Design aspects of the reconstructor for the Gemini Adaptive Optics System (Altair)

Leslie K. Saddlemyer, Glen Herriot, Jean-Pierre Véran, Murray Fletcher

National Research Council of Canada, Herzberg Institute of Astrophysics
5071 W. Saanich Rd., Victoria BC, Canada V8X 4M6

ABSTRACT

The Gemini Telescopes Project is a collaboration to develop two leading edge 8 metre telescopes; Gemini North on Mauna Kea, Hawaii, and Gemini South at Cerro Pachon, Chile. These telescopes will exploit the excellent natural seeing conditions of the sites for observations in the visible and near infrared. To capitalise on the excellent natural image quality that these telescopes will provide, a natural guide star (NGS) adaptive optics (AO) system is being developed for Gemini North.

The Gemini Adaptive Optics System (GAOS), Altair, will be a medium order, Shack-Hartmann (SH) wave-front sensor (WFS) based system running at a control loop speed of at least one kHz. Based on studies of the turbulent layers at Mauna Kea[1] the deformable mirror (DM) has been set conjugate to a fixed altitude of 6.5 kilometres to increase the isoplanatic patch diameter.

The system design choices of a high speed, medium order correction, the instrument location and altitude conjugation have placed challenging demands on the reconstructor for this instrument. We present a here a few of the constraints, demands and design choices for the reconstructor. The paper will conclude with a brief summary of the project status.

Keywords: Adaptive optics, reconstructor, conjugation to altitude, Gemini

1. INTRODUCTION

The Gemini North NGS AO system, Altair, is designed to capitalise on the excellent image quality that the Gemini Telescopes will deliver. This facility instrument must be robust in operation and not require expert AO intervention during normal science observations.

Altair will be a 'medium order' AO system[4]. It will incorporate a 12 x 12 SH WFS driving a 177 element DM. The reconstruction servo-loop rate requirement is one kHz, with a design goal of 1.5 kHz. The WFS incorporates an EEV-39 frame transfer CCD controlled by a San Diego State University generation II detector controller (SDSU II). This is the same hardware configuration as used by the Gemini facility guide CCD's and the Gemini multi-object spectrographs (GMOS) on-instrument wave-front sensors.

A parallel process will monitor the reconstructor control loop statistics and produce optimal reconstruction matrices that will track changes in atmospheric conditions[4]. The reconstruction matrix update rate is expected to be about 0.1 Hz. The entire instrument will reside on the Cassegrain rotator of the Gemini North telescope. This will subject the instrument to a constantly changing gravitational vector and places strict space and mass restrictions on the instrument.

A major feature of Atair is the conjugation of the DM to a fixed altitude of 6.5 kilometres. This will significantly increase the isoplanatic patch diameter and result in a much larger sky coverage than conjugating to ground. This effect is enhanced by minimal ground level turbulence, a condition the Gemini project has strived to attain. However, these design choices for Altair presents challenges for the reconstructor beyond those that already exist for a system conjugated to the telescope pupil.

2. WFS AND DM ILLUMINATION ISSUES

The WFS and DM are optically registered to each other in a classic Fried geometry. The WFS sub-apertures and DM actuators are both arranged in a square pattern. The spots produced by the lenslet array will nominally fall on the intersection of four adjacent pixels, forming an array of sub-aperture spots. These pixel quad-cells are precisely optically aligned to register with the mid-point between each set of four DM actuators.

2.1. WFS CCD geometry

The optical geometry of the WFS is designed to place the spots produced by the lenslet array on the intersections of four CCD pixels. Spot positions will be measured using a standard quad-cell centroiding algorithm after pixel intensities are offset and gain corrected.

