From Brent.Carlson@hia.nrc.ca Wed Nov 1 17:57 MST 2000
Date: Wed, 01 Nov 2000 16:54:56 -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>
CC: ghunt@zia.aoc.NRAO.EDU, pnapier@zia.aoc.NRAO.EDU, dsramek@zia.aoc.NRAO.EDU,
        gvanmoor@zia.aoc.NRAO.EDU, rperley@zia.aoc.NRAO.EDU,
        bclark@zia.aoc.NRAO.EDU, jbenson@zia.aoc.NRAO.EDU,
        ksowinski@zia.aoc.NRAO.EDU, bwaters@zia.aoc.NRAO.EDU,
        Peter Dewdney <Peter.Dewdney@drao.nrc.ca>
Subject: Re: FibreChannel
References: <200010311912.MAA04448@mnemosyne.aoc.nrao.edu>
Content-Type: multipart/mixed; boundary="------------7877B8246CFCE13C634F3132"
X-Lines: 762
Status: RO
Content-Length: 52498

Hi Bill:

Since your email, I've done some more thinking about the correlator back-end and some snooping around the web. I've included a couple of attachments that I will first comment on and then I will answer some of your specific questions and add some comments.

The first attachment is called "PC_Backend_Concept.pdf" and is how I more or less originally envisioned the back-end. The original SCSI concept I had is gone in favour of a CMC (Common Mezzanine Card which is the basis of PMC only without the specific PCI interface) interface such as Systran's FPDP extender (based on some of the info you provided me). Data from multiple Baseline Boards is shipped, via the CMC and FibreChannel/Gbit Ether/FPDP interface, to back-end PCs. Disk and/or tape drives hanging off these PCs store the data and allow (further) back-end image processing computers to access the data. I think if this model is used, we can probably only afford one CMC card per motherboard (although maybe a separate cheap and dirty USB could be used for M&C) and so I think that monitor and control (M&C) functions will have to be included as well. Presumably, the PC will see some sort of image of the LTA and M&C RAM on the board and it will read and write to the RAM. Although exactly how this happens over (for example) FibreChannel is the fuzziest part of this architecture (at least for me anyway).

The second attachment is called "PPMC_Backend_Concept.pdf" and is another possible way of doing it. This embeds a Processor PMC card (PPMC) on the Baseline Board. The PPMC card accesses the LTA memory via the PCI bus, and the M&C is done via a separate I/O bus (part of the PMC standard as I understand it). DMA on the PPMC will allow the processor to be bypassed for high dump rate configurations. I have found several such PPMC boards by different manufacturers just with a quick survey on the web although I never found anything with a FibreChannel or Gbit Ethernet interface (but it would make sense for such a thing to exist I think). I found a VITA draft standard for PPMC (where the CPU on the PPMC card is capable of bus mastering, or slaving) on the web so this appears to be a reasonably common application for PMC (although probably not as common as PMC slave cards). So, the idea here is that the PPMC is the real-time embedded processor providing M&C, data readout, and perhaps some data processing (FFTs etc.). The PPMC card now writes its data via the

(FibreChannel) interface through a switch to waiting disk arrays and/or tape drives. I understand that disk manufacturers are starting to provide FibreChannel interfaces to their devices so this seems like a very real possibility. Image Processing computers can access the data in the disk drives via the network, or off tape archive. For this configuration to be affordable, I think the PPMC cards would have to cost around $2k or $3k each (if they are $3k each it is a $1.2 million expenditure...but no back-end PCs are required) and we could probably only afford to have one per motherboard. I have not checked out manufacturers prices.

The advantage of the second option is that (if its affordable), it potentially provides for much higher performance in the delivered system and is just as easily upgraded if (when) a faster PPMC interface and/or PPMC CPU combination shows up. If COTS PPMC cards are not affordable, then it may be cost-effective to engineer it ourselves--although the whole sticky issue of an O.S. shows up then.

