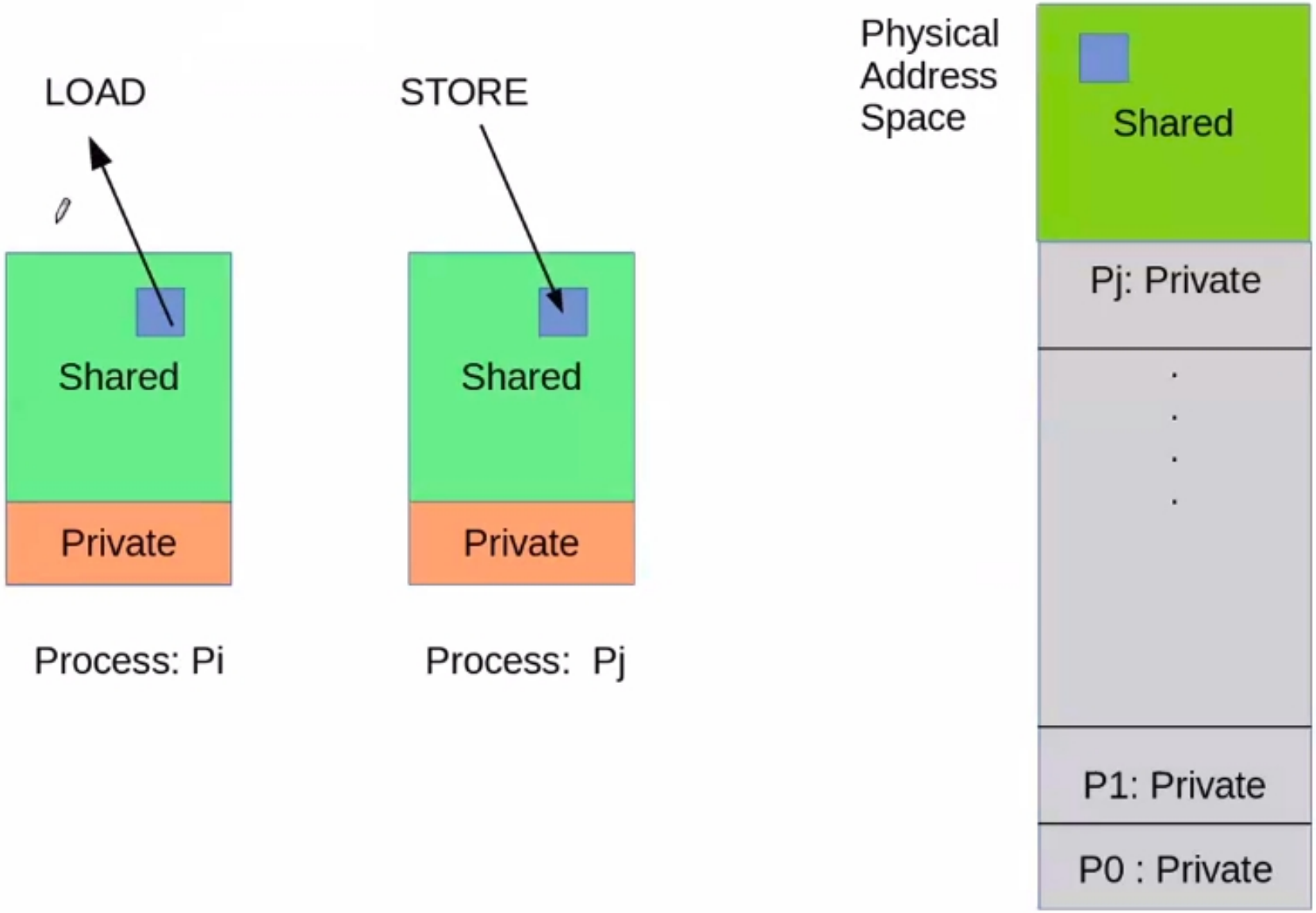
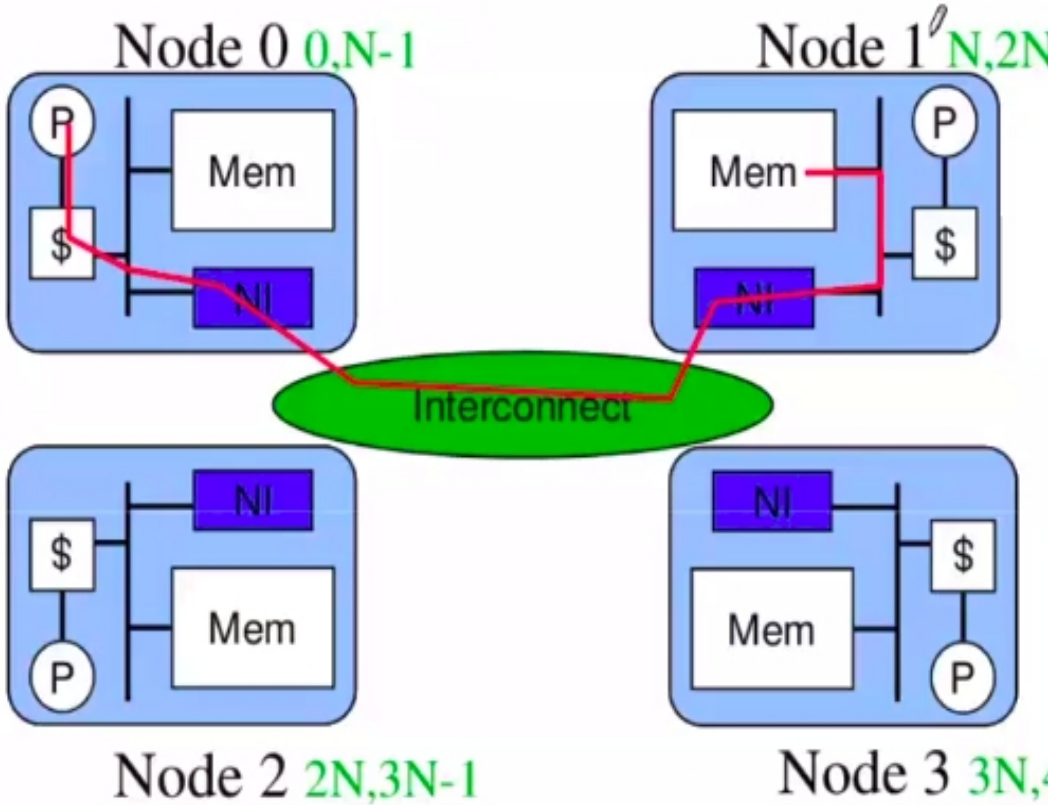
Shared Address Space, Interconnects, Message Passing

