GENERAL ELECTRIC
SPACE DIVISION
PHILADELPHIA

PROGRAM INFORMATION REQUEST / RELEASE

PIR NO.	*CLASS. LTR.	OPERATION	PROGRAM	SEQUENCE NO.	REV. LTR.
	U	1G10	LSD	DMS GEN 005	

*USE "C" FOR CLASSIFIED AND "U" FOR UNCLASSIFIED

FROM Jon Bondy		TO Daryl Lotz	
DATE SENT 8/30/79	DATE INFO. REQUIRED	PROJECT AND REQ. NO.	REFERENCE DIR. NO.

SUBJECT

IGF Disk Throughput Analysis

INFORMATION REQUESTED/RELEASED

Abstract

A study has been started to identify workable Reformatting and Annotation (RAA) Disks for the LANDSAT-D IGF. The baseline devices for this function have changed twice over the past year, but never have included standard DEC products with clearly workable technical solutions. This report attempts to delineate off-the-shelf DEC solutions for the IGF disk functions. The design problem is sufficiently complex that optimization will not be possible within the time constraints of the study.

Conclusions

Neither RP06 nor RM03 drives provide elegant solutions to the IGF requirements, but an RP06 solution appears to be feasible. It can meet both IPC and PGP requirements, but is physically large; its physical size is roughly that of an RM05 solution.

Acknowledgement

These results could not have been produced without the aid of John McBeth, Phil Miller, Bob Novas, Steve Deller, Charlie Gregan, and especially Dick Kaiser.

Task Description

The selected configurations will be evaluated in terms of performance/throughput, cost, schedule impact, software impact, facility impact, VAX buffering requirements, and impact on the reliability/availability of the resulting configuration. Additional considerations will be the minimization of the number of Mass Buss Adaptors (MBA's) required and the flexibility/expandibility of each solution. Included in the description of each configuration studied will be the data formats on disk, an IPC utilization scenario, a PGP utilization scenario, and a description of the configuration.

J. Avery   J. Bondy
E. Beyer   +Lanham Distribution
T. Aepli

PAGE NO.	✓ RETENTION REQUIREMENTS		
	COPIES FOR	MASTERS FOR	
1 OF 41	□ 1 MO.	□ 3 MOS.	
	□ 3 MOS.	□ 6 MOS.	
	□ 6 MOS.	□ 12 MOS.	
	□ MOS.	□ MOS.	
	□	□ DO NOT DESTROY	

A large number of issues combine to make this problem complex. They are each discussed below, after which alternatative solutions are identified and then described.

Problem Description

The RAA disks will serve as image data buffers in the TM IPC, the TM PGP, the MSS RDCP, and the TM HDTA to CCT processes. In TM IPC, data in Band Interleaved by Line format (BIL) will be read from A-Tapes, stored on disk in some intermediate format, and read out in Band Sequential (BSQ) format. In TM PGP, data will be read from P-Tapes in BSQ format, stored on disk in BSQ, and read off of disk in BSQ. MSS RDCP and TM HDTA to CCT are believed to be performance sub-sets of the IPC and PGP processes, and are not analyzed here in detail. (It should be noted that the TM HDTA to CCT process also requires a BIL to BSQ reformatting operation, although at a lower data rate than the TM IPC process requires.) The reformatting required in the IPC case causes the major difficulty, while in the PGP case, the required data rate is the primary challenge. Due to the fact that the processing strings act as backup for each other (see the section on reliability below), similar disk configurations must perform both processes.

Image Data Mapping Routine

The layout of image data on disk, as well as the blocking factors, may change if the disk media or the data formats are modified during the development of the IGF. In order to minimize the impact of any such changes, a single routine should be used to map blocks of image data for a given band to physical locations on an MBA and a disk. In addition to this mapping function, faulty cylinders may be bypassed by storing a 'good cylinder' map on each disk and loading it into the VAX at application startup time. Because a double-cylinder seek takes roughly the same time as a single cylinder seek, single bad cylinders on a pack should not prevent its use. It should be noted that bursts of bad cylinders could cause long seeks between 'adjacent' logical cylinders, and could endanger the application's timeline; a limitation on the number of sequential bad cylinders should be imposed in the program which creates the 'good cylinder' map.

Another function which the mapping routine could provide is to help synchronize the disk read and write processes. If the routine is informed as to whether a mapping request is for a read or a write, it could determine the 'phasing' between reads and writes which were currently being performed. If the read process were advancing too rapidly (perhaps approaching a point where interference with the write process was likely), the mapping procedure could refuse to return control to the read process until sufficient writes were accomplished that the interference was no longer a problem. Implimentation of the routine in this manner could turn out to be non-trivial, as it will have to be re-entrant and to communicate among its current activations.

In any event, if the routine were to not only map the blocks onto disk, but issue the QIO's itself, it could serve as a simple and transparent interface between any user process and the RAA disks.

NSCI/NSCO Startins/Stoppins

There are two scenarios for the three major processes of the IGF, one of which requires that the NSCI and NSCO data rates for all processes be locked tosether (to an unspecified desree), and the other which requires that either input or output tapes be halted temporarily whenever the processins deviates from synchronization. The former implies tishtly coupled R-, A-, and P-Tape formats, while the latter is more flexible in this resard, at the expense of additional loop control software.

For IPC, one could desisn the process with an output rate which was faster than that necessary to keep up with the input. This could be done simply to be conservative, or to allow the P-Tape to be stopped at the end of a scene for an arbitrary period to allow the inputs and outputs to re-synchronize. In editins mode (see below), one could require that either tape be stopped whenever it was waitins for the other process. This would simplify the bufferins stratesies, which would otherwise require that disk read/write phasins be maintained via buffer allocation stratesies and P-Tape preamble-fill. In the past, there was some hesitancy to manipulate the tapes in this fashion, but that appears to be less of a concern now (reference conversations with Dick Kaiser, Bob Novas, and Bob Farrell).

IPC -- General

The A-Tape contains data other than pixel data, much of which must be carried around durins the processins of the imasery. When all of the extra data is taken into account, 6300 bytes are required by each line of imase data (reference 14RO-LSD-DMS-MEMO-040 by John McBeth). The BIL scene size is 6000 lines, so the amount of data for a sinsle band is

6000 lines/band * 6300 bytes/line

or 37.8 Mesa-Bytes (MB). Since there are seven bands per scene, the total storase required for a BIL scene is 264.6 MB. Note that althoush the thermal band (band 6) contains 1/16 of the data of the other bands when downlinked from the spacecraft, it is replicated durins the RDCP process, and is as larse as the other bands on the A-Tape.

The data on the A-Tape is interval data, that is swath data which has not yet been framed to produce scenes. The framins process, which is accomplished durins IPC, causes approximately 5 percent of the data from one scene to also be used in the followins scene. This behaviour is termed 'overlappins'. When overlappins is combined with the band and scene sizes defined above, the BSQ sizes become 39.7 MB/band and 277.8 MB/scene.

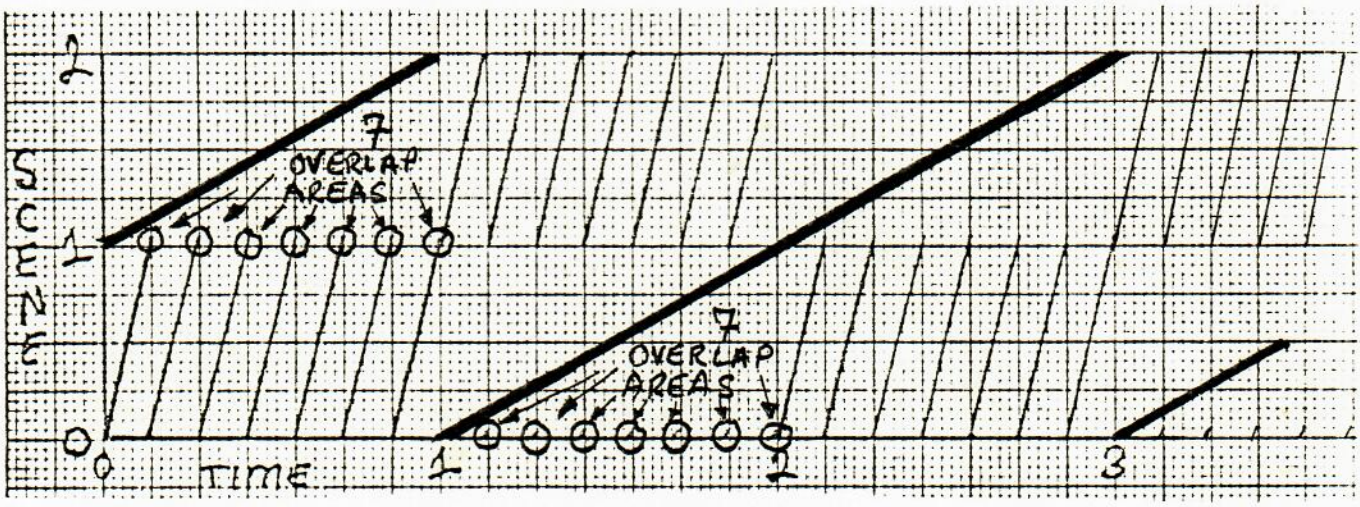
The IPC process will run at 1/16 real time nominally, or 424 seconds/scene. This implies a disk write data rate of 264.6 MB/424 seconds, or 624.1 KB/second averase throushput for the BIL case. Similarly, the BSQ rate is 277.8 MB/424 seconds, or 655.2 KB/second. These performance numbers are the desisn limit for IPC, with 20 percent additional throushput to be added as safety marsin. (If one were to be as conservative as one would like, the problem could become unsolvable.)

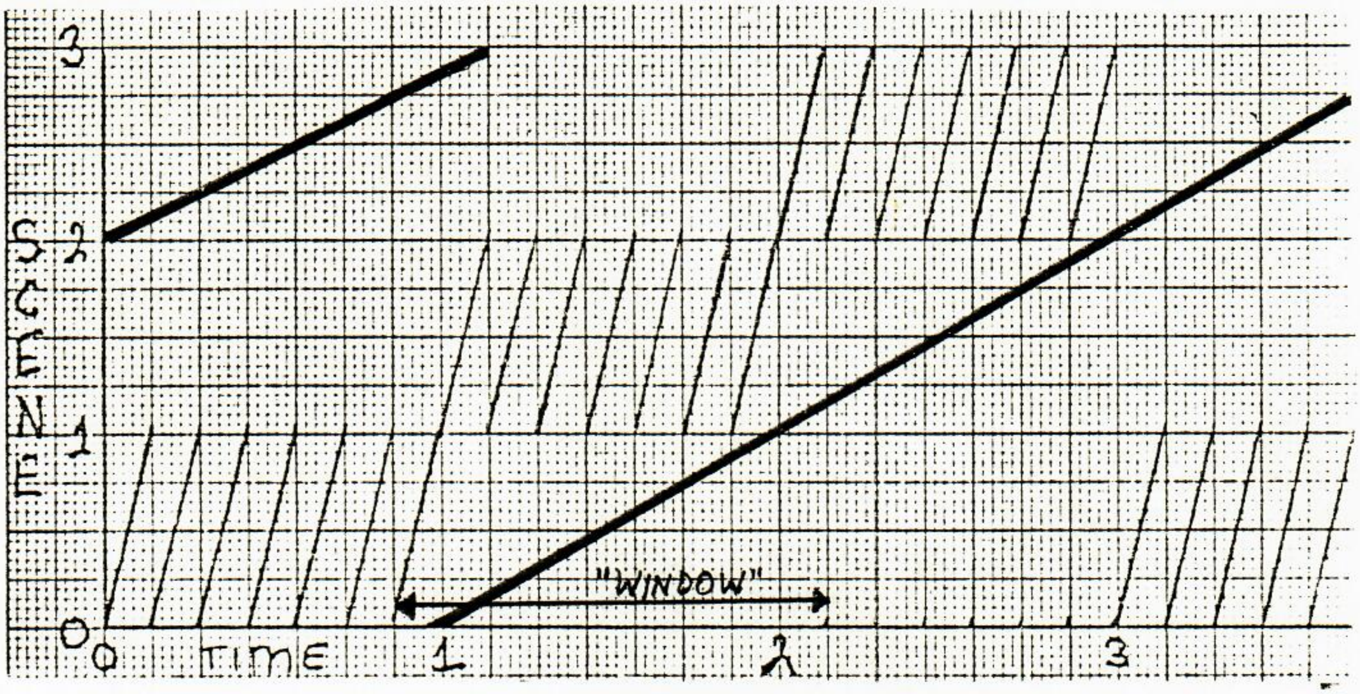
The bottom line is that the conservative rates which we require are 750 KB/seconds for BIL and 786 KB/second for BSQ.

