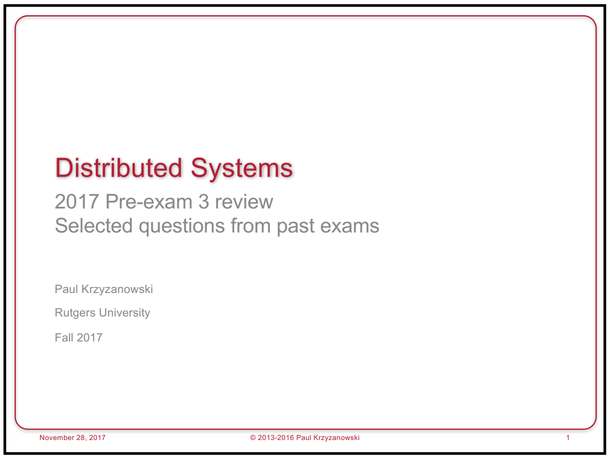


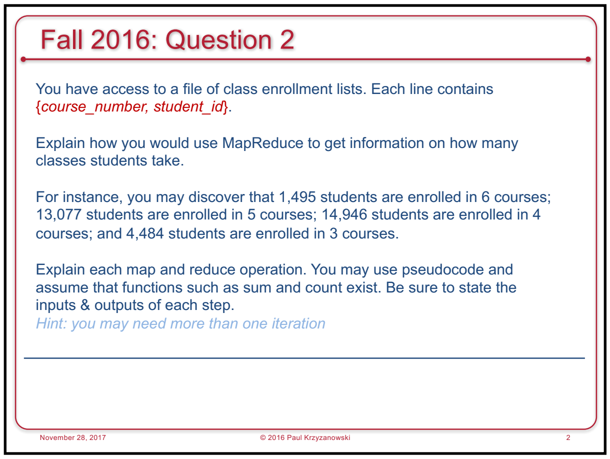
Distributed Systems

2017 Pre-exam 3 review
Selected questions from past exams

Paul Krzyzanowski

Rutgers University

Fall 2017



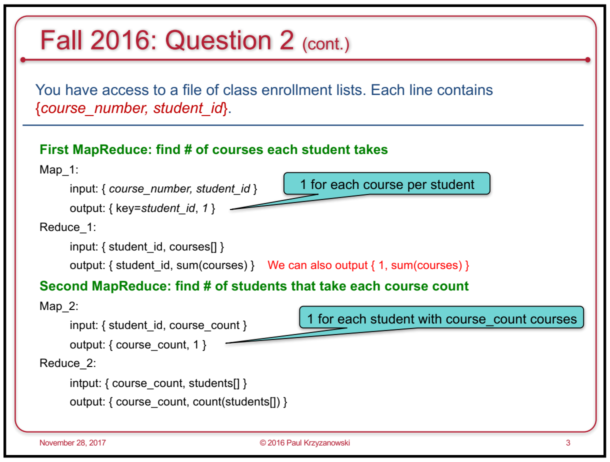
Fall 2016: Question 2

You have access to a file of class enrollment lists. Each line contains *{course_number, student_id}*.

Explain how you would use MapReduce to get information on how many classes students take.

For instance, you may discover that 1,495 students are enrolled in 6 courses; 13,077 students are enrolled in 5 courses; 14,946 students are enrolled in 4 courses; and 4,484 students are enrolled in 3 courses.

Explain each map and reduce operation. You may use pseudocode and assume that functions such as sum and count exist. Be sure to state the inputs & outputs of each step.
Hint: you may need more than one iteration



Fall 2016: Question 2 (cont.)

You have access to a file of class enrollment lists. Each line contains *{course_number, student_id}*.

First MapReduce: find # of courses each student takes

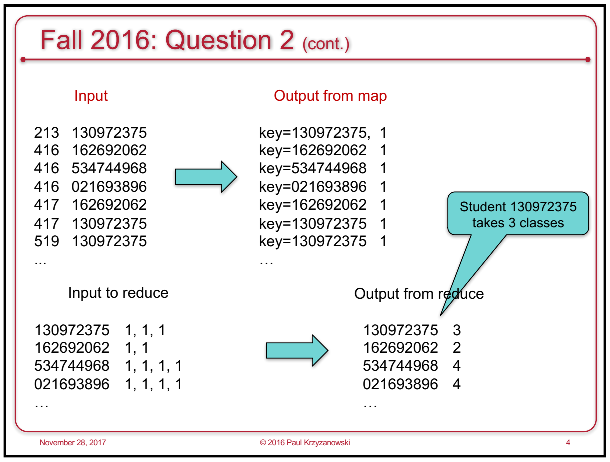
```
Map_1:
    input: { course_number, student_id }
    output: { key=student_id, 1 }
Reduce_1:
    input: { student_id, courses[] }
    output: { student_id, sum(courses) }    We can also output { 1, sum(courses) }
```

1 for each course per student

Second MapReduce: find # of students that take each course count

```
Map_2:
    input: { student_id, course_count }
    output: { course_count, 1 }
Reduce_2:
    intput: { course_count, students[] }
    output: { course_count, count(students[]) }
```

1 for each student with course_count courses



Fall 2016: Question 2 (cont.)

