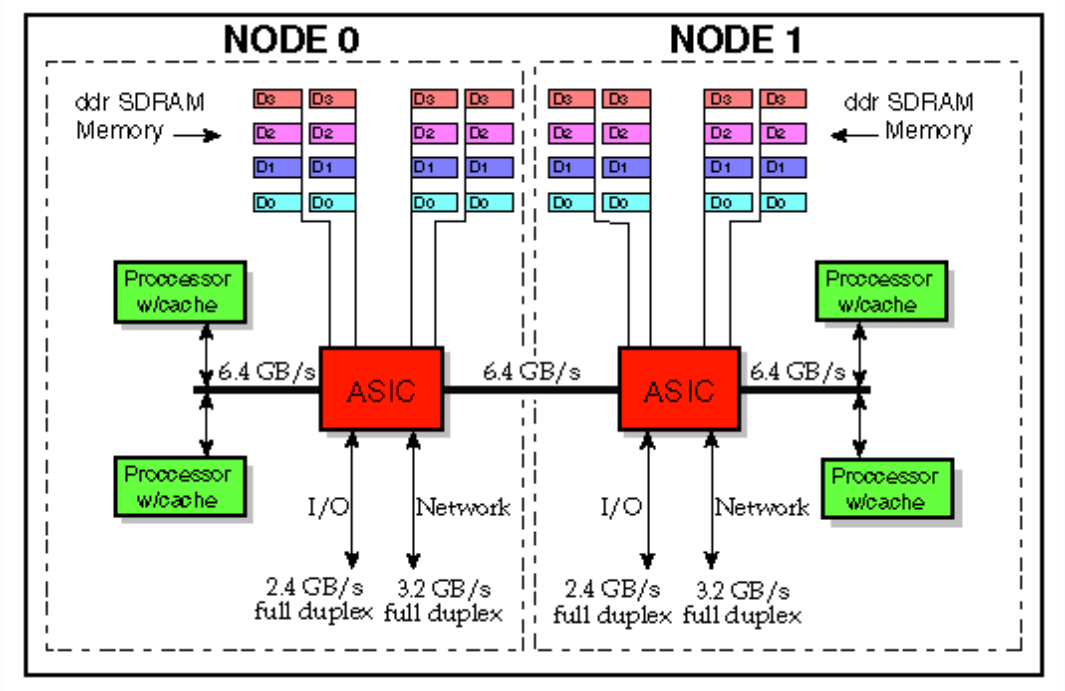
Survey on Shared Memory Multiprocessor Architecture

Manh Huynh, Rohan Dawson
University of Colorado Denver

I. SGI Altix 3000 Global shared-memory architecture

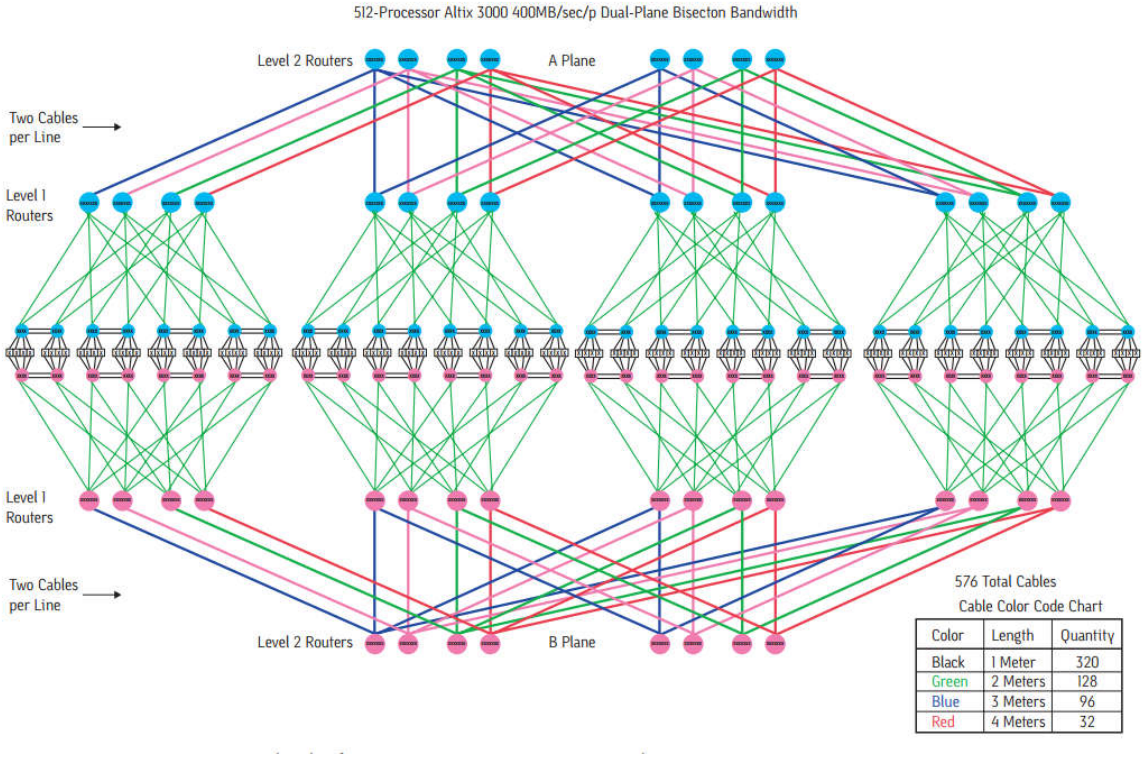
1. **System overview**: SGI Altix 3000 is cache-coherent and shared memory multiprocessor supercomputer system. Altix systems have different versions and all of them are mainly designed for scientific and military applications. The main features of Altix 3000 is consisting of 128 Itanium 2 64-bits processors, and 256GB of global memory. It is based on Distributed Shared Memory architecture, where memory is physically distributed among 32 Computation-brick ("C-bricks"). Each of computation brink consists of 2 pairs of 1.3Ghz Itanium 64-bit. Altix 3000 class computers provide global, cache-coherent and NUMA shared memory. The system is built in 2006.
   The interesting feature of this architecture is NUMAflex using SGI NUMNA protocol implemented directly in hardware using ASIC. This controller can be scaled up independently in processor count, memory capacity and I/O capacity. The figure below shows the architecture of one C-brick consisting of 4 processor connected to 32GB of memory.



