NATIONAL RADIO ASTRONOMY OBSERVATORY
Charlottesville, Virginia

October 25, 1977

To: Addressee

From: M. S. Roberts

Enclosed is a copy of the final post-processing committee report.

Addressees:

D. S. Heeschen
D. E. Hogg
H. Hvatum
B. Balick
W. R. Burns
B. G. Clark
E. Greisen
R. M. Hjellming
F. N. Owen
A. R. Thompson
Library (3)
Committee Members

REPORT OF THE

VLA OFF-SITE DATA PROCESSING AD HOC STUDY GROUP

October 21, 1977

I. PROLOGUE

A study group was formed on March 21, 1977 to consider the general problems related to VLA off-site data processing. This group met seven times to review and analyze the various factors involved in such post-processing. Interested parties were contacted and a variety of written and oral presentations were given the committee. A preliminary version of the report served as basis for discussion for a meeting of VLA personnel and committee members held at the site. Comments were solicited and, where appropriate, included in the final report.

The most remarkable aspect of this study, one that is not readily visible in this report, is the similarity in individual views as to what is necessary and how this might best be accomplished; there is no minority report. However, it is unlikely that such unanimity will be encountered when the specific details of post-processing are considered.

II. THE ON-SITE PRODUCT

As an underlying philosophy we accept the view that, because of its magnitude, detailed re-editing or recalibration of the raw data is neither feasible nor likely in most circumstances. However, the ability for gross editing, e.g., omitting baselines or time intervals, will be requested by some. No distinction is made in the following between line and continuum data since we envision continuum data to be taken on the A array while line data are obtained on the more compact C and D arrays: the amount of data to be handled in either case is comparable.

At the end of an observing run there should be available to the observer:

1) u-v data which has been edited, calibrated, and perhaps sorted. These ungridded data are tagged with baseline, time information. Some display of these data, most likely graphical and perhaps in an averaged form, is highly desirable.

2) A summary of diagnostics which resulted in data rejection or flagging.

3) Since on-line maps have been produced for decision making during the run, hard copy of these maps should be available. For line work, this may consist of maps of selected and/or averaged frequency channels.

and perhaps

4) To speed the throughput it is desirable (but not necessary) that at the end of the run, or in a reasonably short time thereafter, a map of fairly high quality (e.g., corrected, cleaned) be supplied the observer. (Timing here may be determined by map size and number of either line or continuum observations.) This should be available in both hard copy and Fortran readable magnetic tape. Compact (i.e., averaged, gridded) forms of the visibility data should also be available on magnetic tape.

Note to above: This last item regarding maps and magnetic tapes is considered only as "desirable" as part of the on-site product. It involves decisions by the observer (i.e., taper, size, number) that can be made after he is off the telescope without affecting the observations. When requested these maps should be supplied as soon as possible.

A minimum amount of map manipulation and extraction of numerical information may be sufficient for some projects and should therefore be obtainable on-site immediately after the observing run. This point requires some elaboration.

Map manipulation must be possible on-site for staff use and experimentation. If excess capability exists (e.g., nighttime), then usage by the most recent observers should be supported. In a similar vein excess capacity in the computer system is likely during many experiments. Transform operations (on a low priority but still at a useful rate), in addition to those made for the quick-look maps, should also be supported. Such usage must be carefully restricted, but the possibility of limited on-site post-processing for recent observers should be made available if possible.

The u-v data are the fundamental product. As questions arise during examination of the maps, different tapers, fields of view, and source subtraction (for example) will often make retransformation of the u-v data necessary. Gross editing of these data may eventually be little used, but it surely will be in demand initially.

A viable storage, transportation, and archiving system of the u-v data is an important problem for the VLA. It is not addressed in this report.

III. POST-PROCESSING

The output described above will be adequate for some experiments: e.g., detection, variability, position, approximate size. But for many problems further, i.e., post-processing is a necessity. A two-page list of such processing needs is given in the memo "VLA Post-Processing: Phase I", by W. R. Burns and E. W. Greisen (March 25, 1977), see Appendix A. In brief, the astronomer wants (i) to be able to produce results of publishable quality, (ii) to be able to experiment with the data, and (iii) to re-examine the data, especially as new questions arise. And the very nature of the VLA with its movable elements requires a time scale for data reduction (for some observations) which cannot be done in real time. Post-processing is a necessity. It may be performed at various locales:

1) At the user's home institution;
2) At an NRAO computing center;
3) Via time and space rented at a large computing facility;
4) Via a number of small (mini) computers dedicated to VLA processing and located in various high user-density areas; and
5) Some combination of the above.

