Distributed Systems

Fall 2014 Exam 3 Review

Paul Krzyzanowski
Rutgers University
Fall 2014

Fall 2014 - Question 1

Most databases require the use of read locks as well as write locks. Spanner requires locks for writes but offers lock-free reads. Explain how Spanner supports this. After all, you don't want incoming writes to modify some of the data that is being read since that would violate consistency.

Spanner reads snapshot data ≤ a point in time

A new transaction creates a new version of data (like fields in BigTable)

- NOT: uses TrueTime to wait out uncertainty
  - That is done during a transaction (which uses locking) to ensure that the commit timestamp is definitely in the past.

Fall 2014 - Question 2

Over time, you add some newer, more powerful computers to your data center.

Explain how Amazon Dynamo's virtual nodes allow you to handle this hybrid environment of old and new computers.

One computer has one or more virtual nodes attached to it.

Each virtual node represents a point on the logical ring of hash values.

The number of virtual nodes in a system is based on capacity, accounting for heterogeneity in processing capacity.

Fall 2014 - Question 3

What is a potential problem of a *fail-recover* system versus that of a *fail-stop* system?

- A system may recover with old state about the world
- Example: it may try to complete a transaction that was already committed … or multicast messages it had queued for delivery.
- We saw this handled in virtual synchrony:
  - A dead system was kicked out of a group
  - Upon recovery, it has to re-join the group and be the recipient of a state transfer
- NOT: it will fail again
  - *Don't predict the future!*
- NOT: fail-recover systems will exhibit Byzantine faults
  If the system recovered, we assume that it is no longer faulty

Fall 2014 - Question 4

Alice has Bob's X.509 digital certificate. She validated it to ensure that it is legitimate.
How does she now use it to establish a secure communication channel so she and Bob can exchange encrypted messages?

We're *not* asking Alice to validate Bob – just to communicate securely.

By possessing Bob's certificate, Alice has his public key.

1. Alice creates a random session key S.
2. Alice encrypts S with Bob's public key in his certificate.
3. Alice sends the encrypted key to Bob.
4. Bob decrypts the session key using his private key.
5. Alice & Bob now have a shared key and can communicate.

Fall 2014 - Question 4 – Discussion

Alice has Bob's X.509 digital certificate. She validated it to ensure that it is legitimate.
How does she now use it to establish a secure communication channel so she and Bob can exchange encrypted messages?

This is not the question, but…
If Alice first wanted to validate that she's talking with Bob:

1. Alice generates a random string (nonce) and sends it to Bob.
2. Bob encrypts it with his private key and sends the result to Alice.
3. Alice decrypts the received message using Bob's public key (in his certificate). If the result matches the nonce, she is convinced.

Fall 2014 - Question 5

We discussed a two-dimensional content-addressable network (CAN). Consider a one-dimensional version.
Explain how a query is forwarded at a node.

Each node keeps track of its key space: x_{min} and x_{max}.

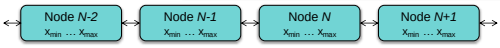
A node also keeps track of its left neighbor and right neighbor

When a node receive a query(key):

h = hash(key)
if (h < x_{min}) forward *query(key)* to the left neighbor
else if (h > x_{max}) forward *query(key)* to the right neighbor
else /* look up locally */

Explain:
1. hash(key)
2. conditions
3. # neighbors

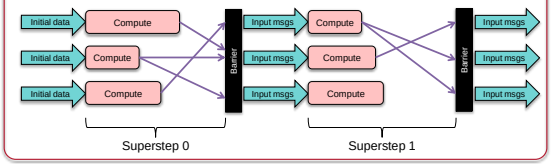
NOT: *if a node does not have a key then the query is forwarded*
NOT: *the node will forward to both the left and right nodes*



Fall 2014 - Question 6

A *superstep* in the Bulk Synchronous Parallel framework is the execution of:

a) User-defined functions in parallel until a point is reached where any one function needs to send a message.
b) User-defined functions in parallel, each starting with reading a set of input messages and progressing until it generates output messages or is done.
c) Several such sequences with a checkpoint at the end.
d) Several such sequences until there is no more input data..



Fall 2014 - Question 7

A Pregel algorithm completes when:

a) A predefined number of supersteps has completed.
b) No vertex outputs any data.
c) No vertex receives any data.
d) Every vertex votes to halt.

- As defined in the paper, algorithm termination is based on every vertex voting to halt.
- Once a vertex votes to halt, the framework will not execute that vertex unless it receives a message.

Fall 2014 - Question 8

Pregel implements this type of group message ordering:

a) Global time ordering
b) Total ordering
c) Partial (causal) ordering
d) Barriers

- Each superstep is a barrier. All messages from superstep N-1 must have been sent and are delivered to each vertex at the start of superstep N.

Fall 2014 - Question 9

How does a *commit wait* differ from a regular *commit*?

a) The transaction waits until all locks are released.
b) The transaction waits until all sub-transactions have completed.
c) The transaction waits until the transaction state has been written to the writeahead log.
d) The transaction waits until the timestamp of the commit is guaranteed to be in the past.