IPC -- Read/Write Interference

Most IPC disk data layouts proposed to date take the BIL format of the A-Tape and bring it closer to the BSQ format of the P-Tape by creating an intermediate 'band interleaved by block' format. When the blocks are actually disk cylinders, the Band Interleaved by Cylinder (BIC) format results. These formats require that all of the mass storage area on which a scene lies be scanned in both the write and read modes, with the read mode reading roughly seven times faster (by skipping 6/7 of the data).

One could represent this graphically, with scene data storage areas on one axis and time on another, as in the following diagrams. Notice the five percent overlap areas in each diagram.

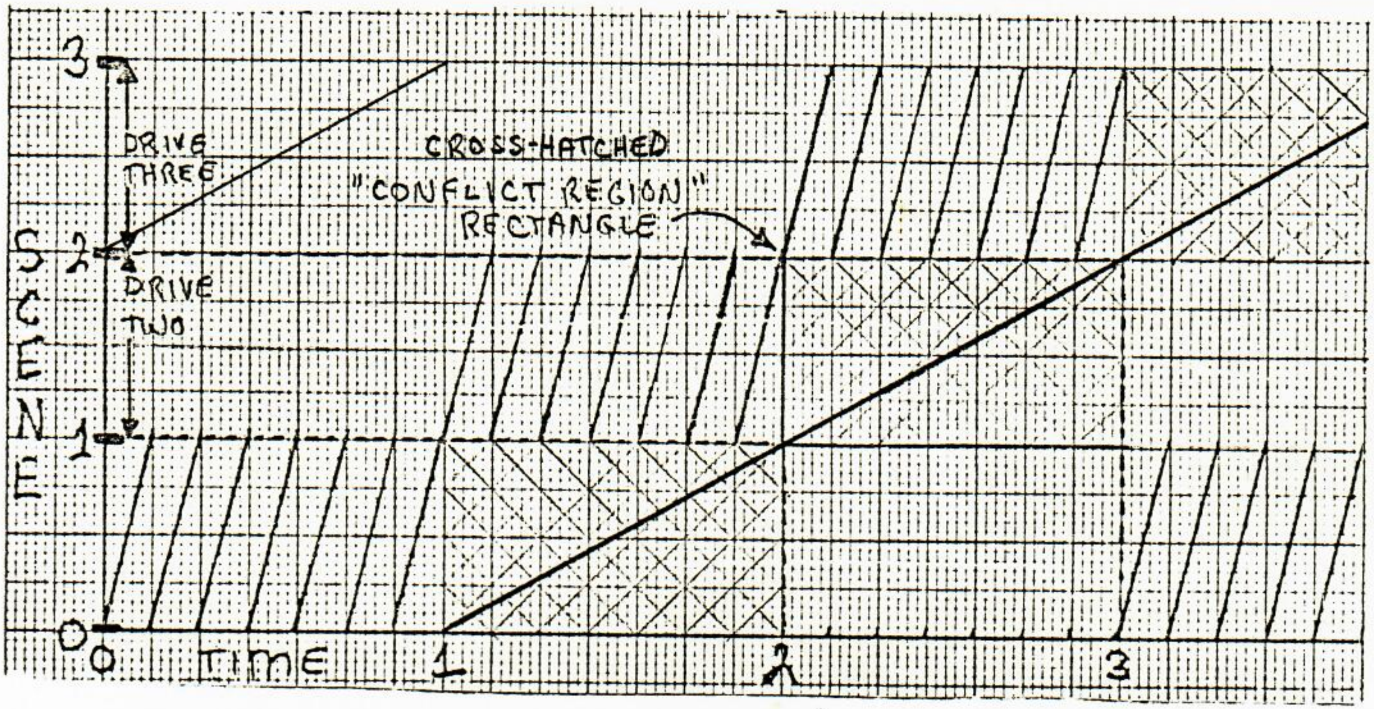




Here one can see the mass storage access pattern for disk buffers with both two and three scenes per buffer. The bursts of seven short-slanted lines represent the access of each of the seven bands for reading, and the one long thick line represents the write access. Since the storage pattern repeats when the buffer is filled (after two scenes for the first diagram, and three for the second), the diagrams need only show three or so scenes for the two-scene buffer case, and five or so for for the three-scene buffer case.

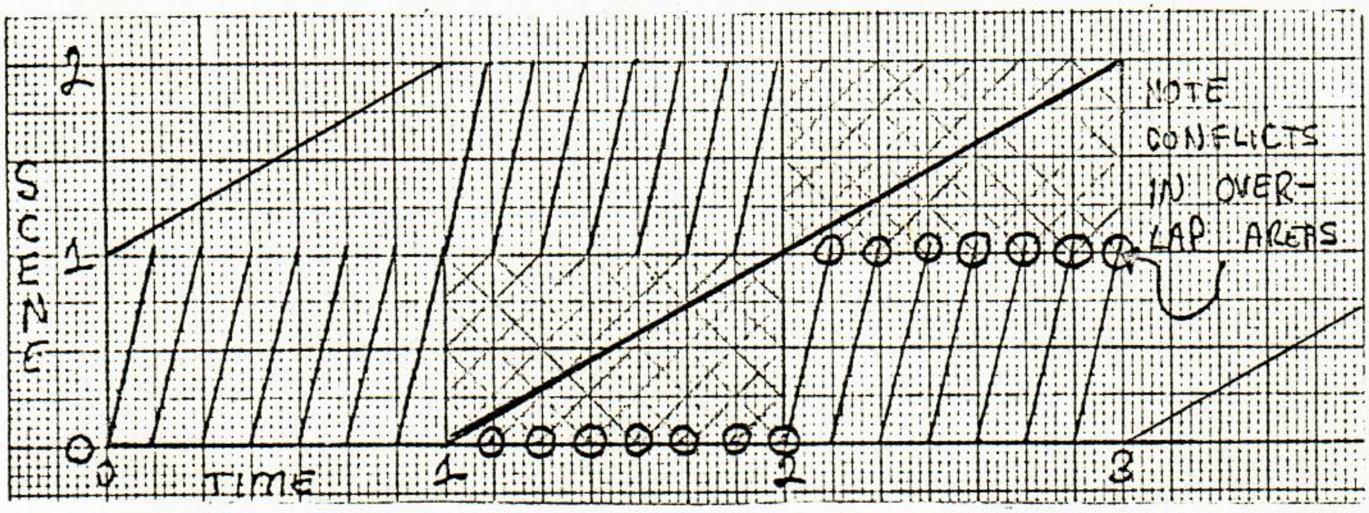
If one is using a disk medium which is being run at close to its rated throughput rate, one cannot tolerate additional seek times other than those accounted for in either the read or the write scenarios. This implies that a given drive cannot be used both for reading and writing data at the same time, which in turn implies that there is a fixed 'window' within which writing can take place. (In the PGP process, the BIL to BSQ reformatting is not performed, and this is not as significant a problem.) As an example, this window is indicated on the second diagram above. When one choses a specific write 'path' within the window, a buffer utilization scenario has been defined, with a read/write phasing delay which characterizes that approach.

Once a specific writing pattern has been chosen, one must allocate scene data areas to disk drives in order to determine if a given drive will be used for both reading and writing at any time during the process. As an illustration, consider a three scene buffer divided among three devices, as given below :



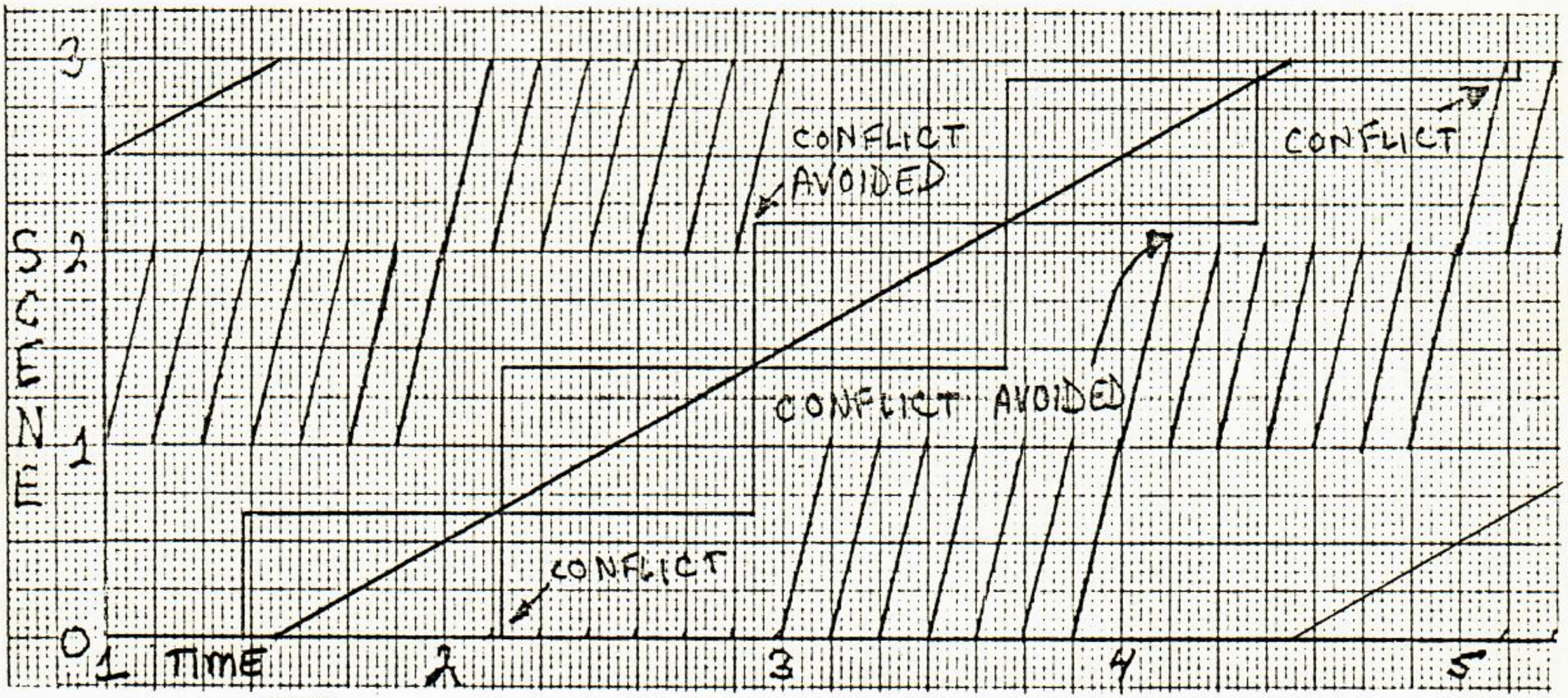
The division of the disk buffer area into drives is shown along the vertical axis. One can project these vertical line segments to the right until they hit the write line, at which point they define time periods when the drives are in use for writing. These time periods can be projected down onto the time axis, or 'conflict regions' can be created as rectangles surrounding the write lines, illustrated as cross-hatched areas above. If a conflict region contains both reading and writing, then there will be a time when a drive will be in use for both purposes, causing excessive seeking and system failure. Notice that as the number of drives increases, the area of the conflict regions decreases, and they tend to lie closer along the write line. For this reason, increasing the number of media can decrease the system's sensitivity to NSCI/NSCO timing mismatches.