Two additional possibilities may be included for completeness: a general astronomical computing center, and on-site post-processing. The former appears too far in the future and would require significant staffing with astronomers familiar with VLA techniques. The latter is not within the purview of this committee.

Each of these possibilities have been considered in detail. The recommendation of this committee is (5); specifically, a combination of (2) and (4). It is described in some detail below. The reasons for rejecting the other alternatives will be given first.

For many observers, the ideal reduction and post-processing procedure would have the user return to his home base to work on his data while involved in other commitments, such as teaching. If he returns with u-v data for just one typical 12-hour observation, he would likely be carrying one or several high-density (6250 bpi) tapes. The handling of these u-v data, even for the minimum situation, would be a major burden for a university computing center. This first step can be bypassed by going directly to a map-tape (as described above).

While some users will require only limited analysis of their map data, we envision that the bulk of the post-processing will require interactive manipulation of these data. To accomplish this a disk drive dedicated to that one observer must be available as well as a display terminal. The latter will require a coaxial cable link to the computer. An additional constraint is the potentially large amount of cpu and "real" time requirements for typical work, See Appndix B for time estimates. Finally, software development assistance from NRAO will be sought; this is a reasonable involvement for NRAO but one that may be difficult to satisfy because of the many different machines and systems that are in operation.

All of these hurdles can be overcome, but it seems to be a grossly inefficient approach. The concept of interactive map manipulation at many university computer centers is not a likely approach to post-processing.

The time-rental concept, unless at some already operating government facility, means that we (the astronomical community) are willing to pay a premium for operating management. The flexibility that may be necessary for VLA post-processing will likely be absent. Going to Los Alamos or Livermore (the former was casually looked at) may, even more, remove the flexibility and will always pose, at some level, a security problem inherent in such installations. At present, 12^h/day shifts on a CDC 6600 are available at Los Alamos. This may increase when their new Cray becomes operative. Renting time is feasible but it cannot be described as desirable in so far as we have been able to ascertain.

Remaining possibilities include an NRAO center, many minis, or a combination of these two. Many minicomputers dispersed about the country without an NRAO center is not viable. Many questions (which we make no attempt to answer) are raised in this extreme situation, e.g.: who develops and maintains the software for this battery or minis? How are special projects and needs handled? Is this a cost-effective (and science-effective) approach? These questions and concerns can be answered satisfactorily by replacing the many minis with a large central computer facility. But other disadvantages can be envisioned: The potential for bureaucratic inefficiency; the necessary lack of flexibility in approach and in programming; and perhaps most important, a computing center would keep users away from their home institutions.

Must we have one extreme or the other? We hope not. A central computer facility for development work and for some visitor (and staff) usage is necessary for a repository of VLA knowledge and expertise; it is a center for both normal and special jobs. In addition to this central (presumably modest sized) facility, several minicomputers would be available for regional use. These computers should have common, NRAO supplied, programs written in Fortran, where practical, and which can be modified by local users. Price estimates of $150,000 to $300,000 for a minicomputer, array processor, and interactive display system have been quoted. These will allow interactive map manipulation. Longer jobs (e.g., retransforming, extensive cleaning) may be run overnight if necessary. Less expensive, more modest systems also appear feasible.

Problems remain. Who gets such systems? Are they regional in the sense that those outside the institution housing such a system will have access to the system? And if so, who will maintain the machine and aid the visitor? We think such problems and the inconveniences they imply will be more than outweighed by having a reduction system in one's own backyard. An added responsibility will go with this added convenience. The user distant from such a local computing system or one with special needs would go to the central NRAO facility which would also be equipped with such minis (or a system which to the user should look like a mini). This NRAO center, in the situation envisaged here, is of modest size, perhaps not much larger (physically) than our present computer center. In addition to a computing facility, such an NRAO center should have auxiliary equipment for the most productive use of the VLA-generated data, e.g., a photographic plate measuring machine, overlay and related photographic facilities. This center must also have a scientific staff to aid the visitor in use of the computer programs and in the problems of data manipulation that will arise.

