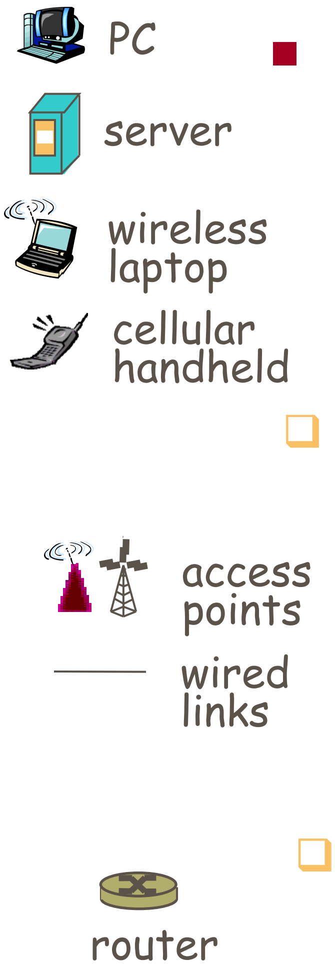
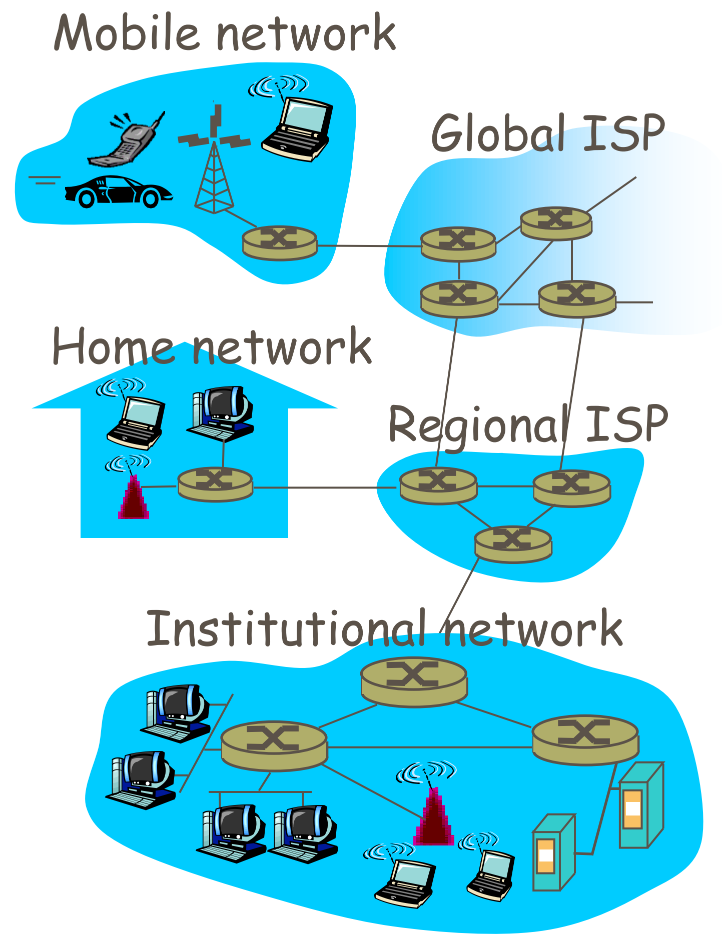
Digital Convergence

- **Separate** networks for ***different media*** transport
  - Internet
  - PSTN (*Public Switched Telephone Network*), PLMN (*Public Land Mobile Network*) (e.g., Cellular GSM) - voice
  - Cable networks – media broadcasting
  - Terrestrial wireless networks – TV broadcasting, radio
- "**Integrated Services**" *Digital* Networks
- Challenges and Issues
  - **QoS** – resource allocation, scheduling, fairness, pricing
  - Mobility
  - Security

Internet: interconnection of networks



- millions of connected computing devices: *hosts = end systems*
  - running *network apps*
- *communication links*
  - fiber, copper, **radio**, satellite
  - transmission rate = *bandwidth (bps)*
- *routers:* forward packets (chunks of data)



All-IP networks

- All media will be transported in IP packets across heterogeneous networks.

Transition of Telecom Industry …

- The roots are traditional *voice* service and the *circuit switched* networks that support it.
- Migrate to a new network and a new, broader set of applications.
- The new network will
  - use internet protocol (IP) and
  - converge voice, *data* and *video* services onto a **single** infrastructure.
- This will
  - significantly reduce the unit cost of delivering services for the operator and
  - create a *flexible*, dynamic technology platform upon which *new* services can be designed and deployed and retried.

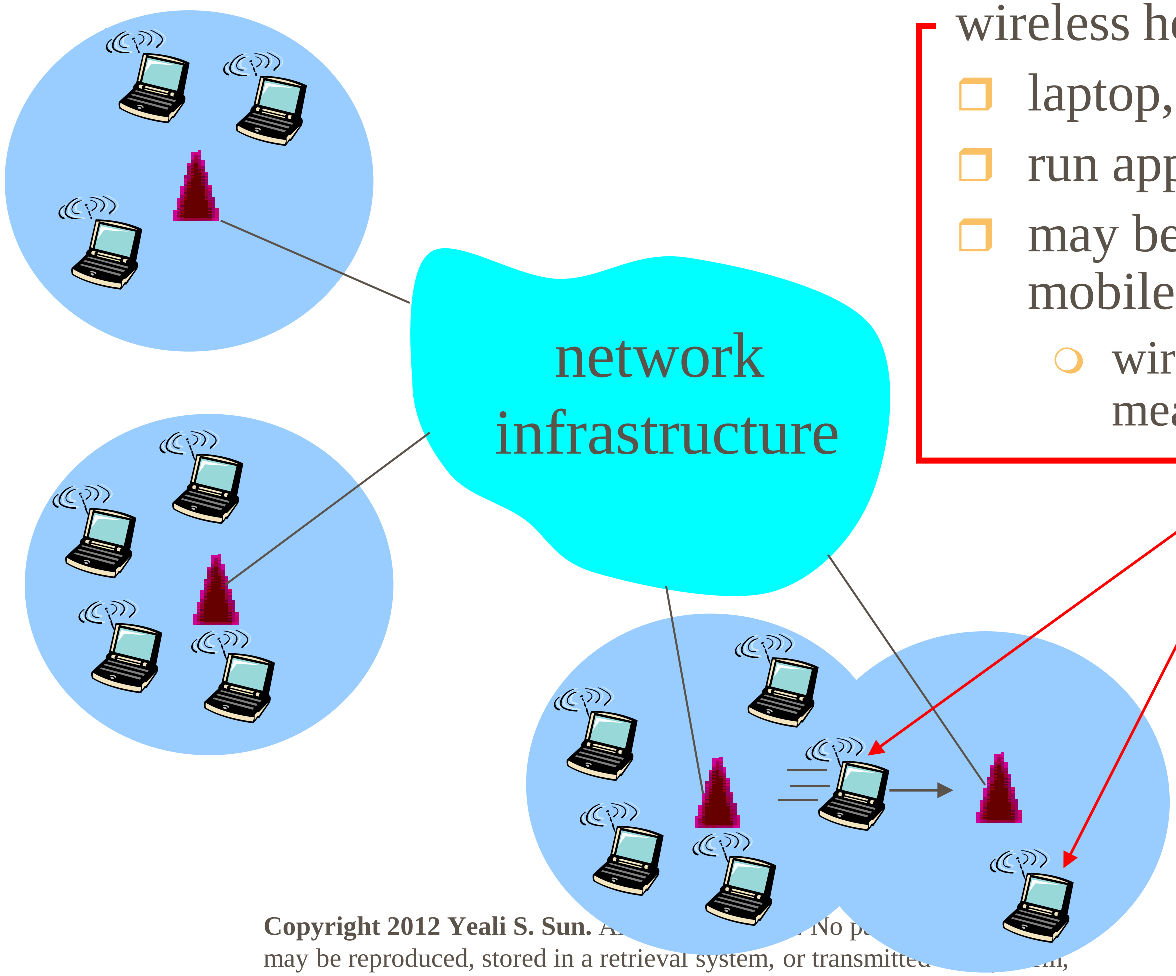
Wireless Cellular Networks in telecommunications: generation (1/2)

- 1G
  - 1981
  - Analog transmission
- 2G
  - 1992
  - Digital transmission
- 3G
  - 2001
  - Multi-media support
  - Spread spectrum transmission and at least 200 kbps
  - HSPA+ standard in WCDMA
    - In 2009: 28 Mbps downstreams and 22 Mbps upstreams without MIMO, i.e. only with one antenna
    - In 2011 up to 42 Mbps downstreams using 2x2 MIMO
  - EV-DO Rev. B in CDMA2000
    - in 2010 and offers 15.67 Mbit/s downstreams

Wireless Cellular Networks in telecommunications: generation (2/2)

- 4G
  - Standards
    - the **IMT-Advanced** (International Mobile Telecommunications Advanced) requirements specified by ITU-R organization in 2009
  - Main requirements
    - **peak speed** at 100 Mbps for high mobility communication (such as from trains and cars) and 1Gbps for low mobility communication (such as pedestrians and stationary users)
    - **all-IP packet-switched** networks
  - Compliant versions, e.g.,
    - Long-term-evolution Advanced (LTE) Advanced
    - WirelessMAN (WiMAX)-Advanced
  - Technologies
    - *OFDMA* and other frequency-domain equalization schemes combined with *MIMO* (Multiple In Multiple Out), e.g., multiple antennas, dynamic channel allocation and channel-dependent scheduling

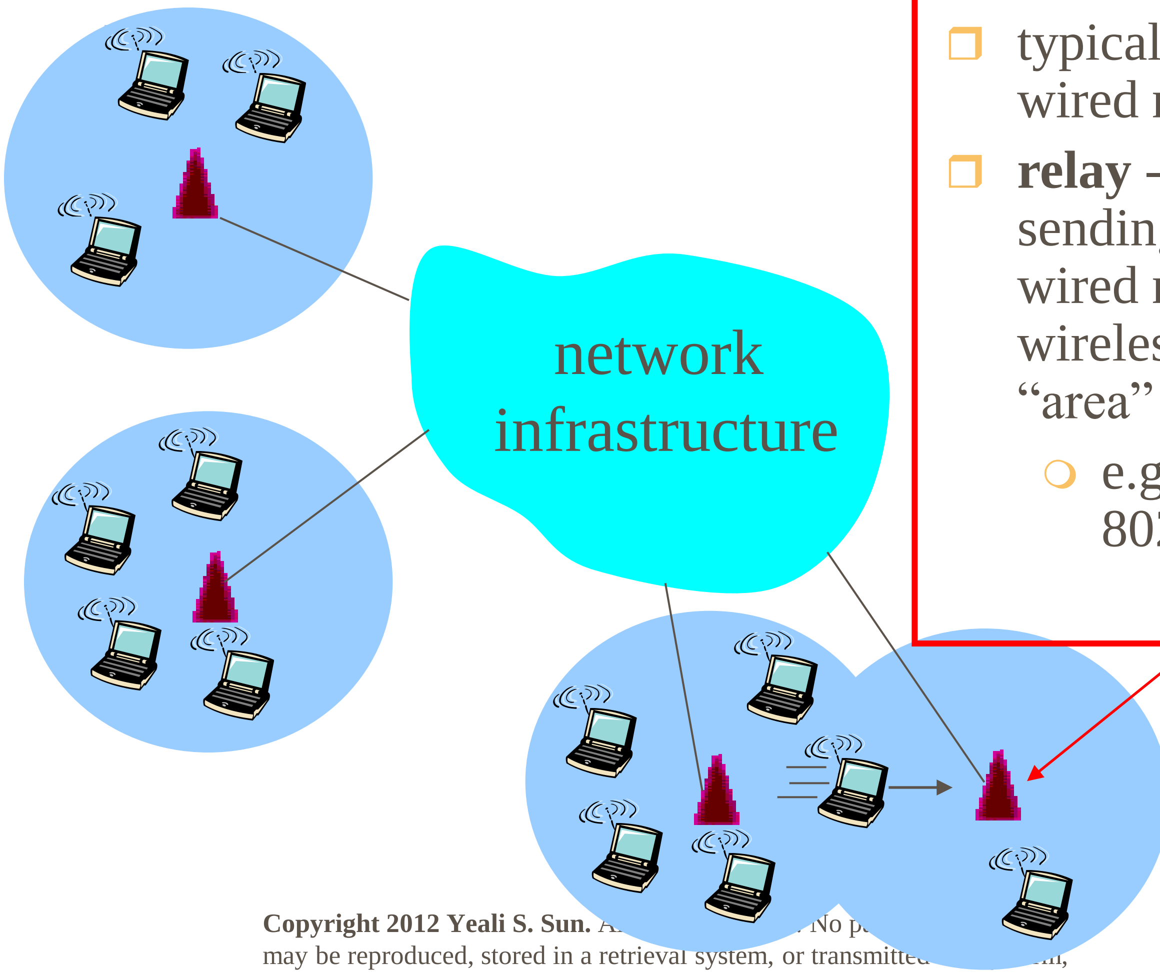
Elements of a wireless network



wireless hosts

- laptop, PDA, IP phone
- run applications
- may be stationary (non-mobile) or mobile
  - wireless does *not* always mean mobility

