

1. Classificação INPE-COM. 3/NTE C.D.U.: 681.3.00		2. Período	4. Critério de Distribuição: interna ☐ externa ☒	
3. Palavras Chaves (selecionadas pelo autor) DEMANDA (COMPUTADORES) PROCESSAMENTO DE DADOS				
5. Relatório nº INPE-1364-NTE/134	6. Data Outubro, 1978		7. Revisado por <i>Jorge de Mesquita</i> Jorge de Mesquita	
8. Título e Sub-Título DETERMINAÇÃO DA DEMANDA DE PROCESSAMENTO DE DADOS DO INPE DE 1978 A 1984			9. Autorizado por <i>Nelson de J. Parada</i> Nelson de J. Parada Diretor	
10. Setor DPD	Código 30.470		11. Nº de cópias 23	
12. Autoria Arry Carlos Buss Filho Iberê Lúcio Ronchetti Teixeira Ubirajara Moura de Freitas			14. Nº de páginas 35	
13. Assinatura Responsável <i>Iberê Lúcio Ronchetti Teixeira</i>			15. Preço	
16. Sumário/Notas Estudam-se as características de funcionamento do centro de processamento de dados do INPE, levantando-se os parâmetros de utilização atual e futura, considerando o crescimento vegetativo dos projetos e programas ora em curso, bem como do impacto da criação de novos projetos. Determinada a demanda, sugere-se uma configuração de equipamentos que a atenda.				
17. Observações				

ÍNDICE

SUMÁRIO	iv
CAPÍTULO I	
CARACTERÍSTICAS DE FUNCIONAMENTO	1
CAPÍTULO II	
UTILIZAÇÃO ATUAL	7
CAPÍTULO III	
ESTIMATIVA DE CRESCIMENTO	17
CAPÍTULO IV	
REQUISITOS DA CONFIGURAÇÃO BÁSICA	29

SUMÁRIO

O objetivo deste trabalho é determinar as características principais que a configuração de equipamentos de processamento de dados do INPE deve possuir, para atender a demanda prevista para os próximos anos. Para atingir este objetivo foi feita uma análise cuidadosa da demanda atual, inferindo-se a demanda futura a partir da previsão de crescimento dos atuais projetos do Instituto, da criação e implantação de novos projetos, e da historicidade de utilização do recurso de P.D. nos anos anteriores.

Divide-se o trabalho em quatro capítulos. No primeiro, faz-se rápido histórico das atividades de P.D. no INPE, e apresenta-se a característica de funcionamento da Divisão de Processamento de Dados (DPD). Este aspecto é discutido em profundidade no capítulo II, que apresenta diversos valores numéricos relativos à situação presente, ficando patente a saturação da configuração atualmente em uso. Passa-se, a seguir, no capítulo III, à estimativa de crescimento de uso tanto pelo atendimento da demanda hoje reprimida quanto pelo surgimento de novas necessidades, oriundas do crescimento vegetativo das atividades atuais e da criação de novas atividades. Finalmente, conhecendo-se em detalhe o volume de serviços, discute-se a maneira de atendê-lo.

CAPÍTULO I

CARACTERÍSTICAS DE FUNCIONAMENTO

O início das atividades de processamento de dados no INPE remonta ao ano de 1964, quando era utilizado o computador IBM 1620 do ITA.

Em fevereiro de 1965 foi instalado o primeiro computador no INPE, um IBM 650, emprestado pela IBM do Brasil, que permaneceu em funcionamento até meados de 1968. Era inicialmente um sistema bastante precário, consistindo apenas de uma leitora/perfuradora de cartões, à velocidade de 100 cpm, uma memória de tambor de 2000 palavras e um processador. As listagens, perfuração dos cartões e o processamento de alguns programas continuaram a ser feitos no ITA até fins de 1965, quando foram instaladas uma perfuradora de cartões IBM 029 e uma listadora contabilizadora IBM 407.

Em 1967 foi feita uma concorrência para a compra de um computador, sendo vencedora a proposta da Burroughs. O B3500, cujas operações se iniciaram em fevereiro de 1968, consistia inicialmente de um processador com 50 KBytes de memória, um disco magnético (tipo fixo) com 2 MBytes, quatro fitas magnéticas lentas (cluster), uma impressora de 1040 lpm, uma leitora de cartões de 800 cpm e uma perfuradora de cartões de 100 cpm. Posteriormente, nos anos de 1969, 1970 e 1971, foram feitas várias expansões, segundo a crescente necessidade de processamento, adicionando-se ao equipamento existente 90 KBytes de memória, 30 MBytes de disco (tipo fixo), quatro unidades de fitas magnéticas rápidas, uma leitora de fita de papel de 1000 cps, duas impressoras de 1100 lpm e uma "plotter" Calcomp 565.

Em 1970 foi adquirido um computador híbrido EAI 590, para a simulação de sistemas contínuos. Sua instalação ocorreu em março de 1971, com a configuração que permanece até hoje, composta de um computador digital EAI 640 com 16 KBytes de memória e um teletipo, um

computador analógico EAI 580 com 16 integradores e uma interface para a ligação dos dois computadores.

Em 1973 foi alugado o núcleo central de um sistema B6700, incorporando-se ao parque de máquinas um processador B6711, 400 KBytes de memória, 60 MBytes de disco (tipo fixo), dois terminais de vídeo (para uso exclusivo da operação) e duas leitoras de cartões de 1400 cpm cada.

Periféricos até então associados ao B-3500 foram transportados para o B-6700 resultando na configuração final relacionada na Tabela I.1.

TABELA I.1
CONFIGURAÇÃO ATUAL DO EQUIPAMENTO PRINCIPAL DA DPD

- processador central B6711
- 400 KB de memória de núcelos
- 2 terminais de vídeo (para uso dos operadores)
- 60 MB em disco fixo
- 4 fitas magnéticas (800 bpi, 72 KB/seg)
- 2 leitoras de cartão (1400 cartões/min)
- 2 impressoras (1.100 linhas/min)
- 1 leitora de fita de papel
- 1 perfuradora de cartão
- "plotter" Calcomp 565 ("off-line")

O B3500, equipado com o restante da periferia, foi transferido, em setembro de 1974, para a cidade de Natal, onde se encontra atualmente.

