail

• A P2P system is distributed
  – No centralized control
  – Nodes are symmetric in functionality

• Large faction of nodes are unreliable
  – Nodes come and go

• P2P enabled by evolution in data communications and technology

• Current challenges:
  – Security (zombie networks, trojans), IPR issues

• P2P systems are decentralized overlays
Characteristics of P2P systems

P2P can be seen as an organizational principle
Applied in many different application domains

Characteristics

Self-organization
Lack of central coordination
Resource sharing
Based on collaboration between peers
Peers are typically equal
Large number of peers
Resilient to certain kinds of attacks (but vulnerable to others)
P2P Volume

Estimates range from 20-40% of Internet Traffic

Latest estimates from Cisco suggest that video delivery is the growing and the share of P2P traffic is becoming smaller

P2P can be used for video delivery as well

.. And voice (Skype, P2PSIP)

Hundreds of millions of people use P2P technology today
Evolution of P2P systems

• ARPA.net had P2P like qualities
  – End-to-end communication, FTP, USENET, ..
  – Today’s BGP is P2P
• Started from centralized servers
  – Napster
    • Centralized directory
    • Single point of failure
• Second generation used flooding (Gnutella v0.4)
  – Local directory for each peer
  – High cost, worst-case O(N) messages for lookup
  – Third generation use some structure (Gnutella v0.7)
• Research systems use DHTs
  – Chord, Tapestry, CAN, ..
  – Decentralization, scalability
• Some recent CDNs and content delivery systems exhibit P2P features (P2P assisted CDN)