1 Performance at Lower Network Layers

1.1 Network Layer - Overview

Goal:

- Transports data from src to dest, across multi. hops
- Every major net. component has net. layer concerns

Design Principles:

- Services provided by network layershould be indep. of net. topology
- Transport layer should be shielded from member, type and topology of net. components
- Network addresses avail. to the transport layer should use uniform numbering plan

Service Types:

- Connectionless
  - e.g. IP: Internet PRotocol
- Connection-oriented
  - ATML Asynch. Transfer Mode Protocol

1.2 Internet

1.2.1 Characteristics

- Initial Design
  - Connection-oriented
  - Peer-peer topology
  - Circuit swithing-based communication
  - e.g. Telegraph and telephone
- Current Design
  - Best Effor Service
    - More appropriate for data transfers
    - No RT requirements
    - End-points can adapt to net. conditions, if they want/need to
  - Connectionless packet-switching based
    - Preferred to circuit-switvhing
    - No set-up delay
    - No blocking (fire and forget, forward if possible, deliver as received)
      - No guarantee that any data reaches dest.

        - Flex. in TX bit rates
            - Circuit-switching usually has few pre-determined bit-rates
        - No stable "path"
            - More reliable (route around problems
        - More efficient use of net. resources when traffic bursty

1.3 IP-based Network Layer

- Based on IP
- Connectionless: uses datagram packet switching
- Open design
- Open implementations
- Open standardisations process
- Independent of physical medium
- Scalable: as evidenced by its growth
- Extensible: protocols have evolved over time, as problems arose and/or req. changed

1.4 IP Router Architecture

1.4.1 Router Basic Functions

- Forwarding
    - Move pkts from routers input to output port
- Routing
    - Determine route taken by pkts from src to dest.
- Routing Activity (When datalink frame arrives)
    - Line card applies datalink layer logic to ensure frame is valid and pkt was successfully received
    - If pkt arrival rate > routers forwarding cap. pkt queued, waits for processing
    - If queue buffers full, pkt discarded
    - If space in queue, after waiting, validity check performed on IP header
    - If dest address non-local host, routing table lookup performed to determine how to forward pkt.
    - Pkts classified into predefined service classes (if defined)
    - TTL field decremented, new header checksum computed, pkt sent to approp. output port
    - Datalink layer login on output ports line card inserts datalink layer header, TX pkt inside a frame
    - If process fails, error msg sent to pkts sender

1.5 Internet Protocol Version 4 (IPV4)

1.5.1 IPv4 Header

- Ver indicates IP V num.
- ID is the same for all fragments of given datagram
  - Fragmentation of a datagram needed when an intermediate net. has a max frame size too small to carry datagram
- Fragmentation offset indicates place of each fragment in datagram
- Time-to-Live decreased by 1 at each hop
- Options field carrys fields that control routing, timing, management, security, etc.
  - Padded to a multiple of 4 bytes

1.5.2 IPv4 Address Format

1.5.3 Special IPv4 Addresses

Loopback useful for debugging/testing network software as IP pkts with loopback addresses are processed by the machine which generated them as if they were incoming pkts

1.6 Internet PRotocol Version 6 (IPV6)

1.6.1 Motivation

- IPv4 Limitations
  - Limited address space
    * Despite subnetting, Network Address Translation (NAT)
    * Classless Inter-Domain Routing (CIDR)
  - No special tratment for real-time traffic
    * Despite IPv4 Type of SErvice field which is ignored by routers
  - No wide usage of security issues
    * Despite IPv4 security ffeatures which are not widely used

1.6.2 Action

- In 90s, following an open design process IPv6 was standardised

1.6.3 ISsues

- IPv6 and v4 not compatible
- IPv6 deployment takes time
  - Both will coexist for a long time

1.6.4 Features

- Increased address space
  - Uses 128 bit address
  - 2^{128} addresses
  - $5 * 2^{28}$ addresses for each person on earth
- Simple pkt header
  - Routers don't do fragmentation
  - No header checksum to be checked
- Support for more options
  - e.g. routing, hop-by-hop, fragmentation, etc.
  - extension headers can be present
- Support for per-flow handling and traffic classes
  - Flows defined by src address, dest addres and flow num.
- Mandatory authentication and security
  - Internet Protocol Security (IPsec) was specially developed (then deployed in IPv4)
- Compatability with existing TCP/IP protocol stack
  - E.g DNS, TCO, UDP, OSPF, BGP, etc.