The development and growth of such a center must proceed in a natural and orderly manner. We already have some experience with post-processing on large computers. Experimentation as to the large computer's role and capabilities should continue. A minicomputer system with necessary peripherals should also be developed and tested. Once this latter system is established, others can copy both the hardware and software.

IV. IN SUMMARY WE PROPOSE:

An NRAO computing center and a number of regional, university-based, minicomputer systems dedicated to VLA post-processing.

The NRAO computing center would contain a modest sized general purpose computer and necessary peripherals for the maintenance of active NRAO programs. It would have facilities to service several VLA post-processing users simultaneously. Re-editing, if requested, could be done on-site, or with the above general-purpose computer, whichever is more feasible. Scientific and technical staff would be located at this center to aid the user. Close communication between this center and on-site personnel is of great importance.

Several (3 to ?) minicomputer systems with adequate memory and interactive displays would be located at supportive universities across the country. An identical minicomputer system would remain at the NRAO computing center for development work. These systems are fast enough to handle many post-processing tasks in a reasonable time.

V. RECOMMENDATIONS

The above is the background and justification for the following recommendations:

1. Identify and delegate responsibilities and authority to a post-processing group to begin detailed planning and implementation of the post-processing system.
2. Begin immediate development of post-processing by a minicomputer system, including software and hardware.
3. Develop preliminary cost projection of the post-processing system, including site location options and begin budgetary action now.

Study Group Members:	R. L. Brown	W. Jaffe
	J. J. Condon	M. S. Roberts (chairman)
	D. S. DeYoung	L. Rudnick
	J. R. Fisher	

VLA POST-PROCESSING: PHASE I

W. R. Burns, E. W. Greisen

March 25, 1977

I. Introduction

In this memorandum we discuss in some detail our ideas concerning what we call Phase I in the development of a VLA post-processing capability. Phase I has already begun and will continue, we anticipate, until 1980. This phase is characterized by low levels of personnel and purchasing commitments and by the use of existing facilities in Charlottesville including the IBM 360/65 and the Modcomp IV mini computer. We divide this memorandum into three parts: current accomplishments, short-term goals, and longer range plans. Phase II, the development of post-processing in the 1980's, is not addressed here.

II. Current Accomplishments

The first goal in the Phase I effort is to provide some support for users of the VLA. In the early stages of Phase I this support will be limited to procedures similar to those in use with the Green Bank interferometer. Of course, as the more advanced capabilities discussed in Section IV of this memorandum are developed, they will be made available to users as well.

At the present time we have developed for the IBM 360/65 a package of PL/1 programs which can at a single frequency per pass, read, edit, calibrate, display, and map the output of the Synchronous Subsystem of the VLA. The adopted format allows for data from as many as 351 baselines. User control of the programs is via Job Control Language (for data sets), PARMS cards (for option selection), and INCLUDE/EXCLUDE cards (for data subset selection). The last of these provides a particularly powerful and versatile tool for editing the data throughout the many processing stages. The program VLATOIBM is used to convert the Synchronous Subsystem tapes into our format and may be used for preaveraging and minor editing and concatanation operations. VLAINDEX is used to provide a summary of the data set. VLACLEAN is an editing program which employs user-set levels in automatically deleting data affected by shadowing or identified as marginal by the reliability flags set by the Synchronous Subsystem. It may also be used to delete data points whose amplitudes deviate excessively from the mean. VLACOPY is used to copy data sets, or to carry out "manual" editing via the INCLUDE/EXCLUDE cards. VLAPOS adjusts phases for changes in phase reference position. VLACORR corrects phases for atmospheric effects and for corrections in antenna station coordinates and corrects amplitudes for elevation effects. VLABCAL solves for corrections to the antenna station coordinates and, optionally, to the positions of the calibration sources. VLACAL fits a time varying complex instrumental gain function to data on calibration sources and corrects the data. VLAPCAL determines the instrumental polarization parameters and corrects the data. The observations may be averaged to selected integration times using VLAAVE, and they may be sorted on a wide variety of parameters using VLASORT. A flexible selection of printer displays are provided by VLASCRIB and VLAVIST. These include plots and a choice of compact or expanded printouts of visibility data as a function of time, printouts of visibility data as a function of projected spacing, and plots of one selected parameter versus another with a wide variety of selections available. There are two u-v plane model fitting programs: VLAPTSRC for unconfused point sources and VMODFIT for models consisting of sums of elliptical gaussians and for the polarizations of such models. The program VLAMODEL allows these models to be added to, or to replace, the observed visibility functions in the data set. VLAMAP sorts, grids, and Fourier transforms the data. A number of map processing routines are also provided including VMCLEAN (for the Högbom CLEAN procedure), VMCOPY (for moving and editing map data sets), VMINDEX (for listing map data sets), VLAPRINT (for printer displays), VLADRAW (for Calcomp contour maps), and VLAPIX (for photographic representations).

