Communication Systems

Benjamin Ding

November 12, 2018

1 Introduction

Communication links

- Simplex, unidirectional
- Half-duplex, bidirectional
- Full-duplex, bidirectional

Metcalfe's Law

Value of a network with N nodes is proportional to the number of connections that may be made between the nodes, or $O(N^2)$.

Network Scales

- PAN(Personal Area Network) Bluetooth
- LAN(Local Area Network) WiFi, Ethernet
- MAN (Metropolitan Area Net) Cable,DSL
- WAN(Wide Area Network) Large ISP

Network Edge

- Applications and hosts
- Both clients and servers.
- In other words, end system

Access Networks

- The physical links that connect an end system to its edge router (first router on path to destination)
- Residential access networks
- Institutional access networks

- Mobile access networks

Network core

- Interconnected routers
- Network of networks

Connecting ISPs

- Connecting each access ISP to each other directly doesnt scale: O(N) connections.
- Connect each access ISP to a global transit ISP
- If there are a number of global ISPs, then they have to be interconnected.

Internet exchange point

- Content providers may run their own network to bring services content close to end users
- Tier-1 commercial ISPs (e.g., Level 3, Sprint, AT&T, NTT), national & international coverage
- Content provider network (e.g, Google): private network that connects its data centers to Internet, often bypassing tier-1, regional ISPs

Wireless Networks

- Wireless LANs
  - Within building (100 ft)
  - WiFi: IEEE 802.11 b/g/n: 11, 54, 600 Mbps
- Wide Area Wireless Access
  - Provided by telco (cellular), 10s of kms
  - Between 1 and 100 Mbps
  - 3G, 4G: LTE

Digital subscriber line (DSL)

- Use existing telephone line to central office DSLAM
- Data over DSL phone line goes to Internet
- Voice over DSL phone line goes to telephone net

Frequency division multiplexing

- Different channels transmitted in different frequency bands

HFC: hybrid fiber coax

- Asymmetric: the downstream/upstream transmission
- Rate ratio as high as 20.

Cable network

- Network of cable, fiber attaches homes to ISP router
- Homes share access network to cable headend
- Unlike DSL, which has dedicated access to central office

Optical Fiber Networks

- Passive Optical Networks
- Optical Network Terminal
- Gigabit-capable Passive Optical Networks (GPON)
- FTTH (expensive)
- FTTN (reuse existing copper infrastructure)
  - Homes share access network to street cabinet, ONU
  - Each home has an ONT
  - OLT at central office

2 Data Transfer In Networks

Protocols

- Format of messages,
- Order of messages sent and received among network entities,
- Actions taken on message transmission and receipt

Client/server model

- Client host requests, receives service from always-on server

Peer-to-peer model

- Minimal (or no) use of dedicated servers

Connection-oriented service

- State is kept at each end, a connection exists!
- An end-to-end path may have to be setup (implies network state)
  - Like an agreed meeting
  - Takes time to setup, overheads, but thorough job

Connection-less service

- Data is just sent towards the destination, ‘no expectations
- Like speaking out of the blue to someone (who may not hear)
- Lightweight, ‘cheap, but no guarantees

Routing

- Determines source-destination route taken by packets.
- Forwarding: move packets from routers input to appropriate router output.

Circuit switching

- Dedicated circuit per call: telephone net
- End-end resources reserved for call
- Call setup required
- Dedicated resources: no sharing
- Circuit-like (guaranteed) performance
- Resource reservation needed (smart nodes)
- Must divide resources into pieces
- Can be wasted (idle) if not needed
- Classic allocation methods: TDM, FDM

Packet-switching

- Data sent thru net in discrete chunks
- Each end-end data stream divided into packets
- User A, B packets share network resources
- Each packet uses full link bandwidth
- Resources used as needed
- Great for bursty data

  - Resource sharing
  - Simpler, no call setup
- Must control congestion: packet delay and loss
  - Protocols needed for reliable data transfer, congestion control
- Need to provide circuit-like bandwidth guarantees for audio/video apps

Resource contention

- Aggregate resource demand can exceed amount available
- Congestion: packets queue, wait for link use
- Store and forward: packets move one hop at a time
- Node receives complete packet before forwarding

Benefits of Packets

- More resilient to errors
- Statistical multiplexing advantage
- Resilience to errors: Error resilience example 60

Statistical multiplexing

- Allows more users to use network! Example 63
- Leveraging statistical multiplexing implies queues to absorb temporary overloads, which generate both delays and losses of packets.

3 Network Layering

Typical communication network tasks

- Find a path through the network
- Activate the direct data communication path
- Ascertain that the destination system is prepared to receive data.
- Application on the source system must ascertain that its counterpart on the destination system is prepared
- Transfers information reliably
- Send as fast as the network allows while sharing the bandwidth
- (optionally) secures information in transit

Layers

- Each instance of a protocol talks virtually to its peer using the protocol
- Each instance of a protocol uses only the services of the lower layer
- Each layer implements a service
  - Via its own internal-layer actions
  - Relying on services provided by layer below
- To handle system complexity
- Explicit structure allows identification, relationship of complex systems
- Modular structure
  - Eases maintenance, updating of system
  - Change of implementation of layers service transparent to rest of system
- High/low layers are usually software/hardware respectively
- Disadvantages
  - Encapsulation brings overhead (more on this later)
  - But minor for long messages
  - Hides information

OSI Model

- Application: network applications, File transfers, web browsing, email
- Presentation: allow applications to interpret meaning of data, e.g., encryption, compression, machine-specific conventions
- Session: synchronization, check pointing, recovery of data exchange
- Transport: transports application-layer messages between application end point (end-end delivery)
- Network: routing network-layer packets from source to destination (routing protocols)
- Link: data transfer between neighboring network elements (Ethernet)
- Physical: bits on the wire

