STRETCH STAFF REPORT
Sept '57

PRODUCT PLANNING STAFF REPORT

Project Stretch September 9, 1957

Contents

STRETCH

Section I - STRETCH Program

The STRETCH program (Project 7000) is basically an advanced development program. High-speed components are utilized to satisfy the computing requirements represented by the Los Alamos AEC contract and the specific data handling requirements represented by the proposed BuShips contract.

Both of these applications have a common requirement of tremendously high operational speed. Certain of the operations which must be performed are apparently also common to both machine systems. For this reason, these common elements have been placed in a package, called the B (for base) machine. The characteristics of the B machine are such that it can perform accounting processing at speeds about 50 times that of the 705. This machine is also capable of performing engineering or scientific calculations at a speed of about 10 times that of the 704. Operationally for this type of calculation, this machine system is in the same performance area as the proposed Sperry Rand LARC.

An additional unit, called H, may be attached to B. H and B combined will satisfy the requirements of the proposed BuShips contract. The combined units would be capable of handling the specialized applications which have been proposed by that agency.

Another unit, called S, may be attached to B. B and S combined must satisfy the requirements of the existing Los Alamos AEC contracts. The combined units would be capable of performing engineering and scientific calculations at speed of about 100 times that of the 704. The specifications for this system are such as to secure for IBM the top-level scientific and engineering computing customers in the 1960-1963 period. These goals, in terms of systems, are illustrated in Figure I-1. The problem is to design B, H, and S in such a way as to satisfy the possibly conflicting requirements of a commercial system, the high-speed computing requirements of Los Alamos AEC, and the specialized processing of BuShips.

The present approach to the problem is a solution which was presented by a combined group of Engineering and Product Planning personnel working in Poughkeepsie during the month of May 1957. This committee was formed "to study the various proposals for several advanced solid-state computers to see if a compromise solution could be achieved, such that IBM could develop three machine systems with a single integrated engineering program". The characteristics of the three machine systems are given in the section entitled,"System Characteristics."

In general, the committee concluded that it would be possible to set forth a workable design incorporating the goals described above, but the question of price was excluded from consideration. It was noted that certain compromises would be necessary to consolidate the engineering effort in the manner proposed. These compromises were not thought to be serious.

The program has been progressing on this approach. Prior to the committee action just described there had been a considerable effort on the part of Engineering to define a complete HARVEST (BuShips) system. Very little engineering effort has been expended upon the specifications of a system capable of satisfying the Los Alamos contract. Product Planning personnel have spent large amounts of time on this part of the problem. Engineering had spent a considerable amount of time in defining a commercial system. Product Planning people felt that such a system could not be derived from the advanced components at any reasonable cost in the near future.

In defining the STRETCH objectives, there can be no question that the program is advancing technology. It is also insuring IBM prestige in the high-speed system area. Considerable uncertainty is introduced when one suggests that these objectives can be met, while at the same time incorporating rather low cost figures for the system. It has been suggested that a minimum but complete commercial system would rent for $60,000 a month. This is speculation and represents an uncertain goal of the STRETCH program. There is no question that this is a desirable goal, but it is not a certain goal.

Figure I - 1

PRESENT STRETCH GOALS

I/O Memory)	+	B	=	System 50 times faster than 705 or 10 times faster than 704
I/O Memory)	+	B + H	=	System Satisfying Buships Requirements
I/O Memory)	+	B + S	=	System Satisfying Los Alamos Requirements at about 100 times 704 speed

Section II - Present Status

PLANNING

Machine Definition A first pass detailing of the B and S machines has just been completed and was presented to Los Alamos August 20. A programmed evaluation of B+S using both classified and unclassified problems is underway.

Planning on the Basic Exchange and the general input-output operations is detailed enough to begin programming. Definition of the medium speed input-output units themselves, however, is further from completion for several reasons. One is an attempt to standardize this equipment across system boundaries; another is that speeds of certain units only recently have been fixed. Input-output instructions for specific units have not been detailed.

A new definition of the memory bus is now under evaluation.