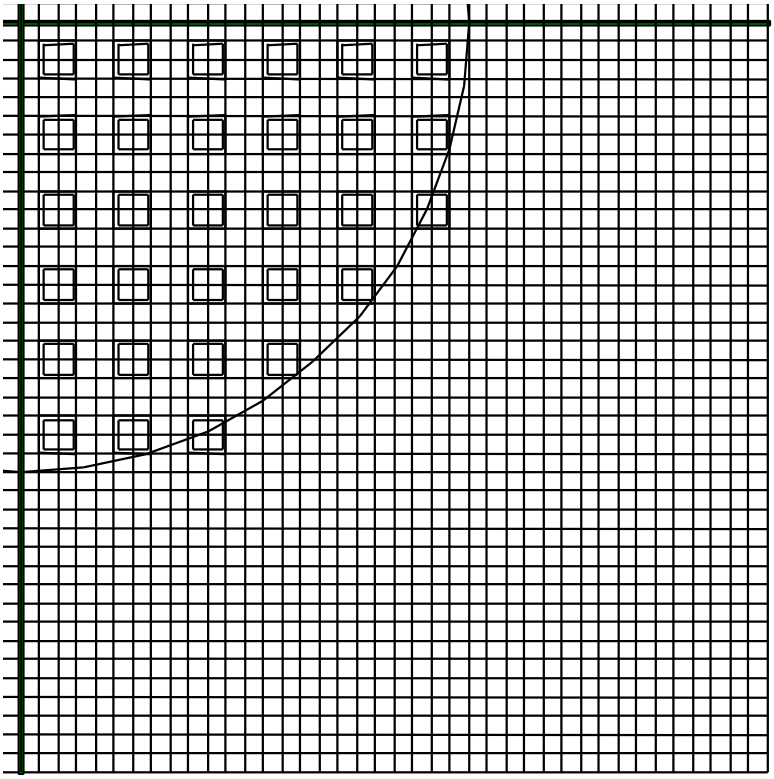

Figure 1 One quadrant of the WFS CCD showing the quad-cells and the illumination limits of an on-axis object

The pixels on the WFS have an image scale of 0.4 arcsec per pixel. To reduce the effects of cross talk, quad-cells are placed on a spacing of every four pixels, producing a pair of guard rows and columns between each quad-cell. To reduce the readout time guard columns and rows are not digitised. The fastest possible readout speed with this configuration is about 600 μsec, with a relatively high read noise. The currently envisioned readout speed of about 750 μsec will produce a read noise of 4 or 5 electrons.

The WFS CCD is read out using all four readout registers, with identical clocking to each quadrant of the CCD. This will always produce a four-way symmetrical image. Figure 1 illustrates one quadrant of the CCD, showing the individual pixels, the digitised quad-cell pixels and the illumination limits of an on-axis point source.

2.2. DM and WFS illumination

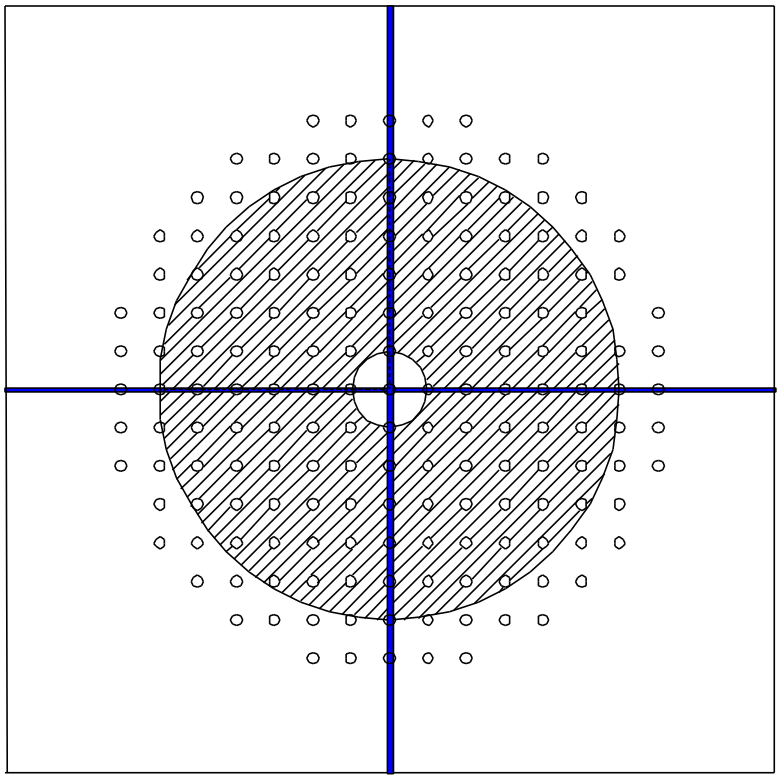

Figure 2 Illumination patch and actuator pattern, on-axis guide and science objects

Figure 2 illustrates the illumination patch and DM actuators for on-axis guide and science objects. With a beam splitter reflecting 400 – 835 nm light to the WFS and transmitting longer wavelengths to the science instruments it is possible to guide on bright science targets. This situation produces a pattern where there are about 120 illuminated SH cells that will be used to close the loop on the illuminated DM cells. The ring of guard actuators will be driven by extrapolating the positions of the illuminated “closed-loop actuators”. This will be done to minimise edge effects. This closely resembles an AO instrument with a DM conjugated to the entrance pupil of the telescope.

Figure 3 illustrates a configuration where the guide and science objects are not coincident. The illumination patch is a function of the field position of the objects. Here the guide object is at the maximum off-axis distance of one arc-minute. Where the illumination patch of the guide object actually falls outside the array of actuators no correction is possible.

