National Radio Astronomy Observatory
Socorro, New Mexico
Very Large Array Program

VLA Computer Memorandum No. 166

UPGRADE OF THE VLA ON-LINE COMPUTER

Gareth Hunt and Ken Sowinski

February 1983

This memorandum contains a proposal to upgrade the VLA on-line (synchronous) computer system. We present here a summary of the current and future requirements for this computer system; and several alternative paths for an upgrade are outlined. From these latter, we have selected two feasible routes as a basis for a decision.

We begin with a general description of the current system.

1. The current system

The current system comprises an array processor (AP) and seven operational CPUs plus an additional one which is a repository of spare parts and which, under normal conditions, is used for program debugging. The AP is a Floating Point Systems model AP120B which is programmed wholly in assembly language. The CPU models are Modcomp II/45 (4), Modcomp II/25 (2), and Modcomp 7810 (2); they all run the MAX III operating system and are connected together with a home-built network.

The two main computers are BOSS, the policy maker, and MONTY, the monitor and control computer. These two CPUs have 1/4 of their maximum possible memory (128k bytes) configured as shared memory for the global common area, which contains the control information (source position, etc.) and the monitor data (current antenna positions, etc.). The remaining 3/4 memory (96k bytes) in BOSS is used up with the real-time tasks (calculation of delays, etc.) and the (heavily overlayed) new source program. The remaining memory in MONTY is used for the DCS communication and for simple interactive displays used by the operators and technicians.

The second pair of computers is the two correlator computers CORA and CORBIN. At present only one of these is used, and this is one of

the reasons for the limited number of baseline-channels available in spectral line mode. This restriction will be lifted as soon as the "pipeline" becomes available, at which time these two computers will be used as the first stage in the sorting process.

The two micro (Modcomp 7810) computers are used only for communications. One (EUNICE) is used as a terminal concentrator, handling all CRT terminals and hard-copy devices for the operators, engineers, and programmers. The second (DEXTER) is used for the real-time communication with the asynchronous system (DEC-10).

The seventh operational computer (SPECTRE) is used to communicate with the correlator and to control the array processor. This computer is in the correlator room, physically separate from the others.

There are three main identifiable bottlenecks in the main two computers:

a) The shared memory is at present almost full; it would already be so were the database descriptors not kept in the non-shared memory of both MONTY and BOSS. Any one of the additional demands for space in this area will require a major redesigning of the data storage algorithms and a major change in the other memory allocation in BOSS.

b) The memory in BOSS is completely used; every small change in this computer requires very careful memory management.

c) In MONTY there is some flexibility available in memory usage, but fully 60% of MONTY's CPU cycles are used up demultiplexing monitor data; the rest is taken up checking this data, generating the commands to the DCS system, and running the interactive (DMT) monitor data displays. There is not nearly enough CPU capacity to do anything additional, e.g. to provide the number or the diversity of displays required to operate the array.

In addition:

d) The I/O bandwidth in BOSS is adequate, but does not have sufficient capacity to allow the addition of faster devices. In addition, no device requiring a direct memory channel can be added, since all of the possible (DMP) channels are now in use.

There are other disadvantages to the current system as a whole:

e) The operating system used in each of the CPUs (MAX III) is very satisfactory for an on-line computer system, but is a clumsy environment for development and debugging. The efficiency of programmers working with this system is far from optimal.

f) Most of the hardware in the on-line system was purchased in 1974. Although the computers have been reliable and should continue to be so, some peripherals are now obsolete and others are obsolescent. We can expect increased maintenance problems and correspondingly higher down time.

