INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392-I

SMR/101 - 22

SECOND COLLEGE ON MICROPROCESSORS: TECHNOLOGY AND APPLICATIONS IN PHYSICS

(18 April - 13 May 1983)

Multiple Processor Systems

- MIMD Systems
- Networks
- Software for Multiple Processor Systems

L.O. HERTZBERGER

NIKHEF-H
P.O. Box 41882
1009 DB Amsterdam
The Netherlands

(1)

MULTIPLE PROCESSOR SYSTEM

GENERAL CLASSIFICATION

NETWORKS

- WIDE AREA
- LOCAL AREA

SOFTWARE FOR MULTIPLE PROCESSORS

(2)

NETWORKS

A.S. TANENBAUM COMPUTER NETWORKS (PRENTICE-HALL)

CERN 82-12 WORKSHOP ON SOFTWARE IN HIGH ENERGY PHYSICS 201-246

MULTIPLE PROCESSOR SYSTEMS

B.A. BOWEN, R.J.A. BUHR THE LOGICAL DESIGN OF MULTIPLE-MICROPROCESSOR SYSTEMS (PRENTICE-HALL)

M. BEN-ARI PRINCIPLES OF CONCURRENT PROGRAMMING (PRENTICE-HALL)

REVIEW ARTICLES

IEEE TRANS. COMPUT VOL C-31/11 (NOV 1982)
VOL C-32/1 (JAN 1983)

IEEE COMPUTER VOL 15/10 (OCT 1982)

E.T. FATHI, M. KRIEGER MULTIPLE MICROPROCESSOR SYSTEMS: WHAT, WHY AND WHEN
IEEE COMPUTER VOL 16/3 (MARCH 1983) 23-34

MULTIPLE PROCESSOR SYSTEMS

(MULTIPLE COMPUTER SYSTEMS)

- **Definition :** SYSTEMS WITH > 1 PROCESSOR
- **Problem :** ALMOST ALL NOWADAYS SYSTEMS !!!

  WE HAVE TO BE MORE SPECIFIC :

  DEALING WITH VON NEUMANN ARCHITECTURES E.G.
  CPU'S CONSISTING OF :

  INSTRUCTION UNIT (CONTROL UNIT)
  DATA UNIT (FUNCTION UNIT)

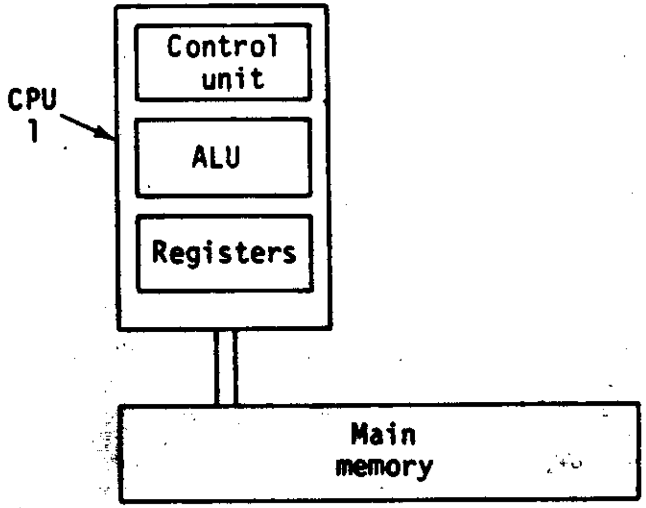
- **Definition: (Flynn)**
  BASED ON MULTIPLICITIES IN INSTRUCTION AND DATA STREAM

  SINGLE INSTRUCTION SINGLE DATA (SISD)
  SINGLE INSTRUCTION MULTIPLE DATA (SIMD)
  MULTIPLE INSTRUCTION SINGLE DATA (MISD)
  MULTIPLE INSTRUCTION MULTIPLE DATA (MIMD)

- A MIMD SYSTEM CAN BE MODELED AS A POOL OF PROCESSORS AND MEMORIES CONNECTED VIA SOME DATA TRANSPORT MEDIUM (MESSAGE SWITCH)

  FURTHER SUBDIVISION BASED ON COUPLING BETWEEN THE PROCESSORS E.G. ASPECTS AS :

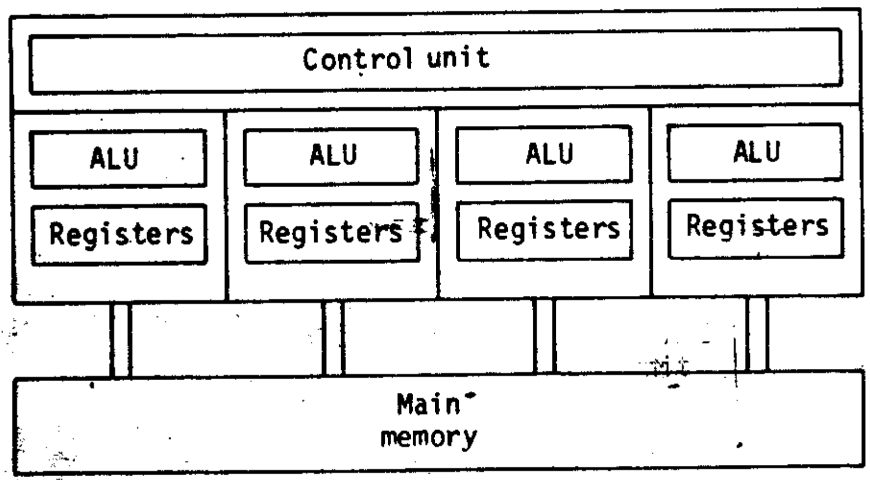
  - DISTANCE BETWEEN PROCESSORS
  - BANDWIDTH COMMUNICATION CHANNEL
  - DEGREE OF PROCESSOR COLLABORATION



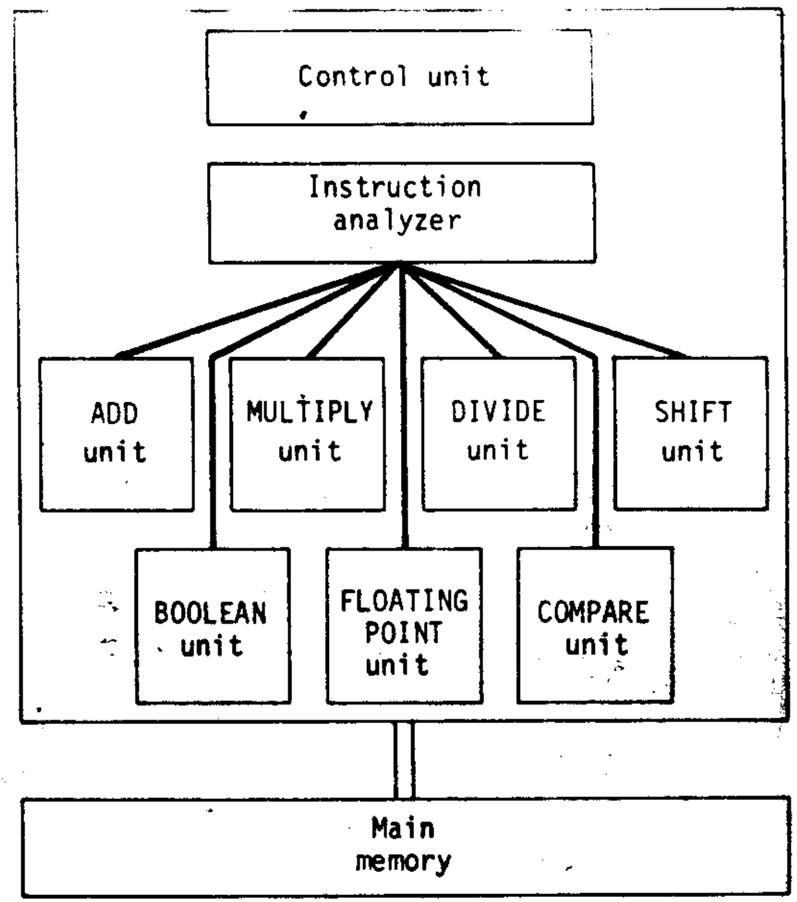
SISD

SINGLE INSTRUCTION (UNIT)

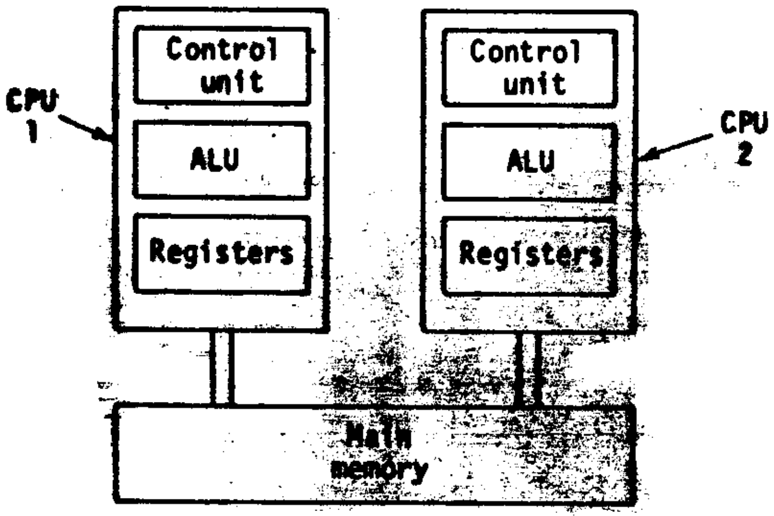
SINGLE DATA (PATH)



SIMD
SINGLE INSTRUCTION MULTIPLE DATA
(ARRAY)



MISD
MULTIPLE INSTRUCTION SINGLE DATA
(PIPELINED PROCESSOR)



MIMD

MULTIPLE INSTRUCTION MULTIPLE DATA

(MULTI PROCESSOR)

L… E…NWIDTH

- PARALLEL COUPLING (BUS) IN GENERAL LIMITED TO NOT MORE THAN SEVERAL METERS
- SERIAL CONNECTION
  - TWISTED PAIR ≤ 50 KBITS/S
  - COAX CABLE 1 – 10 MBITS/S
  - GLASS FIBRES 10 MBITS – ≥ 1 GBITS/S

PROCESSOR COLLABORATION

- PROCESSORS ALL WORK ON SAME TASK (COMMON MEMORY)
- PROCESSORS COMPLETELY INDEPENDENT AND SELF-CONTAINED (MEMORY, I/O, DISCS, TERMINALS ETC)

MORE DETAILS PHYSICAL COUPLING:

LOOSE COUPLING:

- GEOGRAPHICALLY DISPERSED
- LOW BANDWIDTH LINK (< 50 KBITS/S)
- RESOURCE SHARING
- PROCESSOR PERFORMS COMPLETE ALGORITHM

MODERATE COUPLING:

- MORE INTER-COMPUTER ACTIVITY
- HIGH SPEED SERIAL / PARALLEL BUSES
- PROCESSORS WORK CLOSE TOGETHER AND OFTEN HAVE SAME OPERATING SYSTEM
- OFTEN PARTICULAR CLASS OF PROBLEM
- POSSIBILITY FOR RECONFIGURATION

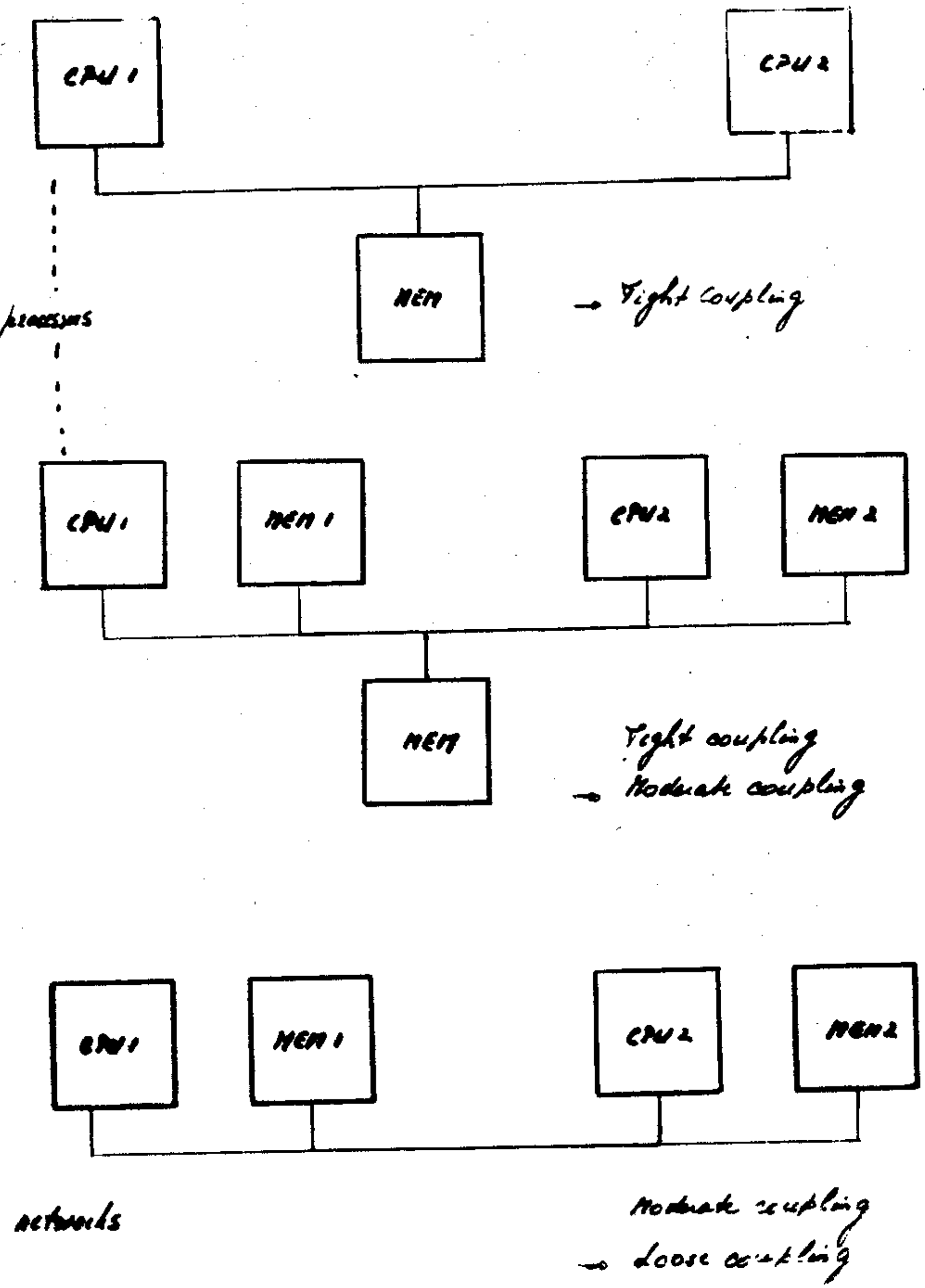
TIGHT COUPLING :