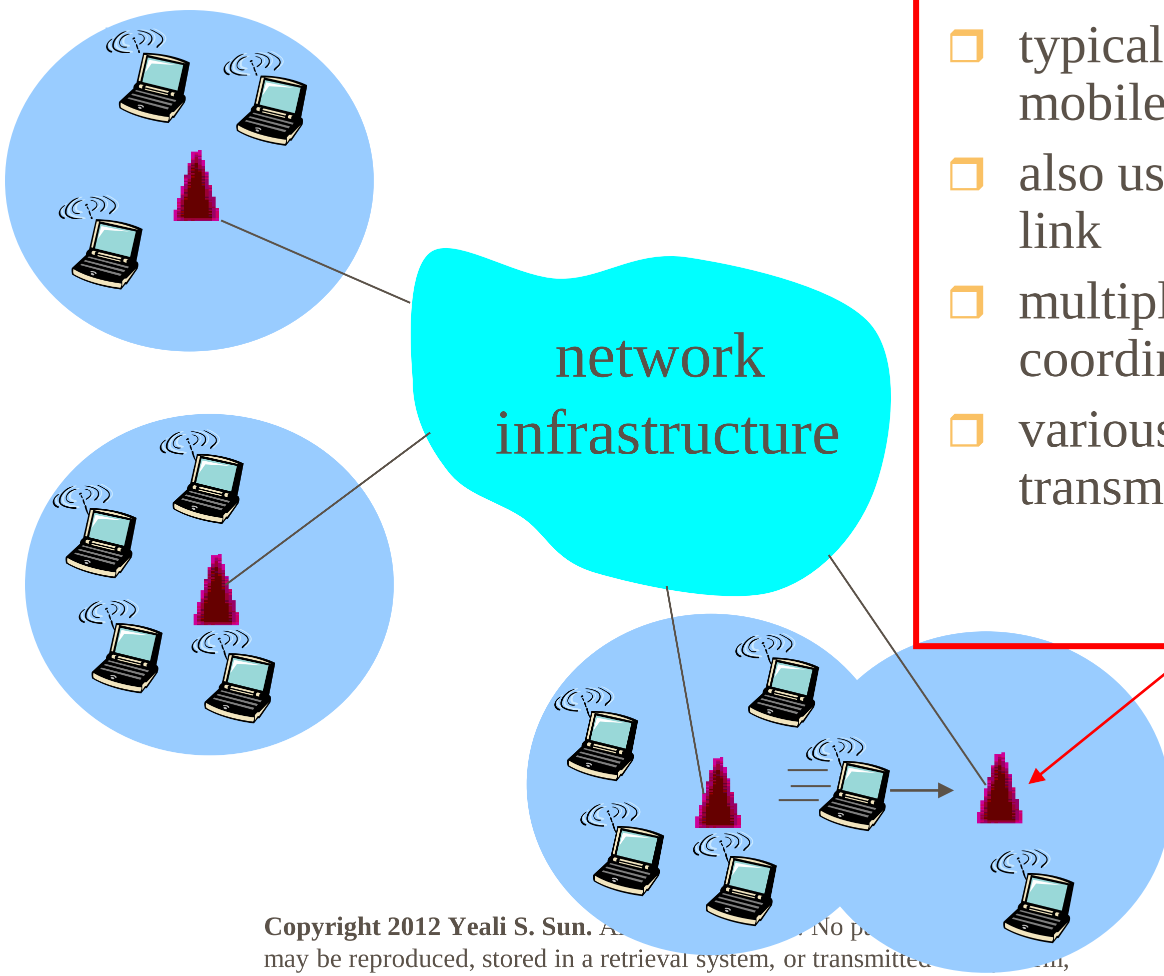
Elements of a wireless network



base station

- typically connected to wired network
- **relay** - responsible for sending packets between wired network and wireless host(s) in its "area"
  - e.g., cell towers, 802.11 access points

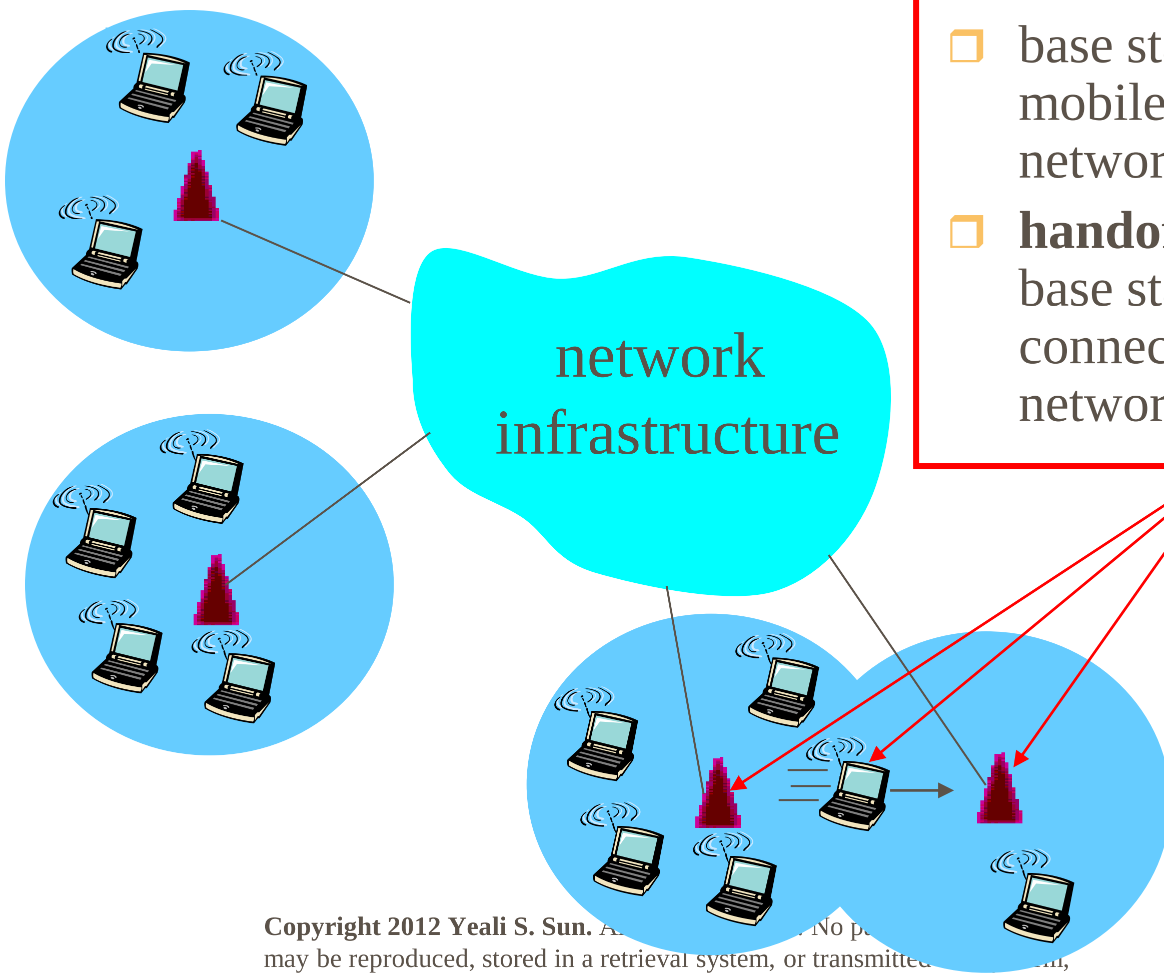
Elements of a wireless network



wireless link

- typically used to connect mobile(s) to base station
- also used as backbone link
- multiple access protocol coordinates **link access**
- various data rates, transmission distance

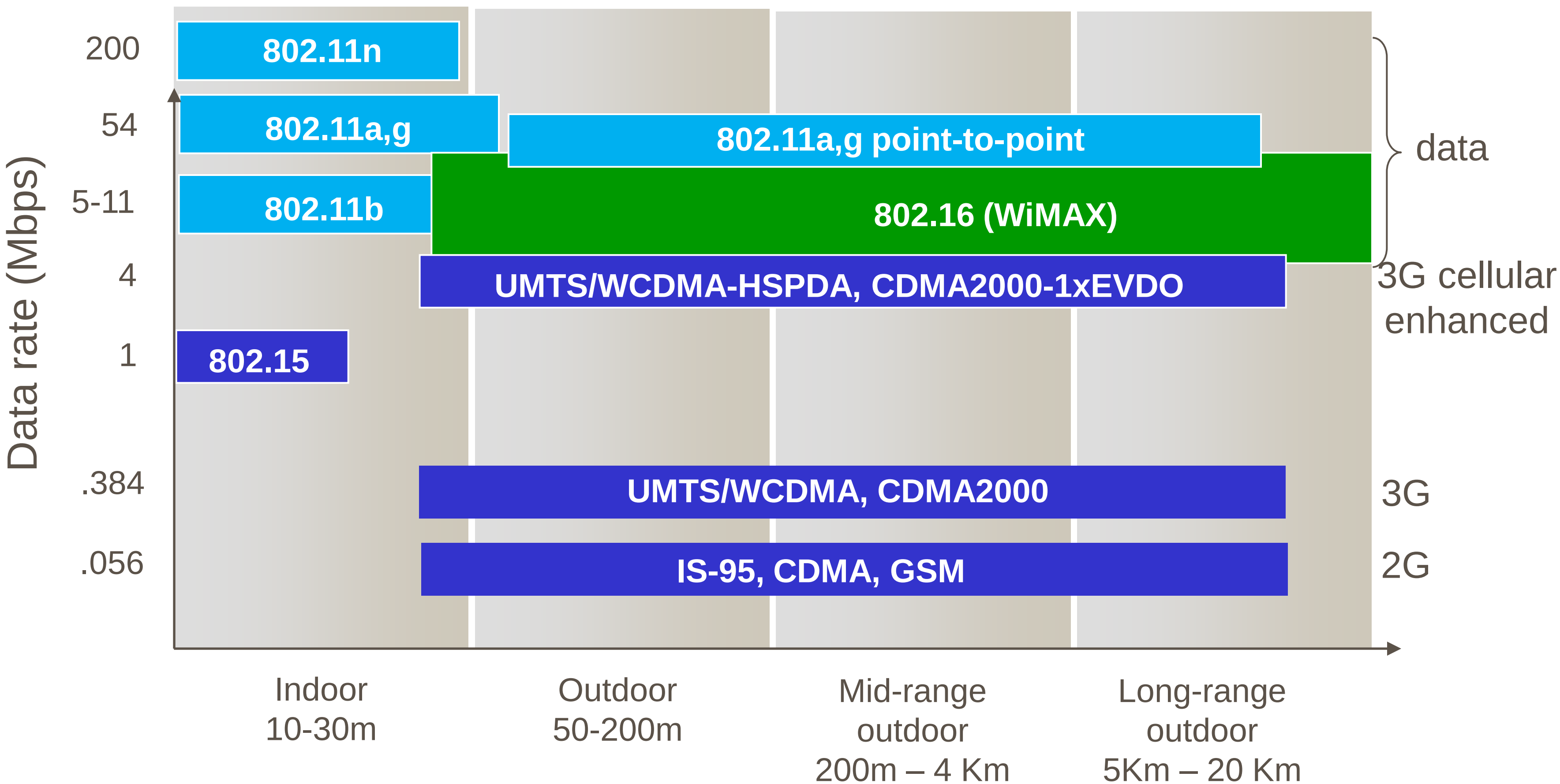
Elements of a wireless network



infrastructure mode

- base station connects mobiles into wired network
- **handoff**: mobile changes base station providing connection into wired network

Characteristics of selected wireless link standards



Quality of Service (QoS) and Resource Management

Professor Yeali S. Sun
National Taiwan University

QoS

- Traffic control
  - admission control
  - traffic shaping and policing
- Resource allocation
- Resource scheduling
- Notion of fairness

- Traffic characteristics
  - Bursty, constant-bit-rate, variable-bit-rate, etc.
  - Throughput, delay, delay jitter, packet loss rate, etc.

Outline

- QoS in ATM (Asynchronous Transfer Mode) networks
  - Service types
  - QoS parameters
- Traffic Shaping and Policing in Traffic Control
  - Leaky bucket
  - Token bucket

References

- S. Tanenbaum, “Computer Networks,” 5th edition, Prentice Hall, 2010.
- S. Tanenbaum, “Computer Networks”, 3rd edition, Prentice Hall, 1996.

Broadband ISDN

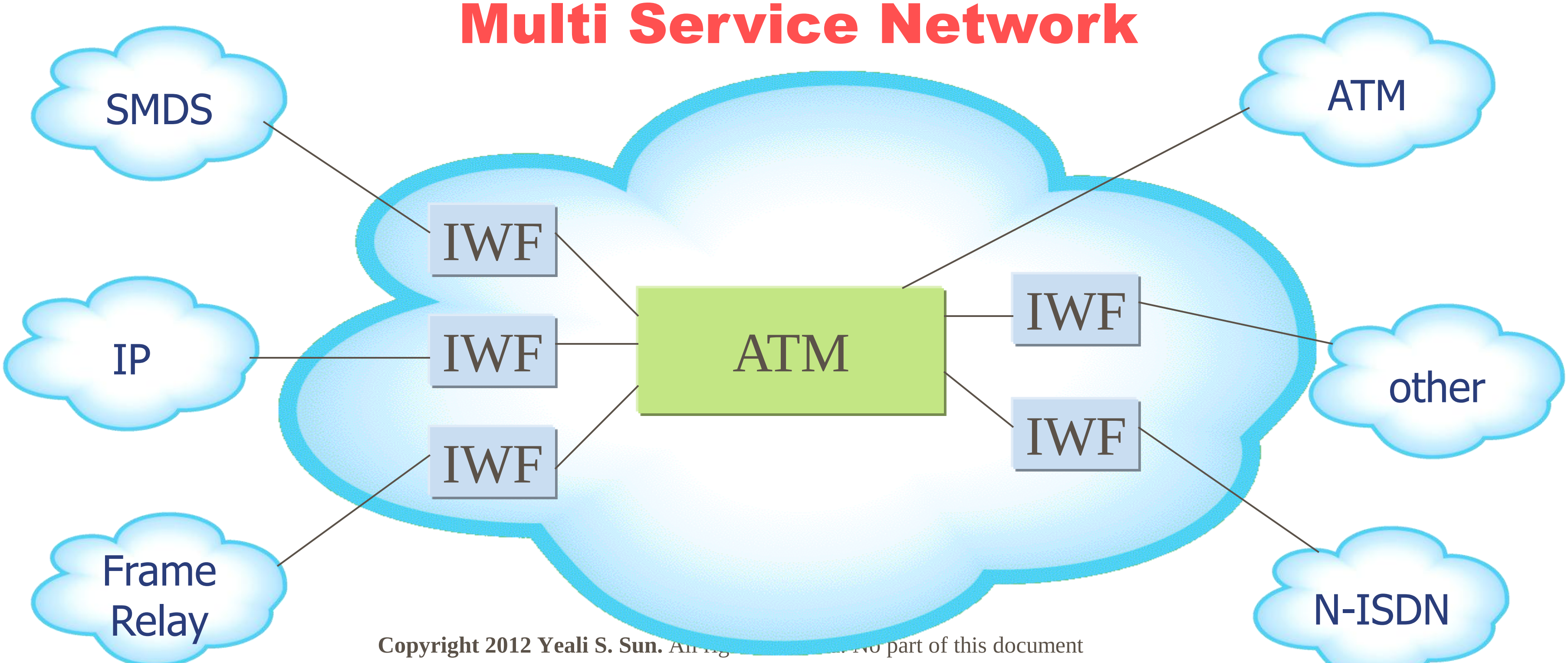
- ISDN – Integrated Services Digital Network
- Main players – telephone companies (teleco)
- The first attempt to build ONE single integrated network for all kinds of information transfer
- Technology of choice
  - **Asynchronous Transfer Mode (ATM)**

Broadband ISDN (cont'd)

- ATM is a technology and a service (visible to the user)
- a.k.a. cell relay
- Why cell switching?
  - High flexibility on handling both constant rate traffic (audio, video) and variable rate traffic (data)
  - *Easier operation for high speed networks* than that using traditional multiplexing techniques

ATM as the backbone infrastructure technology

Multi Service Network



IWF: InterWorking Function

ATM: Introduction

- To transmit all information in small, fixed-size packets (i.e. cells)
  - A cell is of 53 bytes – 5-byte header and 48-byte payload
- Connection-oriented service
  - Three phases: connection establishment, data transfer and connection release

Why ATM cell is size 53 bytes?

- The American standard for DS0 is 64 Kbs, the European standard is 32 Kbs.
- The Americans pushed for an ATM payload of 64 Bytes while the Europeans pushed for 32 Bytes.
- To come to a compromise, the committee
  - added 64 + 32 = 96
  - then divided this sum by 2 for a result of 48.
  - A 5 Byte header was added for a total of 53 Bytes.