2. Additional requirements

In addition to the small changes to the system currently in progress to include such enhancements as the 16 bit floating point format and the B and D IFs, there are many improvements that are essential; many of these cannot be done conveniently (or even at all) with the existing configuration.

a) Real-time diagnosis. When error conditions are detected it is not always obvious what is causing the problem. With the large amount of monitor data available to the on-line computers it should be possible to produce not simply an error message but a complete or partial diagnosis. This requires many more computer cycles in MONTY where they are short plus additional disc storage.

b) Real-time monitor data logging. The visibility data is written on a fixed-head disc and passed to the DEC-10 where a database is constructed. Within a very short time the data is accessible by all of the data analysis programs. Such a system has also been requested for the monitor data. At present all concerned have to wait 2 to 4 days before the monitor data is written into a database. For some failures and for some astronomical experiments (e.g. solar flares) the history of monitor values before, during, and after critical events can be crucial to a correct interpretation.

c) Improved monitor displays. At present there is no communication between the monitor data checker and the appropriate displays. It would make the recognition and diagnosis of errors simpler if (for example) a voltage out of tolerance as reported by the monitor system could be highlighted either in a different color or at least by a flashing message.

d) Graphical displays. The availability of graphical displays of both monitor and visibility data would make the identification of problems much simpler. A real-time historical plot of quantities is much easier to interpret than any row of numbers.

e) High density tape drives. At present the VLA produces about 1,000 archive tapes per year from the on-line system; at the present price of tapes this amounts to about $10k per year (for the cheapest available tapes) in the VLA operating budget. With the addition of the second pair of IFs in 1983 this will double; with the advent of full spectral line capability this may well double again. To keep our tape production down to a manageable level we need to increase the recording density; to use the high density we need large buffers. At present we have neither the memory nor the I/O capacity for this.

f) Ease of programming. At present the programmers can use the spare computer to write and compile programs. Since this is an on-line system the available debugging time is limited. Added to this is the complication of the limited address space: each program cannot grow beyond a certain size. This is not a very efficient way to develop or indeed maintain a system.

g) Simplification of operation. The on-line computer system is used to control the array under the supervision of the operators. It would be a benefit to the smooth running of the array if the system could be re-designed to provide an improved human interface and such routine functions as the maintenance of the observers' log.

h) Reliability. The down-time of the VLA is counted in antenna hours. Since an outage of a central component, such as the correlator or the on-line computer system, causes 27 antennas to be down, the reliability of these components must be very high. Accounted by number of correlators this would be even more emphatic. In the years 1981-2, 41% of the total down-time of the VLA was attributed to the on-line system, much of this being due to network failures (see also section 1f).

i) 512 spectral line channels. The existing VLA correlator is capable of producing 512 channels. The computer hardware was specified to handle 256 channels and in consequence half of the available data is discarded. With some effort, the existing system can be modified to produce 512 channels provided that the integration time is set to a minimum of 20 seconds. Shorter integration times would possibly require a faster array processor, and would increase the memory needed for buffers and the I/O load on the remainder of the on-line system. Such a requirement would also have severe repercussions on the I/O bandwidth and disc capacity specifications for all other computer systems analyzing VLA spectral line data.

j) Short integration times. We can currently observe with an integration time as short as 3 seconds, but there have been requests for yet shorter integration times to observe transient phenomena such as solar flares, flare stars and pulsars, and to make test observations of satellites. It is possible to modify the existing system to make these observations if the number of baselines is restricted, but the changes would not be simple.

k) Scan averaging. At present there is no capacity in the system to average beyond the basic data collection cycle specified by the observer. At present it causes a problem only at Ku and K bands, where the 10 second S/N ratio is insufficient to form an antenna gain solution on all but the very strongest calibrators. It is sometimes impossible to determine whether an antenna is functioning at these bands without integration. More serious is the inability with the current system to be able to form a predictive calibration, which may well be essential to the success of the "pipeline."

l) Additional functions. Any additional function that requires antenna based parameters to be stored cannot be implemented with the current memory restrictions. Into this category come the various solutions to the antenna pointing problems: a polynomial correction for the azimuth bearing defects, temperature dependent models, and an active correction system. Even extensions such as Doppler tracking will increase the amount of control information and decrease the memory available for monitor data.

