Communication Networks
Chapter 11.b: Routing

- Tasks of the Network Layer
- Components and Functions of a Router
- Forwarding and Routing
- Autonomous Systems
- Routing Approaches

The Internet Layer

Transport Layer (TCP, UDP, ...)

Internet Layer

Routing Protocols
- Path finding
- RIP, OSPF, BGP, ...
- AODV, OLSR, ...

Routing Table

IP
- Addressing
- Packet format
- Packet processing

ICMP
- Error/state messages
- Control

ARP/RARP
- Address translation IP ↔ MAC

IGMP / DHCP / NAT / CIDR / RSVP / ...

Network to Host Layer
Repetition: IP Routing Services

- **Source Routing (IPv4)/Routing Header Extension (IPv6)**
  - Sender defines the complete path through the network
- **Problems**
  - Might result in suboptimal paths
  - Security issues in IPv6
    - Direct path definition enables *denial of service* attacks targeting two or more routers
    - Traffic is multiplied by a forced loop between the systems
    - Function defined as deprecated in IPv6 by RFC 5095
- **Alternative routing approaches required**

Repetition: IP Forwarding

- **Definitions**
  - All nodes take forwarding decisions
  - Only the destination IP address is considered traditionally
    - Today, partially additional policies
- **Two general forwarding types**
  - Destination node is part of the same subnet
    - Direct forwarding
  - Destination nodes is reachable via Router/Gateway nodes only
    - In-direct forwarding via multiple Routers
    - Address packet to next Router on layer below IP
What is a Router?

- A device to couple networks on the network layer
- Allows
  - Communication of end systems via one or more subnetworks
  - Routing
  - Forwarding of packets based on unique, ideally hierarchical network addresses
  - Segmentation and reassembly in IPv4 to support different PDU sizes of underlying MAC layer
  - Security mechanisms to control the access to the network based on IP addresses (Firewalls)

Packet Processing
Architecture of a Router

Main features:
- Each network interface has its own layer 1 and layer 2 entity
- The network protocol is usually identical for all network interfaces; in special cases, there might even be several network protocols in a router
- Network layer entity is responsible for forwarding the packets according to their globally unique destination address
- Control entities realize routing protocols, control message protocols or management protocols

General Router Types

**WiFi/DSL Router**
- Used by end-users to provide internet access
- Internet access for private home network
- Support multiple devices within the home network

**Enterprise Router**
- Used for campus area networks (e.g. a university or company)
- Partially employed for network structuring

**Wireless Router Node**
- Used in mobile ad hoc networks
- Might bridge multiple access technologies

**Edge Router**
- Used by Internet Service Providers (ISPs)
- Connects ISPs to customer networks
- Supports different access technologies

**Core Router**
- Used within the core network
- Allows multiple Terabit/s
- High availability of 99.999 % or higher
- Completely redundant hardware/software components
Example Routers

- At home
- Within the core network

Source: https://de.wikipedia.org/wiki/Router

Fundamental Router Tasks

- **Routing**
  - Finding paths to transport packets through a network
  - Mapping reachable subnetworks based on topology
  - Selection of the best path based on IP-Prefixes

- **Forwarding**
  - Selection of next hop for all packets at a node
  - Actually the selection of the appropriate outgoing interface
  - Local decision at each node
  - Has to be done as fast as possible
Control and Data Path

- **Data Path** within the network layer
  - Forwarding data packets
- **Control Path** within the layer above
  - To exchange routing control information
- **Routing Information**
  - Exchanged/colllected by routing protocols
  - Stored in **Routing Table**
- **Routing Algorithm** manages Routing Table
  - Insert/Delete/Update entries
  - Based on gathered Routing Information

Subfunctions of Router Tasks

- **Basic forwarding functions**
  - Header validation
  - TTL verification
  - Address lookup
  - Fragmentation (IPv4 only)
  - Processing of IP options
  - Error notifications via ICMP
- **Complex forwarding functions**
  - Classification
  - Filtering
  - Prioritization
  - Transformation
- **Routing functions**
  - Path calculation
  - Routing table updates
  - Routing protocol execution
- **Management functions**
  - System configuration
  - Monitoring
Router Components

Two Concepts

**Functional perspective:**
modules realizing functions required for packet forwarding

- Common modules:
  - Network interfaces
  - Forwarding engine
  - Queue manager
  - Traffic manager
  - Backplane/switching fabric
  - Route control processor

**Architectural perspective:**
- Hardware-oriented description
- Which hardware components exist and how are they linked?
- Where and how are modules implemented?