Multimedia Communications over ATM

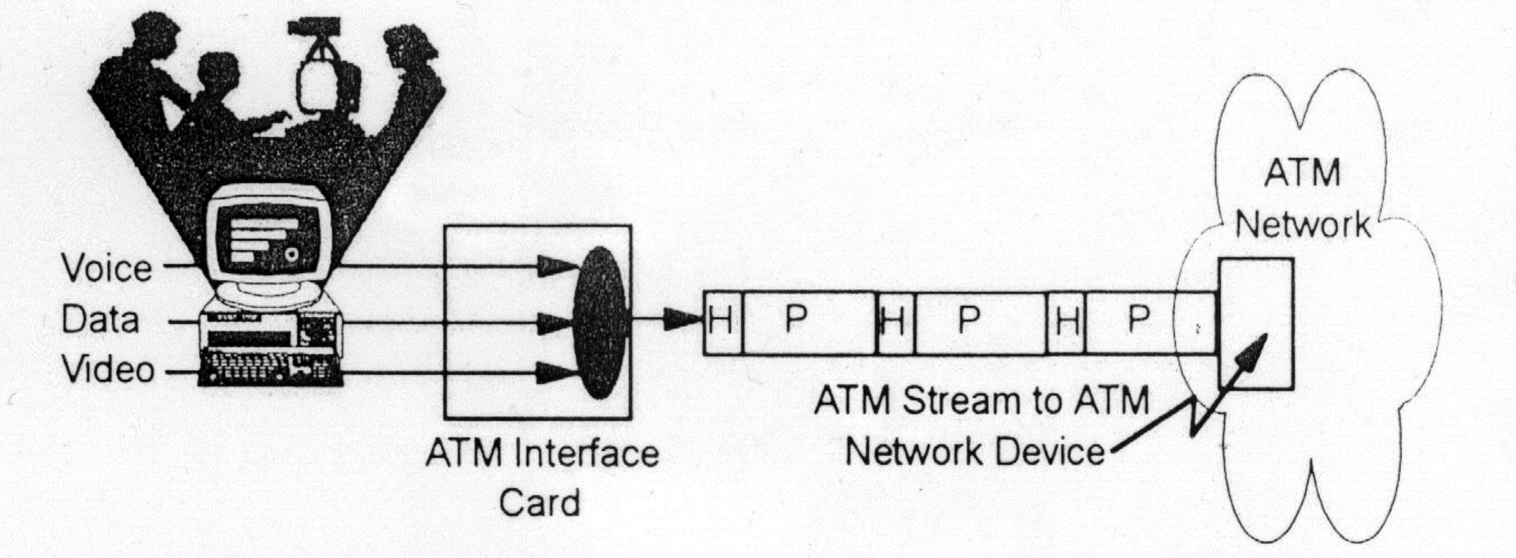


Figure 7.2 Multimedia Communications Example Using ATM

ATM Introduction (cont'd)

- Unreliable cell transfer
  - Designed for fiber optics
  - Leave error control to higher layers
  - For real-time traffic – no need of retransmission of occasional bad cells but just ignoring it.
  - Connection-oriented: cell order is preserved.
- Operating speeds
  - 155.5Mbps (oc-3, STM-1)
    - Compatible with SONET transmission system
  - 622Mbps (oc-12, STM-4)
    - Combine four 155Mbps channels

SONET		SDH	Data rate (Mbps)		
Electrical	Optical	Optical	Gross	SPE	User
STS-1	OC-1		51.84	50.112	49.536
STS-3	OC-3	STM-1	155.52	150.336	148.608
STS-9	OC-9	STM-3	466.56	451.008	445.824
STS-12	OC-12	STM-4	622.08	601.344	594.432
STS-18	OC-18	STM-6	933.12	902.016	891.648
STS-24	OC-24	STM-8	1244.16	1202.688	1188.864
STS-36	OC-36	STM-12	1866.24	1804.032	1783.296
STS-48	OC-48	STM-16	2488.32	2405.376	2377.728

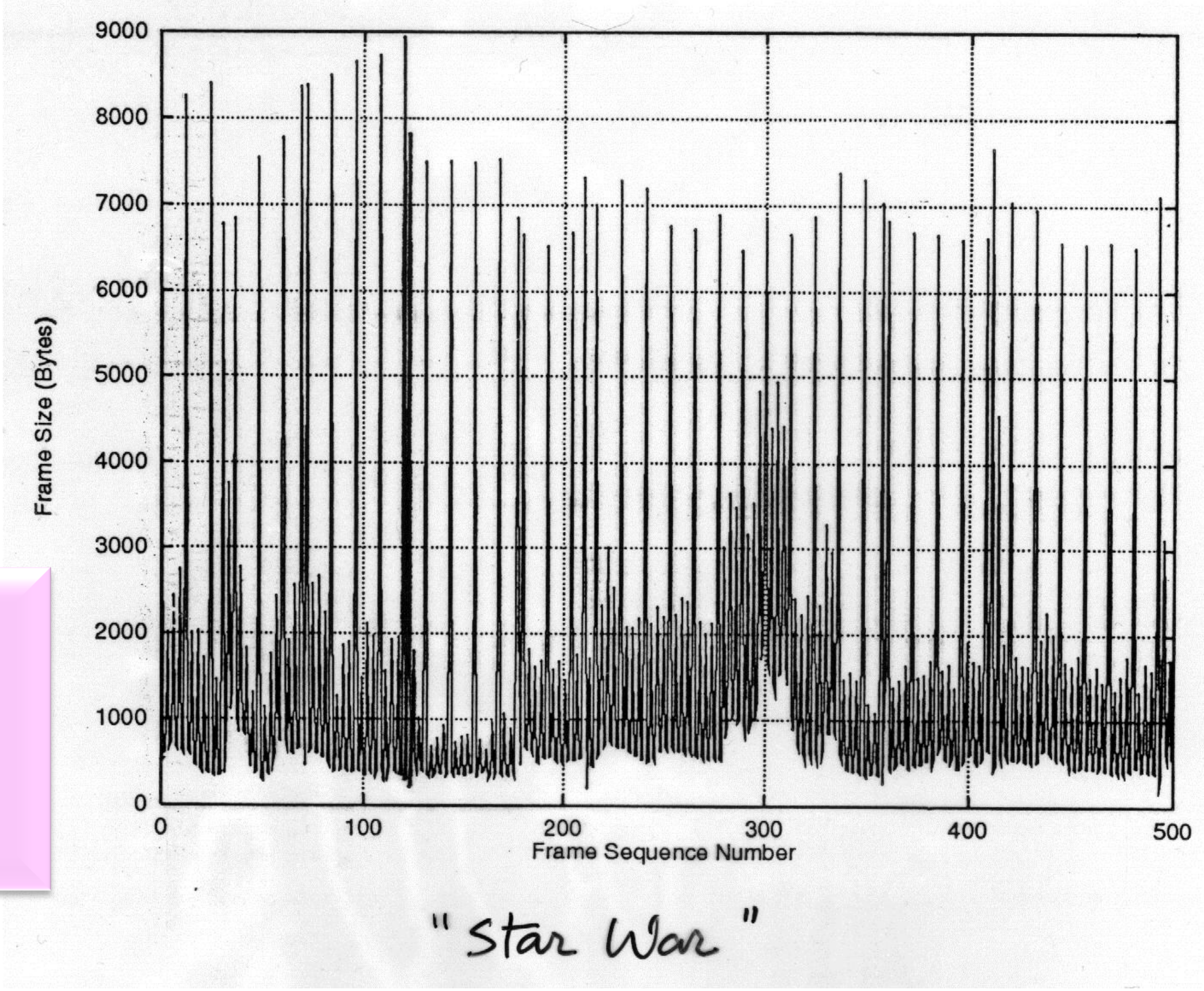
Fig. 2-32. SONET and SDH multiplex rates.

ATM Service Types

- Defined in ATM Forum UNI 4.0 (1988+)
- Five service types

Class	Description	Example
CBR	Constant bit rate	T1 circuit
RT-VBR	Variable bit rate: real time	Real-time videoconferencing
NRT-VBR	Variable bit rate: non-real time	Multimedia email
ABR	Available bit rate	Browsing the Web
UBR	Unspecified bit rate	Background file transfer

Fig. 5-69. The ATM service categories.



- Frame types: I, P and B
- In H.264, I, P, and B slices

Characteristics of the ATM Service Categories

Service characteristics	CBR	RT-VBR	NRT-VBR	ABR	UBR
Bandwidth guarantee	Yes	Yes	Yes	Optional	No
Suitable for real-time traffic	Yes	Yes	No	No	No
Suitable for bursty traffic	No	No	Yes	Yes	Yes
Feedback about congestion	No	No	No	Yes	No

Fig. 5-70. Characteristics of the ATM service categories.

- Bandwidth
- End-to-end Delay
- Statistical multiplexing
- Congestion control

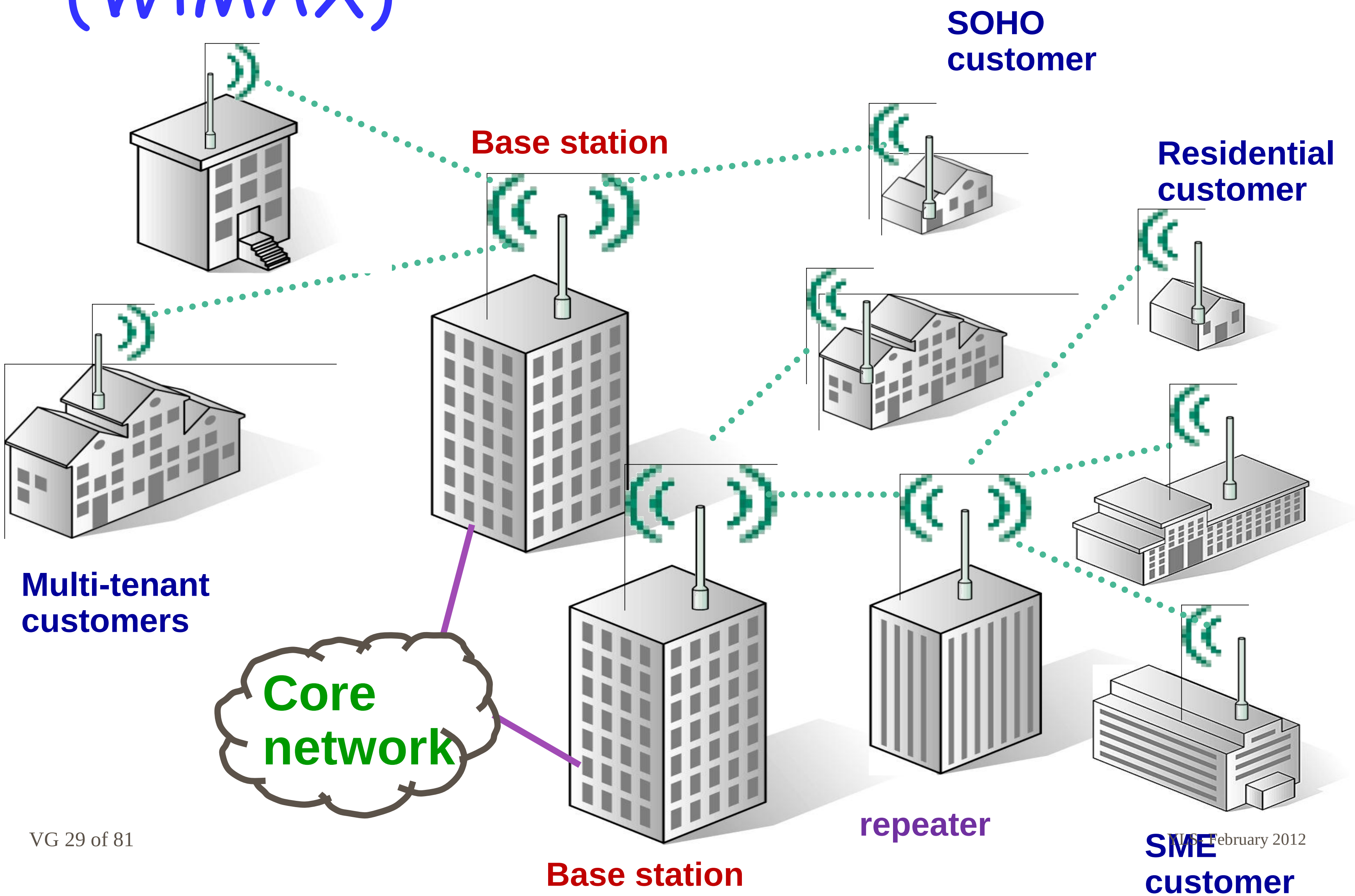
Data Over Cable Service Interface Specification (DOCSIS)

- An international standard for the communications and operation support interface requirements for a data over cable system.
- It is employed by many cable television operators (CATV) to provide Internet access over their existing hybrid fiber coaxial (HFC) infrastructure.
- Specifications
  - Version 1.0 – issued in March 1997,
  - Version 1.1 – adding *Quality of Service (QoS)* capabilities in April 1999.
  - Version 2.0 – *enhancement of upstream transmission speeds* for symmetric services in December 2001.
  - Version 3.0 – enhancement to increase transmissions speeds for both upstream and downstream and support for Internet Protocol version 6 (IPv6) in August 2006.

