
Subject: Re: Mac Clusters & IDL

Posted by [Marshall Perrin](#) on Sat, 12 May 2007 17:36:11 GMT

[View Forum Message](#) <> [Reply to Message](#)

Trae <traewinter@gmail.com> wrote:

- > I do a lot of numerical simulation and image rendering in IDL. This
- > is amazingly CPU intensive and I've been using the new parallel
- > processing capabilities of IDL with good success. (It might not be
- > the most efficient, but man is it easy!)

One question you will need to think about is whether the various parallel nodes in your cluster need to talk back and forth a lot, or whether they can each mostly compute independently, passing only a relatively small amount of data between nodes. This will influence the sort of computing topology that works best for your problem. I suspect that for your case - image rendering - you may often be in the case where each node is working entirely independently to render a different frame or scene (often called "Embarassingly parallel" computing.)

- > I was wondering if anyone on this list has used IDL on a Mac cluster?
- > Specifically, a Pooch cluster?

I have never heard of Pooch before. Apple provides its own clustering technology, Xgrid, which can be used to set up clusters, geared mostly towards the case where each node can work independently. We have an smallish Xgrid cluster (five dual-CPU machines) used for a variety of projects, in my case for running a Monte Carlo radiative transfer code. I used the Xgrid system to distribute jobs to the various nodes, and that works fine, but the nodes don't communicate at all to each other while running the job. More complicated approaches are possible, but just aren't necessary for my particular problem.

- > There are some good threads on other
- > forms of clusters. Mac clusters seem easy to make but expensive.
- > Linux clusters seem to be cheaper hardware wise but harder to keep
- > running.

My personal opinion is that most of the cost savings of Linux compared to Mac are an illusion. Don't look at hardware costs only. Speaking as a former grad student myself, you've got better things to do with your time than administer some cluster (like write papers!). I haven't tried to set up any Linux clusters myself, but from what I've heard from those who have, the learning curve is still pretty steep - and there are many different approaches and software options to sort through along the way.

First off you need to figure out what sort of cluster network topology suits your particular computing needs. Then, if possible, you need to talk your advisor into hiring someone, even just temporarily as a consultant for a few days, who knows what they're doing and can guide you through the setup process. We tried to go the grad-student-sysadmin route for a while, and it was sort of OK, but once we hired a part-time actual sysadmin, things became *much* smoother and more reliable. And the grad student in question suddenly had more time to actually do research! Funny that.

- > Mac cluster? How well does the native IDL multi-threading work on a
- > Mac cluster? Has anyone used IDL_IDLBridge on such a cluster?

Multithreading won't work across different machines in a cluster - all the threads have to run on the same computer. I haven't tried the IDL_IDL bridge so I can't comment on that.

- Marshall

Subject: Re: Mac Clusters & IDL
Posted by [Brian Larsen](#) on Wed, 23 May 2007 20:57:06 GMT
[View Forum Message](#) <> [Reply to Message](#)

Trae,

as another former grad student I would recommend more long walks (preferably with your wife) while you let the code run on a single machine. Any setup you do on computers is a huge one time investment with little payoff for you unless you are planing on sticking around MSU for long enough to make up for the investment. I think the formula is pretty simple

$$\frac{[(\text{decrease in run time}) * (\text{number of times you run it})]}{[(\text{time required to setup and maintain the system}) + (\text{time spent recoding})]}$$

The number on the bottom is large and according to Murphy always comes when you have posters and talks due so if the numerator isn't sufficiently big to make this ratio like 4 I wouldn't bother. Any guess at actual numbers is a guess as I have done lots of linux but very little mac, but you are easily looking at a meltdown once a month that sucks up one whole day, and setup and testing will take the better part of week. So over the course of 6 months this is easily 88 hours (40 + 6*8). So the savings in run time had better be 352 hours or more.

Well enough ranting, papers don't write themselves, spacecraft don't

build themselves, and certainly paperwork doesn't move from one side of my desk to the other by itself.

Cheers,

Brian

--

Brian A Larsen
RBSP-ECT Instrument Suite Scientist

Boston University
Center for Space Physics
725 Commonwealth Ave, Rm 506
Boston, MA 02215-1401