Internet protocol stack

- Missing presentation and session layers
- If needed must be implemented in application

Protocol Data Units (PDU)

- In the layered architecture, user data is passed from layer to layer
- Control information is added/removed to/from user data at each layer (header, trailer)
- Data + header + trailer = PDU
- Each layer has a different PDU

4 Application Layer

Process

- Program running within a host
- Within same host, two processes communicate using inter-process communication (defined by OS)
- Processes in different hosts communicate by exchanging messages
- Client process: process that initiates communication
- Server process: process that waits to be contacted

App-layer protocol

Defines

- Types of messages exchanged e.g., request, response
- Message syntax: what fields in messages & how fields are delineated
- Message semantics: meaning of information in fields
- Rules for when and how processes send & respond to messages
- Public-domain protocols:
  - Defined in 'Requests for Comments (RFCs)
  - Allows for interoperability e.g., HTTP, SMTP
  - Proprietary protocols: e.g., Skype

Transport Layer Security (TLS)

- Cryptographic protocols that provide communications security over a computer network
- Widely used in web applications browsing, email, instant messaging,
- Browsers increasingly expect TLS on the web and treat unencrypted websites as exception.
- Secure Sockets Layer (SSL), the predecessor, is now deprecated by IETF

5 Transport Layer

Sockets

- Process sends/receives messages to/from its socket (analogous to door)
  - Sending process shoves message out door
  - Sending process relies on transport infrastructure on other side of door to deliver message to socket at receiving process
- Also referred to as the API between the application and the transport layers (network).
- The application developer has control of everything on the application-layer side of the socket
- The application developer has little control on the transport-layer side of the socket (can choose transport protocols and specify some parameters)

Socket Primitives

- Page 93

Addressing processes

- To receive messages, process must have identifier
- Host device has unique 32 bit IP(v4) address
- Identifier includes both IP address and port numbers associated with process on host.

6 Physical Layer and Media

Sharing Physical Media

- Coding (represent data)
- Modulation (use a carrier)
- Multiplexing (sharing the medium)
- Frequency Division Multiple Access (FDMA)
- Wavelength Division Multiplexing (WDM)
- Time Division Multiple Access (TDMA)
- Code Division Multiple Access (CDMA)

CDMA

- Allows each station to use the entire allocated frequency spectrum all the time.
- Multiple simultaneous transmissions are separated using coding theory:

- Each station is assigned a unique code called a chip sequence, which are pairwise orthogonal.
- A key measure of information is entropy, representing the average number of bits needed to store or communicate one symbol in a message.
- Channel capacity is the upper bound on the rate of information that can be reliably transmitted over a communications channel.
- Source coding (e.g. for data compression)
- Channel coding (e.g. error-correcting codes)

Data

- Magnetic, optical Media
- Solid-state drives (SSD)

6.1 Guided media

Signals propagate in solid, controlled media:

Twisted Pair (TP)

- Two insulated copper wires
- Category 3: phone wires,
- Category 5: 10 Mbps Ethernet ()
- Category 6: 100Mbps, 1Gbps Ethernet used in DSL, Ethernet

Coaxial cable

- Two concentric copper conductors
- Bidirectional
- Broadband
- Multiple channels on cable
- HFC (Hybrid fibre coax)
- Used for cable TV, Internet
- Note the upstream/downstream asymmetry

Fiber optic cable

- Glass fiber carrying light pulses, each pulse a bit
- High-speed operation
- High-speed point-to-point transmission (e.g., 10s-100s Gbps)

- Low error rate
- Repeaters spaced far apart
- Immune to EM-noise

Passive Optical Network

- Devices
  - OLT (Optical Line Terminal)
  - ONU (Optical Network Unit) multiple subscriber terminal
  - ONT (Optical Network Terminal) single subscriber terminal
- Technologies
  - GPON (Gigabit PON)
  - EPON (Ethernet PON)

6.2 Unguided media

- Signals propagate freely, e.g. radio
- No physical wire
- A publicly regulated, scarce resource!

Spectrum Types

- The Electromagnetic Spectrum
- Radio Transmission
- Microwave Transmission
- Infrared and Millimeter Waves
- Lightwave Transmission

Effects

- reflection
- obstruction by objects
- interference

Radio link types

- Terrestrial microwave
- LAN (e.g., Wifi)

- Wide-area (e.g., cellular)
- 4G, LTE, 5G: 10s to 100s Mbps
- Satellite
  - 270 msec end-end delay
  - Geosynchronous versus low altitude (less delay)

Near Field Communications: RFID and NFC

- The RFID reader always transmits a signal.
- Backscatter is a low-energy way for the tag to create a weak signal of its own that shows up at the reader.
- For the reader to decode the incoming signal, it must filter out the outgoing signal that it is transmitting.
- NFC (Near field communication) is a subset of RFID with a shorter range for security purposes.

Mobile Communications: GSM

- 124 frequency channels
- each uses an eight-slot TDM system

7 Link Layer

- Concerns how to transfer messages over link(s)
- Messages (here) are limited-size frames
- Link layer builds upon the physical layer
- Hosts and routers are nodes
- Communication channels that connect adjacent nodes along communication path are links
- Data unit is a frame, encapsulates packet from layer 3
- Traditionally implemented in "adaptor" (Network Interface Card) along with physical layer

Services

- Framing, link access
  - Encapsulate data packet into frame, adding header, trailer
  - Channel access if shared medium
  - MAC addresses used in frame headers to identify source, destination
- Reliable delivery between adjacent nodes

  - Seldom used on low bit-error link (fiber, some twisted pair)
  - Wireless links: high error rates
- Flow control: pacing between adjacent sending and receiving nodes
- Error detection
- Error correction (without retransmission)
- Half-duplex and full-duplex