The allocation of disk media is complicated by the overlapping of data during reads, as can be seen in the following diagram :



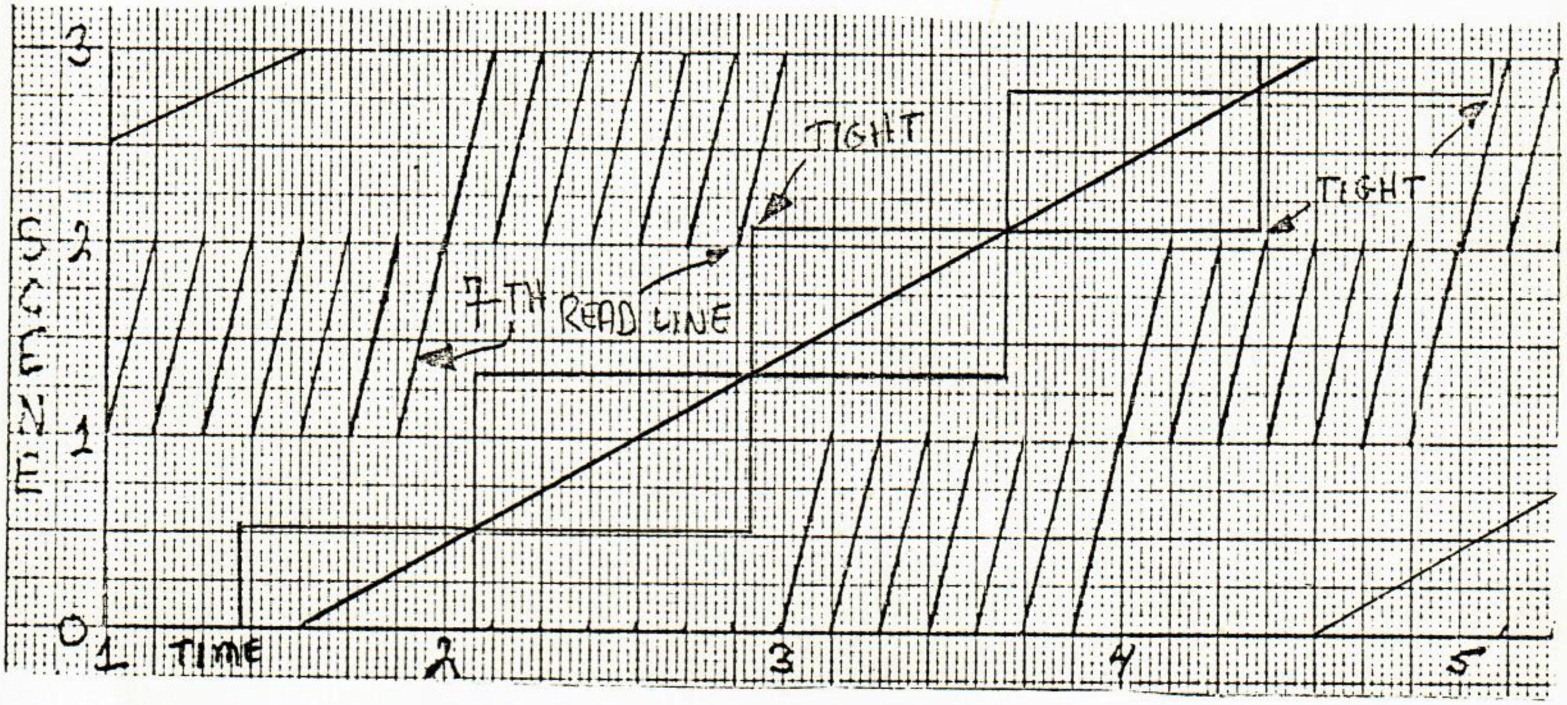
In this case, one can see that it is impossible to allocate the buffer area into two drives without creating a read/write conflict.

As a final illustration, an attempt was made to allocate three scenes across three MBA's with four RM03 drives per MBA, with the data equally distributed. This turned out to be impossible, due to the overlap areas, as illustrated below :

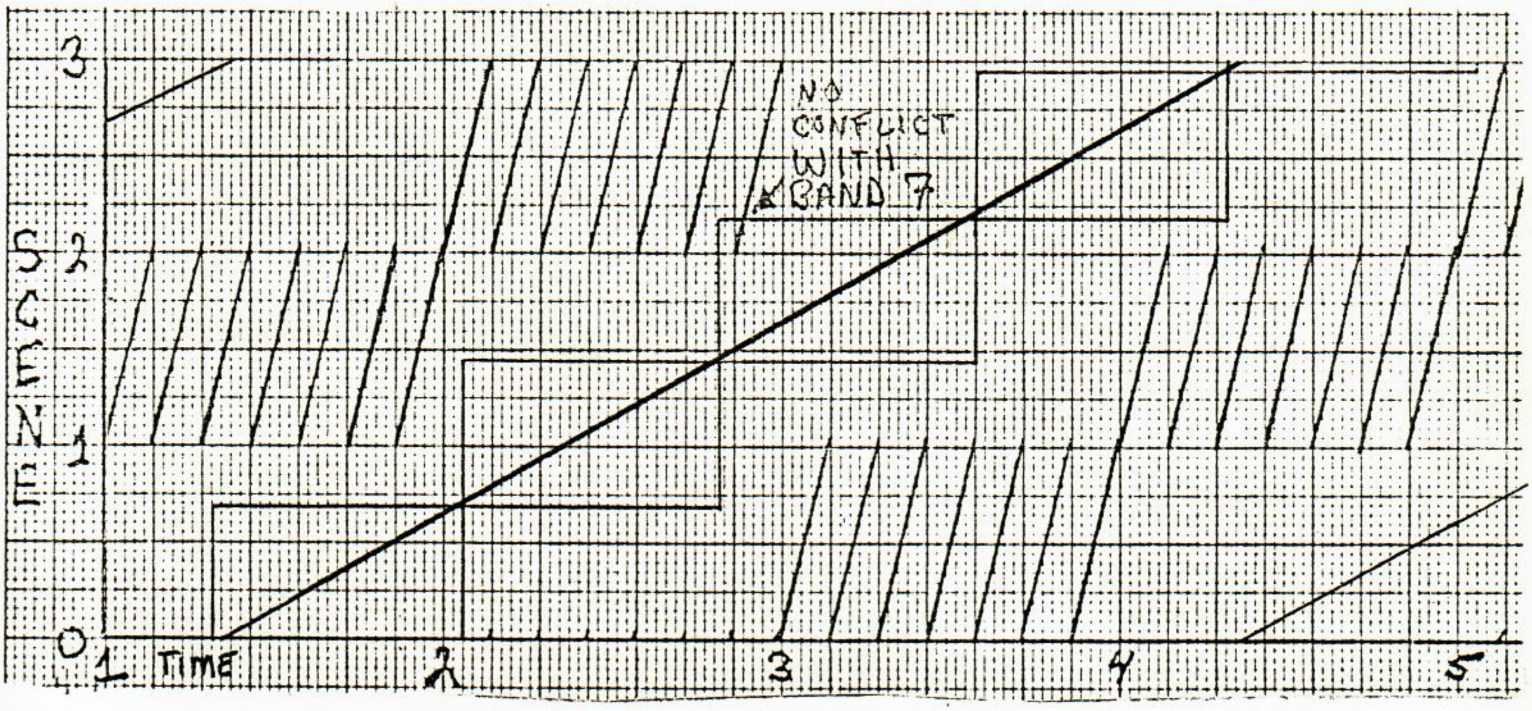


Proving that there is no solution requires 'playing' with the positions of the write line and the phasing of the disk media with respect to the start of scene one for a while. (Note that moving the conflict regions to the left to avoid the conflict near scene time 2 1/7 will force a conflict at time 2 6/7; that starting the write later will move the regions down, forcing a conflict at times 4 1/7 through 4 6/7; and that starting the write earlier will move the regions up, forcing a conflict at time 2 6/7.)

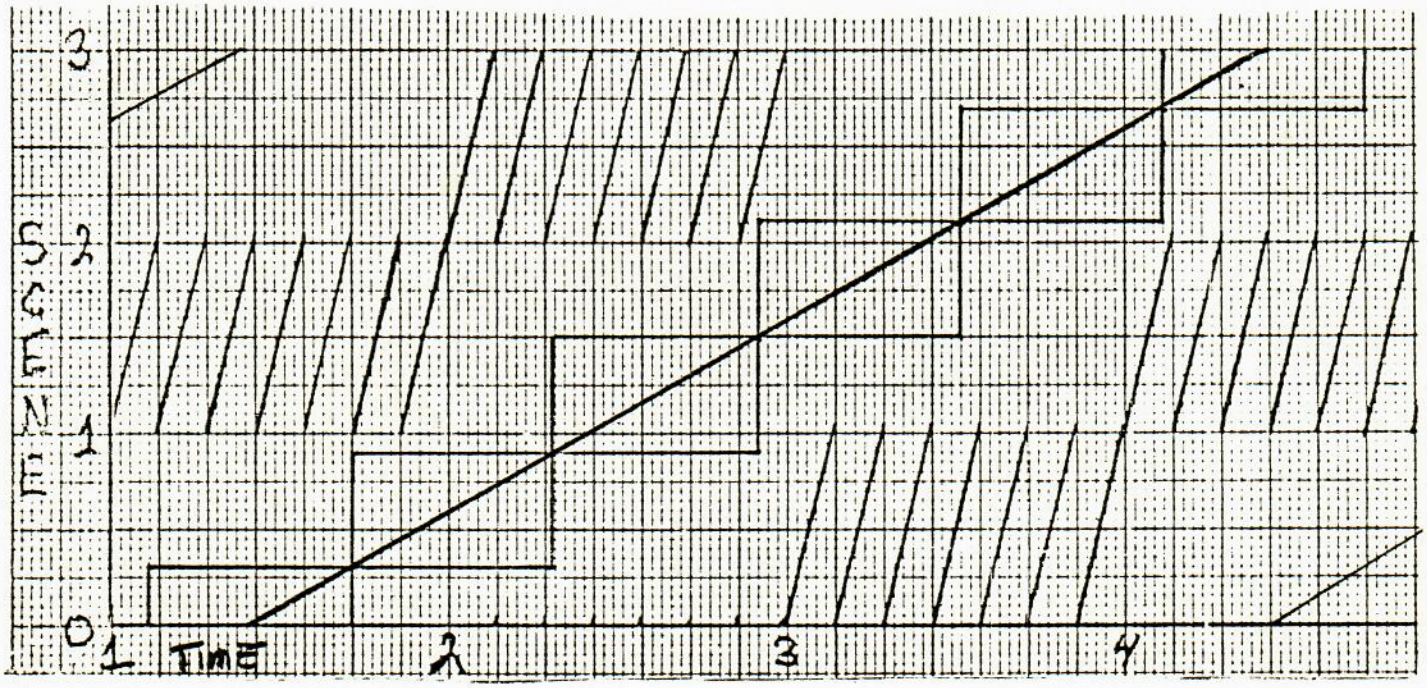
When the requirement that the data be divided up evenly was dropped, the following solution was discovered :



Note that the drive conflict areas are very close to the read lines in a number of places, indicating that minor NSCI/NSCO timing mismatches could cause conflicts, and process failure. If the thermal bands were de-replicated, and placed on a separate drive and MBA, then the seventh read line would disappear (although the system times would not change, (since the thermal data would have to be sent to the FFP in any case) and read/write conflict would exist internal to the separate drive during thermal band reading). This would reduce the conflict areas, allowing the following solution to obtain :



To confirm that a five drive solution exists, the followins diasram was derived. It pertains to both RM03 and RM05 solutions.



It should be clear from the above that the structurins of the disk buffers, the allocation of drives, and the phasins of readins and writins are complex issues, dependins on a variety of factors includins disk data formats, MBA utilization, and drive capacity. Inspection of two- and three-scene buffer diagrams, such as those above, indicate that two-scene buffers will not work at all, and that three-scene buffers will work easily only for configurations where five or more drives are available.

It should be noted that all of the above considers a buffer with three scenes allocated in a fixed fashion. Using the same buffer storage as a circular buffer has not been considered due to the complexity of insuring that read/write interference will not occur when using such a flexible buffer strategy.