1.7 Other Network Layer Protocols

1.7.1 Motivation

- IP is in charge with data transport
- IP reqs. support from other transport layer protocols
  - Control functions (ICMP)
  - Multicast signalling (IGMP)
  - Routing table setup (TIP, OSPF, BGP, PIM)

1.7.2 ICMP

- Internet Control Message PRotocol facilitates:
  - Error reporting
  - Simple Queries
- ICMP msgs. carried by IP datagrams

1.8 Internet Control Message Protocol (ICMP)

1.8.1 ICMP Message Format

- Header (4bytes)
- Fields:
  - Type (1 byte): msg type
  - Code (1 byte): msg subtype
  - Checksum (2 bytes): calculated for entire msg
- Payload (min 4 bytes)
- Content
  - Related to msg type
  - If no data, content includes 4 bytes set to 0
- Note:
  - ICMP msgs have min. length of 8 byte

1.8.2 ICMP Message Types

- Query
  - Request-reply
  - e.g. Echo request, echo reply, timestamp request, timestamp reply, router solicitation, router advertisement
- Error reporting
  - Informs about error in transporting data
  - Sent by routers to hosts or routers
  - Processed at higher layers (e.g. application)

1.9 Routing Algorithms

- Definition
  - Determine route taken by pkts from src to dest
- Desirable Properties
  - Correctness
  - Simplicity
  - Efficiency
  - Robustness
  - Stability - Routing alg. reaches equilibrium in reasonable time
  - Fairness
  - Optimality

- Least-Cost Routing
  - Cost - A value assigned to each link in the net.
  - Cost of a route - Sum of values of all routes links
  - BEst route - Route with lowest cost
- Meaning of cost
  - 1 for each link - best route = fewest hops
  - Financial cost of using link - best is cheapest route
  - delau on link - bes is min.-delay route
  - pkt tx time on link - best is max.-bw route
  - Some comb. of these
- Pkt fprwarding
  - Towards the best route

1.9.1 Types

- Non-adaptive or Static
  - Routing decisions pre-determined, not based on measurements (or estimates) of current net. topology and traffic load
- Adaptive
  - routing decisions may be changed when net. topology and/or trafic load change
    * Extreme case: select new route for each pkt
    * May get info from neighbouring routers, or from net. routers
    * Routes are changed
      · Periodically
      · When topology changes
      · When traffic load changes significantly
- Centralised
  - Routing decisions taken in centralised manner for whole network
- Distributed
  - Routing decisions taken separately by each router, independent from neightbours

1.9.2 Decentralised Routing Algorithm

- Distance-vecor
  - Each router exchanges info about entire net. with neighbour routers at regular intervals
  - Neighbour - connected by direct link
  - Regular interval - every 30 seconds
- Link-state
  - Each router exchanges info about neighbourhood with all routers in net. when there is a change in the topology
  - Neighbourhood of router - set of neighbour routers
  - each routers neighbourhood info is flooded through the net.
  - Change in topology occurs hwen:
    * Neighbouring router not accessibe anymore
    * New router has been added
- Note
  - Link state algorithm converges faster and therefore more widely used

1.10 Distance Vector Routing

Principle

- Assume several LANS represented by "cloudsänd routers/gateways by "boxes"
- Num. in each cloud represents net. ID
- Letter in each box represents router (or gateway) names
- Every router sends its info to its neighbours
- Each neighbor router adds this info to its own and send to its neighbours
- In time all routers learn about the net. struct.
- Each router stores info. about the net. in its routing able
- Routing table includes
  - Net. ID = final pkt dest
  - Cost = Num of hops from this router to final dest.
  - Next hop = neighbouring router to which pkt should be sent
- Initially each router knows net IDs of the net. to which it is directly connected only

