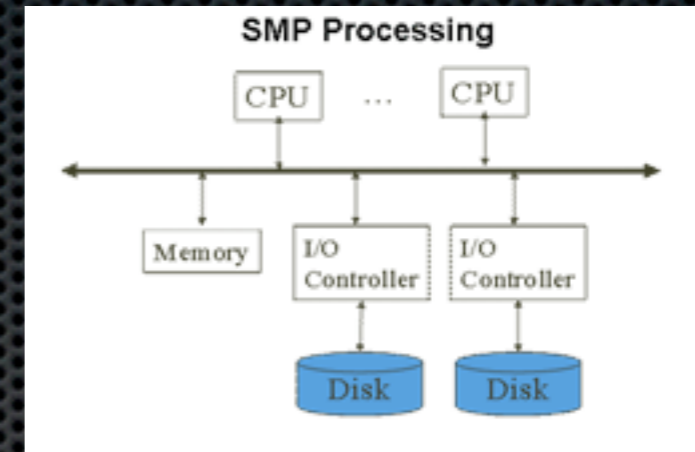


# SMP vs MPP

Bentley Wood and Brian Price

# What is **SMP**?

## Symmetric Multiprocessing



- ✦ Two or More Processors
- ✦ Connected to a Single Shared Main Memory
- ✦ Typically used for up to 8 processors
- ✦ One of the earliest multiprocessor architectures
- ✦ Additional CPUs can be added to improve performance and handle increased loads
- ✦ Also can share disk I/O resources.

# How Does SMP Work?

- ✦ No matter where a task's data is located in memory, any task can be run on any processor.
- ✦ Multiple CPUs attached to a bus
- ✦ CPUs communicate through bus to Main Memory
- ✦ The OS must support SMP

Windows NT & UNIX  
support SMP

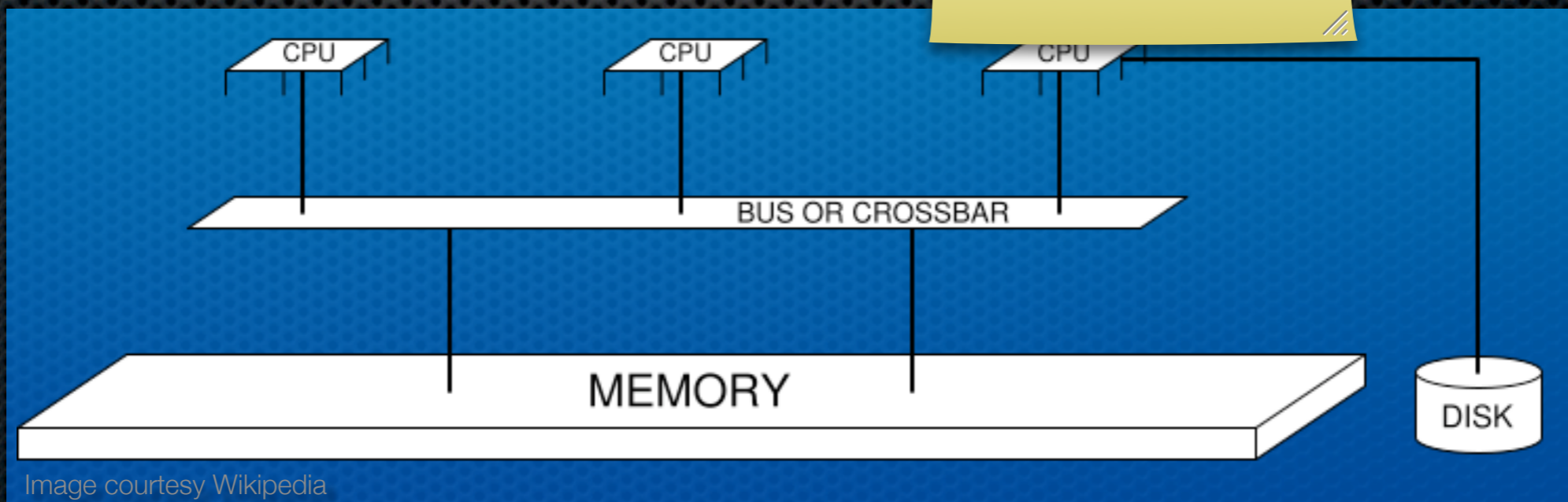


Image courtesy Wikipedia

# Requirements for SMP

- ✦ Hardware must support multiple processors
- ✦ Dual Core CPUs can be used
- ✦ The OS must support SMP
- ✦ Applications can use SMP, but must be multithreaded