Input

```
213   130972375
416   162692062
416   534744968
416   021693896
417   162692062
417   130972375
519   130972375
…
```

Output from map

```
key=130972375,  1
key=162692062   1
key=534744968   1
key=021693896   1
key=162692062   1
key=130972375   1
key=130972375   1
…
```

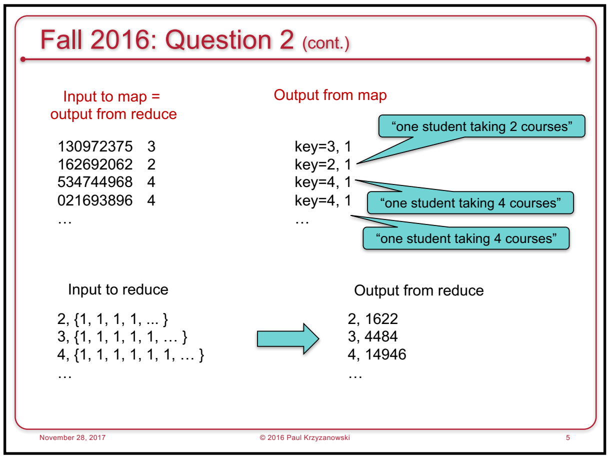
Student 130972375 takes 3 classes

Input to reduce

```
130972375    1, 1, 1
162692062    1, 1
534744968    1, 1, 1, 1
021693896    1, 1, 1, 1
…
```

Output from reduce

```
130972375   3
162692062   2
534744968   4
021693896   4
…
```



Fall 2016: Question 2 (cont.)

Input to map =
output from reduce

```
130972375   3
162692062   2
534744968   4
021693896   4
…
```

Output from map

```
key=3, 1
key=2, 1
key=4, 1
key=4, 1
…
```

“one student taking 2 courses”

“one student taking 4 courses”

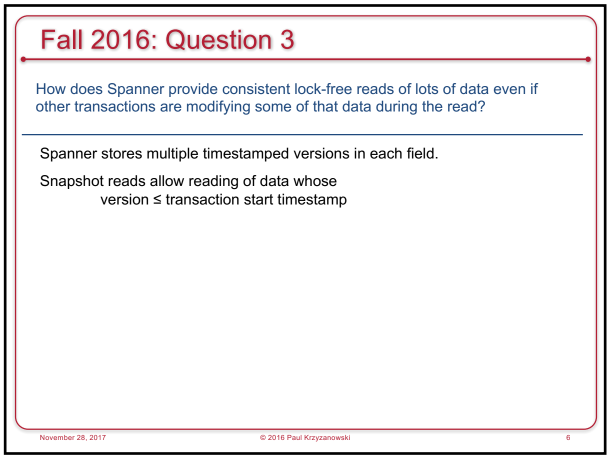
“one student taking 4 courses”

Input to reduce

```
2, {1, 1, 1, 1, ... }
3, {1, 1, 1, 1, 1, … }
4, {1, 1, 1, 1, 1, 1, … }
…
```

Output from reduce

```
2, 1622
3, 4484
4, 14946
…
```

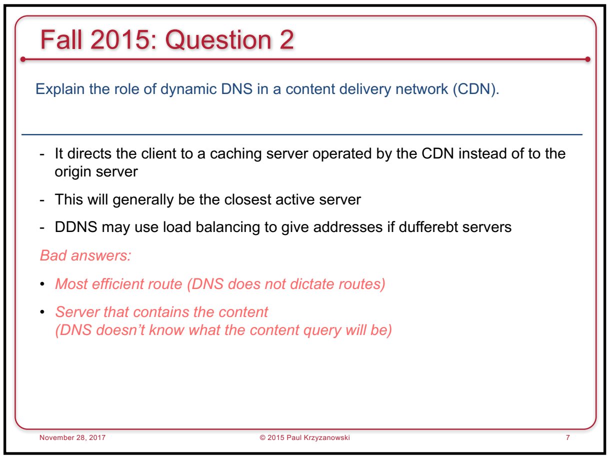


Fall 2016: Question 3

How does Spanner provide consistent lock-free reads of lots of data even if other transactions are modifying some of that data during the read?

Spanner stores multiple timestamped versions in each field.

Snapshot reads allow reading of data whose
version ≤ transaction start timestamp



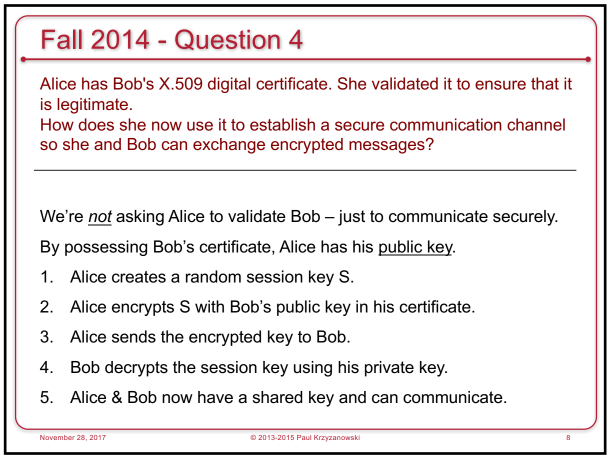
Fall 2015: Question 2

Explain the role of dynamic DNS in a content delivery network (CDN).