These programs and formats are documented in the "User's Guide for VLA Data Reduction in Charlottesville", NRAO Users Manual Series Number 29 by Eric W. Greisen and Frederic R. Schwab. Background information concerning the use of data sets and the Job Control Language and extended discussions of some of the algorithms may be found in "User's Guide to 360/65 Interferometer Data Processing Programs", by L. C. Blankenship, R. M. Hjellming, B. L. Meredith, and F. R. Schwab. These two manuals are maintained on Pandora text editor files in Charlottesville and copies are available upon request.

Some rough edges do remain in the programs and their documentation. This is in part due to our present limited understanding of the instrument. At this writing, the principal problems are:

(1) final implementation and testing of the atmospheric and elevation corrections in VLACORR,

(2) detailed testing of VLAPCAL, and

(3) completion of the documentation.

In fact, the User's Guide is currently handled as a working document. New routines are documented before they are fully tested (and occasionally before they are written). The current status of each program is, however, indicated in tables at the beginning of each chapter.

III. Short-Term Goals

We assume that principal input to a VLA post-processing facility is u-v data which have been edited, calibrated, and partially sorted by the Asynchronous Subsystem of the VLA. In order to acquire early experience with such data and in order to assist in the development of the Asynchronous Subsystem and to support users of that system, we feel that our next task is the creation of, at least, a minimal package of programs to deal with Dec-10 output. Thus, during the next 6 months we expect not only to clear up the problems mentioned in the previous section, but also to develop a capability similar to that described above for output from the Asynchronous subsystem.

The requirements of the Asynchronous package are somewhat simpler than those of the Synchronous package. Counterparts of VLATOIMB, VLACLEAN, VLACORR, VLACAL, VLAPCAL, VLAPCORR, and some parts of VLASCRIB will not be needed. Further more, the map programs VMCLEAN, VMINDEX, VMCOPY, VLAPRINT, VLADRAW, and VLAPIX may be used by both packages. Our current personnel (i.e., Eric Greisen and Fred Schwab) should be enough to accomplish these goals in the stated time. A format has been determined and we expect to begin coding the package soon.

IV. Our Perspective on Future Development and Goals

The program packages described in the preceding sections are designed to meet the ongoing requirements of the VLA during its developmental stage. These packages are expected to neither handle very large volumes of data with practical efficiency nor are they powerful or flexible enough to provide substantial improvements over data reduction procedures now in use with older, 2 to 4 element interferometers. The programs which process Synchronous Subsystem output will fall into disuse quite naturally as the Asynchronous Subsystem becomes more highly developed and more reliable. However, programs to handle the output of the Asynchronous Subsystem will need to become more extensive, flexible, and efficient. This latter development is the subject of the present section of this memorandum.

The project outlined below is a limited one which we feel is suited to Phase I in the development of a VLA post-processing capability. Crash projects, involving large numbers of people, or projects which try to solve problems expected to arrive many years in the future, work even less well in software development than in most other areas. The correct approach, we feel, is to develop programs gradually which work on real data to do useful things for real observers. To make the gradual approach acceptable and to allow close interaction between observers and programmers, the project must begin when the number of observers is still quite small. The intent of the project is to develop as much post-processing capability as possible, but not to worry so much whether it can support multiple users and whether it is as efficient as it might be. The proposed system could be viewed as a test system where the emphasis is to give a single user as powerful and versatile a capability as possible. Early work will be limited to continuum reduction, but spectral-line routines might be added when the need arises.