- The *commit wait* is used in Spanner, Google's globally-distributed database.
- A transaction commit is not complete until the process is certain that the timestamp attached to the transaction is definitely in the past.

Fall 2014 - Question 10

Spanner's TrueTime provides clients with:

a) The exact absolute time regardless of where they are.
b) The exact time adjusted for each client's local time zone.
c) The absolute time along with uncertainty bounds.
d) A globally unique monotonically increasing timestamp.

- We never know the exact time (since there's always uncertainty in clock synchronization and network latency)
- The TrueTime API provides us with two timestamps
  - earliest: the earliest possible time it currently is
  - latest: the latest possible time it currently is

Fall 2014 - Question 11

To ensure that transactions are serialized, Spanner uses:
a) The TrueTime API and commit wait.
b) Strict two-phase locking.
c) Paxos consensus.
d) Two-phase commit protocol.

(a) Commit wait guarantees external consistency

(b) Strict two-phase locking: guarantees serializability

(c) Paxos consensus: guarantees consistent replication

(d) Two-phase commit: guarantees distributed agreement on atomicity but, on its own, does not do locking

Fall 2014 - Question 12

As a rough definition, *consistent hashing* means:
a) As long as the size of the hash table stays the same, a key will repeatedly hash to the same value.
b) Most keys will hash to the same value as the size of the hash table changes.
c) Each key is guaranteed to hash to a unique value, ensuring there are no collisions.
d) If key K_1 is greater than another key K_2 then $hash(K_1) > hash(K_2)$

(a) This is true of any hashing

(c) No hash is guaranteed to be collision free for all inputs

(d) Useless

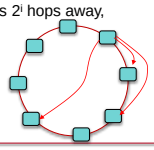
More precisely: consistent hashing with a table of N slots and K keys means that only K/N keys need to be remapped on average when the table size changes. Normal hashing will require just about every key to be remapped

Fall 2014 - Question 13

The purpose of a finger table in the Chord DHT is to:
a) Enable a node to reach its predecessor node as well as its successor node.
b) Reduce the number of hops without storing a list of all machines at every node.
c) Provide a mechanism for fault tolerance, so that <key, value> data can be replicated at N successor nodes.
d) Handle the case where a hash results in a collision and multiple keys must be stored at the same node.

A finger table is a table of *N* succeeding nodes 2^i hops away, where *i=0..N*

entry[0] = 1 node away (neighbor)
entry[1] = 2 nodes away
entry[2] = 4 nodes away
entry[3] = 8 nodes away



Fall 2014 - Question 14

Dynamo uses vector clocks to:
a) Make sure that all client updates to a key are serialized.
b) Provide a multi-version key-value store, allowing a client to retrieve old versions of data.
c) Present the client with multiple values for a key if there were concurrent updates to the key.
d) Provide fault-tolerant replication of a key across multiple nodes.

Vector clocks allow Dynamo to distinguish causally-related from concurrent updates. The application can try to reconcile concurrent changes (for example, merge two concurrent additions to a user's shopping cart).

Fall 2014 - Question 15

A separate network is sometimes used for a cluster heartbeat to:
a) Help distinguish network faults from computer faults.
b) Avoid adding extra traffic to the main network.
c) Ensure that heartbeat messages can go onto the network without collisions.
d) Clearly distinguish signaling traffic from data traffic.

Fall 2014 - Question 16

A *clustered file system* is a:
a) Network file system service that spans multiple computers and provides a file access service.
b) File system that uses a disk that is simultaneously shared across multiple computers.
c) Network file system that resides on a dedicated file server and is used by all members of the cluster.
d) Collection of individual file systems that provides the illusion of one large file system.

- Each operating system has its own file system access code (versus sending network requests to a server)
- The important part is that, since the disk is shared, all systems need to grab locks for any potentially shared blocks (e.g., inodes, free bitmaps, data blocks). This is what the *distributed lock manager (DLM)* does.

Fall 2014 - Question 17

Which system uses *warm failover*?
a) Chubby
b) Bigtable
c) Pregel
d) MapReduce

- Chubby
  - Replicas kept synchronized: hot failover
- Bigtable
  - Replicas kept synchronized: hot failover
- Pregel
  - Workers periodically generate a checkpoint file
- MapReduce
  - A task on a worker restarts from the beginning: cold failover

Fall 2014 - Question 17

Which system uses *warm failover*?
a) Chubby
b) Bigtable
c) Pregel
d) MapReduce

- Chubby
  - Replicas kept synchronized: hot failover
- Bigtable
  - Replicas kept synchronized: hot failover
- Pregel
  - Workers periodically generate a checkpoint file
- MapReduce
  - A task on a worker restarts from the beginning: cold failover

Fall 2014 - Question 18

A overlay caching network such as Akamai uses:
a) Multihoming at the origin server to ensure the original site is connected via multiple ISPs.
b) Proxy servers at the client to cache frequently-accessed content.
c) Load balancers at the origin servers to handle high volumes of requests.
d) Dynamic DNS to map domain names to servers

Fall 2014 - Question 19

A CDN using a caching overlay network does not offer:
a) Caching of static content on edge servers.
b) The ability to replicate the origin server within the network to accelerate dynamic content.
c) Alternate routing choices to the orgin server instead of direct IP routing.
d) Security benefits due isolating the origin and ability to handle high volumes of traffic.

