Hardware Requirements Summary: Clarification of our GC6.2 Submission

An Internal Report

March 24, 1998

M. L. RILEE[1]
RSTX
NASA GODDARD SPACE FLIGHT CENTER
GREENBELT, MD 20771

ABSTRACT

Herein I answer questions about the Solar-Terrestrial Probe (STP) Application. REE/JPL (Remote Exploration and Experimentation) has asked us to clarify a few points we made in our report GC6.2. They have produced and have asked us to comment on a summary of the hardware needs of REE/STP. We also present two executable numeric models and results that summarize our applications (section 3).

1. Plasma Distribution Function (PDF)

A numerical model of the PDF is presented in section 3.1. A potentially confusing point is that in this document and in our report for GC6.2, N_d stands for the number of particle species to be sampled. At the REE meeting at JPL in January, N_d stood for the number of physical particle detectors that sample the plasma; this quantity has been buried in R_t, the sensor's plasma sampling rate. The possibility of confusion is enhanced by the fact that the number of physical devices, six, is about the same as the number of species to be analyzed, six.

1.1. Interprocessor communication

The discrepancy between sections 2.1.2 and 2.3.2 is as follows. In section 2.1.2 of the REE-STP-GC6.2 report, the costs refer to one iteration of the SVD based analysis for one

[1]Please address comments and corrections to michael.l.rilee@gsfc.nasa.gov. Ph. 301-286-4743 Fx. 301-286-1777

particle species. To obtain the total communication cost, one must multiply the SVD based cost as calculated in section 2.1.2 by the number of particle species N_d and the number of iterations $N_{\rm iter}$ required for calibration.[2] Thus the equations of section 2.3.2 are obtained.

The first equation in section 2.3.2 is misleading. In Eq. 3 of the GC6.2 report we have left a factor of $P^{1/2}\zeta$ off of the second term of the right hand side. Since $P \geq 1$ and $\zeta \sim 0.1$ measures the fraction of the data used in the SVD analysis, Eq. 3 of the report still fairly estimates the communication volume for the parallel SVD. However, a more correct Equation is as follows:

$$C_{\rm c-volume} \sim N_{\rm d} \, \left(N_\psi + 14 \, m\zeta N_\psi \, n_{\rm iter} P^{1/2} \log_2 P \right) . \tag{1}$$

Recall that the estimate for the size of the array in the SVD calculation is $\sim m \times N = m \times \zeta N_\psi$. The first term on the right hand side of Eq. 1 is the cost communicating the data from the sensor to the processing element (PE). The second term on the right hand side of Eq. 1 is communication cost of the parallel SVD algorithm (cf. *ScaLAPACK User's Guide*).

The right hand side Eq. 4 of the GC6.2 report is arguably lacking a factor of ζ, and should be written:

$$C_{\rm c-latency} \sim N_{\rm d} \, (17 \, N) \, . \tag{2}$$

1.2. Primary storage

The reason the primary storage is so large is that it presumes we are doing an SVD analysis on data sets of size $\zeta N_\psi \times \zeta N_\psi$. Performing this analysis on one tenth, $\zeta \sim 0.1$, of the data may be extravagant. For *Magnetic Constellation*, the minimum mission requires SVD analyses of size $m \times N$, where m varies between ~ 30 and ~ 200. The ability to analyze larger data sets, between sizes of $m \times N$ and $N \times N$, would allow us to determine basis functions from the data itself using Karhunen-Loève (KL) methods of analysis. In short, if primary storage is in short supply, we impose basis functions on the data that may or may not respect the structure of the plasma's phase space.

If resources are plentiful, we can use methods that extract such structural information from the data itself. Even if it is too expensive to perform KL analyses on PDF data on a regular basis, it is desirable to perform such analysis on the statistical moments derived

[2]The number of iterations may vary from species to species: $N_{\rm iter} \approx 3$ for the electrons, and one iteration may be sufficient for other species.

from the raw PDF data. These analyses will aid the *Constellation* in the detection of coherent structures in the Solar-Terrestrial space environment.

2. Solar Imaging Radio Array (SIRA)

In this section we discuss the *SIRA*. A numerical model of the data burden of this mission is presented in section 3.2. In that section, many parameters of the system and their relationship have been defined. In these estimates we have glossed over important calibrations and image syntheses that may be as expensive as what we present. Finally, for *SIRA*, the computing resources can be devoted to working on this data set. For the electromagnetic field part of the *Magnetic Constellation* type missions, about half the computing power for Science will be available.

