COMP3008J Distributed Systems Tutorial #3.

1. What is routing? What criteria are considered during a 'routing decision'?

Solution: Routing is the process of finding the best path for messages or data to travel from a source to a destination across a some network. Several factors influence the routing decision process:

- Destination
- Bandwidth
- Network Topology
- Routing tables
- Load Balancing
- Network Congestion

2. What is an overlay? Briefly describe what a routing overlay is.

Solution: An overlay refers to a virtual network that is built on top of an existing physical network.

A routing overlay is a virtualised network layer built on top of an existing physical network, specifically designed to manage routing and forwarding decisions independently from the underlying infrastructure. A routing overlay enables more dynamic, scalable, and flexible network routing

3. Below is the Leaf Set, Routing Table and Neighbourhood Set for a Pastry Node with a Globally Unique Identifier (GUID) of 10233102.

   For each of the following network addresses, identify the next node Node 10233102 would forward the message to.

   - 10233001
   - 10211122
   - 10233102
   - 30233123
   - 10233333
   - 01233333

Solution:

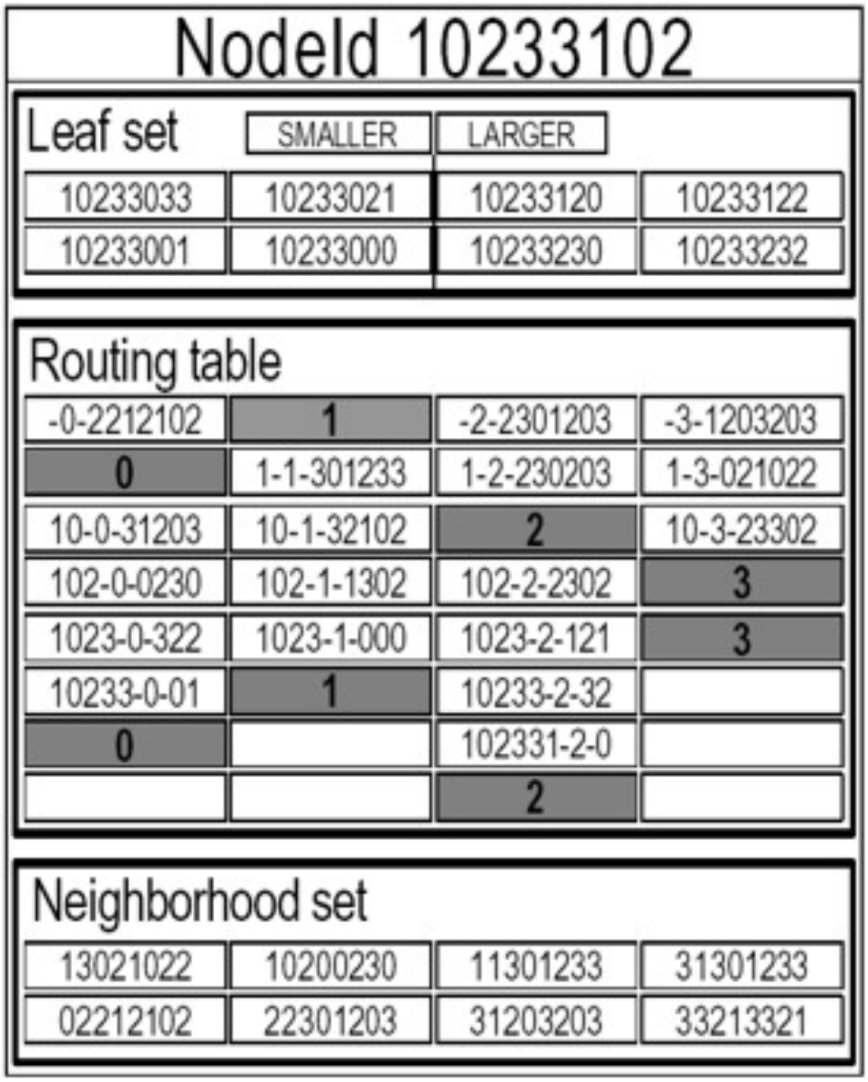

Figure 1: Routing table for Question 3.

