VLB ARRAY MEMO No. 72

(DRAFT 1)

VLBI ARRAY POST PROCESSING REQUIREMENTS

R. D. Ekers

February 3, 1982

1. Changes to Concepts in the VLBA Design Study Report Section 4B.

This is a summary of the differences which have led to an expansion of the required post processing system.

(a) Expansion over current VLBI Processing.

In the continuum the main difference is the assumption that there will be a significant amount of wide field mapping, and work on weak sources where the a priori position is not known to better than a few tenths of a second. Both of these set a similar requirement. Eg to map a half of second field of view at 5GHz would require eight channels for bandwidth synthesis and twelve seconds averaging time. The processing times given in the present report were based on one channel of continuum data with about one minute averaging time. Hence we have more than an order of magnitude increase in the post processing requirement for this particular case.

In the line case the main difference is the larger expection for the total number of channels which will be processed (see VLBA Memo #47). Such higher expectations are justified given the scope of the VLBA project compared with current practice but this can set almost impossible requirements on the post processing system. The extreme H_2O case produces a 130 G Bytes of data in a one day observation. This exceeds the maximum output ever expected from the VLA by a factor of 30, so to be realistic the proposal can't include this kind of observation. The maximum line output data rate specifical for the correlator is 400 Kbtes/sec (VLBA memo #61). This gives 14 Gbytes of data in a 10 hour observation.

The computer system changes in the following table give a maximum storage of 10 Gbytes. It can handle the moderate H_2O case and the OH case in VLBA memo #47, but it still cannot store an entire line observation at the maximum correlator output rate. This data would still have to be dumped on tape and reread for full processing.

(b) Algorithm complexity.

It is sometimes argued that the decreasing cost per unit of computer power with time will alleviate the post processing problem. However present experience indicates that this will be insufficient to offset the expansion in the complexity of algorithms which are being developed. Examples:

1) There has been two orders of magnitude increase in the complexity of problems on which the clean algorithm is being used since its introduction in the 1970's, in the same period the cost per cpu cycles has gone down by a factor_____________.

2) Algorithms to provide an error analysis of CLEAN maps, (or results of other restoration algorithm) are not being implemented because of the excessive requirements on computer time.

3) Self-calibration requirements are increasing because sources of greater complexity are being analyzed, and the increased sensitivity of telescopes is making it possible to decrease the time constant for the self-calbration averaging.

4) Other restoration algorithms such as maximum entropy, optimum deconvolution method, regularization, are likely to come into use. These will require more computer power than CLEAN.

5) Completely new techniques, such as Cornwell's proposal to handle bandwidth synthesis on a source of varying spectral index, are being considered.

(c) Software

In order to minimize operating costs and to improve long term flexibility it is necessary to have an environment in which one can write high level and well structured software. This requirement can be nicely met by the VAX type systems provided they have plenty of memory space and provided we do not depend two heavily on the array processors or other special purpose hardware.

(d) Comparison with VLA post processing computers.

The estimated cost of the current VLA System excluding synchronous computers is about $7.5M. Of this $6M is hardware and $1.5M is software. As a useful comparison for the cost of software development for the VLBA it might be useful to point out that the current AIPS system has cost about $0.5M. It is unlikely that the VLBA development will be much less than that.

The proposed VLBA post processing system is then about half what we now have at the VLA which certainly does not seem excessive.

(e) Outside support.

In order to make this processing problem tractable it is necessary to assume that there will be significant use of outside post processing systems by some users. This is becoming increasingly true in the case of the VLA as there are now 6 large systems being used in the Universities to reduce VLA data.

(f) Correlator intelligence

For some types of observations, eg small continuum sources dominated by a point component, an intelligent correlator can make at least an order of magnitude reduction in the data rate by using delay and fringe rate windowing. This is important because it enables a lot of good science to be done with an insignificant post processing load but unless the majority of VLBA observing is of this type this will not change the post processing requirements.

The use of digital filtering in the processor, as suggested by Ewing, could reduce data requirements on the post processor by a factor of 2 to 3 compared with the estimates in VLBA Memo #47. This effects all observing programs so it is an important function in determining the post processing load.

2. A Proposed Post Prosessing System.

The following table lists the components of the post processing system showing the differences between the proposal in the VLBA design study and our current thinking. It is an attempt to specify a system which assumes that an observer is likely to have visibility data in the computer for about three days, that there is one moderate H_2O user, two easy line users, two wide field continuum mapping users and any number of simple continuum users. The RM05 disks with removeable packs are added to give additional cost effective mass storage on the assumption that users will not need to access all the data all the time.

RDE/er

TABLE 1. VLBA POST PROCESING SYSTEM

COMPONENT	CHANGE	COST
4 Vax-11/780 2 M byte memory	Increase from 3 to 4. (Tape drives and mass store listed separately)	1000 K
4 Additional 2 Mbyte memory	Extra memory on all systems	80
6 TU78 high density tape drives	Increase from 3 to 6 no low density drives	200
4 RP07 disks with controller (500 Mbyte non removable)		150
8 Additional RP07 disks drives	Increase total of directly accessible storage to 10 Gbytes	200
2 RM05 removable disk drives with controllers (256 Mbyte)		100
14 Additional RM05 disks drives		500
40 Disk packs	Add removable disks to provide additional 10G Bytes of mass store without tape I/O	40
4 Array processors (38 bit 64K words)	Increase from 3 to 4	300
2 Printers (600 lpm)	Decrease from 3 to 2 (Use DECNET)	30
12 CRTs (9 text, 3 graphics)	Same	30
2 I^2S image displays	Increase from 1 to 2	140
1 Video disk (Escofier design)	Instead of commercial	20
Optical disk (Mass storage)	Same	40
Misc	Same	50
Software	Same	320
		3200 K