

Distributed Systems

General:

- a system with multiple components located on different machines that communicate and coordinate actions in order to appear as a single coherent system to the end-user
- organizing resources via network with more latency, less bandwidth, and higher error rate
- compared to parallel systems
 - Distributed has multiple (distributed users) vs parallel designed for single user or user process
 - Parallel has less security issues
 - Distributed Systems: cooperative work environment, Parallel Systems: environment designed to provide the maximum parallelization and speed-up for a single task

Definitions:

- task: collection of resources configured to solve a particular problem
 - contains not only the open files and communication channels, but also the threads (a.k.a. processes)
 - a task is a factory, all of the means of production scattered across many assembly lines

Migrating:

- Migrating Computation
 - need to move resources (memory, files)
 - move state of interaction with resources
 - need to restore communication (other side need to know new location)
- Migrating File State
 - keep track of the essential file state within the process and be ready to recreate it
 - migration system guarantees that higher level file operations are atomic, operation opens the file, seeks as needed, performs a read or write, and then closes the file
 - maintain the same file state, but only reopen the file when needed
- Migrating Communication Sessions
 - need to reestablish and map global state, such as network sockets

- higher level network name
- edge cases, such as what happens when a process move mid-transmission

Networking:

- LAN: Local Area Network a homogenous network
 - to communicate broadcast (yell) and all stations hear the broadcast
 - station have station id or LAN Address and messages have sender and receiver id
 - * everyone can hear other stations messages but ignore it unless diagnostic or malicious
 - size is self-limiting
 - * longer wire the weaker the signal
 - * greater the distance through the air the weaker the signal
 - * network can be clogged with collision, can collapse with utilization as low as 30
- bridge
 - connect LANs
 - send message for station only to relevant LAN (hashtable)
 - bridge managed so if station changes LAN
 - bridges (through configuration) create a spanning tree of location to prevent cycles
 - * if bridge fails form a different tree to get around failure
 - can't create bridges for the whole planet
- IP Address
 - IP4
 - * Class A (huge): 8 bits(network) + 24 bits (host) address begin with 0
 - * Class B (big): 16 bits(network) + 16 bits (host) address begin with 10
 - * Class C (small): 24 bits(network) + 8 bits (host) address begin with 110
 - IP6
 - * Classless, first few bits to describe the network/host division
 - * 73.93.0.0/15
 - address with 15 network bits and 17 host bits

incomplete:

- Dynamic Host Configuration Protocol (DHCP) allows IP addresses to be assigned on a temporary or quasi-permanent basis
 - DHCP is great for clients – but not for servers: Servers need to have well-known addresses

sources:

- http://www.andrew.cmu.edu/course/15-440-f14/index/lecture_index.html
- <https://blog.stackpath.com/distributed-system/>