Sending side

- Encapsulates packet (from network layer) in frame
- Adds error checking bits, reliable data transfer, flow control

Receiving side

- Looks for errors, reliable data transfer, flow control
- Extracts data, passes to upper layer at receiving side

8 Error Detection and Correction

- Detection-only has less overhead but requires retransmission. Acceptable when error rate is low, e.g. in wired and fiber.
- Detection and correction requires more overhead but avoids retransmission. Used often in wireless communication links which are naturally more error-prone than wired ones.
- Not 100% reliable!
  - Protocol may miss some errors, but rarely
  - Larger EDC field yields better detection and correction

Coding Basics

- Frame consisting of m data bits (i.e.,message) and r check or parity bits. Define n=m+r.
- In an (n,m) code, the n-bit unit containing data and check bits is referred to as an n-bit codeword.
- The code rate, or simply rate, is the fraction m/n.
- In a block code, the r check bits are computed solely as a function of the m data bits.
- In a systematic code, the m data bits are sent directly, along with the check bits.
- In a linear code, the r check bits are computed as a linear function of the m data bits.
- Exclusive OR (XOR) or modulo2 addition is a popular choice for linear block codes.

Measures of effectiveness

- The minimum distance of the code
- The burst-detecting capability
- The probability that a completely random string will be accepted as error-free

Parity Checking

- Single Bit: Detect single bit errors
- Two Dimensional: Detect and correct single bit errors
- Interleaving: burst error detection (compute parity in different order to transmission)

Hamming Distance

- The number of bit positions in which two codewords differ is called the Hamming distance.
- If two codewords are a Hamming distance d apart, then it requires d single-bit errors to convert one into the other.

Hamming code

- The bits of the codeword are numbered consecutively, starting with bit 1 at the left end, bit 2 to its immediate right, and so on.
- The bits that are powers of 2 (1, 2, 4, 8, 16, etc.) are check bits.
- The rest (3, 5, 6, 7, 9, etc.) are filled up with the m data bits.
- Each check bit forces the modulo 2 sum, or parity, of some collection of bits, including itself, to be even (or odd)
- (11,7), distance 3, i.e. corrects single, detects double.

Internet checksum

- Goal: detect errors (e.g., flipped bits) in transmitted packet (widely used in networks and storage)
- Sender:
  - Treat segment contents as sequence of 16-bit integers
  - Checksum: addition (1s complement sum) of segment contents
  - Sender puts checksum value into checksum field
- Receiver:
  - Compute checksum of received segment
  - Check if computed checksum equals checksum field value
    - NO - error detected

    - * YES - no error detected but could still be errors

Cyclic Redundancy Check

- View data bits, D, as a binary number
- Choose r+1 bit pattern (generator), G
- Goal: choose r CRC bits, R, such that
  - $<D,R>$ exactly divisible by G (modulo 2)
  - Receiver knows G, divides $<D,R>$ by G. If non-zero remainder: error detected!
  - Can detect all burst errors less than r+1 bits
- Widely used in practice (Ethernet, 802.11 WiFi, ATM)
- CRC Example

9 Serial Communications

- Serial: One bit at a time, sequentially
- Parallel: Several bits are sent as a whole, on a link with several parallel channels.

Universal asynchronous receiver/transmitter (UART)

- Converts between parallel and serial data and handles other low-level details of serial communications.
- It is a hardware component (often within CPU or SoC) that implements a variety of serial protocols in embedded systems, e.g. RS232, RS485.
- USART (Universal Synchronous/Asynchronous Receiver/Transmitter) supports synchronous operation
- Transmitter receives data in parallel from data bus adds the start bit, parity bit, and the stop bit(s) to the data frame.
- The entire frame is sent serially from the transmitting UART to the receiving UART.
- The receiving UART samples the data line at the pre- configured baud rate.
- The receiving UART discards the start bit, parity bit, and stop bit from the data frame.
- The receiving UART converts the serial data back into parallel and transfers it to the data bus.
- Advantages
  - Only uses two wires
  - No clock signal is necessary
  - Has a parity bit to allow for error checking
  - Well documented and widely used
- Disadvantages

  - The size of the data frame is limited to a maximum of 9 bits
  - Doesnt support multiple slave or multiple master systems
  - The baud rates of each UART must be within 10% of each other (need accurate clocks).

Serial Peripheral Interface (SPI)

- Small displays, SD card modules, RFID card reader modules, and 2.4 GHz wireless transmitter/receivers all use SPI to communicate with microcontrollers.
- Data can be transferred without interruption. Any number of bits can be sent or received in a continuous stream.
- One master can control more than one slaves.
- MOSI (Master Output/Slave Input)
- MISO (Master Input/Slave Output)
- SCLK (Clock) Line for the clock signal.
- SS/CS (Slave Select/Chip Select)
- Advantages
  - Continuous streaming
  - Full duplex
- Disadvantages
  - 4 wires
  - Single master
  - No error checking
  - No ack

Inter-Integrated Circuit (I2C)

- Data is transferred in messages, which are are broken up into frames of data.
- Each message has an address frame that contains the binary address of the slave.
- Simple bidirectional 2-wire bus
  - SDA (Serial Data)
  - SCL (Serial Clock)

10 Media Access Control (MAC)

Two types of links:

- Point-to-point (DSL, Point-to-point ethernet)
- Broadcast (shared wire or medium like LAN, satellite)

Multiple Access Protocols

- Two or more simultaneous transmissions by nodes: interference leads to collision if node receives two or more signals at the same time
- Distributed algorithm that determines how nodes share channel, i.e., determine when node can transmit
- Communication about channel sharing must use channel itself! No out-of-band signalling channel

Ideal Protocol

- When one node wants to transmit, it can send at rate R, (efficiency).
- When M nodes want to transmit, each can send at average rate R/M (fairness)
- Fully decentralized (no special node, no synchronization of clocks/slots)
- Simple

