CMPU4021 Distributed Systems

Interprocess Communication

Processes

- *Computer program*
  - *Collection of instructions*
    - *process is the actual execution of those instructions*

- An executing instance of a program.

- To facilitate communication between processes, most operating systems support Inter Process Communication (IPC) resources
  - such as pipes and sockets

- IPC is used not just for communication between processes on the same system, but processes on different systems

Interprocess communication

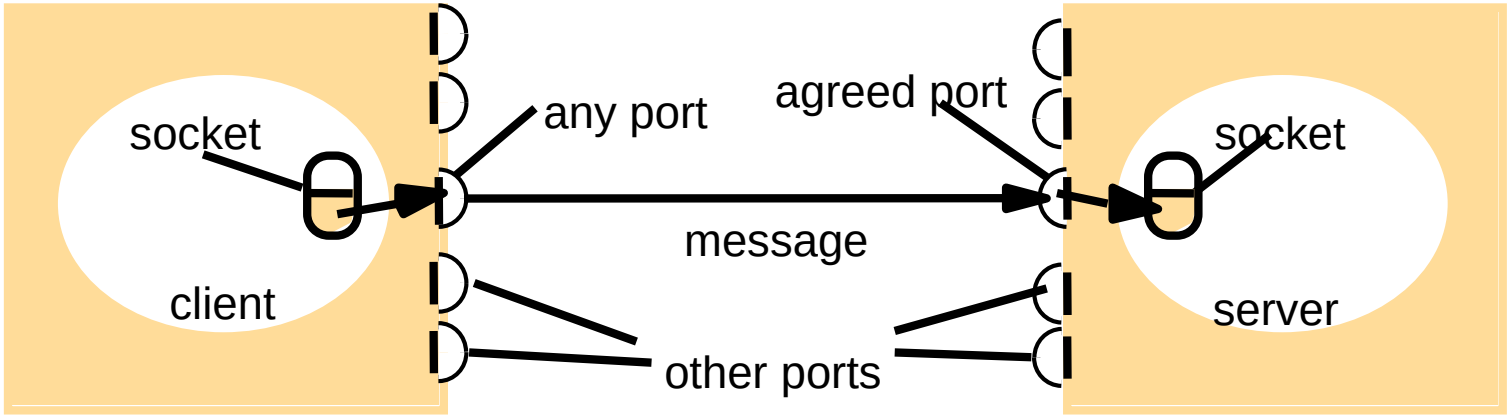
- To communicate
  - one process sends a message (a sequence of bytes) to a destination; and
  - another process at the destination receives the message.
- This activity involves the communication of data from the sending process to the receiving process
  - may involve the synchronization of the two processes.

The API for the Internet protocols

- Network API provides access to the network
  - for application development
- Core services are provided by the operating system that controls access to resources
- Various libraries may handle the rest
- Socket and ports concepts
  - forms the basis of modern network programming.

Ports and Sockets

- Closely associated with the hardware communication links between computers within a network
- Not themselves hardware elements.
- Abstract concepts that allow the programmer to make use of those communication links.
- For a process to receive messages, its socket must be bound to a local port and one of the Internet addresses of the computer on which it runs.



Ports

- The number of available ports will normally greatly exceed the number that is actually used.
- 1 to 1023 port range normally set aside for the use of specified standard services ('well-known' services).
  - E.g. port 80 is normally used by Web servers.
- Application programs wishing to use ports for non-standard services should use a range of 1024 to 65 535.

Protocol Name	**Port Number**	**Nature of Service**
Echo	7	The server echoes the data sent to it.
Daytime	13	Provides the ASCII representation of the current date and time on the server.
FTP-data	20	Transferring files
FTP	21	Sending FTP commands like PUT and GET
Telnet	23	Remote login and command line interaction
SMTP	25	E-mail
HTTP	80	The World Wide Web protocol
NNTP	119	Usenet

Sockets

- Dominant API for transport layer connectivity
- Normally, there are multiple clients wanting the same service at the same time.
  - E.g., multiple browsers wanting Web pages from the same server.
- The server needs some way of distinguishing between clients and keeping their dialogues separate.
  - This is achieved via the use of sockets.
- It is used to indicate one of the two end points of a communication link between two processes.
- Support different protocols and naming conventions
  - *It is not just for the IPs*

Sockets (I)

- When a client wishes to make connection to a server
  - it will create a socket at its end of the communication link.

