



EHB 453, Introduction to Mobile Communications

Lecture 3: All-IP Networking

Prof. Mustafa Ergen





ALL-IP NETWORKING

How it is evolved?

- Circuit switching: a dedicated capacity
- Packet switching: a capacity used on *need* basis.
 - Started as a military project called ARPANET
 - -No end to end dedication
 - -Buffer and forward system in between
 - Utilizes link at maximum extent
 - Traditionally, does not guarantee timely delivery
 - Hence, QoS protocols are introduced for voice, video over IP communication



Outline

 We examine the technology paths to All-IP Networking starting from basics of IP technology and continuing with advanced components of nextgeneration networks.

Today's hierarchical architecture for cellular networks created in the circuitswitched era became inefficient in supporting real-time IP services. The shift to flat IP networks in cellular will deliver substantial cost and flexibility to mobile operation as well as address the increasing requirements of emerging applications.



Orientation

• IP (Internet Protocol) is a Network Layer Protocol.



• IP's current version is Version 4 (IPv4). It is specified in RFC 891.



Application protocol

• IP is the highest layer protocol which is implemented at both routers and hosts









IP Protocol





IP Service

- Delivery service of IP is minimal
- IP provide provides an unreliable connectionless best effort service (also called: "datagram service").
 - Unreliable: IP does not make an attempt to recover lost packets
 - Connectionless: Each packet ("datagram") is handled independently. IP is not aware that packets between hosts may be sent in a logical sequence
 - Best effort: IP does not make guarantees on the service (no throughput guarantee, no delay guarantee,...)
- Consequences:
 - Higher layer protocols have to deal with losses or with duplicate packets
 - Packets may be delivered out-of-sequence



IP Service

• IP supports the following services:



- IP multicast also supports a many-to-many service.
- IP multicast requires support of other protocols (IGMP, multicast routing)



IP Datagram Format

- 20 bytes \leq Header Size $< 2^4 \times 4$ bytes = 60 bytes
- 20 bytes \leq Total Length $< 2^{16}$ bytes = 65536 bytes







Maximum Transmission Unit

- Maximum size of IP datagram is 65535, but the data link layer protocol generally imposes a limit that is much smaller
- Example:
 - Ethernet frames have a maximum payload of 1500 bytes

→ IP datagrams encapsulated in Ethernet frame cannot be longer than 1500 bytes

- The limit on the maximum IP datagram size, imposed by the data link protocol is called **maximum transmission unit (MTU)**
- MTUs for various data link protocols:

Ethernet:1500802.3:1492802.5:4464

FDDI:4352ATM AAL5:9180PPP:negotiated



IP Fragmentation

- What if the size of an IP datagram exceeds the MTU? IP datagram is fragmented into smaller units.
- What if the route contains networks with different MTUs?



- Fragmentation:
 - IP router splits the datagram into several datagram
 - Fragments are reassembled at receiver



IP Address Classification

Table 3.1 IP Address Classification

Class	Address range	Network address field size	Total network addresses	Host address field size	Total host addresses
A	1.0.0.0-126.0.0.0	7	126	24	16,777,214
В	128.0.0.1-191.255.0.0	14	16,383	16	65,534
C	192.0.1.0-223.255.255.0	21	2,097,151	8	254
D	224.0.0.0-239.255.255.255		28		
E	240.0.0.0-255.255.255.255				

CIDR: Classless Inter-Domain Routing Protocol (IP address/Mask), in 1993 with Variable Subnet Masking to introduce arbitrary-length prefixes.

Ex: 196.0.0/21 routes any address in 196.0.0.0 to 196.0.7.0 to the same network since they have the same first 21 bits.



IP Addressing

- An IP address is a 32-bit sequence of 1s and 0s.
- To make the IP address easier to use, the address is usually written as four decimal numbers separated by periods.
- This way of writing the address is called the dotted decimal format.