- It directs the client to a caching server operated by the CDN instead of to the origin server
- This will generally be the closest active server
- DDNS may use load balancing to give addresses if dufferebt servers

Bad answers:

- *Most efficient route (DNS does not dictate routes)*
- *Server that contains the content*
  (DNS doesn't know what the content query will be)



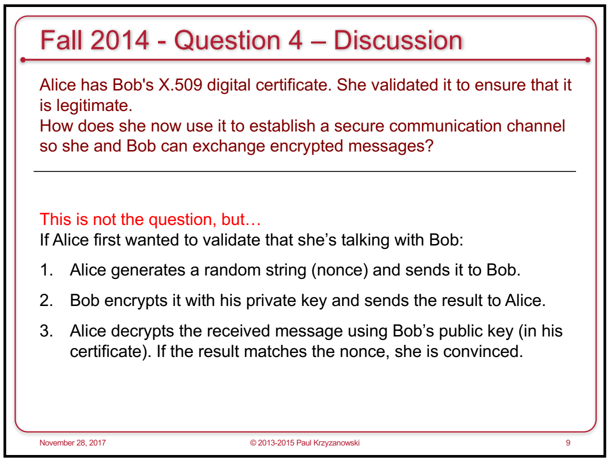
Fall 2014 - Question 4

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.

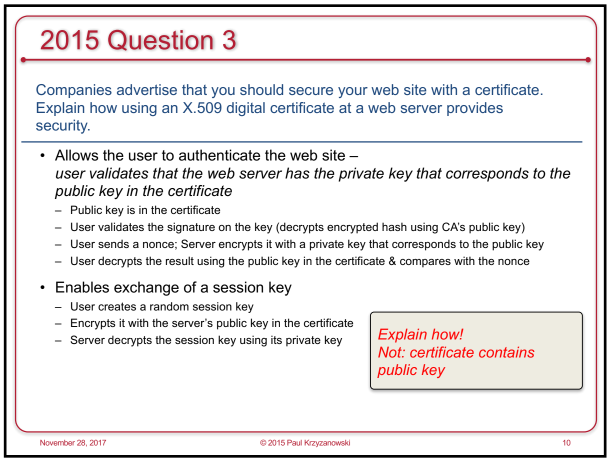


Fall 2014 - Question 4 – Discussion

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.

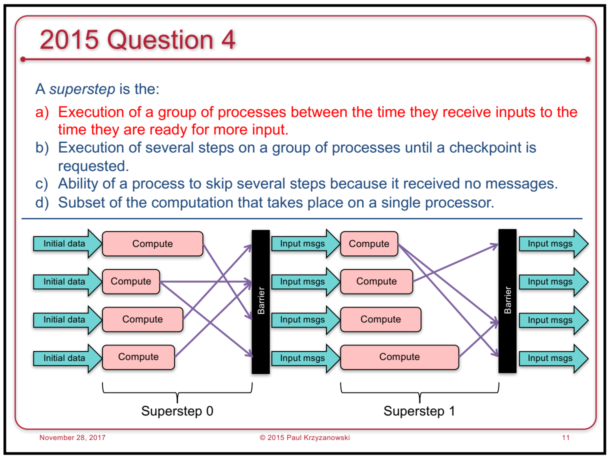


2015 Question 3

Companies advertise that you should secure your web site with a certificate.
Explain how using an X.509 digital certificate at a web server provides security.

- Allows the user to authenticate the web site –
  user validates that the web server has the private key that corresponds to the public key in the certificate
  - Public key is in the certificate
  - User validates the signature on the key (decrypts encrypted hash using CA's public key)
  - User sends a nonce; Server encrypts it with a private key that corresponds to the public key
  - User decrypts the result using the public key in the certificate & compares with the nonce
- Enables exchange of a session key
  - User creates a random session key
  - Encrypts it with the server's public key in the certificate
  - Server decrypts the session key using its private key

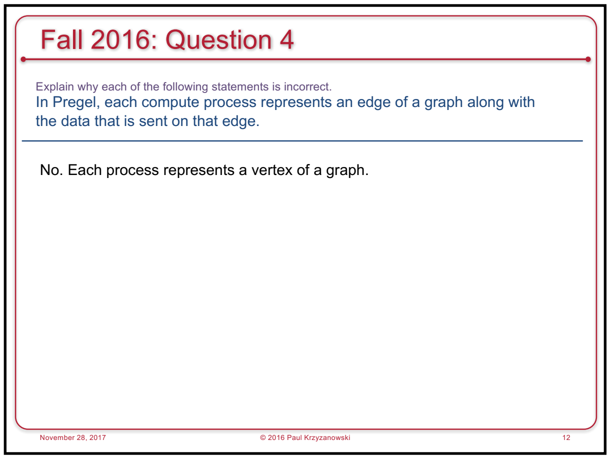
Explain how!
Not: certificate contains public key



2015 Question 4

A *superstep* is the:

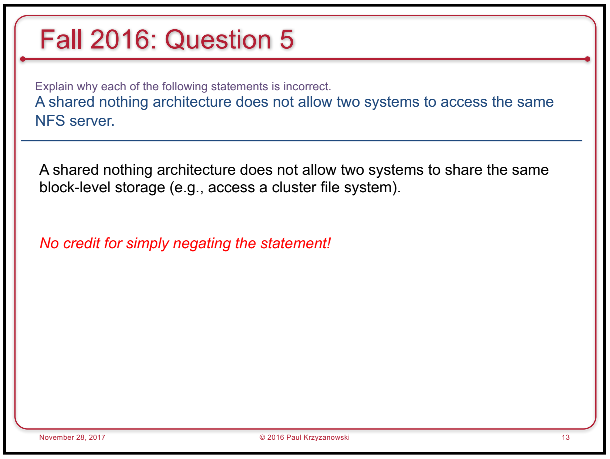
a) Execution of a group of processes between the time they receive inputs to the time they are ready for more input.
b) Execution of several steps on a group of processes until a checkpoint is requested.
c) Ability of a process to skip several steps because it received no messages.
d) Subset of the computation that takes place on a single processor.