1.10.1 Router Table Update

- Distributed Bellman-Ford Algorithm
  - Add 1 to cost of each incoming route (each neighbour 1 hop away)
  - If new dest learned, add its info to routing table
  - If new info received about existing dest.
    * If next hop field is same, replace existing entry with new info even if cost is greater
    * If next hop field not same, replace existing entry with new info if cost is lower

1.11 Link-State Routing

Principle

- Each router sends info about is neighbourhood to every other router
- Each router updates its info about the net. based on info received
- In time, all routers learn about net. struct
- Makes use of link costs (usually a weighted sum of various factors
  - e.g. traffic level, security level, pkt delay
- Link cost is from router to net. connecting it to another router
  - When pkt sent in LAN, every node - including router - can receive it
  - No cost assigned when going from a net. to router

Routing tables

- All routers get their info about their neighbourhood by sending short echo pkts to their neighbours, monitoring response
- All routers share info about their neighbours by sending link-state pkts to all routers in network (flooding)
- Pkts include
  - Advertiser: sending reouter ID
  - Network: Dest. network ID
  - Cost: Link cost to neighbour
  - Neighbour: Neighbour router ID
- Every router prepares a link-state pkt and floods it through the net.
- When router receives all these pkts. it can save the data in a link-state DB
- Assuming that every router receives same pkts, same content will be found in all link-state DBs
- Using info from DB, each router can fill it's routing table

1.11.1 Dijkstra's Shortest Path Algorithm

- ID all link costs in net. either from link-state DB, or using fact that cost of any link from a network to a router is 0
- Build shortest-path spanning tree for router running the alg.
  - Tree has route from router to all possible dest. and no loops
- Router is root of its shortest-path spanning tree
- Node either a net. or a router: nodes connected by arcs
- Algorithm keeps track of 2 sets of nodes and arcs, Temp. and Perm.
- Initially router is in Perm. set and Temp. set contains all neighbour nodes of router itseld, arcs connecting them to router
- Identify Temp. node whose arc has lowest cumulative cost from root: move to Perm set
- All nodes connected to new PErm node and not already in Temp. set along with their arcs, moved to Temp.
- Any node already in Temo set has lower cumulative cost reom root by using a route passing through the new Perm node, this new route replaces the existing one
- Repeat until all nodes and arcs are in Perm set
- NoteL even if all routers link-state DBs are identical, tree determined by routers are different

1.11.2 Building the Routing Table

- Once a router has found its shortest-path spanning tree it can build its routing table
- In large net. memory required to store the link-state DB and the computation time to calc. the link-state routing table can be significant
- In practice, since link-state pkt receptions not synchronised, routers may be using different link-state DBs to build their routing tables
- Result accuracy depends on how different the various routers DB content is

1.12 Routing in the Internet

1.12.1 Centralised Routing

- Initially Internet buiilt around Core system which enabled interconnectivity via core gateways
- Routing info collected, exchanged between core Gateways using Gateway-Gateway protocol

- Routing data processed by Core and result were distrib. back to external routers
- Major weakness of model - Scalability and vulnerability

1.12.2 Decentralised routing

- Internet build as a set of hierarchical inter-connected independent network groups known as Autonomous systems (AS)
- Two level routing
  - Intra-AS: each AS responsibe for own routing
    * Major protocols used in practice
    * Routing Info Protocol (RIP) - based on distance vector alg.
    * Open Shortest Path First (OSPF) - based on link-states alg.
  - Inter-AS: enables routing between AS
    * Major protocol used in practice
    * Border Gateway Protocol

1.13 Multicast Routing

- Definition
  - Delivery of a copy of a packet to a group of receivers
- Types
  - Multicast unicast
    * Mult. pkts travel from src to each dest.
  - Multicast
    * Single pktstravel on common routes
- Multicast in practice
  - Requires multicast addresses for multicast groups
    * Start with "10ïn binary
  - Requires multicast enabled routers
    * Maintain and pass lsit of addresses assoc. with multicast group address
  - Requires multicast routing protocols
    * Distance vector mulicast routing protocols
    * Multicast open shortest path first protocol
  - Requires multicast routers to know about their own multicast groups
    * Internet Group Management Protocol