Quando da instalação do B6700, o centro funcionava durante o expediente normal, isto é, das 8 às 18 horas; eventualmente es

tendia-se este horário até as 22 horas. Com a crescente demanda, passou-se a operar normalmente até às 22 horas em 1975, até às 24 horas em 1976 e, a partir de fevereiro de 1977, adotou-se o funcionamento contínuo, das 7 horas de segunda-feira até às 13 horas de sábado, havendo, ainda, uma disponibilidade de 42 horas, referentes ao período não aproveitado do fim de semana. Esta parada explica-se por problemas administrativos, relativos ao pessoal, e pela falta de usuários, que se fazem presentes apenas durante o expediente normal do Instituto. Ultimamente, entretanto, a demanda tem aumentado de tal forma que soluções para os problemas administrativos estão sendo ativamente procuradas, e pretende-se, ainda este ano, estender o funcionamento do centro a, senão todo, pelo menos parte do fim-de-semana.

O centro opera exclusivamente em "batch", na modalidade de "closed-shop", tendo sido criadas, para disciplinar o acesso ao computador, quatro classes de serviços, com uma correspondente fila de entrada, devendo o usuário caracterizar seu "job", ao submetê-lo para execução, em pequeno, médio, grande ou rotineiro.

Os programas pequenos utilizam pouco recurso: leitora de cartões, impressora e no máximo 90 segundos de processador e 120 segundos de E/S. Para estes programas existem uma leitora e uma impressora, instaladas em um ambiente isolado do restante do sistema, operadas diretamente pelos próprios usuários, e que funcionam, de segunda a sexta-feira, em três períodos de 3,5 horas cada.

Os programas médios podem utilizar todos os tipos de periféricos mas são limitados a um máximo de 8 minutos de processador. Eles são executados sob o controle da operação, de forma a equalizar a utilização dos recursos, de segunda a sexta-feira, em dois períodos de 1,5 horas diárias.

Os programas grandes podem utilizar todos os tipos de periféricos, consumindo mais de 8 minutos de processador. Estes programas são executados diariamente em um período de 4 horas.

Os programas rotineiros estão subordinados a uma esca

la de serviços e posterior controle de qualidade, não sofrendo nenhuma restrição quanto aos recursos utilizados. Para eles estão destinadas 6,5 horas, diariamente.

A Tabela I.2 mostra a distribuição dos horários de funcionamento das diversas classes de serviço, que não se superpõem.

TABELA I.2
DISTRIBUIÇÃO DOS HORÁRIOS DE ATENDIMENTO
DAS CLASSES DE SERVIÇOS

PERÍODO		CLASSE DE SERVIÇOS
INÍCIO	TÉRMINO	
08:00	11:30	Pequenos
11:30	12:30	Rotineiros
12:30	14:00	Médios
14:00	17:30	Pequenos
17:30	19:00	Médios
19:00	22:30	Pequenos
22:30	04:00	Rotineiros
04:00	08:00	Grandes

Outra característica importante da instalação é estar ela voltada principalmente para o atendimento das necessidades dos projetos de pesquisa do INPE, resultando num grande número de usuários, os quais, em sua grande maioria, são os responsáveis diretos pelo desenvolvimento de seus programas. Estas tarefas, que se dividem entre as classes pequena, média e grande, consomem 73% do tempo útil de funcionamento do computador, cabendo às compilações uma parte considerável desta carga.

Os demais 27% do tempo são destinados aos sistemas técnico-administrativos, desenvolvidos e operados pela DPD. Destes, apenas 6% representam tarefas exclusivamente administrativas.

Atualmente o desempenho do sistema está limitado pelas capacidades da memória e do disco. O sistema está sendo mantido em funcionamento efetivo durante mais de 500 horas mensais, sem ociosidade, existindo ainda uma subutilização da CPU, em virtude da pequena capacidade de memória, que força a utilização muito intensa do mecanismo de memória virtual.

CAPÍTULO II UTILIZAÇÃO ATUAL

Neste capítulo estudam-se com detalhe as características das classes de serviço apresentadas no capítulo anterior. Em seguida, os resultados obtidos para cada classe são combinados, chegando-se ao quadro total de utilização do B6700.

A classe dos programas pequenos se caracteriza por compilações e execuções de curta duração, absorvendo a quase totalidade do esforço de desenvolvimento da instalação.

A partir dos dados do "log", isto é, os dados de contabilização gerados pelo próprio sistema, foi construída a figura II.1 (distribuição do tempo "elapsed" esperado), de onde se obteve o valor de 2,5 minutos, como duração média de um "job" pequeno.

Realizando uma amostragem sobre o comprimento da fila de programas pequenos, foi construída a figura II.2, que mostra o comprimento médio da fila em função da hora do dia, ao qual se tentou aplicar a teoria de filas; no entanto o comportamento encontrado não se aproximava dos modelos conhecidos, pois existe um fenômeno de saturação proveniente do fechamento da fila. Os usuários só continuam a chegar à fila até o momento em que o tempo de espera é menor que o tempo que falta para terminar o expediente, gerando uma demanda não atendida. Para contornar este problema, desprezou-se o comportamento da fila nos momentos que antecedem o seu fechamento, considerando somente o período de 8 horas às 10:30 horas e de 14 horas às 16 horas. Também não foi considerado o período noturno, visto que se trata de um uso provocado por uma parte da demanda reprimida durante o dia. Tendo em mente as restrições mencionadas obteve-se uma taxa de chegada de 30 "jobs"/hora.