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

10.1 Channel Partitioning

- Divide channel into smaller pieces (time slots, frequency, code)
- Allocate piece to node for exclusive use

TDMA: time division multiple access

- Access to channel in "rounds"
- Each station gets fixed length slot (length = pkt trans time) in each round
- Unused slots go idle

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

10.2 Random Access Protocols

- When a node has packet to send transmit at full channel data rate R with 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)

- Efficiency: long-run fraction of successful slots (many nodes, all with many frames to send)

slotted ALOHA ALOHA CSMA, CSMA/CD, CSMA/CA

Slotted ALOHA

- Assumptions:
  - All frames same size
  - Time divided into equal size slots 1 frame)
  - Nodes start to transmit only at beginning of slot
  - Nodes are synchronized
  - If 2 ore more nodes transmit in slot, all nodes detect collision
- Operation:
  - When node obtains fresh frame, tries to transmit it in next slot
  - If no collision: it is sent successfully
  - If collision: retransmits frame in each subsequent slot with probability p until success
- Pros
  - Single active node can continuously transmit at full rate of channel
  - Highly decentralized: only slots in nodes need to be in sync
  - Simple
- Cons
  - Collisions waste slots
  - Idle slots
  - Nodes may be able to detect collision in less than time to transmit packet
  - Clock synchronization
- Efficiency:
  - Example p168
  - Probability that any node success $P_{\text{success}} = Np(1-p)^{N-1}$
  - At best, channel useful for transmissions $1/e$ of the time, or 37%

Pure (unslotted) ALOHA

- When frame first arrives transmit immediately
- Collision probability increases by factor of two

- Frame sent at t collides with other frames sent in $[t-1, t+1]$
- Does not require time synchronisation
- Derivation of collision probability

10.3 CSMA (Carrier Sense Multiple Access)

- Listen before transmit:
- If channel sensed idle: transmit entire frame
- If channel sensed busy, defer transmission
- Human analogy: dont interrupt others!

CSMA Variants

- 1-persistent: if the channel is idle, then send the data.
- Nonpersistent: if the channel is already in use, do not continually sense it; wait a random period of time and then repeat the algorithm.
- P-persistent: (slotted system) if the channel is idle, transmit with a probability p; wait for next slot with probability 1-p.

Collisions in CMSA

- Propagation delay means two nodes may not hear each others transmission
- Entire packet transmission time wasted on collision
- Distance & propagation delay in determine collision probability
- The problem of a station not being able to detect a potential competitor for the medium because the competitor is too far away is called the hidden terminal problem.
- Example of above

CSMA/CD (Collision Detection)

- Carrier sensing, deferral as in CSMA
- Collisions detected within short time
- Colliding transmissions aborted, reducing channel wastage
- Collision detection:
  - Easy in wired LANs: measure signal strengths, compare transmitted, received signals
  - Difficult in wireless LANs: received signal strength overwhelmed by local transmission strength
  - Human analogy: the polite conversationalist
  - The minimum time to detect the collision is just the time it takes the signal to propagate from one station to the other.

- CSMA/CD can be in contention, transmission, or idle state.

CSMA/CA (CSMA with Collision Avoidance),

- Conceptually similar to (Ethernets) CSMA/CD
- With channel sensing before sending
- Exponential back-off after collisions.
- However, a station that has a frame to send starts with a random backoff (except in the case that it has not used the channel recently and the channel is idle).
- It does not wait for a collision unlike in CSMA/CD. human analogy: the very polite conversationalist

RTS/CTS

- To avoid hidden terminal problem, can use short Request to Send (RTS) and Clear to Send (CTS) control frames to reserve access to the channel.

11 Birth-Death Processes

- The birth-death process is a special case of a continuous- time Markov chain where the states E_k (or just k) represent k the current size of a population.
- Only transitions from one state to a neighbouring state are permitted!
- A birth may correspond to the arrival of a customer and a death may correspond to the case where a customer leaves
- The system.
- Birth rate λ is the rate at which a birth occurs when the population is of size k.
- Death rate μ is the rate at which a death occurs when the population is of size k.
- Assume that these birth and death rates are independent of time and depend only on current population E_k

Dynamics

- Flow rate in of E_k is $\lambda_{k-1}P_{k-1}(t) + \mu_{k+1}P_{k+1}(t), k \geq 1$
- Flow rate out of E_k is $(\lambda_k + \mu_k)P_k(t), k \geq 1$
- Difference between the two is effective probability flow rate into the state $\frac{dP_k(t)}{dt} = \lambda_{k-1}P_{k-1}(t) + \mu_{k+1}P_{k+1}(t) - (\lambda_k + \mu_k)P_k(t)$

11.1 Equilibrium

Global Balance Equations

- If $p_k = \lim_{t \to \infty} P_k(t)$ exists
- $0 = \frac{dP_k(t)}{dt} = \lambda_{k-1} P_{k-1}(t) + \mu_{k+1} P_{k+1}(t) - (\lambda_k + \mu_k) P_k(t)$
- $0 = -\lambda_0 p_0 + \mu_1 p_1$

Detailed Balance Equations

- Equate flows between states E_k and E_{k+1}
- $p_k \lambda_k = p_{k+1} \mu_{k+1}$
- $p_{k+1} = \frac{\lambda_k}{\mu_{k+1}} p_k$
- Simpler but not as general as global balance equations

Equilibrium Solution

Can be proved via induction

- Observation $\sum_{k=0}^{\infty} p_k = 1$
- $p_k = p_0 \Pi_{i=0}^{k-1} \frac{\lambda_i}{\mu_i + 1}$
- $p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \Pi_{i=0}^{k-1} \frac{\lambda_i}{\mu_i + 1}}$