1.14 Performance Issues

1.14.1 Distance-Vector Routings count-to-infinity problems

- Slow convergence in some conditions
- Slow reaction to link/router failure as info travels in small steps
- Many ad-hoc solns. have been tried, but either also fail to solve count-to-infinity problem or are hard to implement

1.14.2 Link-State Routing Performance

- Link costs can be configured in OSPF (hop, reliability, delay, cost, bw)
- Large mam. req.
- Dijkstra alg. computations are highly processor intensive
- High BW req. if network topology changes often

1.14.3 Need for Intra- Inter-AS routing

- Policy
  - Inter-AS: Concerned with policies
  - Intra-AS: Under same admin, control so policy is less important
- Scalability
  - Inter-AS: Scale for routing among large num. of net.
  - Intra-AS: Scalability less of a concern
- Performance
  - Inter-AS: difficult to focus on performance metrics
  - Intra-AS: highly focused on performance metrics and costs

1.15 Wireless Routing PRotocols

1.15.1 Classification

- Topology-Based Routing
  - Table-Driven (proactive)
  - On-Deman (Reactive)
  - Hierarchical Routing
- Location-based routing
  - Greedy Routing

1.15.2 Table-Driven (Proactive) Routing

- Based on distance-vector and link-state protocols
- Nodes maintain routes to other nodes
- Periodic or event triggered route updates
- Relatively low latency, routes known in advanfe
- Higher overhead and longer route convergence

1.15.3 On-Demand (Reactive) Routing

- Src node inits routing discovery on demand
- Only active routes maintained
- Relative reduced routing overhead
- Long delays when new routes fount

1.15.4 Hierarchical Routing

- Net. divided into clusters
- Nodes talk to cluster head only
- Better scalability (descreases routing overhead), unfair use of resoutces

1.15.5 Location-Based (Geographic Routing

- Routing performed according to position of node
- Routing overhead can be small but optimal routing may not be found

1.15.6 Destination Sequenced Distance Vector (DSDV)

- Proactive routing protocol
- Each node maintains routing table with entries for each node in net.
  - dest addr, seq num., next-hop, hop-count)
- Nodes transmit pkts according to routing table
- Each node has seq num, updated when routing info changes (new node joins, line break)
  - Used to avoid routing loops
- Each node periodically broadcasts routing table updates

1.15.7 Dynamic Source Routing (DSR)

- Reactive protocol
- Srv wamts tp TX, does not know route to dest, inits route discovery
- Route request pkt broadcast, once dest receives, send back route reply, in pkt header ID's each forwarding hop in next node field
- Entire route stored in pkt headers
- At nodfes, route cache used to store most recent routes

1.15.8 Ad Hoc On-Demand Distance Vector (AODV)

- Essentially combo of DSR and DSDV
- DSRs on-demand mechanisms for route discovery and route maintenance
- Uses DSDVs table of precursor, next hop for each route during hop-by-hop routing and sequence numbers (to prevent loops)
- Improve DSR by keeping routing tables at nodes (pkts do not contain entire route)
- Routing table entries have lifetime in contreast to DSR cache

1.15.9 Temporally-Ordered Routing Algorithm (TORA)

- Adaptive routing protovol for muli-hop net.
- Designed to min comms overhead via localization of algorighmic reaction to topological changes (distribd execution)
- Uses directed acyclic graphs instead of shortest path soln
- Each node assigned unique height, pkts flow from high noes to low nodes along path towards dest.

1.16 Performance Issues

1.16.1 Throughput

- Throughput of DSDV decreases drastically with increases in mobnility. DSR outperforms all other protocols

1.16.2 Overhead

- In general routing overhead increases with mobility (topology changes)
- For DSR overhead dependent on num of diff. routes
- For DSDV overhead higher as routing tables need to be maintained

1.16.3 Delay

- DSDV delay lowest, constant (see routing tables)
- DSR high delay (see reactive protocols)
- TORA highest delay (see short-lived and long lived loops)

1.16.4 Optimality

- DSDV and DSR find optimal paths
- TORA and AODV use suboptimal paths even under low mobility

1.17 Quality of Service Support

1.17.1 Buffering

- Significant traffic burstiness when Tx over net.
- Buffering reduces loss, enables control over Tx rate