Como o tempo médio de um "job" é de 2,5 minutos, serão atendidos somente 24 "jobs"/hora, ou seja, a taxa de chegada é maior

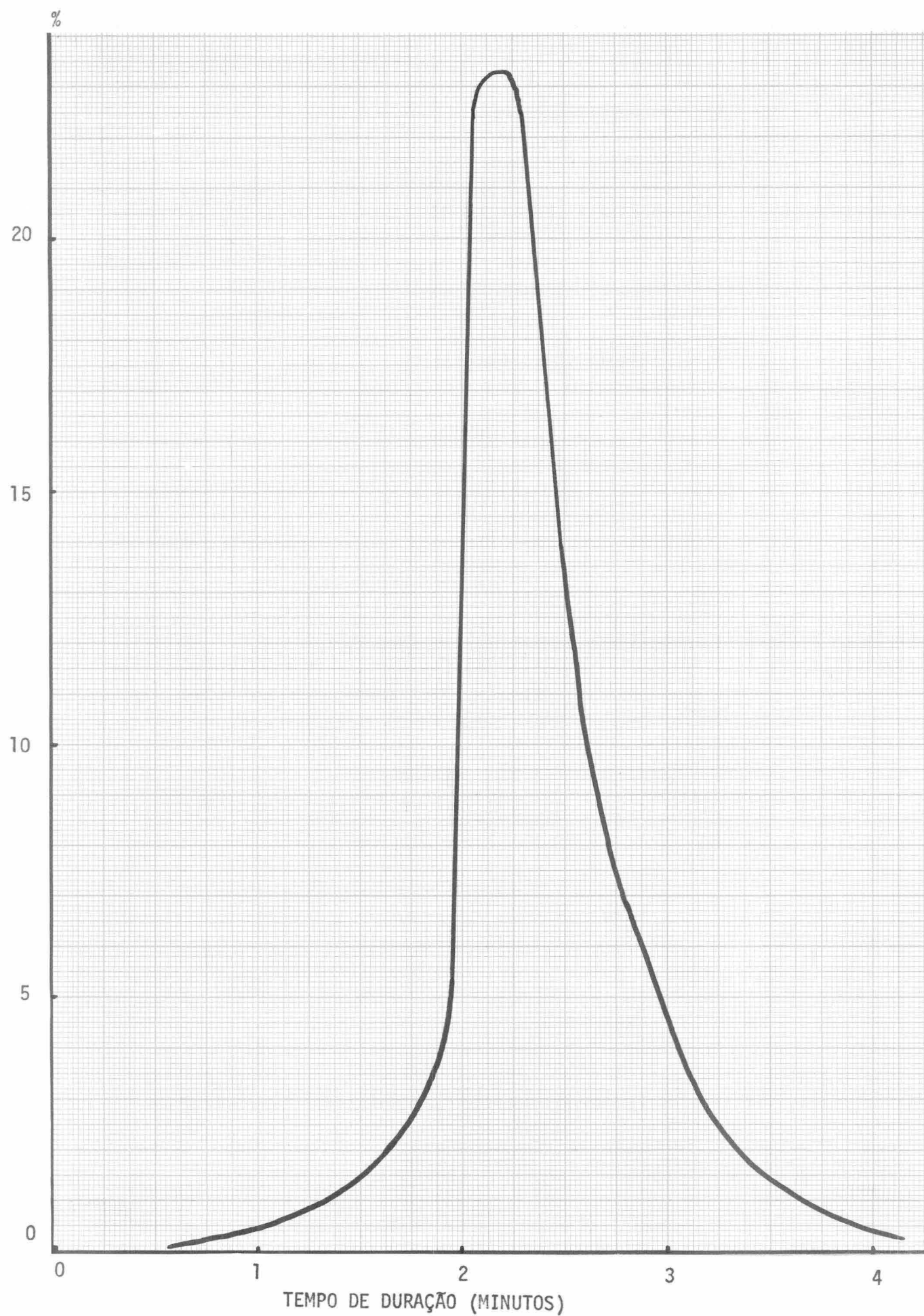


Fig.II.1 - Distribuição da duração dos programas pequenos.