Fall 2016: Question 4

Explain why each of the following statements is incorrect.
In Pregel, each compute process represents an edge of a graph along with the data that is sent on that edge.

No. Each process represents a vertex of a graph.

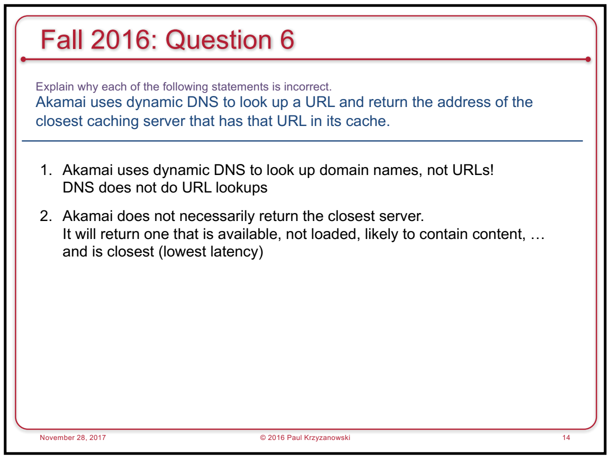


Fall 2016: Question 5

Explain why each of the following statements is incorrect.
A shared nothing architecture does not allow two systems to access the same NFS server.

A shared nothing architecture does not allow two systems to share the same block-level storage (e.g., access a cluster file system).

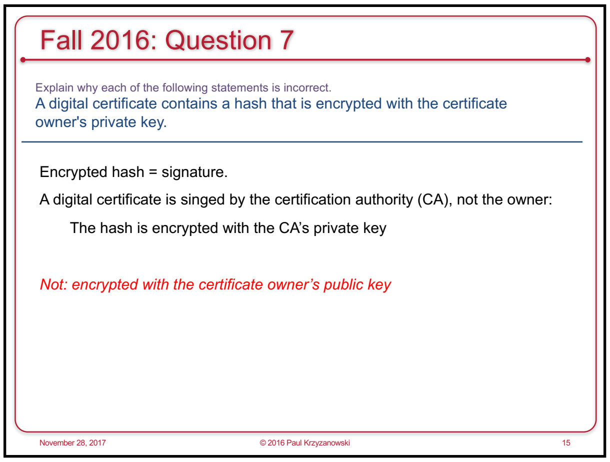
No credit for simply negating the statement!



Fall 2016: Question 6

Explain why each of the following statements is incorrect.
Akamai uses dynamic DNS to look up a URL and return the address of the closest caching server that has that URL in its cache.

1. Akamai uses dynamic DNS to look up domain names, not URLs!
   DNS does not do URL lookups
2. Akamai does not necessarily return the closest server.
   It will return one that is available, not loaded, likely to contain content, … and is closest (lowest latency)



Fall 2016: Question 7

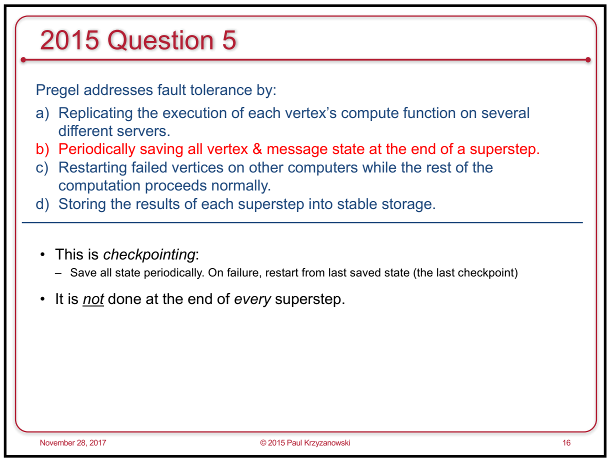
Explain why each of the following statements is incorrect.
A digital certificate contains a hash that is encrypted with the certificate owner's private key.

Encrypted hash = signature.

A digital certificate is singed by the certification authority (CA), not the owner:

The hash is encrypted with the CA's private key

Not: encrypted with the certificate owner's public key

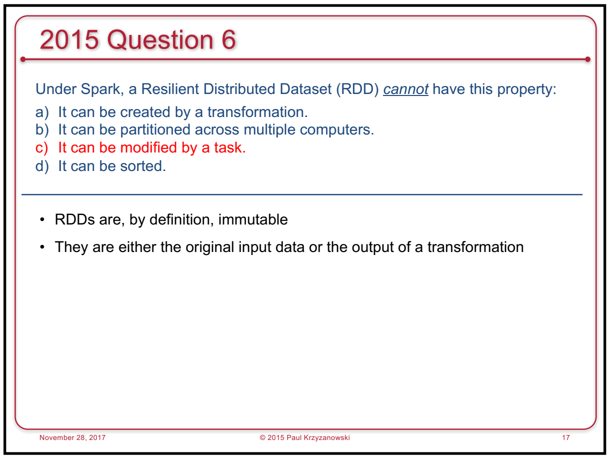


2015 Question 5

Pregel addresses fault tolerance by:

a) Replicating the execution of each vertex's compute function on several different servers.
b) Periodically saving all vertex & message state at the end of a superstep.
c) Restarting failed vertices on other computers while the rest of the computation proceeds normally.
d) Storing the results of each superstep into stable storage.

- This is *checkpointing*:
  - Save all state periodically. On failure, restart from last saved state (the last checkpoint)
- It is *not* done at the end of *every* superstep.