The actuators that fall within the guide object illumination patch will be corrected in closed-loop fashion, including some actuators that were positioned by extrapolation in the on-axis case. However, there is now a large crescent shaped pattern of actuators whose positions must be an extrapolation of the positions of the closed-loop actuators. There are more actuators here that require an extrapolation control than in the strictly on-axis case. Some of these actuators now affect corrections to the science path.

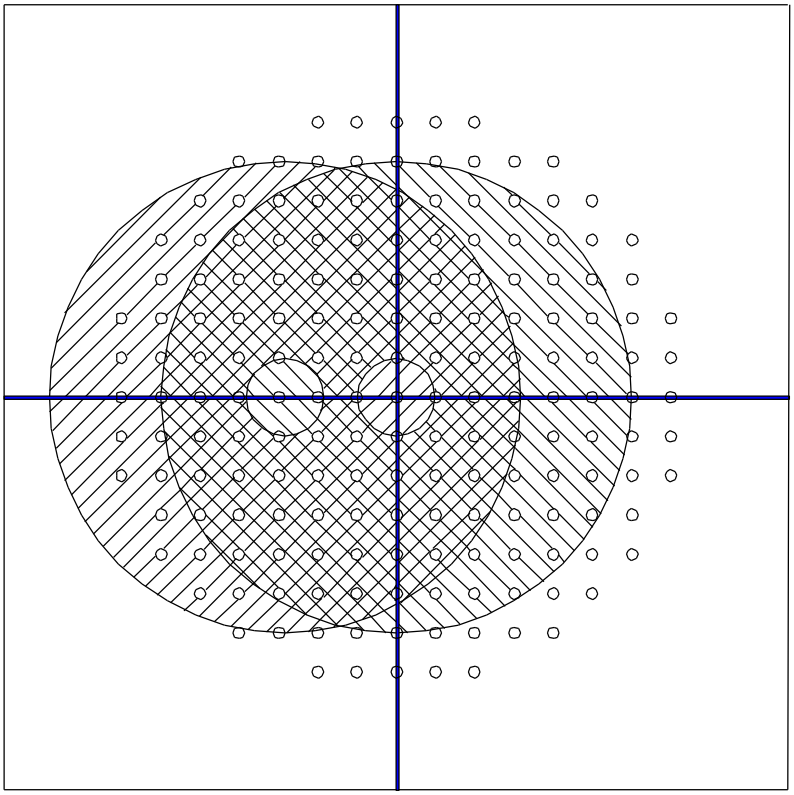

Figure 3 Illumination patch and actuator pattern, on-axis science object, 1 arc-minute off-axis guide object

The crescent shaped portion of the science beam that is beyond the guide object patch will only be corrected in an open-loop extrapolated fashion, and hence sub-optimally. This increases the value of implementing an extrapolation algorithm that not only doesn't introduce edge effects but can actually attempt to extrapolate low order corrections to this portion of the science object beam.

Another side effect of the variable illumination patch position is the increased readout area, and hence readout-time for the WFS with an off-axis guide object. As each quadrant is identically clocked, the over-all read time is increased in order to capture the larger zone that is illuminated. In Figure 3 this is driven by the increase illumination patch in the two left hand quadrants.

2.3. Control issues with non-constant illumination patch position

The variable illumination patch position of the guide object adds an additional set of requirements to the reconstructor and optimising processes over those of a fixed illumination patch position. Between these two processes these include:

1. Selecting from the delivered pixels the set that represents the quad-cell apertures that are illuminated by the guide object.
2. Producing a control matrix that matches these sub-apertures.
3. Mapping the control signals that result from the reconstruction to the appropriate actuators. This will involve not only different actuators but also a different number of actuators as a function of the position of the guide object.
4. Applying an extrapolation algorithm to a variable number of the 'extrapolated' actuators. This may also require a different extrapolation algorithm for the "science extrapolated" and "guard extrapolated" actuators.

3. DATA FLOW

There are three main data flows that are of concern to the core reconstruction processes. These are the highest speed control of the DM and T/T mirrors, the slower off-loading of corrections to the secondary mirror and the slowest task of passing off control loop statistics to the optimising process to produce optimal reconstruction matrices. The interactions of these data flows are illustrated in Figure 4.