Packet Flow Through a Router

[Diagram showing packet flow through a router with labels for control path and data path, and components like RIB, FIB, Buffer Memory, Queue Management, Traffic Management, and Backplane.]

**Key Terms:**
- RIB: Routing Information Base
- FIB: Forwarding Information Base
Router Hardware Components

Cisco CSR-1 Router Architecture


Communication Networks - 11. The Internet

Forwarding Sequence

1. Receive packet from L2 instance of incoming interface
2. Store packet in incoming buffer
3. Validate packet header
4. Process IP options
5. Extract destination IP address
6. Lookup next hop in routing table: destination, outgoing interface(s)
7. Fragmentation \( (\text{if required/possible}) \)
8. Update packet header: decrement TTL, update checksum
9. Forward packet to outgoing interface (possibly multiple ones)
10. Store packet in outgoing buffer
11. Send packet to L2 instance of outgoing interface
12. Send error message \( (\text{if required}) \)
Router Performance Metrics

**Throughput**
- \( D_R = N_I \cdot R_I \)
- \( N_I \): number of interfaces
- \( R_I \): data rate per interface
- unit: bit per second

**Processed packets per second**
- \( PR_R = \frac{D_R}{S_P} \)
- \( S_P \): packet size
- Example: 2 billion packets per second for 640 Gbps throughput and 40 byte packet size
  → Time to finish complete processing: 8 ns per packet

Common packet sizes
- 40 byte: TCP Acknowledgements
- 64 byte: ICMP Echo Requests
- 576 byte: IPv4 Internet Path MTU
- IPv6 minimum MTU
- 1280 byte: IPv6 minimum MTU (RFC 2460)
- 1500 byte: Max. Ethernet Payload

Routing Table Contents

- Path to another network
  - Corresponds to a complete subnetwork
  - Most common entry type
- Path to another node
  - Specific path to that node
- Typically
  - 200,000 – 1,000,000 entries
- Default path
  - Used if no other path is available
  - Default Gateway
- Lookback address
  - Forward packets to the Loopback interface of the current node
  - 127.0.0.1 in IPv4
**IP Forwarding Decision**

- Performed at any participation node
  - Even if not a dedicated router

---

**Routing Table – Principle (1)**

<table>
<thead>
<tr>
<th>Routing Table in Router B</th>
<th>Routing-Table in Router A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>from</strong></td>
<td><strong>via</strong></td>
</tr>
<tr>
<td>A</td>
<td>- (l₁)</td>
</tr>
<tr>
<td>C</td>
<td>- (l₂)</td>
</tr>
<tr>
<td>D</td>
<td>A (l₃)</td>
</tr>
<tr>
<td>E</td>
<td>A (l₄)</td>
</tr>
<tr>
<td>F</td>
<td>- (l₅)</td>
</tr>
<tr>
<td>G</td>
<td>C (l₆)</td>
</tr>
<tr>
<td>H</td>
<td>F (l₇)</td>
</tr>
</tbody>
</table>

**A-H:** Router  
**l₁-l₇:** Incoming/Outgoing Links
### Routing Table – Principle (II)

#### Link Failure

- **Routing Table in Router B**
  - from | via
  - A   | \((l_1)\)
  - C   | \((l_2)\)
  - D   | A \((l_1)\)
  - E   | A \((l_1)\)
  - F   | A \((l_2)\)
  - G   | C \((l_2)\)
  - H   | A \((l_1)\)

- **Routing Table in Router A**
  - from | via
  - B   |
  - C   |
  - D   |
  - E   |
  - F   |
  - G   |
  - H   |

**A-H:** Router
**\(l_1-l_8\):** Incoming/Outgoing Links

### Updating Routing Tables

- **Static**
  - Manual entries

- **Via routing approaches**
  - Based on collected routing information
  - Path selection via specialized routing algorithm

- **Triggered by ICMP messages**
  - In case of errors
  - In case of TTL expiration
Routing Approaches

Task
- Decide which outgoing interface is appropriate for the incoming packet (message)

Forwarding Goals
- Low average packet delay
- High network throughput

Challenges
- Reliable packet delivery
- Low additional load due to routing information exchange
- Fast reaction on topology changes
  - Especially for mobile nodes
- Have up-to-date and complete information on the current network state
- Resource efficiency
- Loop free operation
- Consider application requirements

Routing Approach Aspects

- Used routing information
  - Which addresses are used?