- Upon receiving the client's initial request (on a particular port number),
  - the server will create a new socket at its end that will be dedicated to communication with that particular client.

- *One* hardware link to a server
  - may be associated with *many* ports
- *One* port
  - may be associated with many sockets.

Internet Protocol

- A set of protocols designed to handle the interconnection of a large number of local and wide-area networks that comprise the Internet
- IPv4 & IPv6 (network layer)
- IP-based protocols
  - TCP, UDP, etc.
- IP is connectionless
  - No state needs to be saved at each router
- Support multiple paths for data
  - packet delivery is not guaranteed

Internet Services, URLs and DNS

- Protocol
  - governs the communication that takes place between server and client
- Each end of the dialogue must know
  - what may/must be sent to the other,
  - the format in which it should be sent,
  - the sequence in which it must be sent (if sequence matters); and
  - for ‘open-ended’ dialogues
    - how the dialogue is to be terminated.

Uniform Resource Locator (URL)

- Unique identifier for any resource located on the Internet
- Structure:

  <protocol>://<hostname>[:port]/[<pathname>][/<filename>[<section>]]

- Port may be omitted – for a well-known protocol, and the default port will be assumed.
- If the file name is omitted, then the sever sends default file from the directory specified in the path name. This default file is commonly called *index.html* or *default.html*.
- The ‘section’ part of the URL (not often specified, a name preceded by ‘#’) indicates a named ‘anchor’ in an HTML document.

Domain Name System (DNS)

- A **domain name**, or **host name**, is the user-friendly equivalent of an IP address.
- The individual parts of a domain name do not correspond to the individual parts of an IP address.
  - Domain names do not always have four parts (as IPv4 addresses must have).
- The Domain Name provides a mapping between IP addresses and domain names and is held in a distributed database.
- The IP address systems and the DNS are governed by ICANN (Internet Corporation for Assigned Names and Numbers) – non-profit making organization.
- When a URL is submitted to a browser, the DNS converts the domain name part into its numeric IP equivalent.

Multicast Communication

Multicast communication

- The basic idea is to disseminate information from **one** sender to **multiple** receivers.
- An operation that sends a single message from one process to each of the members of a group of processes
  - usually in such a way that the membership of the group is transparent to the sender.
- Can span multiple physical networks
- There is a range of possibilities in the desired behaviour of a multicast.
- The simplest multicast protocol provides no guarantees about message delivery or ordering.

Multicast

- One-to-many or many-to-many distribution
- In computer networking
  - Group communication where information is addressed to a group of destination computers simultaneously
- Group communication, either
  - *application layer multicast*
  - *network assisted multicast*
    - makes it possible for the source to efficiently send to the group in a single transmission
- Network assisted multicast
  - May be implemented at the Internet layer using IP multicast

Multicast: different ways

- Multicast communication can also be accomplished in other ways than setting up explicit communication paths:
  - Application-level multicasting
  - Flooding-based multicasting
  - Gossip-based data dissemination

Application-level multicasting

- Organize nodes of a distributed system into an overlay network and use that network to disseminate data:
  - Often as a tree, leading to unique paths
  - Alternatively, also mesh networks, requiring a form of routing

Flooding

- Process **P** simply sends a message **m** to each of its neighbors.
  - Each neighbor will forward that message, except to P, and only if it had not seen m before.
  - The more edges, the more expensive
- Flooding messages across the network is extremely robust, but
  - requires special attention if we want to avoid severe waste of resources as nodes may see messages multiple times.

Gossip-based data dissemination

- Gossip-based information dissemination provides ways for multicasting
  - Often less efficient
- Deploys **epidemic** protocols
- Epidemic protocols
  - The main goal of these protocols is to rapidly propagate information among a large collection of nodes using only local information.
  - There is no central component by which information dissemination is coordinated.