Public and Private IP Addresses

- No two machines that connect to a public network can have the same IP address because public IP addresses are global and standardized.
- However, private networks that are not connected to the Internet may use any host addresses, as long as each host within the private network is unique.
- RFC 1918 sets aside three blocks of IP addresses for private, internal use.
- Connecting a network using private addresses to the Internet requires translation of the private addresses to public addresses using Network Address Translation (NAT).

Class	RFC 1918 internal address range
А	10.0.0 to 10.255.255.255
В	172.16.0.0 to 172.31.255.255
С	192.168.0.0 to 192.168.255.255



Introduction to Subnetting

• To create a subnet address, a network administrator borrows bits from the host field and designates them as the subnet field.

Decimal notation for first Host octet	Number of Subnets	Number of Class A Hosts per Subnet	Number of Class B Hosts per Subnet	Number of Class C Hosts per Subnet
.192	2	4,194,302	16,382	62
.224	6	2,097,150	8,190	30
.240	14	1,048,574	4,094	14
.248	30	524,286	2,046	6
.252	62	262,142	1,022	2
.254	126	131,070	510	-
.255	254	65,534	254	-



Obtaining an Internet Address

- Static addressing
 - Each individual device must be configured with an IP address.
- Dynamic addressing
 - Reverse Address Resolution Protocol (RARP)
 - Bootstrap Protocol (BOOTP)
 - Dynamic Host Configuration Protocol (DHCP)
 - DHCP initialization sequence
 - Function of the Address Resolution Protocol
 - ARP operation within a subnet



Address Resolution Protocol (ARP)

- Each device on a network maintains its own ARP table.
- A device that requires an IP and MAC address pair broadcasts an ARP request.





IPv6 Background

- IP has been patched (subnets, supernets) but there is still the fundamental 32 bit address limitation
- IETF started effort to specify new version of IP in 1991
 - New version would require change of header
 - Include all modifications in one new protocol
 - Solicitation of suggestions from community
 - Result was IPng which became IPv6
 - First version completed in '94
- Same architectural principles as v4



IPv4 versus IPv6

- IP version 6 (IPv6) has been defined and developed.
- IPv6 uses 128 bits rather than the 32 bits currently used in IPv4.
- IPv6 uses hexadecimal numbers to represent the 128 bits.



IPv4



IPv6 planned support list

- 128-bit address space
 - This is what it's all about...
- Real-time/QoS services
- Security and authentication
- Autoconfiguration
 - Hosts autoconfig with IP address and domain name
 - Idea is to try to make systems more plug-n-play
- Enhanced routing functionality eg. Mobile hosts
- Multicast
- Protocol extensions
- Smooth transition path from IPv4
 - Can't do it all at once!



IPv6 Packet Format





Packet Format Details

- Simpler format than v4
- Version = 6
- Traffic class same as v4 ToS
- Treat all packets with the same Flow Label equally
 - Support QoS and fair bandwidth allocation
- Payload length does not include header –limits packets to 64KB
 - There is a "jumbogram option"
- Hop limit = TTL field
- Next header combines options and protocol
 - If there are no options then NextHeader is the protocol field
- Options are "extension header" that follow IP header
 - Ordered list of tuples 6 common types
 - Quickly enable a router to tell if the options are meant for it
 - Eg. routing, fragmentation, authentication encryption...



Key differences in header

- No checksum
 - Bit level errors are checked for all over the place
- No length variability in header
 Fixed format speeds processing
- No more fragmentation and reassembly in header
 - Incorrectly sized packets are dropped and message is sent to sender to reduce packet size
 - Hosts should do path MTU discovery
 - But of course we have to be able to segment packets!
 - What about UDP packets?