Shared Address Space



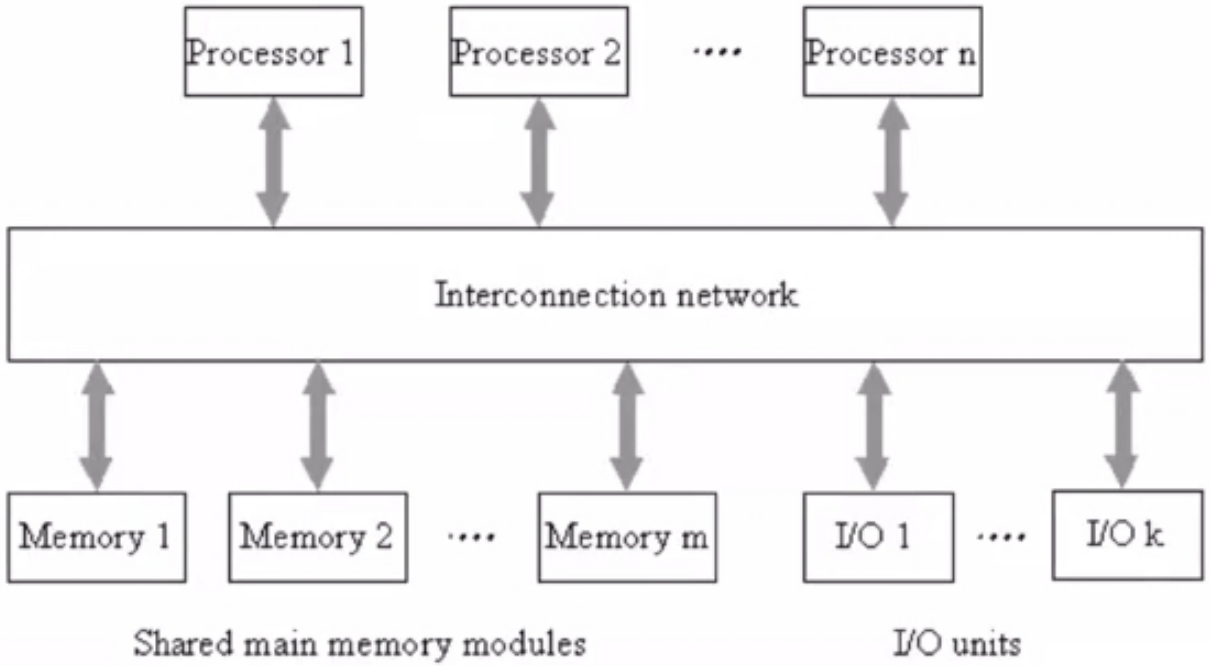
Page Mapping in Shared Memory MP



Keep private data and frequently used shared data on same node as computation

Load by Processor 0 to address N+3 goes to Node 1

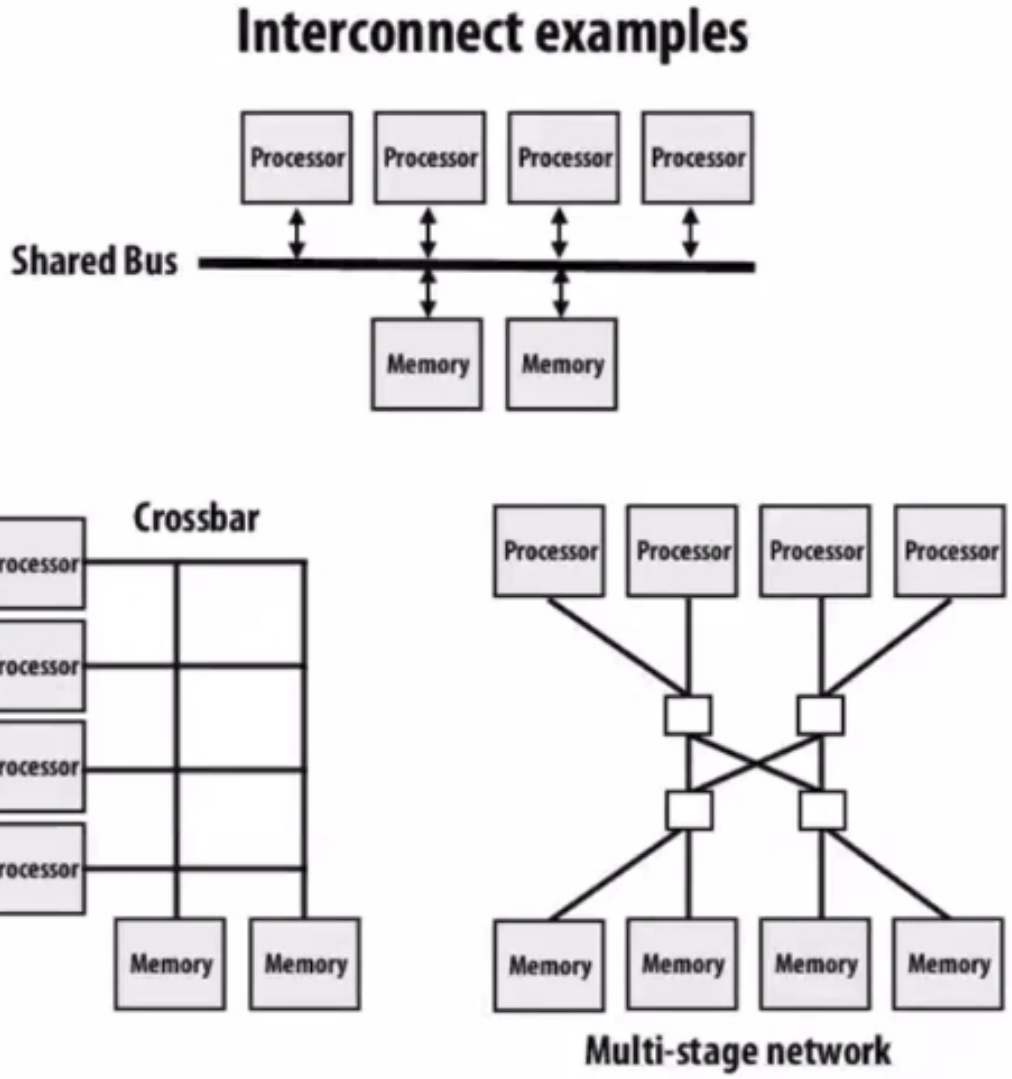
Shared Address Space



Shared Address Space ..

- Communication hardware for shared memory is a natural extension of the memory system
- It has several memory module + I/O Controllers
- Memory capacity can be increased by adding memory modules
- Processing speed can be improved by adding more processors or having faster processors
- Throughput/Processing Capacity can be increased by running many instances of a parallel program or running multiple threads of an application on these multiple processors
- Memory access
  - Single-bank: sequential=>contention
  - Multi-bank+interleaved 

Interconnect types



Interconnect types

- Crossbar
  - Adding processor/memory requires extending theswitch
  - Other structure same