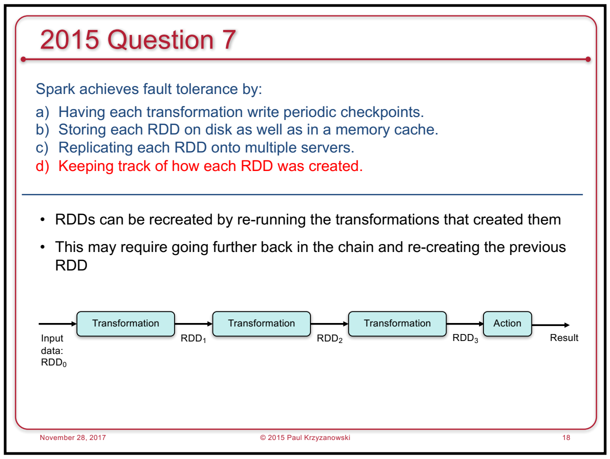


2015 Question 6

Under Spark, a Resilient Distributed Dataset (RDD) *cannot* have this property:

a) It can be created by a transformation.
b) It can be partitioned across multiple computers.
c) It can be modified by a task.
d) It can be sorted.

- RDDs are, by definition, immutable
- They are either the original input data or the output of a transformation

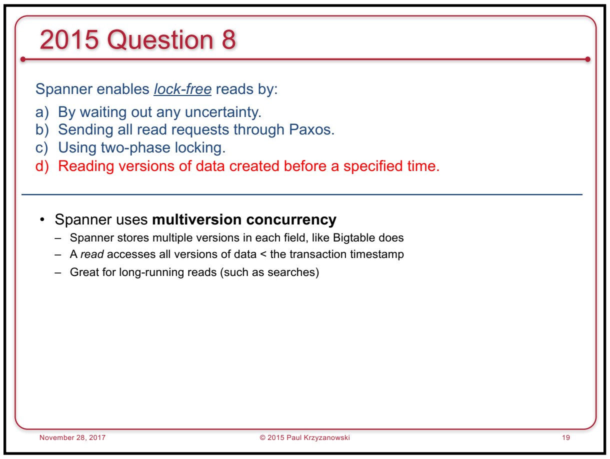


2015 Question 7

Spark achieves fault tolerance by:

a) Having each transformation write periodic checkpoints.
b) Storing each RDD on disk as well as in a memory cache.
c) Replicating each RDD onto multiple servers.
d) Keeping track of how each RDD was created.

- RDDs can be recreated by re-running the transformations that created them
- This may require going further back in the chain and re-creating the previous RDD

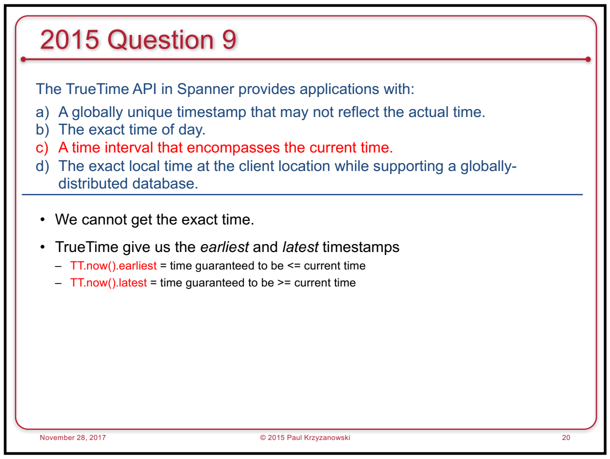


2015 Question 8

Spanner enables *lock-free* reads by:

a) By waiting out any uncertainty.
b) Sending all read requests through Paxos.
c) Using two-phase locking.
d) Reading versions of data created before a specified time.

- Spanner uses **multiversion concurrency**
  - Spanner stores multiple versions in each field, like Bigtable does
  - A *read* accesses all versions of data < the transaction timestamp
  - Great for long-running reads (such as searches)

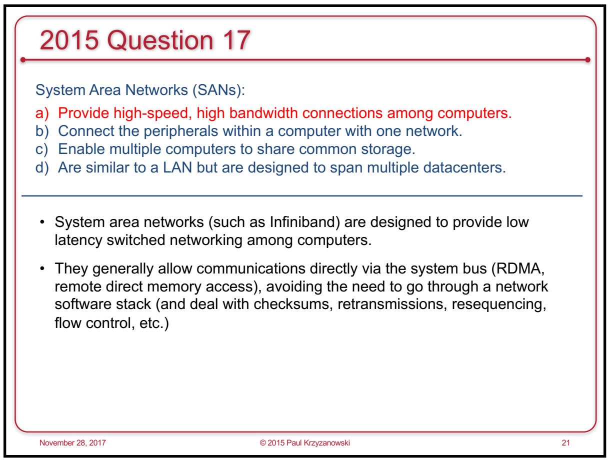


2015 Question 9

The TrueTime API in Spanner provides applications with:

a) A globally unique timestamp that may not reflect the actual time.
b) The exact time of day.
c) A time interval that encompasses the current time.
d) The exact local time at the client location while supporting a globally-distributed database.