- Route determination
  - How are paths obtained?
  - Who identifies a route?

- Collection/exchange of routing information
  - How?
  - When?

- Enforced routing metrics
  - Based on current network state?
  - Without considering the network state?
Route Determination

- How are paths obtained?
  - Static: all paths are predefined and fixed
  - Adaptive: paths can change during network operation

- Who identifies a Route?
  - Possible nodes
    - Source or specialized router
  - Algorithms
    - Distributed
    - Centralized
    - Hierarchical

Exchange of Routing Information

- How?
  - Flooding
  - Selectively

- When?
  - Periodically and after changes → proactive
  - On demand and after changes → reactive
Routing Metrics

General
- Load balancing
- Minimum hop count
- Quality parameters
  - Bit rate, latency, throughput, ...
- Security
- Cost
- Combination of multiple weighted values in one utility function
- Policies
- ...

State dependent
- Fixed cost for possible connections
- Number of packets waiting for transfer
- Error rate
- Packet delay of one link
- Traffic type (dialog, batch)
- Priorities
- Service support by Router

Network Dependent Routing

- Internet
  - Intra Domain Routing (e.g. RIP, OSPF, IS-IS, …)
  - Inter Domain Routing (e.g. BGP, …)
- Mobile Ad-hoc Networks (MANETs)
  - Routing directly between any mobile node
  - e.g. OLSR, AODV, DYMO, DSR, …
- Wireless Sensor Networks (WSNs)
  - RPL
- Delay Tolerant Networks (DTNs)
  - Epidemic, Prophet, Rapid, MaxProp, …
The Structure of the Internet

Autonomous Systems and Routing

Inter Domain Routing
Intra Domain Routing
Autonomous System (AS)

- **Corresponds to an administrative domain**
  - The Internet is build from many separate networks
  - AS build the organizational structure of the Internet
  - Each AS has a globally unique AS number (ASN)

- **Goals**
  - AS use a freely chosen routing protocol internally
  - AS define policies for transit traffic
  - Internal AS structure is hidden

- **Traffic Types**
  - **Local Traffic**
    - Packets from or for nodes within the AS
  - **Transit Traffic**
    - Transferred through an AS

- **AS Types**
  - **Stub AS**
    - Has connections to one other AS only
  - **Multihomed AS**
    - Connects multiple AS
    - Does not handle transit traffic
  - **Transit AS**
    - Connects multiple AS and handles transit traffic

Autonomous Systems in Germany

Intra Domain Routing

- Applied within an AS
- Goals
  - Focus on optimal routes
  - Stable routes

Two fundamental algorithm types

- **Distance Vector algorithms**
  - Require local knowledge only
  - Can produce loops

- **Link State algorithms**
  - Require global knowledge on complete topology

Routing Information Protocol, RIP

- Intra domain routing approach based on distance vector algorithms
  - Bellman Ford Algorithm
  - Supports Split Horizon and Poison Reverse
  - Control message exchange via UDP
- **Idea**
  - Each node stores one entry with the best distance to the destination in its routing table
  - Entry structure: distance, destination, and outgoing interface
- **Algorithm:**
  1. Request the table from all neighbors
  2. Calculate table updates
  3. Send periodic advertisement packets to all neighbors
RIP

**Characteristics**
- adaptive
- distributed path calculation on dedicated router nodes
- Periodic routing information exchange
- Utility function only depending on distance

**Problems (without split horizon, poison reverse)**
- Does not consider current network state
- Path length limited to maximum 15 hops
- Long convergence time in case of updates
  → Count To Infinity problem
- Outdated

---

Open Shortest Path First, OSPF

**Intra domain routing approach based on link state algorithms**
- Dijkstra Algorithm
- Control message exchange via IP

**Algorithm:**
1. Detect neighbors via Hello packets
2. Measure cost to all neighbors via Echo packets
3. Prepare Link State packet containing information about all neighbors
4. Exchange Link State packets with all neighbors
5. Calculate optimal path to all neighbors based on collected information

---

Note: always selects the minimum value

Distance in hops from A
1. exchange
2. exchange
3. exchange
4. exchange
5. exchange
6. exchange
7. exchange
8. exchange