2.1. Array configuration

Currently for the REE-class *SIRA* mission, we envision a large, capable central spacecraft coordinating the activities of small highly specialized spacecraft. A scalable compute engine on the central craft will correlate radio signals from the array elements: this problem is readily parallelizable because of the large number of element pairs. We are currently considering the possibility of using $N_{sc} \sim 64$ array element spacecraft. More array elements allow better quality images to be obtained faster. In fact, the better sampling can *enable* quantitiative methods by reducing the artifacts introduced by image synthesis.

2.2. Interprocessor communication

The main communication steps in *SIRA* occur during the transmission of data from the sensors to the PE (or store) and during the cross multiplication step. The sensor-to-PE transmission has been refered to as *data input*:

$$\mathrm{data\,input}: N_d\, N_{sc}\, N_\psi\, \tau_d^{-1}.$$

For $N_d = 2$, $N_{sc} = 64$, and $N_\psi = 4000\,\mathrm{measurements}$, and $\tau_d = 1\,\mathrm{sec}$,

$$N_d\, N_{sc}\, N_\psi\, \tau_d^{-1} \approx 2\,\mathrm{MB\,sec}^{-1},$$

where we've assumed four 8-bit bytes to the measurement. During the cross multiplication step, reduced data from the sensors is combined (via multiplication) pairwise. There are

$N_{sc}\,(N_{sc}-1)/2$ baselines, i.e. pairs of spacecraft, and for each pair of spacecraft there are N_d^2 combinations of antenna elements to consider. Figure 1 shows the code which we've used to calculate the numbers in this section.

The simplest way to parallelize the $N_{corr} = N_d^2\,N_{sc}\,(N_{sc}-1)/2$ cross multiplication is to broadcast the frequency and time averaged Fourier transformed data across the processing elements. The cost of this communication is:

$$C_{cross-corr} \approx 512\,N_d\,N_{sc} \times\ 4\,\mathrm{bytes}.$$

For $N_d = 2$, $N_{sc} = 64$,

$$C_{cross-corr}/\tau_d \approx 260\,\mathrm{kbytes\,sec}^{-1}.$$

Therefore, the communication burden of this part of the calculation is not terribly demanding. Once the each node has its own copy of the data, one can keep local the task of computing the cross multiplication and inverse Fourier transforming to obtain cross spectra. We do not envision implementing multiprocessor Fourier Transforms for this application.[3] In aggregate, all pairwise cross correlations are computed and stored for transmission or further processing.

```
(let*
  ((signal-size  512)
   (nd           2)
   (nsc          64)
   (bytes        4)
   (td           1)
   (npsi         (/ 125e+3 32))
   (p            nsc))
 (format t "n-corr-commo: ~a~%" (* signal-size nd nsc bytes (/ 1.0 td)))
 (format t "n--psi-commo: ~a~%" (* nd nsc npsi bytes (/ 1.0 td))))
```

Fig. 1.— The lisp code used to calculated communication quantities for *SIRA*.

Finally, we note that the interprocessor communication costs for these correlative analyses are sensitive to parameters in our models. In particular, N_ϕ tells us how much of the cross-correlations to keep; physically this corresponds to pointing the synthetic aperture to

[3]One might argue that we have been trained to avoid multiprocessor FTs by current computing architectures.

N_ϕ different locations in the sky. The number of spacecraft N_{sc} is more important than N_ϕ because the computation and communication costs vary as the number of element-element baselines which scales as $\sim N_{sc}^2$. Likewise, communication and computation costs scale with N_d, the number of antenna elements per spacecraft, but this number is likely to remain fixed at 2. There is an image synthesis and processing element to our application that we have not yet fully assessed. However, our image processing needs are not likely to be much greater than those of the other REE Science Application teams.