- We cannot get the exact time.
- TrueTime give us the *earliest* and *latest* timestamps
  - TT.now().earliest = time guaranteed to be <= current time
  - TT.now().latest = time guaranteed to be >= current time

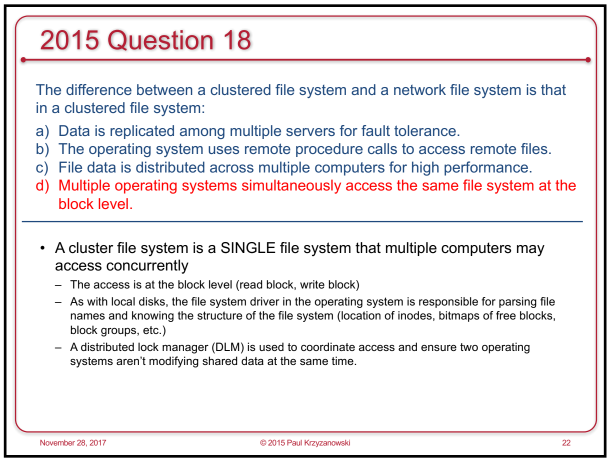


2015 Question 17

System Area Networks (SANs):

a) Provide high-speed, high bandwidth connections among computers.
b) Connect the peripherals within a computer with one network.
c) Enable multiple computers to share common storage.
d) Are similar to a LAN but are designed to span multiple datacenters.

- System area networks (such as Infiniband) are designed to provide low latency switched networking among computers.
- They generally allow communications directly via the system bus (RDMA, remote direct memory access), avoiding the need to go through a network software stack (and deal with checksums, retransmissions, resequencing, flow control, etc.)

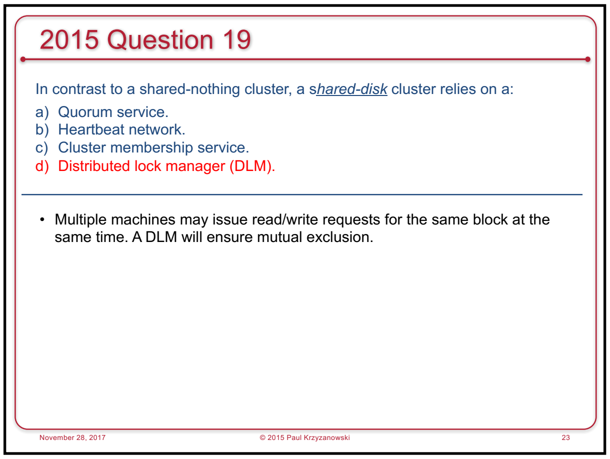


2015 Question 18

The difference between a clustered file system and a network file system is that in a clustered file system:

a) Data is replicated among multiple servers for fault tolerance.
b) The operating system uses remote procedure calls to access remote files.
c) File data is distributed across multiple computers for high performance.
d) Multiple operating systems simultaneously access the same file system at the block level.

- A cluster file system is a SINGLE file system that multiple computers may access concurrently
  - The access is at the block level (read block, write block)
  - As with local disks, the file system driver in the operating system is responsible for parsing file names and knowing the structure of the file system (location of inodes, bitmaps of free blocks, block groups, etc.)
  - A distributed lock manager (DLM) is used to coordinate access and ensure two operating systems aren't modifying shared data at the same time.

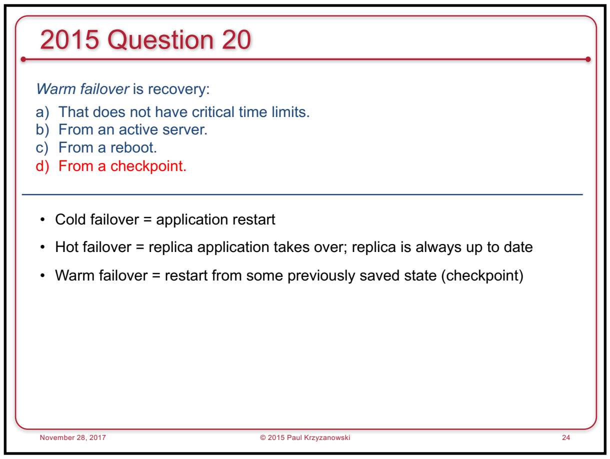


2015 Question 19

In contrast to a shared-nothing cluster, a s*hared-disk* cluster relies on a:

a) Quorum service.
b) Heartbeat network.
c) Cluster membership service.
d) Distributed lock manager (DLM).

- Multiple machines may issue read/write requests for the same block at the same time. A DLM will ensure mutual exclusion.

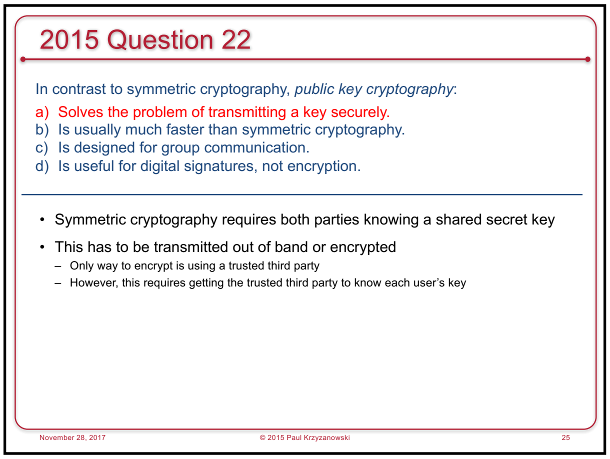


2015 Question 20

Warm failover is recovery:

a) That does not have critical time limits.
b) From an active server.
c) From a reboot.
d) From a checkpoint.