Study of Customer Requirements Product Planning representatives have visited the following potential users of a Stretch system in order to examine their problems and determine their requirements:

United Aircraft
Knolls Atomic Power Laboratory
Douglas Aircraft
North American Aviation
Northrop Aircraft
Lockheed Aircraft
Rand Corporation
Systems Development Corporation
Ramo-Wooldridge Corporation
University of California Radiation Laboratory
Westinghouse Atomic Power Division

A file maintenance problem has been programmed in detail for the B machine. Some of the results of the study are given in the section on Product Planning Evaluation. A neutron diffusion problem has been programmed for B+S.

ENGINEERING

Transistors The following is taken from Mr. Dunwell's report of August 8 to the AEC:

> "we have modified the fabrication process for the PNP switching transistors. The new process has proven to be highly satisfactory from the standpoints of producing transistors with the necessary qualifications for our use. The yield of satisfactory transistors is also good, insuring that the process will be a reliable one in production.

The development of the process for assembly of the NPN switching transistor is about two months behind the PNP, primarily because we started its development somewhat later than the PNP. About one-third of the transistors being produced are NPN units. A sufficient number of power driver transistors is now available for the bench array of the two microsecond memory, and we expect that the array will be operating before my next report."

Circuits New standard circuit cards are now being designed, with completion scheduled in three to four weeks.

Data Flow Model A model using more than five thousand transistors is expected to be in operation in a short time. It will include two registers and an arithmetic device.