In order to determine what programs are likely to be useful, let us consider the common case of a continuum observer faced with a field containing several sources of interest. The data are first calibrated and edited at the site. Maps, probably at less than maximum resolution, would then be made. At the post-processing facility, the observer begins by inspecting his maps for signs of "bad data" (e.g., ripples or other strange regularities). This inspection may be done most quickly and with substantially better sensitivity using television-type display devices. When questionable areas appear on the maps, the observer would review the calibration data used for the source (and possibly also the u-v plane data taken on the source itself) in order to isolate the bad data. Plots of amplitude and phase versus time for the suspect baselines and possibly gray-scale displays of the u-v plane would be needed. Once any questionable data are located, they might be deleted and new maps made.

The next step could involve fine tuning the standard corrections made to the maps. The observer may wish to examine the effects of correcting for edge attenuation due to the primary beam and to the smoothing and resampling of the FFT. He might decide to correct the maps for the delay beam. Whether or not one makes this correction and exactly how this correction is made depends on the nature of the experiment and the quality of the data. In any case, the astronomer will certainly wish to inspect the result to determine whether he wishes to use the corrected map or a smaller portion of the original map. Single-dish data may now be added to the data to fill in the central portions of the u-v plane. The Westerbork experience indicates that this is a delicate process requiring careful calibration of the single-dish data and close inspection of the results of their inclusion in the maps.

Next the astronomer will need to decide on one or more resolutions and map sizes. While we will probably encourage the use of the smallest FFT which will not degrade the final results, the astronomer may note that a confusing source probably lies just outside his field of view and thus may need a larger transform to isolate it. In order to examine extended features the astronomer will wish to subtract any strong point source from his map (via a CLEAN-like program) and then convolve at least a portion of the resulting map to a lower resolution.

Once the astronomer has finished correcting and massaging his maps, he is ready to derive the useful information from his data and to produce the final displays. In the present multi-source continuum example, he will want to isolate the sources on the map and to estimate their positions, sizes, and flux densities. The first step in this modeling process is to search the map plane for sources and to determine good "initial guesses" for their parameters. For a complex field the astronomer will need to assist the programs, guiding them to the sources and providing them with source-dependent control parameters. When a good set of initial guesses has been determined, a fitting routine, which operates on data in the u-v plane, will be used to refine the parameters. The astronomer will again require various methods to compare the model with the data. Should straightforward model procedures be inappropriate for portions of the map, the astronomer will wish to examine such portions in detail and measure parameters such as the widths in various directions, the peak brightness, etc. For many objects, the next stage is an examination of polarization and spectral index maps. A variety of map presentations are required at this point which include good procedures for blanking regions in which the parameters are unreliable. Finally the astronomer will examine his maps to determine which data are required for documentation and publication of his work and which regions, parameters, and methods to use for published displays of his data. His production of these displays (picture, Calcomp, etc.) end his processing task.

The above operations all involve a detailed interaction between the astronomer and his data involving many correction and parameterization steps followed by examinations of the results. These operations could be accomplished through traditional batch mode processing such as is currently done with the Green Bank interferometer. However, we feel that the long waits required in batch mode for the results of each suboperation lengthen the data reduction process enormously and cause the astronomer to lose his concentration and to take shortcuts in the processing. The latter effects can degrade the final results. The combination of real-time displays with a responsive computer system, similar to that used for single-dish processing at NRAO, would allow the astronomer to process his data more rapidly, to maintain his concentration on those data, to check fully the results of his data manipulation, and to discover, more easily, unexpected problems - or results which may be present in his data. The process of data reduction depends heavily on the nature of the particular observations and on the likes and dislikes of the particular observer. However, experience (e.g., TPOWER/CONDARE) has shown that the presence of a responsive, real-time, data display/reduction system enhances the scientific output and makes efficient use of the computer hardware.