Transition from v4 to v6

- *Flag day* is not feasible
- Dual stack operation v6 nodes run in both v4 and v6 modes and use version field to decide which stack to use
 - Nodes can be assigned a v4 compatible v6 address
 - Allows a host which supports v6 to talk v6 even if local routers only speak v4
 - Signals the need for tunneling
 - Add 96 0's (zero-extending) to a 32-bit v4 address eg. ::10.0.0.1
 - Nodes can be assigned a v4 mapped v6 address
 - Allows a host which supports both v6 and v4 to communicate with a v4 hosts
 - Add 2 bytes of 1's to v4 address then zero-extend the rest eg. ::ffff: 10.0.0.1
- Tunneling is used to deal with networks where v4 router(s) sit between two v6 routers
 - Simply encapsulate v6 packets and all of their information in v4 packets until you hit the next v6 router



AS-level Internet Graph

S-level INTERNET GRAPH

IPv4

IPv6



copyright © 2009 UC Regents. all rights reserved.



IP Routing Protocols



3.2 Routing protocols



Hybrid Routing Schemes

- Some parts use static and some parts dynamic routing
 - static routing on the access network
 - dynamic
 routing on
 the core and
 distribution
 network





IGP vs EGP

- Interior Gateway Protocols
 within a single autonomous system
 - single network administration
 - unique routing policy
 - make best use of network resources
- Exterior Gateway Protocols
 - among different autonomous systems
 - independent administrative entities
 - communication between independent network infrastructures

Kind of information that is carried and the way the routing table are calculated •Distancevector protocols

> •Link-state protocols



Distance-Vector vs Link-State

- Distance-vector protocols
 - Each router
 periodically sends to
 his neighbors
 - how far is the destination
 - the next hop to get there
 - Install routes directly in tables

Bellman-Ford Algorithm [RFC1058]

- Link-state protocols
 - Each router sends information about
 - Links to which it is attached
 - State of the links
 - It is flooded throughout the network
 - Every router calculates its routing table

Dijkstra's Algorithm



RIP

- Packets are sent every 30 se necessary
- Route is considered down i sec. (distance set to infinity)
- Two kinds of messages
 - request
 - Response
- The metric is a hop-count
 - The value of 1 to 15 infinity)

Problems:

Split-horizon

- the information about destination routed on the link is omitted Poison reverse
 - the corresponding distance is set to infinity if the destination is routed on the link

IGP, distance-vector protocol First used in XNS (Xerox Network Systems) Designed as a component of the networking code for the BSD release of UNIX

incorporated in program "routed" (rote management daemon) First documented in RFC 1058



Example:



Routing table for node A

	Dest.	Link	Нор
Dest. Link Hop	A	local	0
A local 0	В	1	1
B 1 1	С	1	2
C 1 2	D	1	3
E 2 1	Ε	2	1
	F	1	3

Dest.	. Link	Нор
Α	local	0
В	1	1
С	1	2
D	1	3
E	2	1
F	1	3
G	1	4



RIP: Pros and Cons

RIP II is documented in RFC-1287, RFC-1388 and RFC-2453

- Updates
 - A timer is associated with each entry in the routing table
 - much longer than the period of transmission of information
 - Triggered updates
 - request nodes to send messages as soon as they notice a change in the routing table
- Advantages
 - Simple to implement
 - Low requirement in processing and memory at the nodes
 - Suitable for small networks
- Disadvantages
 - Slow convergence
 - Bouncing effect
 - Counting to infinity problem
- Limitations
 - Maximum hop count of 15

- RIP is not alone! IGRP and EIGRP
- restricts the use of RIP in larger networks, but prevents the count to infinity problem (endless loops)
- Difference in links speed is not reflected in the hop-count metrics
 - congested links can be still included in the best path



OSPF

- Link state or SPF technology
- Developed by OSPF Working Group of IETF (not proprietary)
- Designed for TCP/IP Internet environment
- Documented in RFC 1583, RFC 2178


OSPF - Link State Protocol

- Link
 - an interface on the router
- Link state
 - description of the interface and the neighboring routers
 - IP address, mask, type, routers connected to
- Link state database
 - collection of link state advertisement for all routers and networks



How OSPF Works?