IPC -- Editing

Scene editing consists of processing to P-Tape only a sub-set of the un-framed scenes on an A-Tape, and skipping over the other A-tape data at high speed. Suppose that we wish to geometrically correct scenes 1 and 4 of an interval. Normally, the overlapped data area between scenes one and two is obtained from the start of scene two without any special considerations. During editing, that extra data must be read as a special case, using part of the buffer area for the second scene. In order to maintain a fixed-length and -position buffering scheme, the start of scene four must be located in the third buffer, providing further justification for the choice of three-scene disk buffers. It should be noted that the system control timing of the editing scenario is complex and has not been dealt with here in detail.

IPC -- Annotation Data

Due to the triplication of annotation data major frames on the A-Tape, annotation data will come off of the A-Tape at one third the rate of image data, and so must be sent to the annotation disk(s) at approximately 210 KB/second. It is read off of the disks at the start of geometric correction of each scene, and sent to the FFP, but this is an infrequent operation which could be handled at a low data rate if some VAX buffer room were dedicated to this operation. For the purposes of the study, the annotation data will be treated as a throughput problem when coming from A-Tape, but not as such when being transmitted to the FFP.

In the past, the annotation data was placed on a separate RAA disk, since the use of the image data disks for this purpose could cause interference while reading the image data. The fact that the annotation data rate is somewhat low is not sufficient to permit the use of the image disks, since the seeking activity which would result would reduce the image data read rate below acceptable levels. It appears that a separate disk drive is required for the annotation data, although it could be placed on an image disk MBA if it were used at a lower priority than the image data. In both the RM03 and RP06 solutions discussed below, it is assumed that the annotation data is written to a separate drive on an MBA which is also used for image data, and that annotation data is both read and written at a low priority.

PGP

The PGP scenario also runs at 1/16 real time, but the Laser Beam Recorder's (LBR's) data format causes some data bursting which makes these data rates higher than during IPC. An additional contributor to the higher data rates in PGP is that the line lengths have increased to 7100 bytes (reference update to 14R0-LSD-DMS-MEMO-110 by John McBeth). The LBR normally runs at a fixed number of output lines per minute, whether image data is being created or not. This, combined with the approximately 24 percent of the LBR output which is not image data, means that the data rate during image creation is proportionally higher. Even using fast forwarding during the inter-image spaces, the data rate during image generation is 852.2 KB/second (same reference).

After adding a 20 percent safety factor, we get 1022.6 KB/second. It should be noted that only two scenes are required for the PGP process, making it possible to use fewer drives on string three.

Schedule

There are two schedule milestones which involve the disks, those being the start of IPC and PGP software integration, and the Operational Readiness Test (ORT) at Goddard. It is conceivable that a temporary disk configuration could be designed which would get the project through software integration with degraded throughput, but would then be replaced by a final configuration just prior to ORT; the software impact of this approach could be significant if not planned for carefully. Initial software integration is scheduled to start as early as 1/80, although it is not clear that the software associated with the RAA disks will be integrated as early as that; a temporary disk solution would have to be be available in the 1/80 to 3/80 time frame. Since ORT is scheduled for fall of 80, the final disk configuration should be in place by 7/80 at the latest, in order to allow final software integration.

Current DEC delivery schedules indicate seven month delays in delivery for RM03 and RP06 drives, implying that drives ordered in 9/79 would be delivered in 4/80. This implies that temporary solutions will fail due to schedule unless some delivery relief is given by DEC.

Mass Bus Adaptors (MBA's)

The number of MBA's available for accessing the RAA disks in the baseline configuration is two. It appears that there is a physical limitation on the VAX which permits only a total of eight MBA and UBA interfaces. In any event, no one with whom I was able to speak advocated the addition of more MBA's, even if there were a way to connect them. Since a dedicated MBA is required for each of the NSCI, NSCO, and FFP, and a single UBA is required for the miscellaneous peripherals on the system, only four MBA's remain to be allocated to the disks. In order to insure that debugging activities cause no interference with the application, the system disk(s) should have their own dedicated MBA. This leaves at most three MBA's for use with the RAA disks, and the third MBA would only be available by 'stealing' it from the tape drive. It is expected that the tape drive, which will not be in use during the IPC process, could be made to share an MBA with the NSCO. The conclusion which is reached here is that although two MBA's will be available, and three could be obtained with some minor difficulty, four will NOT be available. This conclusion is in contrast to some previous scenarios which were more optimistic about MBA availability.

QIO's

QIO's enter into this analysis in two manners, as CPU overhead induced by the QIO calls, and as the latency (or turn-around time) between I/O completions and the initiation of the next I/O in the queues.

I/O for the VAX is somewhat more complex than for computers where virtual paging does not occur. Specifically, each time an I/O is requested, the system must check and insure that all pages in the I/O buffers are in physical memory (not out on disk), and that they are locked in memory so that they are not paged out during I/O. Although it is anticipated that the application tasks will be locked in memory during processing, the normal VMS procedure would be to check for page lock during each I/O, on a per page basis. This implies that large I/O's would require some more setup time than short I/O's (although the setup time per byte would decrease with larger I/O's). After I/O, the normal procedure would be to unlock the pages which were locked for the I/O.

In general, I/O requests require three time periods in order to complete, those of queuing, turn-around, and de-queuing. Queuing involves the page checking described above, as well as the insertion of the request in the appropriate queue (one per MBA/UBA) according to task priority. Turn-around involves the accession of completion codes from the MBA/UBA for the previous I/O and the setup (including loading of mapping registers) and initiation for the next I/O. De-queuing involves the (possible) unlocking of I/O locked pages and the returning of completion codes to the calling process.

Total time for these three phases is given by
    1.2 + (0.050 * pages) milliseconds
where 'pages' is the number of pages to be transmitted by the I/O (reference 14R0-LSD-GS-MEMO-067 Rev B by Phil Miller). The 50 microsecond page time is primarily taken up in locking and unlocking pages, although there is some time allocated to setting up the mapping registers in the MBA's and UBA's. The turn-around time, which is a component of the above, is given by
    300 + (3 * pages) microseconds
where 'pages' is as above (same reference). The per page time here is purely MBA mapping register time.

It is possible to 'jam' I/O operations, by doing all of the queuing operations before the I/O complete is encountered by the system, and performing all of the de-queuing operations after the next I/O has been initiated. This would allow the time between I/O completions and the next initiation to be simply the turn-around time, not the total QIO time.

In addition, it should be possible to modify the standard system drivers in order to reduce the page locking/unlocking time during the IPC/PGP applications, as the entire application will be locked in memory in any case. This may turn out to be important, as the total CPU time dedicated to QIO processing in the simplest IPC case is about 30 percent (see Appendix I). This loading reduces to around 18 percent when all page locking/unlocking is removed.

The upshot of all of this is that small QIO's require about 300 microseconds of turn-around time, while large QIO's require about 1000 microseconds. These times reflect the best case time if no other I/O processing is pending when the I/O complete interrupt occurs. During IPC, the highest priority I/O operations will be on the NSCI and NSCO, since they cannot recover from data overruns. (A parallel situation occurs in PGP, with the LBR replacing the NSCO.) If two MBA's are dedicated to the image data, there could be as many as three I/O's to turn around before a given RAA disk I/O is processed. This problem may be significant, since NSCI I/O initiate rates will be on the order of 100/second (one per line). Because of this, we will be using 2 milliseconds as the turn-around design time, to take I/O queuing effects into consideration. (It should be noted that at 3 milliseconds, we start having some disk format problems and performance degradation.) This time will be taken into account when laying data out on disk, so that disk rotations are not missed due to turn-around delays.

QIO Chaining

Since some of the QIO computations refer to entering the requests in the queues themselves, chaining of related QIO requests (i.e., having one QIO entry cause multiple I/O initiations) might save some CPU time. This could help the NSCI, which will be doing scatter writes into different buffers in memory for each line read. It also could help the disks, since optimal disk utilization may involve filling cylinders with data. As the cylinder sizes on the disks under consideration are greater than 64 KB, this would require more than one I/O initiation per cylinder.

Blocking Factors

Clearly as the number of lines of data per disk block (per QIO) decreases, the total number of QIO's increases, potentially choking the CPU with QIO overhead computations. Use of the largest block sizes which optimize storage utilization is to be encouraged.

Use of QIO chaining to fill cylinders sounds good on the basis of maximizing disk throughput, but it can backfire if the requisite VAX memory buffer space increases dramatically. As an example, filling each cylinder on an RP06 with a band implies buffer space for

4 blocks/cylinder * 8 lines/block * 6300 bytes/line * 7 bands

or 1.4 MB of buffer. Since one would need to double buffer the data, this approach would require almost 3 MB of buffer for this I/O path alone, and is clearly impractical.

Disk Data Formats

Since the disk data formatting problem for PGP is much less complex than that of IPC, it will be dealt with later, in the discussions of each individual solution. Although the formats of the IPC data to and from the disks are fixed (BIL and BSQ, respectively), the format of the data on the disk is not. Clearly, as the BSQ data rate is larger, one should attempt to optimize the data layout on disk for that case. However, a 'natural' layout on disk would simplify the application and the debugging process. These two criteria tend to trade off against each other.

The limitation that VAX I/O block lengths be less than 64 KB implies that multiple blocks must be written to fill a cylinder. The I/O turn-around times given above indicate that blocks which are to be read or written consecutively should not be adjacent on the disk, but should be separated by enough sectors to allow for this latency without missing a disk revolution. In addition, since seeks require significant time to complete, inter-cylinder data layouts should be skewed by some amount, again to avoid dropping a revolution. Inter-cylinder write patterns differ from inter-cylinder read patterns in IPC, and the data layout will necessarily trade off read rates versus write rates. Since the exact data layout on disk is subject not only to the above considerations, but also to the physical properties of the drives, each drive must be analyzed separately in order to optimize throughput. Appendix II gives some tables which characterize each drive according to its basic properties, and Appendices III and IV give data packing densities for each drive/blocking-factor/latency combination, for IPC and for PGP, respectively.

Thermal De-replication

The thermal band actually contains only 1/16 of the data of the other bands, but is 'replicated' up to the data quantities of the other bands during the RDCP process. It must be de-replicated on its way to the FFP in order that correct resampling take place in the GCO, so de-replicating it earlier (i.e., on its way to the disks, instead of from the disks) would do no harm. In fact, the reduction of scene size from 264.6 MB to 236.3 MB would help greatly in the RM03 case (the de-replication in this case is by lines, so the reduction is from 7 bands to 6.25 'bands'). Using an optimal packing density for RM03's of 92.29 percent of the disk (to be derived later in this report), it takes 4.25 RM03 packs to hold a normal scene, while it takes 3.8 packs to hold a de-replicated scene. This could save as many as three disk drives in the long run, while complicating the format of the data on disk, since band six isn't always there. The use of de-replicated data also reduces the disk write data rate to 557.2 KB/second (668.6 KB/second with 20 percent margin), giving us more margin while running this process. When reading the data, the thermal band's data rate then drops to 163.8 KB/second.

These rates are low enough that storage of the de-replicated thermal band on a separate drive was considered, due to 'nicer' data formats for the remaining bands. It was not retained because contention still exists on the thermal band between reading and writing for IPC, and a single drive was hard pressed to perform adequately during the conflict periods.

Buffer Lengths and Disk Errors

Given the high data rates and constant use which will be expected of the IGF RAA disks, it is not surprising that soft disk errors are likely to occur roughly once per day (unable to re-locate reference). Because of this, it would be good if the final configuration were able to sustain small numbers of re-tries without causing system failure. A number of areas, including disk formatting, can increase our ability to provide this capability, but the primary method is that of increased disk buffers in VAX memory. If re-tries are incorporated into the system design, system performance changes due to re-tries should be monitored, and therefore the re-tries should be logged in some safe (i.e., non-interfering) fashion.