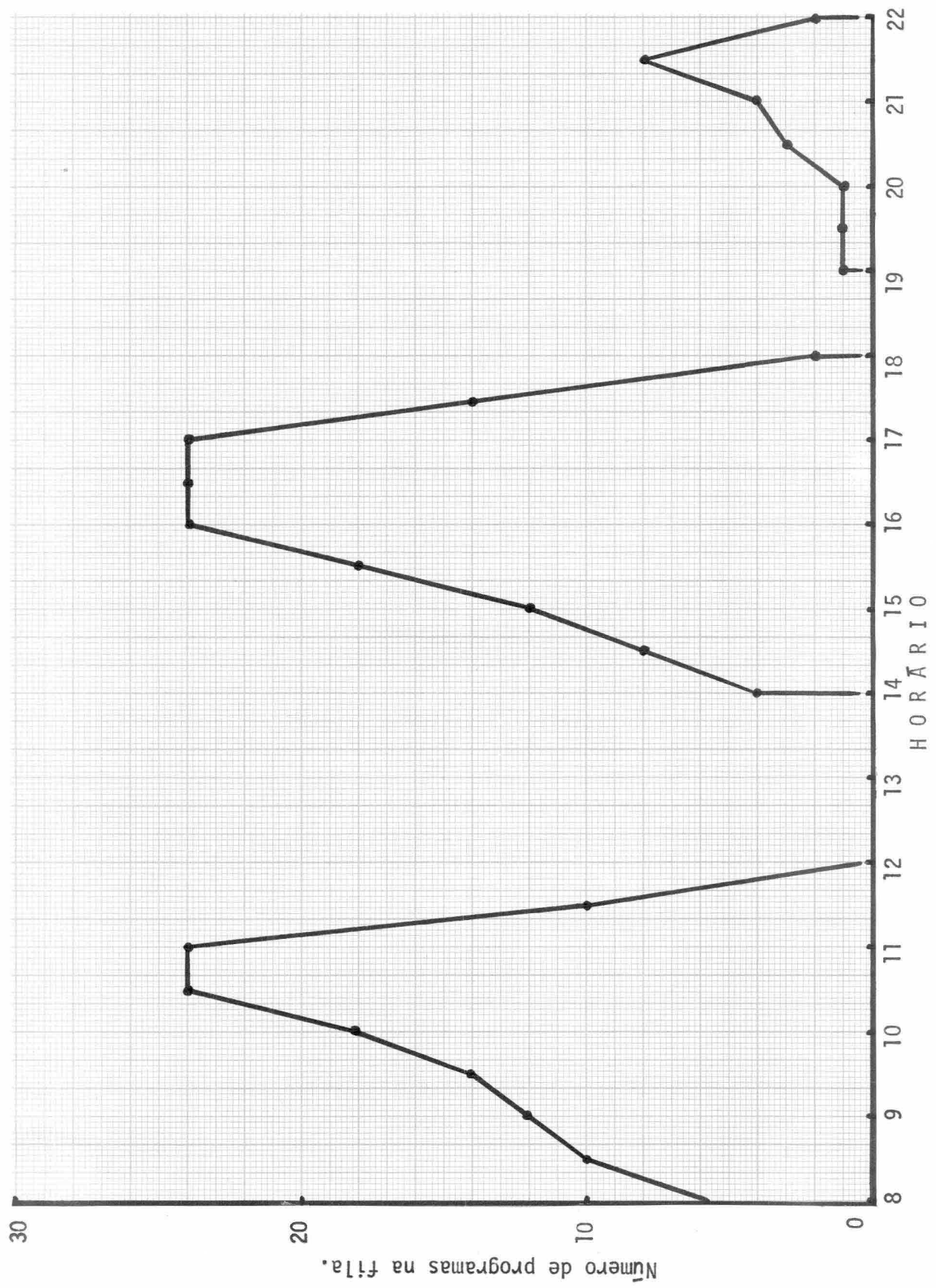


Fig.II.2 - Comprimento da fila de programas pequenos.

que a taxa de atendimento e, conseqüentemente, a fila tende a crescer indefinidamente, fato que realmente acontece até o instante do fechamento da fila. Em conseqüência, o tempo de espera é em geral superior a 60 minutos.

A classe dos programas médios se caracteriza pela execução de programas que necessitam outros periféricos além de leitora e impressora, bem como uma utilização média de processador. Além dos serviços normalmente destinados a esta categoria, aqui se incluem também aqueles programas pequenos que, por necessitarem de mais recursos do que os disponíveis na sua classe, notadamente fitas magnéticas, são obrigados a se transformarem em programas médios. Em conseqüência, os valores médios dos "jobs" que rodam nesta categoria estão bem abaixo dos valores máximos permitidos, ao contrário dos pequenos, que utilizam, em média, porcentagem bem maior do limite permitido.

Devido ao número de usuários desta classe, e ao tempo destinado ao processamento deste tipo de programa, consegue-se que a espera seja inferior a 12 horas.

Os programas grandes se caracterizam pela grande utilização de processador e média utilização de outros recursos, podendo algumas vezes necessitar também de uma grande quantidade de memória. Quando esta última característica é verificada, os programas chegam a permanecer na máquina até 7 vezes o tempo de CPU usado, pois devido à baixa capacidade de memória do sistema, em vez do programa consumir o recurso de processador em taxa elevada, passa-se a ter o fenômeno de "trashing" causado pelo uso intenso do mecanismo de memória virtual.

Para se determinar a frequência de utilização de um usuário desta classe, levantou-se, com precisão, o tipo de serviço executado, e encontraram-se as seguintes características:

- . Um programa utiliza quase que exclusivamente o recurso de processador.
- . Um pedido de execução normalmente é acompanhado da compilação

do programa.

- . O programa não é executado mais de uma vez por dia.
- . O mesmo programa é executado um número grande de vezes.

Para estes serviços foi feito um levantamento sobre uma amostra representativa (cerca de 50%) dos usuários, determinando-se para cada um, os valores do tempo de espera de uma tarefa, do tempo de análise dos resultados, do número de execuções mensais e do tempo médio de processador. Estes dados estão mostrados na Tabela II.1, tirando-se dali que o tempo médio de espera de uma tarefa é de 48 horas, podendo em alguns casos, atingir 140 horas.

A classe dos programas rotineiros se caracteriza pela execução dos programas técnico-administrativos sujeitos a uma programação rígida, devendo ser entregues em dias previstos. Por exemplo, as previsões de passagem de um satélite devem estar processadas antes de sua passagem. Por este motivo, estes usuários são atendidos de forma preferencial.

Para quantificar a demanda de utilização dos diversos tipos de serviços, qualificados anteriormente, levantaram-se, com dados obtidos do "Log", os seguintes parâmetros: tempo de processador, tempo de E/S, cartões lidos e linhas impressas. A Tabela II.2 caracteriza o perfil de uso da instalação, relacionando o uso dos diversos recursos em um mês médio de 1978. Não foi incluído, neste levantamento, o uso médio de memória, porque este recurso está sendo usado em toda a sua capacidade, e os valores medidos são distorcidos pelo mecanismo de memória virtual.

A partir das informações acima se podem deduzir as características médias de uma tarefa, que são apresentadas na Tabela II.3.

