4DM4 Assignment # 1, 2013
Exascale Computing

Posted: Wednesday Sept. 18, 2013
Due: Monday, Sept. 30, 2013

(Assignment #1 electronic submissions are due on the day specified, by 6 pm.
Late penalty = 15% per day.
Solutions are discussed in the tutorial, so submissions are due before the tutorial.)

Many governments worldwide have started programs to achieve Exa-scale computing by 2020, including the USA, China, Europe, Japan, and India. The goal is to achieve a sustained performance of 1 exaflop per second. The India givernment aims to achieve 132 Exaflop performance by 2017. A 'Flop' is a floating point operation, on a 64-bit number. Here is a table of performance scales:

1 Gigaflop = 10^9 flops/sec = 1 billion flops
1 Teraflop = 10^12 flops/sec
1 Petaflop = 10^15 flops/sec
1 Exaflop = 10^18 flops/sec = 1 billion Gigaflops

A true exaflop-scale machine will achieve a performance of 1 exaflop on several different benchmarks, i.e., a very large FFT, a very large sorting problem, a very large N-body problem, etc. Each year the world's most powerful machines are posted at this web-site:

https://www.top500.org

The goal of this assignment is to do some 'back-of-the-envelop' calculations, to estimate the scale, size, and power requirements of such a machine. Assume you can use AMD or Intel processors. You can find a comparison of processors, performance, and power at the web-site at the end of this assignment. Assume you can use IBM Ethernet switches. You can find data on switches at the web-site at the end of this assignment.

The Algorithm: Bubble Sort

We will use a much simpler Bubble-Sort algorithm, where the communications requirements between processors is minimized. In essence, we are building a special-purpose exascale computer catering to bubble-sort, rather than a fully programmable exascale computer.

A bubble-sort of N=8 numbers is shown in the next fig. There are N/2=4 iterations of processing, where each iteration has 2 stages of processing. This figure shows that after N/2 iterations, the largest number has moved to its final position. Some thought will establish that the final sequence is sorted.

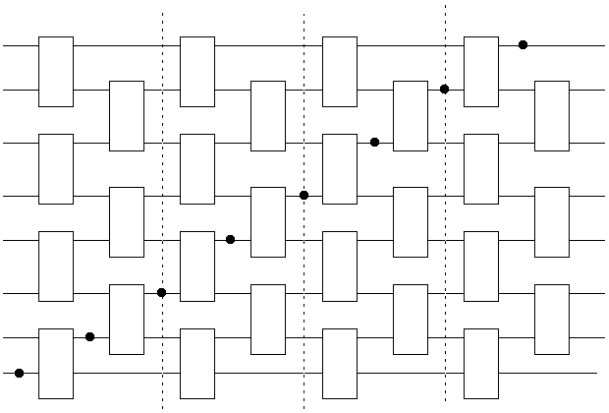

Fig. 1. Bubble sort.

In this assignment, assume 8 ADM or Intel processors can fit on one PC board (printed circuit board). Each board can have several 10 Gbit Ethernet ports for data IO (Input/Output). Assume parallel data arrives from the outside world, on many 10Gb Ethernet ports. The data can arrive in waves, so that the above bubble sort hardware can be pipelined. Multiple waves of data can reside in the hardware at the same time. Each wave eventually exists and is sorted. Assume parallel sorted data is sent back to the outside world, on many 10Gb Ethernet ports.

You can pick the number of PC boards, the number of processors, the types of processors, the number of switches, the types of switches. Please provide the data-sheet for any component you select. You can change the sorting algorithm to use a different pattern of interconnections for bonus marks (but be forewarned: an optimal parallel sorting algorithm, which can sort N numbers with O(N) processors in O(logN) time, has never been found.) You can use the sorting 'bubble' algorithm shown above to sort batches of numbers, rather than single numbers. If you try this, then clearly explain your algorithm.

Please explain your design clearly. Explain the number of processors, the type of processors, the peak GFLOP/sec per processor, the power per processor.
Explain the number of PC boards, the number of 10Gb Ethernet switches, the total power for processors (before cooling considerations), and the power for switches (before cooling considerations).

There is considerable design freedom here. You will have to make some design decisions and proceed, the same as all the other engineers working on these problems around the world. You may have to make some simplifying assumptions. Please make sure to clearly state all assumptions.

------- URLs ----------
You can find a comparison of AMD processors, performance, and power at this web-site:

http://sites.amd.com/us/business/products/server/server-selector/Pages/AMDServerComparisonTool.aspx

You can find data on IBM rackswitches (Ethernet Switches) at this web-site:

http://www-03.ibm.com/systems/networking/switches/rack.html