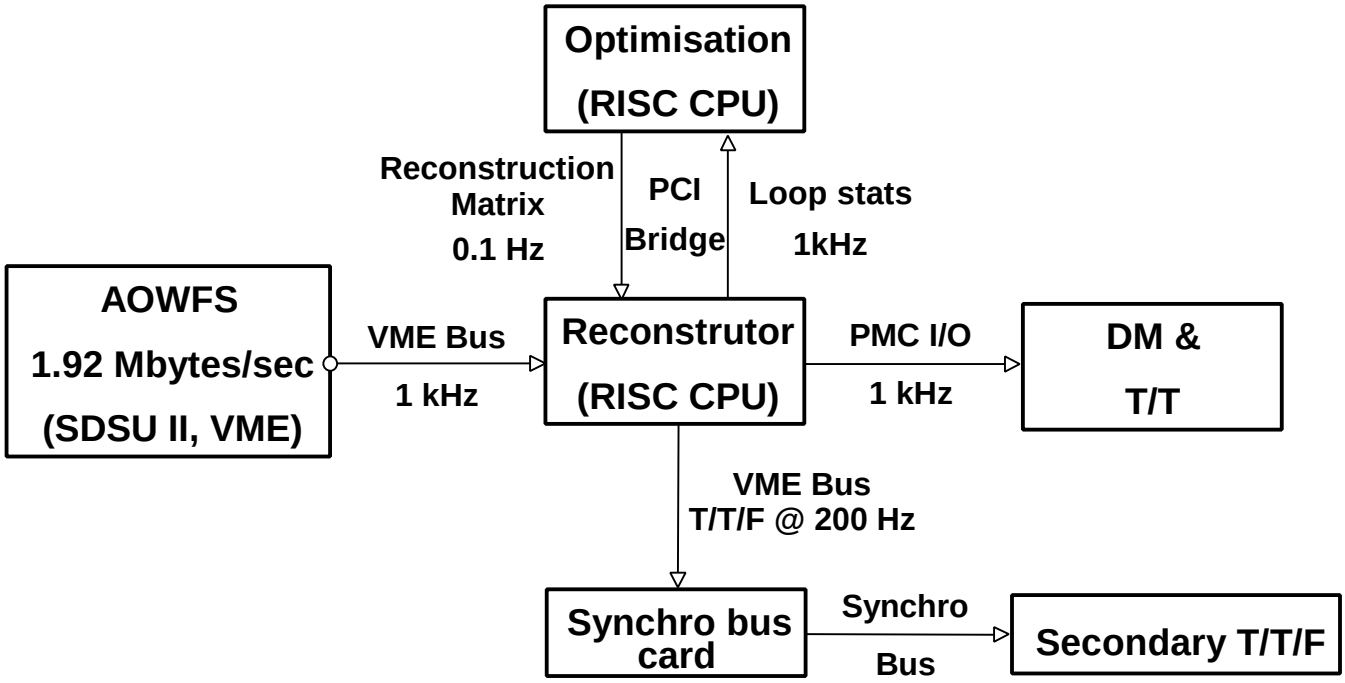

Figure 4 Reconstruction data flows

3.1. DM and T/T data flow

The most critical, highest speed data flow is the flow of WFS CCD data through to the delivery of the DM and T/T mirror drive signals. This must occur at a speed up to 1 kHz (with a goal of 1.5 kHz). A high powered RISC based CPU board will be dedicated to this task. This can be viewed as consisting of these seven sequential tasks:

1. The CCD readout electronics deliver pixel values directly to shared memory on the reconstructor CPU hardware over the VME bus.
2. The pixel values are offset and gain compensated and X and Y slope values computed for each quad-cell to produce a vector of slope errors.
3. The slope error vector is multiplied by the reconstruction matrix to produce an actuator error vector. A calibration vector is then added in to compensate for systematic errors.
4. The actuator error vector is fed into a servo controller to produce a vector of servoed drive signals.
5. The vector of servoed drive signals is extrapolated to produce the extrapolated drive signal vector.
6. The pair of actuator signal vectors is mapped to physical actuators and output to the DM and T/T electronics over the private PCI mezzanine card (PMC) I/O bus.
7. The necessary statistics are delivered to the optimisation process. This involves writing signal values into a circular buffer in the optimisation processor shared memory.

3.2. Optimisation data flow

The process that produces optimised reconstruction matrices requires statistics from the reconstruction control loop process. This is expected to include the vector of slope errors, the vector of actuator errors and the vector of servoed drive signals. This process will reside on a separate processor from the reconstruction process, so these values will be delivered over a private PCI bus between the two processor boards.

The optimisation process will deliver a new reconstruction matrix to the reconstruction processor memory over the private PCI at a rate of about 0.1 Hz. A notification will accompany each new matrix delivery to signal the reconstruction process to hot swap in the new matrix.

3.3. Secondary mirror T/T/F data flow

Any persistent T/T/F terms that are present on the DM and T/T mirrors will be filtered and off-loaded to the telescope secondary mirror. Any system that requires communication with the secondary control system (SCS) is equipped with a connection to the SCS dedicated low latency bus through a VME interface card. Altair will be capable of sending update signals to the SCS at a rate of 200 Hz.