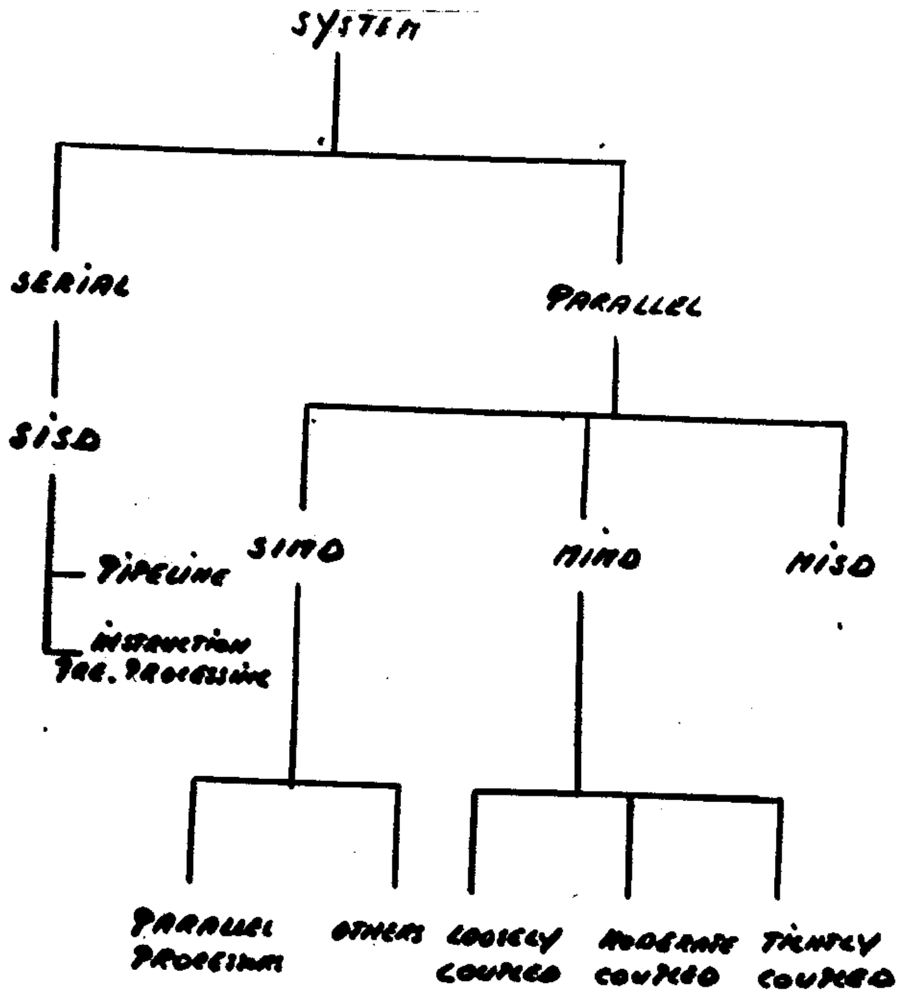
- PHYSICAL CLOSENESS PROCESSORS
- HIGH BANDWIDTH INTERCONNECTIONS
- GENERALLY MEMORY SHARING
- PROCESSORS WORK CLOSE TOGETHER ON SAME ALGORITHM OR TASK
- PROCESSORS HAVE SAME OPERATING SYSTEM multi processors

MULTI PROCESSOR SYSTEMS

BOUNDARIES BETWEEN MULTIPLE AND MULTI PROCESSOR SYSTEMS DIFFUSE

ESPECIALLY FOR MICROPROCESSOR BASED SYSTEMS !!!





FLYNN'S COMPUTER CLASSIFICATION

NETWORKS

WHY ???

- RESOURCE SHARING INDEPENDENT OF LOCATION
  (DATA/INFORMATION STORED IN MANY COMPUTERS)
- REDUNDANCY CONSEQUENTLY RELIABILITY
  (IF ONE COMPUTER GOES DOWN ANOTHER TAKES OVER)
- PRICE OF COMPUTERS VERSUS COMMUNICATION
  (COMPUTERS (µP-BASED) ARE CHEAP CAN BE PLACED EVERYWHERE)

TENDENCY:

- BECAUSE µP ESPECIALLY µP-BASED COMPUTERS ARE CHEAP IT IS ATTRACTIVE TO CONNECT THEM TO INCREASE COMPUTING POWER AND TO SHARE EXPENSIVE PERIPHERALS
  (DISCS, PRINTERS, TAPES ETC)

RESULT:

DISTRIBUTED COMPUTER SYSTEM

MULTI(PLE) PROCESSOR SYSTEM
LOCAL AREA NETWORK

TENDENCY:

BECAUSE COMMUNICATION BECOMES CHEAP IT IS ATTRACTIVE TO CONNECT COMPUTERS AT DIFFERENT PLACES

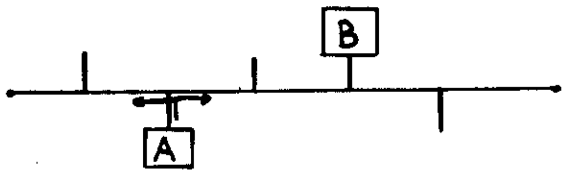
RESULT:

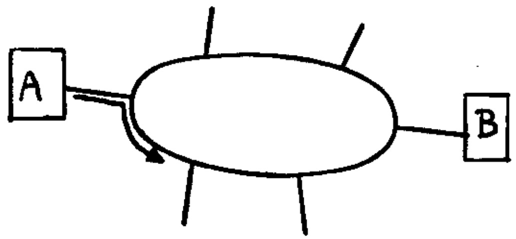
COMPUTER NETWORKS

WIDE AREA OR LONG HAUL NETWORKS !!

Local Area Network

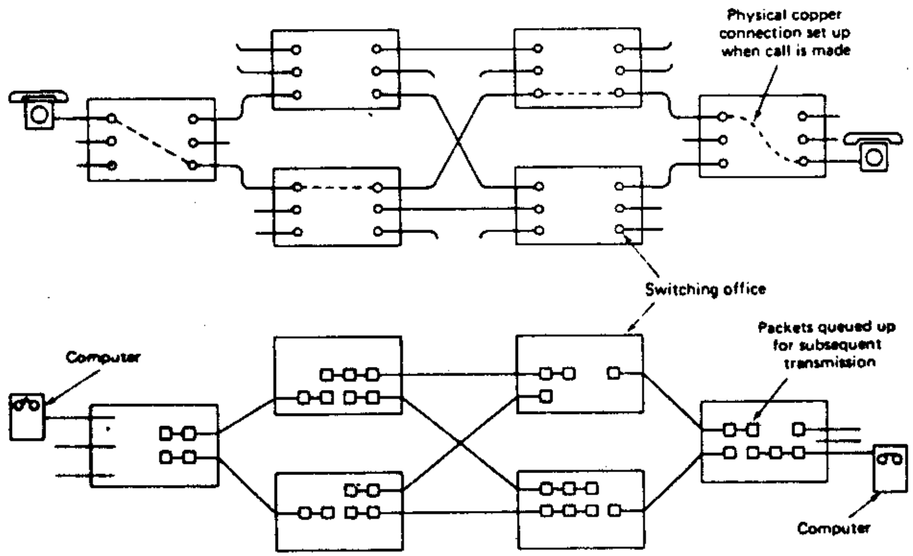
distance : ≤ 2 km
speed : 1 - 10 Mbaud
examples : - Ethernet (CSMA/CD)
- Cambridge ring (free slot)
- IBM ring (token)





CIRCUIT SWITCHING

: FIRST CONNECTION MADE
- AFTER THAT ANALOGUE LINE ALL THE TIME AVAILABLE



PACKET SWITCHING

- PACKETS ARE ROUTED OVER DIFFERENT LINES
- AT THE END STATION ASSEMBLED INTO MESSAGES

Network Services:

- terminal traffic
- file transfer
- remote job entry/manupilation
- electronic mail
- central data bases
- interactive graphics
-

public telephone:

1. A calls B
2. PTT makes connection (hardware)
3. private line during call
4. disconnect

leased lines:

2. fixed connection between A and B
3. permanent private line

PPS network:

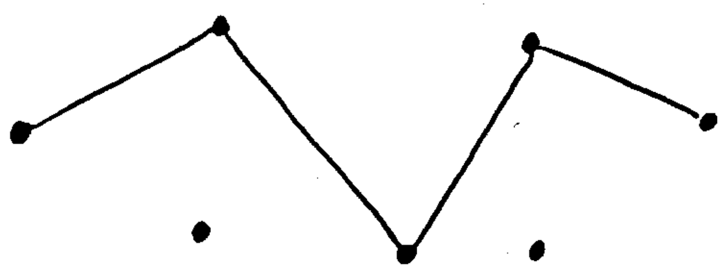
1. A calls B
2. PTT makes connection (soft/hardware)
3. private line during transfer of <u>one</u> packet of data.
4. disconnect

Wide Area Network

distance : ≤ 25000 km

speed : ≤ 64 kbaud

examples : - public telephone
- Leased Lines
- public packet switched network



HOP BY HOP
(ROUTING)

NETWORK STRUCTURE

- FOLLOW TERMINOLOGY FIRST MAJOR NETWORK ARPANET
- IN A NETWORK THERE ARE COMPUTERS RUNNING ALL TYPES OF USER SOFTWARE
  THESE MACHINES WILL BE CALLED HOSTS
- HOSTS ARE CONNECTED BY A COMMUNICATION SUBNET OR TRANSPORT SYSTEM OR TRANSMISSION SYSTEM
- JOB SUBNET CARRYING MESSAGES FROM HOST TO HOST
  RESULT IS THAT COMMUNICATION ASPECTS (SUBNET) ARE SEPERATED FROM APPLICATION ASPECTS (HOST)

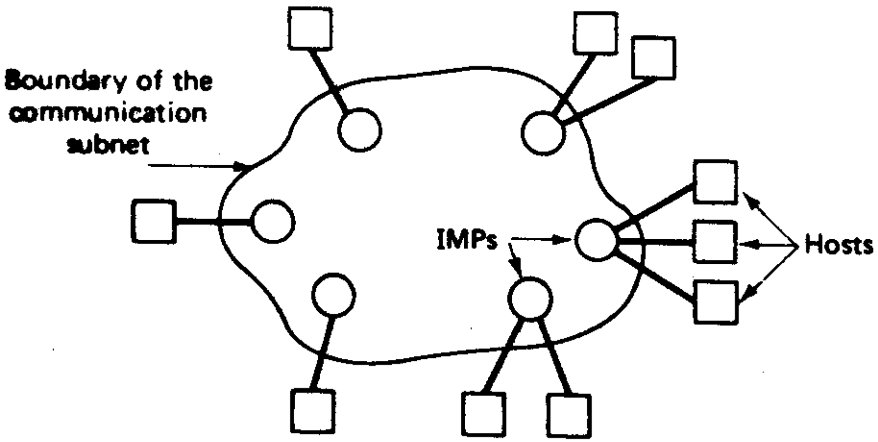
  COMMUNICATION SUBNET CONSISTS OUT OF TWO BASIC COMPONENTS :

- SWITCHING ELEMENT (INTERFACE MESSAGE PROCESSOR IMP, COMMUNICATION COMPUTER, SWITCH NODE)

