CPE 633
Chapter 7 – Case Studies

Dr. Rhonda Kay Gaede

UAH

Electrical and Computer Engineering

7.1 NonStop Systems

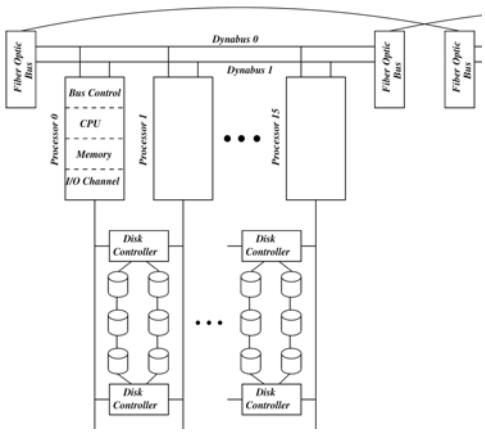
- The main use of these fault-tolerant systems has been in online transaction processing, where a reliable response to inquiries in real time must be guaranteed.
- The systems, and their fault-tolerant features, have been evolving since 1976 through several generations.

7.1 NonStop Systems
- Key Design Principles

- Modularity – Hardware and software are constructed of fine grain modules. These modules constitute units of failure, diagnosis, service, and repair.
- Fail-Fast Operation – A fail-fast module either works properly or stops. Hardware checks and software consistency checks support fail-fast operation.
- Single Failure Tolerance – For processors, this means that a second processor is available. For storage modules, it means that the module and the path to it are duplicated.
- Online Maintenance - Hardware and software modules can be diagnosed, disconnected for repair and then reconnected, without disrupting the entire system's operation.

7.1.1 NonStop Systems
- Architecture

- The system consists of clusters of up to 16 processors.
- Each custom processor has a CPU, local memory with its own operating system copy, a bus control unit and an I/O channel.



7.1.1 NonStop Systems - Architecture

- CPU Error Detection
  - Datapath
    - All units that don't modify data propagate (buses, registers) the parity bits.
    - Other units that alter data (arithmetic, counters) use special circuits that predict parity based on data and parity inputs, doesn't work well for multiply.
    - For multiply, a second multiplication is performed with one operand shifted.
  - Control
    - Parity checks, illegal state detection, self-checking logic.
  - Scan path support.

7.1.1 NonStop Systems - Architecture

- Memory Error Detection
  - No shared memory for single point of failure.
  - Cache and main memory have Hamming code capable of single-error correction and double error detection for data and single-error-detection parity code for address.
  - The cache performs retries to take care of transient faults.
  - There is a spare memory module that can be switched in if permanent failures occur.
  - The cache supports a write-through policy, guaranteeing the existence of a valid copy of the data in main memory.
  - A parity error in the cache forces a cache miss followed by a fetch from main memory.

7.1.1 NonStop Systems - Architecture

- The 16 processors operate independently and asynchronously and communicate via messages sent over the dual Dynabuses.
- The Dynabus interface is designed so that a single processor failure will not disable both buses.
- Similar duplication occurs in the I/O systems, in which a group of disks is controlled by dual-ported controllers connected to I/O buses from two different processors.
- One of the ports is primary with the other backup.
- All data transfers are parity-checked, and a watchdog timer detects that a controller stops responding or that a nonexistent controller address.
- The controllers achieve fail-fast by using dual lock-stepped microprocessors with self-checking logic.
- The two independent ports are physically separated circuits.

7.1.1 NonStop Systems - Architecture

- Power and cabling also matter. Parts of the system are redundantly powered from two different power supplies and battery backups are provided.
- The system supports disk mirroring which, along with 4 different controller device paths, provides eight paths for data read and write operations.
- The disk data is protected by end-to-end checksums in which the processor calculates a checksum and appends it to the data written to the disks.
- Checksum – error detection
- Disk mirroring – data recovery

7.1.2 NonStop Systems
- Maintenance and Repair Aids

- A maintenance and diagnostic processor
  - Communicates with all processors and remote service center
  - Collects failure related information
  - Capable of some automated failure diagnosis
  - Allows engineers at remote center to run diagnostic tests
  - Can reconfigure the system
- Each computing processor diagnostic unit
  - Monitors status, memory, Dynabus interface, I/O channel
  - Reports to the maintenance processor
  - Can force the computing processor to single-step and collect diagnostic information using the scan-paths
  - Generate pseudo-random tests

7.1.3 NonStop Systems
- Software

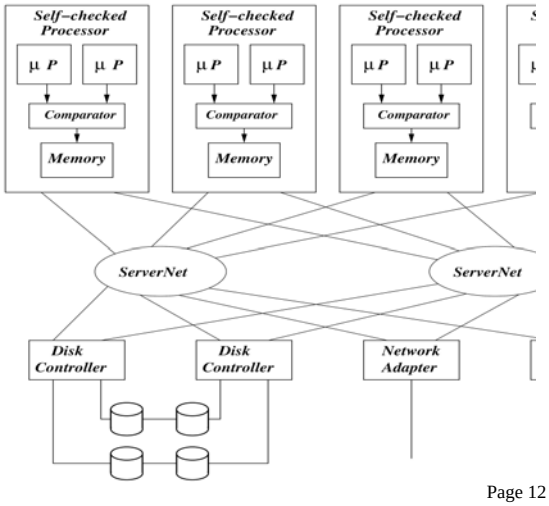
- The hardware redundancy present (redundant communication buses) contribute to the performance of the fault-free systems.
- Most of the burden of the system fault-tolerance is borne by the operating system (OS).
- The OS detects failures of processors or I/O channels and performs the necessary recovery.
- When a new process starts, the OS generates a clone on backup processor. This clone waits for messages from the OS or the primary.
- Checkpoints taken on the primary are sent to the backup, process sate updated by OS.
- If the primary process fails, the OS orders the backup to start execution from the last checkpoint.

