Microserver architectures and system software

With the end of Moore's Law, hardware architectures are becoming increasingly complex. We research low-level software mechanisms at the boundary between system architecture and system software to mitigate this complexity.

Summary

Hardware architectures are becoming more complex, with heterogeneous multi-core processors, accelerators and reconfigurable logic, and deeper memory hierarchies. Techniques for energy efficiency, including near threshold, DVFS, and energy-efficient interconnects bring additional complexity and timing variability. Manual optimization of the data layout, placement, and caching will become uneconomic and time consuming, and will, in any case, soon exceed the abilities of the best human programmers.

Runtime systems and system software must hide or mitigate this complexity with more and more intelligence across the programming environment. There needs to be a change in mentality from programming “heroism” towards trusting the compiler and runtime system (as in the move from assembler to C/Fortran).

This group is working on multiple levels of the low-level software stack, from the hypervisor to the run-time system. Our main research lines are:

- Hypervisor support for shared memory resources
In **EUROSERVER**, we are using Xen's transcendent memory (tmem) abstraction to share memory across multiple coherence islands.

- **Energy proportional interconnects**
  In EUROSERVER and Mont-Blanc 2, we are researching low-level techniques for energy proportionality in the interconnects, to reduce the energy consumption without incurring a significant performance overhead.

- **Run-time data management**
  In ExaNoDe, we are extending the OmpSs programming model and Nanos+Clusters run-time system to manage shared memory with coherence islands.

---

**Objectives**

- Investigate energy proportionality, including intelligent power management for Ethernet and InfiniBand
- Provide hypervisor support for programming model abstractions, shared resources and isolation
- Run-time system support for novel architectures

---

Barcelona Supercomputing Center - Centro Nacional de Supercomputación