Tandem

- Tandem
  - from

- to HP
Tandem

- Background
- Tandem NonStop Systems - Cyclone
  - Commercial Database Systems with very long MTTF
  - Modularity
    » units of service, failure, diagnosis, repair, growth
    » fault containment regions
    » expandable for performance
Tandem

- Fail Fast Mode
  - terminate operation immediately after error detection
  - reduces error propagation
  - single error corrections/ double error detection
  - ECC, data coding
  - hardware self checking
  - software and firmware consistency checks
  - after failure OS distributes processors applications on remaining processors
  - load balancing is transparent to user
Tandem

- **Architecture** (Overview in Pra96. Fig. 4.1)
  - loosely coupled MIMD, up to 16 processors
  - dual processors, independent & asynchronous
  - heavy use of low-level dual redundancy
  - multiple, physically separate sections
  - each section: up to 4 processors, communication via Dynabus
  - write through cache
  - mirrored disks

- **Processor Pair**
  - primary/backup approach
  - primary sends checkpoints
  - when primary proc. fails:
    - backup becomes primary
Tandem

- **Hardware Fault Tolerance**
  - single fault tolerance
  - primary objective to prevent single fault to bring down system
  - redundant hardware: processors, busses, I/O controllers, disks, power supplies
  - spare RAM chips
  - each processor has own power supply

- **Software Fault Tolerance**
  - processors can detect other halted processors
  - “I'm alive" protocol
  - GUARDIAN 90 OS maintains idle backups of user processes
  - Processor consistency check via checkpoint messages
Tandem

- Networking and I/O
  - Networks
    - Dynabus: 40 MB/s = 2 independent 20 MB/s buses
    - Dynabus+: 4 unidirectional fiber optics,
      - up to 50m physical separation
      - robust to electro-magnetic interference
  - I/O
    - processor can support 2 I/O systems
    - each system has 2 channels
    - each channel supports up to 32 I/O devices
    - burst data of 5 MB/s = 10 MB/s per processor
    - DMA I/O
    - mirrored disks (dual ported)
Tandem

- On-line Maintenance
  » Field replaceable units (FRU)
    ■ processors
    ■ I/O controllers
    ■ fans
    ■ power supplies
    ■ can be installed/replaced by user
  » Warm swaps of FRU
  » Effective MTTR = milliseconds
    => very high Availability
Tandem

- Tandem - Himalaya
  - Main features
    - loosely coupled massively parallel computer
    - 2 to 4080 processors
    - cross-coupled MIPS R4400 RISC processors
      - one logical processor
      - both processors operate in lockstep
    - 32K primary cache, 4MB secondary cache
    - up to 256 MB RAM
    - 4 independent I/O channels
    - fiber-optic TorusNet
      - horizontal controller => 4 sections (each section = 4 processors)
      - vertical controller => 14 nodes = domain
      - depth controller => 16 domains
Tandem

- Himalaya 2000
Tandem

K2000SE server
Tandem

optional expansion cabinet

optional I/O expansion cabinet
Tandem

- TorusNet
  » section
  » node
  » ring
Tandem

- K200, K2000, K20000 Servers Spec. Features:
  - Target: online transaction processing
  - Standard RISC technology
  - Loosely coupled architecture
  - Dual interprocessor buses
  - Dual-ported controllers
  - Fault-tolerant power subsystem
  - In case of power outage server memory is preserved via integrated battery backup
Tandem

- NonStop Operating system
  » core of Tandem’s open systems environment
  » enables operation to run primary and backup processes
  » before performing any critical function, sends backup process a checkpoint message containing data and status information
  » kernel supports end-to-end integrity features
  » micro-kernel is message-based (parallel processing software)
  » kernel supports application program and operations control interfaces called “personalities”
  » these personalities support applications from different platforms
  » e.g. relational database management personalities applications can be developed using:
    - SQL, Data Access Language (Macintosh), SQL Server (Microsoft/Sysbase), ODBC (Microsoft), Oracle Tools (Oracle)
Tandem

- other personalities are transaction processing personalities allows parallel transaction processing services for different systems
- “guardian services” allow compatibility to Tandem applications
- “open systems services” supports UNIX
Tandem