- Each router generates link-state advertisements for its links
- When no OSPF areas are configured, link-state advertisements are flooded to all routers
- It is crucial that all routers have identical link state database
- Shortest path three is calculated by all routers and routing tables are derived



Example: Choosing an Optimal Path





The Link Metric

- Possible metrics
 - –hop count
 - inverse of the link bandwidth
 - -delay
 - dynamically calculated
 - administratively assigned
 - combination
- Traffic should be monitored and metrics adjusted



Example for Bad Metrics



Link State Advertisement (LSA)

- Generated periodically or in response to any change
- Contains:
 - source identification
 - -sequence number
 - link state age
 - –list of neighbors



Bringing up Adjacency

- Synchronizing databases via comparison of sequence numbers
- "Interesting records" the sequence numbers are different or not present in database
- Client-server relationship is established first



The Flooding Protocol

- Used to securely deliver LSAs
 - Every node sends the LSA on every link except the one from where it received it
 - Very fast and very reliable, but wastes bandwidth
 - Messages sent only when there is a change or every 45 minutes
 - Each node compares the newly received LSA with the entry in the data base. If it is newer the database is updated



Securing the Map Updates

- Flooding procedure includes hop-by-hop acknowledgments
- Database description packets are transmitted in a secure fashion
- Each link state record is protected by a timer and is removed from the database if a refreshing packet does not arrive in due time
- All records are protected by checksum
- Messages can be authenticated, e.g. by passwords



Shortest Path Algorithm

- Places the router at the root of the tree
- In each iteration adds the router that is closest to it (smallest cumulative metric of the path)
- Finished when all routers are added and the shortest path tree is generated



Shortest Path Tree and Routing Table for R6





Scaling OSPF

Rule of thumb

- no more than 150 routers / area

Reality

- no more than 500 routers/area

- Backbone area is an area that glue all the other areas
 always marked as area 0
- proper use of areas reduces bandwidth
 - summarized routes

-instability is limited within the area



OSPF Advantages

- No limitation on hop count
- Supports classless routing
- Routing updates sent only when there is a change or very rarely
- Faster convergence
- Better load balancing
- Logical definition of areas
- Authentication and external routes tagging



RIP vs OSPF

- More complex than RIP
 - the documentation is five times the
 - -the management needs more infor
 - the implementation needs more co
- Why design such complex procedure?
 - -routing is important
 - requires less "signalization" messages
 - compute better routes

OSPF is not the protocol **IS-IS protocol** is part of OSI routing framework for CLNP similar in design to **OSPF** uses different termino Ogy



Internet Structure

Original idea





Internet Structure

Today





Route Propagation in the Internet

- Autonomous System (AS)
 - corresponds to an administrative domain
 - examples: University, company, backbone network
 - assign each AS a 16-bit number
- Two-level route propagation hierarchy
 - interior gateway protocol (each AS selects its own)
 - exterior gateway protocol (Internet-wide standard)
- Routes information is propagated at various levels
 - hosts know local router
 - local routers know site routers
 - site routers know core router
 - core routers know everything



Popular Interior Gateway Protocols

- RIP: Route Information Protocol
 - -distributed with BSD Unix
 - -distance-vector algorithm
 - -based on hop-count (infinity set to 16)
- OSPF: Open Shortest Path First
 - -recent Internet standard
 - uses link-state algorithm
 - supports load balancing
 - supports authentication



EGP: Exterior Gateway Protocol

- Overview
 - Original standard for Internet routing protocol (c 1983)
 - designed for tree-structured Internet
 - Single backbone
 - concerned with *reachability*, not optimal routes
- Protocol messages
 - neighbor acquisition: one router requests that another be its peer; peers exchange reachability information
 - neighbor reachability: one router periodically tests if the another is still reachable; exchange HELLO/ACK messages;
 - uses a k-out-of-n rule: ¼ to stay up, ¾ to establish
 - routing updates: peers periodically exchange their routing tables (including route weights) using a basic distance vector method
 - There can be multiple connections between ASs