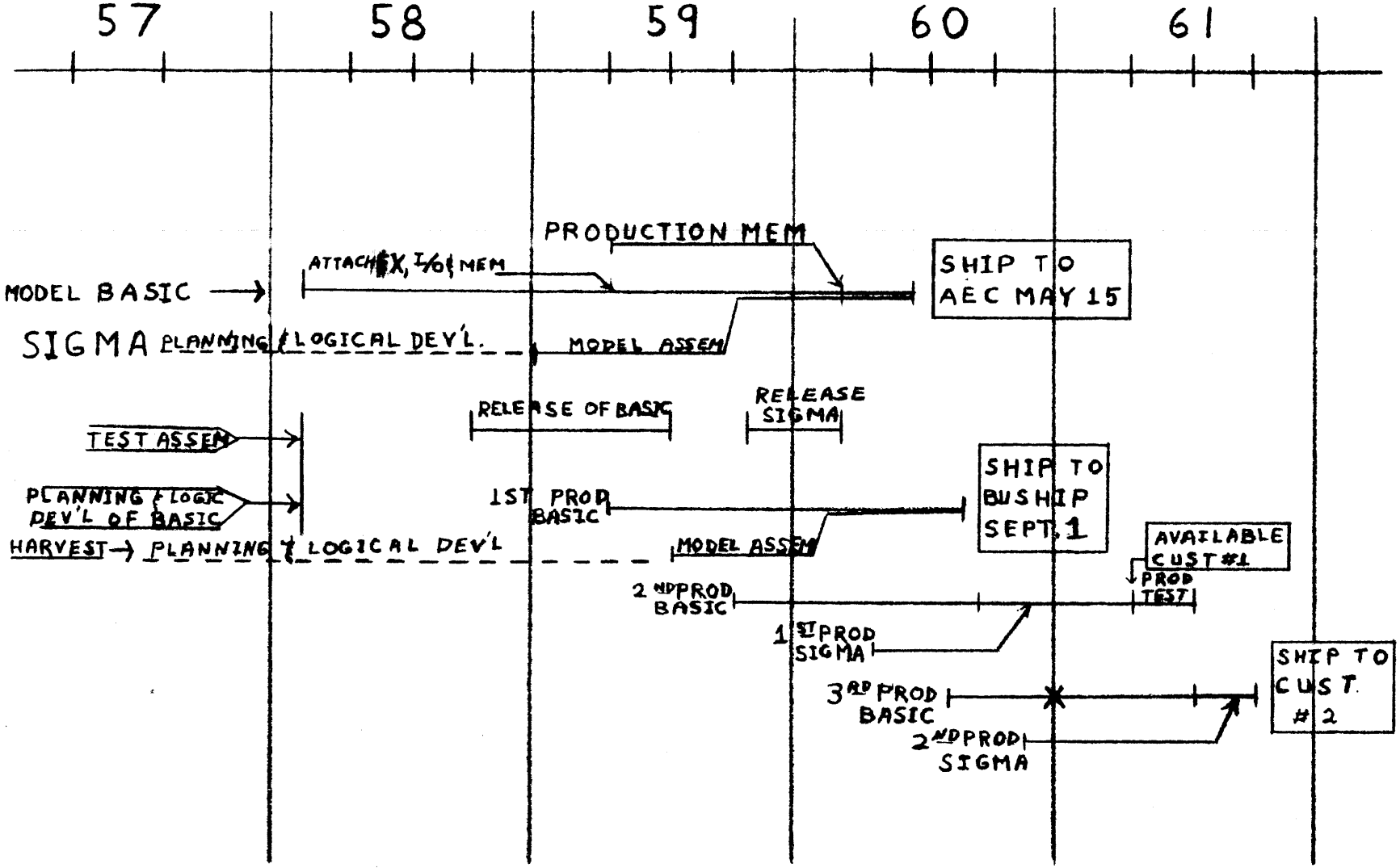
Section III - Schedules

The schedule for meeting the AEC and BuShips commitments is shown in Figure III-1. This figure also indicates tentative release periods for the B and S machines and the schedule for the first two production machines. The immediate goal is, of course, to deliver a B plus S system to Los Alamos, New Mexico on or before May 15, 1960. Under the contemplated plans for BuShips, a B plus H system is to be delivered to that organization on or before September 1, 1960.

Both B and S systems could be released during 1959 and 1960 so that production machines could be available in either the B (basic) system or in the B+S configuration during the second quarter of 1961.

Each of the machine systems B, B+S and B+H are to be specified prior to January 1, 1958. A preliminary manual to satisfy Harvest (B+H) requirements has been written. This manual was published on May 1, 1957. A manual describing a system capable of satisfying AEC requirements will be written by October 1, 1957. Agreement on the specifications must be obtained with Los Alamos AEC personnel by November 17, 1957.

A preliminary version of the B system exists in the Harvest manual. This version presumably will be modified to some extent as the AEC manual is written. Engineering will write a revised manual for this machine. The manual was scheduled for completion on July 1, but a more detailed approach was taken and indications are that this manual will be finished October 1.



STRETCH SCHEDULE

Section IV - Financing

The only existing machine-construction contract in the STRETCH area is a contract with the AEC (Los Alamos) for a device capable of handling their calculation problems. This contract is for a fixed cost of $4,300,000. At the time this contract was signed, it was estimated that the total cost of the program would require IBM to expend something like an additional $6,000,000. It was anticipated that some of this money would be derived from other smaller, but contributing, research and development contracts. To date there are the following contracts:

Project	Source	End Date
Transistor Circuits	AFCRC	7/31/57
Silo (memory)	BuShips	8/31/57
Plantation or Rancho	BuShips	10/31/57
Tractor (tapes)	BuShips	10/1/57

The following contracts are proposed:

Project	Source	End Date
Rancho Extension	BuShips	10/31/57
Harvest System	BuShips	9/1/60

The present plan is to obtain a fixed cost contract for the B machine from BuShips and a cost-fee contract for the H attachment. The Product Development people are planning on obtaining this contract prior to November 1. If this is not done, additional financing will probably be required by IBM.

A summary of the Stretch program cost and recovery is on the following page. (the figures given are estimates from Engineering which have not been verified by the Price Analysis Department). The figures do not include costs for release to production. Estimated cost for release is $7,500,000 of which $1,500,000 to $1,800,000 will be required in 1958. This is in addition to the $700,000 currently in the Product Development budget for Stretch for 1958.