12 Queueing Theory

Notation

- C_n nth customer to enter the system
- τ_n arrival time for C_n
- t_n inter-arrival time between C_{n-1} and C_n $(\tau_n - \tau_{n-1})$
- x_n service time for customer C_n
- w_n amount of time spent in queue (before service)
- $s_n = w_n + x_n$ total time in system
- $\alpha(t)$ number of customers who arrive in time (0, t)
- $\delta(t)$ number of customers who leave the system in time (0, t)
- $N(t)$ number of customers who in the system at t
- $\gamma(t)$ total customer seconds, total area between $\alpha(t)$ and $\delta(t)$ curves
- $\lambda(t) = \frac{\alpha(t)}{t}$ average arrival rate during interval (0, t)

- $\mu(t)$ average service rate during the interval (0, t)
- $T_t = \frac{\gamma(t)}{\alpha(t)}$ average time customer spent in system during interval (0, t)
- $\bar{N}_t = \frac{\gamma(t)}{t} = \lambda_t T_t$ the average number of customers in queueing system during interval (0, t)

12.1 Basic characteristics of queueing systems

Arrival Patterns

- Assume independence in many cases (i.i.d)
- $A(t) = P[\text{time between arrivals} \leq t]$

Service Patterns

- Assume independence in many cases (i.i.d)
- $B(x) = P[\text{service time} \leq x]$

Queueing discipline

- First come, first served (FCFS), most common; also known as FIFO
- Last come, first served (LCFS), e.g. stacks; also known as LIFO
- Random selection for service (RSS)
- A variety of priority schemes, e.g. VIPs

System Capacity

- In some queueing processes, there is a physical limitation to the amount of waiting room, so that when the line reaches a certain length, no further customers are allowed to enter until space becomes available as a result of a service completion.
- E.g. Router buffer capacity

Number of Service Channels

- Number of service channels = number of parallel service stations which can serve customers simultaneously.

Number of Service Stages

Queueing Notation

- Queueing processes can be described using the shorthand A/B/X/Y/Z
- A: indicates the interarrival time distribution

- B: indicates the service time distribution
- X: indicates the number of parallel service channels
- Y: indicates the restriction on system capacity
- Z: indicates the queueing discipline
- E.g. M/D/2/ /FCFS indicates a queueing process with exponential interarrival times (M stands for Markovian or memoryless), deterministic service times, 2 parallel service stations, no restrictions on maximum number allowed in the system and first come, first served.

Measuring System Performance

- Some measure of the waiting time that a typical customer might be forced to endure
- An indication of the manner in which customers may accumulate, i.e. queue size.
- A measure of the idle time of the server

Customer waiting times

- The time a customer spends in the queue (before service)
- The total time a customer spends in the system

Customer accumulation measures

- The number of customers in the queue (not being served)
- The total number of customers in the system

Idle service measures

- Can include the percentage of time any particular server may be idle, or the time the entire system is devoid of customers

12.2 Littles Law

- Assume arrival rate λ_t converges to λ
- Assume arrival rate T_t converges to T
- Then $\bar{N}_t$ will converge to $\bar{N} = \lambda T$
- It holds for almost every queueing system that reaches a steady state!
- Applies to subsets of systems: the average number of customers in the queue is equal to the average arrival rate of customers times the average time a customer spends waiting in the queue. $\bar{N}_q = \lambda W$

Average time in system

$T = \bar{x} + W$

Average number of customers in service

$$\begin{aligned}\bar{N}_s &= \bar{N} - \bar{N}_q \\ &= \lambda T - \lambda W \\ &= \lambda(T - W) \\ &= \lambda \bar{x}\end{aligned}$$

Utilization factor

- Ratio of rate at which work enters the system to maximum rate (capacity) at which system can perform work
- $\mu = \frac{1}{\bar{x}}$ average service rate
- For single server $\rho = \frac{\lambda}{\mu} = \lambda\bar{x}$
- For multiple (m) servers $\rho = \frac{\lambda}{m\mu} = \frac{\lambda\bar{x}}{m}$
- $\rho < 1$ is a stable system
- Can be interpreted as expected value of fraction of busy servers
- $P[\text{idle}] = P[\text{no customers}] = p_0 = 1 - \rho$ in a single server system

12.3 PASTA Property

Poisson arrivals see time averages

- There are several situations where we need a probabilistic characterization of a queueing system as seen by an arriving customer.
- The steady-state occupancy probabilities upon arrival need not to be equal to the corresponding unconditional steady-state probabilities
- Suppose the interarrival times are independent and uniformly distributed between 2 and 4 sec, while customer service times are equal to 1 sec (deterministic).
- An arriving customer always finds an empty system!
- On the other hand, the average number in the system as seen by an outside observer looking at the system at a random time is 1/3 by Littles Law.
- Property holds under very general conditions for queueing systems with Poisson arrivals regardless of the distribution of the service time.

12.4 M/M/1 Queue

Equilibrium Probability

$p_0 = 1 - \frac{\lambda}{\mu}$
$p_k = (1 - \rho)\rho^k$

Average number of customers in system

$\bar{N} = \frac{\rho}{1-\rho}$

Average time spent by customer in system

$T = \frac{\bar{N}}{\lambda} = \frac{1}{\mu-\lambda}$

Variance of number of customers in system

$Var[N] = \frac{\rho}{(1-\rho)^2}$

Probability of finding at least k customers in system

$P[\geq k \text{ in system}] = \rho^k$

Average number of customers in the queue

$\bar{N}_q = \frac{\rho^2}{1-\rho}$

Average time spent in the queue

$W = \frac{\rho}{\mu-\lambda}$

12.5 M/M/m Queues

Maximum of m servers

Equilibrium Probability

$p_k = \begin{cases} p_0 \frac{(m\rho)^k}{k!} & k \leq m \\ p_0 \frac{(m\rho)^k}{m! m^{k-m}} & k > m \end{cases}$
$\rho = \frac{\lambda}{m\mu} < 1$
$p_0 = \big[\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \big(\frac{(mp)^m}{m!}\big)\big(\frac{1}{1-\rho}\big)\big]^{-1}$