Depending on upper limit
**OSPF Algorithm Details**

- Neighbor discovery
  - HELLO packet
- Measurement of delay / costs for each neighbor
  - ECHO packet
- All information is entered in a LINK STATE packet
  - Sender, list of neighbors with the according delay, age
  - periodical or event-driven (e.g. new neighbor, failure, etc.) transmission
- Transmission to all neighbors
  - Usually flooding, but with enhancements: deletion of duplicates, discarding old packets etc.
- Computing the shortest path to all routers (e.g. Dijkstra)
  - Quite complex, but there are optimizations

---

**OSPF**

- **Characteristics**
  - adaptive
  - distributed path calculation on dedicated router nodes
  - periodic routing information exchange
  - utility function depending on current state of the link
  - path determination to all other routers

- **Pros**
  - considers current network state
  - authenticated control messages
  - supports Areas for improved scalability

- **Cons**
  - requires expensive recalculation after every update
Border Gateway Protocol, BGP

- Inter domain routing approach based on path vector algorithm
  - store list with paths to all other AS
  - control message exchange via TCP

- Idea
  - routes correspond to path to other AS
  - internal details of AS are unknown
  - routing table contains aggregated paths to all other AS
  - routing algorithm not defined
    - Selected based on policies
    - Examples: Hot-Potato, Cold-Potato, Shortest Path, …

BGP

- Characteristics
  - (adaptive)
  - distributed path calculation on dedicated router nodes
  - no periodic routing information exchange, only event based updates due to failures
  - utility function depending on number of AS in the path

- Cons
  - does not consider current link state
  - does not support load balancing – always selects one path only
  - very large routing tables
  - security issues due to missing authentication
Additional Algorithms for BGP

**Hot-Potato**

- **Goals:**
  - Forward transit traffic fast to another AS
  - Preserve internal resources

- **Approach:**
  - Forward packet to interface with shortest outgoing queue

**Cold-Potato**

- **Goals:**
  - Keep control over traffic as long as possible
  - Enforce *Quality of Service* parameters

---

**BGP Policy Examples**

- prohibiting foreign packets to be traversed through the network
  - although the path might be shorter
- political restrictions
- company policies
  - company X does not want to pay for packet transmission through the network of company Y
Mobile Networks

**Ad Hoc Networks (MANETs)**
- Direct wireless communication between mobile nodes
- Dynamic network topology
- No previous planning possible/required
- High flexibility
- Complex end-systems have to provide all functions
- Decentralized management → self-organization

**Wireless Infrastructure Networks**
- Wireless communication between mobile end system and base station / access point
- Base stations connected via wired network → connecting multiple cells to a logical network
- Requires careful network planning
- Infrastructure supports additional services and QoS guarantees
- Complex base stations, but simple end nodes
- Centralized management

Routing Challenges in MANETs

- Mobile nodes
  - Possibly different node speeds
- Frequent, fast topology changes
  - Neighbor discovery
  - Age of routing information
- Limited resources
- No dedicated outgoing interfaces
  - Incoming and outgoing packets via the same wireless interface
  - Broadcast to all neighbors
**Ad hoc On-Demand Distance Vector, AODV**

- Intra domain / Ad hoc routing approach based on distance vector approach
  - Dijkstra Algorithm
  - Control message exchange via UDP

**Algorithm:**
1. Broadcast *Route Request* (RREQ) for path towards destination
2. Unicast *Route Reply* (RREP) to inform requesting node about path
3. Store route in Routing Table
4. Intermediate nodes can answer RREQ packets if their Routing entry is still valid
5. *Route Error* (RERR) to inform nodes about missing neighbors

**AODV**

- **Characteristics**
  - adaptive
  - distributed path calculation on dedicated router nodes
  - reactive routing information exchange on demand

- **Pros**
  - low additional network load since path are searched on demand only

- **Cons**
  - higher initial delay until path is found
Optimized Link State Routing (OLSR)

- Intra domain / Ad hoc routing approach based on link state approach
  - Control message exchange via UDP

- **Algorithm:**
  1. Discovery of 1 or 2 hop neighbors via HELLO packets
  2. Distribute Routing Information throughout the network via TC packet (Topology Control)
  3. Select Multipoint Relays for each node N
  4. N can reach all 2-hop-neighbors via Multipoint Relays
  5. Multipoint Relays are responsible for the distribution of Broadcast messages and TC messages