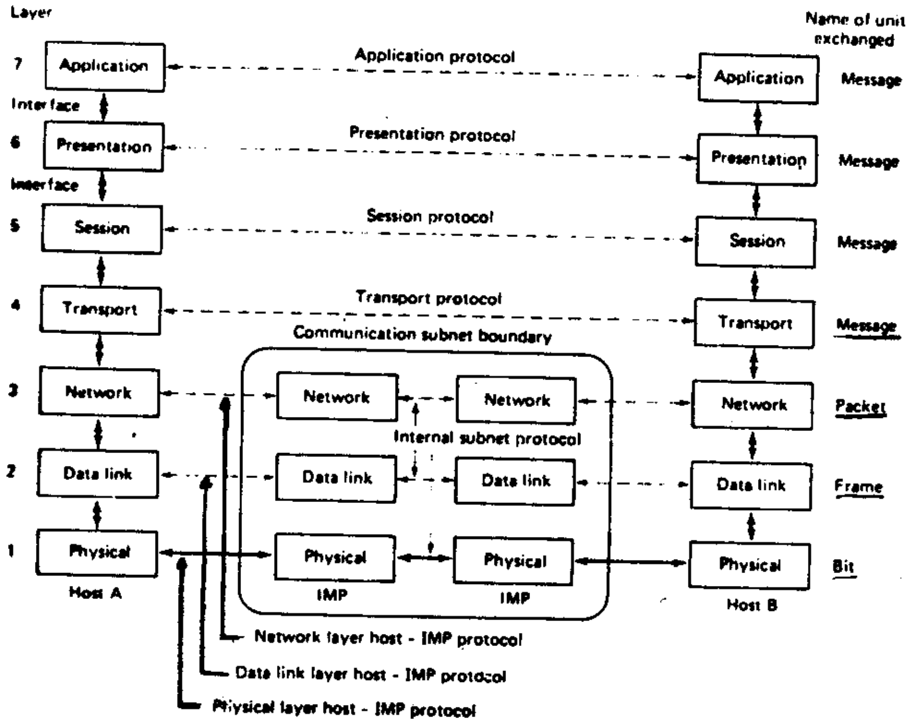
  TRANSMISSION LINE (CHANNEL, CIRCUIT)
  - TWISTED PAIR CABLE
  - COAXIAL CABLE
  - GLASS FIBRE

  BUT ALSO COMMUNICATION CHANNELS GROUND STATIONS AND SATELITES

INTERFACE MESSAGE PROCESSOR



ISO MODEL AND INTERFACE MESSAGE PROCESSOR (IMP)



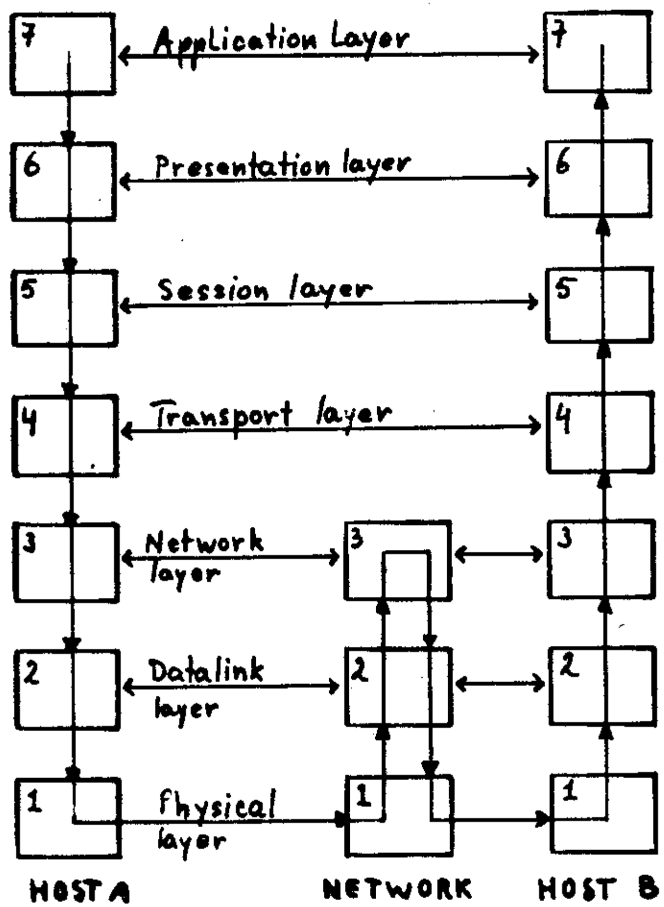
WHY ???

- TO REDUCE DESIGN COMPLEXITY OF NETWORKS
- NETWORK ORGANIZED AS NUMBER OF <u>LAYERS</u> OR <u>LEVELS</u>
- NUMBER OF LAYERS IN NETWORK DIFFER
- PURPOSE OF LAYER IS TO OFFER CERTAIN SERVICE TO NEXT HIGHER LAYER (SHIELDING OF DETAILS HOW LAYER IS IMPLEMENTED)
- LAYER n CARRIES CONVERSATION WITH LAYER n IN OTHER MACHINE
- <u>RULES</u> AND <u>CONVENTIONS</u> USED IN CONVERSATION ARE CALLED LAYER <u>n PROTOCOL</u>
- ENTITIES COMPRISING CORRESPONDING LAYERS ARE CALLED PEER PROCESSES (PEER PROCESSES COMMUNICATE VIA PROTOCOL)
- EXEPT FOR LOWEST LAYER NO DATA IS DIRECTLY TRANSFERRED BETWEEN LAYERS
- EACH LAYER PASSES DATA TOWARDS LOWER ONE

COMMUNICATION :

- PHYSICAL COMMUNICATION BETWEEN MACHINES AT LOWEST LEVEL
- VIRTUAL COMMUNICATION BETWEEN HIGHER LEVELS
- INTERFACE BETWEEN LAYER:
  - PRIMITIVE OPERATIONS
  - SERVICES

<u>OSI model of ISO</u>



LAYERS

PHYSICAL LAYER

RESPONSIBLE FOR TRANSMISSION SO QUESTIONS THAT HAVE TO BE SOLVED ARE :

- VOLTAGE FOR DIGITAL REPRESENTATION
- LENGTH OF SIGNALS (BANDWIDTH)
- UNIDIRECTIONAL / BIDIRECTONAL CHANNEL
- ALSO CONNECTOR TYPES

DATA LINK LAYER

RESPONSIBLE TO TRANSFORM RAW BIT STREAMS INTO RECOGNIZABLE FRAMES FREE OF TRANSMISSION ERRORS

- INPUT DATA IS BROKEN UP INTO DATA FRAMES (OFFICIAL DATA_LINK_SERVICE_DATA UNIT)
- IT RECOGNIZES FRAME BOUNDARIES
- PERFORMS ERROR DETECTION AND CORRECTION
- PERFORMS DATA BUFFERING
- IN BIDIRECTIONAL LINES IT TAKES CARE THAT TRANSMISSION A → B IS FINISHED BEFORE STARTING B → A

SUMMARIZE :

BREAK INPUT DATA INTO DATA FRAMES WHICH ARE SEQUENTIALLY TRANSMITTED, PROCESS ACKNOWLEDGEMENT FRAMES RECEIVED AND ASSEMBLE INTO OUTPUT DATA

SOMETIMES CALLED COMMUNICATION SUBNET LAYER

RESPONSIBLE FOR CONTROL OF OPERATION OF COMMUNICATION SUBNET

- DETERMINES HOW PACKETS (UNITS OF INFORMATION) ARE ROUTED IN SUBNET, WHICH IS THE MAIN CHARACTERISTIC OF THE IMP – HOST INTERFACE
- MAJOR ISSUE IS DIVISION LABOR BETWEEN IMP AND HOST (WHO IS RESPONSIBLE THAT PACKAGES ARRIVE AT DESTINATIONS, AND IN PROPER ORDER)
- TOO MANY PACKAGES WILL CAUSE CONGESTIONS, TRAFIC BOTTLENECKS

TRANSPORT LAYER

ALSO CALLED HOST TO HOST LAYER

- DATA FROM SESSION LAYER IS SPLIT UP IN PACKAGES
- CREATES A CONNECTION FOR EACH TRANSPORT (POSSIBLE MORE CONNECTIONS FOR ONE TRANSPORT)
- DETERMINES SERVICE PROVIDED TO SESSION LAYER (SOURCE TO DESTINATION)
- MULTIPLEXING MESSAGES INTO ONE CHANNEL

LAYERS 1 TO 3 ARE CHAINED 4 TO 7 END TO END

RESPONSIBLE TO PROVIDE <u>USER INTERFACE</u> INTO THE NETWORK

- USER NEGOCIATE FOR CONNECTION WITH ANOTHER COMPUTER
- CONNECTION BETWEEN USERS IS A SESSION
  THE OPERATION TO SET UP A SESSION IS CALLED A BINDING
- SESSION ALLOW USER TO LOG INTO REMOTE HOST OR TO TRANSPORT A FILE

PRESENTATION LAYER

RESPONSIBLE TO PERFORM FUNCTION FOR THE USER WHICH ARE REQUESTED SO OFTEN TO WARANT GENERAL SOLUTION

- FUNCTIONS ARE PROVIDED AS LIBRARY ROUTINES
- FUNCTION CAN BE TEXT COMPRESSION
  (USE CODE IN WHICH MOST FREQUENT OCCURING WORDS GET SMALLEST BIT REPRESENTATION)
- ALSO ENCRIPTION TO PROVIDE SECURITY

APPLICATION LAYER

RESPONSIBLE TO PERFORM FUNCTION CREATED BY USER

- DIVIDING APPLICATION OVER SEVERAL MACHINES
- BANKING AND AIRLINE RESERVATION ALGORITHMS
  (E.G. ACCESS TO OTHER ... DATA BASE)

WIDE AREA NETWORKS
(PTT'S — PPS NETWORKS)

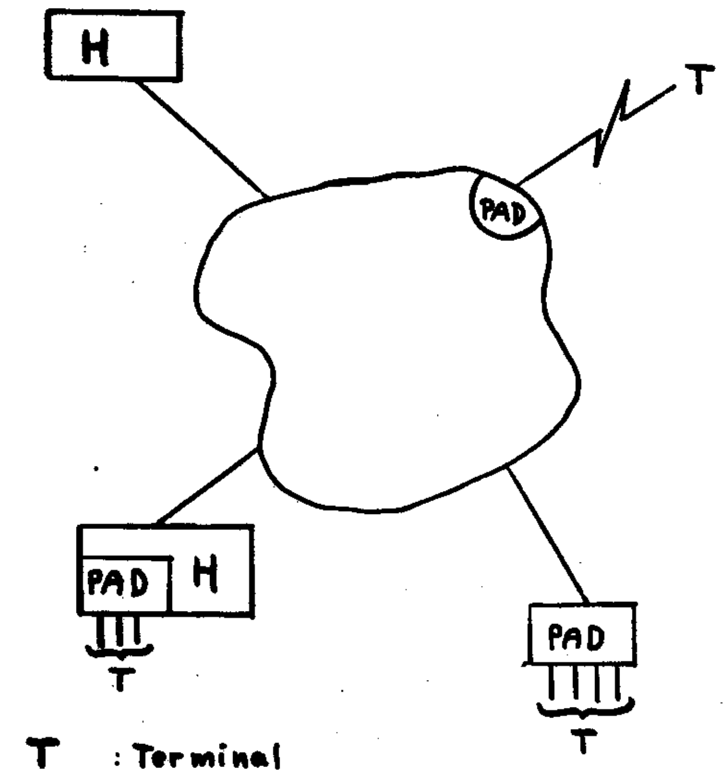
- STANDARD X.25 COVERS: PHYSICAL, DATALINK, NETWORK LAYERS
- HOST INTERFACE: <u>D</u>ATA <u>T</u>ERMINAL <u>E</u>QUIPMENT (DTE)
- IMP INTERFACE: <u>D</u>ATA <u>C</u>IRCUIT-TERMINATING <u>E</u>QUIPMENT (DCE)
- IMP IS CALLED <u>D</u>ATA <u>S</u>WITCHING <u>E</u>XCHANGE (DSE)