MBA Utilization Approaches

Previous thinking used one MBA to write to disks, and the other to read from them. By using suitably large media, one could get all of the data on less than eight drives (an MBA/controller limitation). By dual-porting all of the drives to both MBA's, one could access any given portion of the data for reading OR for writing, as long as a drive was never accessed for both. The fact that large capacity, high speed media are not available at the time of this writing makes the above approach impractical, since either the data won't fit on eight drives (RM03) or the data rates with 2 MBA's are too low (RP06). In addition, the dual-porting scheme won't work as simply when more than eight drives or more than 2 MBA's are employed.

Variations on the dual-porting technique have been proposed, but they tend to be more difficult to analyze, less flexible, not as easy to expand, and have a complex pattern of inter-drive interference due to bussing layouts which are not general purpose. Some solutions have required rapid switching between the two ports on a dual-ported drive, and an example of such a configuration is included as an alternate RP06 configuration.

Reliability

Availability models for the IGF are just now being generated, and it was not clear how to evaluate the alternative solutions on this criterion. The approach taken for IPC was to specify enough disks on strings one and two so that either could perform the process if the other string were to lose a disk. For PGP, enough disks were specified on string three to perform the process unless a drive failed, at which point it would be performed by whichever string out of one and two was not being used by IPC. Since the PGP process only needs two scenes worth of disk, this approach appears to cause no problems if only single failures are under consideration.

Ping-Ponging

In discussions with Bob Novas and Dick Kaiser, the concept of 'ping-ponging' the MBA's was invented. This method provides that, rather than dedicating a single MBA for disk writes and another for disk reads, both are used, more or less evenly, for both processes. One would write all bands of all scenes to all MBA's. In this way, the overlap constraint on MBA utilization would disappear, since the pattern of MBA utilization while reading disks during overlap would not differ from that during the other phases of reading, except as to which drive was being referenced. Note that this scheme does NOT ameliorate the requirement that drives be dedicated to reading or writing for a given time period.

The scheme would work as follows. The disk write process would accumulate lines of imagery until it was ready to write blocks for all seven bands, at which time it would issue seven QIO's. Those for bands 1, 3, 5, and 7 would be issued to MBA 1, and those for bands 2, 4, and 6 would be issued to MBA 2. The next time, the QIOs for the even bands would be issued to MBA 1, and those for the odd bands to MBA 2. The following diagram illustrates this QIO issue sequence, where 'Bi BLj' means band 'i' and block 'j'.

MBA 1	MBA 2
B1 BL1	B2 BL1
B3 BL1	B4 BL1
B5 BL1	B6 BL1
B7 BL1	B1 BL2
B2 BL2	B3 BL2
B4 BL2	B5 BL2
B6 BL2	B7 BL2
B1 BL3	B2 BL3
etc	etc

Because the bursts of QIO's would be issued quite rapidly, it is likely that all of the writes would be queued sequentially, and inter-QIO latencies would be small enough not to drop a revolution. Reading would proceed by issuing a series of reads on alternating MBA's for data from the same band, with a similar clustering effect taking place, minimizing lost rotational latencies. As the reads are for a different pair of drives than are the writes, we can mix read and write QIO's on each MBA as long as read bursts are not allowed to pre-empt writing to the point that buffer overflow occurs (or visa-versa). It should be noted that if a ping-pong approach is chosen as baseline, that approach should be simulated in order to be certain that the queues which can form at various parts of the system can be comfortably accomodated. Also, note that should a band be stored more heavily on one MBA than on the other(s), the throughput in read mode would become unsatisfactory.

The weakness with this scheme is that, as one switches from one MBA to another, one switches from one drive to another, and the relative positions of the platters on the two drives are unknown, introducing a possible rotational latency. If one can determine that the additional latency is not significant from a throughput standpoint, then the ping-pong technique should work. It is for this reason that BIC format is recommended for the RM03, since it will maximize the amount of time spent transmitting data between seeks.

There are a number of benefits of this technique over dedicated MBA techniques. Addition of a third MBA can increase RAA disk throughput by a factor of 1.5 with little modification to the data layout schemes. The layout of data on the disks, and thus the data access algorithms, are simple, especially so when compared with other, non-symmetric 3-MBA configurations. More drives can be added to the MBA's (up to a limit of eight per MBA), providing some flexibility. Design parameters for this technique are the number of bands per MBA and the number of MBA's.

Identification of Solutions

There are only two 'reasonable' disk drives which are current DEC products, the RM03 and the RP06. There appears to be only one configuration of RM03's which can be considered seriously, while there are two RP06 configurations which bear consideration. A section on a preliminary RM05 solution is provided for comparison of the RM03 and RP06 solutions with a 300 MB disk configuration.

RP06 Solutions

Disk Data Format and Performance -- IPC

Given a design criterion of 2 milliseconds for I/O latency, we find that for the RP06, this corresponds to 2.6 sectors, or 3 sectors. Looking at Appendix III, we find optimal packing density of 94.20 percent with blocks which are eight lines long, and four blocks per cylinder. With this utilization factor, each RP06 holds 164.3 MB, or 0.62 scenes (or 1.61 drives/scene). Mapping the blocks and 'gap' sectors onto the cylinder, we find the following pattern :

	SECTOR																					
TRACK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	g	g	g	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	2	2	2	g	g	g	s	s	s	s	s	s	s	s	s	s	3	3	3	3	3	3
11	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
12	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
13	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
13	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
15	3	3	3	3	3	g	g	g	4	4	4	4	4	4	4	4	4	4	4	4	4	4
16	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
17	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
18	4	4	4	4	4	4	4	4	4	4	4	.4	4	4	4	4	4	4	4	4	4	4
19	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	g	g	g

where '1' through '4' represent the respective data blocks, 'g' represents I/O turn-around gap sectors, and 's' represents seek gap sectors.

With a seek time of 7 milliseconds for a single cylinder seek, the equivalent number of sectors is 9.2, or 10 sectors. Adding the normal I/O latency duration of 3 sectors to this means that we require 13 sectors between data on inter-sector transfers.

We will allocate two bands to each cylinder (two blocks per band), in order to keep the VAX buffer sizes reasonable. Since raw RP06 data rates are too low for either IPC or PGP, we will use a three MBA ping-pong approach, with approximately two bands written to each active disk. The diagram below shows the location of blocks for each band on each MBA and cylinder. Notice that there are four blocks per MBA, and that the band data rotates around the MBA's in order to distribute the read accesses evenly; this pattern repeats every 7 cylinders.

```
                    C Y L I N D E R
                                      1 1 1 1 1 1
                1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 ....

        M       1 7 6 5 4 3 2 1 7 6 5 4 3 2 1 ....
        B       1 7 6 5 4 3 2 1 7 6 5 4 3 2 1 ....
        A       2 1 7 6 5 4 3 2 1 7 6 5 4 3 2 ....
        1       2 1 7 6 5 4 3 2 1 7 6 5 4 3 2 ....

        M       3 2 1 7 6 5 4 3 2 1 7 6 5 4 3 ....
        B       3 2 1 7 6 5 4 3 2 1 7 6 5 4 3 ....
        A       4 3 2 1 7 6 5 4 3 2 1 7 6 5 4 ....
        2       4 3 2 1 7 6 5 4 3 2 1 7 6 5 4 ....

        M       5 4 3 2 1 7 6 5 4 3 2 1 7 6 5 ....
        B       5 4 3 2 1 7 6 5 4 3 2 1 7 6 5 ....
        A       6 5 4 3 2 1 7 6 5 4 3 2 1 7 6 ....
        3       6 5 4 3 2 1 7 6 5 4 3 2 1 7 6 ....
```

Since sequential blocks are spaced far enough apart that rotational misses are not likely to occur, writing data to the disk will cause full cylinder writes. The seek sectors are placed between blocks 2 and 3 for reasons which will be explained in the next paragraph. Because of this, blocks should be written to cylinders in the sequence (3,4,1,2,seek), in order to minimize time wasted during seeks. This probably will result in a cylinder-write-and-seek time of 19 revolutions, but, assuming a miss (worst case), we find that the effective throughput rate for a single MBA will be

(4 blk * 8 li/blk * 6300 by/li) / (20 revs * 0.016666 sec/rev)

or 604.8 KB/second. Using the third MBA one half of the time will cause an effective increase in write throughput to 907.2 KB/second, meeting the requirements.

When reading data, we will be reading two blocks from each cylinder on a particular drive. We will assume that blocks 1 and 2 will be read together, and 3 and 4 will be read together. Looking at the layout of blocks on the MBA's above, we see that we must read blocks 1 and 2, seek a cylinder, read blocks 3 and 4, and then do a long (7 cylinder) seek. It is important that the reading of blocks 3 and 4 follow onto the reading of 1 and 2 without a rotational delay, since dropping a revolution is less likely then than during the long seek. Because of this, the extra sectors in the cylinder have been positioned in between blocks 2 and 3. This results in four blocks read in 19 rotations plus a missed rotation, or 20 rotations total, yielding

(4 blk/cyl * 8 li/blk * 6300 by/li) / (20 revs * 0.01666 sec/rev)

or 604.8 KB/second. Using the third MBA to full advantage results in a read data rate of 907.2 KB/second, meeting the requirements.

Disk Data Format and Performance -- PGP

Again given the 2 millisecond I/O latency, using Appendix IV, we find that storage optimization produces a blocking factor of 7, a utilization factor of 92.89 percent, and four blocks per cylinder. The layout for this data format is similar to that for IPC, and is as follows :

	SECTOR																					
TRACK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	g	g	g	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
9	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
10	2	g	g	g	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
11	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
12	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
13	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
14	3	3	3	3	3	3	3	3	3	3	3	3	3	3	g	g	g	4	4	4	4	4
15	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
16	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
17	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
18	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
19	4	4	4	4	4	g	g	g	s	s	s	s	s	s	s	s	s	s	s	s	s	s

where the symbols are the same as for the IPC case. As there is no reformatting during this process, data can simply be written as sequential blocks of lines, and read in the same fashion. This makes the placement of the seek sectors in the middle of the cylinder unnecessary.

As in the IPC case, the inter-cylinder sector gap should be 13 sectors. Happily, the gap has increased, insuring cylinder reads without missing a rotation. The resulting effective data rate, for both reads and writes (no reformatting for PGP) is

(4 blk/cyl * 7 li/blk * 7100 by/li) / (19 revs * 0.01666 sec/rev)

or 627.8 KB/second. When using three MBA's, this would effectively increase to 941.7 KB/sec, giving reasonable margin for the process while not exactly meeting our 20 percent margin goal.

Configuration

We are forced to use three MBA's in ping-pong fashion in order to increase the RP06's basically low data rate. Due to arguments given in the 'IPC -- Disk Scene Buffer Phasing' section above, there seems to be no simple way to avoid having three drives per MBA, yielding a total of nine drives. An additional drive is required for the annotation data. The third string only requires storage for two scenes, and therefore requires only six drives.

The resulting configuration is given below

```
+---------+
:         :
: string  :--MBA--RP06--RP06--RP06
:         :
:   one   :--MBA--RP06--RP06--RP06
:         :
:   VAX   :--MBA--RP06--RP06--RP06--RP06
:         :
:_________:

+---------+
:         :
: string  :--MBA--RP06--RP06--RP06
:         :
:   two   :--MBA--RP06--RP06--RP06
:         :
:   VAX   :--MBA--RP06--RP06--RP06--RP06
:         :
:_________:

+---------+
:         :
: string  :--MBA--RP06--RP06
:         :
: three   :--MBA--RP06--RP06
:         :
:   VAX   :--MBA--RP06--RP06
:         :
:_________:
```

Buffer Sizes

For IPC, VAX buffer requirements for the NSCI-to-disk transfer will be
      2 blks * 8 li/blk * 6300 by/li * 7 bands * 2 buffers
or 1411 KB. For the disk-to-FFP transfer, they will be
      16 lines * 6300 by/line * 2 buffers
or 202 KB.

For PGP, both the read and write buffer sizes will be
      3 blocks * 7 lines/block * 7100 bytes/line * 2 buffers
or 298 KB.

Schedule

The quoted DEC delivery for RP06's now is 7 months. If an order is placed for the drives during September, delivery is quoted in 4/80. With a little pressure on DEC, delivery of enough disks for one string could be accomplished at an earlier time, meeting the software integration deadline.

