

From Brent.Carlson@hia.nrc.ca Wed Nov 22 13:17 MST 2000
Date: Wed, 22 Nov 2000 12:16:35 -0800
From: B Carlson <Brent.Carlson@hia.nrc.ca>
Organization: National Research Council of Canada
X-Mailer: Mozilla 4.5 [en] (WinNT; U)
X-Accept-Language: en
MIME-Version: 1.0
To: Bill Sahr <bsahr@zia.aoc.NRAO.EDU>
Subject: Re: EVLA correlator back-end
References: <200011221842.LAA23740@mnemosyne.aoc.nrao.edu>
Content-Type: multipart/mixed; boundary="-----F078AAB127AACFC4C98A4BC0"
X-Lines: 113
Status: RO
Content-Length: 5158

Hi Bill:

> Food for thought (as opposed to thoughts of food) over the
> Thanksgiving weekend - 240 computers accepting data from the
> correlator baseline boards just doesn't seem workable. Some
> percentage of the computers would always be down, out-of-synch
> or some such problem. Is there any way to reduce the number ?
> Some scheme for "steering" the data from the LTAs? Some of
> our thoughts re the backend interface go in that direction,
> i.e. the possibility of onboard PPMC cards using an interface
> that is switchable, but we're not "there" yet.

I have been working on refined (i.e. more detailed) architecture diagrams for the 3 main boards in the correlator (Station, Baseline, Phasing Boards). I have produced two sets of drawings...a) one set with an FPDP interface and no processor on the board, and b) one set with a PPMC card and a Gbit FibreChannel interface.

Here's my current thoughts on the backend for both scenarios.

a) (FPDP interface). Commercial PC backends with multiple FPDP PCI cards in each PC. So, each PC can handle ~4 boards for a total of ~100 PCs (at \$2k each = \$200k). Is 100 networked PCs a problem? I don't know...I've heard of Beowulf clusters that are bigger...and how many millions of people use cheap and dirty PCs everyday? Except for the Gates-OS, they don't break down all the time. We have a Beowulf cluster here consisting of 16 commercial dual-Pentium 450 machines running Linux which runs like snot all the time. The only time it breaks down is when Willis does things like chews up all system memory or hammers it with some bizarre "DRAMA-ism". What happens when one PC does crash? Tony Willis says that the system should be able to be configured so that if a PC crashes, then only it must be rebooted (or, if it fails, only it has to be replaced without touching or affecting the others). So, in this case, we would lose an antenna or 4 Baseline Boards until the affected PC is replaced. Because of FPDP cable length limitations, this would probably have to be a rack-mount PC (although there are cost and ease-of-replacement advantages to a desk-top if a suitable installation configuration could be devised).

a).1. In each correlator rack exists a CompactPCI rack (or two). In each one, a minimum of 1 CPU and maybe more depending on performance requirements. This CompactPCI rack contains multiple FPDP interface

boards (presumably one for each correlator board). A FiberChannel network card in the CompactPCI rack is the gateway for data to the outside world. I'm looking at Motorola CompactPCI cards and I'm going to price out this option. The system performance can be upgraded by replacing CPU cards and/or adding more CPU cards.

b) All of the real/quasi-real time processing is done by the PPMC cards. They write data to big network drives via network switches. Image processing/archiving machines access these drives via the network switches as well. To replace the PPMC cards (i.e. for upgrade) requires pulling each correlator board out, popping out the old, and installing the new. It may be difficult to get an affordable PPMC card with FibreChannel. The latest soon-to-be-released Motorola PPMC card (which has tons of performance, memory etc -- the PrPMC800) has a 100 Mbit Ethernet interface and sells for ~\$1700 (x 400 boards = \$680k). Still, 100 Mbit Ethernet is 2.4 Gbytes/sec system wide. FibreChannel would provide 24 Gbytes/sec system wide.

> The ALMA correlator has some sort of steering that may
> address this issue. The data converges onto a total of 16
> backend interface ports. I will try to get additional
> information.

Ok, but I'm still going on the premise that we want the correlator to have a wide data pipeline going out and only the backend CPU H.P. limits its performance. In that case, option a) is probably the best. Also, I don't think there data handling requirements (after the LTA) are as severe because of dump times and fewer spectral channels.

> Is Thanksgiving a Canadian holiday ? If yes, have a happy
> one. (I will be out of the office Thur & Fri, 11/23-24.)

We have thanksgiving in October...by the end of November, its pretty much winter conditions in Canada (except where we are...the snow usually doesn't show up until Xmax...if at all).

I'll see you in December and we can talk about the back-end some more.

Have a good one!

Brent.