4. TOP LEVEL DESIGN

The Gemini telescopes control system is based on the infrastructure provided by the Experimental Physics and Industrial Control System[2] (EPICS). This requires within Altair a VME based computer that is responsible for communication with the other telescope systems and the motion control requirements within Altair. Altair will be mounted to the cassegrain rotator on the Gemini telescope. This places strict mass, volume and cabling restrictions on the system. This led to the desire to minimise the components comprising the reconstructor. Fitting as many of the components of the reconstructor within this existing VME chassis as possible was desired to alleviate these restrictions. This and the previously discussed issues led to the physical design illustrated in Figure 5. The AO wave-front processing system (AOWPS) performs the atmospheric corrections.

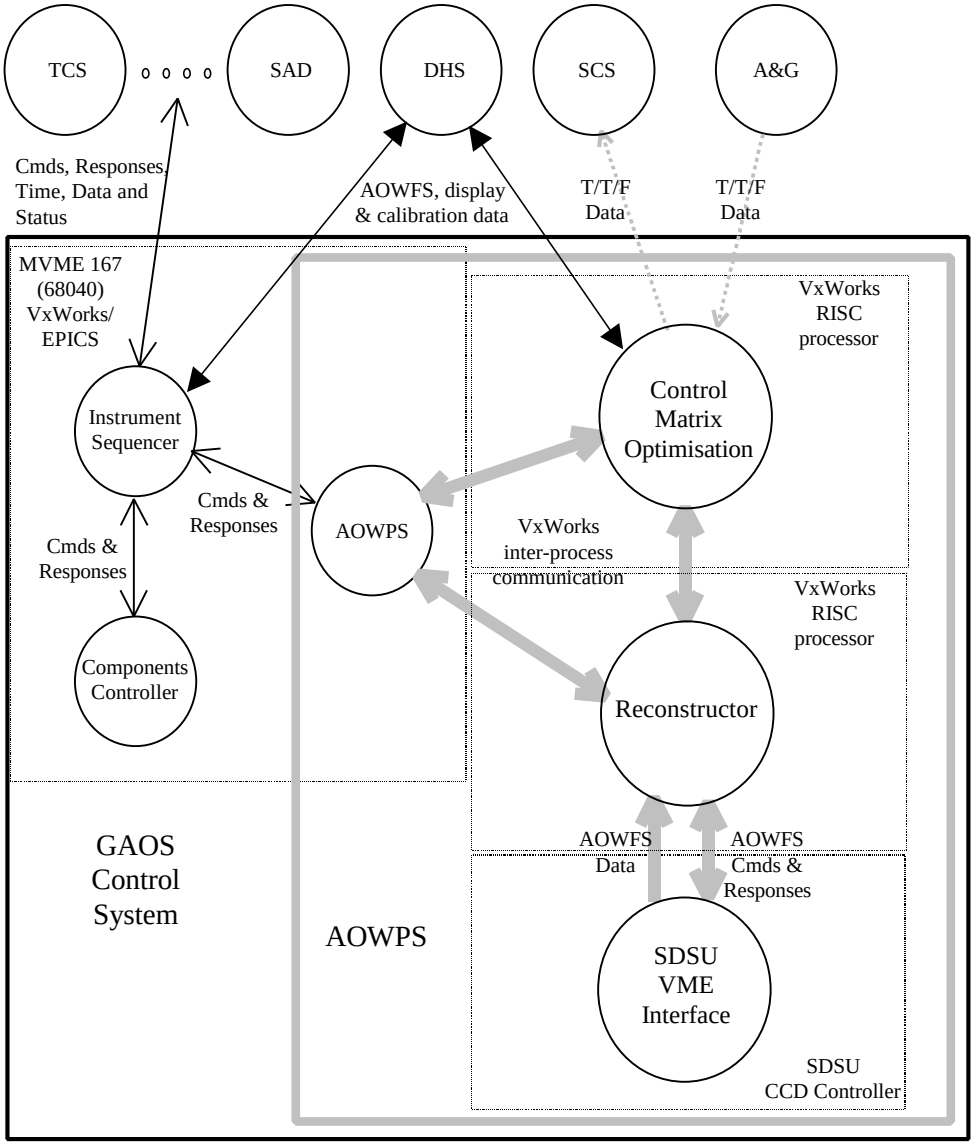

Figure 5 Altair system physical design

There are four main boards that comprise the Altair control system, including the reconstructor. These are:

1. The MVME 167. This board runs the EPICS control system and manages all communications with other telescope systems such as the telescope control system, the status and alarms database (SAD) and the data handling system (DHS). It is responsible for the motion control of devices within Altair other than the DM and T/T mirrors.
2. The control matrix optimisation processor. This is a high performance RISC based processor. It is responsible for processing the reconstruction loop statistics and delivering updated reconstruction matrices. It will also deliver to the DHS any data for real-time display or for storage and off-line analysis.
3. The reconstructor processor. This will be a board identical to the optimisation processor. It will be responsible for the reconstruction and driving of the DM and T/T mirrors. It will also accept T/T/F information from the acquisition and guidance (A&G) system from other WFSs and send T/T/F information to the SCS.
4. The SDSU II controller for the Altair WFS. This will present the WFS CCD values to the reconstructor processor.