Software Impact

Software impact of this configuration is probably minimal, given that a single 'mapping' routine is used by all applications programs to determine where (MBA/disk/track/sector) each data line is stored.

Facility Impact

The baseline facilities configuration is for 5 RM05's, but without the double footprint taken into account. Placing ten RP06's where five were planned for will not be possible at the Lanham facility, at least not for all three strings simultaneously. The Goddard facility will be able to hold the drives if a partition is moved, but power and air-conditioning may still be a problem, as these figures will essentially double for the disk sub-system (reference conversations with Charlie Gregan). It should be noted that an RM05 solution would be roughly as large, and would suffer from the same spacial problems at both facilities. Seen in this light, the RP06 facility problems may not be so severe as to prevent its use as a backup configuration, or even as the final configuration.

Alternate RP06 Solution

An alternate RP06 configuration has been proposed by Dick Kaiser as certainly being physically smaller, and thus less costly, and perhaps being feasible to implement. He undertook a short study to better understand this configuration, the results of which are discussed below.

The configuration which was studied has only 6 dual-ported RP06's with three MBA's and four drives per MBA, and is shown below

```
              1
        |-----0-----|
        |           |
        |     2     |
        |-----0-----|
        |           |
        |  3     5  |
        |--0--|--0--|
        |     |     |
        |  4  |  0  |
        |--0--|--0--|
        |     |     |
        |     |     |
      MBA1  MBA2  MBA3
```

An alternative way of looking at this is via its 'logical' configuration, as represented below

```
       drive 4                drive 5
   MBA 1---0---2 MBA    MBA 2---0---3 MBA     Scene 1

       drive 2                drive 3
   MBA 3---0---1 MBA    MBA 1---0---2 MBA     Scene 2

       drive 0                drive 1
   MBA 2---0---3 MBA    MBA 1---0---3 MBA     Scene 3
```

It is intended that the six drives be separated into three pairs, and scenes be buffered onto these pairs. This implies that at any one time, two drives will be being written to and two read from, with two idle. Since there are only three access paths to the drives, a fixed allocation of MBA's to drives will not provide the throughput required. One must allocate paths to the drives dynamically, and one must insure that the requests for the drives are evenly distributed. The logical configuration diagram above is intended to show which drives are paired as scene buffers.

In order to optimize throughput, we require that MBA's are allocated and cylinders are written in a burst of QIO's so that rotational delays are not encountered within a cylinder. It turns out to be best if three blocks are used per cylinder, with 8 lines/block for bands 1-5, and 10 lines/block for a combination of band 7 and a compressed band 6 (simplifying the data layout). Although this only uses 71 percent of the disks, each drive pair still has 247 MB of useful storage, while only 237 MB are required for a scene of de-replicated data.

If we assume that a cylinder can be written with only a single rotational delay, then the data rate for an RP06 is

(3 blk/cyl * 8 li/blk * 6300 by/li) / (15 revs * 0.01666 sec/rev)

or 604.8 KB/second. With the extra MBA taken into account, this becomes 907.2 KB/second, clearly providing adequate margin.

If we assume that a block can be read with the same single rotational penalty, then we get a read transfer rate of

(1 blk * 8 li/blk * 6300 by/li) / (5.7 revs * 0.016666 sec/rev)

or 530.5 KB/second. When the extra MBA is taken into account, the effective read rate is 795.8 KB/second, meeting our goals.

The dynamic allocation of MBA's is the crux of this configuration. If a read or write request is encountered while a path to the desired drive is available, the allocation is trivial.

Suppose, however, that we are writing to drives 0 and 1 and reading from drives 4 and 5. In this case, we could have MBA's 1, 2, and 3 allocated to drives 4, 5, and 0, respectively. The only other drive to which data could be transmitted is drive 1, but the MBA's to that drive (MBA 1 and MBA 3) are tied up. Suppose that the next MBA to become free is in fact MBA 2. In this case, rather than waiting for MBA 1 or 3 to become free, we must initiate I/O on MBA 2 to the other drive (in this case drive 0), in order to keep the aggregate throughput up. By keeping a large buffer of data ready for transmission, we can transmit data whose "turn hasn't come yet" and transmit the data which was blocked by the MBA conflict at a later time. As long as data to a given drive is written in order, this will not cause seek delays. It is estimated that less than 0.5 milliseconds will be required to schedule a burst of three block transmissions.

The buffering requirements for this configuration are more severe than for the others, with triple buffering (1.89 MB) being required at a minimum, and quadruple buffering (2.52 MB) perhaps being necessary. Although it seems that triple buffering is sufficient, it should be pointed out that there may not be enough memory to allow quadruple buffering should it prove to be necessary.

In summary, this technique trades off cost and facility problems for increased complexity in the software. If the MBA allocation routine is included as part of the Image Data Mapping Routine (discussed above), then the use of this configuration could be transparant to the applications programmer.

RM03 Solutions

Disk Data Format and Performance -- IPC

Given a design criterion of 2 milliseconds for I/O latency, we find that for the RM03, this corresponds to 3.8, or 4 sectors. Looking at Appendix III, we find optimal packing density of 92.29 percent with blocks which are 6 lines long. With this utilization factor, each RM03 holds 62.2 MB, or 0.235 scenes (or 4.25 drives/scene). If thermally de-replicated data were to be used, the figures would change to 0.263 scenes per drive, or 3.799 drives/scene. It is because of this reduction in the number of drives required, that thermal de-replication is suggested when RM03's are used. Mapping the blocks and 'gap' sectors onto the cylinder, we find the following

SECTOR

TRACK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	g	g	g	g	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	g	g	g	g	s	s	s	s

where the symbols are as defined above for the RP06 case.

With a seek time of 8 milliseconds for a 6 cylinder seek (estimated for reading), the equivalent number of sectors is 16, or half a track. Addition of 4 sectors for I/O turn-around brings the total up to 20. As can be seen above, the gap between records 1 and 2 is too small to absorb the seek time. Since only two blocks can be written per cylinder, the number of different cylinder formats is small. No advantage can be gained by changing formats on subsequent cylinders, so a full rotational delay penalty must be absorbed. Were it possible to reduce the block lengths so that the seek sectors were exactly equal to 20, the throughput might increase, although the packing density would decrease. This alternative has not been fully investigated.

Maximum sustainable throughput for this system would be a cylinder read with a full rotational miss prior to the read. This is the same as

(2 blk/cyl * 6 li/blk * 6300 by/li) / (6 revs * 0.01666 sec/rev)

or 756.0 KB/second. Since this does not meet the IPC plus 20 percent read goal, ping-ponging on three MBA's will be required in order to increase throughput to 1134 KB/second, meeting both the read and write requirements.

We will store a band per cylinder on each of the three active drives, writing two cylinders per MBA every time a buffer fills. The de-replicated thermal band will be written every twelve times the other buffers fill, in order that all three MBA's can write thermal data, and thus simplify the layout of data on disk. The diagram below shows this layout, with the band stored on each cylinder indicated for each MBA. Notice that the pattern repeats starting with cylinder 26, and that the bands rotate among the MBA's in order to distribute the data when it is being read.

CYLINDER

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
MBA 1	1	4	3	7	2	5	1	4	3	7	2	5	1	4	3	7	2	5	1	4	3	7	2	5	6	1	4	3	7	2	5
MBA 2	2	5	1	4	3	7	2	5	1	4	3	7	2	5	1	4	3	7	2	5	1	4	3	7	6	2	5	1	4	3	7
MBA 3	3	7	2	5	1	4	3	7	2	5	1	4	3	7	2	5	1	4	3	7	2	5	1	4	6	3	7	2	5	1	4

For disk writes, full cylinders will be written, producing the throughput figure calculated above. For disk reads, every sixth cylinder will be read. Since there will be a rotational miss in any event, the additional seek time (less than 2 milliseconds) will be lost in the rotation, again making the throughput figure reduce to that calculated above.

Disk Data Format and Performance -- PGP

Using the same 2 millisecond I/O latency as for the IPC case, and looking at Appendix IV, we find that optimal use of the RM03's for PGP require a blocking factor of 5, and a utilization factor of 86.67 percent. The reduction in effective storage for this process is not significant, as only two scenes need be stored on disk. The data layout is straightforward, and is given below :

SECTOR

TRACK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	g	g	g	g	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	g	g	g	g	s	s	s	s	s	s	s	s	s	s	s	s	s	s

where the symbols are as defined previously. Since only 17 sectors are available for seek delays, with 20 required, we can always anticipate cylinder reads with rotational latency penalties. Throughput is then given by

(2 blk/cyl * 5 li/blk * 7100 by/li) / (6 revs * 0.016666 sec/rev) or 710 KB/second. With the third MBA, this effectively increases to 1065 KB/second, meeting the requirements.

Buffer Sizes

For IPC, VAX buffer requirements for the NSCI-to-disk transfer will be
2 blks * 6 li/blk * 6300 by/li * 6 bands * 2 buffers
plus
6 blks * 6 li/blk * 6300 by/li * 1 band * 2 buffers
or 1361 KB. For the disk-to-FPP transfer, they will be
16 lines * 6300 by/line * 2 buffers
or 202 KB.

For PGP, both the read and write buffer sizes will be
     3 blks * 5 lines/blk * 7100 bytes/line * 2 buffers
or 213 KB.

Configuration

Since three MBA ping-ponging will be required in order to meet the PGP requirements, and since storage for three scenes must be available on disk for IPC, and since a de-replicated scene fits on 4 RM03 drives, it would seem that four drives per MBA will be required. As discussed under 'IPC -- Read/Write Interference' above, this actually must increase to five per MBA in order to avoid read/write interference. (A solution with four drives per MBA was discovered which requires tight NSCI/NSCO time matching and therefore is not recommended.) These additional drives imply that thermal de-replication is not required, and the above data layouts could be modified accordingly if desired. One more drive is required for the annotation data. String three requires only nine drives, since it need only store two scenes, and need not perform the reformatting operation.

The resulting configuration is given below

```
 +---------+
 |         |
 | string  |--MBA--RM03--RM03--RM03--RM03--RM03
 |         |
 |   one   |--MBA--RM03--RM03--RM03--RM03--RM03
 |         |
 |   VAX   |--MBA--RM03--RM03--RM03--RM03--RM03--RM03
 |         |
 |_________|

 +---------+
 |         |
 | string  |--MBA--RM03--RM03--RM03--RM03--RM03
 |         |
 |   two   |--MBA--RM03--RM03--RM03--RM03--RM03
 |         |
 |   VAX   |--MBA--RM03--RM03--RM03--RM03--RM03--RM03
 |         |
 |_________|

 +---------+
 |         |
 | string  |--MBA--RM03--RM03--RM03
 |         |
 |  three  |--MBA--RM03--RM03--RM03
 |         |
 |   VAX   |--MBA--RM03--RM03--RM03
 |         |
 |_________|
```

Schedule

The quoted DEC delivery for RM03's now is 7 months. If an order is placed for the drives during September, delivery is quoted in 4/80. With a little pressure on DEC, delivery of enough disks for one string could be accomplished at an earlier time, meeting the software integration deadline.

Software Impact

Software impact of this configuration is probably minimal, given that a single 'mapping' routine is used by all applications programs to determine where (MBA/disk/track/sector) each data line is stored.

Facility Impact

The baseline facilities configuration is for 5 RM05's, but without the double footprint taken into account. It appears that finding space for 16 RM03's per string is essentially impossible for both the Lanham and Goddard facilities.

RM05 Solutions

A rough analysis of a configuration for RM05's is given below for the purposes of comparison with the other configurations, and to indicate the analytical method used to derive the configuration. The size of the RM05 makes the number of reasonable solutions quite large, and this increased flexibility makes selecting even a mildly optimized solution quite difficult.