TABELA II.1
LEVANTAMENTO DAS CARACTERÍSTICAS DE UTILIZAÇÃO
DOS USUÁRIOS DE PROGRAMAS GRANDES

Número do usuário	Número médio de execuções por mês	Tempo médio de espera (dias)	Tempo médio de análise de resultados (dias)	Tempo médio de processamento de uma tarefa (min)
1	6,59	1,8	2,8	24
2	1,00	2,0	30,0	-
3	1,00	3,0	30,0	-
4	9,23	2,3	1,0	28
5	1,00	4,0	30,0	-
6	3,06	1,8	8,0	24
7	3,19	4,0	5,4	24
8	2,00	1,2	14,0	40
9	4,62	2,0	4,5	41
10	1,00	3,0	30,0	83
11	0,83	6,0	30,0	42
12	14,63	1,8	0,2	31
13	11,67	2,0	0,6	28
14	6,29	2,2	2,6	21
15	11,24	2,7	0,0	14
16	0,86	5,0	30,0	22
17	3,75	4,0	4,0	28
18	1,46	5,5	15,0	41
19	1,00	5,0	30,0	-
20	8,22	1,4	2,2	21
21	7,50	3,0	1,0	41
22	6,59	1,8	2,7	42
23	2,94	4,2	6,0	10
24	8,82	2,2	1,2	41
25	4,62	3,5	3,0	21

NOTA: Amostra de 25 usuários num universo de 52 usuários.

TABELA II.2
PERFIL DE UTILIZAÇÃO MENSAL MÉDIA DE 1978

Tipo de tarefa		Nº de tarefas	Tempo de CPU(horas)	Tempo de E/S(horas)	Cartões lidos (mil)	Linhas impressas (mil)
Compilações	PEQUENA	5.280	5,87	9,30	1.115,2	1.346,4
	MÉDIA	440	0,47	0,73	90,2	110,0
	GRANDE	132	0,36	0,58	70,4	83,6
Execuções	PEQUENA	4.400	19,80	3,00	160,5	963,6
	MÉDIA	660	10,52	7,75	125,3	836,0
	GRANDE	132	47,64	6,64	4,4	158,4
	ROTINEIRA	4.180	40,34	77,00	484,0	3.630,0
Total da Instalação		15.224	125,00	105,00	2.050,0	7.128,0

TABELA II.3
CARACTERÍSTICA MÉDIA DE UMA TAREFA

Tipo de tarefa		Tempo de CPU (minutos)	Tempo de IO (minutos)	Cartões lidos	Linhas impressas
Compilações	PEQUENA	0,07	0,10	211	255
	MÉDIA	0,07	0,10	205	250
	GRANDE	0,16	0,27	533	633
Execuções	PEQUENA	0,27	0,04	36	219
	MÉDIA	0,96	0,71	190	1.266
	GRANDE	21,65	3,03	33	1.200
	ROTINEIRA	0,58	1,10	116	868

Analisando as características de utilização da atual instalação conclui-se que o tempo de processador constitui o melhor parâmetro para avaliar a demanda. Os demais fatores possuem comportamento complexo, dependendo não só das características dos programas como também das restrições de equipamento.

Além dos fatores apresentados, uma análise retrospectiva do uso de processador, apresentada na Tabela II.4, indica, que apesar dos esforços para aumentar o tempo de utilização de processador, este aumento só se verificou após um aumento proporcional do tempo de funcionamento. O motivo da baixa taxa de utilização (de ordem de 23% do tempo total) é justificado pela quantidade de memória existente, confirmando a utilização de processador e memória em toda a sua capacidade permitida. Fenômeno semelhante não é verificado para os demais parâmetros, caracterizando uma utilização não crítica.

Em vista disto levantou-se, de maneira detalhada, o perfil de utilização de processador da instalação, tabela II.5, que servirá de base para os estudos posteriores.

TABELA II.4
COMPARAÇÃO ENTRE O TEMPO DE FUNCIONAMENTO E UTILIZAÇÃO DE PROCESSADOR

Ano	Total de horas de funcionamento	Total de horas de processador utilizadas	Taxa de utilização
74	230	43	18.7
75	322	72	22.4
76	435	101	23.2
77	552	127	23.0
78 *	552	127	23.0

* Valor estimado considerando o primeiro semestre de 1978.

TABELA II.5
PERFIL MENSAL DE UTILIZAÇÃO DE PROCESSADOR PARA 1978

Tipo de tarefa	Número de tarefas	Tempo de CPU (horas)
<u>C O M P I L A Ç Õ E S</u>	5.852	6,70
PEQUENA	5.280	5,87
Fortran	2.574	2,57
Algol	2.112	2,20
Cobol	484	0,73
Outras	110	0,37
MÉDIA	440	0,47
Fortran	198	0,18
Algol	198	0,22
Cobol	44	0,07
GRANDE	132	0,36
Fortran	132	0,36
<u>E X E C U Ç Õ E S</u>	9.372	114,30
PEQUENA	4.400	19,80
Fortran	2.134	9,60
Algol	1.760	7,96
Cobol	396	1,76
Outras	110	0,48
MÉDIA	660	10,52
Fortran	308	5,24
Algol	308	5,24
Cobol	44	0,04
GRANDE	132	47,64
Fortran	132	47,64
ROTINEIRA	4.180	40,34
Algol	1.650	22,73
Cobol	2.530	17,61
TOTAL:	15.224	125,00

CAPÍTULO III

ESTIMATIVA DE CRESCIMENTO

Até o momento sô foi mostrada uma fotografia da situação atual do centro de processamento, caracterizando os diversos tipos de usuários e a utilização dos recursos existentes. Pode-se perceber, em função do exposto, a existência de uma demanda não atendida, para determinadas classes de serviços.

Para estimar a demanda até 1984, seguiram-se as seguintes etapas:

- . Avaliação da demanda não atendida em 1978.
- . Estimativa de crescimento vegetativo, considerando os valores históricos existentes.
- . Previsão de uso de grandes projetos em implantação.

Todas as estimativas foram baseadas no processador B6711 atualmente instalado no INPE, cuja velocidade foi avaliada em 0,5 MIPS (MIPS= 10^6 instruções por segundo).

Do estudo da demanda não atendida, resulta uma necessidade de 120 horas mensais de processador, obtidas da forma descrita a seguir.