Hypertreading is not true dual core and may have issues with SMP

# Requirements cont...

SMP or Node Machine Size	Machine Configuration
Small	16 or less CPUs, non-dual core 18GB RAM or less 3 dedicated I/O fiber channels or less 8GB or less "RAM cache" on disk units <300MB per second throughput on each I/O channel
Medium	16 to 24 CPUs, non-dual core 18 to 48GB RAM 5 to 8 dedicated I/O fiber channels 12 to 24GB RAM cache on disk units 400MB to 2GB per second throughput on each I/O channel
Large	24 to 64 CPUs, dual-core or not 48 to 128GB RAM (or more) 12 to 64 dedicated I/O fiber channels 24 to 168GB RAM cache on disk units 2GB per second or better throughput on each I/O channel

From B-Eye Network

# Advantages of SMP

- ✦ Running programs not designed for SMP see a slight performance increase.
- ✦ Hardware interrupts are sent to an idle CPU by the kernel (no need to context switch)
- ✦ Any task can run on any CPU
- ✦ Large Scalability
- ✦ RAM & CPUs Must be same make and size
- ✦ Extremely Fast

# Disadvantages of SMP

Each CPU adds less performance gain than the previous one; until eventually no performance is gained.

Cause is resource contention.

- ✦ Programs must be multithreaded to get the most out of SMP
  - ✦ Most games and word processing applications will not get much out of SMP
- ✦ Serialized RAM Access and Cache Coherency causes system lag
- ✦ Loss in hardware efficiency as the number of jobs being processed increases
- ✦ Limit to number of processes
- ✦ Once 32 CPUs are reached, require 1.3x RAM per CPU