Disk Data Format and Performance -- IPC

The 2 millisecond I/O latency figure corresponds to a four sector delay for RM05's. Looking in Appendix III, we chose a blocking factor of nine with a utilization factor of 91.07 percent and 5 blocks per cylinder. This results in the following layout of data on a cylinder

	SECTOR																														
TRACK										1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3	3	3
	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	g	g	g	g	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
8	2	2	g	g	g	g	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
9	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
30	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
11	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	g	g	g	g	4	4	4	4	4	4	4
12	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
13	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
14	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
15	4	4	4	4	4	4	4	4	g	g	g	g	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
16	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
17	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
18	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	g	g	g	g	x
19	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

where the symbols are as before, only 'x' indicates that the sector is not used. As with the RM03, seek latency is 13 sectors, and we would like to find a way to write data to consecutive cylinders so that we can minimize the wasted time. After we write block five of the first cylinder, we must determine which block will next pass under the heads after the seek. By determining the number of sectors between the ends of each block and the start of all other blocks, we can attempt to minimize the lost time. Such data in the form of a table is given below.

		... to start of block number				
		1	2	3	4	5
f e o b n	1	17	4	23*	10	29
r n f l u	2	30	17	4	23*	10
o d o m	3	11	30	17	4	23*
m c .	4	24*	11	30	17	4
k .	5	3	24*	11	30	17

It can be seen that the smallest numbers in the matrix which are greater than 13 are those with stars after them. They represent the 'hops' which must be taken between consecutive cylinders when laying out the data on the RM05. Taking this into account, we find that the order in which data is written onto an RM05 is as given in the following table.

	CYLINDER																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Block 1	1	7	13	19	25	26	32	38	44	50	51	57	63	69	75	76	...
Block 2	2	8	14	20	21	27	33	39	45	46	52	58	64	70	71	77	...
Block 3	3	9	15	16	22	28	34	40	41	47	53	59	65	66	72	78	...
Block 4	4	10	11	17	23	29	35	36	42	48	54	60	61	67	73	79	...
Block 5	5	6	12	18	24	30	31	37	43	49	55	56	62	68	74	80	...

If we use the above pattern to write band data to each disk, with one block of data per band and no thermal de-replication, we obtain the following pattern of bands on the disk. Note that the pattern repeats after cylinder 21.

	CYLINDER																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Block 1	1	7	6	5	4	5	4	3	2	1	2	1	7	6	5	6	5	4	3	2	3
Block 2	2	1	7	6	7	6	5	4	3	4	3	2	1	7	1	7	6	5	4	5	4
Block 3	3	2	1	2	1	7	6	5	6	5	4	3	2	3	2	1	7	6	7	6	5
Block 4	4	3	4	3	2	1	7	1	7	6	5	4	5	4	3	2	1	2	1	7	6
Block 5	5	6	5	4	3	2	3	2	1	7	6	7	6	5	4	3	4	3	2	1	7

If we write one cylinder at a time, we lose 24 sectors (worst case) during the seek interval. This resolves to a throughput rate of

(5 blk/cyl*9 li/blk*6300 by/li) / (18.75 revs*0.01666 sec/rev)

or 907.2 KB/second, well within our goal.

If we read the data, we read one block and then seek at least one cylinder. The data layout shown above indicates that quite frequently we will be attempting to read sequential blocks on adjacent cylinders, forcing a rotational latency. If we investigate this worst case, we find that the total number of 'sector-times' which it takes to read the block is 111 (for the block) plus 13 (for the seek) plus 32 (for the missed rev), or

(9 li/blk * 6300 by/li) / ((111 + 32 + 13) * 0.000521 sec)

or 697.6 KB/second. Although this will survive during IPC, it is not conservative. If we take an 'average' rotational penalty of 32

sector-times, then we get
(9 li/blk * 6300 by/li) / ((111 + 32) * 0.000521 sec)
or 761.0 KB/second, somewhat better.

A data layout optimized for reading would perform better, and in fact increasing the number of blocks written per band might help. One could dedicate cylinders to each band (as in the RM03 case) without an increased VAX memory buffer penalty if one scatter-wrote data to seven cylinders, one block per cylinder at a time. Some of the data layouts which result when additional blocks are written per band are more complex, and have not been investigated fully at this time.

Another way to increase read throughput would be to increase the number of 'gap' sectors from 4 to 8, thus increasing the values in the inter-block distance table above by 4 (mod 32). This would make it feasible to use seek sector distances of 14 and 15, instead of 23 and 24, reducing wasted time and changing the layout of bands on the disk. This approach has not been fully investigated at this time.

Disk Data Format and Performance -- PGP

Using the 2 millisecond turn-around time, Appendix IV indicates that a blocking factor of 8 with five blocks per cylinder is optimal. It turns out that the block lengths in this case are identical to the IPC case, and the cylinder data layout diagram need not be repeated for this case.

Read and write rates are the cylinder rates, or
(5 blks/cyl * 8 li/blk * 7100 by/li) / (19 revs * 0.01666 sec/rev)
or 896.8 KB/second, which is marginal.

Buffer Sizes

For IPC write buffers, we get
1 blk * 9 li/blk * 6300 by/li * 2 buffers
or 113 KB. For IPC read buffers, we get
2 blocks * 9 lines/block * 6300 bytes/line * 2
or 227 KB.

PGP buffers are
2 blocks * 8 lines/block * 7100 bytes/line * 2 buffers
or 227 KB.

Configuration

The RM05's can be hung from two MBA's with all drives dual-ported to allow access from either MBA. One MBA will be dedicated to reading and the other to writing, with the dual-port feature being used to allocate drives to each function.

Tradeoff Matrix

	RM03	RP06	RM05	
Dereplicate Thermal	yes	no	no	
Thermal on RAA disk	yes	yes	yes	
Annotation on RAA disk	no	no	no	
Annotation on RAA MBA	yes	yes	yes	
Num drives, String 1	16	10	5	
Num drives, String 2	16	10	5	
Num drives, String 3	9	6	3	
Number of spare drives	0	0	0	
IPC blocking factor	6	8	9	
IPC blocks/cylinder	2	4	5	
IPC blocks/band	2	2	1	
IPC bands/cyl	1	2	5	
IPC write rate	1134	907.2	907.2	KB/second
IPC read rate	1134	907.2	761	KB/second
IPC write buffer	1361	1411	113	KB
IPC read buffer	202	202	227	KB
PGP blocking factor	5	7	8	
PGP read/write rate	1065	941.7	896.8	KB/second
PGP buffers (each)	213	298	227	KB
Number of MBA's	3	3	3	
Facility impact	extreme	signif	signif	

VAX CPU Loading Due to QIO Compute Time During IPC

From	To	QIO's	Block	Length	Length	Page	QIO	QIO	QIO	QIO
		per	Factor	Bytes	Pages	Lock	Lock	Lock	Some	None
		sec				Time	Unit	Total	Total	Total
						(ms)	(ms)	(ms)	(ms)	(ms)
NSCI	VAX	96	1	6300	13	0.65	1.85	177.6	73.9	73.9
VAX	DISK	16	6	37800	74	3.7	4.9	78.4	78.4	20.8
DISK	VAX	17	6	37800	74	3.7	4.9	83.3	83.3	22.1
VAX	FFP	17	6	37800	74	3.7	4.9	83.3	22.1	22.1
FFP	VAX	12	9	63900	125	6.25	7.45	89.4	15.6	15.6
VAX	NSCO	18	6	42600	84	4.2	5.4	97.2	23.4	23.4
TOTAL								609.2	296.7	177.9

Notes:

Column 1 : data source.

Column 2 : data destination.

Column 3 : number of QIO's requested per second, based on required data rate (per Dick Kaiser's unpublished throughput study) and the estimated block size.

Column 4 : blocking factor.

Column 5 : block length, in bytes.

Column 6 : block length, in pages.

Column 7 : the time it would take to lock the pages of the block into memory at 0.050 milliseconds per page.

Column 8 : the total time it would take to perform a QIO for this phase of the processing, consisting of the page lock time from Column 7 plus 1.2 milliseconds general QIO overhead.

Column 9 : total CPU utilization if page locking is performed on all QIO's. Multiplies the per-QIO time from Column 8 by the number of QIO's from Column 3.

Column 10 : total CPU utilization if page locking is performed only by standard DEC (i.e., disk) drivers, and QIO chaining is used for groups of ten NSCI records. NSCI CPU time is then given by ((96/10) * 1.2 ms) + (96 * 0.65 ms).
Note: this is the IPC baseline figure, per Phil Miller.

Column 11 : total CPU utilization if most page locking is eliminated. Multiplies the number of QIO's from Column 3 by 1.3 milliseconds for each QIO.

Disk Drive Characteristics

		RM03	RM05	RP06	UNITS
1	SOURCE	CDC9762	CDC9766	MEM 677	
2	CYLINDERS	823	823	815	
3	SURFACES	5	19	19	
4	SECTORS/TRACK	32	32	22	
5	USABLE TRACKS	4114	15636	15485	
6	USABLE SECTORS	131648	500352	340670	
7	USABLE K-BYTES	67404	256180	174423	
8	ROTATION SPEED	3600	3600	3600	RPM
9	AVERAGE LATENCY	8.3	8.3	8.3	MS
10	MAXIMUM SEEK	55	55	50	MS
11	AVERAGE SEEK	30	30	28	MS
12	ONE CYL. SEEK	6	6	7	MS
13*	AVG SECTOR WAIT	521	521	758	MICROSECONDS
14	ONE CYL. SEEK	11.52	11.52	9.23	SECTORS
15*	MAX XFER RATE	1212	1212	806	KB/SEC
16*	AVG BURST RATE	983040	983040	675840	BYTES/SEC
17*	AVG MLT CYL RAT	819200	933888	642048	BYTES/SEC
18*	AVG MLT CYL RAT	1600	1824	1254	SECT/SEC
19	SECTOR SIZE	512	512	512	BYTES
20	TRACK SIZE	16384	16384	11264	BYTES
21	CYLINDER SIZE	81920	311296	214016	BYTES
22	SECTORS/CYL	160	608	418	
23	CYL READ TIME	83.3	316.6	316.6	MS
24	CYL READ RATE	12	3.16	3.16	CYLS/SECOND
25	HEIGHT	40	?	47	INCHES
26	DEPTH	33	?	32	INCHES
27	WIDTH	22	?	31	INCHES
28	POWER DISSIPATE	1210	?	1285	WATTS

NOTES :

LINE 13 : THE AVERAGE TIME FOR A SECTOR TO ROTATE UNDER THE HEADS OF THE GIVEN DISK. CALCULATED AS FOLLOWS :
1 / (60 TRACKS/SECTOR * ? SECTORS/TRACK)

LINE 15 : THE MAXIMUM DATA RATE DURING A SECTOR READ OR WRITE. THE AVERAGE RATE IS LOWER, DUE TO INTER-SECTOR PAUSES. OBTAINED FROM THE LITERATURE.