Para os programas pequenos foram feitos dois levantamentos: o primeiro baseado no "log" do sistema e o segundo através de entrevistas com os próprios usuários. O objetivo era obter o número de execuções conseguidas e o número de execuções desejadas de uma mesma tarefa por dia, sendo encontrados os valores médios de 2,5 e 5,5 respectivamente. Estes resultados indicam uma demanda de 2,2 vezes o uso real da máquina. Usando este valor e a utilização atual de processador, para programas pequenos, extraída da Tabela II.2, encontrou-se uma demanda não atendida de 30 horas de processador.

Para os programas médios, o atendimento atual foi considerado satisfatório, e, por este motivo, concluiu-se não existir uma demanda não atendida.

Para os programas grandes, foram feitas as seguintes suposições:

- . Os programas seriam recebidos, processados e devolvidos no máximo em 24 horas.
- . O tempo de análise dos resultados, por parte dos usuários, seria mantido igual a média atual observada.
- . O tempo médio de processador por programa permaneceria constante.

Em consequência das condições anteriores e da Tabela II.1, foi possível determinar o número de execuções mensais desejadas por usuário e, por diferença, determinar a demanda não atendida de processador. Estes valores são apresentados na Tabela III.1, de onde se obteve uma demanda não atendida de 63 horas mensais.

Para os programas rotineiros, não existe uma demanda não atendida, pela forma preferencial com que são processados.

Pelo exposto anteriormente pode-se verificar a existência de dificuldades no tocante ao desenvolvimento de novos programas e à execução de programas que exigem grandes quantidades de tempo de processador. Estas conclusões confirmam a insatisfação externada por muitos usuários. Em consequência, existem tarefas cujo desenvolvimento é considerado inviável no contexto atual. Supondo a inexistência das limitações, estas tarefas seriam imediatamente iniciadas, e o seu volume de utilização é estimado em 27 horas, ou seja, cerca de 30%, em relação a demanda não atendida (93 horas).

A partir destes valores pode-se compor o perfil de utilização desejado para 1978, apresentado na Tabela III.2, e que representa um uso total de 245 horas mensais de processador.

TABELA III.1
DEMANDA NÃO ATENDIDA DE PROGRAMAS GRANDES

Número do usuário	Tempo médio de processamento de uma tarefa (minutos)	Nº médio de execuções por mês	Nº médio de execuções desejadas por mês	Demanda não atendida de processador (minutos)
1	24	6,59	8,00	34
2	-	1,00	1,00	0
3	-	1,00	1,00	0
4	28	9,23	15,00	162
5	-	1,00	1,00	0
6	24	3,06	3,33	7
7	24	3,19	4,69	36
8	40	2,00	2,00	0
9	41	4,62	5,45	34
10	83	1,00	1,00	0
11	42	0,83	0,47	6
12	31	14,63	24,00	291
13	28	11,67	19,11	208
14	21	6,29	8,33	43
15	14	11,24	30,00	262
16	22	0,86	0,97	2
17	28	3,75	6,00	63
18	41	1,46	1,88	17
19	-	1,00	1,00	0
20	21	8,22	9,23	21
21	41	7,50	15,00	307
22	42	6,59	8,00	60
23	10	2,94	4,29	14
24	41	8,82	13,64	198
25	21	4,62	7,50	60
			SOMA:	1.825

NOTA: Amostra de 25 usuários num universo de 52 usuários.

Demanda não atendida da amostra = 1.825 minutos.

Demanda não atendida total = 3.796 minutos.

TABELA III.2
PERFIL MENSAL DE UTILIZAÇÃO REQUERIDA PARA 1978

Tipo de tarefa	Número de tarefas	Tempo de CPU (horas)
<u>C O M P I L A Ç Õ E S</u>	14.096	16,23
PEQUENA	13.301	14,79
Fortran	6.484	6,47
Algol	5.321	5,54
Cobol	1.219	1,84
Outras	277	0,94
MEDIA	440	0,47
Fortran	198	0,18
Algol	198	0,22
Cobol	44	0,07
GRANDE	355	0,97
Fortran	334	0,91
Algol	21	0,06
<u>E X E C U Ç Õ E S</u>	16.280	228,77
PEQUENA	11.085	49,88
Fortran	5.376	24,19
Algol	4.434	20,05
Cobol	998	4,43
Outras	277	1,21
MEDIA	660	10,52
Fortran	308	5,24
Algol	308	5,24
Cobol	44	0,04
GRANDE	355	128,03
Fortran	334	120,50
Algol	21	7,53
ROTINEIRA	4.180	40,34
Algol	1.650	22,73
Cobol	2.530	17,61
TOTAL:	30.376	245,00

A estimativa de demanda, para 1984, levando em consideração apenas o crescimento vegetativo é de 716 horas, obtida da seguinte forma: foi admitido que, para os próximos anos, o crescimento seria linear e igual ao verificado de 1974 a 1978, considerando a demanda total requerida, tabela III.2. Deste modo a utilização de processador cresce de 51 horas por ano, alcançando o valor de 551 horas em 1984. A este valor, adicionou-se uma parcela de 165 horas (30% de 551), correspondente ao aumento da demanda relacionada com expansão prevista de pessoal da área de pesquisa.

A estimativa de uso para os projetos em desenvolvimento, que utilizam o processador de forma intensiva, é de 344 horas mensais, obtidas da forma descrita a seguir.

O Departamento de Meteorologia está desenvolvendo projetos para previsão numérica de tempo e modelagem atmosférica. O primeiro já se encontra em fase de testes, devendo estar operacional a partir de 1979. A previsão de tempo para um período de 24 horas, utiliza em média, 4 horas de processador, e a demanda total é estimada em 34 horas mensais.

O projeto de modelagem atmosférica será iniciado em 1979 e estima-se que em 1982 entrará em operação um modelo de complexidade intermediária que utilizará, em média, 10 horas de processador para simular 24 horas, resultando numa demanda mensal de 140 horas.

As estimativas acima se referem apenas ao desenvolvimento, testes e aprimoramento dos modelos. A utilização destes sistemas para previsão diária de tempo e simulação de grandes períodos (1 a 2 anos) exige computadores muito rápidos. Alguns centros, que possuem aplicações nesta área, operam computadores com velocidades superiores a 100 MIPS.