Limits of EGP

- At first glance, EGP seems like a distance vector protocol since updates carry lists of destinations and distances – but distances are NOT reliable.
- EGP was designed to support tree topologies, not meshes
 - False routes injected by accident can have really bad consequences (black holes) – there is no easy way for dealing with this problem
 - Loops can easily occur all we are doing is forwarding routing tables
- EGP was not designed to easily support fragmented IP packets

 all data is assumed to fit in MTU.
- Solutions to these and other EGP problems were all manual



BGP-4: Border Gateway Protocol

- BGP-1 developed in 1989 to address problems with EGP.
- Assumes Internet is an arbitrarily interconnected set of ASs
- AS traffic types
 - Local
 - starts or ends within an AS
 - Transit
 - passes through an AS
- AS Types
 - stub AS: has a single connection to one other AS
 - carries local traffic only
 - multihomed AS: has connections to more than one AS
 - refuses to carry transit traffic
 - transit AS: has connections to more than one AS
 - carries both transit and local traffic



BGP-4 contd.

- Each AS has:
 - one or more border routers
 - Handles inter-AS traffic
 - one BGP speaker for an AS that participates in routing
 - BGP speaker establishes BGP sessions with peers and advertises:
 - local network names
 - other reachable networks (transit AS only)
 - gives path information including path weights (MEDs)
 - withdrawn routes
- BGP goal: find loop free paths between ASs
 - Optimality is secondary goal
 - It's neither a distance-vector nor a link-state protocol
- Hard problem
 - Internet's size (~12K active ASs) means large tables in BGP routers
 - Autonomous domains mean different path metrics
 - Need for flexibility



BGP Example

- Speaker for AS2 advertises reachability to P and Q
 - network 128.96, 192.4.153, 192.4.32, and 192.4.3, can be reached directly from AS2



- Speaker for backbone advertises
 - networks 128.96, 192.4.153, 192.4.32, and 192.4.3 can be reached along the path (AS1, AS2).
- Speaker can cancel previously advertised paths



Some BGP details

- Path vectors are most important innovation in BGP
 - Enables loop prevention in complex topologies
 - If AS sees itself in the path, it will not use that path
- Routes can be aggregated
 - Based on CIDR (classless) addressing
- Routes can be filtered
- Runs over TCP
- Most of the same messages as EGP
 - Open, Update, Notify, Keepalive
- BGP session have only recently been made secure



BGP in practice

- 10-20 "tier 1" ASs which are the Internet backbone
- Clearly convergence is an issue why?
- Black holes are always a potential problem
- There are lots of BGP updates every day!
- BGP is really the heart of the Internet
- BGP is a means by which network operators control congestion in the Internet.
- BGP is really a big problem!



Multicast IP

- The Internet Group Management Protocol (IGMP) is used by IPv4 systems (hosts and routers) to report their IP multicast group memberships to any neighboring multicast routers. Note that an IP multicast router may itself be a member of one or more multicast groups, in which case it performs both the "multicast router part" of the protocol (to collect the membership information needed by its multicast routing protocol) and the "group member part" of the protocol (to inform itself and other, neighboring multicast routers of its memberships).
- IGMP is also used for other IP multicast management functions, using
- message types other than those used for group membership reporting.



IGMP through versions

- Version 1, specified in [RFC-1112], was the first widelydeployed version and the first version to become an Internet Standard.
- Version 2, specified in [RFC-2236], added support for "low leave latency", that is, a reduction in the time it takes for a multicast router to learn that there are no longer any members of a particular group present on an attached network.
- Version 3 adds support for "source filtering", that is, the ability for a system to report interest in receiving packets *only* from specific source addresses, or from *all but* specific source addresses, sent to a particular multicast address.