COST UNDER PROPOSED HARVEST CONTRACT

	COST	RECOVERY
	(in thousands)	
AEC Commitment (includes $3,168 Basic Dev.)	$12,078	$4,300
AFCRC Transistor Circuits	---	380
BuShips Support	---	496
Present Commitment	$12,078	$5,176
Basic Development - Support by BuShips	$ 1,914	
Basic System for Harvest Assembly	800	$5,000
Basic System for Harvest Hardware	2,286	
Harvest Engineering and Assembly	1,385	1,385
Harvest Hardware	964	964
Delivery	193	193
Harvest Commitment	$ 7,542	$7,542
Estimated Total Commitment	$19,620	$12,718
1957 # 7000 Expense	- 3,405	
Cost Plus activities	- 2,542	
Total Subject to Range	$13,673	
50% Range	6,837	
Estimated Net Cost		$ 6,902
Estimated Maximum Net Cost		$13,739

Above includes $5,321,122 of direct transistor cost

CES:jdk
9/5/57

Figure IV-2

OUR PRESENT COMMITMENT TO THE AEC

1	500 CPM Card Reader
1	155 CPM Card Punch
1	1,000 LPM Wire Printer
4	Type 727 Tape Units
1	Typewriter and console
1	1,048,576 Word, 4 Microsecond Transmission Rate, Disc Memory
16	0.2 Microsecond registers
2	512 Word 1/2 Microsecond Memories
4	8,192 Word 2 Microsecond Memories
1	Exchange
1	Input/Output Computer Frame*
1	High Performance Computer Frame*

For delivery in May 1960

Fixed Price $4,300,000

*IBM is planning that these two items will become one computer as defined by B+S.

Figure IV-3

HARVEST PROPOSAL TO BUSHIPS

1	250 CPM Card Reader
1	100 CPM Card Punch
1	500 LPM Wire Printer
4	729-I Tape Units
1	Inquiry Station
2	1024 Word 1/2 Microsecond Memories
4	16, 384 Word 2.0 Microsecond Memories
1	Basic Computer Frame
1	Harvest Supplementary Frame

(High performance tape units to be quoted on later this year)

For delivery late in 1960

If sold at cost $8, 785, 000

6/20/57

V. Market Analysis

There has been no official market analysis, because of lack of specifications and cost figures. The large scientific and engineering computing area has been examined by Mr. B. L. Sarahan in April 1957 on a customer-by-customer basis. An estimated rental of $200,000 per month (sales price of $10,000,000) was used. The size of the B+S market was thought to be between 11 and 29 systems over a 5 year period. The average of 20 systems has been used as an initial estimate. The dominating characteristic of this market is a need for high computational speed. These customers would insist on a reasonable price, but there probably would be little change in the market potential if rentals were increased by as much as 25%.

The market potential for the B machine alone as a commercial data processing system is largely dependent upon price. There is no question that the internal processing speed of the B machine is considerably faster than the 705 Model 3. If the minimum configuration of the B system (main frame, exchange, bus system, 16,384 words of 2.0 microsecond memory, a combination of ten 729 Model 3 and Model 1 tape drives, high-speed printer, high-speed card reader, and high-speed punch) should have a rental of $60,000 per month, the system would have an adequate market potential as an advancement for 705 and 709 systems. A systems estimate might be on the order of 125 systems over a 5-year period assuming an average rental of $80,000 per month. This is again an unofficial estimate.

Market Analysis personnel are just completing a qualitative evaluation of the market potential in this area. The results of this survey should be available in September. Preliminary reports indicate field interest in a commercial system in the monthly rental range of $60,000 to $85,000. The system is expected to be two to three times the internal arithmetic speed of the 705 with heavy emphasis on very large memories and with a very high combined effective character rate for the associated input-output equipment. A lesser interest in a system 50-200 times the speed of present systems is shown.

VI. The Data Processing Line 1957-1962

In Figure VI-1, an attempt has been made to show the magnetic tape processing equipment line as it exists in 1957. The line is projected to 1962. An assumption is made that IBM will be required for competitive reasons to supersede all existing equipment by transistorized or some other form of solid-state machines. It is further assumed, and more significantly, that there will be a decrease in the cost per operation to the customer.

The upper portion of the figure shows the price ranges of announced equipment (650, 701, 702, 704, 705, and 709) and soon to be announced equipment (660 and 705 Model 3). It can be seen that the 660 already indicates an area where the 705 Model 1 and Model 2 performance can be duplicated at lesser cost to the customer. The 705 Model 3 will also guarantee lower job costs, although system costs will be somewhat higher. It is shown that when all equipment is announced, that the price range from $10,000 to $80,000 per month will be covered by at least one IBM system. In some price areas, several different systems will be available. This is a flexible arrangement certainly, but it is probably costly to IBM and somewhat confusing to our customers and salesmen. In some cases, specialized functions will eliminate the seeming duplication as only one machine will be applicable to that given situation.