We have alluded in the discussion above to a variety of software routines required for continuum data reduction and analysis. Below we will present a condensed list of the capabilities which should be present, eventually, in the proposed software system. We do not discuss spectral-line problems here because they are somewhat further down the road and because the discussion given in "The VLA Spectral Line System: A Progress Report" (NRAO, January 1976, pages II.B-3 through II.B-15) covers the subject well.

A "wish list" for continuum VLA post processing:

A. Data integrity examination and correction

1. Map routine for large fields with minimal resolution and just adequate accuracy

2. u-v plane display routines including gray-scale (source data) and baseline-time (source or calibrator data)

3. Bad data deletion routine

4. Minor numerical routines for estimation of amplitude and phase variations and noise.

B. Map correction

1. Gridding (e.g., sin x/x)

2. Primary beam

3. Delay beam

4. 3-D smearing

5. Center hole in u-v plane

6. Polarization beam

7. Internal calibration by observed point source response.

C. Map restoration

1. Include data from other telescopes particularly to fill center hole in u-v plane

2. Point source subtraction

3. Clean

4. Maximum entropy on small fields

5. Map combination including fields at different phase centers

D. Convolution of restored maps

1. Gaussian

2. Beam of another radio telescope

3. Extended source enhancement

E. Parameter estimation

1. Position, flux, angular size fitting for simple source components in map plane

2. u-v plane model refining

3. Detailed examination at high resolution of small fields surrounding source components

4. One dimensional fits across extended sources at arbitrary angles

F. Map comparison

1. Polarization

2. Spectral index

3. Rotation measure

G. Special problem software (low priority)

1. Astrometry/proper motion

2. Circular polarization

3. Pulsars

4. Planetary system

H. Display

1. Gray scale (TV and, for hard copy, Dicomed) including labeling, contrast enhancement, pseudo coloring, contour or polarization overlay, et al.

2. Contour and line drawing outputs in both publishable and quick and dirty forms

3. Printer output for processing documentation and detailed results.

Proposed Development System

The capabilities listed above can only be offered with a great deal of work and with a small, but not insignificant, capital expenditure. In the latter area, because of the fundamental nature of map display in the data processing procedure, acquisition of a display system is of high priority.

Experience has shown that unresponsive (i.e., very slow) interactive systems frustrate the user and hence, are ineffective. Thus, when hardware bottlenecks can be identified, corrective hardware (such as an array processor or more core or disk) should be acquired.

We propose to use a combination of the IBM 360/65 and the Modcomp IV both as stand alone computers and as a connected system (at least to the RJE level). Items which involve the Modcomp will normally be made available on a one-user-at-a-time basis, while items involving only the 360 will, of course, be multi-user. The proposed project schedule includes; during the first year, the addition of a graphics capability to the Modcomp (not later than August 1977), with the completion of at least initial versions of roughly half of the above capabilities by mid 1978. In the first year, effort will be concentrated on simply producing the capability. In some cases, its use will be cumbersome because of a lack of a particular piece of hardware or simply because attention was focused on acquiring capabilities rather than making them efficient. In the second year (1978-1979), we expect to refine the algorithms (e.g., to account for 3-D effects and the like), to speed up those that are particularly inefficient, and to offer nearly all of the capabilities listed above. Hardware requirements during the second year will probably include additional core and disk for the Modcomp. During the third and final year of Phase I an array processor would be added to speed up the programs, the algorithms would be extended and refined, and some capabilities peculiar to spectral line processing added. By mid 1980 the post-processing station should be as flexible and as fast as we know how to make it, though it will still have a very limited throughput. An integral part of the entire program will be the close interaction with observers to improve and develop the conceptual basis as well as the tools of map analysis.

The total proposed project cost, to be completed in 1979, is $225K. The immediate step is a microprocessor-based display system ($75K) which should be procurred by June 1977 and installed by August. It is felt that the particular order of items during the following two years should be determined as hardware bottlenecks are found in the use of the system. It is not now clear, for example, whether a little more core or a lot more disk would buy the greater improvement in throughput. The tradeoffs will become clearer when we see which routines are most heavily used and which are the most frustrating.

