Introduction to the BEOWULF Design

Robert G. Brown
Duke University Physics Department

Parallel Computation

- Many kinds of useful work can be done in parallel and be completed faster. Doing work in parallel also allows one to do more than one can tackle alone. "'Many hands make light work."
- The same is true for work done on computers. Some kinds of things can be done much faster or more thoroughly if done in parallel on many computers at once instead of on one computer.
- However, certain things need to be true for this possibility to become a reality.
- One needs an operating system capable of chewing gum and talking at the same time. Figuratively speaking, of course.
- One needs a communications mechanism (a network) between the participating computers.
- One needs various tools to actually write a parallel program and make it all work, or alternatively one needs a program written by others to execute in your parallel environment.

IF these are all available, it is possible to build a parallel supercomputer out of commodity parts – desktop computers, cheap networks. In some cases, the "supercomputer" can be nothing more complicated than all of the existing desktop computers in a department working together.

They are available, and have been for some time. The name of the design is the Beowulf. It was invented (by this name) by Don Becker and Thomas Sterling at Nasa Goddard (CESDIS) about five years ago, although many others contributed before and since. A beowulf is:

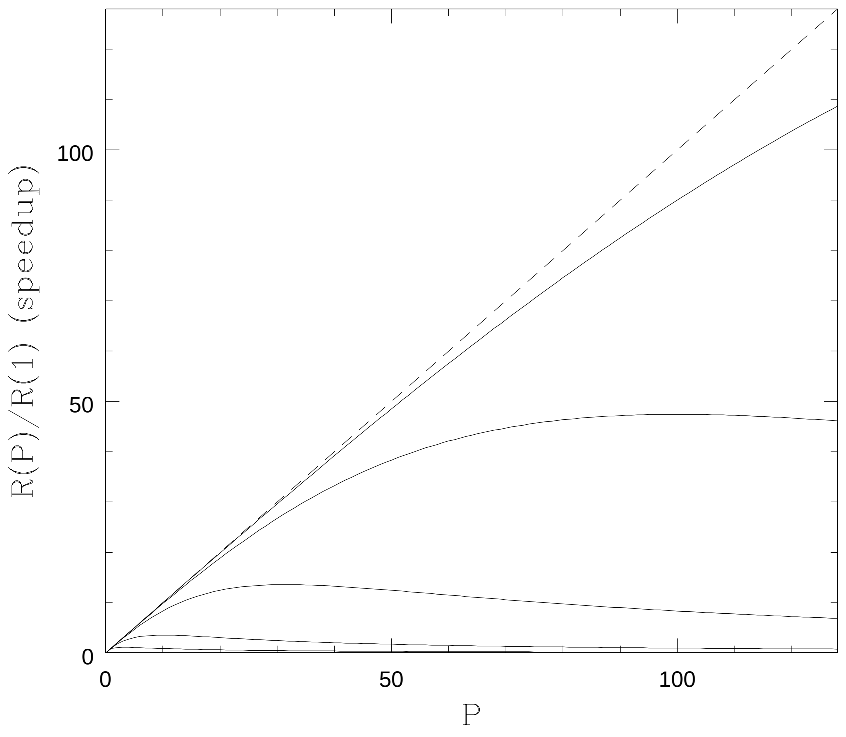
- **Cheap. Anywhere from "free" to a few tens of thousands of dollars.**
- **Powerful. As powerful as an SP2 or SP3 costing ten times more.**
- **Useful.**

What Can One Do with a Beowulf?

Many things of interest to law enforcement or security personnel:

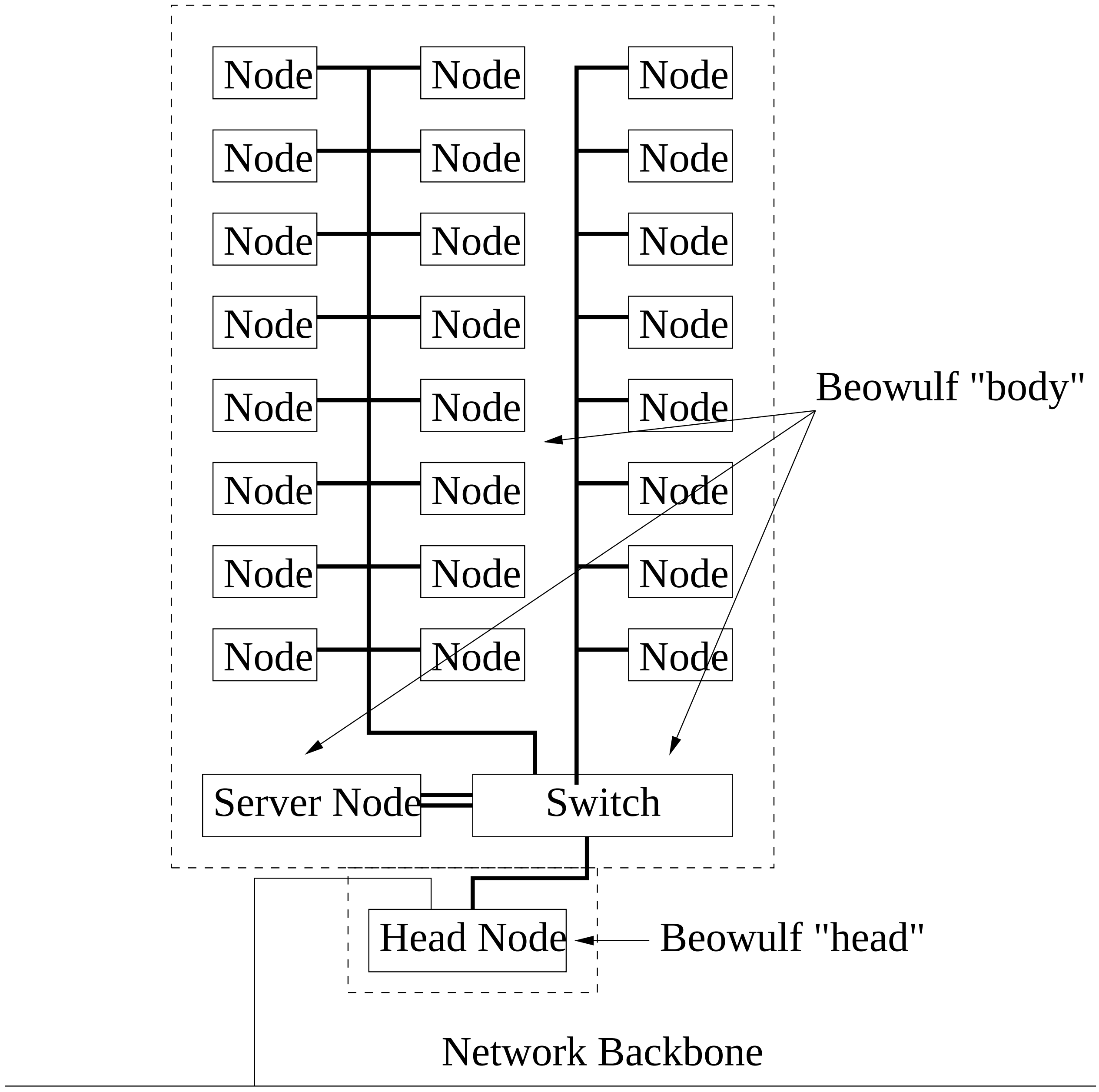
- Crack passwords on seized machines or on machines being secured (to test their quality).
- Decrypt (probably poorly) encrypted material. Note that there are certainly plenty of encryption schemes that will defy any effort to break them, but criminals are not known for being smart...
- Build advanced predictive models (using e.g. neural network technology designed to run on beowulf-class supercomputers).
- Perform other computationally intensive operations that can be split into many mostly independent parts.