The AP will be used at about a third of its capacity for four IF continuum observations, and will be used to at least half its capacity for

full line work. Data memory in the AP can be easily expanded from the present 32k words to only 64k words. At present, only about one half of the available program memory is used. Adding any additional functions here, such as interference detection and suppression, may tax the machine to its limit.

m) Additional hardware. There is enough memory capacity to include the 2cm and 21cm FET amplifiers with their increased number of monitor points. This will not be true when any other hardware is added. In particular the following are planned: 1cm FETs, 6cm FETs, a 327MHz system, a 600MHz system, an X-band system for the Neptune encounter, and a new focus-rotation unit. To include these either the monitoring of some values will have to be dropped or a major effort will have to be made to reorganize the monitor database.

3. Future requirements

In addition to the Very Long Baseline Array (VLBA) proposal, several proposals were made at the Future Instrumentation Workshop at Greenbank in October 1982 to build instruments which, if realized, may have considerable impact on the operation of the VLA.

a) The VLBA. We assume that the operation of this will be independent of the VLA operation except in the cases where all or part of the VLA is linked as an element to the VLBA. This is adequately handled by the existing system. Enhancement to this operation should be routine.

b) The "inrigger/outrigger" antenna. The correlator will have to be augmented to provide the larger delays necessary to correlator the outrigger antenna with the existing VLA antennas. Apart from a non-trivial software effort to handle these delays, the special equipment (e.g microwave links) in such an antenna compounds the problems raised in section 2m.

c) The Neptune encounter. Discussions are still in progress to determine the form of necessary modifications to the VLA to allow its use as a deep space receiving station. At present it seems that the impact on the on-line system will be small, except for the additional monitor points required for the new receivers (see section 2m). A more modern, adaptable system would make such experiments easier to implement.

d) The millimeter array (VHFA). This independent instrument has been discussed by the NSF committee on the future of millimeter astronomy (Millimeter Array memo no. 2). The operation of this will be independent of the VLA, but the current proposal is that it will be operated by the same person who operates the VLA. If this is pursued, then the on-line systems of the VHFA and the VLA may need to communicate to provide a simplified human interface to the operator extending the requirements stated in section 2g.

e) The very low frequency (75 MHz) array (VLFA). This addition to the VLA has been discussed by Rick Perley and Bill Erickson in VLA Scientific Memorandum No. 146. As proposed, it will "piggy back" on the current VLA system. The proposal is to use the existing waveguide to transfer the command data, the monitor data, and the received signals in much the same way as is used by the VLA itself. It will, therefore, certainly use the existing DCS system and on-line system for monitor and command communication. The actual control of the VLFA may be done in an entirely separate computer, and the communication with this computer should be relatively easy, needing few resources. However, the processing of the monitor data and the presentation of it in a unified way to the operator and the technicians will certainly require more resources and a simpler human interface (sections 2g and 2m).

4. Modifications to the existing system

The system as it exists now has reached its limits of growth; something has to be done. The most obvious first attempt is to augment the existing system. If we constrain ourselves to keeping the current system, several patches can be proposed to help us limp through a few more years.

a) Swap the physical CPUs called SPECTRE and EUNICE. Since SPECTRE is a model 11/45 it can be connected to the remaining port of shared memory and access global common. Barry Clark has discussed this in his memo of 20th December 1982 entitled "Modcomp Enhancement Kludges." The physical arrangement of this solution would make maintenance of the three connected CPUs inconvenient. Additional hardware changes will have to be made to EUNICE for the real-time interrupt and the auto-load at system initialization. Some or all of the copies of the display program (DMT) could be run in SPECTRE to help alleviate the CPU cycle crunch in MONTY; and we propose that the antenna gains for bands not currently in use be made disc resident, which will free some global common memory for the necessary expansion of the monitor data mentioned in section 2m.