  - Less scalable. Up to small number of cores
  - Therefore use multi-stage
- Multi-stage
  - +ve = cost increases more slowly with the ports
  - +ve: It has ability to access all memory directly from each processor
    - This allows any process or to run any process or handle any I/O event
    - Data structures could be shared within the OS
  - -ve: increased latency
  - -ve: Decreased bandwidth (per port if all used at once)

Interconnect types

- Bus
  - When processor, cache, MMU could fit on a single board or few boards
  - It was organised around central memorybus
  - +ve:bus access allows any process or to access any location
  - Same access latency for all processors frommemory
  - Called Symmetric MultiProcessor (SMP)
  - -ve: limiting factors: number of processors that can be connected as aggregate bandwidth decreases
  - But caches helped in solving memory access latency
  - However,cache consistency maintainance became an issue

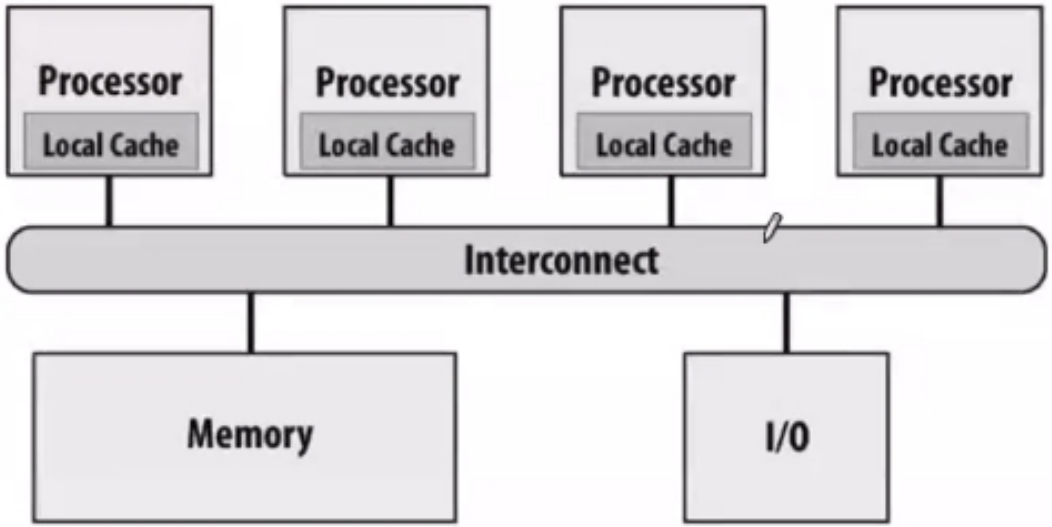
Building scalable Shared Memory Machines