   The NUMAflex allows to connect up to 512 processors, each can support 6.4 GB per second of bandwidth.

2. **Interconnection Network**

   NUMAflex network for Altix is configured in a fat tree topology. It provides dual plane (parallel networks) for increased bisection bandwidth. The topology is shown as below. All square in the middle is processor, circle is router (R-brick). T



3. **3. Other features:**

   Operating system: Linux

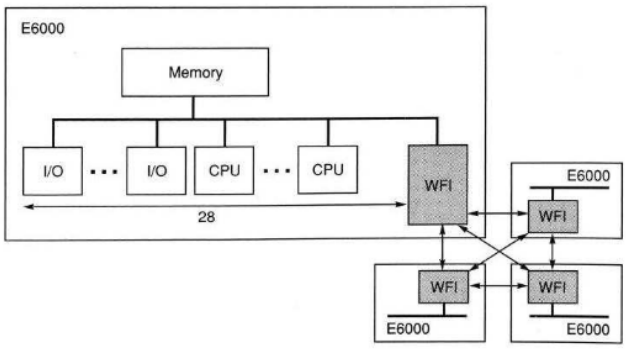
   Programming language: XPMEM API consist of a library that abstracts the underlying kernel module. We also have SGI Message Passing Toolkit. The code sample is given on the paper [1].

4. **Conclusion:**

   There is certain features that we may know before such as NUMA, however the parallel interconnection network is interesting that we haven't mention in class. In overall, it is said that this altix 3000 is powerful server and supercluster architecture for HPC applications.

II. Sun Microsystem's Wildfire architecture

WildFire is the result of taking existing Symmetric Multiprocessing (SMP) technology, and trying to make it scalable. It achieves this by linking 1-4 compatible SMPs (Sun Microsystems E6500/E5500/E4500/3500) together through a WildFire Interface (WFI) board so that the OS can treat them as one single unified system. The total number of cores and memory depends on the loaded components, but the WildFire supports up to 112 UltraSparc processors. Each processor is linked to a shared portion of memory, as seen in the diagram below.



Although the default behaviour is to maximise performance by keeping a given process to a given node, the hardware has the capability to allow a single process to span as many of the nodes as it needs to.

The architecture operates on a modified version of Solaris 2.6, and it does not appear to be restricted to any particular language. It was built in 1997 as a feasibility study into distributed shared memory systems, and as such is not still in use today.

The interconnection network within each SMP is a single long bus that connects all the processors, but the connectivity between SMPs of course happens directly through the WFI.

Although the full WildFire architecture is not too similar to the examples from the lectures, a strong comparison can be made between a single SMP module instead.

References:

[1] Woodacre, Michael, et al. "The SGI® AltixTM 3000 global shared-memory architecture." Silicon Graphics, Inc. (2005).

[2] https://en.wikipedia.org/wiki/Altix

[3] http://sc.tamu.edu/help/altix/architecture.shtml

[4] E. Hagersten and M. Koster, "WildFire: a scalable path for SMPs," *High-Performance Computer Architecture, 1999. Proceedings. Fifth International Symposium On*, Orlando, FL, 1999, pp. 172-181.

http://www.cs.utah.edu/~rajeev/cs7820/papers/hpca99.pdf