IEEE 802.16 (WiMAX)

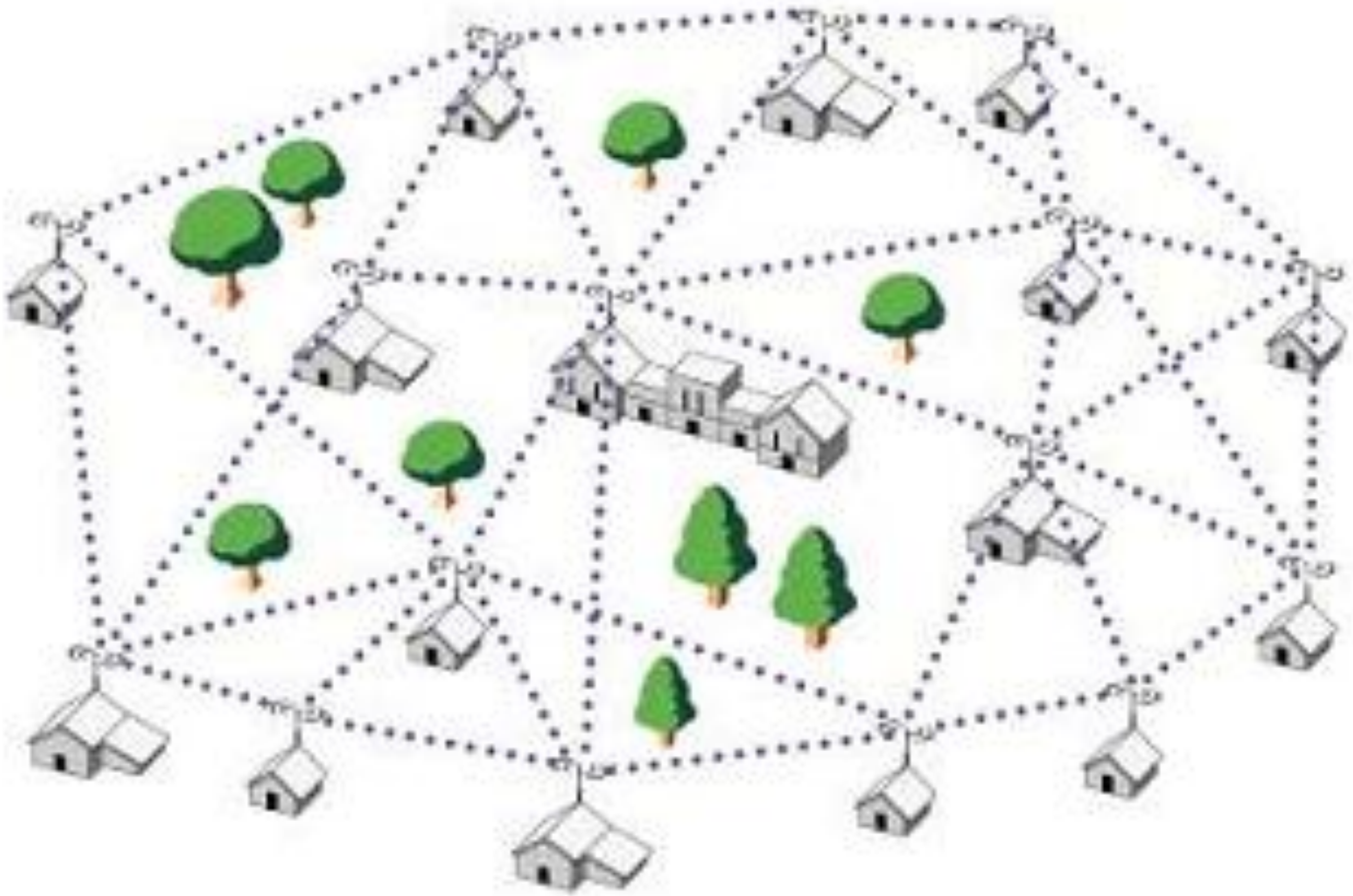
- The IEEE 802.16-e 2005 standards for *mobile broadband wireless services* with the geographical coverage scale of *a metropolitan area*.

Broadband Wireless: IEEE 802.16e (WiMAX)

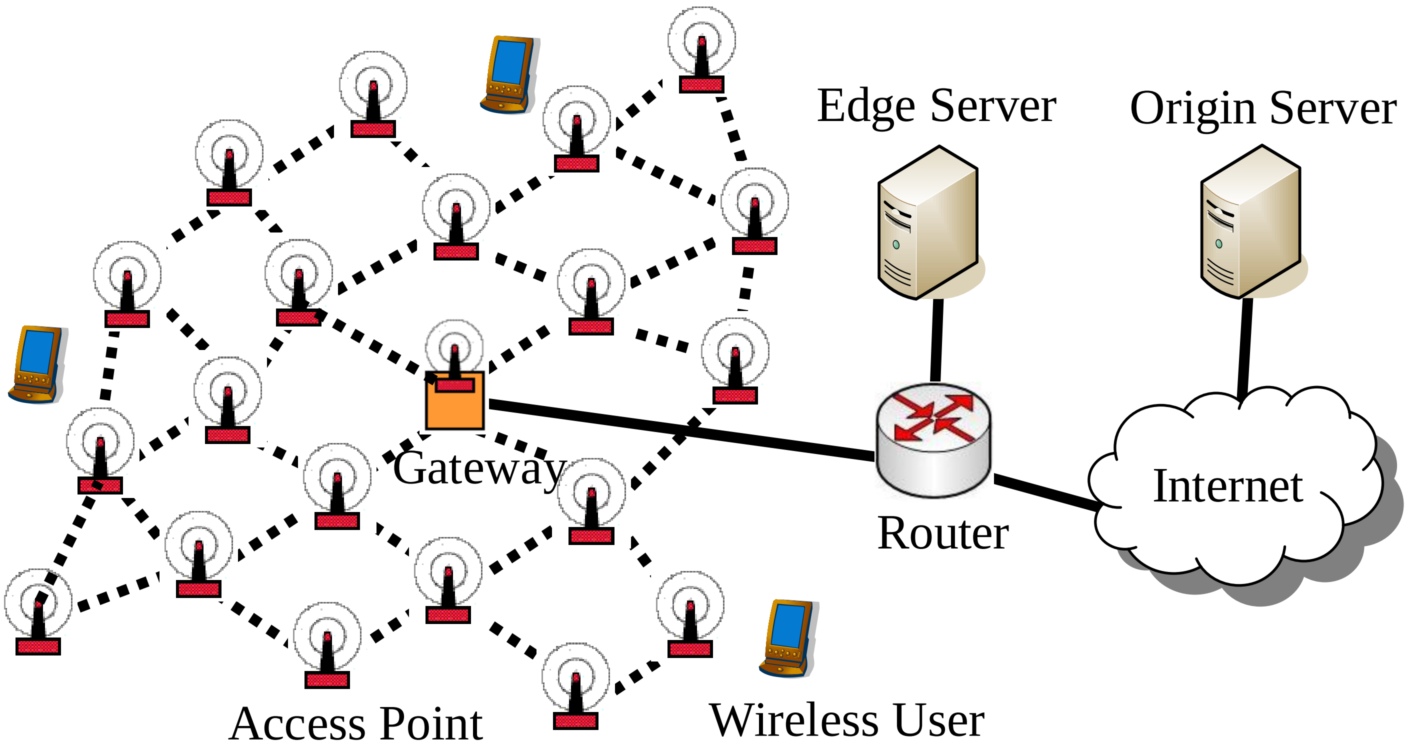


Source: Nokia Networks

Broadband Wireless: IEEE 802.16e (WiMAX)



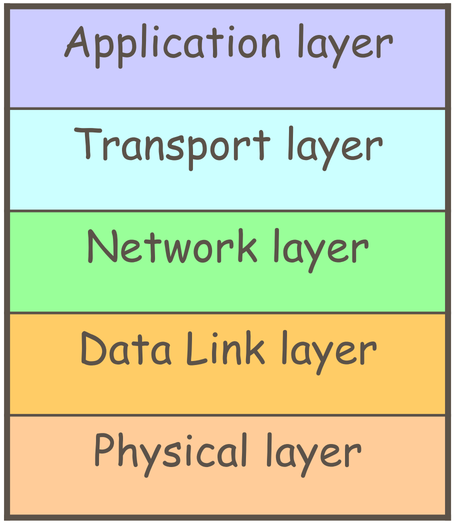
Wireless Mesh Network



- Packet relay (forwarding) routes

Resource Management in WiMAX

- The physical layer technology is based on the Orthogonal Frequency Division Multiple Access (OFDMA).
- In the MAC layer, WiMAX frames are constructed in two dimensions: subchannels in the *frequency* domain and OFDMA symbols (or time slots) in the *time* domain.
- User data are carried on areas called *bursts*, each consisting of a group of subchannels and associated OFDMA symbols.

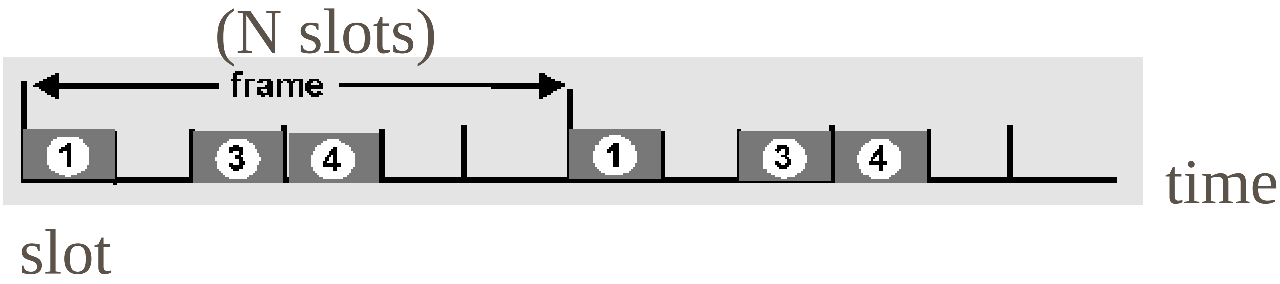


MAC (Medium Access Control) Protocols: a taxonomy

Three broad classes:

- Channel Partitioning
  - divide channel into smaller "pieces" (time slots, frequency, code)
  - allocate piece to node for exclusive use
- Random Access
  - channel not divided, allow collisions
  - "recover" from collisions
- "Taking turns"
  - Nodes take turns, but nodes with more to send can take longer turns

Channel Partitioning MAC protocols: TDMA



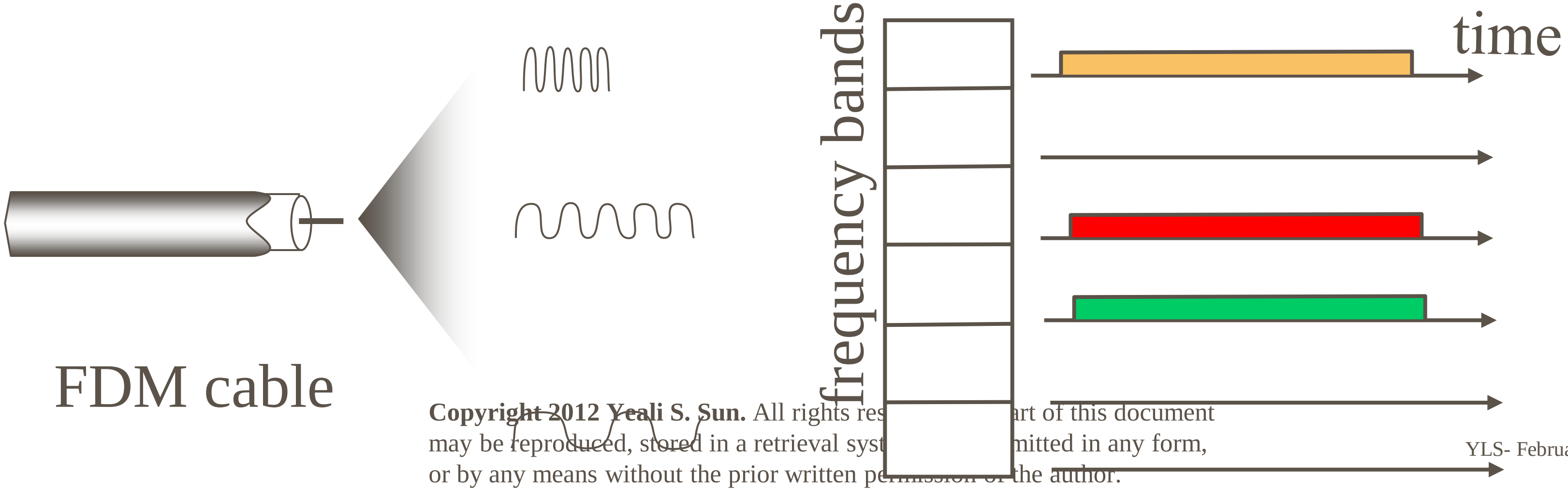
TDMA: time division multiple access

- channel divided into N time slots (frames)
- access to channel in "rounds"
- each station gets fixed length slot (length = pkt trans time) in each round
- unused slots go idle
- example: 6-station LAN, 1,3,4 have pkt, slots 2,5,6 idle
- inefficient with low duty cycle users and at light load.

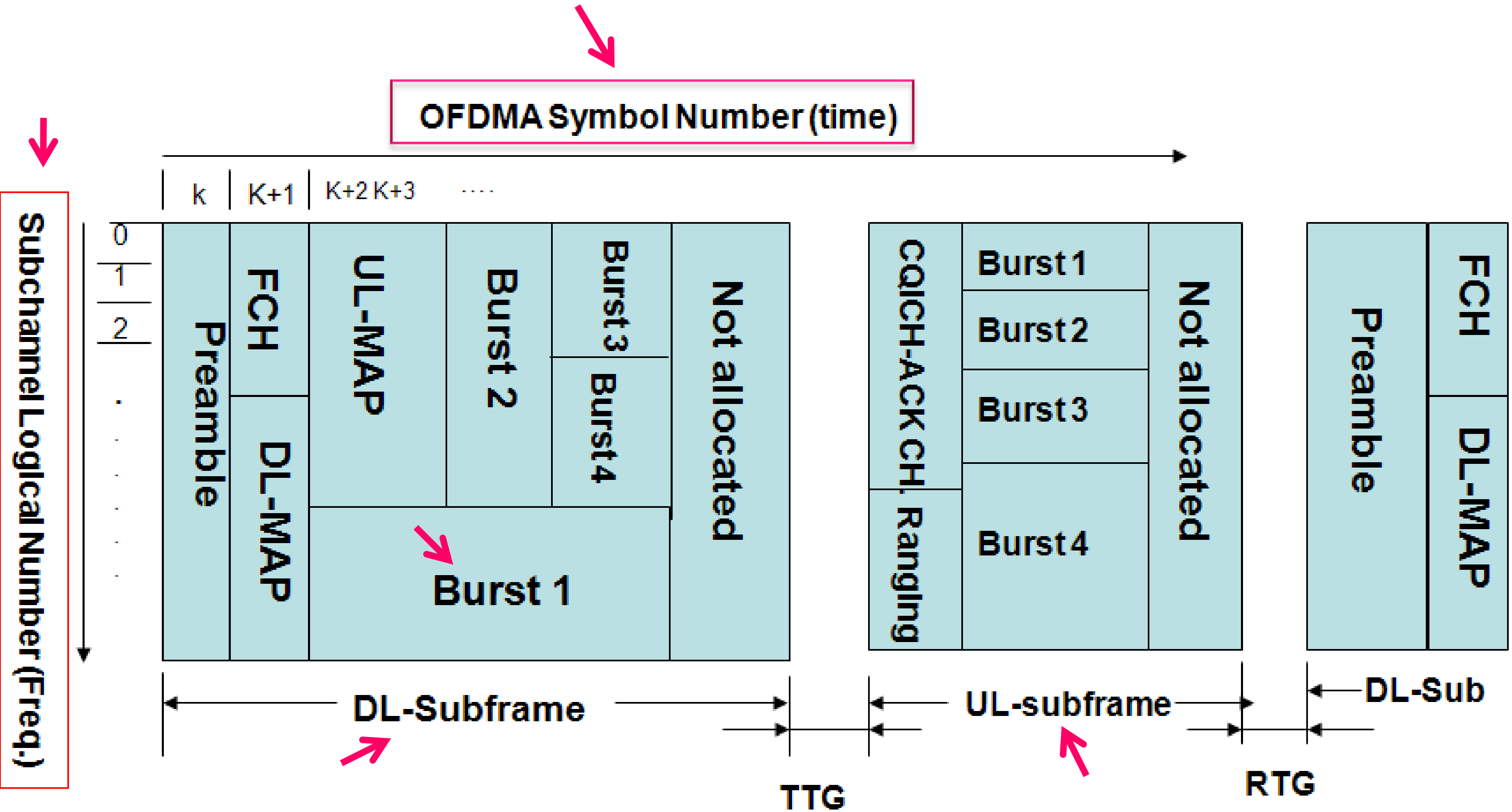
Channel Partitioning MAC protocols: FDMA

FDMA: frequency division multiple access

- channel spectrum divided into frequency bands
- each station assigned fixed frequency band
- unused transmission time in frequency bands go idle
- example: 6-station LAN, 1,3,4 have pkt, frequency bands 2,5,6 idle



WiMAX Frame Structure



Resource Management in WiMAX

- Data carried on the *same burst* are coded using the *same* coding and modulation, randomization, FEC coding and bit interleaving techniques.
- In WiMAX, the radio transmission quality is measured by CINR (Carrier to Interference plus Noise Ratio) or SINR (Signal)
- The *better the channel quality*, the data can be transmitted using *higher modulation methods* (e.g., 64QAM (3/4)), and *less number of OFDMA symbols* with *more data bytes per symbol*.
- Seven modulation options are specified in the standard.
- Each mode carries a different amount of data per subchannel and requires different maximum number of concatenated slots to carry a certain amount of data.