- Basic processor component is well suited for computation task
- Problem exists with interconnect
- Bus does not scale as it has fixed aggregate bandwidth
- Crossbar does not scale well because its cost increases as the square of the number of ports
- Other scalable interconnection networks exist such that aggregate bandwidth increases+cost is not much
  - But we need to be careful about the resulting increase in latency, as processors must not stall for memory access

Scalable Network

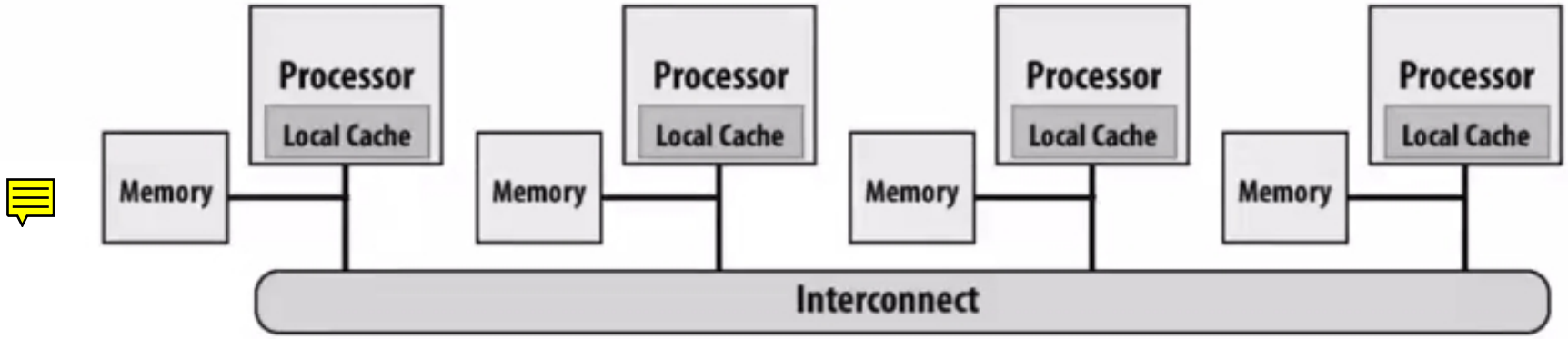
- Option-1: Scalable memory system having uniform memory access (UMA)
  - Ex:dance-hall
  - But each access requires round-trip network latency + a large bandwidth must be provided to each processor
- Option-2: NUMA
  - Each processor has its own local memory and can access global/remote memory when required
  - Local = Fast
  - Remote = Slow
  - Therefore NUMA
  - Bandwidth demand reduces

Scalable network



Dance-hall organisation

NUMA



Scalable network

- NUMA architecture
- Local memory = data moved closer to processor that accesses it
- Memory storage hierarchy = allows data to be migrated closer to accessing processor
- Expressing communication in terms of the storage = allows shared data to be migrated closer to accessing cores
- Migration + Replication of data across general purpose interconnect presents a unique set of challenges