IGMP v1 - Behaviour





IGMP v2 - enhancements

- Version 1, specified in [RFC-1112], was the first widelydeployed version and the first version to become an Internet Standard.
- Version 2, specified in [RFC-2236], added support for "low leave latency", that is, a reduction in the time it takes for a multicast router to learn that there are no longer any members of a particular group present on an attached network.
- Version 3 adds support for "source filtering", that is, the ability for a system to report interest in receiving packets *only* from specific source addresses, or from *all but* specific source addresses, sent to a particular multicast address.



IP QoS Defined

 The goal :

 Provide some level of predictability and control beyond the current IP "best-effort" service

Fundamental principle
 Leave complexity
 at the "edges" and
 keep network
 "core" simple

Internet today

- Provides "best effort" data delivery
- Complexity stays in the endhosts
- Network core remains simple As demands exceeds capacity, service degrades gracefully (increased jitter etc.)

Delivery delays cause problems to real-time applications

Performance attributes Service availability Delay Delay variation (jitter) Throughput Packet loss rate Vary according to Service Level Agreement (SLA)



QoS Protocols

- ReSerVation Protocol (RSVP)
- Differentiated Services (DiffServ)
- Multi Protocol Labeling Switching (MPLS)
- Subnet Bandwidth Management (SBM)

QoS can be achieved by : Resource reservation (integrated services) Prioritization (differentiated services)

QoS can be applied :

Per flow (individual, uni-directional streams) Per aggregate (two or more flows having something in common)



RSVP

- Implementation

Sender

PATH message containing traffic specification (bitrate, peak rate etc.) Receiver **RECV** message containing the reservation specification (guaranteed or controlled) the filter specification (type of packets that the reservation is made for)

Attributes

- The most complex of all QoS
- technologies
- Closest thing to circuit emulation
- on IP networks
- The biggest departure from "besteffort" IP service
- Provides the highest level of QoS in terms of :
 - Service guarantees
 - Granularity of resource allocation Detail of feedback to QoS-enabled applications





DiffServ

- Implementation

Two traffic classes are available : Expeditied Forwarding (EF) - 1 codepoint Minimizes delay and jitter Provides the highest QoS Traffic that exceeds the traffic profile is discarded

Assured Forwarding (AF) - 12 codepoints 4 classes, 3 drop-precedences within each class Traffic that exceeds the traffic profile is not delivered with such high probability



MPLS

- Label Switching
- Used to establish fixed bandwidth routes (similar to ATM virtual circuits)
- Resides only on routers and is protocol independent
- Traffic is marked at ingress and unmarked at egress boundaries
- Markings are used to determine next router hop (not priority)

The aim is to simplify the routing process ...



MPLS

- Implementation



Fig. 3.6 MPLS Operation: MPLS has 32-bit header which contains the label (20-bits), the Class of Service (CoS) field (3-bits) to implement service classes, the Stack (S) field (1-bit) to support hierarchical label stack for routing packets through LSP Tunnels, TTL (time-to-live) field (8-bits) as in conventional IP TTL



MPLS

- Conclusions
- Labels can be "stacked"
 - -This allows MPLS "routes within routes"
- Label Distribution Protocol (LDP)
 - Distributes labels across MPLS-enabled routers
 - Ensures they agree on the meaning of labels
 - -Usually transparent to network managers
- Implication :
 - Define a policy management that distributes labels


Summary

- IP Addressing
 - Started with IPv4 and introduced IPv6
 - Still IPv6 use is limited
- IP Routing
 - -Interior and Exterior protocols
- IP QoS
 - Providing more than best effort traffic



References

- Mobile Broadband, Ergen
- IMS, J. Rafferty
- Internet Telephony based on SIP, H. Sinnreich, A. Johnston
- A Multi-gigabit Rate Deep Packet Inspection Algorithm using TCAM, J-S Sung, et. al.
- CS40 Lecture 6: Security, R. Johari
- SIP, N. V. Pandrye
- Security Evolution on the Edge, W. Wilkening
- QoS in Data Networks, O. Ruso