Future - Phase II

In the 1978-80 time frame NRAO must specify and develop whatever post-processing facility will be used in the 1980's. This system must service the needs of the full, normally operating VLA including both continuum and spectral line observations. Many questions remain unanswered regarding this future facility, not the least of which is the question of where NRAO's basic responsibility ends with regard to the extent of supported data processing. Also unclear is the level of funding NRAO will have at its disposal. These difficult questions are not addressed in this report. It is important to recognize, however, that regardless of what level of post-processing the NRAO decides to support, the two or three year program outlined here can be viewed as a necessary prerequisite. The kind of facility in need will not evolve by simply signing a contract with IBM or the like. Rather, at least several man years of difficult work is first required. It is that endeavor which is proposed here.

NOTE ON MINICOMPUTER TIMES FOR VLA POST-PROCESSING

W. R. Burns and F. R. Schwab

July 26, 1977

General Comments on Minicomputer Versus Big Computer

For purposes of this discussion I define a minicomputer as a computer system which in total cost is less than $250K. Such a system might well have a 500 ηsec cycle time, 1 megabyte of memory, a 1600 bpi tape drive, 200 m byte of disk, and an interactive grey scale display device. It is clear that such a system would be adequate for solving many of the data processing requirements of the VLA continuum system. Some specifics in this regard are outlined in the following section. A statement of caution, however, should be given. Most of the cost of the multimillion dollar large systems is in the software, not the hardware. The conveniences that this more highly developed software brings one are real, but are not always apparent in the present type of comparison. This is largely because the better software buys flexibility primarily and when specific tasks are compared, flexibility normally has little merit. In the real computer world, however, flexibility is a very significant factor which should not be underestimated. Additionally, the larger computers in general have much more rugged hardware, but reliability has little relevance in the kind of comparison made here. In brief, if a mini system is used by only one or two people at a time, and if it is used to do rather specific, well-defined and preset things, it can be very helpful. If it is pushed to do rather more than these tasks, it will probably do nothing well.

Timing Comparisons

What is the real time required to run a particular program on a minicomputer of the type one would buy in the next few years? This naturally depends on the type of program, the amount of money and disk available, the kind of hardware, etc. An assumption which is probably no worse than most others is that the real time in question would correspond well to the real time used by the program if it were run on the IBM 360/65 in small partition on a stand alone basis. As such we have tried to estimate the real time for a number of continuum interferometer programs under these conditions.

The program selected are the FFT, contour map drawing, map linear interpolation and scaling for the Dicomed and the clean algorithm. For most of these programs the run time could be either increased or decreased by making changes in the algorithm. The run time of a contour package, for example, depends very heavily on the type of contour algorithm used. As such the times here are given without regard to whether or not probable changes would be made in the algorithm. This is not unreasonable since the particular algorithms timed have evolved over many years of use.

1. Fast Fourier Transform - FFT

   (a) 256 x 256 complex - single precision.
   Estimated ratio of real/CPU time = 2.5
   CPU time = 40 sec
   Real time = 100 sec

   (b) 1024 x 1024 complex - single precision
   CPU time = 13 minutes
   Real time = 32 minutes

2. Contour Map

GPCP algorithm - dirty map of fairly complex source

512 x 512 points

15 levels - 10% contours

Estimated ratio of real/CPU time = 4

CPU time = 9-1/2 minutes

Real time = 38 minutes

3. Prepare Dicomed Tape

a) Output grid = multiple of input grid

Input array = 128 x 128

Output array = 850 x 850

Estimated ratio of real/CPU time = 3

CPU time = 1.5 minutes

Real time = 4.5 minutes

b) Input grid does not match output grid

Input array arbitrary

Output array = 820 x 820

Estimated ratio of real/CPU time = 2.5

CPU time = 2 minutes

Real time = 5 minutes

4. Clean Algorithm

a) Map size = 128 x 128 (resident in core)

Beam is read in 4 rows at a time

Estimated ratio of real/CPU time = 2

CPU time = .8 sec per iteration

Real time = 1.6 sec per iteration

Time changes as square of linear dimension, i.e., for 500 iterations of a 1024 x 1024 map we get 500 x 8^2 x 1.6 = 14 hours. (requires $\approx$ 4 megabyte of core)

These times could probably be reduced by a factor of 25 with the addition of an array processor on the minicomputer at an additional cost of $50K, i.e., a nominal total system cost of $300K.