» Transaction Manager (NonStop TM/MP) deals with effects of incomplete transactions, system failures and network failures.
» Remote Duplicate Facility allows data to be located remote to shield from environmental disaster.
» Safeguard security management facility deals with security issues
» Network support includes TCP/IP, IPX/SPX, NETBIOS, AppleTalk, SNA, OSI and ATM
Tandem

- Maintenance
  » key data logged and evaluated by expert-system to identify potential problem
  » can dial automatic to online support center
  » field replaceable units can be exchanged by warm swaps
# Himalaya processor specifications

<table>
<thead>
<tr>
<th>Configurations</th>
<th>K200</th>
<th>K2000</th>
<th>K20000 TorusNet Node</th>
<th>K20000 TorusNet Domain</th>
<th>K20000 Multidomain TorusNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microprocessor type</td>
<td>MIPS R4400/125 MHz</td>
<td>MIPS R4400/125 MHz</td>
<td>MIPS R4400/200 MHz</td>
<td>MIPS R4400/200 MHz</td>
<td>MIPS R4400/200 MHz</td>
</tr>
<tr>
<td>Maximum processors</td>
<td>4</td>
<td>16</td>
<td>16</td>
<td>224</td>
<td>4,080</td>
</tr>
<tr>
<td>Relative performance range</td>
<td>2 to 4</td>
<td>2 to 16</td>
<td>3.9 to 31.5</td>
<td>3.9 to 440.9</td>
<td>3.9 to 8,030.5</td>
</tr>
<tr>
<td>Cache (per processor)</td>
<td>1 MB</td>
<td>1 MB</td>
<td>4 MB</td>
<td>4 MB</td>
<td>4 MB</td>
</tr>
<tr>
<td>Maximum main memory</td>
<td>512 MB</td>
<td>4,096 MB</td>
<td>4,096 MB</td>
<td>57 GB</td>
<td>1,044 GB</td>
</tr>
<tr>
<td>Maximum disk storage</td>
<td>184 GB</td>
<td>3,232 GB</td>
<td>65 TB</td>
<td>918 TB</td>
<td>16,711 TB</td>
</tr>
<tr>
<td>Maximum I/O channels</td>
<td>4</td>
<td>16</td>
<td>64</td>
<td>896</td>
<td>16,320</td>
</tr>
<tr>
<td>Multifunction controllers</td>
<td>2/proc. cabinet</td>
<td>2/proc. cabinet</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Minimum battery hold-up time</td>
<td>3.75 hr</td>
<td>2 hr</td>
<td>2 hr</td>
<td>2 hr</td>
<td>2 hr</td>
</tr>
</tbody>
</table>
HP Integrity NonStop

- Tandem was bought by Compaq which merged with HP in 2002.
  - HP Integrity NonStop server: www.hp.com/go/nonstop
  - HP NonStop Advanced Architecture (NSAA)
    » Software fault tolerance
    » Hardware fault tolerance
    » Scalability from 2 to 4,080 processors
    » Scalability to 65 TB of main memory
    » Online database and application manageability
    » Query processing while maintaining transaction response times
    » Data integrity
    » Leverage of industry standards in hardware and software
**HP Integrity NonStop servers**

Ultra-robust servers that deliver 24x7 continuous availability, unrivaled data integrity, and virtually unlimited scalability—ideal for demanding, transaction-intensive applications.