Bill Sahr wrote:

> Hello Brent,
>
> I plan to break this response to your email into two replies. In the
> first, this email, I hope to begin setting up a framework and schedule for
> arriving at a decision. In the second, probably with a somewhat different
> distribution, I want to go into technical issues.
>
> Since the issue you raise could involve dollar amounts in the $1-$2 Million
> range, I want to expand the distribution list for this email. So, to
> recapitulate a bit for benefit of those who have not seen our earlier
> emails:
>
> The issue is the choice of interface to be used for 1) getting the data out
> of the backend of the correlator (the LTAs) and into computers for further
> processing, and 2) the monitor and control data from/to the correlator. This
> matter has risen to the surface at this time becuase you must soon write
> "the PLAN" for the correlator, i.e., "a detailed architecture, schedule and
> cost document" which must "define a reasonable plan for this interface". As
> we have discussed, the computers which will receive this data are likely to
> use Wintel hardware for reasons of cost, and may run Linux as the OS.
>
> So, the 1st question is "when"? People here are likely to want
> several weeks (or months, if they can be had) to consider this matter.
> By when must this decision be made in a form that is specific enough for
> inclusion in the PLAN? Would it be acceptable for your document to present
> 1 or 2 alternatives rather than a single approach ?

A hard decision is not required for the document and 1 or 2 alternatives could certainly be included. I want to have a definition of some interface(s) that looks like it will meet system requirements and a reasonably detailed cost estimate for it as well. This, I think, is much better than drawing a blank box and labeling it "interface" with some cost pulled out of the air.

Originally I wanted to have this "plan document" ready for the mid-December NSF visit but it doesn't look like it will be possible to do it. Although, depending on what is required from our end, we would at least like to have the meat of the document ready (refined architecture, cost, schedule). The "plan

document" will have to be ready for NRAO and NRC to sign off on before concerted design and construction can begin.

>
> Next is the issue of candidate interfaces, and sufficient technical detail
> to allow investigation. I see 4 candidate interfaces that may be worth
> some research - 1) Fibre Channel 2) gigabit ethernet, 3) USB, and 4) Firewire.
> As to the technical detail - what form factors/bus specifications will
> be needed. I'm on slightly shaky ground here as my knowledge of current
> hardware standards is a bit out of date. However, I would assume that
> we would want a PCI card for the computer side, but if it should turn out
> that Wintel hardware is not used, it would be nice to have VME (6U) &
> CPCI (PMC) cards also available, just to keep our options open.

I guess when it comes to CPCI and VME I am concerned about cost. Maybe PPMC (option 2 above) or CMC with PC backend (option 1 above) will be ok for cost. Although, I guess the choice depends on who pays for it and/or what performance is required.

> What do you
> need for the correlator side - you've mentioned CMC, but I am not familiar
> with that spec. Fibre Channel or gigabit ethernet may turn out to be best
> for the transfer of data in the LTAs. Both are network approaches which
> fit the approach of using switches to converge the data on some number of
> computers with the option of removing the switch at a later date to get
> the full interface bandwidth for each baseline board. (More about this
> approach in the email on technical considerations.)
>
> Costs. We should probably attempt to obtain ballpark figures from at
> least two vendors, preferably three. We may find that a company like
> VMIC, for example, offers products of use to us and is significantly
> less expensive than Systran.

Yes. I've got one more memo to write in the next few days describing the simulation tests I ran on the phasing subsystem signal processing and then I want to work on the refined architecture and refined costs. Then I will start looking for costs on PPMC cards (for option 2.) and for CMC cards (for option 1.).
>
>
> Who pays ? Does your group pay for the correlator side, and we pay for
> the computer side ? That seems the most likely arrangement.

We committed to pay for the PC back-end and this would naturally include the interface hardware. However, if the back-end architecture changes, who pays for what will have to be discussed.