- Cold failover = application restart
- Hot failover = replica application takes over; replica is always up to date
- Warm failover = restart from some previously saved state (checkpoint)

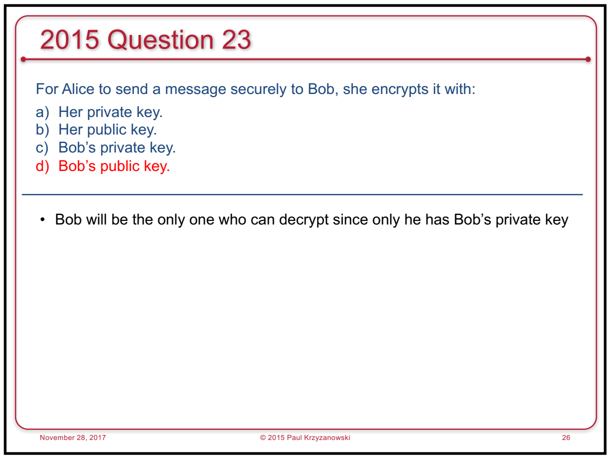


2015 Question 22

In contrast to symmetric cryptography, *public key cryptography*:

a) Solves the problem of transmitting a key securely.
b) Is usually much faster than symmetric cryptography.
c) Is designed for group communication.
d) Is useful for digital signatures, not encryption.

- Symmetric cryptography requires both parties knowing a shared secret key
- This has to be transmitted out of band or encrypted
  - Only way to encrypt is using a trusted third party
  - However, this requires getting the trusted third party to know each user's key

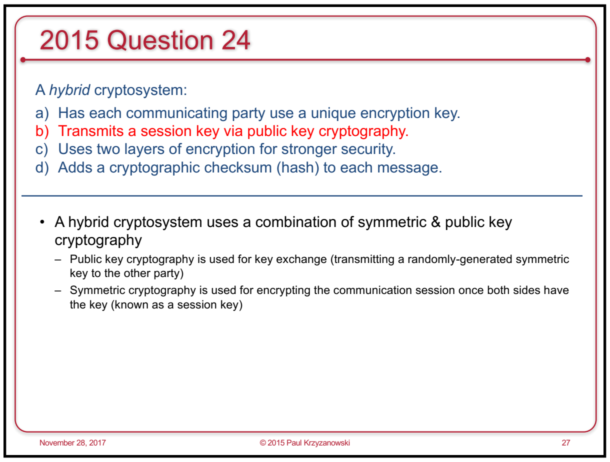


2015 Question 23

For Alice to send a message securely to Bob, she encrypts it with:

a) Her private key.
b) Her public key.
c) Bob's private key.
d) Bob's public key.

- Bob will be the only one who can decrypt since only he has Bob's private key

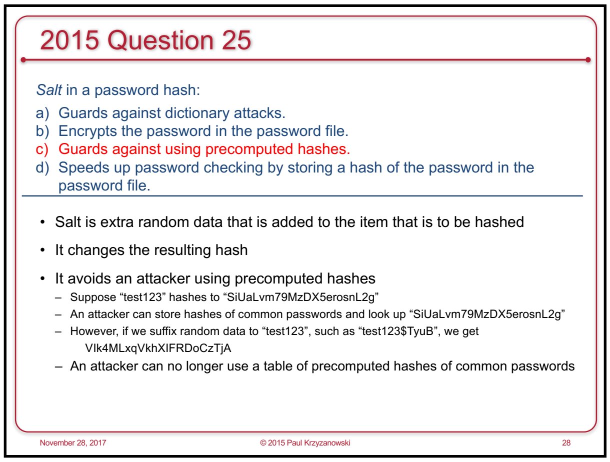


2015 Question 24

A *hybrid* cryptosystem:

a) Has each communicating party use a unique encryption key.
b) Transmits a session key via public key cryptography.
c) Uses two layers of encryption for stronger security.
d) Adds a cryptographic checksum (hash) to each message.

- A hybrid cryptosystem uses a combination of symmetric & public key cryptography
  - Public key cryptography is used for key exchange (transmitting a randomly-generated symmetric key to the other party)
  - Symmetric cryptography is used for encrypting the communication session once both sides have the key (known as a session key)

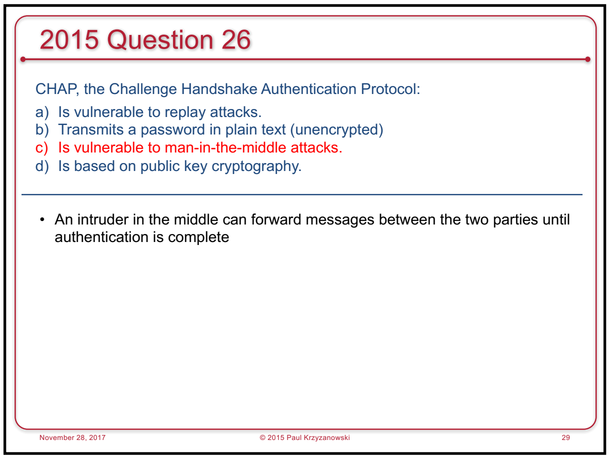


2015 Question 25

Salt in a password hash:

a) Guards against dictionary attacks.
b) Encrypts the password in the password file.
c) Guards against using precomputed hashes.
d) Speeds up password checking by storing a hash of the password in the password file.