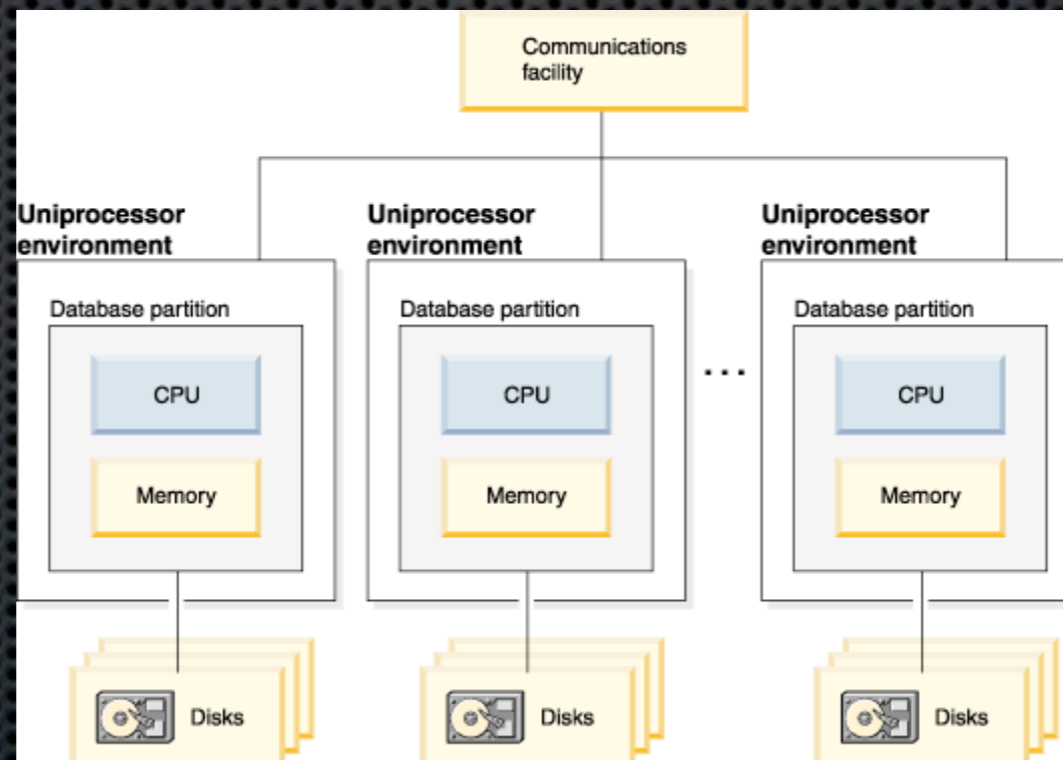
# MPP

## Massive Parallel Processing

### What is MPP?

[A] term used in computer architecture to refer to a computer system with many independent arithmetic units or entire microprocessors, that run in parallel. The term massive connotes hundreds if not thousands of such units.

- Via Wikipedia



# How does it work?

Each system in a MPP is just that, a separate system. MPP takes SMP systems and distributes them, attaching them to one main control unit via hi-speed fibre

- In an MPP system, each CPU contains its own memory and copy of the operating system and application. Each subsystem communicates with the others via a high-speed interconnect. In order to use MPP effectively, an information processing problem must be breakable into pieces that can all be solved simultaneously. In scientific environments, certain simulations and mathematical problems can be split apart and each part processed at the same time. In the business world, a parallel data query (PDQ) divides a large database into pieces. For example, 26 CPUs could be used to perform a sequential search, each one searching one letter of the alphabet.

# How does it work?

Cont..

- ✦ To take advantage of more CPUs in an MPP system means that the specific problem has to be broken down further into more parallel groups. However, adding CPUs in an SMP system increases performance in a more general manner. Applications that support parallel operations (multithreading) immediately take advantage of SMP, but performance gains are available to all applications, simply because there are more processors. For example, four CPUs can be running four different applications.