In an effort to show the relation of the STRETCH program to our future business, an idealized product line for 1962 is given in the lower portion of the figure. Four machines cover a wider rental and performance range than eight machines cover today. The goals of the 750 program are aimed at the Intermediate Data Processing goals. The specifications of this program are briefly: A cost of $15,000 to $20,000 per month for a minimum system; performance of main frame equivalent to 705 Model 3; suitable provision for use of main frame as computing element for peripheral in-line input-output equipment; and provision for the latest types of input-output equipment. The basic rental of $15,000 per month should be capable of being magnified a factor of 4 times through increased storage and increased numbers of input-output units. This system must be capable of handling both commercial and scientific computing problems in an efficient, economical manner. It is noted that a performance similar to the 705 Model 3 must be achieved at roughly half the cost of the system.

The Basic STRETCH computer has essentially the same characteristics as the Intermediate machine, except that the rental of a minimum system should be $60,000 per month. Its performance must be at least ten times the performance of the 705 Model 3, with an increase in rental of only a factor of two. Performance on computing problems should be at least ten times the performance of the 704. This system must be capable of handling both commercial and scientific problems in a simple straight-forward manner.

A rental expansion of a factor of at least three should be possible through increased memory and increased number of input-output units. There need be no program compatibility between the intermediate and Basic Stretch systems. Every advantage in the intermediate area must be directed to lower costs.

Compatibility between the Basic and the High Speed Stretch system is probably necessary from the way the system is derived. This is felt to be of some advantage but if it tends to slow down the high speed configuration, complete compatibility could be sacrificed. The immediate goal of the high speed system is highest possible computational performance. The system should have 100 times the performance of the 704 system. Cost-wise this system could have a rental of five times that of the 704. Volume memories and high-speed input-output devices must be a part of this advanced system.

One area remains and that is at the low end of the scale. The Intermediate Data Processing system has price-wise covered a good part of the existing low performance area. In line with the lower cost of operation, it seems essential to do the same thing for the $5,000 to $15,000 rental area. This goal could be accomplished by a new system or by reducing the price of the 650 tape system.

In a very optimistic way, one may consider the immense problem with which we are dealing. From this chart, assume the following:

System	No. of Systems	Average Rental	Extended Rental
Small	3,000	$10,000	$30,000,000
Intermediate	1,000	40,000	40,000,000
Basic Stretch	125	80,000	10,000,000
High Speed Stretch	20	200,000	4,000,000
	Total Monthly Rental		84,000,000
	Total Yearly Rental		$1,008,000,000

An extrapolation from the Market Analysis Product Line Forecast (May 1957) indicates that this total number of points is possible during the first half of 1964. Categories F (650 All Systems) and G (DP, 700 series and RAMAC exclusive of 650) were combined in making the extrapolation. An allowance was made for the RAMAC systems.

The manufacturing and financial problems connected with such a program are staggering. The figures have, of course, only been assumed, but the problems involved, if the transistor emerges as a low-cost unit, are fairly graphically illustrated. These figures are meant to be for the 1962-1967 area.

It should also be noted that the area of greatest revenue is again (as now) in the lower cost areas. The phenomena of the diminishing number of customers with increased cost is probably realistic. This simply indicates that the intermediate and small data processing areas are extremely important to IBM's future. The STRETCH area (particularly the high speed goals) is important to IBM from the prestige point of view.