SOMETIMES WE ONLY WANT TO CONNECT TERMINALS TO X.25 NETWORK

<u>P</u>ACKET <u>A</u>SSEMBLER <u>D</u>ISASSEMBLER
THREE X'S

X.3 DEFINES PAD PARAMETERS
X.28 DEFINES TERMINAL - PAD INTERFACE
X.29 DEFINES PAD - HOST (DTE) INTERFACE

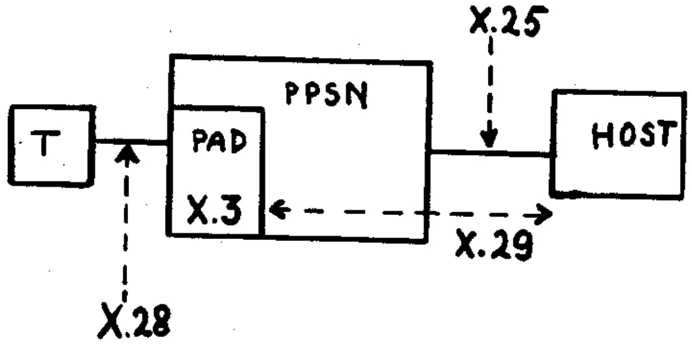
X25 PPS NETWORK



T : Terminal
H : Host
PAD : Packet Assembler/Dissembler

X SERIES PROTOCOLS

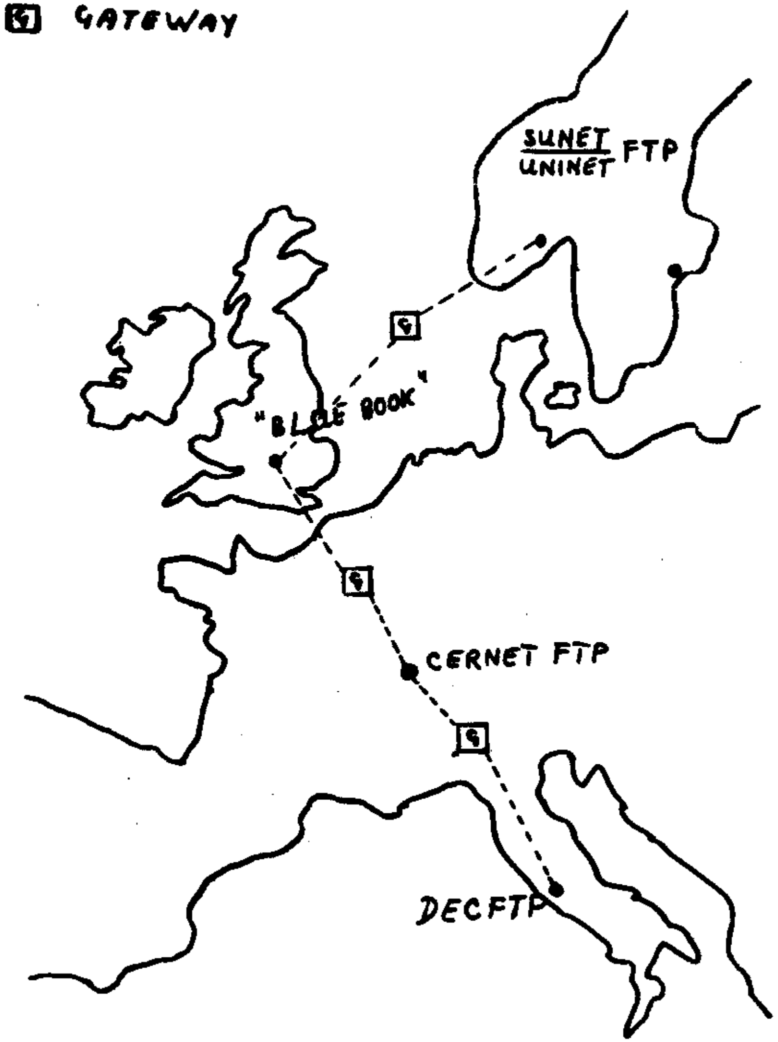
- terminal traffic



- file transport : no standard

FTP

G GATEWAY

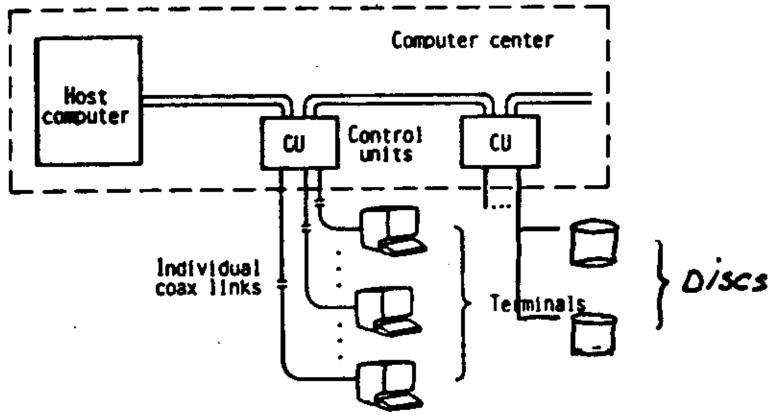


COMPUTER CENTRE (OLD)

- TERMINALS NO LOCAL INTELLIGENCE
- TIGHTLY COUPLED VIA I/O CHANNELS
- CONTROL UNITS (CU) PERFORM CONTROL FUNCTION

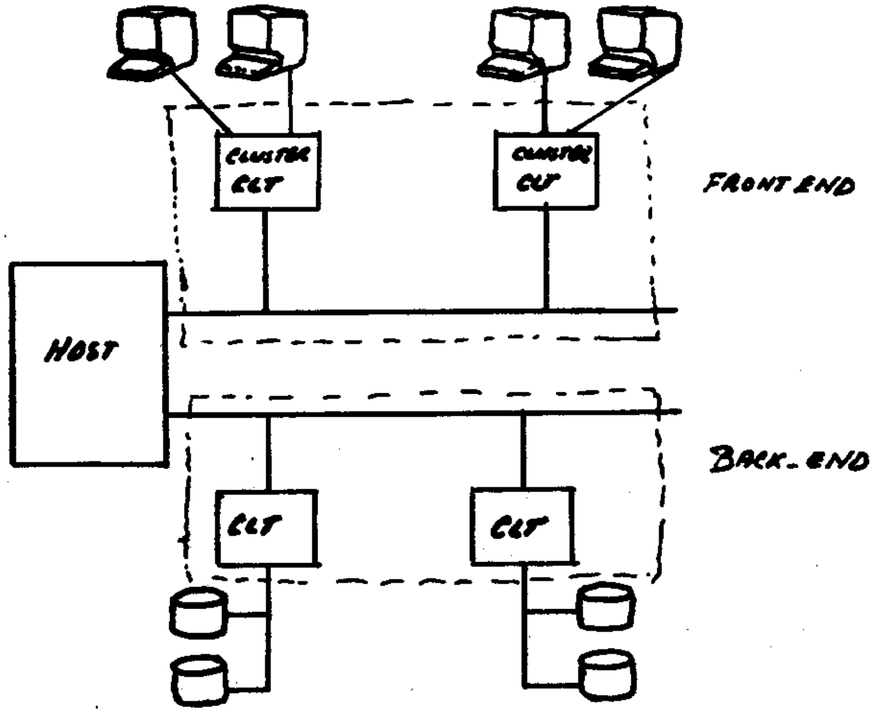
PROBLEMS / SHORTCOMINGS

- CONTROL UNITS HAVE TO BE IN CENTRE
- INDIVIDUAL CABLES TO ALL TERMINALS
- TERMINAL SWITCHING FROM ONE COMPUTER TO OTHER DIFFICULT



CU'S COUPLED VIA PARALLEL CHANNEL

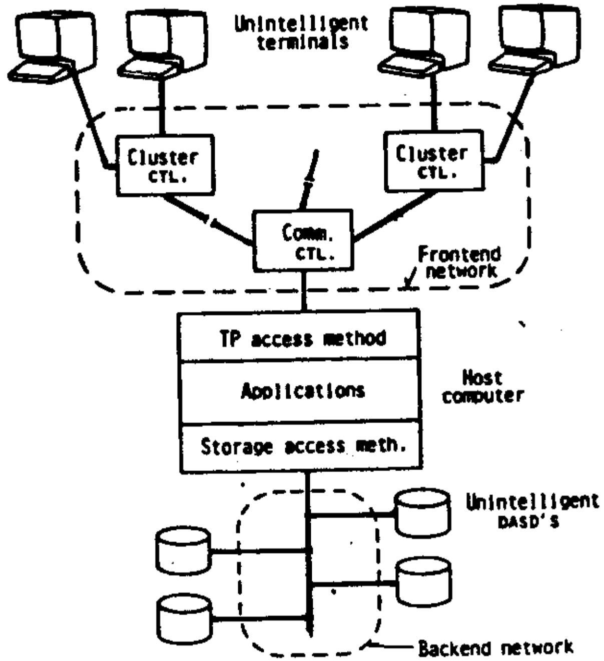
- DIFFERENT PARALLEL BUSSES FOR TERMINALS AND STORAGE DEVICES
- ONLY TERMINAL CONNECTIONS SERIAL



- IF WE COULD MAKE PARALLEL CONNECTION SERIAL TIGHT COUPLING → LOOSE COUPLING (NETWORK)
- THE CONNECTION HAS TO BE FAST ESPECIALLY FOR BACK-END (32 ms/16 BIT)
- POSSIBLE WITH LOCAL AREA NETWORKS 1 — 10 Mbit/s

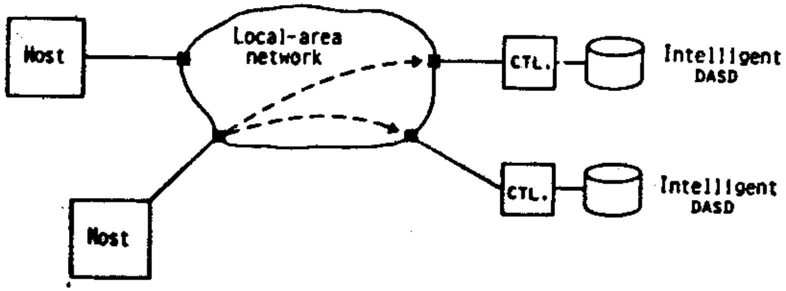
CL... CENTRE (NEW)

- INTELLIGENT CLUSTER CONTROLLERS CONNECTED INTO FRONT-END NETWORK (CONNECTIONS SERIAL)
- MASS STORAGE DEVICES CONNECTED INTO BACK-END NETWORK (SERIAL CONNECTIONS)



LOCAL AREA BACK-END NETWORK

- DIRECT ACCESS STORAGE DEVICES (DASD) NEED INTELLIGENCE TO ACCESS NETWORK
- EACH HOST CAN COMMUNICATE WITH EACH DASD

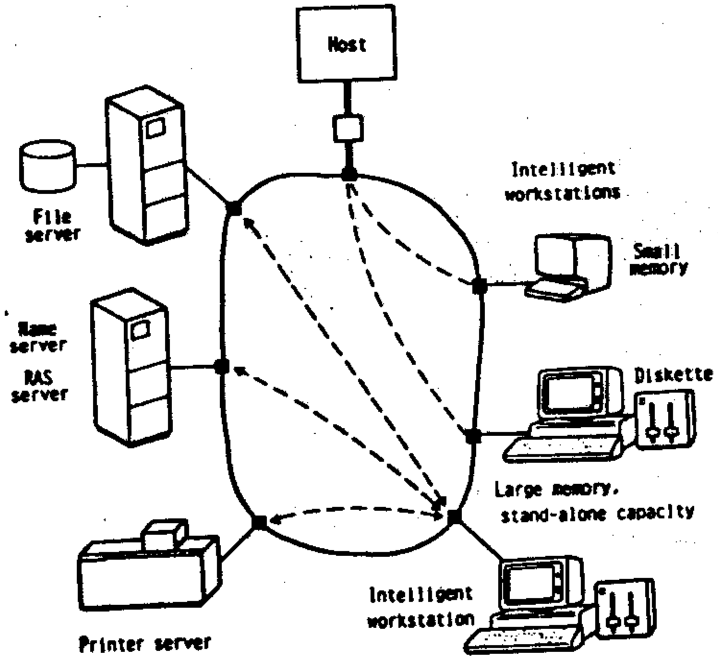


CLIENT-SERVER MODEL

- WORKSTATIONS (CPU, MEMORY, TERMINAL, SOMETIMES SMALL DISC) ARE CLIENTS

COMMUNICATE WITH

- INTELLIGENT PRINTERS, DISCS (FILE SERVER, DATA BASE) WHICH ARE CALLED SERVERS



LOCAL AREA NETWORKS

WIDE AREA NETWORK:

- SPEED TELEPHONE LINE SMALL (≤ 50 Kbit/s) COMPARED WITH SPEED NETWORK NODE
- CONSEQUENCE NODE CAN DO THE SWITCHING AND ROUTING (STORE AND FORWARD, HOP BY HOP)
- HAVE TO DEAL WITH PUBLIC TELEPHONE SERVICES (TRANSMISSION LINES ARE EXPENSIVE SO SHARING WITH OTHERS)

LOCAL AREA NETWORK

- PROCESSORS NOT ANY MORE FAST COMPARED WITH SPEED TRANSMISSION LINES (≥ 1 Mbit/s)
- CONSEQUENCE HOP BY HOP REPLACED END TO END
- NO LIMITATION OF PUBLIC TELEPHONE SERVICES

CONSEQUENCE

- LOCAL AREA NETWORKS REPLACE FRONT-END AND BACK-END NETWORKS BY ONE DISTRIBUTED SYSTEM.

L[illegible] NETW[illegible]

- TERMINALS
- MASS STORAGE DEVICES (DISC, [illegible] TAPES)
- PRINTERS & PLOTTERS
- COPIERS & TELECOPIERS
- MONITORING & CONTROL EQUIPMENT
- BRIDGES & GATEWAYS TO OTHER NETWORKS

LOCAL AREA NETWORK PROVIDE SERVICES AS:

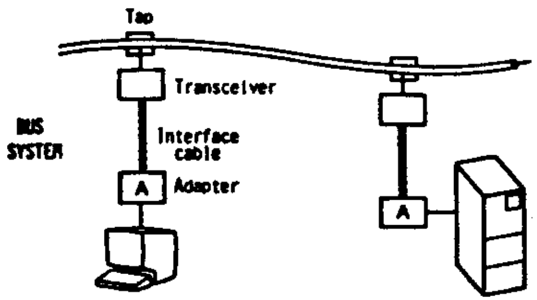
- FILE TRANSPORT AND ACCESS (OLD O.S)
- REMOTE DATA BASE ACCESS
- GRAPHICS, WORD PROCESSING, ELECTRONIC MAIL
- PROCESS CONTROL
- DIGITIZED VOICE TRANSMISSION

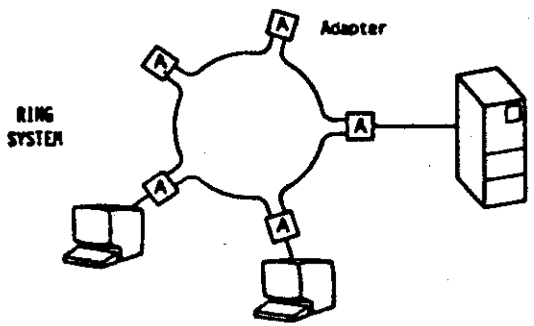
FUNCTIONAL REQUIREMENTS NETWORK

- SUPPORT <u>AT LEAST</u> 200 DATA DEVICES
- LENGTH <u>AT LEAST</u> 2000 METERS
- DATA RATE IN RANGE 2 - 20 Mbits/s
- ERROR CAUSED BY ADDITION OR DELETION OF DEVICES <u>AT LEAST</u> < 1 SECOND
- MEAN TIME BETWEEN UNDETECTED ERRORS ONE YEAR
- FAULTY DEVICE SHOULD NOT HANG-UP REST NETWORK

