DISTRIBUTED SYSTEM STRUCTURES

NETWORK OPERATING SYSTEMS

- The users are **aware** of the physical structure of the network.
- Each site has its own OS and some protocol (i.e. FTP or Telnet) provides an interface to those OS.
- Users must know machine and directory structure in order to find a file.

DISTRIBUTED OPERATING SYSTEMS

- The users may be **UNaware** of the physical structure of the network.
- Data and process usage appears seamless.
- This seamlessness results because of the following actions:

DATA MIGRATION

- Send the file (whole or partial) to be worked on at the requested site.
- Must send any modified portion of the file back to the originator.

COMPUTATION MIGRATION

- Do work on the machine where file is located.
- Requires remote procedure calls. Send message to the remote machine that invokes a process to handle the request.

PROCESS MIGRATION

- Move the whole operation to the "best" location. This may be done for

a) Load balancing - distribute processes across network to even the workload.
b) Computation speedup - subprocesses run concurrently on different sites.
c) Hardware preference - process execution may require specialized processor.
d) Software preference - required software might be available at only a particular site.
e) Data access - run process remotely, rather than transfer all data locally.

REMOTE SERVICES

Requests for access to a remote file are delivered to the server. Access requests are translated to messages for the server, and the server replies are packed as messages and sent back to the user.

- A common way to achieve this is via the **Remote Procedure Call (RPC)** paradigm.

- Messages addressed to an (RPC) daemon listening to a **port** on the remote system contain the name of a process to run and the parameters to pass to that process. The process is executed as requested, and any output is sent back to the requester in a separate message.

- A port is a number included at the start of a message packet. A system can have many ports within its one network address to differentiate the network services it supports.

- The (RPC) scheme requires binding client and server port.

- Binding information may be pre-decided, in the form of fixed port addresses.

  - At compile time, an (RPC) call has a fixed port number associated with it.

  - Once a program is compiled, the server cannot change the port number of the requested service.

- Binding can be done dynamically by a rendezvous mechanism.

  - Operating system provides a rendezvous daemon on a fixed (RPC) port.

  - Client then sends a message to the rendezvous daemon requesting the port address of the (RPC)it needs to execute.

- A distributed file system (DFS) can be implemented as a set of (RPC) daemons and clients.

  -The messages are addressed to the (DFS) port on a server on which a file operation is to take place.

  -The message contains the disk operation to be performed (i.e., **read, write, rename, delete, or status**)

  -The return message contains any data resulting from that call, which is executed by the (DFS) daemon on behalf of the client.

Threads

- Threads can send and receive messages while other operations within the task continue asynchronously.

- **Pop-up thread** - created on ‘‘as needed’’ basis to respond to new RPC.

  - Cheaper to start new thread than to restore existing one.

  - No threads block waiting for New work; no context has to be saved, or restored.

  - Incoming (RPC)s do not have to be copied to a buffer within a server thread.

- RPC’s to processes on the same machine as the caller are made more lightweight via shared memory between threads in different processes running on same machine.

Robustness

To ensure that the system is robust, we must:

- **Detect failures.**
- **Reconfigure** the system so that computation may continue.
- **Recover** when a site or a link is repaired.

Failure Detection -

To detect link and site failure, we use a **handshaking** procedure.

- At fixed intervals, sites A and B send each other an **I-am-up** message.
- If site A does not receive this message within a predetermined time period, it can assume that:

  a) site B has failed,

  b) that the link between A and B has failed,

  c) or that the message from B has been lost.

- At the time site A sends the Are-you-up? message, it specifies a time interval during which it is willing to wait for the reply from B.
- If A does not receive B's reply message within the time interval, A may conclude that one or more of the following situations has occurred:

  a) Site B is down.

  b) The direct link if one exists from A to B is down.

  c) The alternative path from A to B is down.

  d) The message has been lost.

Reconfiguration -

- A procedure that allows the system to reconfigure and to continue its normal mode of operation.

- If a direct link from (A to (B has failed, this information must be broadcast to every site in the system, so that the various routing tables can be updated accordingly.

- If it is believed that a site has failed because it can no longer be reached then every site in the system must be so notified, so that they will no longer attempt to use the services of the failed site.

Recovery from Failure -

- When a failed link or site is repaired, it must be integrated into the system gracefully and smoothly.

- Suppose that a link between A and B has failed. When it is repaired, both A and B must be notified. We can accomplish this notification by continuously repeating the handshaking procedure.

- Suppose that site B has failed. When it recovers, it must notify all other sites that it is up again. Site B then may have to receive from the other sites various information to update its local tables.

Consider These Design Issues:

Transparency and locality -

- Distributed systems should look like conventional, centralized system and not distinguish between local and remote resources.

User mobility -

- Brings user's environment (i.e., home directory) to the machine where the user logs in.

F**ault tolerance -**

- System should continue functioning, perhaps in a degraded form, when faced with various types of failures.

Scalability -

- System should adapt to increased service load.

Large-scale systems -

- Service demand from any system component should be bounded by a constant that is independent of the number of nodes.

Process structure of the server -

- Servers should operate efficiently in peak periods; use of light-weight processes or threads.