1.17.2 Packet Scheduling

- Enables selection of packets for differentiated Tx
- Arrival based: First in First Out (FIFO), Last In First Out (LIFO)
- Priority based: Src or pkts have different priority
- Weight based: Weighted Fair Queueing (WFQ)

1.17.3 Traffic Shaping

- Controls flow of data
- Time-based: Leaky bucket alg
- Token-based: Token bucket alg

1.17.4 Admission Control

- Enables access if certain performance metrics are met
- Otherwise refuses admission
- Maintains certain level of performance e.g. quiality

1.18 Traffic Engineering

1.18.1 Motivation

- Network traffic highly dynamic
- Network resources variable
- No network control
- No guaranteed QoS
- No efficient use of network resources
- No guaranteed security, reliability, resilience, etc.

1.18.2 Definition

- TE is concerned with optimizing performance of telecomms network by dynamically analyzing, predicting, and regulating the bhavour of dat TX over that network

1.18.3 Goal

- Optimisation in terms of efficiency (i,e, costs and quality)

1.18.4 Major Solutions

- Intergrated Services
  - Focus on providing QoS delivery guarantees per-flow (using resource reservation: RSVP)
  - Concerns on: complexity, scalabity, business model, etc.
- Differentiated Services
  - Focus on providing QoS support per-class
  - Routers on the network differentiate traffic treatment based on it's class, ensuring preferential treatment for higher priority traffic
  - No advance setup, no reservation, no negotiations for each flow, easier to implement
- Other solutions: MLPS

1.19 Performance of Datalink and Physical Network Layers

- Wireless PANs (BT - IEEE802.15)
  - v. low range
  - wireless connection to printers etc
- Wireless LANs (WiFi - IEEE 802.11)

  - Infrastructure as well as ad-hoc net. possible
  - Home/office net.

- Wireless MANs (WiMAX - IEEE 802.16)
  - Large scale network
  - Base station-based infrastructure

1.20 WiMax IEEE 802.16

- Group formed in 98
- Standards
  - Air-interface for wireless broadband
  - Line of Sight comms
  - Ooperates in 10-66GHz range
- 802,16a amendment
  - Included Non LOS version in 2-11 GHz freq. band
  - PHY layer uses orthogonal freq. division multiplexing
  - MAC layer supports Orthogonal Frequency Division Multiple Access
- 802.16d
  - Further amended standard
  - Formed basis for first WiMax soluions
- Above standards supported fixed wireless applications
  - No mobility support
- 802.16e - 2005
  - added support for mobility and improved performance
  - Enable soft and hard handover between base stations
  - Introduce scalable OFDMA
    - Enables higher sppectrum efficiency in wide channels
    - Cosst reduction in narrow channels
  - Improces coverage using
    - Antenna diversity schemes
    - Hybrid ARQ (hARQ)
  - Improving capacity and coverage using
    - Adaptive Antenna Systes (AAS)
    - Multiple Input Multiple Output(MIMO) tech
  - Into'd high performance coding techniques to enhance security and NLOS performance
    - Turbo coding

    * Low density parity check
  - Intro's downlink sub-channelization alling admins trade coverage for capacity or vice versa
  - Increases resistance to multipath interference using enhances FFT alg. which can tolerate larger delay spreads
  - Adds extra QoS classs (enhanced rtPS) more appropriate for VoIP apps

1.21 WiMax 802.16e Service Classes

1.21.1 Unsolicited Grant Service (UGS)

- Fixed size pkt carried periodically without requiring explicit req. for bw allocation every time. Real-time high bw (T1) CBR applications (e.g. VoIP)

1.21.2 Extended Real-Time Poling Service (ertPS)

- Newly intro'd scheduling service in 802.16e complement periodic bw allocations with possibilities for mobile stations to req. additional resources during original allocation. Supports applications whose bw req. vary in time (e.g. VoIP, streaming)

1.21.3 Real Time Polling Service (rtPS)

- Intro'd to support real time services with variable zide data pkts fenerated periodically, such as MPEG video delivery applications. Frequent unicast polling opps are provided such as movile stations can req. bw and satisfy their timing req.