TECHNICAL APPROACHES

- BUS
- RING





LOCAL AREA NETWORK & ISO REFERENCE MODEL

PROPOSAL FOR STANDARDIZATION

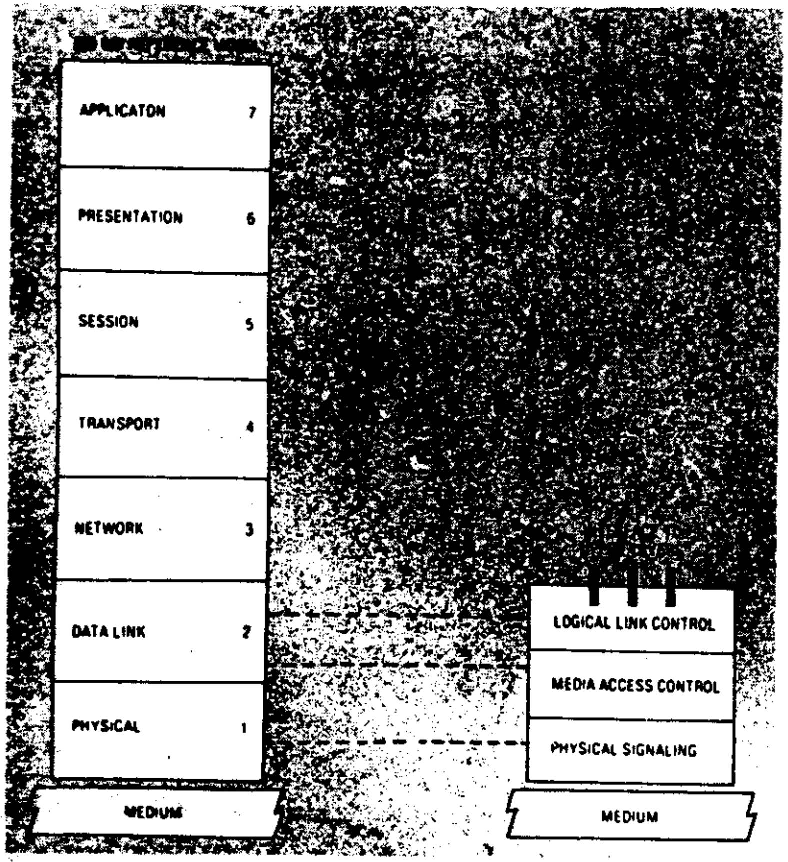
- IEEE 802
- ONLY LEVEL 1 & 2 INVOLVED
- MAPPING OF TWO LAYERS OF ISO MODEL INTO THREE LAYERS OF 802 STANDARD

IN PROPOSAL STANDARIZATION

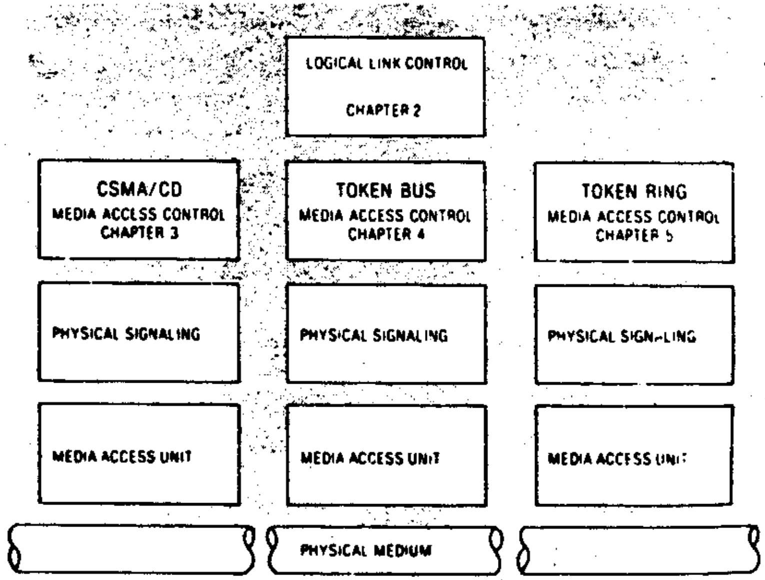
- LAYER SPECIFICATION eg. A WAY HOW LAYER IS IMPLEMENTED
- SERVICE SPECIFICATION eg. SERVICE PROVIDED TOWARDS NEXT HIGHER LAYER

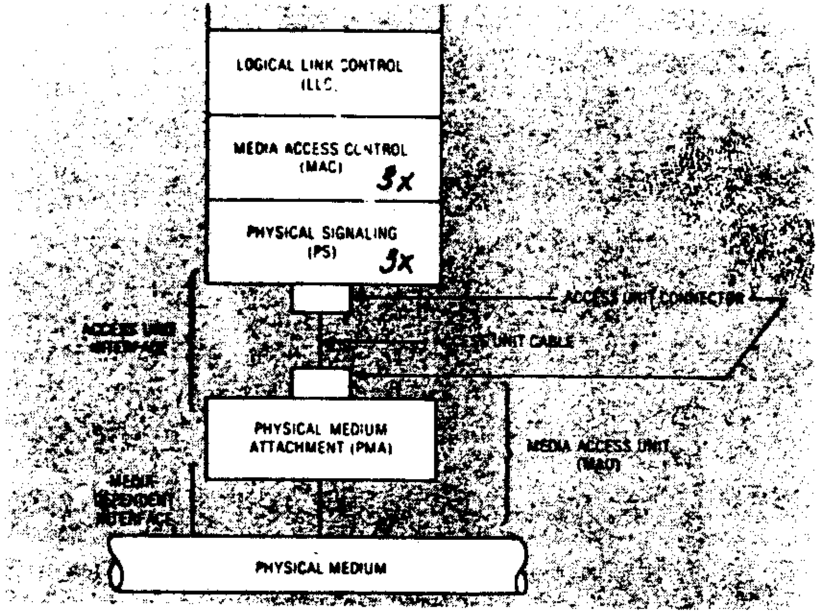
MEDIUM ACCESS CONTROL

- DESCRIBES THREE PARALLEL LAYERS BECAUSE THE DIFFERENCES IN ACCESS
    - CARIER_SENSE_MULTIPLE_ACCESS WITH COLLISION DETECT (CMSA/CD)
    - TOKEN BUS
    - TOKEN RING



802 PROPOSAL FOR CSMA,
TOKEN BUS
TOKEN RING





ACCESS:

CSMA / CD
TOKEN BUS
TOKEN RING

THERE EXIST OTHER ACCESS METODS

ACCESS METODS

BASIC PHILOSOPHY

- EITHER ACCESS IS CONTROLLED AND NO COLLISION WILL OCCUR OR IT IS RANDOM AND COLLISIONS TAKE PLACE
- IN LAST CASE WE NEED RECOVERING METHOD

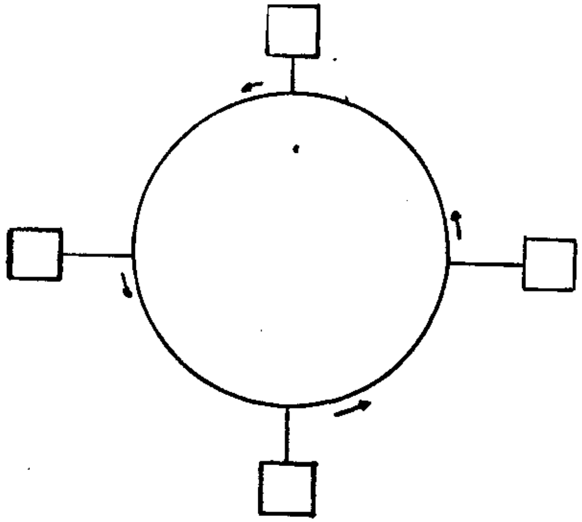
CONTROLLED ACCESS:

TOKEN PASSING RING

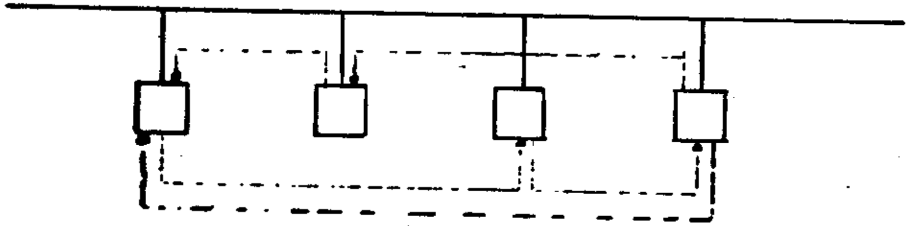
- ACCESS CONTROLLED BY TOKEN CIRCULATING IN RING IF NO MESSAGE TOKEN IS PASSED ON
- IF A STATION WANTS TO TRANSMIT IT FIRST HAS TO CAPTURE IT AND CHANGE IT TO BUSY
- BLOCK LENGTH MESSAGE ARBRITARY
- STATION SENDING MESSAGE HAS TO REMOVE IT FROM RING.

TOKEN PASSING BUS

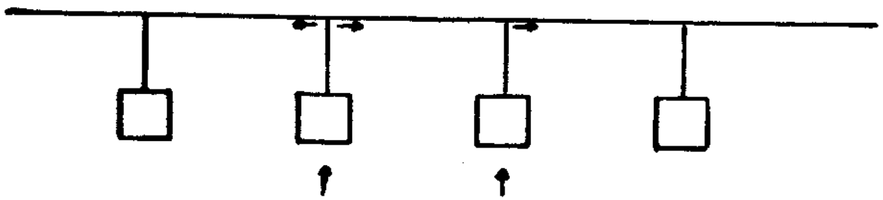
- ADDRESSED TOKEN PASSED FROM ONE STATION TO THE OTHER
- CAPTURE OF TOKEN AS WITH RING
- AT THE END OF TRANSMISSION TOKEN IS PASSED ON BY TRANSMITTING STATION
- NO REMOVAL OF MESSAGES REQUIRED



TOKEN RING



TOKEN BUS



CSMA/CD

CSMA WITH COLLISION DETECT

- IF STATION WANTS TO SEND FIRST LISTEN BEFORE TRANSMIT TO FIND OUT TRANSMISSION CONDITION
- IF TRANSMISSION IN PROGRESS STATION WAITS
- IF FREE, STATION SENDS MESSAGE

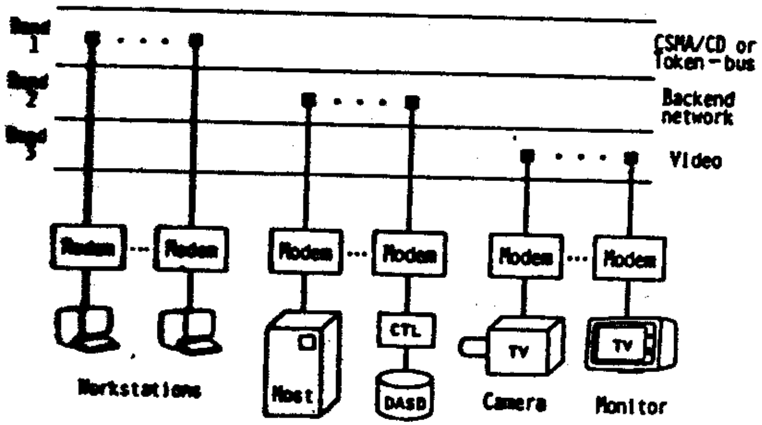
PROBLEM AT SAME TIME ANOTHER STATION CAN ALSO SEND ⟶ COLLISION

- SENDING STATION CAN DETECT THIS BY COMPARING TRANSMITTED WITH RECEIVED DATA
- IN CASE OF COLLISION TRANSMISSION ABORTED TRIES AGAIN AFTER RANDOM TIME
  (RANDOM TIME = N X BUS DELAY TIME)

CLASSIFICATION BASED ON TRANSMISSION

- BASEBAND SYSTEMS
  FULL BANDWIDTH USED FOR TRANSMISSION
- BROADBAND SYSTEMS
  FREQUENCY MULTIPLEXING INTO DIFFERENT BANDS SO ONE CABLE FOR
  - TOKEN RING / BUS OR CSMA/CD
  - VOICE
  - VIDEO

USE OF A BROADBAND SYSTEM



MIMD SYSTEMS AND μP

WHAT IS LEFT ??