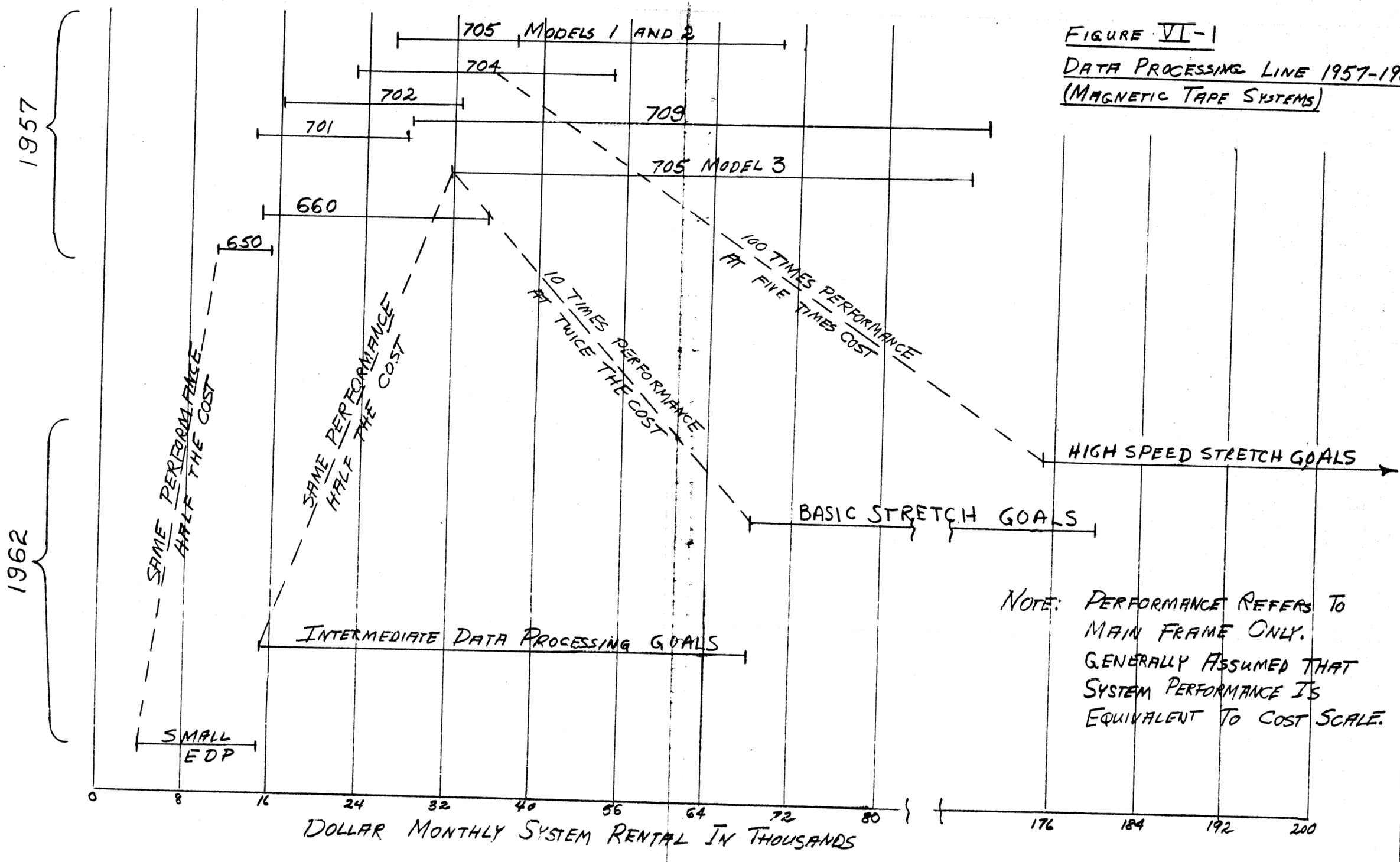
The chart and the numbers just presented also indicate the need for maintaining a very close price objective on these several systems. Engineering man power and money are critical. If the Basic Stretch system should miss by a factor of two in rental, a fifth machine program would be necessary to cover the area between the intermediate and Basic. Even worse, the number of potential systems might drop to the point where the STRETCH program would be seriously questioned, if both Basic and High-Speed were affected.

These ideas will be expanded further and checked against other future estimates to verify accuracy. The importance of a planned line of equipment and the inter-relation of this equipment has been stressed. Programs exist for obtaining the Intermediate, and Stretch systems; no program has been formalized to obtain the small data processing system.

FIGURE VI-1

DATA PROCESSING LINE 1957-1962

(MAGNETIC TAPE SYSTEMS)



NOTE: PERFORMANCE REFERS TO MAIN FRAME ONLY. GENERALLY ASSUMED THAT SYSTEM PERFORMANCE IS EQUIVALENT TO COST SCALE.