<table>
<thead>
<tr>
<th></th>
<th>HP Integrity NonStop NS21000 server</th>
<th>HP Integrity NonStop NS22000 servers</th>
<th>HP Integrity NonStop BladeSystem NB54000c servers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entry-class, high levels of availability and data integrity in a cost-effective commercial server</td>
<td>Excellent price-performance, software fault-tolerant solution for small to mid-size enterprises</td>
<td>Industry-leading 24x7 availability, scalability, and data integrity with a high level of performance</td>
</tr>
<tr>
<td><strong>Processors supported</strong></td>
<td>Intel® Itanium® processor 9300 series</td>
<td>Intel Itanium processor 9300 series</td>
<td>Intel Itanium processor 9300 series</td>
</tr>
<tr>
<td><strong>Clustering</strong></td>
<td>Expand-over-IP (does not support ServerNet clustering)</td>
<td>Expand-over-IP (does not support ServerNet clustering)</td>
<td>NonStop BladeCluster solution ServerNet clustering</td>
</tr>
<tr>
<td><strong>Number of processors</strong></td>
<td>2 – 4</td>
<td>2 – 4</td>
<td>2 – 16 per node</td>
</tr>
<tr>
<td><strong>Licensable cores per processor</strong></td>
<td>1 (fixed)</td>
<td>2 (fixed)</td>
<td>2 or 4 (user’s choice)</td>
</tr>
<tr>
<td><strong>Maximum number of logical processors per cluster</strong></td>
<td>1020 (Expand-over-IP)</td>
<td>1020 (Expand-over-IP)</td>
<td>4080</td>
</tr>
<tr>
<td><strong>Operating systems supported</strong></td>
<td>HP NonStop OS (J-Series)</td>
<td>HP NonStop OS (J-Series)</td>
<td>HP NonStop OS (J-Series)</td>
</tr>
<tr>
<td><strong>Memory (cluster)</strong></td>
<td>32 TB</td>
<td>32 TB</td>
<td>256 TB</td>
</tr>
<tr>
<td><strong>ServerNet processor connectivity</strong></td>
<td>Versatile I/O (VIO)</td>
<td>Versatile I/O</td>
<td>BladeSystem ServerNet switches</td>
</tr>
<tr>
<td><strong>I/O infrastructure</strong></td>
<td>IP CLuster I/O Module (CLIM), Storage CLIM, Telco CLIM</td>
<td>IP CLIM, Storage CLIM, Telco CLIM</td>
<td>IP CLIM, Storage CLIM, Telco CLIM, I/O Adapter Modular Enclosure (IOAME)</td>
</tr>
<tr>
<td><strong>Maximum CLustered I/O Modules (CLIMs)</strong></td>
<td>6</td>
<td>6</td>
<td>48</td>
</tr>
<tr>
<td><strong>Maximum I/O Adapter Module Enclosure (IOAME) adapters</strong></td>
<td>0</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td><strong>Maximum disk drives</strong></td>
<td>200</td>
<td>200</td>
<td>3248 FC disks with all IOAME-based storage and IP connectivity (no CLIMs) 2300 SAS disks with all CLIM-based storage and IP connectivity (no IOAMEs)</td>
</tr>
<tr>
<td><strong>Management</strong></td>
<td>OSM, HP SIM, HP Insight Control for NonStop, HP NonStop Essentials, Web ViewPoint, ASAP, HP IT Performance Suite</td>
<td>OSM, HP SIM, HP Insight Control for NonStop, HP NonStop Essentials, Web ViewPoint, ASAP, HP IT Performance Suite</td>
<td>OSM, HP SIM, HP Insight Control for NonStop, HP NonStop Essentials, Web ViewPoint, ASAP, HP IT Performance Suite</td>
</tr>
<tr>
<td><strong>Rack height (EUA unit)</strong></td>
<td>36U and 42U</td>
<td>36U and 42U</td>
<td>42U</td>
</tr>
<tr>
<td><strong>Telco options</strong></td>
<td>None</td>
<td>NS2200T: -48V DC (commercial rack) NS2200ST: -48V DC (seismic rack)</td>
<td>NB54000c-cg: -48V DC Carrier Grade (seismic rack) NEBS Level 3 Certified</td>
</tr>
</tbody>
</table>
**HP Integrity NonStop**

- from: “HP NonStop Advanced Architecture A business white paper”

Figure 2. The NonStop Blade Complex using NonStop Advanced Architecture for Integrity NonStop servers

The NonStop Blade Complex within Integrity NonStop servers, with both Dual Modular Redundancy (DMR) and Triple Modular Redundancy (TMR) configurations, is shown in figure 2.

* Note: LSU = logical synchronization unit
HP Integrity NonStop

- **XP12000 Disk Array**
  - *Reliable*: Provides extreme reliability with no single point of failure
  - *Available*: All components are redundant and hot-swappable
  - *Scalable*: Scalable to 165 TB of internal capacity, 14 PB of external capacity, and 128 GB of cache
  - *Performance*: Uses an advanced crossbar fault-tolerant architecture, resulting in outstanding levels of random and sequential I/O operations for database and OLTP workloads
  - *Heterogeneous*: Supports multiple operating systems
  - *Open Connectivity*: Supports Fibre Channel, FICON, and ESCON connectivity
  - *Flexible*: Supports a mixture of disk drives configured as RAID 1 and RAID 5