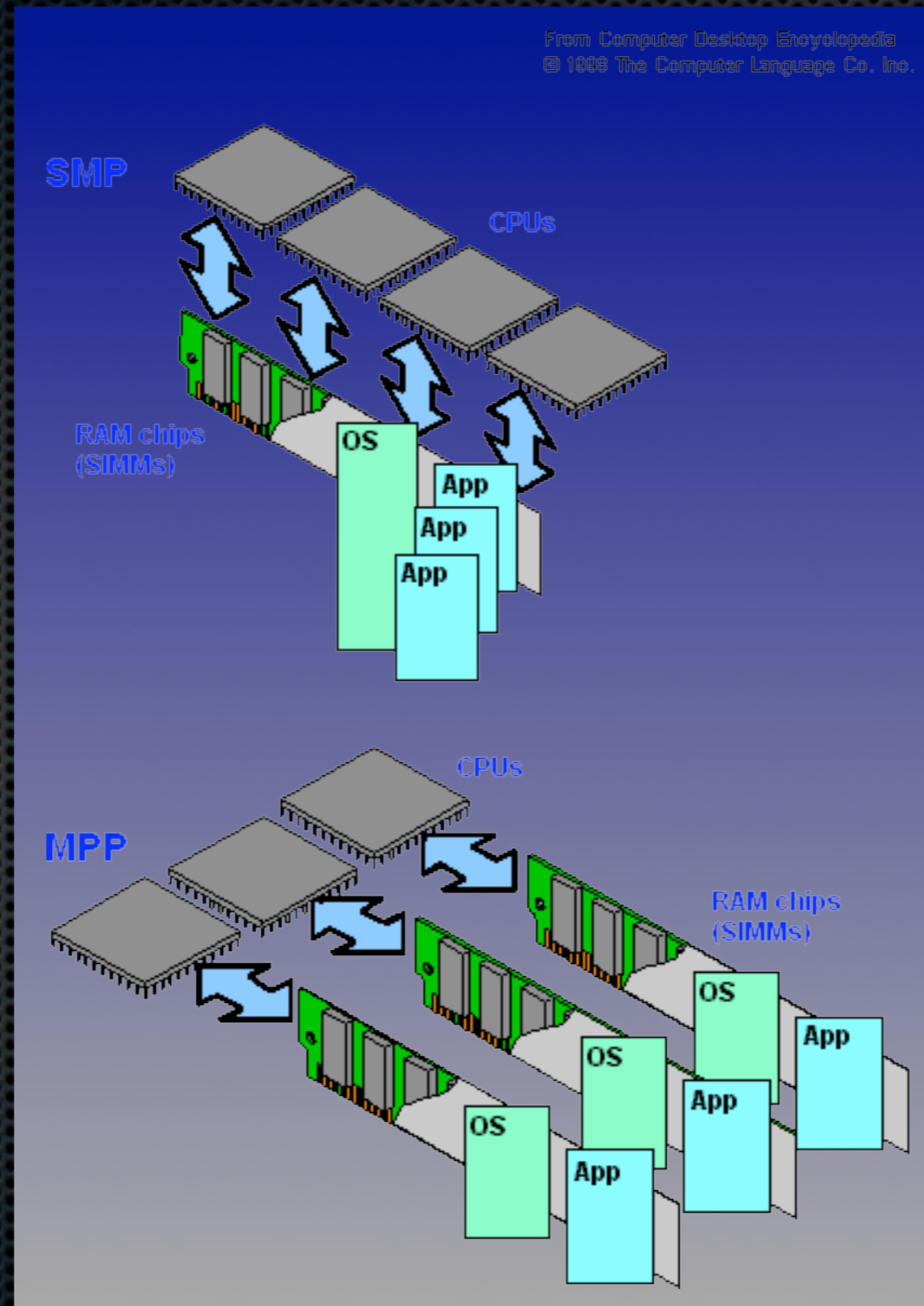
# Advantages

- ✦ Single Point of Management
- ✦ Everything is run in parallel
- ✦ Each system is a totally separate unit

# Disadvantages

- ✦ Adding another node means reseeding your application to each system
- ✦ COSTLY. Multiple SMP systems \$\$\$
  - ✦ Building Entire Systems
  - ✦ Establishing and upgrading the infrastructure

# SMP and MPP



# SMP and MPP Differences

## ✦ SMP

- ✦ Single Computer / OS
- ✦ Multiple CPUs
- ✦ Each CPU shares its RAM with the other CPUs; operates under one kernel
- ✦ Bus limits number of CPUs and Memory

## ✦ MPP

- ✦ Multiple computers
- ✦ May have one or more CPU in each computer
- ✦ Each “node” has its own CPU, I/O, RAM and Application
- ✦ Each node doesn't require big-iron (lots of CPU, RAM, I/O Disk)

# Relation to Databases ...

- ✦ Very Large Data Warehouses (and DWs in general) are not measured by “dead data”
- ✦ Instead, measured by:
  - Data loaded per hour
  - How much data which is queried and utilized in the database
  - And a mix of the previous two
- ✦ DW can be very resource intensive applications. The more power, the better.

# Amazon Example



- ✦ 61TB Query Data Warehouse
- ✦ 100,000 Queries / week
- ✦ 2 identical Data Warehouses
  - 16 Node RAC/Linux Cluster
  - Oracle 10g R1 RAC on Red Hat Enterprise Linux 3
  - 16 HP DL580s, each with 4x 3GHZ CPUs
- ✦ Metrics
  - Each holds 51TB raw data (growth rate doubles per year)
  - Each is 61TB total database size with 2TB indexes
  - 71TB total disk for each

# Sun Fire Oracle SMP

- ✦ A Server built for SMP:
  - ✦ 72 Dual-Core CPUs
  - ✦ 576GB RAM
  - ✦ 72 Hot swap I/O Channels
  - ✦ Up to 60TB DB storage (largest in 2003 was 30TB)



# The Question

- Name 3 differences between SMP and MPP.

# Resources

- ✦ Following resources used:

- ✦ [Answers.com](#)
- ✦ [DataSynapse.com](#)
- ✦ [DMReview.com](#)
- ✦ [Wikipedia](#)
- ✦ [Google](#)
- ✦ [B Eye](#)