- Salt is extra random data that is added to the item that is to be hashed
- It changes the resulting hash
- It avoids an attacker using precomputed hashes
  - Suppose "test123" hashes to "SiUaLvm79MzDX5erosnL2g"
  - An attacker can store hashes of common passwords and look up "SiUaLvm79MzDX5erosnL2g"
  - However, if we suffix random data to "test123", such as "test123$TyuB", we get Vlk4MLxqVkhXlFRDoCzTjA
  - An attacker can no longer use a table of precomputed hashes of common passwords

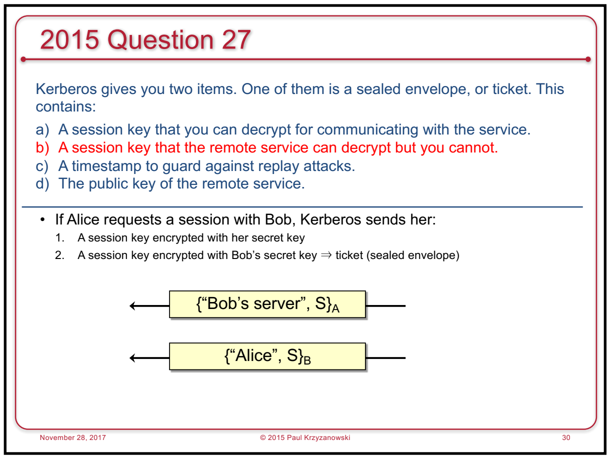


2015 Question 26

CHAP, the Challenge Handshake Authentication Protocol:

a) Is vulnerable to replay attacks.
b) Transmits a password in plain text (unencrypted)
c) Is vulnerable to man-in-the-middle attacks.
d) Is based on public key cryptography.

- An intruder in the middle can forward messages between the two parties until authentication is complete



2015 Question 27

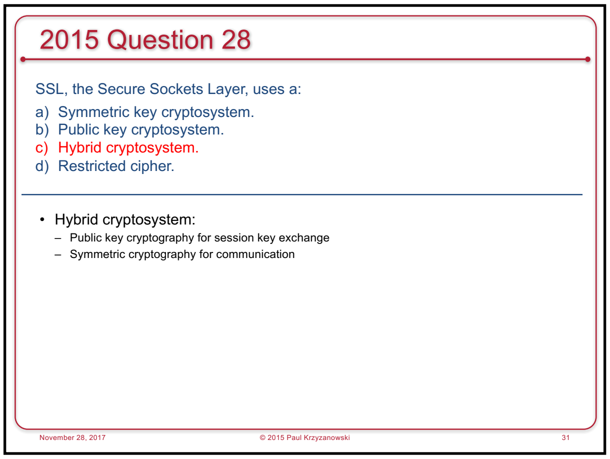
Kerberos gives you two items. One of them is a sealed envelope, or ticket. This contains:

a) A session key that you can decrypt for communicating with the service.
b) A session key that the remote service can decrypt but you cannot.
c) A timestamp to guard against replay attacks.
d) The public key of the remote service.

- If Alice requests a session with Bob, Kerberos sends her:
  1. A session key encrypted with her secret key
  2. A session key encrypted with Bob's secret key ⇒ ticket (sealed envelope)

$\{\text{“Bob’s server”}, S\}_A$

$\{\text{“Alice”}, S\}_B$

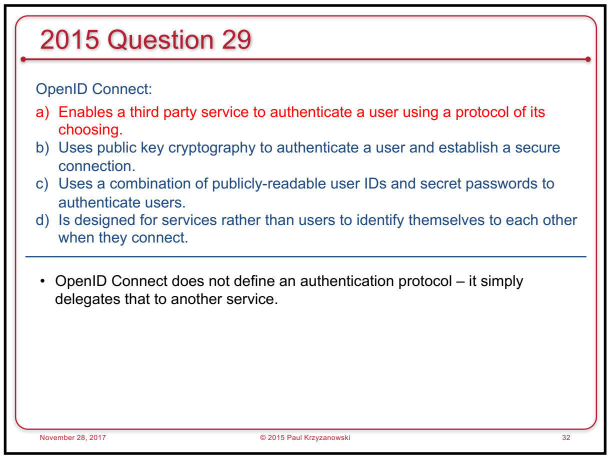


2015 Question 28

SSL, the Secure Sockets Layer, uses a:

a) Symmetric key cryptosystem.
b) Public key cryptosystem.
c) Hybrid cryptosystem.
d) Restricted cipher.

- Hybrid cryptosystem:
  - Public key cryptography for session key exchange
  - Symmetric cryptography for communication

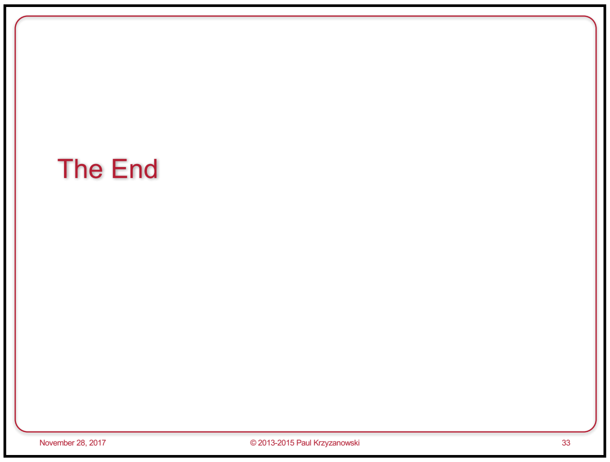


2015 Question 29

OpenID Connect:

a) Enables a third party service to authenticate a user using a protocol of its choosing.
b) Uses public key cryptography to authenticate a user and establish a secure connection.
c) Uses a combination of publicly-readable user IDs and secret passwords to authenticate users.
d) Is designed for services rather than users to identify themselves to each other when they connect.

- OpenID Connect does not define an authentication protocol – it simply delegates that to another service.



The End