Erlangs C Formula

- Steady state probability that an incoming call waits in the queue (all servers full)
- $P[\text{queueing}] = \frac{\left(\frac{(mp)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)}{\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \left(\frac{(mp)^m}{m!}\right)\left(\frac{1}{1-\rho}\right)}$

Average number of customers in queue

$\bar{N}_q = \sum_{k=m}^{\infty}(k-m)p_k = \left(\frac{(mp)^m \rho}{m!(1-\rho)^2}\right)p_0$

Average waiting time in queue

$W = \frac{\bar{N}_q}{\lambda}$

Average time in system

$T = \frac{1}{\mu} + W$

Average number of customers in system

$\bar{N} = \lambda T$

12.6 M/M/∞ Queues

Infinite number of servers

Equilibrium Probabilities

$p_0 = e^{-\lambda/\mu}$
$p_k = \frac{(\lambda/\mu)^k}{k!}e^{-\lambda/\mu}$

Average number of customers in system

$\bar{N} = \frac{\lambda}{\mu}$

Average time spent by customer in system

$T = \frac{\bar{N}}{\lambda} = \frac{1}{\mu}$

12.7 M/M/1/k Queue

- Poisson input turned off when queue is full
- Do not need $\frac{\lambda}{\mu} < 1$

Equilibrium Probabilities

$\lambda/\mu \neq 1$

$p_0 = \frac{1-\lambda/\mu}{1-(\lambda/\mu)^{K+1}} = \frac{1-\rho}{1-\rho^{K+1}}$

$p_k = p_0 = \begin{cases} \frac{1-\lambda/\mu}{1-(\lambda/\mu)^{K+1}}\left(\frac{\lambda}{\mu}\right)^k = \frac{1-\rho}{1-\rho^{K+1}}\rho^k & 0 \leq k \leq K \\ 0 & \text{otherwise} \end{cases}$

$\lambda/\mu = 1$

$p_k = \begin{cases} p_0 = 1/(K+1) & 0 \leq k \leq K \\ 0 & \text{otherwise} \end{cases}$

Blocking Probability

$P[\text{block}] = \frac{1-\rho}{1-\rho^{K+1}}\rho^K$

Average number of customers in queue $\lambda/\mu \neq 1$

$\bar{N}_q = \sum_{k=1}^{K}(k-1)p_k = \frac{\lambda/\mu}{1-\lambda/\mu} - \frac{(\lambda/\mu)[K(\lambda/\mu)^K+1]}{1-(\lambda/\mu)^{K+1}} = \frac{\rho}{1-\rho} - \frac{\rho[K\rho^K+1]}{1-\rho^{K+1}}$

Average number of customers in system

$\bar{N} = \sum_{k=0}^{K} kp_k = \sum_{k=1}^{K} kp_k = \bar{N}_q + (1-p_0)$

Average time in system (when entered)

$T = \frac{\bar{N}}{(1-p_K)\lambda} = \frac{\bar{N}}{\lambda_{eff}}$

Effective arrival rate

$\lambda_{eff} = (1 - P[\text{blocking}])\lambda = (1-p_K)\lambda$

Little's Law for systems with blocking

$\bar{N} = \lambda_{eff}T$

12.8 M/M/1/1 Queue

Special case of M/M/1/k

Equilibrium probabilities

$$p_k = \begin{cases} \frac{1}{1+(\lambda/\mu)} & k = 0 \\ \frac{\lambda/\mu}{1+(\lambda/\mu)} & k = 1 = K \\ 0 & \text{otherwise} \end{cases}$$

12.9 M/M/m/m Queue

- m servers
- Each newly arriving customer gets a private server
- If all servers are occupied, customer is lost

Equilibrium Probabilities

$$p_0 = \left[\sum_{k=0}^{m} \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}\right]^{-1}$$

$$p_k = \begin{cases} p_0 \Pi_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} = p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} & k \leq m \\ 0 & k \geq m \end{cases}$$

Erlang's B Formula (blocking probability)

$$p_m = \frac{(\lambda/\mu)^m/m!}{\sum_{k=0}^{m} (\lambda/\mu)^k/k!}$$

12.10 M/G/1 Queues

- Arrival rate is a Poisson process λ
- Service times have a general distribution (not necessarily exponential)
- Non-Markovian

P-K Formula

Average waiting time is $W = \frac{\lambda \bar{X^2}}{2(1-\rho)}$

- $\bar{X} = E[X] = \frac{1}{\mu}$: Average service time
- $\bar{X^2} = E[X^2]$: Second moment of service time
- $\rho = \lambda/\mu = \lambda \bar{X}$
- Can have $\rho < 1$ but infinite W if the second moment of X is infinite
- Valid for any order of servicing customers as long as order is determined independently of service time

Average time in system

$T = \bar{X} + \frac{\lambda \bar{X^2}}{2(1-\rho)}$

Average number of customers in queue

$\bar{N}_q = \lambda W = \frac{\lambda^2 \bar{X^2}}{2(1-\rho)}$

Average number of customers in system

$\bar{N} = \lambda T = \rho + \frac{\lambda^2 \bar{X^2}}{2(1-\rho)}$

Vacations

Waiting time is now $W = \frac{\lambda \bar{X^2}}{2(1-\rho)} + \frac{\bar{V^2}}{2\bar{V}}$