IP Multicast

- Built on top the Internet Protocol
- An implementation of multicast communication
- IP packets are addressed to computers
  - ports belong to the TCP and UDP levels
- IP multicast allows the sender to transmit a single IP packet to a set of computers that form a multicast group.

IP Multicast

- At the application programming level
  - IP multicast is available only via UDP.
  - An application program performs multicasts by sending UDP datagrams with multicast addresses and ordinary port numbers.
  - It can join a multicast group by making its socket join the group, enabling it to receive messages to the group

- At the IP level
  - A computer belongs to a multicast group when one or more of its processes has sockets that belong to that group.
  - When a multicast message arrives at a computer, copies are forwarded to all of the local sockets that have joined the specified multicast address and are bound to the specified port number.

IP Multicast

- Multicast addresses may be permanent or temporary.
- Permanent groups exist even when there are no members – their addresses are assigned by IANA
- The remainder of the multicast addresses are available for use by temporary groups, which must be created before use and cease to exist when all the members have left.
- When a temporary group is created, it requires a free multicast address to avoid accidental participation in an existing group.

Multicasting

- Broader than *unicast*
  - *(*one sender and one receiver, point-to-point communication) but
- Narrower and more targeted than broadcast communication.

- Sends data from one host to many different hosts, but not to everyone;
  - the data only goes to clients that have expressed an interest by joining a particular multicast group.

- Used for ‘public meetings’ on the Internet
  - a multicast socket sends a copy of the data to a location close to the parties that have declared an interest in the data;
  - the data is duplicated only when it reaches the local network serving the interested clients;
  - the data crosses the Internet only once.)

Multicasting (I)

- Most multicast data is audio or video or both (relatively large and robust against data loss).
- Multicast data is sent via UDP
  - unreliable – but can be as much as three times faster than data sent via connection-oriented TCP.
- A multicast group is specified by a class D IP address and by a standard UDP port number.
- Class D IP addresses are in the range 224.0.0.0 to 239.255.255.255, inclusive.
- The address 224.0.0.0 is reserved and should not be used (the address of the Network Time Protocol distributed service – assigned the name ***ntp.mcast.net***)

Multicasting (II)

- Better than broadcast where everyone receives traffic
  - And broadcasting is limited.
- With multicasting, data duplicated only where necessary
- Routers typically form trees to efficiently distribute information
- The IP multicast protocol is relying on a few routers that support IP multicast routing to determine the routes, but transmitting IP packets through other routers using standard IP addresses.

Multicast Group

- A set of Internet hosts that share a multicast address.
- Dynamic membership
  - Machine can join or leave at any time
- Any data sent to the multicast address is relayed to all the members of the group. Membership in a multicast group is open; hosts can enter or leave the group at any time.
- Groups can be either permanent or transient:
  - Permanent groups have assigned addresses that remain constant, whether or not there are any members in the group; typically given names
  - Most multicast groups are transient – exist only as long as they have members (in the 255 -- 238 range).

IP multicasting

- No restriction on number of hosts in a group
- Machine does not need to be a member to send messages
- Efficient: Packets are replicated only when necessary

IP Multicast Addresses

- Internet Assigned Numbers Authority
  - IPv4 addresses: http://www.iana.org/assignments/multicast-addresses/multicast-addresses.xml
  - IPv6 addresses: https://www.iana.org/assignments/ipv6-multicast-addresses/ipv6-multicast-addresses.xhtml
- Similar to ports – service-based allocation
- For multicast, we have:
  - 224.0.0.0: reserved
  - 224.0.0.1 (*all-hosts group*): contains all multicast-capable hosts on this subnet
  - 224.0.0.2 (*all-routers group*): contains all multicast routers on this subnet
  - 224.0.0.3: unassigned

IP Multicast Use

- Initially
  - Internet radio, NASA shuttle missions, collaborative gaming
- IPTV
  - Cable TV networks are moving to IP delivery
  - Multicast allows one stream of data to be sent to multiple subscribers using a single address

References

- Chapter 4: Coulouris, Dollimore and Kindberg, Distributed Systems: Concepts and Design, 5/E

- Chapter 4: Maarten van Steen, Andrew S. Tanenbaum, Distributed Systems, 3rd edition (2017)