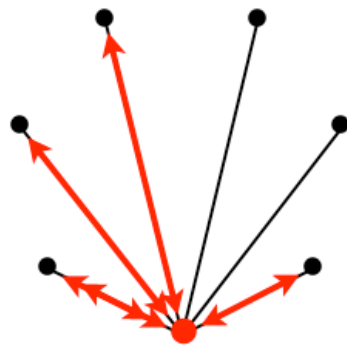
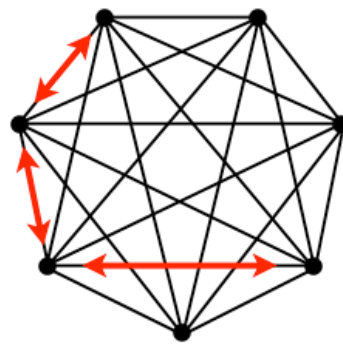


SEM & gridMathematica comparison - pingpong benchmark

The Supercomputing Engine for *Mathematica* (SEM) is a technology inspired by the way that modern supercomputers perform work. SEM supports an “all-to-all” communications network between *Mathematica* kernels, much like a complete telephone network. It is well known in high-performance and scientific computing that the communications network ultimately limits the addressable problem size, so those practitioners opt for the most capable communications network available.



gridMathematica's topology

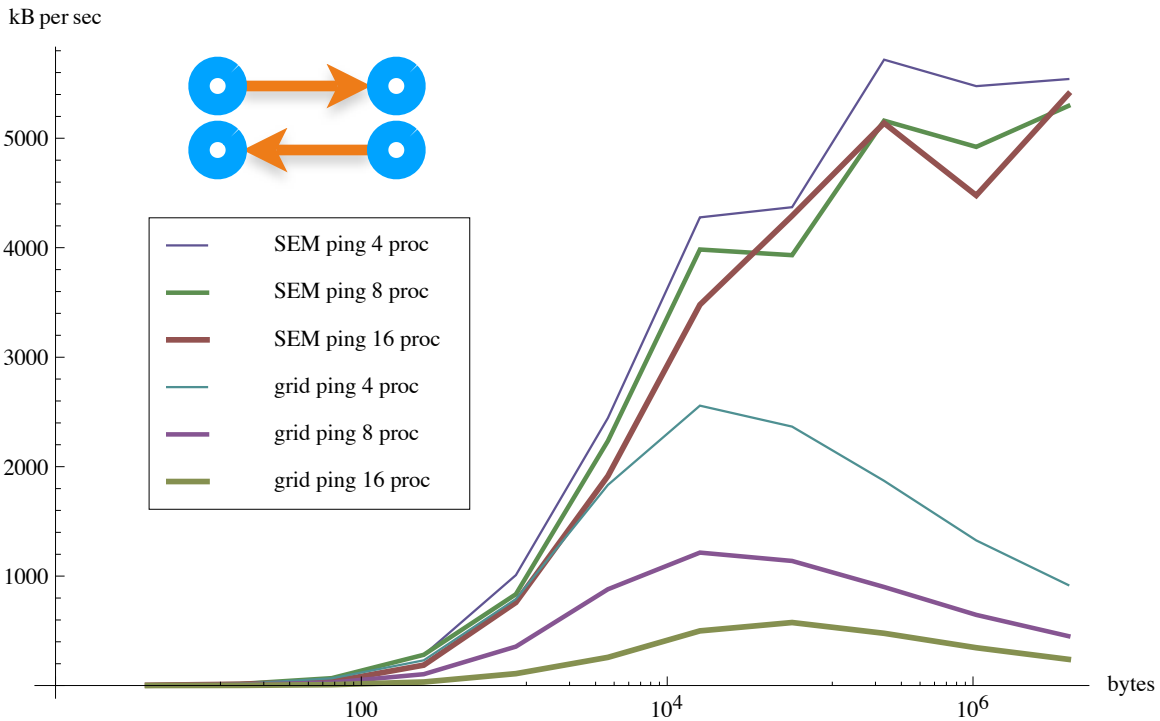


SEM's topology

gridMathematica's major distinction from the approach common to supercomputers is that only one *Mathematica* kernel can communicate directly with all other kernels. Its “master-worker” approach limits gridMathematica's usefulness to a subset of what SEM can do, particularly for solutions where different parts of the calculation communicate. We were asked to provide a way to compare how these technologies' capabilities are different.

We drew from a high-performance computing benchmark called “pingpong”, developed by Viktor K. Decyk of the UCLA Plasma Physics Group. It stresses the network of a parallel computers in two stages after dividing the processing elements into the even ones (0, 2, 4, ...) and the odd ones (1, 3, 5, ...). The pingpong stage has each even one send a message to an odd one (0 to 1, 2 to 3, ...) and back again. The message is varied in size from a few bytes to many megabytes while each operation is timed and averaged. The second stage performs a “swap”, meaning each pair of processes sends and receives a message from each other simultaneously, if possible. Again the message size is varied and the results are timed and averaged.

