1.5 Distributed Systems

A network, in the simplest terms, is a communication path between two or more systems. Distributed systems depend on networking for their functionality. By being able to communicate, distributed systems are able to share computational tasks, and provide a rich set of features to users.

Networks vary by the protocols used, the distances between nodes, and the transport media. TCP/IP is the most common network protocol, although ATM and other protocols are in widespread use. Likewise, operating-system support of protocols varies. Most operating systems support TCP/IP, including the Windows and UNIX operating systems. Some systems support proprietary protocols to suit their needs. To an operating system, a network protocol simply needs an interface device-a network adapter, for example-with a device driver to manage it, and software to package data in the communications protocol to send it and to unpackage it to receive it.

Networks are typecast based on the distances between their nodes. **A** local-area network (LAN), exists within a room, a floor, or a building. A wide-area network (WAN), usually exists between buildings, cities, or countries. **A** global company may have a WAN to connect its offices, worldwide. These networks could run one protocol or several protocols.

The media to carry networks are equally varied. They include copper wires, fiber strands, and wireless transmissions between satellites, microwave dishes, and radios. When computing devices are connected to cellular phones, they create a network. Even very short-range infrared communication can be used for networking. At a rudimentary level, whenever computers communicate they use or create a network. These networks also vary by their performance and reliability.

1.5.1 Client-Server Systems

As PCs have become faster, more powerful, and cheaper, designers have shifted away from the centralized system architecture. Terminals connected to centralized systems are now being supplanted by PCs. Correspondingly, user-interface functionality that used to be handled directly by the centralized systems is increasingly being handled by the PCs. As a result, centralized systems today act as **server systems** to satisfy requests generated by **client systems.** Server systems can be broadly categorized as compute servers and file servers.

- **Compute-server systems** provide an interface to which clients can send requests to perform an action, in response to which they execute the action and send back results to the client.
- **File-server systems** provide a file-system interface where clients can create, update, read, and delete files.

1.5.2 Peer-to-Peer Systems

The growth of computer networks-especially the Internet and World Wide Web (WWW)-has had a profound influence on the recent development of operating systems. Many PCs became connected to computer networks. With the introduction of the Web in the mid-1990s, network connectivity became an essential component of a computer system.

Virtually all modern PCs and workstations are capable of running a web browser for accessing **hypertext** documents on the Web. Operating systems now also include the system software that enables a computer to access the Internet via a local-area network or telephone connection. Several include the web browser itself, as well as electronic mail, remote login, and file-transfer clients and servers.

In contrast to the tightly coupled systems, the computer networks used in these applications consist of a collection of processors that do not share memory or a clock. Instead, each processor has its own local memory. The processors communicate with one another through various communication lines, such as high-speed buses or telephone lines. These systems are usually referred to as **loosely coupled systems** (or **distributed systems).**

Some operating systems have taken the concept of networks and distributed systems further than the notion of providing network connectivity A **network operating system** is an operating system that provides features such as file sharing across the network, and that includes a communication scheme that allows different processes on different computers to exchange messages.

A computer running a network operating system acts autonomously from all other computers on the network, although it is aware of the network and is able to communicate with other networked computers. A distributed operating system is a less autonomous environment: The different operating systems communicate closely enough to provide the illusion that only a single operating system controls the network.

1.6 Clustered Systems

Like parallel systems, **clustered systems** gather together multiple CPUs to accomplish computational work. Clustered systems differ from parallel systems, however, in that they are composed of two or more individual systems coupled together. The definition of the term clustered is not concrete; many commercial packages wrestle with what a clustered system is, and why one form is better than another. The generally accepted definition is that clustered computers share storage and are closely linked via LAN networking.

Clustering is usually performed to provide **high availability.** A layer of cluster software runs on the cluster nodes. Each node can monitor one or more of the others (over the LAN). If the monitored machine fails, the monitoring machine can take ownership of its storage, and restart the application(s) that were running on the failed machine. The failed machine can remain down, but the users and clients of the application would only see a brief interruption of service. In **asymmetric clustering,** one machine is in **hot standby mode** while the other is running the applications. The hot standby host (machine) does nothing but monitor the active server. If that server fails, the hot standby host becomes the active server. In **symmetric mode,** two or more hosts are running applications, and they are monitoring each other. This mode is obviously more efficient, as it uses all of the available hardware. It does require that more than one application be available to run. Other forms of clusters include parallel clusters and clustering over a WAN. Parallel clusters allow multiple hosts to access the same data on the shared storage. Because most operating systems lack support for this simultaneous data access by multiple hosts, parallel clusters are usually accomplished by special versions of software and special releases of applications.

In spite of improvements in distributed computing, most systems do not offer general-purpose distributed file systems. Therefore, most clusters do not allow shared access to data on the disk. For this, distributed file systems must provide access control and locking to the files to ensure no conflicting operations occur. This type of service is commonly known as a **distributed lock manager (DLM).**

Cluster technology is rapidly changing. Cluster directions include global clusters, in which the machines could be anywhere in the world (or anywhere a WAN reaches). Such projects are still the subject of research and development.

Clustered system use and features should expand greatly as **storage-area networks (SANS)**. SANs allow easy attachment of multiple hosts to multiple storage units. Current clusters are usually limited to two or four hosts due to the complexity of connecting the hosts to shared storage.