- TIGHT COUPLED
- MODERATE COUPLED

PROBLEM

- NO MODEL LIKE ISO AVAILABLE
- NO STANDARIZATION

HOWEVER WE COULD USE THE ISO MODEL TO DEFINE AT WHICH LEVEL WE TALK

- ABOUT AT THE DATA LINK LAYER
- BECAUSE INTERCONNECTIONS ARE DISCUSSED

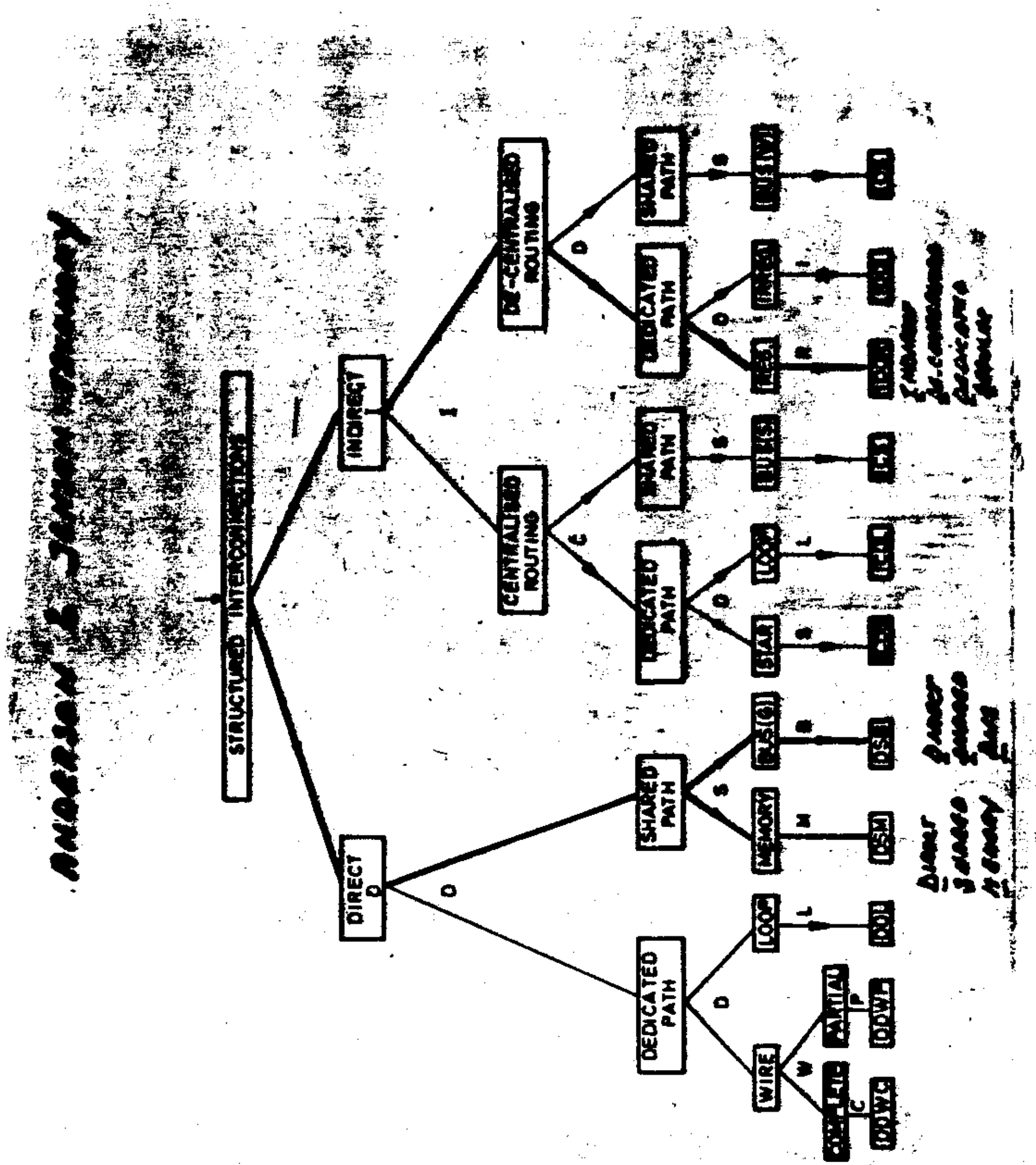
WHY μP BASED MIMD SYSTEMS??

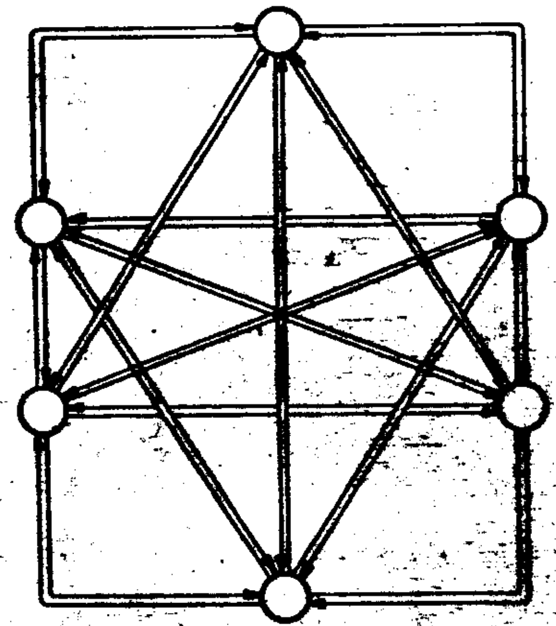
ADVANTAGES :

- MODULAR / FLEXIBLE
- REDUNDANT / RELIABLE
- QUICK RESPONSE TIME
- SYSTEM OPTIMALISATION

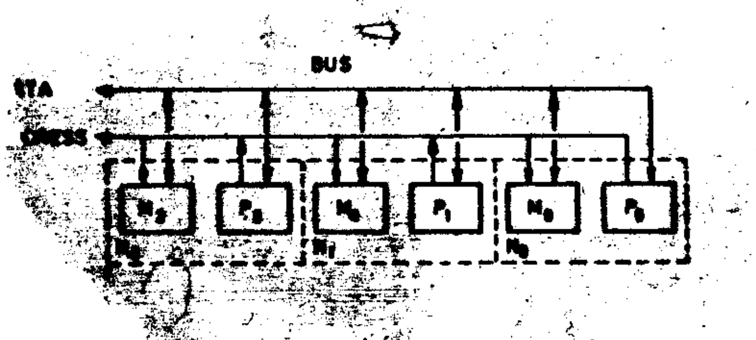
DANGER : PERFORMANCE REDUCTION

- SYNCHRONISATION
- COMMUNICATION
- RESOURCE SHARING





DIRECT DEDICATED WIRE DIRECT



DIRECT SHARED BUS

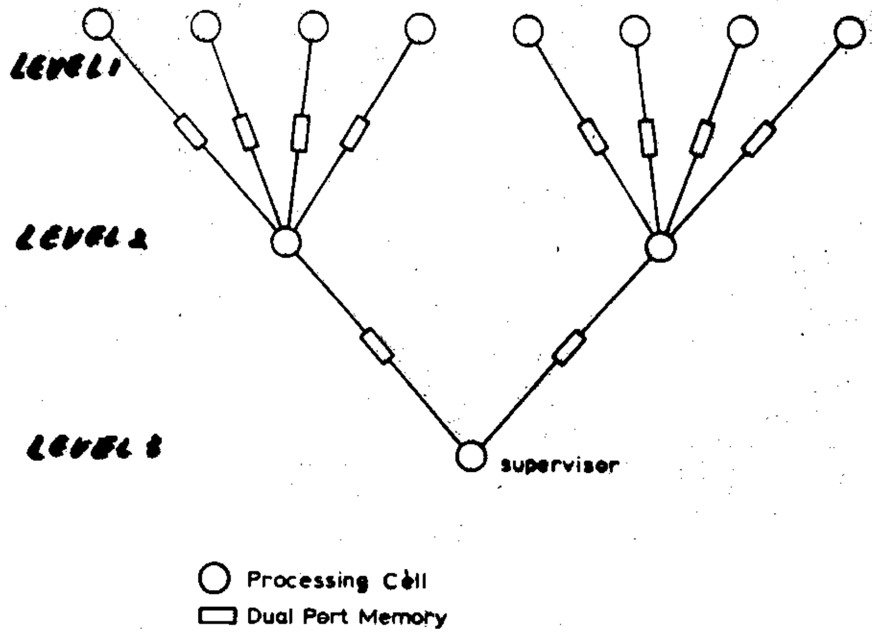
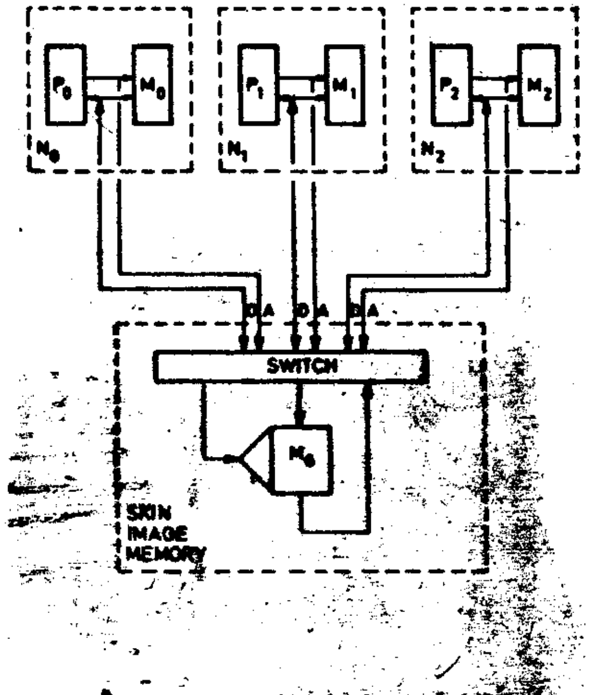


FIG. 2A VERTICAL SYSTEM EXPANSION



DIRECT SHARED MEMORY

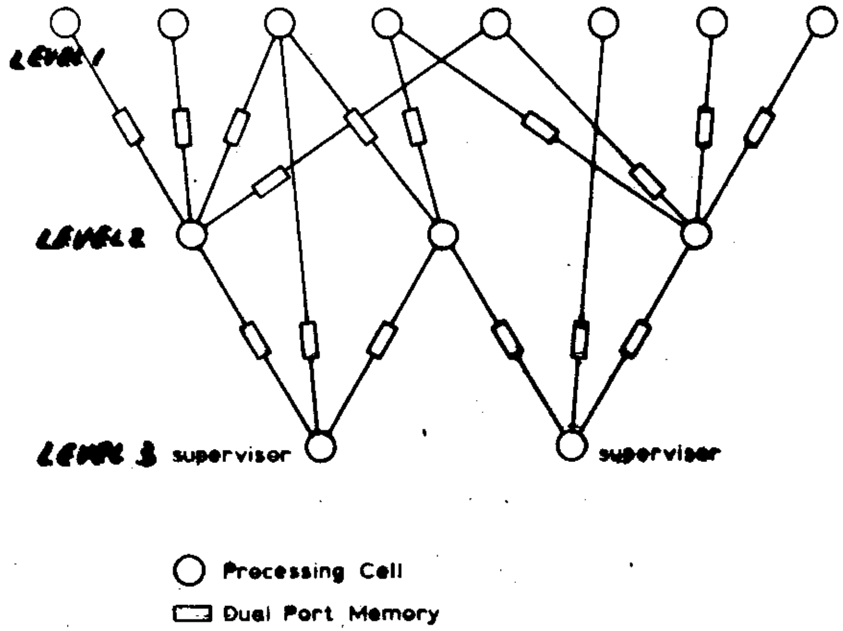


FIG.2B VERTICAL SYSTEM EXPANSION WITH TWO SUPERVISORS

STRUCTURES REALIZED

- DIRECT SHARED BUS
  - DEDICATED BUS
  - TIMESHARED BUS
  - CROSSBAR SWITCH
- DIRECT SHARED MEMORY
  - MULTIPORT MEMORY
- INDIRECT DE-CENTRALIZED DEDICATED REGULAR
  - TREE STRUCTURE
- COMBINATIONS

→ ARBITRATION

IF ≥ 2 PROCESSORS / PROCESSES } WANT SAME RESOURCES

AT THE SAME TIME

PROTECTION WITH FLAG : SEMAPHORE

HARDWARE BUS/MEMORY ARBITRATOR
SOFTWARE CRITICAL REGION (TAS)

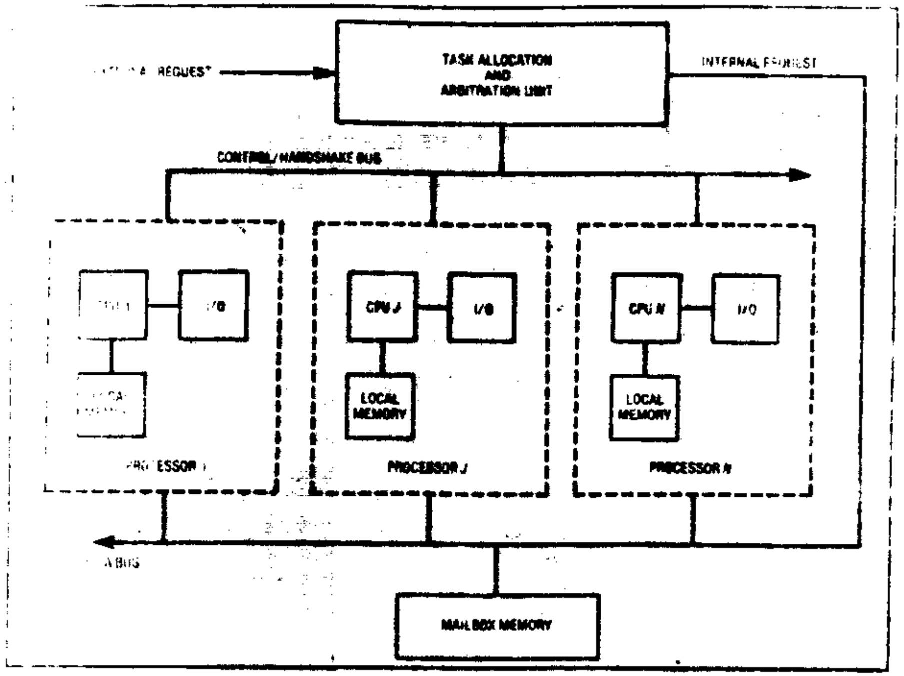


Figure . ...mplified block diagram of a task-driven multi-microcomputer system.