A fim de ilustrar o crescimento das aplicações nesta área, cita-se o NMC (National Meteorological Center - USA) que aumentou a sua capacidade instalada em 100 vezes, num período de 10 anos,

o que corresponde a uma taxa anual de crescimento de aproximadamente 58%.

O Departamento de Sistemas Espaciais, através do grupo de dinâmica orbital, está desenvolvendo sistemas relacionados ao projeto de lançamento de um satélite nacional.

O controle de um ou mais satélites exige um conjunto de computadores de pequeno porte, para controle das missões em tempo real, e o apoio de um sistema de grande porte, para a determinação de posição, previsão de órbitas e, dependendo do tipo de satélite, processamento das informações recebidas.

Atualmente, para atender às necessidades do Centro de Produção de Imagens, foi desenvolvido um sistema para a previsão de órbitas dos satélites da série Landsat. Baseado na experiência adquirida pode-se dizer que 24 horas de previsão de órbita utilizam entre 7 e 100 minutos de processador, em função do modelo geopotencial e intervalo de integração adotados.

Considerando que serão controlados diversos satélites, utilizando modelos de complexidade diferentes, e que a confiabilidade dos sistemas exige um contínuo desenvolvimento e teste de programas, é estimada uma demanda de 120 horas mensais de processador, para 1984.

O controle de um satélite doméstico aumenta consideravelmente os requisitos de processamento, exigindo a execução imediata dos programas relacionados ao controle da missão, bem como a necessidade de facilidades de "backup". Como ilustração, cita-se o sistema canadense, que dispõe de um computador de uso compartilhado, com capacidade de processamento de aproximadamente 1.15 MIPS, para controle de seus dois satélites de comunicação domésticos. Durante as fases críticas da missão o sistema é usado sob regime de alta prioridade, interrompendo os demais serviços. Como "backup" estão associados ao sistema dois computadores idênticos, ligados através de linhas telefônicas.

O Departamento de Sensoriamento Remoto, para os projetos voltados ao processamento de imagens, necessita também uma utilização intensiva de processador. Para este fim o Instituto dispõe de equipamentos especializados, a saber: o GE I-100 e o Bendix MDAS. Estes sistemas possuem características de "*hardware*" que aceleram o processamento das imagens. No entanto, existem alguns métodos que não utilizam estas características, fazendo uso apenas da capacidade de processamento convencional dos sistemas PDP-11/45. Neste caso enquadra-se a classificação de imagens por máxima verossimilhança. Utilizando este método está, atualmente, em desenvolvimento, um projeto para classificação de culturas de cana de açúcar, no estado de São Paulo. Posteriormente este projeto deverá ser estendido para a classificação sistemática de outras culturas em uma região maior, implicando em um volume de informação 50 vezes superior ao tratado pelo projeto piloto, e que exigirá a utilização de uma máquina de grande porte.

A classificação de uma foto do Landsat (512 x 512 pontos) consome, em média, 4 minutos de processador, e, em vista disto, espera-se uma demanda total de 50 horas mensais quando o projeto estiver totalmente implantado.

A partir das estimativas apresentadas anteriormente, estabeleceu-se o uso total de processador em 1060 horas mensais, em 1984, resultando no perfil de utilização mostrado na Tabela III.3.

A demanda de 1060 horas de processador também pode ser obtida aplicando-se uma taxa constante de 27% ao ano sobre a demanda requerida em 1978. Esta estimativa pode ser considerada como bastante conservadora em virtude da sistemática de avaliação adotada, e será considerada como a demanda mínima.

Por outro lado analisando o crescimento verificado entre a demanda real em 1974 e a demanda requerida em 1978, encontra-se uma taxa anual média de 50%. Admitindo-se a manutenção desta taxa, a demanda em 1984 seria de 2870 horas mensais de processador. Este valor será considerado como um limitante superior, neste estudo. A utilização

TABELA III.3
PERFIL DE UTILIZAÇÃO MENSAL ESTIMADA PARA 1984

Tipo de tarefa	Número de tarefas	Tempo de CPU (horas)
<u>C O M P I L A Ç Õ E S</u>	42.110	50,0
PEQUENA	38.860	43,0
Fortran	18.950	18,9
Algol	15.550	16,1
Cobol	3.560	5,4
Outras	800	2,6
MÉDIA	1.270	1,4
Fortran	570	0,6
Algol	570	0,6
Cobol	130	0,2
GRANDE	1.980	5,6
Fortran	1.860	5,2
Algol	120	0,4
<u>E X E C U Ç Õ E S</u>	48.510	1.010,00
PEQUENA	32.400	146,0
Fortran	15.710	70,8
Algol	12.960	58,6
Cobol	2.920	13,0
Outras	810	3,6
MÉDIA	1.930	31,0
Fortran	900	15,4
Algol	900	15,4
Cobol	130	0,2
GRANDE	1.980	715,00
Fortran	1.860	671,0
Algol	120	44,0
ROTINEIRA	12.200	118,0
Algol	4.800	66,0
Cobol	7.400	52,0
TOTAL:	90.620	1.060,00

ção mensal pode aproximar-se deste valor limite, por exemplo, se for necessária a utilização, em base sistemática, dos programas de modelagem atmosférica. Apenas este projeto pode requerer mais de 1000 horas mensais de processador.

A figura III.1 visualiza as taxas de crescimento mencionadas.

Até aqui, foram estimadas as necessidades em relação ao processador. Com base no levantamento apresentado na Tabela II.3 (característica média de uma tarefa) podem deduzir-se coeficientes de correlação entre o tempo de CPU e os demais parâmetros (tempo de E/S, cartões lidos e linhas impressas), apresentados na Tabela III.4. Admitir-se-á que as características médias dos programas não sofrerão grandes alterações ao longo do tempo. A análise dos valores conhecidos indica que a utilização total de processador tem crescido a uma taxa superior à verificada para os demais recursos. Desta forma os coeficientes de correlação atuais estimam a demanda futura máxima destes recursos.