SINR

- First, for successful transmission, individual pair's SINR value must satisfy the capture threshold constraint, i.e.,

$$SINR = \frac{P_v}{noise + I} = \frac{P_u / d^{\alpha}(u,v)}{\eta + \sum_x \frac{P_x}{d^{\alpha}(x,v)}} \geq \gamma$$

  - where P_u and P_x are the transmission power of node u and node x, respectively.

Shannon Capacity

- The channel capacity C: the theoretical tightest upper bound on the rate of clean (or arbitrarily low bit error rate) data that can be sent with a given average signal power S through an analog communication channel subject to additive white Gaussian noise of power N, is:

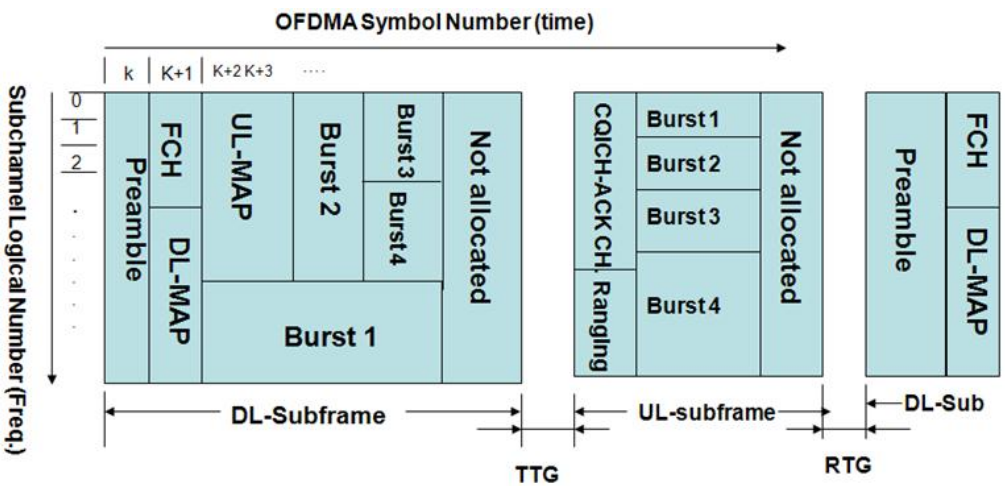
$$C = B \log_2 (1 + SINR)$$

where

- **C** is the channel capacity (bps);
- **B** is the bandwidth of the channel (Hz);

Network Resource - Allocation in WiMAX

- User data are carried in areas called **bursts**,
  - Each consisting of a group of subchannels *f* and associated OFDMA symbols *m*
  - Channel condition : Mode *k*



$$r_i^{WiMAX}(k) = B_k \times \sum_{j=1}^{n_i} (f_{b_j} \times m_{b_j}) \times F$$

- n_i: the total burst number
- b_j: the burst index
- F: the number of frame in a second

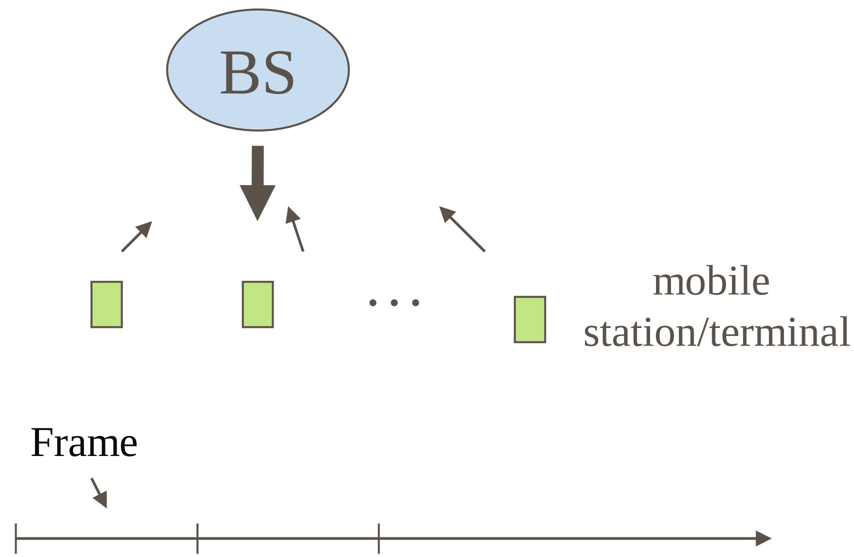
Mode (k)	Modulation	Useful data per slot (B_k)	Max number of concatenated slots	Max data payload
0	QPSK(1/2)	6 bytes	6	36 bytes
1	QPSK(2/3)	9 bytes	4	36 bytes
2	16QAM(1/2)	12 bytes	3	36 bytes
3	16QAM(3/4)	18 bytes	2	36 bytes
4	64QAM(1/2)	18 bytes	2	36 bytes
5	64QAM(2/3)	24 bytes	1	24 bytes
6	64QAM(3/4)	27 bytes	1	27 bytes

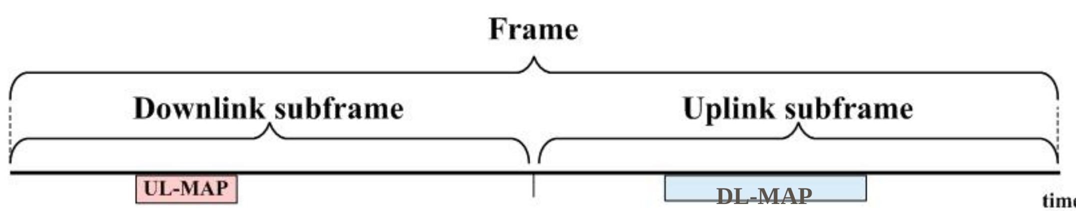
Broadband Wireless: IEEE 802.16d

- 5 different scheduling services to meet the QoS requirements of multimedia applications
  1. *UGS (*Unsolicited Grant Service)
  2. *rtPS* (real-time Polling Service)
  3. *ertPS* (extended real-time Polling Service)
  4. *nrtPS (*non-real-time Polling Service)
  5. *BE (*Best Effort)

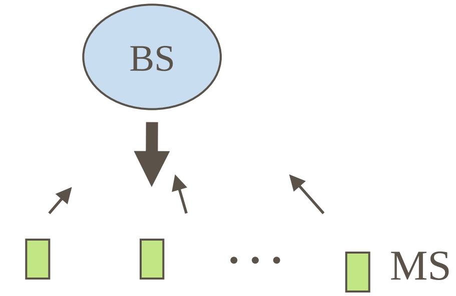
UGS - Unsolicited Grant Service

- Target for applications that generate **fixed-size** data packets at **periodic intervals**
  - T1/E1
  - VoIP without silence suppression
- An uplink service flow is granted by the BS at fixed intervals without additional polling or interaction.



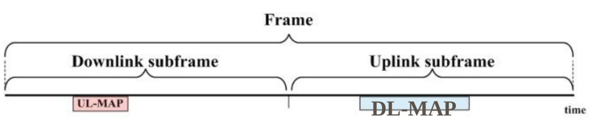


rtPS - real-time Polling Service



- Target for applications that generate variable-size data packets at periodic intervals with stringent latency and throughput requirements.
  - MPEG (Moving Pictures Experts Group) video
  - VoIP with silence suppression

- The fixed polling interval is typically short to meet the service flows' real-time needs.
- Key QoS parameters
  - Minimum reserved traffic rate
  - Maximum latency

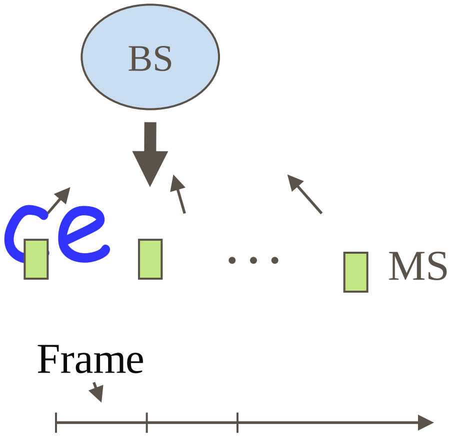


- rtPS connections are required to notify the BS of their current bandwidth requirements.
- The BS *periodically* grants unicast polls to rtPS connections
  - Polling period may be specified as an optional QoS parameter

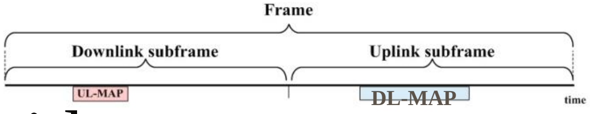
ertPS – extended-real-time Polling Service

- Additional service type defined in IEEE 802.16e
- Periodic fixed grants *unless* mobile stations notifies BS for grant *change*.

nrtPS - non-real time Polling Service



- Target for applications that
  - do ***not*** have any specific **delay requirements.**
  - require high throughput or need variable-sized data grants on a regular basis, such as high-bandwidth FTP.
- The scheduler sends unicast polls to mobile stations on a fixed interval to determine whether data is queued for transmission on a particular service flow.



- If data is queued, the scheduler provides a transmission grant for the *service flow*.

BE - Best Effort

- The scheduler grants transmit opportunities on a FCFS basis.
- Difference between nrtPS & BE
  - nrtPS is reserved a minimum amount of bandwidth
- nrtPS & BE both request bandwidth by
  1. Piggybacking
  2. Broadcast Polls

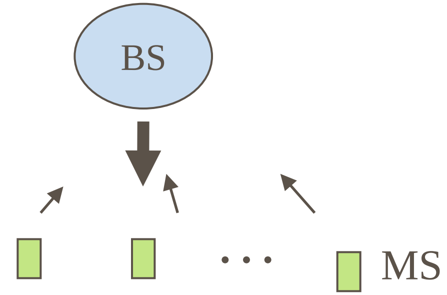
Summary: service and channel access scheduling

	Periodicity	Grant(s)	Operation mode
UGS	yes	fixed	
rtPS	yes	variable	User tells its current needs
ertPS	yes	fixed & variable	Allow user to change its needs
nrtPS	no	variable	BS polls, user tells
Best effort	no	no	Contend for telling BS their needs

Random Access Protocols

- When node has packet to send
  - transmit at full channel data rate R.
  - no *a priori* coordination among nodes
- two or more transmitting nodes ➜ "collision",
- random access MAC protocol specifies:
  - how to detect collisions
  - how to recover from collisions (e.g., via delayed retransmissions)
- Examples of random access MAC protocols:
  - slotted ALOHA
  - ALOHA
  - CSMA, CSMA/CD, CSMA/CA

QoS functions



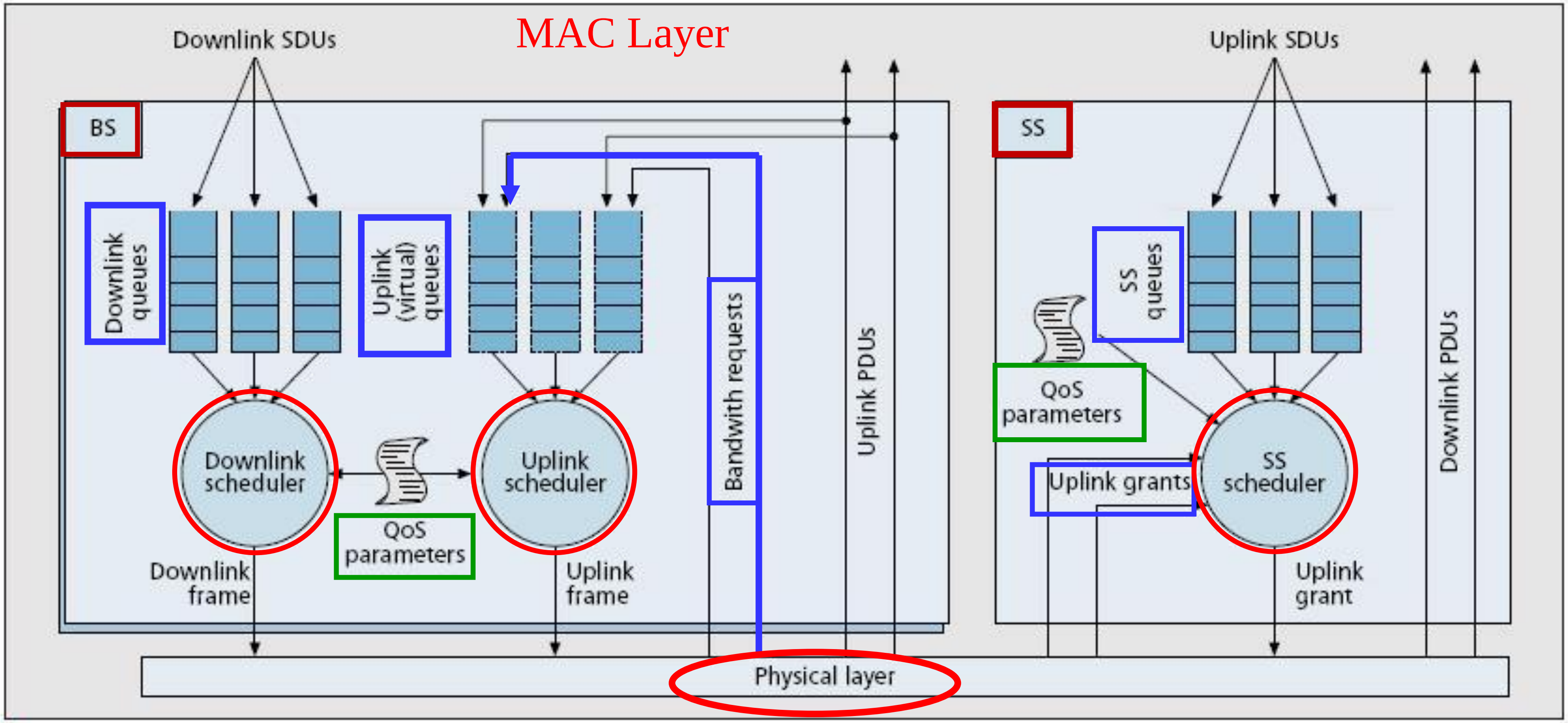


Figure 2. *QoS functions within the BS and SSs.*

UMTS (3G) QoS Traffic Classes

Traffic Class	Conversational Class	Streaming Class	Interactive Class	Background Class
Fundamental characteristics	Preserve time relation (variation) between information entities of the stream	Preserve time relation (variation) between information entities of the stream	Request response pattern	Destination is not expecting the data within a certain time
	Conversational pattern (**stringent and low delay**)		Preserve payload content	Preserve payload content
Example application	Voice	Streaming Video	Web Browsing	Background download of e-mails

UMTS (*Universal Mobile Telecommunications System*)

Quality of Service (QoS)

- An important provisioning in ATM networks to support *real-time continuous media*.
- A *traffic contract* (or service level agreement (SLA)) defines the service between **service user** and **service provider**
- " … identification of *parameters* that can be *directly observed and measured* … " at the *service access point*.
- The defined parameters apply to cell (packet) streams *conforming* to the negotiated traffic contract.
  - What about cells not conforming with the traffic contract?