This design raises some performance questions. These include VME bandwidth and latencies for the delivery of WFS CCD data and the use of a single high-powered CPU to perform the time critical reconstruction. We consider these next.

4.1. VME bandwidth

With the worst-case scenario of a one arc-minute off-axis guide object the SDSU CCD controller will be delivering about 150 quad-cells worth of data to the reconstructor CPU for each control loop; With 4 pixels/quad-cell and 2 bytes/pixel, this is:

$$150 \times 4 \times 2 = 1200 \text{ bytes/WFS readout}$$

At the maximum readout speed of about 600 μsec, this leads to a maximum bandwidth requirement of:

$$\frac{1200 \text{ bytes/readout}}{600\ m\text{sec}} \cong 2 \text{ Mbytes/sec}$$

This is about 5% of the theoretical bandwidth of the VME bus[5] for D32 transfers and still only about 20% of a more cautious 10 Mbyte/sec value. As well, this occurs only during 600 out of each 1000 μsec, leaving about 400 μsec unutilised by the reconstruction processes.

4.2. VME latencies

The VME bus will be utilised by other, unsynchronised, processes such as the EPICS CPU communicating with motor controllers, time and digital and analogue I/O boards in the VME chassis. The concern is that the SDSU CCD controller will randomly be denied access to the VME bus for a long period of time introducing variable and possibly unacceptable latencies.

The VME bus is a priority based bus. It will be necessary to ensure that the SDSU CCD controller and reconstruction processors have the highest priority in the system. However, once another device has the bus, all other requests for the bus must wait. A block transfer on the bus by another board can occupy a considerable period of time. It is therefore vital to know the maximum time that the CCD controller will have to wait for the bus. Based on a pessimistic data transfer rate of 5 Mbytes/sec for a D8 transfer (as opposed to the VME theoretical rate of 10 Mbytes/sec) and assuming a 16 byte transfer without releasing the bus, we can compute a theoretical delay.

$$\frac{1}{5 \text{ Mbytes/sec}} \times 16 \text{ bytes} = 3.2\ m\text{sec/transfer}$$

With about 100 μsec allocated to reconstruction latency, 3 μsec is an insignificant impact. To test the realism of this value two investigations were made. The first was to inspect the EPICS drivers for the set of boards that would be used for the motion control, and produce the only other traffic on the VME bus. The second was to monitor the VME bus on an operating EPICS system to monitor the behaviour of a functioning system.

The code inspection revealed that all communication with auxiliary boards was one access at a time. The transfer of an array of data, such as a command string to a motor driver board, occured one character at a time. Therefore one would expect that maximum bus latencies would be significantly lower than calculated above.

The EPICS system that was assembled for GMOS was borrowed to perform this inspection. The system was populated with the following boards:

- the EPICS processor
- a bc635VME Bancomm time bus interface card
- a VME44 Oregon Microsystems stepper motor driver card
- a VME8-8 Oregon Microsystems stepper motor driver card
- a XVME-240 XyCom TTL I/O card and
- a XVME-566 XyCom analogue input card

The BBSY* back-plane signal was monitored. This signal is asserted any time that the VME bus is in use[5]. This provides an indication of the time the bus is not immediately available. Two motors were set moving and this line was monitored for three minutes. The access times were always 200 nanoseconds. This is consistent with a single transfer operating at a bandwidth of 5 Mbytes/sec.Hence it is not expected that other traffic on the VME bus will significantly affect the transfer of CCD data across the bus.

4.3. Single CPU reconstructor

An array of DSP's is the most common choice of reconstructor processing for a medium order AO system. However, the performance of RISC CPU's has increased dramatically in the past few years. In the interests of accommodating the mass and volume constraints and to simplify the software task, especially considering the issues listed in Section 2.3, it was decided to investigate the possibility of using a single high-speed CPU processor.

To investigate the capabilities of these processors, a benchmark program was written. The algorithm for this benchmark is representative of a real reconstruction except that instead of an extrapolation algorithm to derive positions for the extrapolated actuators, all actuators signals were determined as though closed loop controlled. This was considered a worse case scenario for these actuators. In the benchmark each actuator signal was determined by forming the dot product of the 240 element slope vector with one row of a reconstruction matrix, and then performing a simple integral servo controller function. An extrapolation will more likely be based on some function of the nearest few actuator positions that are under closed-loop control and will involve fewer instructions.