- 10233001 → 10233001
- 10211122 → 10211302
- 10233102 → 10233120
- 30233123 → 31203203
- 10233333 → 10233232
- 01233333 → 02212102

4. X, Y and Z are processes that use Reliable Multicast for communication between them. The following is a summary of a sequence of messages that are sent by these processes. Sometimes these are not received by one or both of the other processes. We assume that the sequence number for all processes begins at 0.

   1. Y sends message. X and Z both receive.
   2. X sends message. Only Y receives.
   3. Y sends message. X and Z both receive.
   4. Z sends message. Only Y receives.
   5. Z sends message. X and Y both receive.

   - For each message, what is the sequence number and the acknowledgement?
   - What other messages will be sent to make this multicast reliable (you should include these messages in the correct position in the sequence)?

Solution:

1. Y sends message. X and Z both receive.
   $\{S_Y = 0, \{\}, \text{message}\}$.
2. X sends message. Only Y receives.
   $\{S_X = 0, \{ < Y, 0 >\}, \text{message}\}$.
3. Y sends message. X and Z both receive.
   $\{S_Y = 1, \{ < X, 0 >\}, \text{message}\}$.
   To make it reliable: Z detects it has missed a message from X, will send a NAK 0 to X. X will resend message 2 to Z.
4. Z sends message. Only Y receives.
   $\{S_Z = 0, \{ < X, 0 >, < Y, 1 >\}, \text{message}\}$.
5. Z sends message. X and Y both receive.
   $\{S_Z = 1, \{ < X, 0 >, < Y, 1 >\}, \text{message}\}$.
   To make it reliable: X detects it has missed a message from Z, it will place this message in a hold-back and will send a NAK 0 to Z. Z will resend message 4 to X.

5. Briefly explain the gossip architecture. Give an example of when it may be necessary to sacrifice consistency for high availability.

Solution: The Gossip architecture is a framework for implementing highly available services. Data is replicated close to the groups of clients that need it. Replica Managers then exchange "gossip" messages periodically (based on time or number of updates to the file) in order to convey the updates they have each received from clients. The system makes two guarantees: each client obtains a consistent service over time and relaxed consistency between replicas. Example: Taobao, food ordering.

6. What are Symmetric and Asymmetric Encryption algorithms?

Solution:

- Symmetric encryption uses a single key that needs to be shared among the people who need to receive the message while asymmetrical encryption uses a pair of public key and private key to encrypt and decrypt messages when communicating.
- Symmetric encryption is an old technique. Asymmetric is relatively new.
- Asymmetric encryption was introduced to complement the inherent problem of the need to share the key in symmetrical encryption model, eliminating the need to share the key by using a pair of public-private keys.
- Asymmetric encryption takes relatively more time than symmetric encryption.

7. What are the main differences between capability lists and access lists?

Solution: Access control lists are created by dividing the resource access matrix column-wise, whereas a capability list is created by dividing the access matrix row-wise.

ACLs are good when users manage their own files, protection is data-oriented and easy to change rights to a resource.
Capability lists are easy to delegate, easy to add/delete users but more difficult to implement.

8. What do we mean by public key cryptography algorithms? Why are they generally used?

Solution: Each user has 2 keys, public key and private key. Knowledge of one does not reveal the other
Public key used to encrypt messages and verify signatures.
Private key **only known by the owner**. Used to decrypt messages and create signatures.

Used to ensure secrecy and integrity of messages, authentication of messages and to creates digital signatures.

9. Define the Bell-LaPadula security model.

Solution: The Bell-LaPadula model uses security labels (or classifications) assigned to both subjects (e.g., users or processes) and objects (e.g., files, databases). These labels dictate the access control rules that determine who can read or write to certain objects, depending on their classification. BLP deals with confidentiality in order **to prevent unauthorized reading**.

10. Explain grid computing.

Solution: Grid computing is a type of distributed computing where multiple computer systems, often geographically dispersed, are networked together to work collaboratively on complex tasks. It allows for the sharing, selection, and aggregation of resources (like processing power, storage, or data) from different organizations or locations, enabling the execution of large-scale computations that would be difficult for a single machine or data center to handle. Resources supported by heterogeneous computer hardware, operating systems, programming languages and applications.

11. What are the primary requirements of a computing grid?

Solution:

- Remote Access to Resources
- Processing where it is stored and managed
- Dynamic Service Creation
- Metadata
- Directory Services
- Management Software