---

OLSR

- **Characteristics**
  - adaptive
  - distributed path calculation on dedicated router nodes
  - periodic routing information exchange
  - proactive path determination

- **Pros**
  - route immediately available
  - no additional path finding delay

- **Cons**
  - higher load on network due to proactive control message exchange
Hybrid Routing

In general

- Combination of two or more different aspects of a routing approach

Routing for Specialized Use Cases

- Geocast
  - Routing based on node positions
  - Requires exact positioning information (e.g. via GPS/GLONASS)

- Multicast
  - Simultaneous delivery of messages to multiple receivers
  - Requires group membership information

- Opportunistic Approaches
  - Exploit „fastest“ delivery of the packet
  - Based on broadcast nature of the medium
Geocast Routing

- Idea
  - Forwarding based on node position
  - Select closest node to destination as next hop

- Variants
  - Destination is a node
  - Destination is a geographic region → receivers are all nodes within that region

- Problems
  - Age of positioning information
  - Dead ends possible (e.g. closest node cannot forward the message further)

Multicast Routing

- Idea
  - Each packet contains group information
    - List of required receivers
    - Bit identifier allowing to identify all receivers
    - Group address
    - Determination of one or multiple links to reach all group members

- Problems
  - Who is identifying group members / receivers?
  - Age of routing information
Opportunistic Routing

- **Idea**
  - Flood messages to all neighbors
  - Only the "fastest" neighbor forwards message again by flooding again

- **Problems**
  - What happens if multiple nodes forward the message?
  - What happens if the fastest node forwards the message in the wrong direction?

Delay Tolerant Routing

- **Network**
  - Frequent disruptions due to mobility, destruction, energy saving
  - Long delay links due to distance
  - No end-to-end connections possible

- **Idea**
  - Store messages until new contact occurs → message switching
  - Flood messages or use other prediction methods

- **Problems**
  - How to determine the next contact?
  - How to handle scarce resources?
  - What happens if no contact opportunities arise?
Epidemic Routing

- **Idea**
  - Flood messages to all nodes in the network

- **Algorithm in case of a contact**
  - Exchange a summary of all stored messages
  - Compare summary of peer with own list
  - Exchange missing messages bi-directionally

Epidemic

- **Characteristics**
  - adaptive
  - distributed path calculation on dedicated router nodes
  - no routing information exchange
  - uses names instead of IP addresses as routing information

- **Pros**
  - simple algorithm
  - ensures minimum delay

- **Cons**
  - high resource consumption in terms of memory and unnecessary transfers
Sensor Networks Challenges

- Limited node resources
  - Processing capabilities
  - Memory
- Typically tree-like topology
  - Dedicated node types: source and sink
  - Partially fixed topology
  - Load distribution not balanced
  - Supports large number of nodes → scalability
  - Different node density possible
- Different network goal
  - Not to transport as much data as possible
  - Reliably detect events
  - Reliably report sensor reading regularly

Routing Protocol for Low-power and Lossy Networks (RPL)

- Idea
  - Transfer packets to a sink node
  - Build target-oriented weighted graph to find an optimal path
- Algorithm
  - Child nodes inform parents which other node are reachable via this child
  - Select „best“ parent based on distance vector algorithm
RPL

- **Characteristics**
  - adaptive
  - distributed path calculation on dedicated router nodes
  - periodic routing information exchange
  - periodic status recalculation

- **Pros**
  - supports energy saving mechanisms like sleeping

- **Cons**
  - requires much memory on the nodes
  - optimized to support central sink nodes
  - increased overhead for direct source communication due to multiple trees

RFCs

- **MILLS, D.L.:** Exterior Gateway Protocol formal specification, April 1984 (RFC 904)
- **ORAN, D.:** OSI IS-IS Intra-domain Routing Protocol, February 1990 (RFC 1142)
- **MALKIN, G.:** RIP Version 2 Protocol Analysis, November 1994 (RFC 1721)
- **MALKIN, G.:** RIP Version 2 Protocol Applicability Statement, November 1994 (RFC 1722)
- **HAWKINSON, J.; BATES, T.:** Guidelines for Creation, Selection, and Registration of an Autonomous System (AS), 1996 (RFC 1930)
- **MOY, J.:** OSPF Version 2, April 1998 (RFC 2328)
- **MOY, J.:** OSPF Standardization Report, April 1998 (RFC 2329)
- **MALKIN, G.:** RIP Version 2, November 1998 (RFC 2453)
RFCs


References