QoS Parameters

ATM Cell Transfer

- “Successful” Cell transfer
  - No error and received within a specified time $T_{max.}$
- “Errored” Cell
  - Received within T_{max} but with error
- “Lost “ Cell
  - No cell is received within T_{max}, e.g., “never showed up” or “*late*”

ATM Cell Transfer (cont'd)

- "Misinserted" Cell
  - A received cell for which there is no corresponding transmitted cell
- "Severely-Errored" Cell Block
  - When M or more Lost/Misinserted/Errored cells are observed in a received cell block of N cells transmitted consecutively on a given connection. (M out of N)

Quality of Service (QoS) - Different Perspectives (1/3)

- User-level QoS
  - End-to-end QoS
  - Qualitative description
  - *User's quality of experience* (*QoE*)
    - how to improve the quality of offered services, as perceived by users
- Application-level QoS
  - Quantitative description
  - Application-oriented
  - Translation of User's QoS to Application's QoS is still an open research issue

Application Level QoS: example - H.264 (1/3)

- A codec technology employs ***layered coding***.
- It has been chosen to support *robust content delivery* on **a wide variety of networks and channel-type environments**
  - ranging from very low bit rate, low frame rate, and low resolutions for *mobile devices* to high bit rate and high resolution of *HDTV*, broadcast, DVD storage, RTP/IP packet networks, and ITU-T multimedia telephony systems.
- In order to enable **multi-vendor end devices** to successfully interwork with each other, H.264 defines the *profiles* and *levels* to assure all H.264 encoders and decoders (software and hardware) conform to the standard specification and are interoperable.
- It is useful for *streaming media applications*.

Application Level QoS: example - H.264 (2/3)

- A *level* is defined to place limits on the parameters such as sample processing rate, picture size and *coded bit rate* of user devices.
  - The former two parameters are related to the performance of user devices.
  - The latter specifies the *network QoS requirement*.
- *Sixteen* levels from 1 to 5.1 have been defined in H.264/AVC with the maximum bit rate ranging from 64kbps to 240Mbps.
- For the current 3G and WiMAX networking environments, *six* level numbers are considered.
- A H.264 video streaming service request must specify the requested level number.

H.264: different video quality levels for video streaming service (3/3)

Video Quality Level number	Max/Min bit rate (r_{max}/r_{min})
2	2048/768kbit/s
1.3	768/384kbit/s
1.2	384/192kbit/s
1.1	192/128kbit/s
1.b	128/64kbit/s
1	<=64kbit/s

VoIP QoS Criteria

- PESQ (ITU-T P.862)
  - MOS (Mean Opinion Score) value

- ITU G.107
  - E-Model

$$R = R_o - I_s - I_d - I_{e\text{-}eff} + A$$

  - R_o: basic signal-to-noise ratio
  - I_s: *simultaneous* impairment factor
  - I_d : *delay* impairment factor
  - $I_{e\text{-}eff}$: *equipment* impairment factor including *packet loss*
  - A : advantage factor

User/application Level QoS: example - VoIP PESQ (ITU-T P.862) - MOS (mean opinion score) value

R-value Lower limit	MOS	Speech Quality Category	User Satisfaction
90	4.34	Best	Very Satisfied
80	4.03	High	Satisfied
70	3.6	Medium	Some users dissatisfied
60	3.1	Low	Many users dissatisfied
50	2.58	Poor	Nearly all users dissatisfied

Quality of Service (QoS) - Different Perspectives (2/3)

- Transport-level QoS (end-to-end)
  - TCP layer QoS
  - UDP layer QoS
  - IP layer QoS
  - Physical layer (e.g., ATM connection, wireless channel, etc.) QoS
- Network Performance
  - Network Element Performance
  - Per-Hop Behavior

Quality of Service (QoS) - Different Perspectives (3/3)

- The issue is "*How to map users requirement to network support Quality of Service (QoS) service?*"

About Network Control and Management

- Planning
  - Time scale – a range of weeks, months, years
- Statistics vs. raw data
- Individual vs. aggregate
  - per user, per session, per flow, etc.
- Service vs. individual components (network elements) availability

Service Level Agreement (SLA)

- When a virtual circuit (connection, session, association, etc.) is established, both service user and service provider must *agree on* a contract defining the service, including
  - Source Traffic Description
    - To characterize the offered load
  - Requested QoS Class
    - QoS desired by the user and accepted by the provider
    - Measurable quantities
  - The Compliance Requirements
    - What if service user violates the contract?
    - What if service provider violates the contract?

QoS Parameters in ATM (1/6)

- Worst-case performance parameters
  - The service provider is required to **meet or exceed** it.
- Source traffic description
  - Peak Cell Rate (PCR)
    - **T=1/PCR -** the minimum spacing between cells
  - Sustained Cell Rate (SCR)
    - e.g., for CBR SCR=PCR; for other service categories, SCR << PCR

QoS Parameters in ATM (2/6)

- Minimum Cell Rate (MCR)
  - For ABR, MCR>=0
- (PCR/SCR) ratio has been used as one measure of traffic burstiness

"Source traffic characteristics specify how fast the user wants to send data."

- Statistical values – average over a certain period of time

QoS Parameters in ATM (3/6)

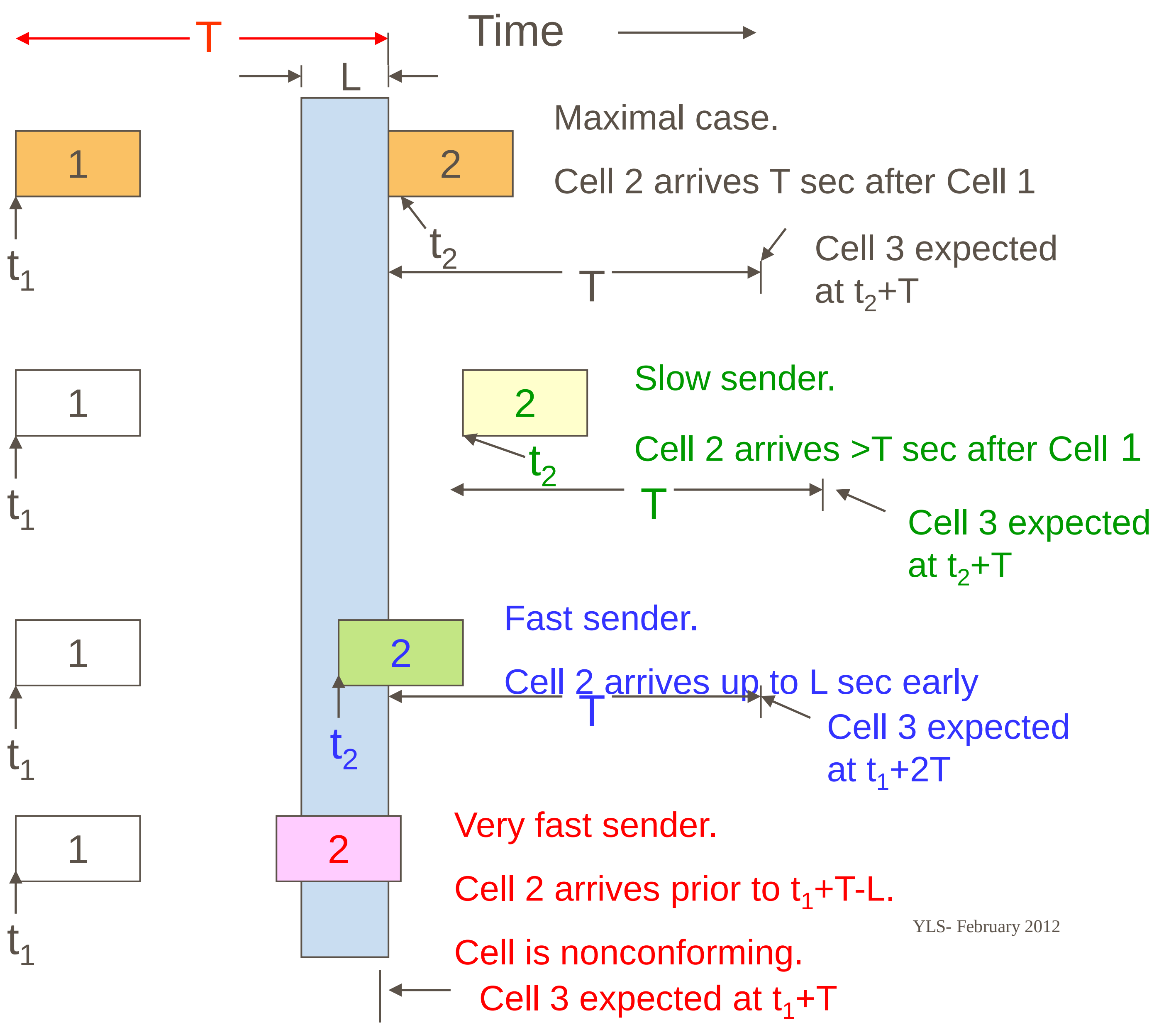
- Conformance check
  - Cell arrivals may be *statistically* described by an average arrival rate such as PCR or SCR.
  - In reality, cell arrivals may vary in time.

QoS Parameters in ATM (3/6)

- CVDT (Cell Variation Delay Tolerance)
  - It is used to *control the amount of variability acceptable*
  - It specifies how much variation will be present in cell arrivals at a network node
    - *conformance check*
    - *traffic shaping*
  - It is measured at the receiver

Conformance Check: cell arrival variation control

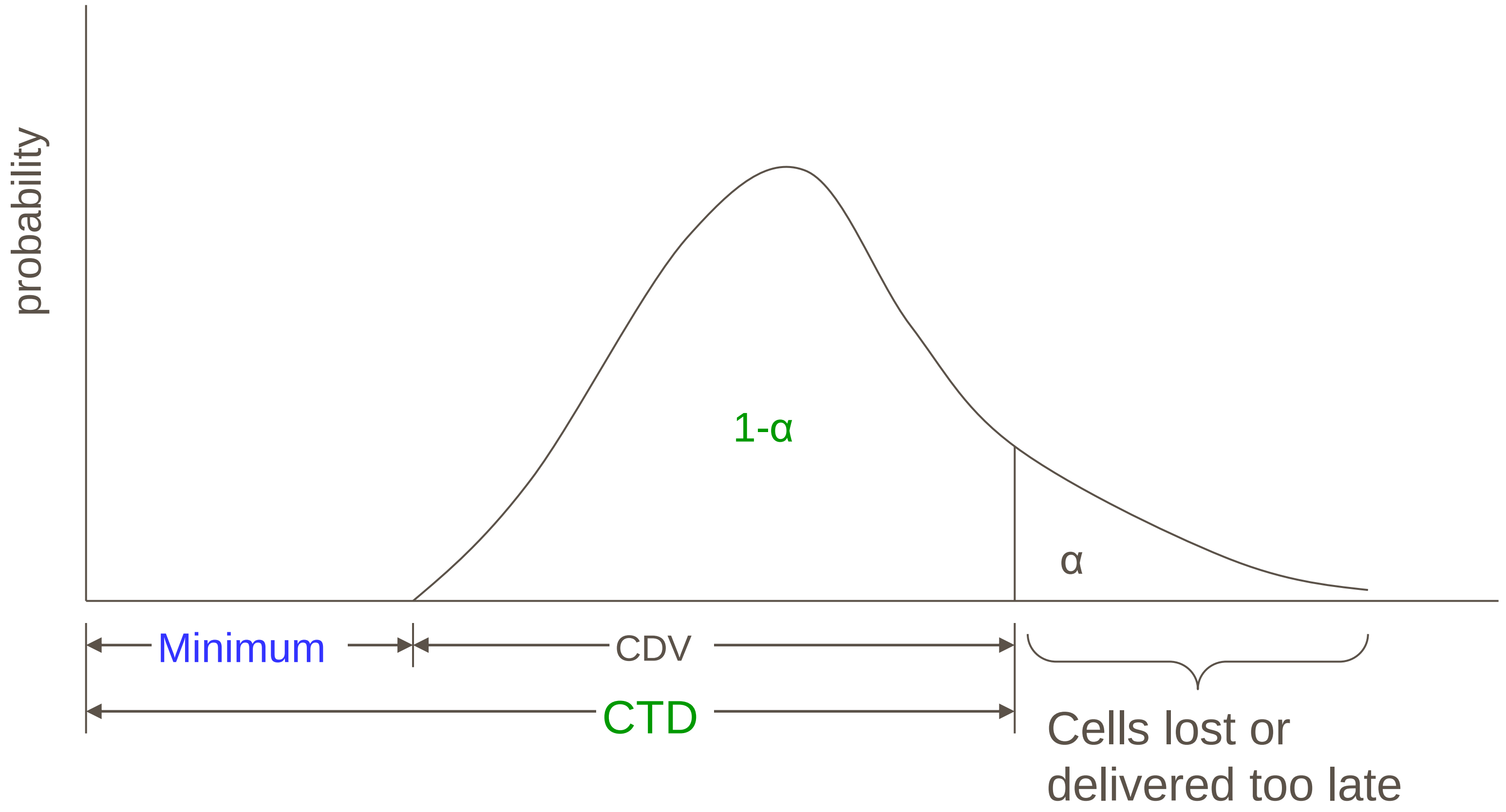
e.g., PCR
(T=1/PCR)



QoS Parameters in ATM (4/6)

- Specify the *characteristics* of network delivery service
  - Cell Loss Rate (CLR) (1-Cell_successfully_received/Cell_sent)
  - Cell Transfer Delay (CTD) ($t_{received}$ - t_{sent})
  - Cell Delay Variation (CDV) – how uniformly the cells are delivered

Cell Delay: distribution probability function



QoS Parameters in ATM (5/6)

- Cell Delay Jitter

$$D_{i+1} - D_i = (R_{i+1} - S_{i+1}) - (R_i - S_i)$$
$$= (R_{i+1} - R_i) - (S_{i+1} - S_i)$$
= inter-arrival time at receiver -
inter-departure time at sender