.

Architecture	Processor speed [MHz]	Operating System	Compiler Directives	Execution Time [msec]
Sparc (Sun Workstation) (sun4m)	50	SunOS 5.5.1	-Xc -fast -native -dalign	3.96
UltraSparc clone (sun4u)	200	SunOS 5.5.1	-Xc -fast -native -dalign	0.65
AXPVME (21066)	160	Digital Unix 4.0	Unknown compiler options	0.533
UltraSparc clone (sun4u)	300	SunOS 5.5.1	-Xc -fast -native -dalign	0.55
Motorola MVME167	33	VxWorks	-O4	31

Table 1 Table of benchmark results for various processor/operating systems

It can be seen in Table 1 that the highest performing processors are comfortably able to complete the computations necessary for a reconstruction. However, the benchmark is not able to simulate the other latencies due to things such as data transfers across the bus and I/O to the DM and T/T mirror electronics. This is only possible by testing real hardware in a realistically operating situation.

4.4. Reconstruction processing algorithms

The relatively large period of time when the SDSU controller is digitising and transferring data across the VME bus presents an opportunity to perform a significant amount of processing.. Figure 6 represents a simple timing diagram for the reconstructor, illustrating time lines for two different processing approaches. Approach two will be implemented within the Altair reconstructor. The two variants each have their benefits.

1. Processing option one is the more classic approach. The entire vector of slopes is assembled (pixel pre-processing and centroiding occur as the pixel values arrive in shared memory) and then multiplied by the reconstruction matrix.. This has the benefit that the dot products occur with maximum efficiency of the CPU. DSP processors are very efficient with this model due to the single cycle multiply-accumulate instructions and automatic array increment registers. However, all of the processing must occur within the relatively small period of about 300 μsec.
2. Processing option two has the CPU perform the matrix multiplication in steps, column wise. As each centroid is determined, this value is multiplied down the appropriate column of the reconstruction matrix, updating the result vector. This adds more total instructions to the process. The entries of the result vector must be loaded and saved many times. There is now, however, about 850 μsec of processing time available. With the multiple instruction units and massive pipelining within modern RISC processors this is an efficient approach.

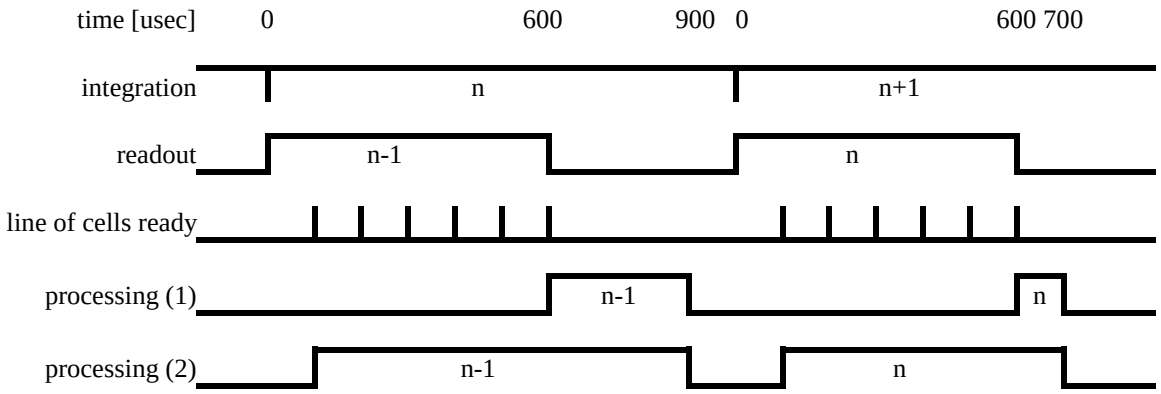

Figure 6 Simplified timing diagram for reconstruction process

4.5. VME slot arrangement

The VME bus is a priority-based bus. By appropriate selection of the bus grant levels, arbiter algorithm and slot positions of the various system components, it is possible to ensure that the SDSU II controller and reconstruction CPU have the highest priority access to the bus[5].

Figure 7 illustrates the VME slot assignments for the most time critical components in Altair. The reconstructor board becomes the system arbiter, as opposed to the MVME167 EPICS board. The SDSU II controller board is in slot 2. This, along with the use of VME bus request BR3* will ensure that these two communicate at the absolute highest bus priority[5]. The other data paths that involve the reconstructor can be seen here to be private channels.