VII. System Characteristics

The problem of attempting to consolidate the three distinct classes of applications into one is illustrated by the following lists of system characteristics. These items were used by the joint Engineering - Product Planning group to arrive at the initial determination of the three systems. Sets of characteristics are given for each machine system.

"B" System Characteristics (Commercial problems only)

1. Data is highly structured in hierarchies.
   a. Basic unit is the byte, a collection of bits representing a character, digit, or control datum.
   b. Fields are composed of variable numbers of bytes, all of one size.
   c. Records are groups of fields.
   d. A file is a collection of all records of a given type.
2. Arithmetic is performed on fields of integers, often unsigned, and the ratio of maltiplications to additions is comparatively low.
3. On the average, few arithmetic operations are performed per item brought into memory.
4. Non-arithmetic operations of importance include editing for input-output formats, code conversion, tabular transformations, and decision operations.
5. Taken together, the non-arithmetic operations outnumber arithmetic operations in frequency of execution.
6. The amount of information processed at a time, including programs and frequently consulted tables, currently ranges from ten to five hundred thousand bytes.
7. The amount of information which must be available for consultation within short periods ranges from a few hundred thousand to several million bytes.

8. Indexing is important primarily for relative addressing rather than array scanning.

9. No universal and well-defined symbolism exists for describing problems, and few useful macro-operations are known.

10. Most machine time is used on a few different programs. These are usually run many times with frequent minor changes. These usually are changes in the problem and not in its method of solution.

11. Most programmers are full-time but relatively unsophisticated.

12. Sorting and file maintenance are generally performed upon long records with short key fields.

13. File consultation is not well-ordered; therefore sequentially arranged files have low activity rates. This low activity applies both to records within the file and to fields within the record.

14. Satisfaction of random inquiries is necessary; problems may be viewed as slow - scale real-time.

15. Problems are usually ill-defined enough so that the intervention of human judgment at some point is necessary.

"B" and "B+S" System Characteristics (Technical problems only)

1. The basic data units are signed numbers which may vary over a wide range in precision and a very wide range in magnitude. For problem control, fields of varying size consisting of integers and logical operands are usually introduced.

2. Arithmetic is usually performed on signed floating - point words, and the ratio of multiplications to additions is comparatively high.

3. On the average, many arithmetic operations are performed per item brought into memory.

4. Non-arithmetic operations are present but to a lesser degree than in commercial requirements. Logical operations facilitate complex program control.

5. The arithmetic operations outnumber the non-arithmetic operations in frequency of execution.

6. The amount of information processed at a time currently ranges from four to two hundred thousand words.

7. Large amounts of information are commonly held in readiness for sequential consultation.

8. Indexing is important for both array-scanning and relative addressing, especially subroutine control.

9. Usually there exists a universal and well-defined symbolism for describing problems, and many useful macro-operations are known.

10. Much machine time is used on short runs of many different problems. These are usually run only a few times between changes which arise as experiments in the method of solving a problem.

11. Most programmers are part-time users of the machine as a tool for their investigations. However, these programmers are usually sophisticated and competent.

12. Sorting and maintenance of long-term files is of secondary importance.

13. File consultation is well-ordered and activity rates are high (usually 100%)

14. Random inquiry during a calculation is rare.

15. Problems are presently so specified that human intervention is performed between machine runs. More sophisticated operation in the future will probably involve more intervention.

"B+H" System Characteristics (BuShips only)

1. The basic data unit is the byte; integer counts represent a second unit. Bytes are presented for operation in well-defined hierarchical patterns.

2. Arithmetic is performed on counts; modular arithmetic and logical operations are performed on individual bytes in long sequences.