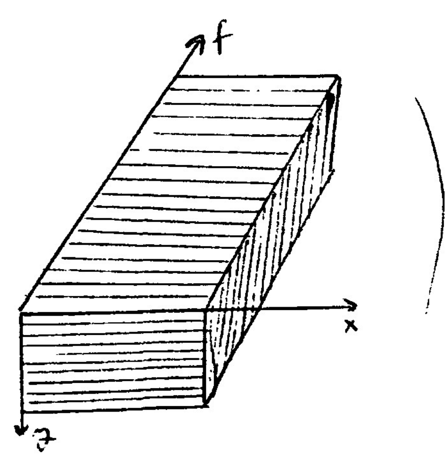
b) Same as (a) but with map stored on disc instead of core.

Increase ratio of real time/CPU time from 2 to 6.

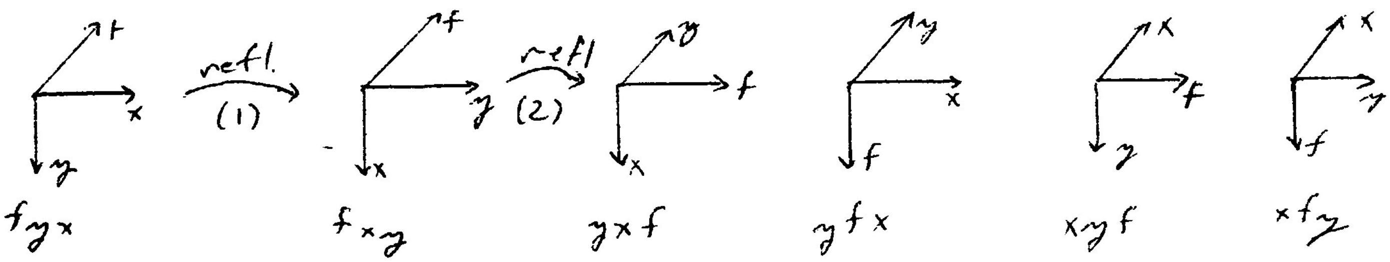
5. Transposition of Spectral Line Maps

In spectral-line aperture synthesis the output of the Fourier transform and map correction step is a series of spatial maps (on RA-Dec or x-y axes), one for each desired velocity. In computer terms, the map values are in the order {fyx} (when the right-most coordinate varies most rapidly). For purposes of spectral analysis one needs to rearrange the data to obtain the spectrum at each map point (i.e., to place the map data in {yfx} order). Such a transposition is also useful in generating maps, familiar to users of the 300-foot telescope, in which the axes are right ascension-velocity at fixed values of declination. The computer operations required to perform this transposition are described below.

A timing run was performed on the CV 360/65 for transposing (reflecting) simulated spectral line maps held on disks in model 3330 disk drives. The data that were operated on can be thought of as cubes and other right-parallelepipeds of numbers (16 bit integers), stored sequentially as rows from successive planes lying parallel to one of the faces of the parallelepiped.



The data sets used in this test were of a type in which map rows could be accessed either sequentially or in a random-access mode. The input/output operations performed in the test achieved reflections of the axes, as drawn above, so that sequential access of a new data set would produce rows parallel to a different axis than in the prior data set, but so that one axis would remain fixed.



The timing information tabulated below comprises the following operations:

(a) Generating a file of test data in fyx order.

(b) Performing the reflection (1) to generate a fxy data set.

(c) Performing the reflection (2) to write a yxf data set.

(d) Reading chunks of test data from each file to verify that the program performed correctly.

Operations (a) and (d) required negligible time in comparison to the total.

The reflections were achieved by reading single entire planes (matrices) of data into core, transposing in core, and writing the data out again. Note that step (b) requires only sequential access of the fyx data set, while step (c) requires random access of the fxy data set. Unbuffered, unblocked Regional (1) data sets were used, with record lengths sufficient to hold only one row of data. Data access was always to two different disk units. Chained scheduling was used in all cases to obtain priority in disk access.

Map Dimensions	Real Time (min.)	CPU Time (min.)	Real/CPU	Core (bytes)
255x128x128	60.08	13.07	4.6	64K

During this run, no other jobs were allowed to run in the computer.