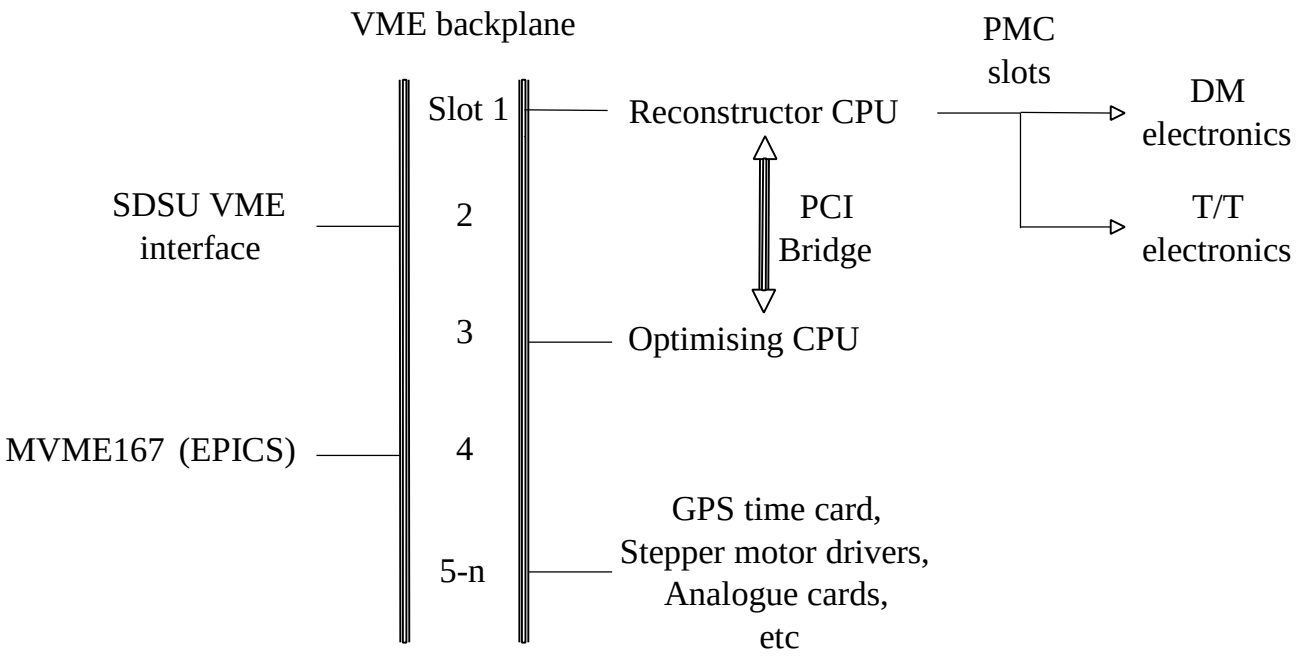

Figure 7 VME slot assignments

5. Summary

Some of the constraints, requirements and issues that have affected the reconstructor for Altair have been presented. Many of these are due to the decision to conjugate the DM to the 6.5 km turbulent layer above Mauna Kea and to incorporate the entire system on the cassegrain rotator. Some of the theoretical and test results have been presented that led to the current control system design, including the reconstructor. However, these must be confirmed with actual throughput tests.

The necessary components to test the time critical main reconstruction process have been ordered. The final components are expected by middle to late April 1998. These will include:

- A high-speed RISC based single board computer running the VxWorks real-time operating system
- An SDSU II CCD controller.
- A PMC digital I/O daughter card.

These will be integrated to run a full-length data flow, from SDSU II controller to outputs on the PMC board. The SDSU II controller card can be programmed to deliver simulated data at the proper rates or real data from a WFS CCD head. No optics will be present to deliver sub-aperture spots as would be present in the final product. The PMC digital I/O card will be used to simulate the data stream output to the DM and T/T mirror electronics. It will also be used to provide timing signals for a logic analyser to monitor the timing delays at various points in the reconstruction.

The CPU will run code representative of the final code. This will include all of the steps listed in section 3.1 except for the transfer of data to the optimising processor. Most of the code is ready for these tests. Results of these tests are expected to be available by late April or early May 1998.

6. References

1. R. Racine, B. Ellerbroek, “Profiles of night-time turbulence above Mauna Kea and isoplanatism extension in adaptive optics"” SPIE Vol. 2534, pp. 248-257. 1995
2. http://www.atdiv.lanl.gov/aot8/epicshm.htm, Los Alamos National Lab site for EPICS information
3. EPICS Input/Output Controller (IOC) Application Developer’s Guide, M. Kraimer, Argonne National Laboratory, December 1994
4. G. Herriot, et. al., *Altair Preliminary Design Review Documents*, National Research Council Canada
5. W. D. Peterson, *The VMEbus Handbook*, VFEA International Trade Association, Scottsdale, AZ, 1993