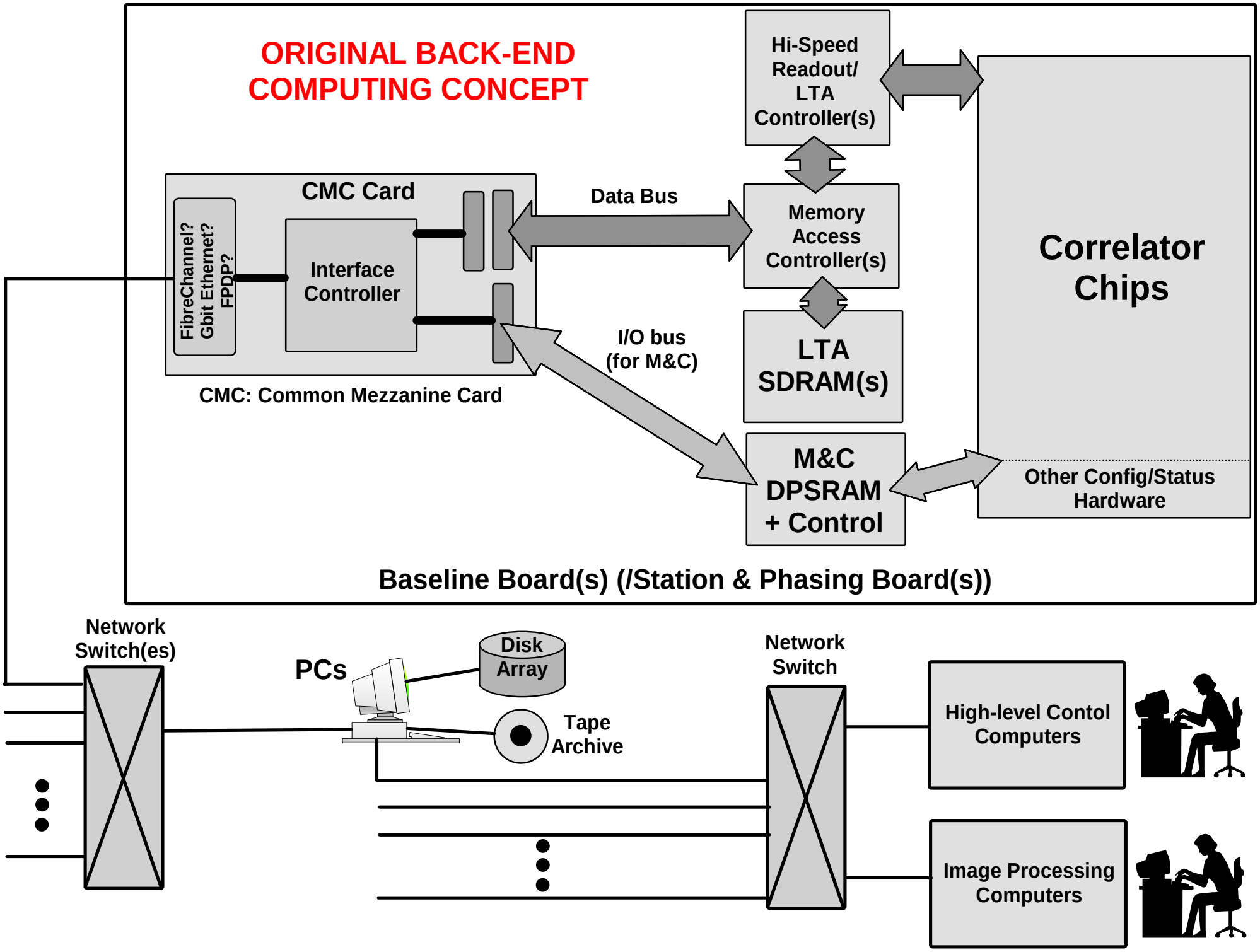
>
> Finally, there is the issue of availability over time. Will Fibre Channel
> or the other interfaces still be around when it's time to build the correlator
?
> Will it still be the interface of choice ? Can the correlator be built and
> the computers be chosen in such a way that we can substitute a different
> interface, at a later time, with a minimum of cost & work?