1.21.4 Non-Real-Time Polling Service (nrtPS)

- Similar with rtPS, unicast polling opps less freqent. Contention based polling can also be used to req. bw resources

1.21.5 Best Effort Service (BE)

- Designed for services with no strict QoS requirements such as email and web apps. Mobile stations use contention based polling to request resources

1.22 WiFi IEEE 802.11

- First std published in 97 for WLAN comms
- Since, various extensions proposed to address different issues - higher bit rate, QoS support, security
- Tech gained popularity because of low deployment and maintenance cost, as well as relatively high bitrate
- IEEE 802.11 - 1997

  - supports data rates up to 2Mbps, initially developed for best effort traffic only
  - Each host connects to an IEEE802.11 access point
  - Wireless medium shared with other nodes associated with same AP point
  - Contention for medium access which determines increased collision rates and consequently lower data rates especially when num of mobile hosts involved in sumultaneous data comms increases
  - IEEE 802.11 MAC layer provides mech. for medium access coordination:
    * Distributed Coordination Function (DCF) - distributed
    * Point coordination Function (PCF) - partly centralised

- IEEE 802.11b
  - Increased max data rate to 11Mbps, operating in 2.4 GHz freq. band
- IEEE 802.11g
  - Increase max. data rates to 54Mbps
- IEEE 802.11a
  - Data rates up to 54Mbps operating in 5GHz freq. band
- IEEE 802.11e
  - QoS extesion provided by two new mechanisms
    * Hybrid Coordination Function (HCF) - PCF extension
    * Enhanced Distributed Coordination Function (EDCF) - DCF extension

1.22.1 Other 802.11 extensions

- IEEE 802.11n
  - Higher bitrates up to 600Mbps
  - QoS support similar with 802.11e
- IEEE 802.11p
  - Wireless comms in vehicular environments
  - Short to medium range comms at high data transfer rates
- IEEE 802.11ac VHT
  - Offers data rates up to 1Gbps for low velocity mobile hosts

1.23 WiFi IEEE 802.11 Issues

1.23.1 Hidden Station Problem

- Consider that station B has TX range indicated by left oval, C by right oval. Any station in these ranges can hear TX from B and C respectively
- Station C, outside TX range of B cannot hear B, likewise B cannot hear C. Station A in range of both. Assuming B TX to A, C cannot hear B, believes medium is free, also TX data to A. Neither TX successful
- RST and CTS frames introduced to solve this problem. Before sending, B sends RTS to A (includes duration of TX). A hears RTS and replies with CTS which also includes TX duration info. As C is in range of A it gets CTS msg, learns of hidden station will be using channel, refrains from TX.

1.24 WiFi IEEE 802.11e

1.24.1 QoS Support

- Access Class (AC)
  - AC_VO (Voice), AC_VI (video), AC_BE (best effort) and AC_BK (Background)
- Transmission Opportunity (TxOP)
  - Time duration dring which station is allowed Tx turst of data frames
- Arbitration Interframe Space (AIFS[AC])
  - Period of time a wireless node has to wait before alloewd TX next frame
  - Dependent on access class
- Contention Window (CWmin, CWmax)
  - Deoendent on access class

1.25 WPAN

1.25.1 IEEE 802.15

- Bluetooth
  - Interconnects various protale devices and their accessories
  - 2.4GHz band
  - Data rates of up to 1Mbps (BT v1.0) and up to 3Mbps (v2.0)
  - Future rates expected to be between 53Mbps and 480Mbos
  - IEEE802.15.1 based on BT v1.1
- IEEE 802.15.4
  - Low-range, low-power wireless network comms

- Based on this standard, Zigbee protocol defines the network layer specialized on ad-hoc networking and the application layer targeting wireless sensor networks as well as other monitoring and control applications
- IEEE 802.15.4/Zigbee offers data rates up to 250 Kbps in the 2.4GHz band

- Ultra-Wideband (UWB)
  - WiMedia Alliance defined UWB wireless comms tech sipporting wide range of data rates from 53Mbps to 480Mbps over short range using low power transceivers
  - PHY/MAC protocols developed by WiMedia became ECMA 368 standard and later on ISO/OEC 26907
- Wibree
  - Ultra-low power wireless net. comms tech
  - Ranges up to 10m and data rates of 1Mbps