- At the end of each busy period the server goes on vacation
- New arrivals msut wait until end of vacation period
- If the system still idle at completion of vacation, new vacation starts immediately
- Let V_i be the durations of successive vacations taken by the server,
  - Assume i.i.d and independent of customer interarrival and service time,
  - Mean $\bar{V}$
  - Second moment $\bar{V^2}$
- Example: slotted FDM

12.11 G/G/1 Queueing

Average waiting time in queue

$W \leq \frac{\lambda(\sigma_a^2+\sigma_b^2)}{2(1-\rho)}$

- σ_a^2: Variance of interarrival times
- σ_b^2: Variance of service times
- λ: Average arrival rate
- ρ: Utilization factor λ/μ
- Upper bound becomes tight as $\rho \to 1$

12.12 Nonpreemptive Priority Queueing

- A customer undergoing service is allowed to complete service without interruption even if a customer of higher priority arrives in the meantime
- A separate queue is maintained for each priority class
- When the server becomes free, the first customer of the highest nonempty priority queue enters service
- Shown that average delay per customer is reduced when shorter service times are given higher priority

Mean residual time

$R = \frac{1}{2} \sum_{i=1}^{n} \lambda_i \bar{X_i^2}$

Average time spent in queue for class k

$W_k = \frac{R}{(1-\rho_1-\cdots-\rho_{k-1})(1-\rho_1-\cdots-\rho_k)}$

Average time spent in system for class k

$T_k = \frac{1}{\mu_k} + W_k$

Average time spent in queue over all classes

$W = \frac{\lambda_1 W_1 + \cdots + \lambda_n W_n}{\lambda_1 + \cdots + \lambda_n}$

Average time spent in system over all classes

$T = \frac{\lambda_1 T_1 + \cdots + \lambda_n T_n}{\lambda_1 + \cdots + \lambda_n}$

Minimizing total cost

$\frac{\bar{X_1}}{c_1} \leq \frac{\bar{X_2}}{c_2} \leq \cdots \leq \frac{\bar{X_n}}{c_n}$

12.13 Preemptive Resume Priority Queueing

- Preemptive resume priority - A customer is interrupted when a higher priority customer arrives and is resumed from the point of the interruption once all customers of higher priority have been served
- Preemptive repeat priority - interrupted customer needs to restart service when it gets served again

Average time spent in system for class k

$T_k = \frac{\frac{1}{\mu_k}(1-\rho_1-\cdots-\rho_k)+R_k}{(1-\rho_1-\cdots-\rho_{k-1})(1-\rho_1-\cdots-\rho_k)}$

13 Networks of Queues

- Interarrival times at later queues can become correlated with service times of earlier queues.
- Assume Poisson packet arrivals, all packets same length, transmission time proportional to packet length
- Example

Kleinrock Independence Approximation

- Idea: merging packet streams has an effect similar to restoring independence of interarrival times and packet lengths.
- Adopt an M/M/1 model for each communication link on the network (this is an unrealistic assumption!)
- It was thought to be a reasonably good approximation for systems with Poisson stream arrivals at entry points, packet lengths nearly exponentially distributed, a densely connected network, and moderate-to-heavy traffic.

Burkes Theorem

- For M/M/1, M/M/ and M/M/m systems with arrival rate λ, assume that $P_k(0)$ is chosen such that $P_k(0) = p_k$ (i.e systems in steady state at all times)
- Then, the queuing system (Markov chain) is reversible if and only if its stationary distribution and transition rates satisfy detailed balance equations. This is true for birth-death processes.
- In the reversed queueing process arrivals and departures correspond to those of the originals in reverse.
- With the assumptions above:
  - The departure process is Poisson with rate λ
  - At each time t, the number of customers in the system is independent of the sequence of departure times prior to t. This is counter-intuitive!
- Example

Jacksons Theorem

- Assume that each customer will eventually exit the system with probability 1.
- Assume that service times of customers at queue j are exponentially distributed with mean $1/\mu$, are mutually independent and independent of the arrival process.
- Assume $\rho_j = \lambda_j/\mu_j < 1 \forall j$
- Jacksons Theorem: $\forall n_1, \ldots, n_k \geq 0, P(n_1, \ldots, n_K) = P_1(n_1)P_2(n_2)\ldots P_k(n_K)$ where $P(n_1, \ldots, n_K)$ is the probability of having n of having n_1 in queue 1, n_2 in queue 2, $\ldots n_k$ in queue K, and $P_j(n_j) = \rho_j^{n_j}(1 - \rho_j)$

- Jacksons theorem says that the number of customers in the networks queues acts as if each queue is an independent M/M/1 queue
- However they are not actually independent
- Can show that the total arrival process at each queue is in general not Poisson.

14 Internet Routing

Hierarchical Routing

- Aggregate routers into regions, autonomous systems (AS)
- Routers in same AS run same routing protocol (intra-AS)
- Routers in different ASs can run different intra- AS routing protocols
- Gateway router: direct link to router in another AS at edge of its own AS
- Forwarding table configured by both intra- and inter-AS routing algorithms
- Hierarchical routing saves table size, reduces update traffic
- Intra-AS
  - Sets entries for internal and external destinations
  - Single admin, so no policy decisions needed
  - Can focus on performance
- Inter-AS
  - Set entries for external destinations
  - Admin wants control over how its traffic routed, who routes through its net.
  - Inter-AS: policy may dominate over performance
  - Must prpoagate which destinations are reachable to all routers in its AS

Interior Gateway Protocols (IGP)

Intra-AS routing on the internet

- RIP: Routing Information Protocol
- OSPF: Open Shortest Path First
- EIGRP: Enhanced Interior Gateway (Cisco proprietary)
- IS-IS: Intermediate System to Intermediate

Routing Information Protocol (RIP)

