IFISC

NUREDDUNA
IFISC's Computational Cluster

Nuredduna is the main computational infrastructure of IFISC. It is a GNU/Linux based computer cluster designed and built at IFISC to perform simultaneously a large number of scientific computations as quickly and efficiently as possible (High Throughput Computing).

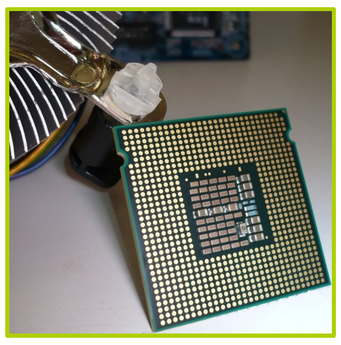
The cluster is made with out-of-the-shelf components from the personal computer market. A choice which offers, probably, the best performance/price ratio in intensive computation

Nuredduna is configured in a fully transparent way for the user which have the same home directory as in all IFISC desktops. The user sees the cluster as a single Linux computer.

HARDWARE CONFIGURATION

• *Nodes (computational units)*

- Nuredduna has currently 120 nodes and a grand total of 332 computational cores. That makes it the biggest cluster at UIB campus.
- 74 nodes with Intel Core2 E6600 dual-core processors running at 2.4 Ghz on Intel DQ965GF and Asus P5B-VM microATX motherboards with Intel 965 chipsets
- 46 nodes with Intel Core2 Q6700 quad-core processors running at 2.66 GHz on a Gigabyte GA-G33M-DS2R microATX motherboard with Intel G33 chipset



- The amount of memory of each node ranges from 2 to 8 GB of DDR2 running at 667 MHz, for a total of about 280 GB of RAM
- Nodes are diskless. That implies less power consumption, less heat generation, more reliability and less cost.
- Nodes are boxed in slim form factor cases Aopen H360 with 300 W PFC power supply.

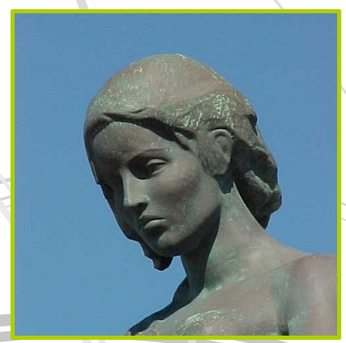
•*Communications*

- We use three 48-port Gigabit Ethernet switches. One HP Procurve 2448 and two D-Link DGS-1248T. Each of them have a switching capacity of 96Gbps.
- Switches are connected among them through the IFISC Vlan on the campus network.
- Node 1 is directly connected to the IFISC Vlan.
- The other nodes are connected to one of the switches using the Gigabit Ethernet interface integrated on the motherboard.
- The disk server Dolsa is also directly connected to the IFISC Vlan through two Ethernet cables in channel bonding configuration, providing higher speed and more reliability.

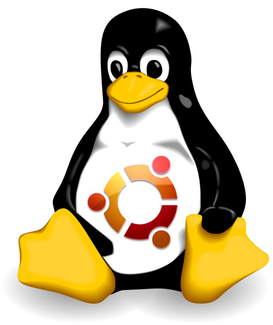
- The whole cluster is located on floor 0 of the “Instituts Universitaris” building, in a room specially designed for it, including technical floor for cabling. We have a 40A triphasic electrical supply backed up by UPS and a power generator. The cooling capability is 20000 frigories

WHY “NUREDDUNA”?

Nuredduna is the Mediterranean counterpart of the anglosaxon myth on the warrior Beowulf. She is a fiction heroine in the poem "La deixa del geni grec" (1900) by the Majorcan writer Costa i Llobera. The character is inspired on a prehistoric epic story during the first invasions of the island of Mallorca by the greeks. Granddaughter of a Highest Priest, Nuredduna was the Sybil of a talayotic tribe in the caves of Arta. Unlike Beowulf's relying in his physical force, her strength was love. Giving her life to save Homero's, Nuredduna became the symbol of the union of Mallorca with the foreign greek civilization.



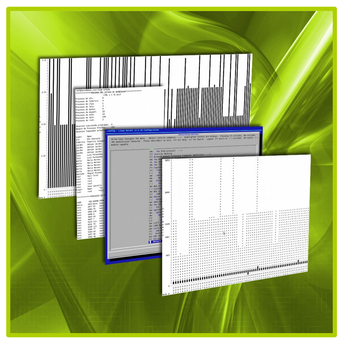
SOFTWARE CONFIGURATION



• *Operating System*

- Operating system for nuredduna is Ubuntu Server 64 bit edition with the last available Linux kernel patched with MOSIX.
- All nodes boot via the Ethernet network and load the Operating System from the IFISC disk server dolsa, so all of them share the same settings.
- Network is set up so all nodes can be accessed directly from any computer inside IFISC VLAN.
- Users have the same home directory as in all the other IFISC computers including desktops and servers.

• *MOSIX*



- MOSIX (http://www.mosix.org) is a patch for the Linux kernel plus some utilities running in the user space that enhance the Linux kernel with cluster computing capabilities.
- Trough that software layer, users see the cluster as a single Linux computer.
- MOSIX allows applications to run in remote nodes, as if they where running locally:
  - Kernel patch intercepts all system-calls.
  - Migrates processes without interrupting the calculation.
  - For migrated processes, most of the system-calls are forwarded to home-node, performed there, and results are sent back to the process.
  - A few system-calls can be performed directly on the remote node.
- MOSIX distributes the computational load among the nodes dynamically:
  - Each node calculates its load: # programs / (#cores*speed).
  - Every few seconds each node compares its load with another one randomly.
  - The load of these two nodes is balanced by migrating one or several programs.
  - Over the time this averages the load automatically without any “central authority”.
- If one node has to be shut down, MOSIX first migrates to another node the programs that have been sent there. That eases maintenance tasks.

NUREDDUNA DIAGRAM