3. On the average, many operations are performed per item brought into memory.

4. Besides editing and code conversion for input-output, particularly facile tabular transformations are required.

5. Byte-wise operations considerably outnumber arithmetic operations upon fields.

6. The amount of information processed at a time is of the same order as in commercial problems.

7. Large information capacity is needed for sequential consultation.

8. Indexing is important for relative addressing and especially for array scanning.

9. Few useful macro-operations are known.

10. Long and short, frequently and infrequently changing problems all occur.

11. Programmers are fairly sophisticated.

12. Sorting is generally performed on short records with long keys.

13. File activity rates are high.

14. Satisfaction of random inquiries during calculation is rare.

15. Intervention of human judgment usually occurs between machine runs.

16. Special conditions must be able to interrupt byte-wise operations.

17. Simulation of counter banks is necessary.

18. The accumulation of sums and counts in parallel with byte-wise operations is very desirable.

Section VIII - Product Planning Evaluation

Evaluation of B+S As stated earlier, the primary objective of this machine is the solution of large scale scientific problems at a rate 100 times that of the 704. This is consistent with Los Alamos requirements. If we assume that Product Development can meet the speeds statdd in the Los Alamos contract (i. e. arithmetic operations of 0. 6 usec floating and, 1. 2 usec floating multiply, and 1. 8 usec floating divide; and memory speeds of 0. 5 usec and 2. 0 usec) then, we believe that B+S will be about 75 times as fast as the 704.

The performance of the entire system is very dependent upon organization, e. g. Stretch has memories only 6 times and 24 times as fast as the 704, yet we expect 75 times 704 speed. Any major slip or series of minor slips could seriously hamper the performance. Simulation programs have been written and as Product Development releases more accurate times, these will be evaluated against typical programmed examples to gauge performance.

Although the original goals may not be completely reached, we believe that the above speeds will be considered as satisfactory both by Los Alamos and the large scale scientific computing market. The extent of the market is determined largely by the following factors:

a. Major reduction in job cost - the airframe companies, for example, have many large problems that they do not even try to solve now on 700 series simply because the cost is prohibitively high. A reduction in job cost by a factor of ten may be attained with B+S and this will immediately bring a whole new class of problems within range of economical attack.

b. Reduced net elapsed time - some classes of problems are excluded from 700 series today not because of cost but because the time required to obtain solutions attenuates the value of results. Frequently, faster methods, such as special purpose analog devices, are available. The B+S system will clearly reduce the computation time for all problems, over the 704. However, it is not clear that the net elapsed time will be greatly shortened for small problems.

c. Improved reliability - If performance objectives of 30 hours mean free error time are realized, a new class of real time control problems will be within reach of B+S.

Evaluation of B Assuming an average operation time of 4 usec for B, Product Planning has compared the 705-3, 709, and B machines with respect to a file maintenance problem (as described in Technical Report #1). The 729-3 tapes with a rate of 60,000 characters per second were assumed. As expected, the example proved to be completely tape limited, with about 2/3 of the compute time available for other uses. The computing speeds of B were 18 times as fast as the 705, 13 times as fast as the 705-3, and 9 times as fast as the 709.

As a commercial data processor, the B machine appears to have little advantage over present 700 series, other than the computing speeds mentioned above. Other factors, such as number of instructions required, number of information bits needed, and ease of programming seem roughly equivalent to the 705-3 and 709. It is essential that we proceed with parallel development of suitable auxiliary memory devices for B to afford proper utilization of its increased compute speeds. Until such units are available, B would probably be marketed mainly for those applications which are strongly compute limited on the 705-3 and 709. Multiprogramming might be a partial answer to the unbalance referred to above. It should not be considered so until at least one 700 series installation has demonstrated its practicality.

Conclusion and Recommendations We believe that the separation of functions represented by B and S is the best known solution to our need for an extremely fast scientific machine for Los Alamos and similar customers as well as our ultimate requirements for an extension beyond the 700 series speeds. While it is realized that the performance of the B+S configuration may suffer and that B as well as B+S will require more components than would have been needed if separate systems had been built, nevertheless the compromise still appears worthwhile in view of the savings that Product Development has stated must accrue from such an integrated design effort.

We are continuing to improve the accuracy of our estimates, both with respect to performance and costs. No matter how reasonable the present course of action appears, we must allow for the possibility of error by remaining flexible. We should not commit ourselves contractually to deliver any exactly specified system which we subsequently will release to production. From this standpoint, the proposed BuShips contract must be as vague on computer details as the Los Alamos contract.