- Uses distance vector algorithm (# of hops)
- RIP advertisements

  - Distance vectors: exchanged among neighbors every 30 sec via Response Message (also called advertisement)
  - Each advertisement: list of up to 25 destination subnets within AS
  - Example P531
- Failure and Recovery
  - If no advertisement heard after 180 sec then neighbor/link declared dead
  - Routes via neighbor invalidated
  - New advertisements sent to neighbors
  - Neighbors in turn send out new advertisements (if tables changed)
  - Link failure info quickly (?) propagates to entire net
  - Poisoned reverse used to prevent ping-pong loops (infinite distance = 16 hops)
- RIP routing tables managed by application-level process called route-d (daemon in Unix)
- Advertisements sent in UDP packets, periodically repeated

OSPF (Open Shortest Path First)

- open: publicly available
- Uses Link State Algorithm
  - LS packet dissemination
  - Topology map at each node
  - Route computation using Dijkstras algorithm
- OSPF advertisement carries one entry per neighbor
- Advertisements flooded to entire AS directly over IP (rather than TCP or UDP)
- Advanced Features
  - All OSPF messages authenticated (to prevent malicious intrusion)
  - Multiple same-cost paths allowed (only one path in RIP)
  - For each link, multiple cost metrics for different services (e.g., satellite link cost set low for best effort traffic; high for real time traffic)
  - Integrated uni - and multicast support:
- Hierarchical OSPF
- Two-level hierarchy: local area, backbone.
  - Link-state advertisements only in area
  - Each nodes has detailed area topology; only know direction (shortest path) to nets in other areas.
- Area border routers: summarize distances to nets in own area, advertise to other Area Border routers.
- Backbone routers: run OSPF routing limited to backbone.

- Boundary routers: connect to other ASs, serve gateway function.

Carrier Ethernet

Wide Area Network (WAN) providers want to provide their customers with Ethernet services and make use of the volume and cost advantages of Ethernet technologies in their networks.

MPLS

- Packet-forwarding technology which uses labels in order to make data forwarding decisions (for any network layer protocol).
- Layer 3 header analysis is done just once (when the packet enters the MPLS domain). Label inspection drives subsequent packet forwarding.
- Helps with Virtual Private Networking (VPN), Traffic Engineering (TE), Quality of Service (QoS), Any Transport over MPLS (AToM)
- Decreases the forwarding overhead on the core routers.
- Label is a four-byte locally-significant identifier used in order to identify a Forwarding Equivalence Class (FEC), which is a group of IP packets which are forwarded in the same manner, over the same path, and with the same forwarding treatment.

14.1 BGP-Border Gateway Protocol

- De facto standard; extremely complex
- Allows subnet to advertise its existence to rest of Internet: I am here
- BGP provides each AS a means to:
  1. eBGP (external): obtain subnet reachability information from neighboring ASs.
  2. iBGP (internal): propagate reachability information to all AS-internal routers.
  3. Determine good routes to subnets based on reachability information and policy.
- BGP sessions: pairs of BGP routers (peers) exchange BGP messages
  - Advertising paths to different destination network prefixes (path vector protocol)
  - Exchanged over semi-permanent TCP connections.
- When AS2 advertises a prefix to AS1:
  - AS2 promises it will forward datagrams towards that prefix.
  - AS2 can aggregate prefixes in its advertisement

BGP messages

- OPEN: opens TCP connection to peer and authenticates sender
- UPDATE: advertises new path (or withdraws old)

- KEEPALIVE: keeps connection alive in absence of UPDATES; also ACKs OPEN request
- NOTIFICATION: reports errors in previous msg; also used to close connection
- FSM Diagram P547

Path attributes & BGP Routes

- When router learns of new prefix, it creates entry for prefix in its forwarding table.
- Advertised prefix includes BGP attributes. prefix + attributes = route
- Two important attributes:
  - AS-PATH: contains ASs through which prefix advertisement has passed: e.g, AS 67, AS 17
  - NEXT-HOP: the interface on the gateway router in the next- hop AS that leads to the gateway router in the current AS (this interface is still on a subset routable by the intra-AS routing).
- When gateway router receives route advertisement, uses import policy to accept/decline
  - E.g.: never route through AS x
  - Policy-based routing

BGP route selection

- Router may learn about more than 1 route to some prefix, it must select one.
- Elimination rules:
  - Local preference value attribute: policy decision
  - Shortest AS-PATH
  - Closest NEXT-HOP router: hot potato routing
  - Additional criteria

15 Congestion Control

- Too many sources sending too much data too fast for network to handle
- Different from flow control (flow control is per-connection)
- Manifestations
  - Lost packets (buffer overflow at routers)
  - Long delays (queueing in router buffers)

Sender

- Limits transmission rate
  - Last byte Sent - Last Byte Acked $\leq$ cwnd

- rate = $\frac{\text{cwnd}}{RTT}$ Bytes/sec
- cwnd is dynamic, function of perceived network congestions

- Perceiving congestion
  - Loss event = timeout or 3 duplicate acks
  - TCP sender reduces rate (cwnd) after loss event
  - Three mechanisms:
    * Slow start
    * AIMD
    * Conservative after timeout events

TCP Slow Start

- When connection begins, increase rate rapidly until first loss event or ssthresh reached:
- Double cwnd every RTT
- Done by incrementing cwnd for every ACK received
- Initial rate is slow but ramps up exponentially fast until hit ssthresh or problem
- When connection begins, cwnd= 1 MSS
- Example:
  - MSS = 500 bytes
  - RTT = 200 msec
  - Initial rate = 20 kbps
  - Available bandwidth may be ¿¿ MSS/RTT

16 DNS

Dynamic DNS

- Standardized method Of dynamically updating domain name server record
- s
- Dynamic DNS providers offer a software client program that automates the discovery and registration of the client system's public IP addresses.
- Included by many home networking/routers in firmware
- Basically a constantly changing public IP

A small number of Web sites receive massive traffic and a vast