3. Computational Models

3.1. PDF Model

The following code applies to section 1.

```
;;; Start a closure for a probe.
;;; These numbers are for the PDF experiment.
(let*
    ((td    3)      ; measurement cadence (seconds)
     (rt    800)    ; measurement rate (per-second)
     (rx    100)    ; energy resolution (energy-bins)
     (nd    6)      ; number of species
     (nsc   64)     ; number of spacecraft
     (m     30)     ; number of moments/modes to find
     (npsi          ; amount of data per measurement
      (* rt rx td))
     (nhat  m)      ; reduced data per measurement
     (zeta  0.1)    ; fraction of data used for SVD-like analyses
     (n             ; data used for SVD-like analyses
      (* zeta nhat))
     (niter 3)      ; number of iterations for calibration
     (p     2)      ; number of processors (on a spacecraft)
     (cflop         ; floating point operations required
      (* nd (+
     (* 3.0 (+ (* 13.0 m m n) (* 50.0)))
     (* 100.0 npsi))))
     (tp            ; time interval between downloads
      (* 100 (* 24.0 3600)))
     (rdl   1.0e+6) ; downlink data rate
     (xi    3)      ; fudge factor for peak memory usage
     (bytes 4)      ; bytes per word

     (primary-storage-high (* xi nd zeta zeta npsi npsi bytes))
     (primary-storage-low  (* xi nd m zeta npsi bytes))
     (secondary-storage    (* nd nhat bytes (/ tp td)))
     (data-in              (/ (* nd npsi bytes) td))
     (data-out             (/ (* nd nhat bytes) td))

     (ipc           ; interprocessor communication, dom. by SVD (fr. scalapack)
      (/ (* nd
      (+ npsi
         (* 14.0 (sqrt p) (log p 2) niter m zeta npsi))
      bytes)
   td)))
  (defun printer (x y units) (format t  "~a: ~8,2,1,1,,,E ~a~%" x y units))
  (defmacro printit (x units) `(printer ',x ,x ,units))
  (defun hline () (format t "~%"))
  (hline)
  (format t "PDF Model.~%")
  (hline)
  (printit TD "")
  (printit RT "")
  (printit RX "")
  (printit ND "")
```

```
  (printit NSC "")
  (printit M "")
  (printit NPSI "")
  (printit NHAT "")
  (printit ZETA "")
  (printit N "")
  (printit NITER "")
  (printit P "")
  (printit CFLOP "")
  (printit TP "")
  (printit RDL "")
  (printit XI "")
  (printit BYTES "")
  (hline)
  (printit primary-storage-low 'bytes)
  (printit primary-storage-high 'bytes)
  (hline)
  (printit secondary-storage 'bytes)
  (hline)
  (printit data-in 'bytes-per-second)
  (printit data-out 'bytes-per-second)
  (hline)
  (let ((n-procs p))
    (printit n-procs 'pe))
  (printit ipc 'bytes-per-second)
  (hline)
  (format t
  (concatenate
   'string
   "Note  that the particle instruments "
   "will form~%only about half of the Science Application data "
   "burden of a plasma~%sampling probe.~%~%")))

PDF Model.

TD:  3.00e+0
RT:  8.00e+2
RX:  1.00e+2
ND:  6.00e+0
NSC:  6.40e+1
[GC threshold exceeded with 2,003,024 bytes in use.  Commencing GC.]
[GC completed with 1,029,744 bytes retained and 973,280 bytes freed.]
[GC will next occur when at least 3,029,744 bytes are in use.]
M:  3.00e+1
NPSI:  2.40e+5
NHAT:  3.00e+1
ZETA:  1.00e-1
N:  3.00e+0
NITER:  3.00e+0
P:  1.00e+0
CFLOP:  1.45e+8
TP:  8.64e+6
RDL:  1.00e+6
XI:  3.00e+0
BYTES:  4.00e+0

PRIMARY-STORAGE-LOW:  5.18e+7 BYTES
```

PRIMARY-STORAGE-HIGH: 4.15e+10 BYTES

SECONDARY-STORAGE: 2.07e+9 BYTES

DATA-IN: 1.92e+6 BYTES-PER-SECOND
DATA-OUT: 2.40e+2 BYTES-PER-SECOND

N-PROCS: 1.00e+0 PE
IPC: 1.92e+6 BYTES-PER-SECOND

Note that the particle instruments will form only about half of the Science Application data burden of a plasma sampling probe.

PDF Model.