SOFTWARE FOR MULTIPLE PROCESSOR SYSTEMS

MOST DIFFICULT PART OF MULTIPLE PROCESSOR SYSTEMS COMMUNICATION BETWEEN CONCURRENT PROCESSES

CLASSIC EXAMPLE PRODUCER/CONSUMER PROBLEM

- PRODUCER PROCESS DATA TO BE UTILIZED BY CONSUMER PROCESS
  (PRODUCER READS FROM DISC OR TAPE)
- A FULL DATA BUFFER FROM PRODUCER TO CONSUMER AN EMPTY BUFFER IS RETURNED

A NUMBER OF PROBLEMS CAN OCCUR :

MUTUAL EXCLUSION

- IF TWO PROCESSES WANT TO HAVE THE SAME RECOURCES AT THE SAME TIME ONE WINS THE OTHER HAS TO WAIT
  (E.G. BOTH WANT A TAPE, OR DATA FROM A DUAL PORT MEMORY)

SYNCHRONISATION

- COÖPERATING PROCESSES MAY REQUIRE SOME FORM OF SYNCHRONISATION ON OPERATING ON SHARED RESOURCES
- FIRST DATA HAS TO BE PRODUCED BEFORE IT CAN BE CONSUMED

PROC. / TASK

- SOFTWARE MODULE (SUBROUTINE)
- AMOUNT OF CODE PERFORMING CERTAIN TASK
  - CAN BE INDEPENDENTLY TESTED
  - HAS OWN LOCAL VARIABLES
- BEHAVES AS ACTIVE ELEMENT
- CAN ACCESS GLOBAL (COMMON RECOURCES)
- MODERN LANGUAGES TRY TO LIMIT AMOUNT OF CODE TO MINIMUM

```
TASK CALCULATE
    A LOCAL
    B LOCAL
  A = A+B
END

TASK MULTIPLY
TASK DEVIDE
```

- OCCURS IN SITUATION WHERE PROCESSES ARE MUTUALLY WAITING FOR EVENTS CAUSED BY THE OTHER PROCESS

DETERMINACY

- PROCESSES WHICH CONCURRENTLY ACCESS COMMON DATA COULD ALTER THIS DATA BY AN INTERLEAVED SEQUENCE OF OPERATIONS (DATA INVALIDATION)

COMMUNICATION BETWEEN PROCESSES CAN AVOID THESE PROBLEMS BY TWO CONCEPTS:

- MONITOR CONCEPT
- MESSAGE PASSING CONCEPT

HARDWARE CONSIDERATIONS

- MONITOR CONCEPT REQUIRES EXISTENCE OF COMMON ADDRESSABLE MEMORY
- MESSAGE PASSING DOES NOT REQUIRE THIS EXISTENCE (NETWORK)

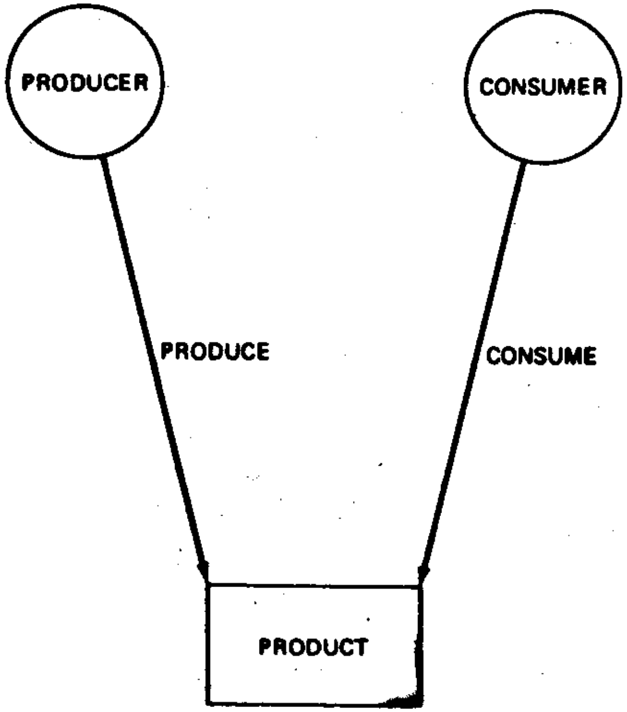
MONITOR

- USES CRITICAL REGION CONCEPT
  A CRITICAL REGION OF CODE IN A CONCURRENT PROCESS IS CONTROLLED SUCH THAT ONLY ONE PROCESS CAN ENTER AT THE TIME
- COMPARE WITH RAILWAY TRACK IN WHICH ONLY ONE TRAIN CAN ENTER AT THE TIME ENTRANCE AND EXIT ARE PROTECTED BY SEMAPHORES
- A TRAIN APPROACHING A CRITICAL REGION WAITS IF SEMAPHORE IS UP, IF SEMAPHORE DROPS TRAIN CAN PROCEED

SEMAPHORE

- CODE TEST SEMAPHORE AND INCASE NOT SET, SETS IT AND ENTERS CRITICAL REGION (TEST AND SET ARE INDIVISIBLE)
- AT THE END SEMAPHORE HAS TO BE RESET AGAIN
- IN CASE SEMAPHORE IS SET PROCESS HAS TO WAIT IN READY QUEUE (SCHEDULING)

A MONITOR IS A CRITICAL REGION IN WHICH SCHEDULING AND INTERPROCESS COMMUNICATION IS PERFORMED



SEMAPHORE

A SEMAPHORE IS A COUNTER S INITIALIZED S:=1

- IF A PROCESS WANTS TO ENTER CRITICAL REGION:

  IT PERFORMS A WAIT

  - IF COUNTER S>0 IT ENTERS CRITICAL REGION AND S:=S-1
  - ELSE PROCESS SUSPENDED IN READY TO RUN (RTR) QUEUE

- IF PROCESS LEAVES CRITICAL REGION

  IT PERFORMS A SIGNAL

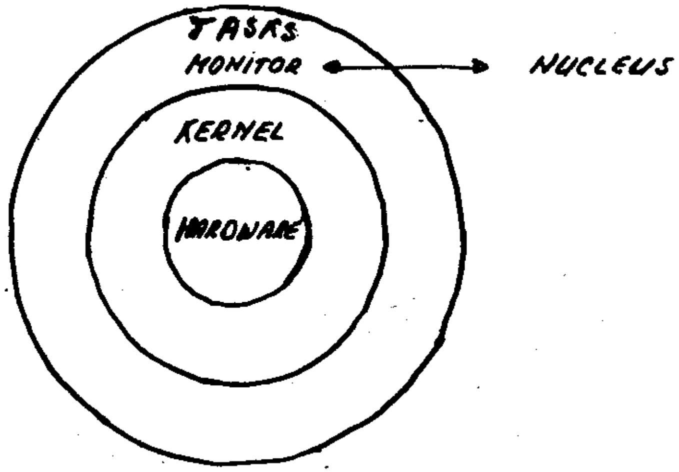
  - IF COUNTER S≤0 NEXT PROCESS FROM RTR QUEUE IS STARTED
  - ELSE S:=S+1

SYNCHRONISATION

- USING A CRITICAL REGION A BUFFER FULL BUFFER ENTRY/POINTER CAN BE PASSED FROM PRODUCER TO CONSUMER

SCHEDULING

- ASSOCIATION OF STATUS INFORMATION WITH CRITICAL REGION
- A BLOCK IS A WAIT ON A CONDITION RECORDED IN STATUS INFORMATION
- PROCESS SUSPENDED IN QUEUE ASSOCIATED WITH STATUS INFORMATION



TASKS : USER PROGRAMS

MONITOR
NUCLEUS } SYNCHRONISATION

KERNEL   SERVICE TASKS
          BOOTSTRAP

SYNCHRONISATION

SEMAPHORE (DIJKSTRA, 1968)
- COMMUNICATION COMMON DATA
- COMMON MEMORY
  (NOT STRUCTURED)

MONITOR (HOARE, 1974)
- INITIALISATION PROCEDURES
  (LOCAL DATA)
- COMMON PROCEDURES (FOR TASKS)
- SYNCHRONISATION WAIT/SIGNAL
  ONLY ONE TASK ACTIVE AT A TIME

EXAMPLES ON SINGLE PROCESSING SYSTEMS
- CONCURRENT PASCAL (1975)
- MODULA (1977)

AFTER 1975 :
- PARALLEL CONCEPTS TO BE IMPLEMENTED ON MULTIPLE PROCESSORS
- TASKS EXECUTING CONCURRENTLY ON MORE PROCESSORS
- MONITOR IN SHARED MEMORY
  ↳ HOWEVER :
  TOO SEVERE HARDWARE CONSTRAINT BECAUSE SHARED MEMORY

NEW DEVELOPMENTS

TASK SYNCHRONISATION

- SEMAPHORES (MONITORS)
  (ORIGINAL ON SHARED MEMORY)
- MESSAGE PASSING (NUCLEUS)
  (ORIGINAL ON NETWORKS)
  - SEND PRIMITIVE
  - REPLY PRIMITIVE
  - RECEIVE PRIMITIVE

- BRINCH HANSEN
  MONITOR IMPLEMENTATION WITHOUT NEED FOR SHARED MEMORY
- HOARE
  MESSAGE PASSING AS SIMPLE INPUT OUTPUT OPERATIONS

WHY

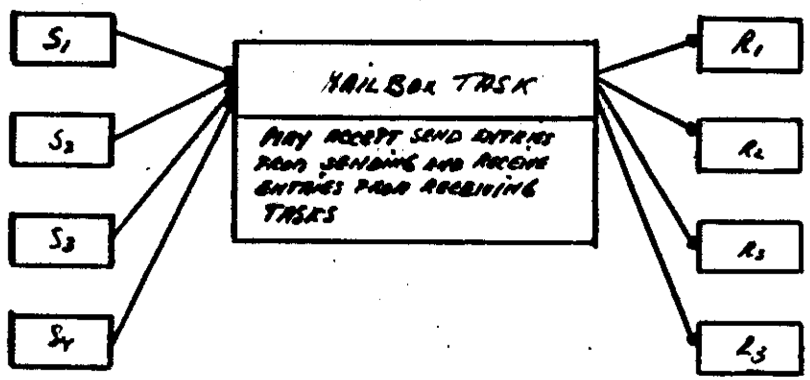
- HARDWARE INDEPENDENCE
- OPTIMAL USE CONCURRENCY
- MINIMUM INTERACTION HIGHER LEVEL LANGUAGE
- OS HANDLES RT ASPECTS

MESSAGE PASSING

FOR SYNCHRONISATION AND COMMINICATION TWO DIFFERENT STRATEGIES :

- PRIMITIVES THAT PERFORM SYNCHRONISATION AND COMMUNICATION IN A PASSIVE WAY (MONITORS)
- PRIMITIVES THAT PERFORM SYNCHRONISATION TOGETHER WITH COMMUNICATION IN A ACTIVE WAY (MESSAGE PASSING)
- MESSAGES SEND, RECEIVE, REPLY

MAILBOX



SENDING TASK SENDS MESSAGES VIA OUTMAIL
RECEIVING TASK RECEIVES MESSAGES VIA INMAIL
MAILBOX TASK TRANSMITTING

RENDEZVOUS

↓ SEND ↓ SENDING TASK

↓ ACCEPT SEND (INMAIL) ↓ MAILBOX TASK

- IF SEND IS PERFORMED BEFORE ACCEPT SEND IS REACHED: SEND WAITS
- IF ACCEPT SEND IS PERFORMED BEFORE SEND WAS REACHED: ACCEPT SEND WAITS
- IF SEND IS PERFORMED AND ACCEPT SEND IS REACHED THERE IS A RENDEZVOUS

↓ RECEIVE ↓ RECEIVING TASK

↓ ACCEPT RECEIVE (OUTMAIL) ↓ MAILBOX TASK

- INTERRUPT IS ACCEPT RECEIVE

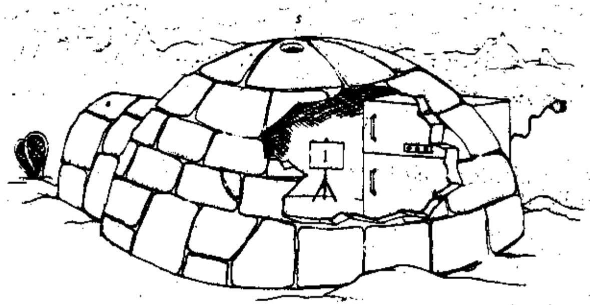


Fig. 4.2.

MONITOR CONCEPT

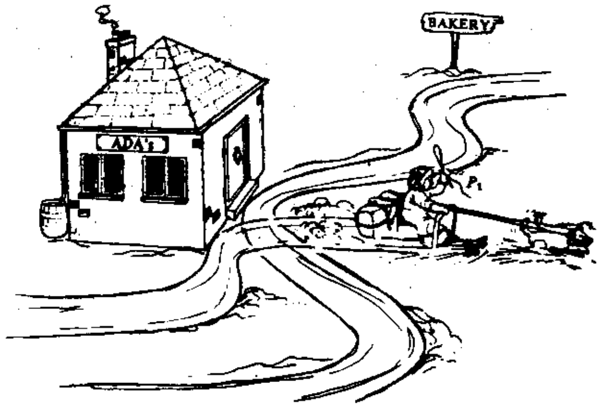


Fig. 6.1.

