BLASTing Off With Green Destiny

Aaron Darling and **Wu-chun Feng**
Computer & Computational Sciences Division
Los Alamos National Laboratory

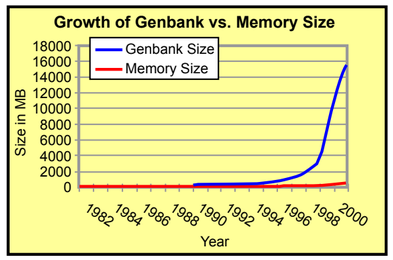
What is BLAST?

- Sequence database-search program that looks for similarities between a query sequence and a large database of sequences.
- Computationally intensive algorithm that is fundamental to bioinformatics.

Why Should BLAST Be Parallelized?

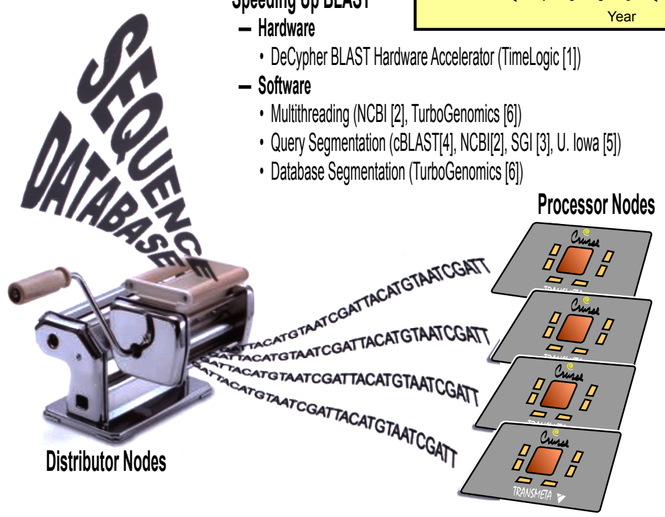
- Sequential algorithm is long running but embarrassingly parallel.
- Sequence databases are growing exponentially in size.
- Sequence databases are usually larger than a single node's memory size, thus causing a lot of disk I/O.

Size in MB	DB name	Description
5700	nt	non-redundant nucleotide DB
2200	Human EST	Human expressed sequence tag DB
1100	Mouse EST	Human expressed sequence tag DB
510	nr	non-redundant amino acid DB



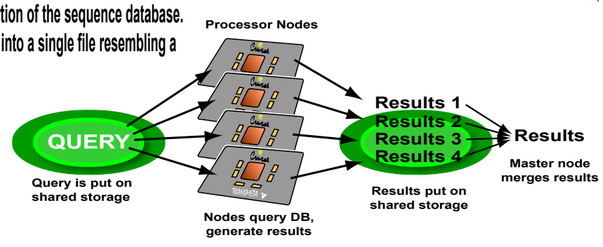
Speeding Up BLAST

- Hardware
  - DeCypher BLAST Hardware Accelerator (TimeLogic [1])
- Software
  - Multithreading (NCBI [2], TurboGenomics [6])
  - Query Segmentation (cBLAST[4], NCBI[2], SGI [3], U. Iowa [5])
  - Database Segmentation (TurboGenomics [6])



Our Approach: Combine Multithreading and Database Segmentation

- Distribute a portion of the sequence database to each cluster node.
- Each cluster node searches a query against its portion of the sequence database.
- Results are reported to a master node and merged into a single file resembling a standard BLAST output.
- Benefit
  - Database fragments are small enough to stay in the buffer cache, eliminating disk I/O.



Our MPI Implementation

- Format, segment, and distribute sequence databases amongst nodes in a cluster.
  - Execute an MPI wrapper for the standard BLAST formatdb called mpiformatdb.
- Submit BLAST queries.
  - Execute an MPI wrapper for the standard blastall program from the NCBI BLAST distribution.
- Aggregate the "search results" files from each cluster node into a single user-specified output file.

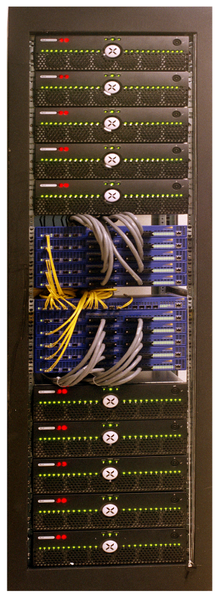
Availability

- Open-source distribution based on MPI [10].