These hardware and software changes would probably not be considered if a complete upgrade were imminent, but they represent the best interim solution otherwise. Other possible kludges are as follows:

b) Simple reassignment of tasks in the various computers (load leveling). This is essentially the same as section 4a but without taking advantage of the additional memory port. The data in the global common area (32k bytes) would have to be written into the memory of the newly positioned SPECTRE every 10 seconds at least. This would be acceptable for most purposes, but would make the investigation of a quickly fluctuating (<10s) phenomenon impossible. As noted by Barry, neither this nor 4a allow would provide space to accommodate a background task for program development or for real-time monitor listings.

c) Move the archive I/O from BOSS into CORA/CORBIN. Even if a 6250 bpi tape drive were available for the Modcomp II, it is clear that BOSS has neither memory to accommodate the larger buffers necessary to take advantage of the higher density nor sufficient I/O bandwidth. The easiest way to achieve this would be to add a task in CORA to collect the tape records together and write large records onto the fast drive. At the cost of channel-baselines, CORA has the buffer space in the present system. In the final "pipeline" configuration the additional I/O burden on CORA would be an unwelcome interference; the buffer space will not be available.

d) Increase the size of shared memory. This would make use of paging registers; it is discussed in detail by Barry Clark (op. cit.) in response to a suggestion by Sandy Weinreb.

e) Add a micro-computer based demultiplexor for the DCS monitor data. The proposal is to interpose a computer, either a Modcomp or a

fast micro, between the DCS and MONTY whose purpose would be to demux monitor data and maintain the monitor data base in its own memory. This is a slight variant of the plan mentioned in 4b, the variation being that the demuxor does not have shared memory, and so must send data buffers on demand via CPU link(s). This addresses two of the major problems namely the global common crunch, and the lack of available CPU cycles in MONTY. Difficulties abound however. Some are similar to the problems with the paged memory kludge addressed by Barry as his fifth point; in addition we have to worry about yet more computer links, perhaps to a foreign computer. We also have to worry about how and where to do the software development. Besides the bits of software mentioned by Barry LOG, CORMON, and SYSGAP may have to be extensively modified. With the micro holding the majority of the monitor data some memory in global common will become available, but probably not sufficient to include an on-line ANTSOL with scan averaging.

5. Upgrade

None of the stopgap solutions detailed in section 4 addresses all of the problems and they all represent a substantial effort either in software or hardware or both in a system that is at best obsolescent. A major upgrade will have to be faced sooner or later. It can be argued that the current system ought to be frozen at or near its present level and effort be directed towards designing and implementing the future system.

a) The preferable option is to recognize our software investment in Modcomps and commit ourselves to upgrading the entire on-line system to the new Classic model, or perhaps the new, yet to be announced, 32-bit product line. The software costs are small since much of the code can be transferred directly and the operating system that would be used (MAX IV) is structurally very similar to the current one (MAX III). The hardware development costs are also small since I/O devices are upward compatible. We hope that, using this approach, an upgraded on-line system could be available early in 1985 if a purchase order for the new hardware were placed in mid 1983.

The simplest upgrade would be to use two Modcomp Classics to replace BOSS and MONTY and a third to replace CORA/CORBIN. A fourth CPU to be used both for software development and also as a repository of spares would complete the system.

A possible configuration is shown in figure 1.

b) The fall-back option is to patch the current system using the proposal of section 4a to enable it to be usable for as long as possible and then go through a complete procurement procedure to find a suitable replacement system. We would propose investigating chip-based micros having approximately the same capacity of the Modcomps mentioned in the previous section (5a). It will require rewriting all software and re-designing and building all our special purpose interfaces. The main disadvantage of this is that the installation date of a new system could not be before 1987, but this may be outweighed by the decreasing cost of suitable computer hardware.

c) Upgrade of the array processor. The proposals we have outlined above all assume that the existing array processor will remain in place and be interfaced to any future system. Despite the fact that the architecture is relatively old (mid 1970s) and the technology and packaging is decidedly inferior to today's systems, it is surprising that there is currently no other similar device that performs significantly better. We assume that this will change soon. The existing capacity seems to be adequate for most reasonable requests for additional capabilities required of the rest of the on-line system in the next few years (but see section 2I). This suggests that the AP will eventually have to be replaced with a more capable system, but we propose that this improvement be considered and be carried out independently of the on-line system upgrade being discussed here.