LINE 16 : THE AVERAGE DATA RATE TO/FROM A DRIVE OVER A TRACK. CALCULATED AS FOLLOWS :
512 BYTES/SECTOR * ? SECTORS/TRACK * 60

LINES 17 AND 18 : THE AVERAGE TRANSFER RATE IF ONE ROTATION IS LOST PER CYLINDER READ. CALCULATED AS FOLLOWS :
(512 BYTES/SECTOR * ? SECTORS/TRACK * ? SURFACES) /
(16.66666 MILLISECONDS/TRACK * (? SURFACES + 1))

COST (LIST PRICES) :

RM03 SINGLE/DUAL ACCESS CONTROLLER + DRIVE = REM03A/BA = $25.0K/$33.0K
RM03 SINGLE/DUAL ACCESS DRIVE ONLY = RM03A/BA = $19.0K/$21.0K
RP06 SINGLE/DUAL ACCESS CONTROLLER + DRIVE = REP06A/BA = $44.0K/$56.6K
RP06 SINGLE/DUAL ACCESS DRIVE ONLY = RP06A/BA = $34.0K/$39.2K

```
PROGRAM THRU(TTY);
  (* PROGRAM TO CALCULATE IGF UTILIZATION FACTORS OF RAA DISKS *)
CONST
(* BLOCKS PER CYLINDER FOR ALL DISKS *)
  RM03SPC = 160;
  RM05SPC = 608;
  RP06SPC = 418;
  (* NUMBER OF BYTES PER LINE OF A-TAPE DATA *)
  LINLEN = 6300.0;

VAR
  BFACT, BLENS, ESECT, RM03, RM05, RP06 : INTEGER;
  BLENB, PUTIL, RM03PU, RM05PU, RP06PU : REAL;
  OFILE : TEXT;

BEGIN
REWRITE(OFILE);
FOR ESECT := 0 TO 7 DO BEGIN
  WRITELN(OFILE);
  WRITELN(OFILE,
  ' BF    BLOCK  BLK E PERCNT RM PERCNT RM PERCNT RP PERCNT');
  WRITELN(OFILE,
  '       LNGTH  LNG M UTIL   03 UTIL   05 UTIL   06 UTIL');
  WRITELN(OFILE,
  '       (BYTE) (S) P');
  WRITELN(OFILE);
  FOR BFACT := 1 TO 10 DO BEGIN
    WRITE(OFILE,BFACT:3);
    BLENB := BFACT * LINLEN;
    WRITE(OFILE,BLENB:9:1);
    BLENS := TRUNC((BLENB / 512) + 0.99999);
    WRITE(OFILE,BLENS:4);
    WRITE(OFILE,ESECT:2);
    BLENS := BLENS + ESECT;
    PUTIL := (BLENB * 100.0) / (BLENS * 512.0);
    WRITE(OFILE,PUTIL:7:2);
    RM03 := TRUNC(RM03SPC / BLENS);
    WRITE(OFILE,RM03:3);
    RM03PU := (RM03 * PUTIL * BLENS) / RM03SPC;
    WRITE(OFILE,RM03PU:7:2);
    RM05 := TRUNC(RM05SPC / BLENS);
    WRITE(OFILE,RM05:3);
    RM05PU := (RM05 * PUTIL * BLENS) / RM05SPC;
    WRITE(OFILE,RM05PU:7:2);
    RP06 := TRUNC(RP06SPC / BLENS);
    WRITE(OFILE,RP06:3);
    RP06PU := (RP06 * PUTIL * BLENS) / RP06SPC;
    WRITE(OFILE,RP06PU:7:2);
    WRITELN(OFILE);
    END; (* FOR BFACT *)
  END; (* FOR ESECT *)
END.
```

RAA Disk Utilization Factors for IPC

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	0	+94.65	12	+92.29	46	+93.09	32	+94.20
2	+12600.0	25	0	+98.44	6	+92.29	24	+97.14	16	+94.20
3	+18900.0	37	0	+99.77	4	+92.29	16	+97.14	11	+97.14
4	+25200.0	50	0	+98.44	3	+92.29	12	+97.14	8	+94.20
5	+31500.0	62	0	+99.23	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	0	+99.77	2	+92.29	8	+97.14	5	+88.31
7	+44100.0	87	0	+99.00	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	0	+99.43	1	+61.52	6	+97.14	4	+94.20
9	+56700.0	111	0	+99.77	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	0	+99.23	1	+76.90	4	+80.95	3	+88.31

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	1	+87.89	11	+84.59	43	+87.02	29	+85.37
2	+12600.0	25	1	+94.65	6	+92.29	23	+93.09	16	+94.20
3	+18900.0	37	1	+97.14	4	+92.29	16	+97.14	11	+97.14
4	+25200.0	50	1	+96.51	3	+92.29	11	+89.05	8	+94.20
5	+31500.0	62	1	+97.66	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	1	+98.44	2	+92.29	8	+97.14	5	+88.31
7	+44100.0	87	1	+97.88	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	1	+98.44	1	+61.52	6	+97.14	4	+94.20
9	+56700.0	111	1	+98.88	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	1	+98.44	1	+76.90	4	+80.95	3	+88.31

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)

RAA Disk Utilization Factors for IPC

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	2	+82.03	10	+76.90	40	+80.95	27	+79.48
2	+12600.0	25	2	+91.15	5	+76.90	22	+89.05	15	+88.31
3	+18900.0	37	2	+94.65	4	+92.29	15	+91.07	10	+88.31
4	+25200.0	50	2	+94.65	3	+92.29	11	+89.05	8	+94.20
5	+31500.0	62	2	+96.13	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	2	+97.14	2	+92.29	8	+97.14	5	+88.31
7	+44100.0	87	2	+96.78	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	2	+97.46	1	+61.52	6	+97.14	4	+94.20
9	+56700.0	111	2	+98.00	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	2	+97.66	1	+76.90	4	+80.95	3	+88.31

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	3	+76.90	10	+76.90	38	+76.90	26	+76.54
2	+12600.0	25	3	+87.89	5	+76.90	21	+85.00	14	+82.42
3	+18900.0	37	3	+92.29	4	+92.29	15	+91.07	10	+88.31
4	+25200.0	50	3	+92.87	3	+92.29	11	+89.05	7	+82.42
5	+31500.0	62	3	+94.65	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	3	+95.88	2	+92.29	7	+85.00	5	+88.31
7	+44100.0	87	3	+95.70	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	3	+96.51	1	+61.52	5	+80.95	4	+94.20
9	+56700.0	111	3	+97.14	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	3	+96.89	1	+76.90	4	+80.95	3	+88.31

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)

RAA Disk Utilization Factors for IPC

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	4	+72.38	9	+69.21	35	+70.83	24	+70.65
2	+12600.0	25	4	+84.86	5	+76.90	20	+80.95	14	+82.42
3	+18900.0	37	4	+90.03	3	+69.21	14	+85.00	10	+88.31
4	+25200.0	50	4	+91.15	2	+61.52	11	+89.05	7	+82.42
5	+31500.0	62	4	+93.22	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	4	+94.65	2	+92.29	7	+85.00	5	+88.31
7	+44100.0	87	4	+94.65	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	4	+95.57	1	+61.52	5	+80.95	4	+94.20
9	+56700.0	111	4	+96.30	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	4	+96.13	1	+76.90	4	+80.95	3	+88.31

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+6300.0	13	5	+68.36	8	+61.52	33	+66.79	23	+67.71
2	+12600.0	25	5	+82.03	5	+76.90	20	+80.95	13	+76.54
3	+18900.0	37	5	+87.89	3	+69.21	14	+85.00	9	+79.48
4	+25200.0	50	5	+89.49	2	+61.52	11	+89.05	7	+82.42
5	+31500.0	62	5	+91.83	2	+76.90	9	+91.07	6	+88.31
6	+37800.0	74	5	+93.45	2	+92.29	7	+85.00	5	+88.31
7	+44100.0	87	5	+93.62	1	+53.83	6	+85.00	4	+82.42
8	+50400.0	99	5	+94.65	1	+61.52	5	+80.95	4	+94.20
9	+56700.0	111	5	+95.47	1	+69.21	5	+91.07	3	+79.48
10	+63000.0	124	5	+95.39	1	+76.90	4	+80.95	3	+88.31

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)

RAA Disk Utilization Factors for PGP

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	0	+99.05	11	+95.34	43	+98.07	29	+96.21
2	+14200.0	28	0	+99.05	5	+86.67	21	+95.79	14	+92.89
3	+21300.0	42	0	+99.05	3	+78.00	14	+95.79	9	+89.57
4	+28400.0	56	0	+99.05	2	+69.34	10	+91.23	7	+92.89
5	+35500.0	70	0	+99.05	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	0	+99.05	1	+52.00	7	+95.79	4	+79.62
7	+49700.0	98	0	+99.05	1	+60.67	6	+95.79	4	+92.89
8	+56800.0	111	0	+99.94	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	0	+99.84	1	+78.00	4	+82.11	3	+89.57

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	1	+92.45	10	+86.67	40	+91.23	27	+89.57
2	+14200.0	28	1	+95.64	5	+86.67	20	+91.23	14	+92.89
3	+21300.0	42	1	+96.75	3	+78.00	14	+95.79	9	+89.57
4	+28400.0	56	1	+97.31	2	+69.34	10	+91.23	7	+92.89
5	+35500.0	70	1	+97.66	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	1	+97.89	1	+52.00	7	+95.79	4	+79.62
7	+49700.0	98	1	+98.05	1	+60.67	6	+95.79	4	+92.89
8	+56800.0	111	1	+99.05	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	1	+99.05	1	+78.00	4	+82.11	3	+89.57

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)

RAA Disk Utilization Factors for PGP

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	2	+86.67	10	+86.67	38	+86.67	26	+86.26
2	+14200.0	28	2	+92.45	5	+86.67	20	+91.23	13	+86.26
3	+21300.0	42	2	+94.55	3	+78.00	13	+88.95	9	+89.57
4	+28400.0	56	2	+95.64	2	+69.34	10	+91.23	7	+92.89
5	+35500.0	70	2	+96.30	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	2	+96.75	1	+52.00	7	+95.79	4	+79.62
7	+49700.0	98	2	+97.07	1	+60.67	6	+95.79	4	+92.89
8	+56800.0	111	2	+98.17	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	2	+98.27	1	+78.00	4	+82.11	3	+89.57

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	3	+81.57	9	+78.00	35	+79.83	24	+79.62
2	+14200.0	28	3	+89.47	5	+86.67	19	+86.67	13	+86.26
3	+21300.0	42	3	+92.45	3	+78.00	13	+88.95	9	+89.57
4	+28400.0	56	3	+94.01	2	+69.34	10	+91.23	7	+92.89
5	+35500.0	70	3	+94.98	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	3	+95.64	1	+52.00	6	+82.11	4	+79.62
7	+49700.0	98	3	+96.11	1	+60.67	6	+95.79	4	+92.89
8	+56800.0	111	3	+97.31	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	3	+97.50	1	+78.00	4	+82.11	3	+89.57

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)

RAA Disk Utilization Factors for PGP

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	4	+77.04	8	+69.34	33	+75.27	23	+76.30
2	+14200.0	28	4	+86.67	5	+86.67	19	+86.67	13	+86.26
3	+21300.0	42	4	+90.44	3	+78.00	13	+88.95	9	+89.57
4	+28400.0	56	4	+92.45	2	+69.34	10	+91.23	6	+79.62
5	+35500.0	70	4	+93.70	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	4	+94.55	1	+52.00	6	+82.11	4	+79.62
7	+49700.0	98	4	+95.17	1	+60.67	5	+79.83	4	+92.89
8	+56800.0	111	4	+96.47	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	4	+96.75	1	+78.00	4	+82.11	3	+89.57

BF	BLOCK	BLK	E	PERCNT	RM	PERCNT	RM	PERCNT	RP	PERCNT
	LNGTH	LNG	M	UTIL	03	UTIL	05	UTIL	06	UTIL
	(BYTE)	(S)	P							
1	+7100.0	14	5	+72.99	8	+69.34	32	+72.99	22	+72.99
2	+14200.0	28	5	+84.04	4	+69.34	18	+82.11	12	+79.62
3	+21300.0	42	5	+88.51	3	+78.00	12	+82.11	8	+79.62
4	+28400.0	56	5	+90.93	2	+69.34	9	+82.11	6	+79.62
5	+35500.0	70	5	+92.45	2	+86.67	8	+91.23	5	+82.94
6	+42600.0	84	5	+93.49	1	+52.00	6	+82.11	4	+79.62
7	+49700.0	98	5	+94.24	1	+60.67	5	+79.83	4	+92.89
8	+56800.0	111	5	+95.64	1	+69.34	5	+91.23	3	+79.62
9	+63900.0	125	5	+96.00	1	+78.00	4	+82.11	3	+89.57

Column 1 : blocking factor (number of lines per block)
Column 2 : block length (bytes)
Column 3 : block length (sectors)
Column 4 : number of empty/gap sectors allocated after each block
Column 5 : percent utilization of storage within a block (including the empty/gap sectors)
Column 6 : number of blocks per cylinder on an RM03
Column 7 : percent utilization of storage within a cylinder for an RM03 (and therefore, percent utilization of the disk)
Column 8 : number of blocks per cylinder on an RM05
Column 9 : percent utilization of storage within a cylinder for an RM05 (and therefore, percent utilization of the disk)
Column 10 : number of blocks per cylinder on an RP06
Column 11 : percent utilization of storage within a cylinder for an RP06 (and therefore, percent utilization of the disk)