ADA RENDEZ-VOUS

OTHER SOFTWARE FUNCTIONS

SCHEDULING

- LOGIC SCHEDULING DEFINE TASK THAT CAN RUN CONCURRENLY (USER, PARTLY COMPILER)
- PHYSICAL SCHEDULING DETERMINATION WHICH TASK SHOULD RUN (PREEMTIVE SCHEDULING)

DIAGNOSTICS

- EACH CPU MODULE MUST HAVE FACILITIES TO TEST HIS NEIGHBOURS
- EACH MODULE SHOULD HAVE FULL SET OF TEST PROGRAMS THAT CAN BE ACTIVATED FROM OUTSIDE
- ACTIVATION CAN BE TASK OF SCHEDULER

DEBUGGING

- PROGRAM DEBUGGING ON MACHINE ON WHICH THEY ARE WRITTEN (HOST) (SYMBOLIC DEBUGGER, SIMULATION)
- SPECIAL DEBUGGING TOOLS FOR MULTIPLE PROCESSORS DIFFICULT

MULTIPLE μP SYSTEM:

- PREFERENCE MODERATE COUPLING
- BECAUSE TENDENCY TO CONCURRENTLY EXECUTE A NUMBER OF TASKS THEY ARE CALLED
  MULTIPLE TASK MULTIPLE DATA STREAM (MTMD)

CONCURRENCY NOW PROVIDED BY

- OPERATING / RUNTIME SYSTEM
  TASKS DEFINED BY USER OPERATING SYSTEM PROVIDES PRIMITIVES FOR: SYNCHRONISATION AND COMMUNICATION

CONCURRENCY IN NEAR FUTURE PROVIDED BY

- LANGUAGE E.G. COMPILER (ADA, MODULA-2)

CONSEQUENCE FOR ARCHITECTURE

- INDEPENDENT MONOBOARD COMPUTERS (CPU, MEM, I/O) WITH SOMETIMES COMMON ADDRESSABLE MEMORY
  INTERCONNECTION:
  - TIME SHARED BUS
  - COAX CABLE (NETWORK)
  - DUAL PORT MEMORY

THIS FOR EXCHANGE OF:

- MESSAGES
- DATA

FUNCTIONAL MODULARI... (69)

- SYSTEMS CAN BE DISTRIBUTED IN FUNCTIONS WITHOUT NEED FOR TIGHT CENTRAL SYNCHRONIZATION E.G.
  - PRE_PROCESSING OF DATA
  - LOCAL REAL_TIME RESPONSE (CONTROL)
  - LOCAL DATA TAKING

RELIABILITY

- REDUNDANCY EASIER TO CREATE IN DISTRIBUTED THEN CENTRALIZED SYSTEM
- GRACEFUL DEGRADATION IN CASE A COMPONENT FAILS

RESOURCE SHARING

- HARDWARE & SOFTWARE
- DATA & DATA BASES
- COMPUTATATIONAL POWER

RESPONSE

- REAL TIME SITUATIONS DEMAND OFTEN CRITICAL BUT SIMPLE RESPONSE (SCANNER, COMPUTER CONTROLLED MACHINE)
- LOCAL (DISTRIBUTED INTELLIGENCE EASIER IN THIS CASE

WHERE TO USE MTMD SYSTEMS (70)

PROCESS CONTROL

- DIFFERENT COMPUTATIONAL DEMANDS
- REAL_TIME CONSTRAINTS (E.G THE COMPUTER HAS TO RESPOND WITHIN CERTAIN AMOUNT OF TIME)
- ALSO ROBOTICS IS SUCH A FIELD
- ALSO MONITORING OF COMPLEX PHYSICS DETECTOR

SYSTEMS WITH EXTENSIVE I/O PROCESSING

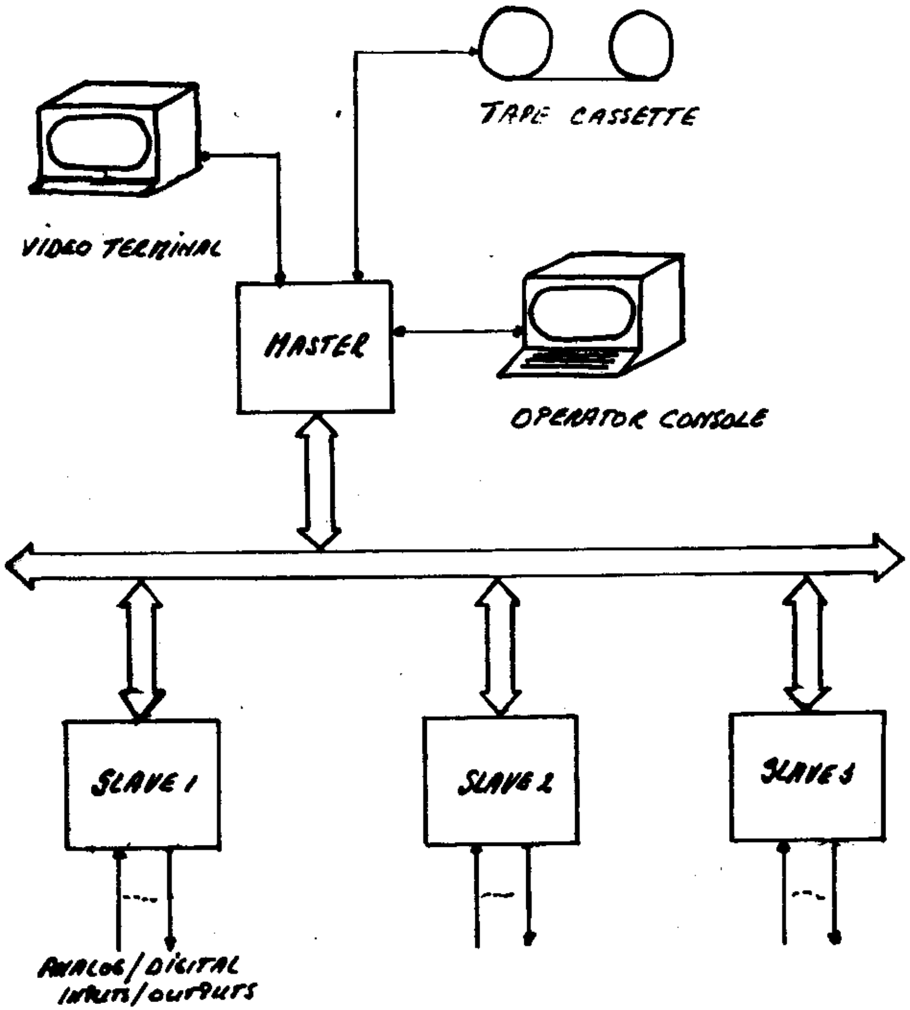
- DATA AQUISITION SYSTEMS WITH DATASTREAMS COMING FOR MORE THAN ONE SOURCE
- DOING THIS WITH ONE MONOBOARD SYSTEM CAUSES SEVERE SYSTEM DEGRADATION (CONTROL)

SYSTEMS REQUIRING HIGH RELIABILITY

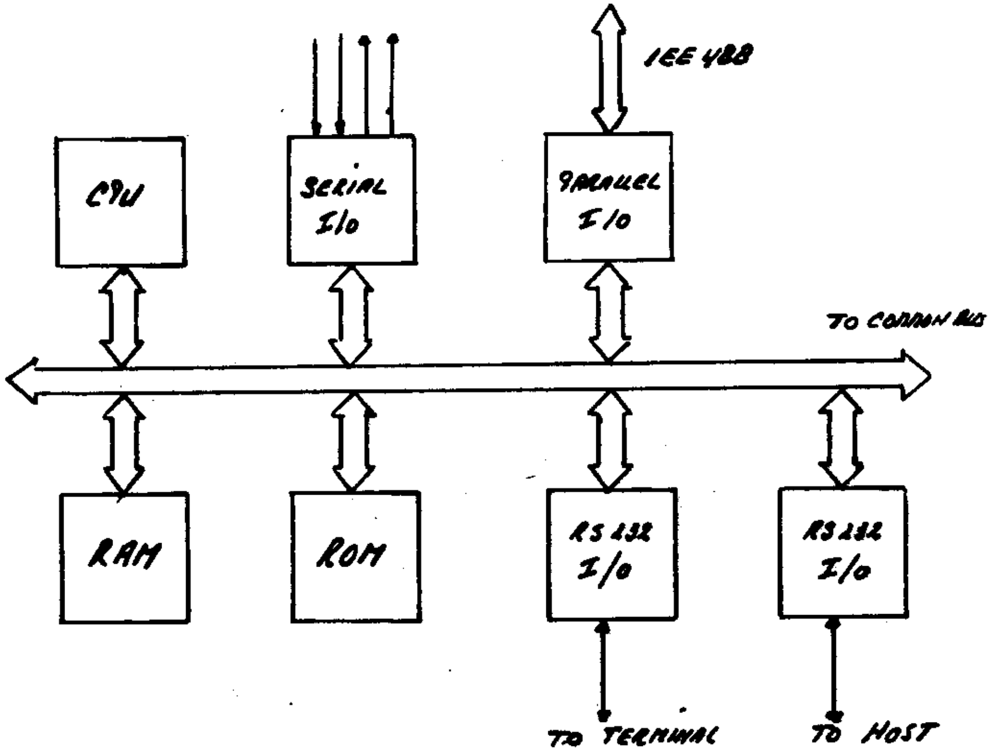
- MILITAIRY
- FLIGHT CONTROL
- PROCESS CONTROL

BE CAREFULL TO USE MTMD SYSTEMS IF:

- ONLY MORE PROCESSING POWER IS REQUIRED (FOR TIME BEING)
- GLOBAL TASK CANNOT BE PARTITIONED IN RELATIVELY INDEPENDANT SUBTASKS



HYDRAULIC PRESS CONTROL



MONOBOARD COMPUTER WITH I/O

TO BE USED MASTER / SLAVE } COMPUTER

HYDRAULIC PRESS

PURPOSE

- YIELDING METAL SHEETS IN A MOULD
- THIS IS DONE BY COORDINATED MOVEMENT OF CLAMPS AND MOULD HOLDER.
- CLAMPS AND MOULD ARE CONTROLLED BY STEPPING MOTOR TYPE CONTROL

TWO MODES OF OPERATION

- OFF LINE PREPARATION ON CASSETTE (LEARNING PHASE)
- ON LINE CONTROL MACHINE CYCLES

SYSTEM STRUCTURE & FUNCTIONS

- MASTER AND THREE SLAVES
- INFORMATION FLOW <u>BETWEEN SLAVES</u> RESTRICTED TO EMERGENCY INTERRUPTS
- HANDLER OFF LINE COMMANDS
- DEFINITION OF MACHINE CYCLES
- → EXECUTION MACHINE CYCLES
- → CONTROL MACHINE STATUS DISPLAY HANDLERS
- → SYNCHRONISATION BETWEEN MASTER & SLAVES
- FILE SYSTEM ON CASETTE

NUMBER OF SIGNALS

- 85 DIGITAL INPUT (PLUNGERS & SELECTORS)
- 56 DIGITAL ALARMS & DIAGNOSTIC SIGNALS
- 40 DIGITAL OUTPUTS (LIGHTS & VALVE CONTROLS)
- 12 CONTROL LINES PRESS MOVEMENTS

SOFTWARE REALIZATION

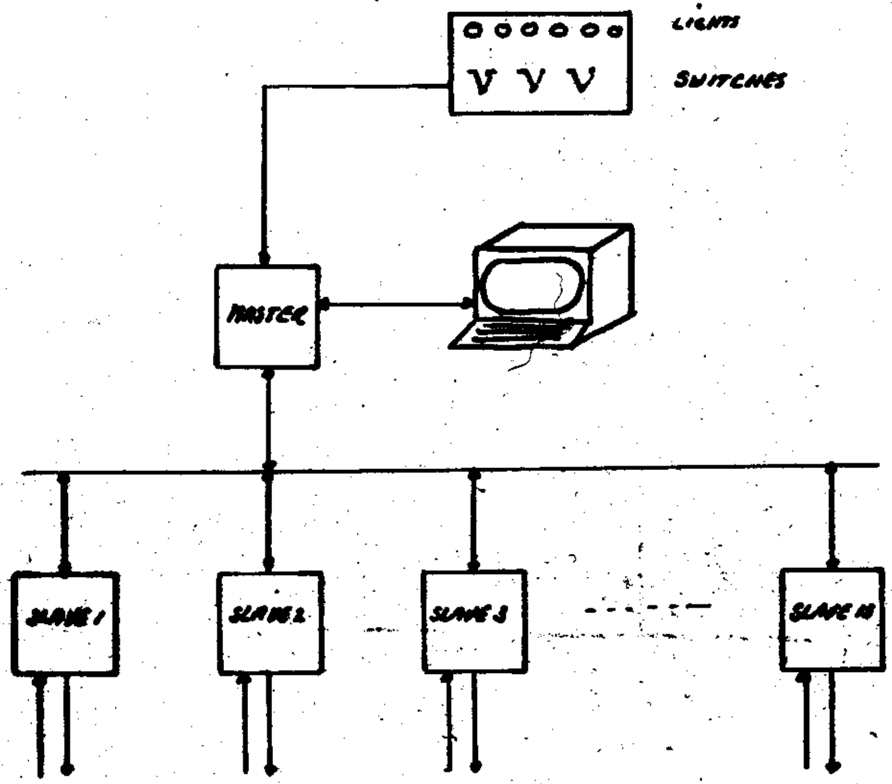
SLAVES

- AUTO DIAGNOSTICS DURING IDLE PERIODS (LOCAL SCHEDULER)
- ALARM TASKS BASED ON HIGH PRIORITY INTERRUPTS (LOCAL ACTION)
- TASKS TO CONTROL SENSORS AND ACTUATORS (SYNCHRONIZATION REQUIRED)

MASTER

- COMMUNICATION TASK WITH VIDEO MONITOR
- TASK FOR CONTROL OPERATOR CONSOLE
- TASK FOR CONTROL CASSETTE
- COMMAND INTERPRETER TASK
- MANUAL CONTROL TASK }
- AUTOMATIC CONTROL TASK }

- MASTER ISSUES COMMANDS TO SLAVES
- EACH COMMAND HAS TO BE COMPLETED BEFORE NEXT COMMAND IS ISSUED
- SIMPLE PRODUCER RELATION (SEND, RECEIVE, REPLY MECHANISME)



BOTTLING LINE CONTROL SYSTEM