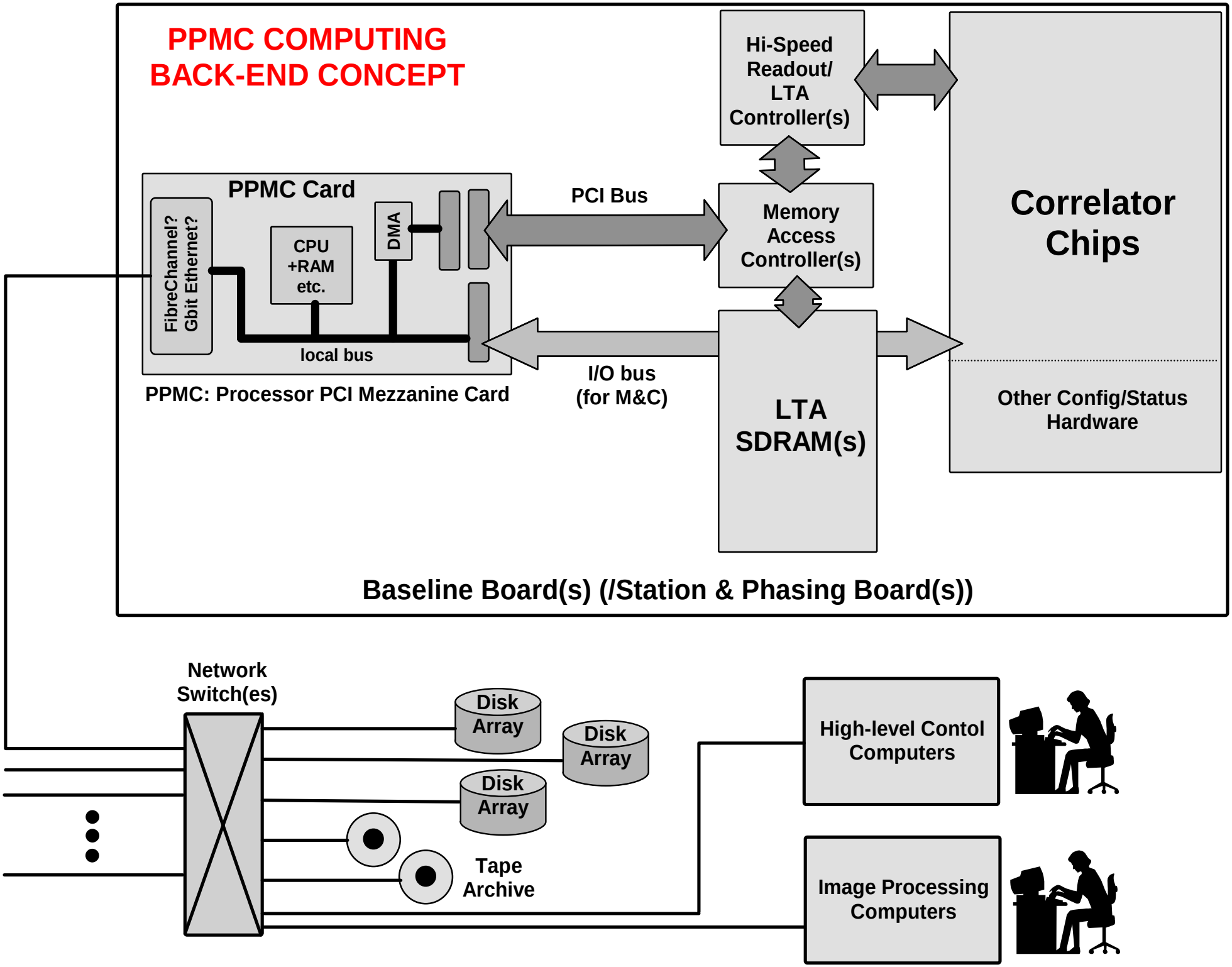
That's why I like the CMC and/or PPMC approach...the system can be upgraded with a new interface, and fancier processors without requiring the motherboards to be redesigned (provided that PMC remains a standard!).

> So, issues have been raised. I'll try to get an email containing much
> more technical detail put together over the next week or two (or three,
> if your timelines permit it).

Ok, I look forward to seeing what you put together.

Cheers,
Brent.


ORIGINAL BACK-END
COMPUTING CONCEPT
Hi-Speed
Readout/
LTA
Controller(s)
CMC Card
FibreChannel?
Gbit Ethernet?
FPDP?
Interface
Controller
Data Bus
Memory
Access
Controller(s)
Correlator
Chips
I/O bus
(for M&C)
LTA
SDRAM(s)
CMC: Common Mezzanine Card
M&C
DPSRAM
+ Control
Other Config/Status
Hardware
Baseline Board(s) (/Station & Phasing Board(s))
Network
Switch(es)
PCs
Disk
Array
Tape
Archive
Network
Switch
High-level Contol
Computers
Image Processing
Computers


PPMC COMPUTING
BACK-END CONCEPT
Hi-Speed
Readout/
LTA
Controller(s)
Correlator
Chips
PPMC Card
FibreChannel?
Gbit Ethernet?
CPU
+RAM
etc.
DMA
local bus
PCI Bus
Memory
Access
Controller(s)
PPMC: Processor PCI Mezzanine Card
I/O bus
(for M&C)
LTA
SDRAM(s)
Other Config/Status
Hardware
Baseline Board(s) (/Station & Phasing Board(s))
Network
Switch(es)
Disk
Array
Disk
Array
Disk
Array
Tape
Archive
High-level Contol
Computers
Image Processing
Computers