TD: 3.00e+0
RT: 8.00e+2
RX: 1.00e+2
ND: 6.00e+0
NSC: 6.40e+1
M: 3.00e+1
NPSI: 2.40e+5
NHAT: 3.00e+1
ZETA: 1.00e-1
N: 3.00e+0
NITER: 3.00e+0
P: 2.00e+0
CFLOP: 1.45e+8
TP: 8.64e+6
RDL: 1.00e+6
XI: 3.00e+0
BYTES: 4.00e+0

PRIMARY-STORAGE-LOW: 5.18e+7 BYTES
PRIMARY-STORAGE-HIGH: 4.15e+10 BYTES

SECONDARY-STORAGE: 2.07e+9 BYTES

DATA-IN: 1.92e+6 BYTES-PER-SECOND
DATA-OUT: 2.40e+2 BYTES-PER-SECOND

N-PROCS: 2.00e+0 PE
IPC: 3.44e+8 BYTES-PER-SECOND

Note that the particle instruments will form only about half of the Science Application data burden of a plasma sampling probe.

3.2. SIRA Model

The following code applies to section 2.

```
;;; A model for the Solar Imaging Radio Array

;;; A model for the Solar Imaging Radio Array

(let*
    ((td      1)      ; measurement cadence
     (nd      2)      ; number of detectors--dipole antennas here.
     (nsc     64)     ; number of spacecraft in the array (16-128)
     (nbase           ; number of baselines between array s/c
      (* 0.5 nsc (1- nsc)))
     (ncorr           ; number of correlations to cross-multiply
      (* nd nd nbase))
     (nphi    5)     ; number of sky positions to work on
     (rt      4000)   ; rate of data input
     (npsi            ; number of bytes read in: (* Rt td)
      (* rt td))
     (nhat            ; amount of reduced data
      (* ncorr nphi))
     (bytes 4)        ; number of bytes per word (or measurement)
     (cfft-1          ; flop count needed by fourier transform of raw data
      (* 40.0 npsi (log npsi 2)))
     (cfft-2          ; flop count needed for cross-correlation analyses
      (* 40.0 512 (log 512 2)))
     (cpu-speed       ; across the entire array of sattelites
      (/ (+ (* nsc nd cfft-1) (* ncorr cfft-2)) td))
     (primary-storage ; main memory of processing elements
      (*
       (+
         (* 2.0 nsc nd npsi) ; raw data
         (* ncorr 2.0 512)   ; perform this many cross-correlations
         (* ncorr nphi))    ; keep around this many skypoints per correlation
       bytes))
     (secondary-storage ; mass-storage
      (/ (* 8.64e+4 nhat bytes) td))
     (data-in         ; sensor -> pe
      (/ (* nd nsc npsi bytes) td))
     (data-out        ; pe -> mass-store
      (/ (* 0.5 nphi nd nd nsc bytes (1- nsc)) td))
     (ipc-corr-commo  ; broadcast reduced data to PEs
      (* 512 nd nsc bytes (/ 1.0 td))))
  (defun printer (x y units) (format t  "~a: ~8,2,1,1,,,E ~a~%" x y units))
  (defmacro printit (x units) `(printer ',x ,x ,units))
  (defun hline () (format t "~%"))
  (hline)
  (format t "SIRA Model.~%")
  (hline)
  ;;
  (printit TD "")
  (printit ND "")
  (printit NSC "")
  (printit NBASE "")
  (printit NCORR "")
  (printit NPHI "")
```

```
  (printit NPSI "")
  (printit NHAT "")
  (printit CFFT-1 "")
  (printit CFFT-2 "")
  ;;
  (hline)
  (printit cpu-speed 'flop-per-second)
  (hline)
  (printit primary-storage 'bytes)
  (printit secondary-storage 'bytes-per-day)
  (hline)
  (printit data-in 'bytes-per-second)
  (printit data-out 'bytes-per-second)
  (hline)
  (printit ipc-corr-commo 'bytes-per-second))

;;; Sample output

SIRA Model.

TD:  1.00e+0
ND:  2.00e+0
NSC:  6.40e+1
NBASE:  2.02e+3
NCORR:  8.06e+3
NPHI:  5.00e+0
NPSI:  4.00e+3
NHAT:  4.03e+4
CFFT-1:  1.91e+6
CFFT-2:  1.84e+5

CPU-SPEED:  1.73e+9 FLOP-PER-SECOND

PRIMARY-STORAGE:  3.73e+7 BYTES
SECONDARY-STORAGE: 1.39e+10 BYTES-PER-DAY

DATA-IN:  2.05e+6 BYTES-PER-SECOND
DATA-OUT:  1.61e+5 BYTES-PER-SECOND

IPC-CORR-COMMO:  2.62e+5 BYTES-PER-SECOND
```