

CSE 124

Review

March 10, 2016, UCSD

Prof. George Porter

Outline

- Overview of topics
- Small group activity:
 - Working example problems similar to the final
- Open discussion for project related questions

1. Topics

Network programming fundamentals

- Network sockets API: open(), connect(), send(), recv(), etc?
- How names are resolved to IP addresses in DNS
- To query a dns server, you need to know at least one IP address without using DNS. explain.
- Internet Protocol (IP) What is the IP service model?

Network programming fundamentals

- End to end protocols
 - Move from host-to-host to process-to-process communication model
 - UDP – thin layer on top of IP
 - TCP – provide abstraction of reliable in-order byte stream on top of IP protocol
- Signals and timeouts
- Concurrency, multi-tasking, multiplexing
- Locking, mutexes, sharing state between threads/processes

Protocol design and analysis

- Framing vs. parsing
- Delimiter vs. length-value
- Graphing data, axes labels, error bars

Data centers and CDNs

- Round-robin DNS vs load balancers. advantages and disadvantages of each
- Replication vs partitioning: advantages and disadvantages
- Terms: MTTR, MTBF, availability, yield, harvest, DQ principle
- Terms: top of rack switch, PUE, SPUE
- Tail latency vs. average latency
- Energy and power

Application-layer protocols

- E-Mail
 - Purpose of the MTA, UA, MIME, SMTP, POP, IMAP
 - What are the two parts of an email?
- HTTP
 - Terms for http: proxy server, request header, response header, status code
 - Protocol structure/syntax

RPCs

- Explain concept of 'idempotent'
- Maybe vs at least once vs at most once semantics; how to implement each of these?
- Thrift
- Role of stub compiler, RPC runtimes
- Discuss whether the following operations are idempotent:
 - Pressing a lift (elevator) request button
 - Writing data to an offset in a file
 - Appending data to the end of a file (assuming there are no other writers in the system)

Overlay networks, P2P, Chord

- Compare and contrast aspects of flooding queries, supernodes, vs structured. Tradeoffs--which is better for joining, leaving, advertising content, querying for content.
- Assume you have a Chord system of 5 nodes with identifiers 0, 1, 2, 6, and 7.
 - Draw the identifier circle and show which nodes the following keys will be assigned to 0, 4, 5, 8. You can assume the identifier space is 0-9.
- Each chord node must maintain routing state.
 - Describe exactly what routing state must be maintained at each node to ensure correct function. Show what this state would be for node 6. What is the expected lookup time of an object?
 - Describe what routing state must be maintained at each node to ensure fast lookup times. Show what this state would be for node 6. what is the expected lookup time of an object?

IPSec, VPNs, Tunnels

- How does IPSec ensure authentication?
Encryption of traffic?
- How can you use tunnels to connect two networks together?
- What is IP-in-IP encapsulation?

2. Example problems and solutions

(Get into groups of $\approx 2-3$ students)

Network programming

What is the main function of the **network** layer?

- a. Node-to-node delivery
- b. Process-to-process delivery
- c. Synchronization
- d. Updating and maintenance of routing tables

What is the main function of the **transport** layer?

- a. Node-to-node delivery
- b. Process-to-process delivery
- c. Synchronization
- d. Updating and maintenance of routing tables

Network programming

Which of the following is an application layer service?

- a.** Remote log-in
- b.** File transfer and access
- c.** Mail service
- d.** All of the above

Network programming

For each of the following operations on a remote file server, write whether they are more likely to be delay sensitive or bandwidth sensitive:

- a. Open a file _____
- b. Read the entire contents of a file _____
- c. List the contents of a directory _____
- d. Retrieve the size of a file _____

Data centers

What are two reasons why we need a DNS system, when we can directly use an IP address?

Data centers

- In a horizontally partitioned system, node failure:
 - Always reduces harvest
 - Never reduces harvest
 - Sometimes reduces harvest
- In a horizontally partitioned system, node failure:
 - Always reduces yield
 - Never reduces yield
 - Sometimes reduces yield

Data centers

- In a vertically partitioned system, node failure:
 - Always reduces harvest
 - Never reduces harvest
 - Sometimes reduces harvest
- In a horizontally partitioned system, node failure:
 - Always reduces yield
 - Never reduces yield
 - Sometimes reduces yield

Application layer protocols

Which of the following protocols is used for sending an email?

- a. POP
- b. UA
- c. SMTP
- d. MTA

Which of the following protocols is used for retrieving an email?

- a. POP
- b. UA
- c. SMTP
- d. MTA

What is the purpose of an E-Mail MTA?

Application-layer protocols

Which of the following is present in both HTTP requests and HTTP responses?

- a. HTTP version number
- b. URL
- c. Status code
- d. Status phrase

Describe two ways in which HTTP 1.1 request pipelining improves network performance when running over TCP.

RPCs

Put the following remote procedure call (RPC) events into the correct order by writing a number in each blank. The first event should get a 1, the second event should get a 2, etc.

- _____ Client code calls the client stub. This call is a local procedure, with parameters pushed on to the stack in the normal way.
- _____ The local OS on the server passes the incoming packets to the server stub.
- _____ The client stub marshalls the parameters into a message and makes a system call to the OS to send the message.
- _____ The server stub unmarshalls the parameters from the message.
- _____ The client's local OS sends the message from the client machine to the server.
- _____ The server stub calls the server procedure.

Overlays and P2p

A disadvantage of centralized p2p networks is:

- a.** They have a single point of failure (the centralized index)
- b.** Each client in the system must contact every other node in the system any time it wants to retrieve content
- c.** Both A and B
- d.** Neither A nor B

In the fully distributed p2p network Gnutella, there is no central index, and instead, queries are flooded amongst the nodes. What is a common mechanism to limit the propagation of queries in a flooding network?

- a.** Content Distribution Networks (CDNs)
- b.** Randomly evicting clients from the p2p network with probability ρ
- c.** A Time-to-live (TTL) field in the query
- d.** Photon torpedoes

Overlays and p2p

You live in California, and often play a video game with a friend of yours who lives in New York, which is about 3,000 miles away. The performance of this game is highly dependent on the round-trip time (RTT) latency, sometimes called the “ping time”. When you measure the RTT to your friend’s computer, it is very slow (250 milliseconds). There is a company that says that they can cut your ping time in half without requiring that you or your friend to change anything about your Internet connections or buy any new hardware. Is this possible? Explain why or why not.

Protocol design 1

- http://cseweb.ucsd.edu/~gmporter/classes/wi16/cse124/exams/example_protocol_problem.pdf

Protocol design 2

- Tele-surgery is when a remote surgeon controls a robot to perform a procedure. The movements of the robot must be accurate, and must occur in the proper order
- Assume a surgical procedure is an ordered list of operations, and each operation is 100 bytes long (binary data)
- Design a protocol that (1) executes the correct steps in the correct order, or (2) executes a prefix of the steps then stops before doing anything wrong
 - Assume the channel is lossy and has a high error rate
- (Just sketch the solution with your group)

3. Open discussion for project questions