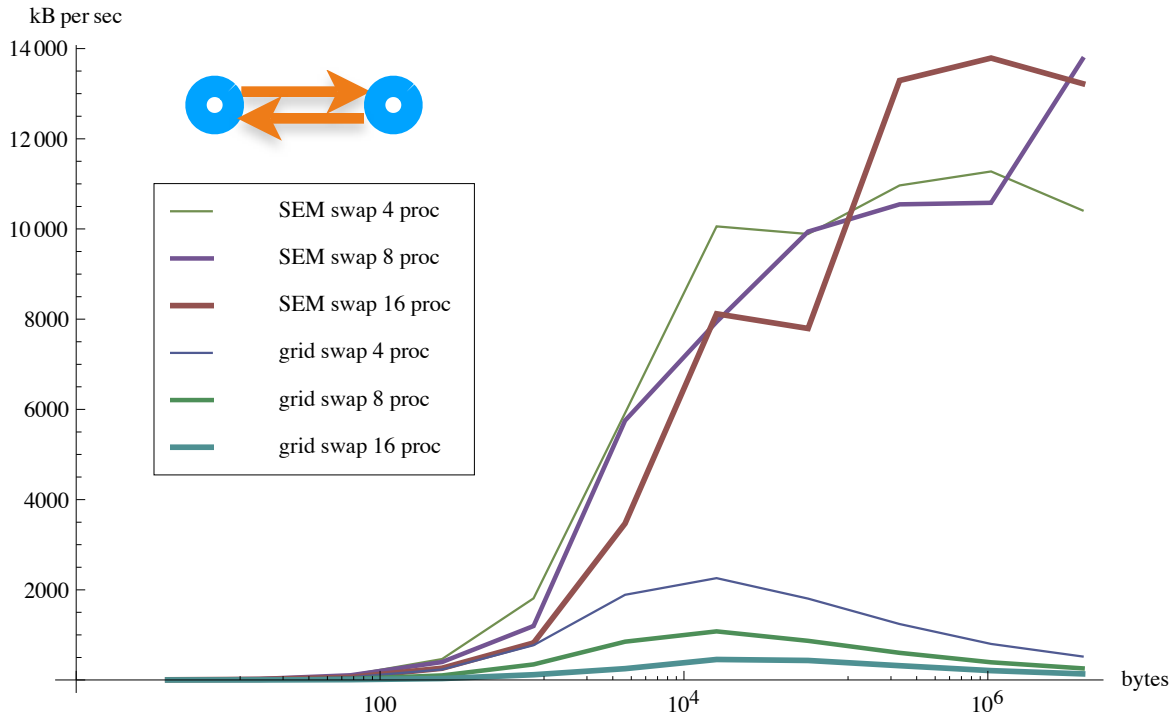
Below is the result comparing the pingpong test for SEM and gridMathematica.



The horizontal axis is message size in bytes, and the vertical axis is bandwidth in kB/s. This test shows 4-, 8-, and 16-processor cases. When communications does not matter for the computation, that corresponds to where the message size is small in this figure. However in most problems communications is important, for example when message sizes are large (the right side of the figure), the SEM plots show higher performance than the gridMathematica plots.

Also note that the gridMathematica performance for 8 processors about half that of its 4-processor case, then half again for 16 processors. This is explained because there is still only one processor (or kernel) that acts as “traffic cop” for all messages on the system. It can only handle so much bandwidth by itself, so if one were to go to 32 or more processors the performance of gridMathematica will be still poorer.

The following graph shows results of the swap test comparing SEM and gridMathematica.



We show 4-, 8-, and 16-processor tests here. The difference is more dramatic here because the “master” in gridMathematica can only handle one message at any moment. We see a similar trend here. When communications is involved, gridMathematica decreases in performance with increasing processor count. The single master is a bottleneck to data. This is a consequence of the fundamental topology of gridMathematica’s “master-worker” approach that is not experienced with proper use of SEM. In fact, for large message sizes, SEM performance does not degrade in this test with processor count.

We used two 8-core Intel Xeon Mac Pros connected with a Netgear GS116 Gigabit Switch produce these 16-core results, but future hardware will use more and more processing cores. The most recent evidence is that Intel recently announced the new Nehalem chip, which can have up to 8-cores on a single chip. Hardware makers will likely use Nehalem to produce 16-core machines, up from 8 last year. This trend is clear and is commonly discussed by the industry, not only by pundits but by Intel, AMD, and IBM. Our recommendation is to get ahead of this curve and anticipate how to best address these issues. We plan to update this benchmark test on future hardware, but our conclusion is that performance available via SEM exceeds, and over time will increasingly exceed, that of gridMathematica.