7.1.3 NonStop Systems - Software

- Processors continuously ping each other and themselves every second and check for responses from all every two seconds.
- If a processor does not hear from another processor, all outstanding communications with it are cancelled.
- An important OS component is the disk access process, it is implemented as a primary/backup process pair.
- To deal with software failures, numerous consistency checks are included in every software module.
- Upon detection of a problem, the processor is halted and the backup process initiated.

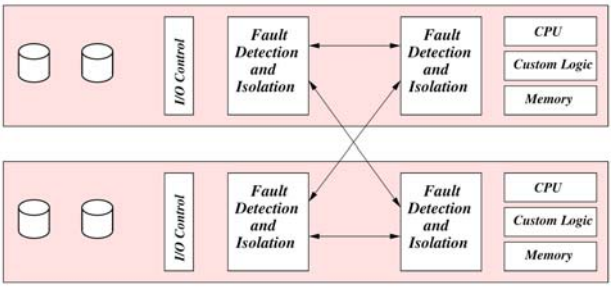
7.1.4 NonStop Systems - Modifications to the Architecture

- Use commercial processors
- No self-checking so lock-step operation
- Use packet switching, protect each packet with CRC
- Lock-stepping no longer practical, only compare I/O operations
- Allows TMR



7.2 Stratus Systems

- Like the NonStop system, every unit is replicated (at least once), including processors, memory units, I/O controllers, disk and communication controllers, buses, and power supplies.
- Unlike the NonStop system, off-the-shelf software need not be modified to consist of primary/backup process pairs.
- Stratus uses the pair and spare principle, in which each pair consists of two lock-step processors.

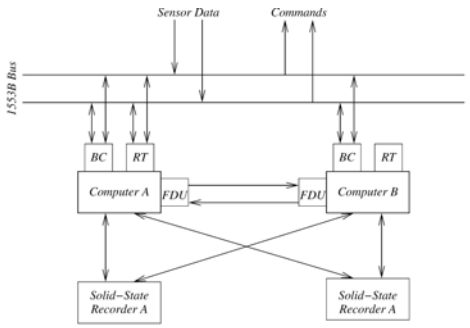


7.2 Stratus Systems

- More recent versions have loosened the lock-step requirement to matching I/O operations.
- Like NonStop, Current Stratus systems can be configured to use TMR structures with voting to detect or mask failures.
- Unlike NonStop, the memory unit is also duplicated allowing the contents of the main memory to be preserved through most system crashes.
- The I/O and disks are duplicated as well, with redundant paths connecting individual I/O controllers and disks to the processors. The disk systems use disk mirroring.
- The processors, memories and I/O units have hardware error-checking.
- Special attention is given to device drivers.
- On a crash, one CPU is kept offline to dump memory to disk, it rejoins after the dump is complete.
- Every fault is reported to a remote Stratus support center.

7.3 Cassini Command and Data Subsystem

- Fault-tolerance is key to an 11 year mission and is provided by a dual-redundant system.
- The heart of the CDS is a pair of flight computers, each with tiny memories (512K) and two 2 Gbit solid-state recorders.
- Communication between the computers and storage occurs via a dual-redundant bus.
- Faults handled first by ground control, then by autonomous recovery.



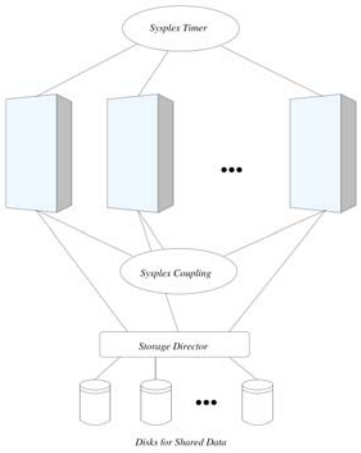
FDU Fault Detection Unit: Manages CDS redundancy
RT Remote Terminal: Permits communication with the other computer
BC Bus Controller: 1553B controller. Only one is active at any time.

7.4 IBM G5

- I/O System
  - There are multiple, dynamically switched paths from the processor to the I/O devices.
  - Inline error checking is provided.
  - Channel adapters prevent propagation of interface errors.
- Processor
  - Duplicated I-unit and E-unit, pairs work in lock-step
  - R-unit registers are protected by ECC
  - Store buffer is protected by ECC
  - Writes to the L1 are written through to the L2
  - L2, main memory data and buses protected by (72, 64) SEC/DED Hamming code, L1 only has parity
  - L2 lines can be invalidated, special logic detects multiple failures of a single line, spare line swapped in
- Memory
  - Address bus has 1 parity bits per 24 bits
  - Memory scrubbing, reading and correcting entire contents

7.5 IBM Sysplex

- The IBM Sysplex is the current generation IBM mainframe.
- Up to 32 single- or multi-processor units inter-connected with a global timer.
- Storage devices are equally shared.
- Units periodically emit heartbeats, absence indicates failure, must be sure of failure for restart
- Automatic Restart Manager restarts affected tasks, it is aware of global system state.



7.6 Itanium

- L1I and L1D tag and data arrays are protected by error-detecting parity
- L1D has byte-wise parity for fine-grain stores
- Bits from adjacent cache lines are physically interleaved on silicon
- L2 has SEC/DED coding or data array and parity for tag array
- L3 tag and data arrays have SEC/DED coding
- If error not hardware-correctable, hardware containment may be required and error handling initiated.
- Error handling proceeds from hardware layer to processor abstraction layer to system abstraction layer, to operating system
- Processor used for current NonStop systems