Testing Platform

Green Destiny: A 240-Node Bladed Beowulf Cluster in One Cubic Meter [7-9]

- First cluster prototype of the "Supercomputing in Small Spaces" project (http://sss.lanl.gov)
- Compute Node
  - 667-MHz Transmeta TM5600 (Intel-compatible), 640-MB RAM, and 20-GB hard disk.
- Network Interconnect
  - 100-Mb/s (Fast Ethernet)
- Operating System
  - Linux 2.4.x
- Selected Press Coverage
  - "At Los Alamos, Two Visions of Supercomputing," The New York Times, June 25, 2002.
  - "Bell, Torvalds Usher Next Wave of Supercomputing," CNN, May 21, 2002.

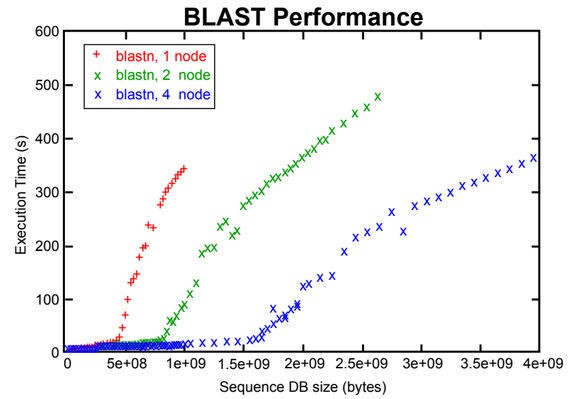


SUPERCOMPUTING in SMALL SPACES

Los Alamos NATIONAL LABORATORY

Traditional BLAST vs. mpi-BLAST

- As database size increases → disk I/O increases → performance decreases.



- mpi-BLAST run time for 300 kB query against nt:

Nodes	Runtime(s)	Speedup over 1 node	Speedup/Nodes ratio
1	80774.93	1.00	1.00
4	8751.97	9.23	2.31
8	4547.83	17.76	2.22
16	2436.60	33.15	2.07
32	1349.92	59.84	1.87
64	850.75	94.95	1.48
128	473.79	170.49	1.33

Future Work

- A more efficient mpirun implementation, which distributes the BLAST queries to the compute nodes.
  - Current implementation severely affects performance despite the 170-fold speed-up over 128 processors.

References

1. A. Shpuntof, AFS Informatics, Christopher Hoover, TimeLogic. *Personal Communication*, Aug. 2002.
2. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. "Basic Local Alignment Search Tool." J. Mol. Biol. 215:403-410, 1990.
3. N. Camp, H. Cofer, and R. Gomperts. "High-Throughput BLAST," *SGI White Paper*, Sept. 1998.
4. cBLAST. http://wwwcbbc.murdoch.edu.au/research/cluster/.
5. R.C. Braun, K.T. Pedretti, T.L. Casavant, T.E. Scheetz, C.L. Birkett, C.A. Roberts. "Parallelization of Local BLAST Service on Workstation Clusters," *Future Generation Computer Systems*, 17:745-754, 2001.
6. R.D. Bjornson, A.H. Sherman, S.B. Weston, N. Willard, J. Wing. "TurboBLAST: A Parallel Implementation of BLAST Based on the TurboHub Process Integration Architecture," From http://www.turbogenomics.com/products/turboblast.pdf
7. W. Feng, M. Warren, E. Weigle, "The Bladed Beowulf: A Cost-Effective Alternative to Traditional Beowulfs," To appear at *IEEE Cluster 2002*, Sept. 2002.
8. M. Warren, E. Weigle, W. Feng, "High-Density Computing: A 240-Node Beowulf in One Cubic Meter," To appear at SC2002: *High-Performance Networking and Computing Conference*, Nov. 2002.
9. G. Johnson, "At Los Alamos, Two Visions of Supercomputing," *The New York Times*, June 25, 2002.
10. W. Gropp, E. Lusk, A. Skjellum, "Using MPI: Portable Parallel Programming with the Message Passing Interface," 2nd edition, 1999.