An Important Figure

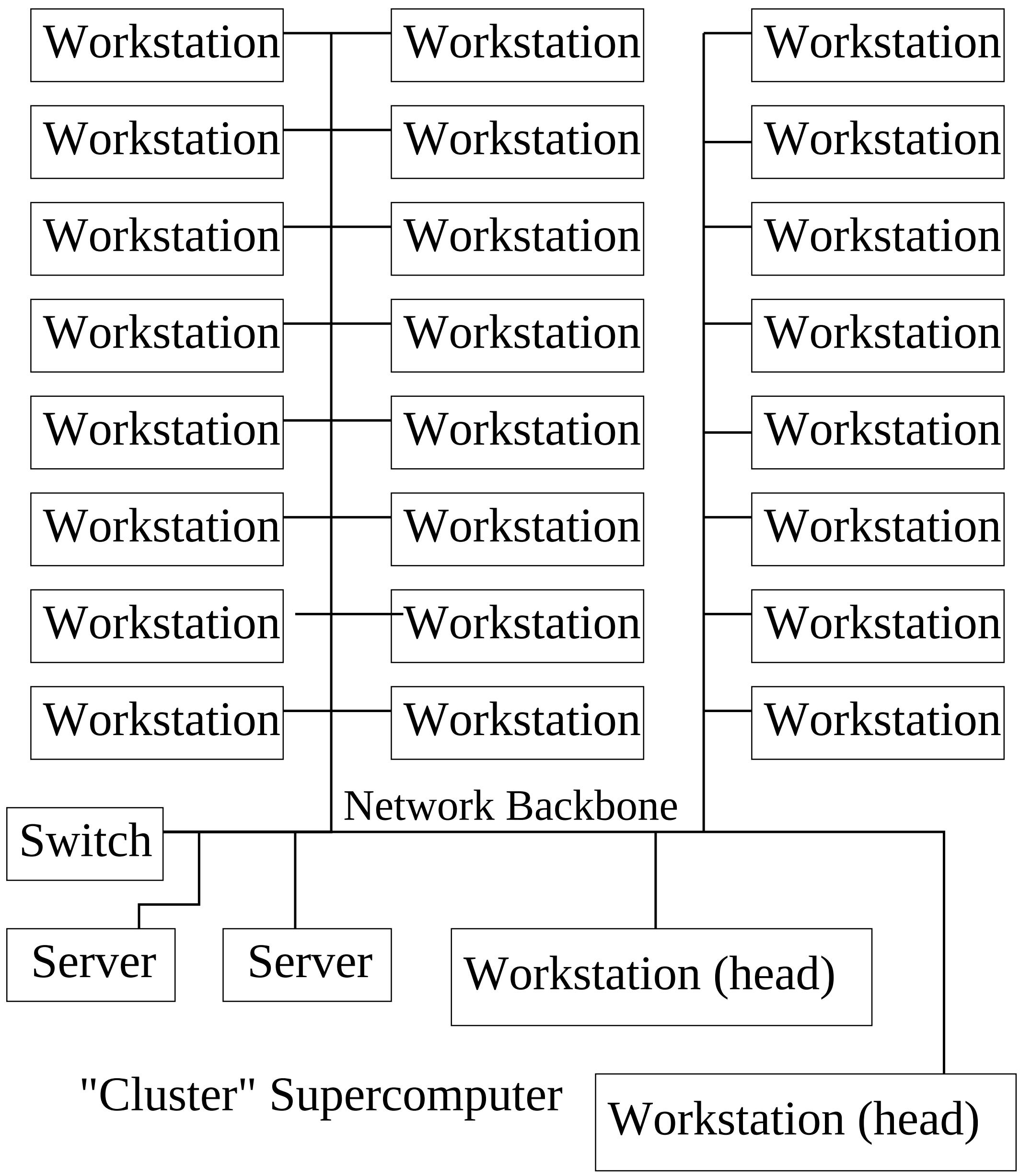


For various ratios of parallel work done to communications time. Law enforcment type parallel applications are mostly "embarrassingly parallel", and get the *best* speedup scaling (with benefit from as many as hundreds or even tens of thousands of nodes, like the RC5 DES decryption project).

Schematic for a Beowulf



A Workstation Cluster



Recipe for a Beowulf

For the kinds of problems likely to be encountered in law enforcement and security:

- Obtain as many inexpensive ($500-1000) PC's as you can afford. No need to get monitors, keyboards, mice.
- Get a fast ethernet switch with as many ports as you have systems (or more).
- Put Linux on each node. No, WinXX won't do (remember about chewing gum and talking).
- Put the parallel software and libraries on each node.
- Connect them together.
- Run your application(s).

This procedure is more likely to have a happy outcome if one buys e.g. "How to Build a Beowulf" by Sterling, Salmon, Becker and Savarese and possibly my book, or obtains a (still cheap) turnkey beowulf from any one of many vendors.

Surprise!

If one is running Linux within one's department on systems on a mutual network, they can *already* **function like a "beowulf" on e.g. password cracking or related problems. Linux can easily handle desktop usage and background computation.**

Linux also (almost) never crashes. No more "blue screen of death"!

Conclusions

- Beowulfery is a cheap, efficient way of applying supercomputer class power to certain law enforcement problems.
- Easy to do "at home" (I'm showing you my home beowulf)!
- Lots of commercial and consultative support. FREE consultative support on the beowulf mailing list.
- Demos of password cracking, neural network pattern matching, and a "fun" parallel application can be demonstrated with the small beowulf here.
- http://www.beowulf.org
- http://www.phy.duke.edu/brahma
- Contact: Robert G. Brown, rgb@duke.edu