6. Necessary Resources

Here is a rough estimate of the manpower and capital necessary to achieve some of the options.

Project	Capital Cost ($k)	Software Manpower (man yr)	Hardware Manpower (man yr)	Installation Date
4a	10	1/2	1/2	1983
4e	25	1	1	1984
5a	600	4	1	1985
5b (in addition to 4a or 4e)	350	10	2	1987
5c	100	1	1/2	1990?

Appendix

To: G. Hunt

From: B. Clark

Subj: Modcomp enhancement kludges

Date: Dec 20,1982

The first possible kludge would address the current shortage of shared memory. This would call for replacing the 16K Modcomp page with something a bit more modern, and replete with a paging register. Below are a few thoughts about how it should be done.

First, it would be too much to try to teach the I/O system about the pager. We would have to forbit I/O directly to paged memory. This essentially restricts us to trying to page the shared Boss/Monty memory, which is indeed a sore spot.

Second, it seems to me that separate paging registers are needed for Monty and Boss--I can see no way of keeping a single register synchronized without unacceptable overhead.

Third, the I/O busses, especially on Boss, are very near their length limits. I would not like to see the paging registers put on the respective I/O busses. I suggest that the memory trap a specific address, and use numbers stored in it as a paging register, or rather two addresses, one for Monty, one for Boss.

Fourth, the paging registers must somehow be protected from context switching. The alternatives are to access the paged memory with the Taskmaster locked out, or to have the Context Switcher save and restore the paging words. The former implies that the long transfers into the dataset 4 areas from Boss would lock out the real time geometry routines for a total of about 2.5 milliseconds. The latter implies that the context switcher would have to know whether it is running in Boss or Monty (unless the hardware can write protect these locations, which I doubt it can do without upsetting the timing). Neither is nice, but the latter is better.

Fifth, there are so many references to the various control blocks throughout our software that I am loath to try to locate all of them and make sure that they are properly paged. It is very much more convenient if the control blocks can be maintained in shared global common without paging, as they are now. Paging a portion of the shared area, to extend the DCS storage areas, seems to be a reasonable thing to do. Code for accessing these areas is much more localized; I can recollect only two places it occurs, in the REX,GDD service and in DCS itself. The data allocations, as of December 14, were:

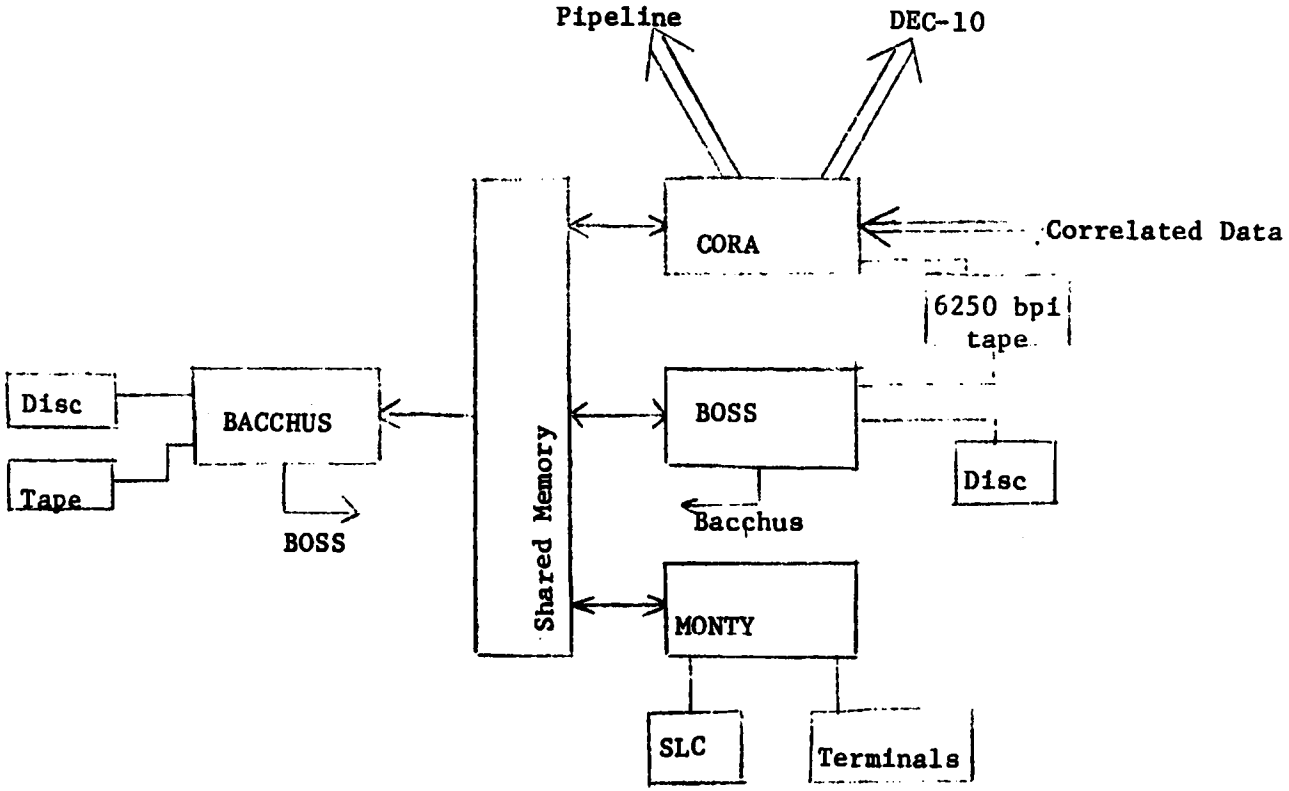

Figure 1. Block diagram of a possible configuration. This configuration is dominated by shared memory, all computers have access to it, the spare computer however will have read-only access. In addition each of the computers will have one megabyte of private memory. For redundancy, there are two large discs and two tape systems. Ignored here are such details as how the computers are to be booted and whether or not we need co-processors to run the AP, terminals and the DEC-10. It is assumed that unique peripherals will be switchable between their normal host and the spare computer.

Control Blocks	15A5 (hexadecimal)
DS 0	2BB
DS 1	A40
DS 2	7C8
DS 3	35B
DS 4	81D
DS 5	5FB
Antenna 0	28E

The suggestion is to allocate 3000 Hex words to Control Blocks and antenna 0 data, which provides a comfortable room for expansion of both, and 1000 Hex words of data paged eight deep, allocated by dataset. The tightest page is for DS 1. Even in this case the additional room avalible, 5C0 words, is sufficient to allocate a buffer for the 38 currently unused multiplexor lines, add four additional digital lines, and install averaging on five lines which do not currently have it.

In addition to the memory finagle, there is available à way to get a few more CPU cycles. This involves hanging an additional port on the shared memory, for another Modcomp II/45, and another of our traditional "everybody move one job slot left" shuffles. Spectre is a Modcomp II/45, bought that way in anticipation that some such move might be wanted in the future, and his job is simple enough that it probably could be handled by a 7810, say Eunice. We would then move Eunice into the screen room and Spectre into the main room, in sufficient contiguity to Monty and Boss to share the new, shared, paged memory. Task distribution would then be as follows:

Boss	GEO,G10,DUM,LOG
Monty	DCS,CHK,B
Spectre	SB,DMn,B
Eunice	CON,LIN,SIN,SB

The background, sitting either behind all the DMT's or behind DCS, would probably run slower than it does now.

Notice that neither of the kludges discussed above address the problem of making room for another background in which an updated version of REVIEW or MLIST could be run, nor do they provide an environment so copious in memory that a real-time ANTSOL, with scan averaging and listing of closure defects, becomes a natural development.

copies: W. Randolph, S. Weinreb, R. Ekers, P. Dooley