- Measured at the *receiver* …

- Specify the characteristics of the network
  - **CER (Cell Error Rate):** Fractions of cells delivered with error
  - **SECBR (Severely-Errored Cell Block Ratio):** Fraction of blocks garbled
  - **CMR (Cell Misinsertion Rate):** Fraction of cells delivered to wrong destination

QoS Parameters: summary (6/6)

Parameter	Acronym	Meaning
Peck Cell Rate	PCR	Maximum rate at which cells will be sent
Sustained Cell Rate	SCR	Long-term average cell rate
Minimum Cell Rate	MCR	Minimum acceptable cell rate
Cell Delay Variation Tolerance	CDVT	Maximum acceptable cell jitter
Cell Loss Ratio	CLR	Fraction of cells *lost* or delivered too *late*
Cell Transfer Delay	CTD	How long delivery takes (mean and maximum)
Cell Delay Variation	CDV	The variance in cell delivery times
Cell Error Rate	CER	Fractions of cells delivered with error
Severely-Errored Cell Block Ratio	SECBR	Fraction of blocks garbled
Cell Misinsertion Rate	CMR	Fraction of cells delivered to wrong destination

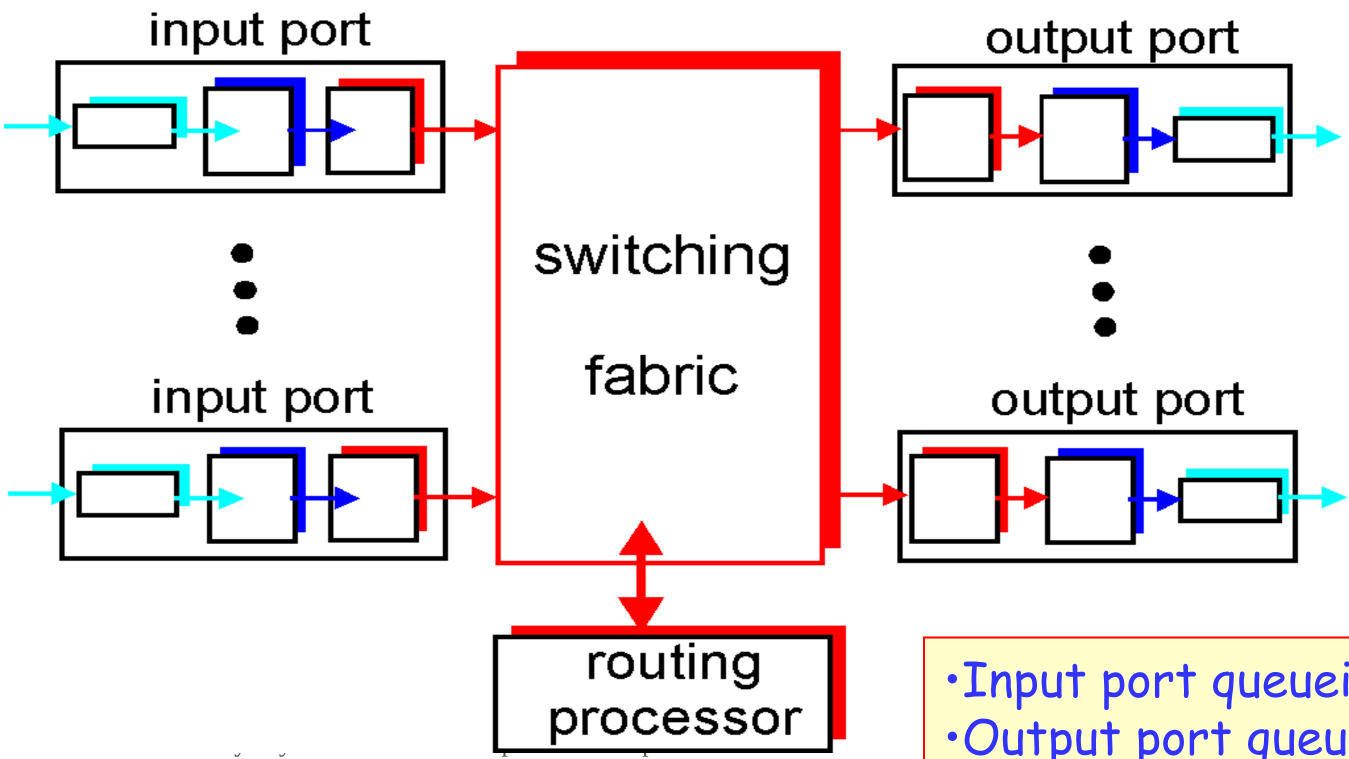
Sources of QoS Degradation (1/5)

- Propagation Delay
- Media Error
  - Including random and/or *burst* bit errors
- Switch architecture
  - Switch matrix design, e.g., blocking and non-blocking

Router Architecture Overview

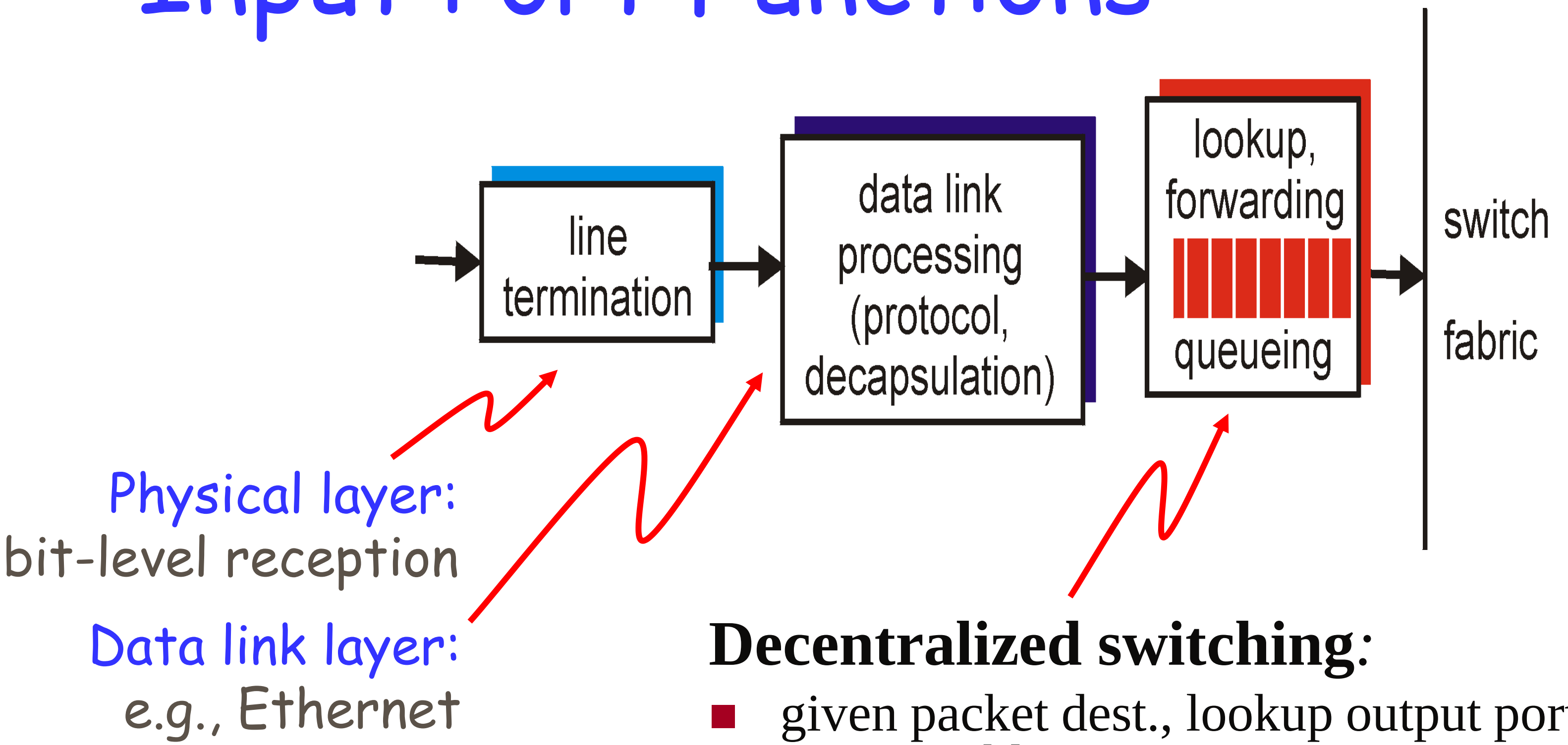
Two key router functions:

- run routing algorithms/protocol (RIP, OSPF, BGP)
- *switching packets* from incoming to outgoing link



- Input port queueing
- Output port queueing

Input Port Functions

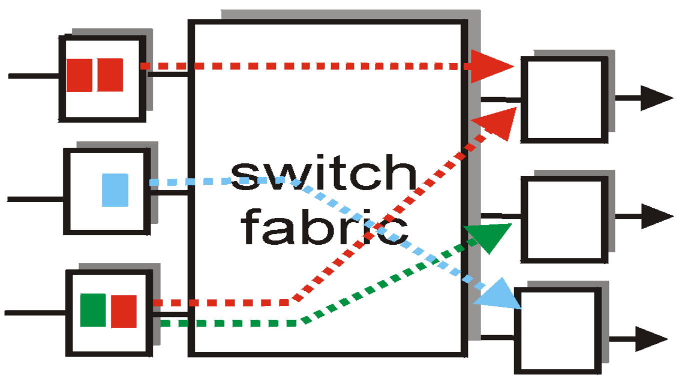


Decentralized switching:

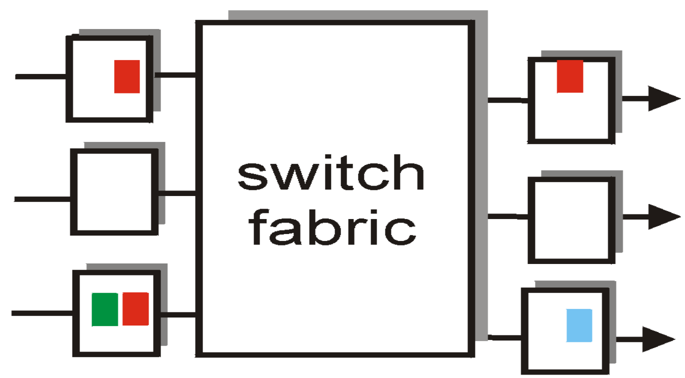
- given packet dest., lookup output port using routing table in input port memory
- goal: complete input port processing at 'line speed'
- queuing: if packet arrive faster than forwarding rate into switch fabric

Input Port Queueing

- Fabric slower than input ports combined -> queueing may occur at input queues
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward
- *queueing delay and loss due to input buffer overflow!*

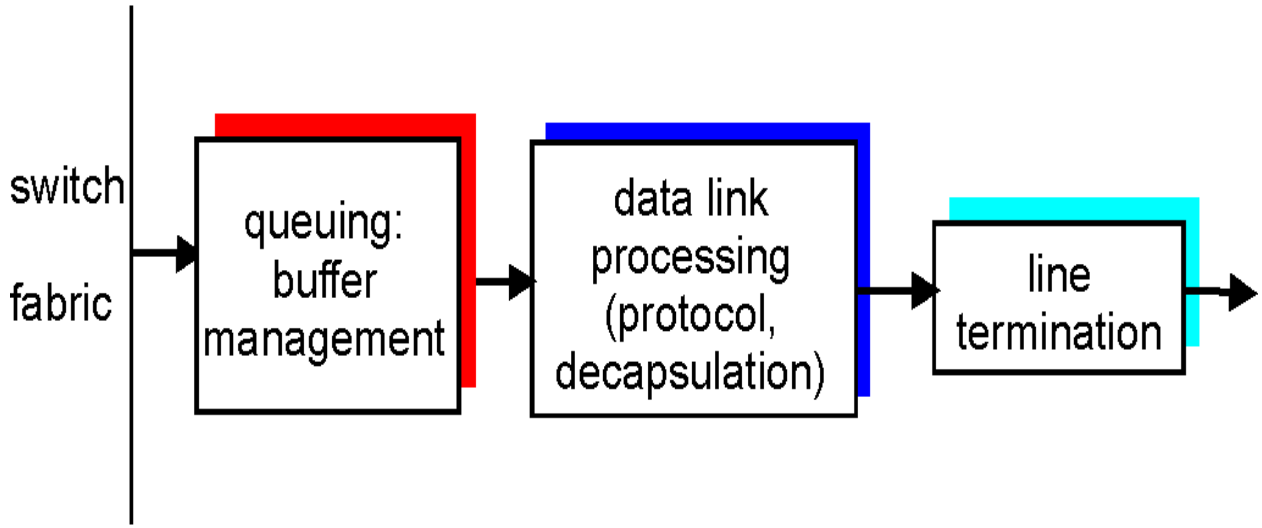


output port contention at time t - only one red packet can be transferred



green packet experiences HOL blocking

Output Ports

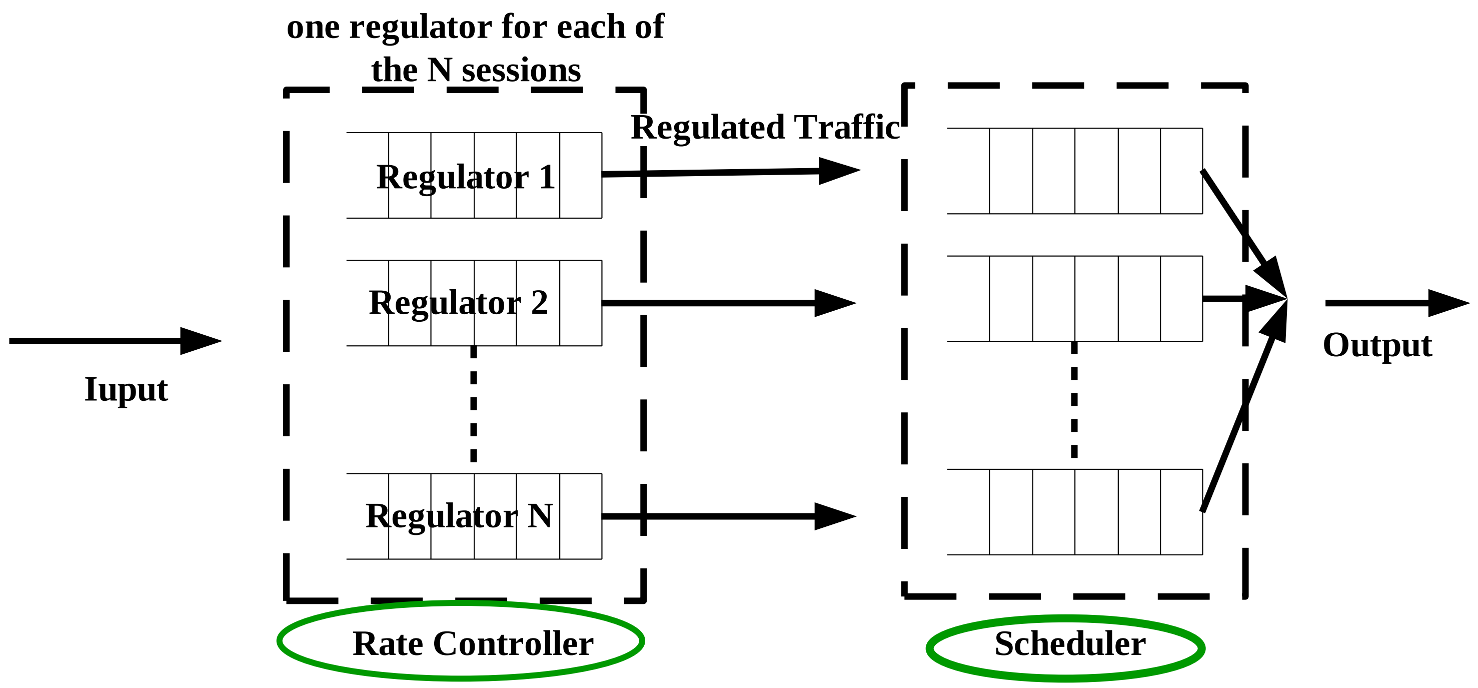


- *Buffering* required when packets arrive from fabric faster than the transmission rate
- *Scheduling discipline* chooses among queued packets for transmission