Fall 2014 - Question 20

The *transport system* in Akamai's CDN:
a) Encrypts content for secure transport.
b) Finds the most efficient routes to the origin servers.
c) Queues content requests for delayed access if the network is too congested.
d) Uses multiple network links concurrently to achieve higher bandwidth.

Fall 2014 - Question 21

Suppose you can test all combinations of a 32-bit key in one second.
How long will it take you to test all combinations of a 40-bit key?
a) 40/32 = 1.25 seconds
b) (40-32) = 8 seconds
c) 2^8 = 256 seconds
d) $(2^{40} - 2^{32}) = 1.09 \times 10^{12}$ seconds

- Each bit in a key doubles the search space
- Additional 8 bits (40-32) = 2^8 = 256 times as long

Fall 2014 - Question 22

An advantage of using a public key algorithm for secure communication is:
a) A message can be securely delivered to many recipients at once.
b) It is far more secure than using a symmetric algorithm.
c) It is much faster than using a symmetric algorithm.
d) You do not need to have a shared secret key.

(a) No. You will need to encrypt for each recipient

(b) No.

(c) No, it's slower

Fall 2014 - Question 23

A hybrid cryptosystem:
a) Combines encryption with digital signatures.
b) Relays messages through a trusted third party.
c) Applies two layers of encryption for added security.
d) Uses public key cryptography to encrypt a session key.

- Uses public key cryptography to exchange a session key
- Uses symmetric cryptography to encrypt the communication session (encrypting and decrypting with the session key)

Fall 2014 - Question 24

Hashed passwords in the Password Authentication Protocol (PAP):
a) Reduce the value of stealing a password file.
b) Enable a password to be sent securely over a network.
c) Ensure that a password has not been maliciously modified.
d) Protect the user from a man-in-the-middle attack.

(b) No – passwords are sent in plain text

(c) No – the hash is not a signature for the password

(d) No – passwords are sent in plain text & can be relayed via a man in the middle

Fall 2014 - Question 25

With Kerberos, how does Bob know that Alice is authorized to talk with him?
Alice sends Bob a ticket that:
a) Is encrypted with Bob's secret key.
b) Is encrypted with Bob's public key.
c) Contains an authorization token that Bob can present to Kerberos.
d) Is encrypted with Alice's private key.

Bob can decrypt the ticket to extract the session key

Fall 2014 - Question 26

OAuth:
a) Relies on HTTP redirection to allow a user to specify access permissions to a particular service.
b) Is a centralized authorization service that manages access rights among services from multiple vendors.
c) Is a protocol to securely obtain user IDs and passwords to services that an application needs to access.
d) Is a protocol designed to authenticate users using multi-factor authentication prior to accessing a service.

(b) No. It's designed to authorize services it offers.

(c) No. It does not provide passwords.

(d) Authentication mechanisms are not defined in Oauth/OpenID

Fall 2014 - Question 27

How does a service validate an OpenID token that is returned after a user authenticates with the Identity Provider?
a) It decrypts it with its private key.
b) It sends it to the identity provider (authorization server).
c) It asks the user to decrypt it.
d) It asks the user to validate the signature.

- The service contacts the identity provider to request an access token

Fall 2014 - Question 28

You need stateful inspection in a packet filtering firewall to:
a) Drop all packets from the Internet whose source address matches that of an internal computer.
b) Allow only UDP requests to a streaming media server.
c) Permit return traffic only in response to outbound messages.
d) Deny all TCP traffic to port 80 (HTTP) on a specific machine.

(a) (b) (d) – Do not require state

Fall 2014 - Question 29

Which services would not run on a computer in the DMZ?
a) Web server for serving your organization's web pages.
b) Mail server for accepting incoming email for your organization.
c) Web caching proxy for outbound web surfing.
d) FTP server for providing file access to external users.

Services that accept external requests go in the DMZ

Fall 2014 - Question 30

Transport mode differs from tunnel mode in that:
a) Tunnel mode provides a bidirectional communication channel.
b) Transport mode is responsible for optimally routing the packet to its destination.
c) Tunnel mode encapsulates the entire IP packet and prepends a new IP header.
d) Transport mode encrypts the entire IP packet while tunnel mode encrypts only the data.

(a) Any VPN is bidirectional

(b) Routing is handled by IP

(c) Yes – Transport mode is host-to-host. Tunnel mode uses tunneling – packet encapsulation

(d) No: Tunnel mode supports encrypting the entire packet. Transport mode cannot encrypt the headers that are needed for routing

Fall 2014 - Question 31

Which statement is not generally true about VPNs?
a) Using a VPN is faster than not using one.
b) A VPN enables multiple LANs (local area networks) to communicate over the public Internet.
c) Using a VPN is cheaper than using a leased line (private network).
d) VPNs encrypt packet data.

(a) Extra software to encrypt & validate data

The End