Sources of QoS Degradation (2/5)

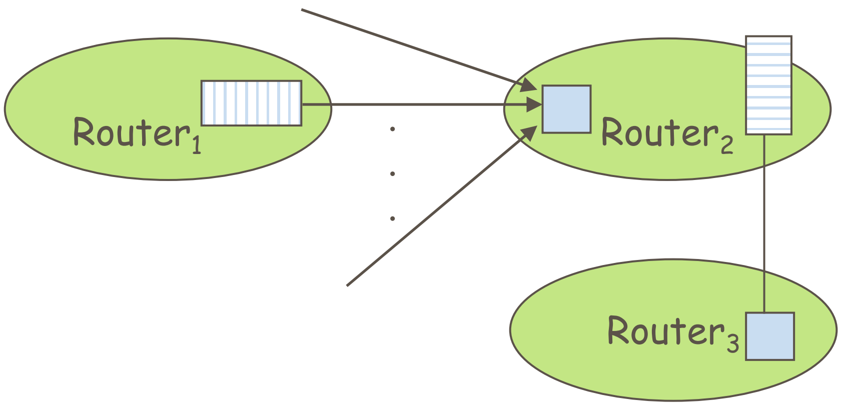
- Buffer strategy
  - Buffer space management: dedicated buffer per port, shared buffer pool or some combination;
- Packet transmission scheduling
  - Scheduling is to determine which packet to send next and are used primarily to manage the allocation of bandwidth among flows.
  - FIFO, priority or complex queueing disciplines (e.g., weighted fair queueing (WFQ))

Scheduling with traffic rate control



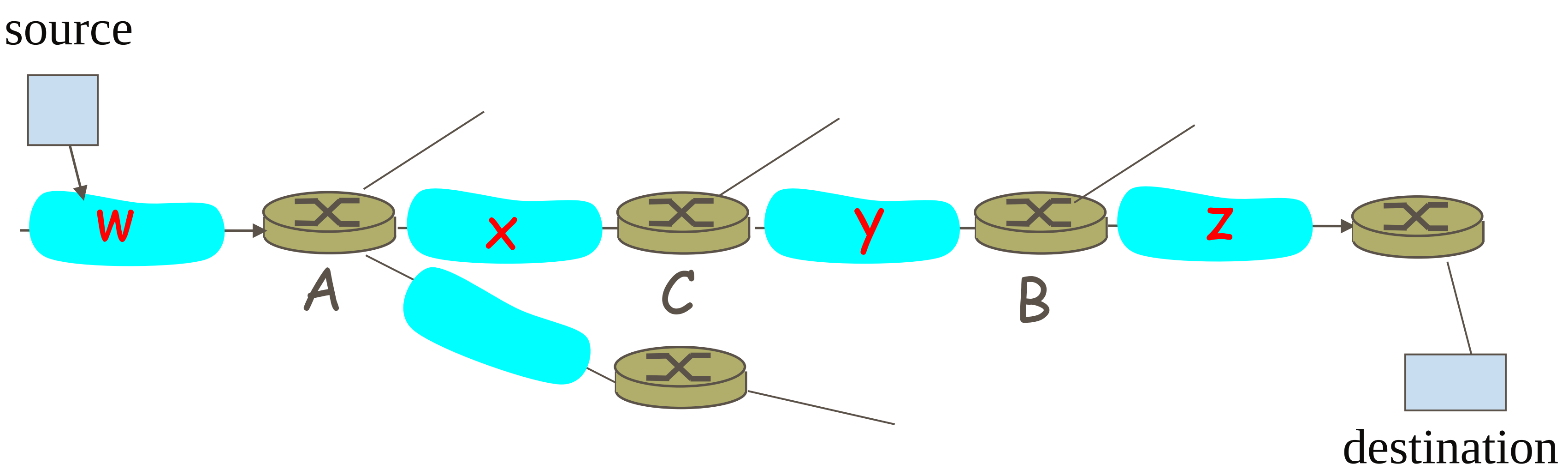
Sources of QoS Degradation (3/5)

- Traffic Load
  - The load offered on the route



Sources of QoS Degradation (4/5)

- Congestion
  - Bottleneck link
  - Congestion control
    - TCP
  - Queue management (Tail drop vs. Random Early Drop (RED))



Queue Management

- It is to manage the length of packet queues by dropping packets when necessary or appropriate.
- Active queue management
- The traditional technique for managing router queue lengths is "tail drop".
  - A max length (in terms of packets) is set for each queue.
  - Incoming packets are accepted for the queue until the max length is reached, then drop subsequent incoming packets until the queue decreases.

The need for active queue management

- Two important drawbacks:
  - Lock-out : in some situations tail drop allow a single connection or a few flows to monopolize queue space, preventing other connections from getting room in the queue.
  - However, this does not take into account that packet bursts play in Internet performance.

Sources of QoS Degradation (5/5)

- Resource allocation
  - Per packet, session, class, link, etc.
- Failures
  - Events that impact availability, e.g., port failures, switch failures or link failures

Impact of QoS Degradation on Performance Parameters

Attributes	**CER**	**CLR**	**CMR**	**MCTD**	**CDV**
Propagation Delay				X	
Media Error Statistics	X	X	X		
Switch Architecture		X		X	X
Buffer Capacity		X		X	X
Number of Tandem Nodes	X	X	X	X	X
Traffic Load		X	X	X	X
Failures		X			
Resource Allocation		X		X	X

Traffic Control

Introduction

- Traffic is policing at the edges of the network and packets are either permitted to go or dropped
- Traffic Contract
  - **Source traffic description**
  - **Requested QoS class**
- Conformance Check
  - User Parameter Control (UPC) - identify conforming or non-conforming packets
  - Actions for non-conforming traffic
  - Mechanisms, e.g., Leaky bucket, GCRA

Introduction (cont'd)

- Call Admission Control (CAC)
  - To control the overall traffic *permitted* to enter the network

- *CAC and UPC are network operator specific*

To be continued. ☺

TCP

- Congestion Control
- Random Early Drop (RED)

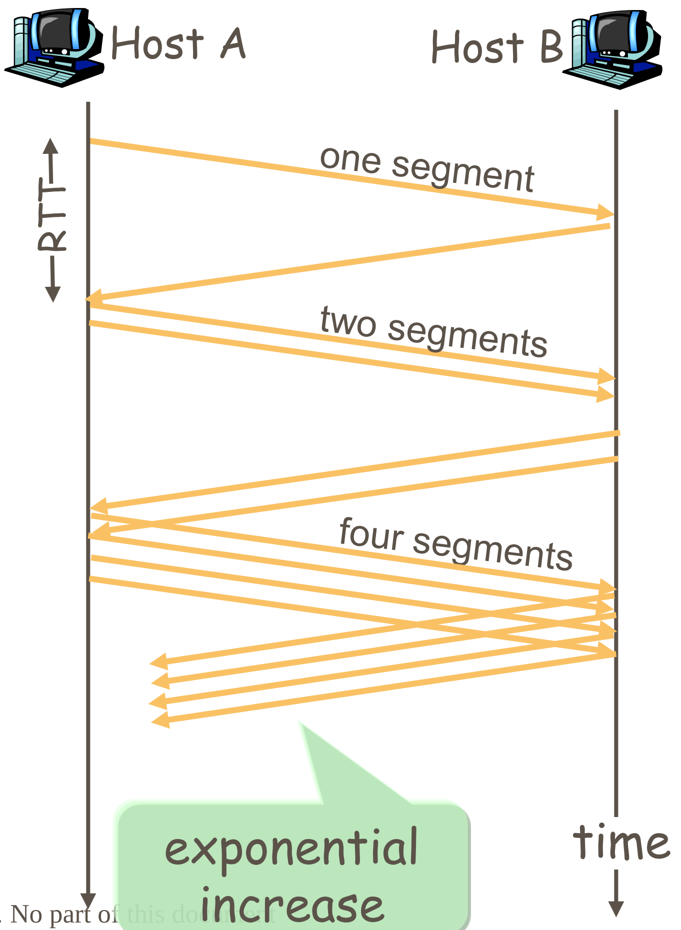
TCP Congestion Control

- Slow start
- Congestion avoidance

TCP Slow Start (1/3)

-> PROBE network's maximum "throughput"!

- When connection begins, increase rate exponentially until first loss event:
  - Double **cwnd** every RTT
  - done by incrementing `cwnd` for every ACK received
- Summary: initial rate is slow but ramps up exponentially fast



TCP Slow-Start (2/3)

- To get data flowing there must be acks to clock out packets; but to get acks there must be data flowing.
- Maintain a per connection state variable in the sender – "congestion window" *cwnd*
- "When to enter Slow-Start Phase?"
  - When a connection begins
  - After a timeout

TCP Slow Start (3/3)

- Algorithm –
  - When starting or restarting after a loss, set cwnd=1 packet.
  - Each time an ACK is received, *cwnd* is incremented by one segment size, i.e. one ack for each new data,

    cwnd =cwnd+1.
  - When sending, send the min(receiver's_advtiseWin, cwnd)
- *cwnd* is maintained in bytes.
  - The *segment size* is announced by the *receiver*.

Congestion Avoidance

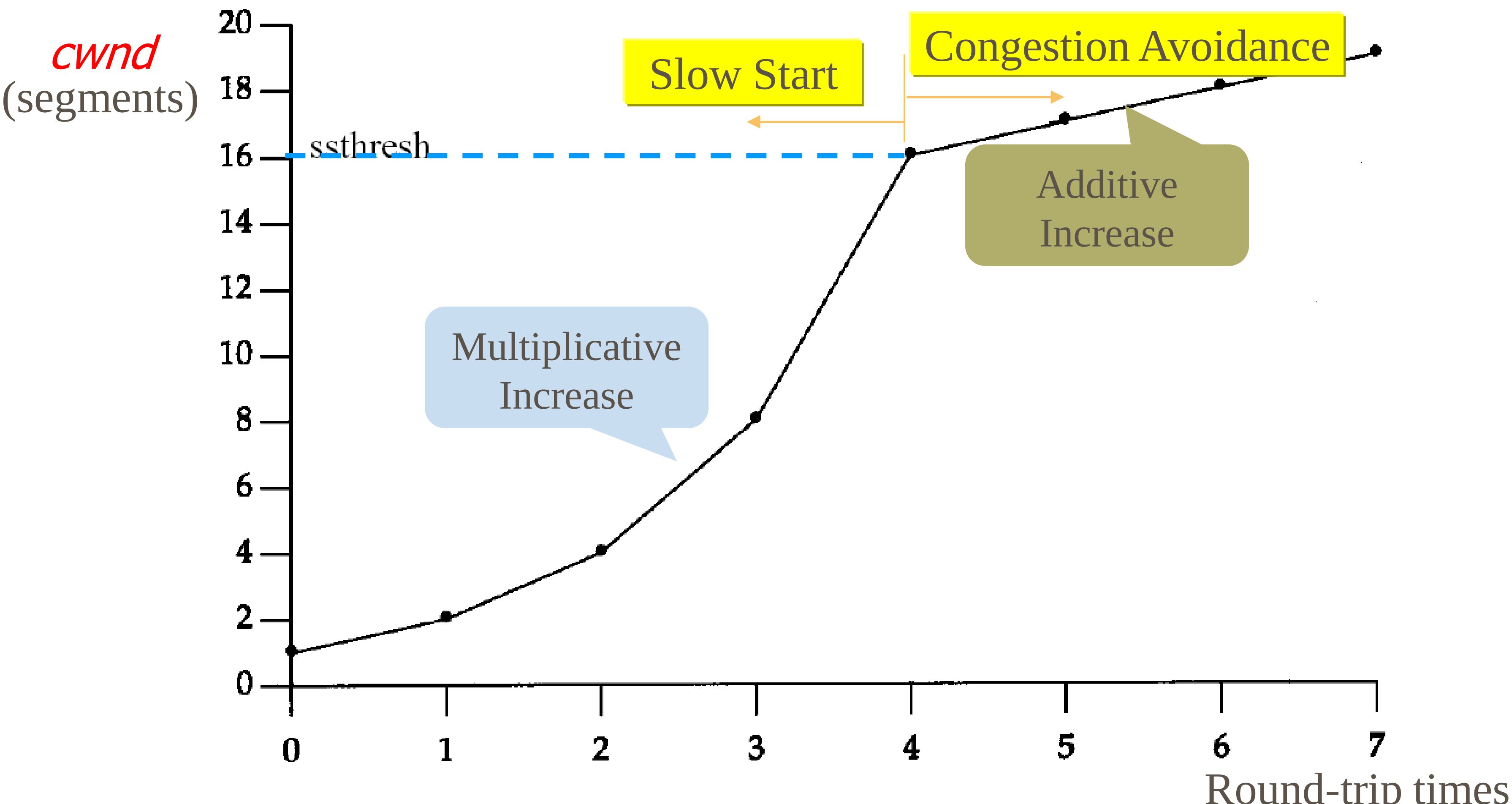
- Congestion is indicated by a *timeout* or the reception of *three* duplicate ACKs.

- The goal is to avoid increasing the window size too quickly and causing additional congestion.

Congestion Avoidance Algorithm

- Slow start phase
  - When a connection begins: *cwnd* is one segment and *ssthresh* (slow start threshold) is 65,535 bytes.
  - When congestion occurs, *ssthresh=cwnd/2*, cwnd=1
- Once *cwnd=ssthresh*, the connection enters the congestion avoidance phase.
  - On each ack for new data, cwnd=cwnd+1/cwnd (additive increase)
  - When sending, send the min(receiver's AdvertiseWinow, cwnd)

AIMD: *additive increase, multiplicative decrease*



※Visualization of slow start and congestion avoidance.

Summary of TCP congestion control

- *Approach:* increase transmission rate (window size), probing for usable bandwidth, until loss occurs
  - *additive increase:* increase **CongWin** by 1 MSS every RTT until loss detected
  - *multiplicative decrease*: cut **CongWin** in half after loss

Saw tooth behavior: probing for bandwidth