De posse destas informações se pode estabelecer o perfil de utilização dos demais recursos para 1980 e 1984, como mostrado nas Tabelas III.5 e III.6. Deste perfil, verifica-se a existência de um grande número de compilações e execuções de programas, desenvolvidos por uma coletividade de usuários bastante numerosa. Para que este tipo de serviço possa ser executado com eficiência, surge a necessidade de descentralização dos meios de entrada e saída, como por exemplo a utilização de um sistema de "time-sharing". Em consequência parte da demanda de cartões e linhas será reduzida.

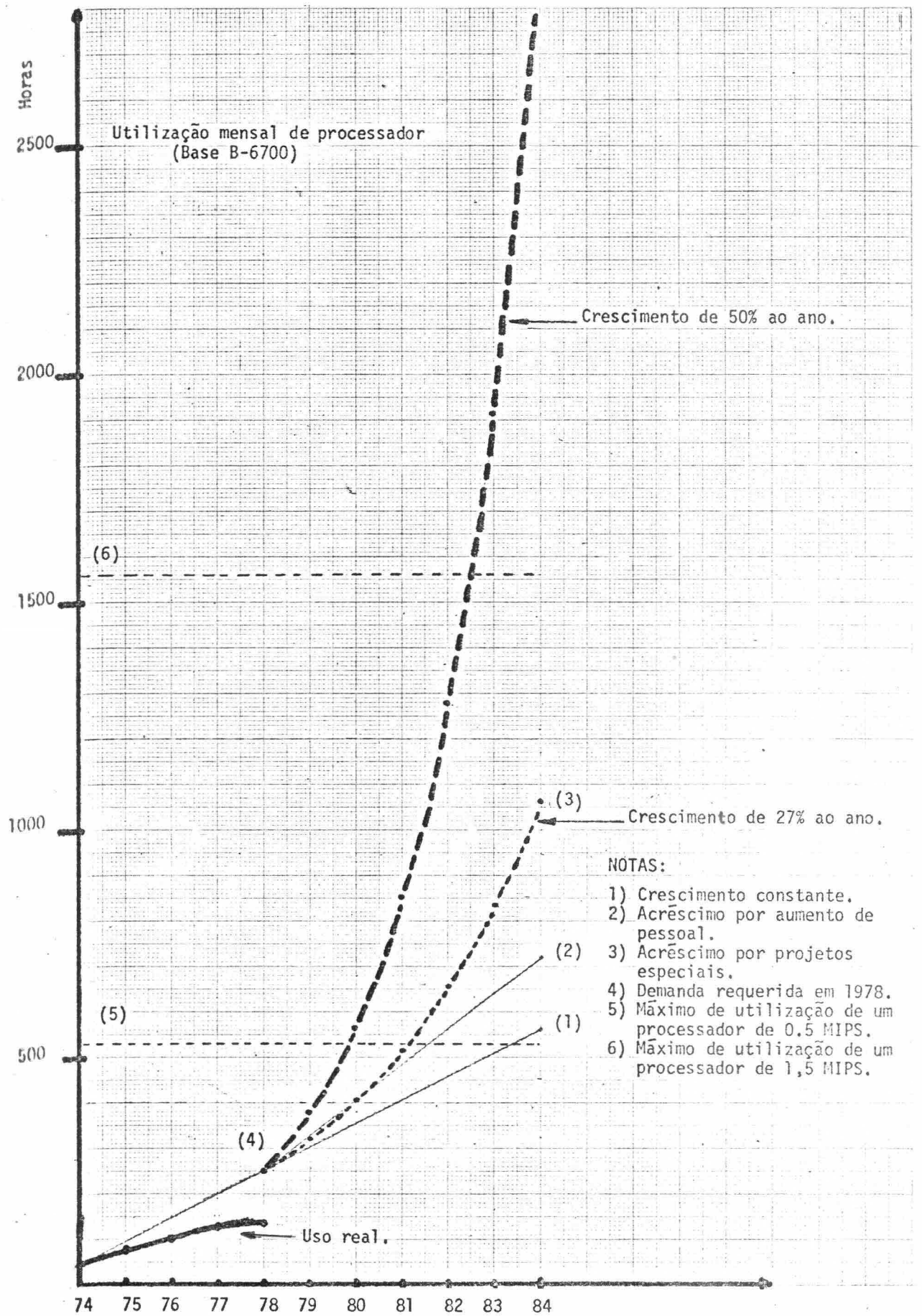


Fig.III.1 - Uso de processador (real e estimativa futura).

TABELA III.4
CORRELAÇÃO ENTRE OS DIVERSOS PARÂMETROS DE UMA
TAREFA E UM MINUTO DE TEMPO DE CPU

	Tipo de Tarefa	Tempo de I/O em minutos	Cartões Lidos	Cartões Impressos
Compilações	PEQUENA	1,43	3.014	3.642
	MÉDIA	1,43	2.928	3.571
	GRANDE	1,68	3.331	3.956
Execuções	PEQUENA	0,15	133	811
	MÉDIA	0,74	198	1.318
	GRANDE	0,14	152	55
	ROTINEIRA	1.89	200	1.496

TABELA III.5
PERFIL DE UTILIZAÇÃO MENSAL MÉDIA PARA 1980

	Tipo de tarefa	Numero de tarefas	Tempo de CPU(horas)	Tempo de E/S (horas)	Cartões lidos(mil)	Linhas im pressas(mil)
Compilações	PEQUENA	18.830	21,0	30,0	3.800	4.600
	MÉDIA	620	0,7	1,0	120	150
	GRANDE	500	1,4	2,4	280	330
Execuções	PEQUENA	15.700	71,0	10,6	570	3.450
	MÉDIA	930	14,9	11,0	180	1.180
	GRANDE	500	181,0	25,0	20	600
	ROTINEIRA	5.920	57,0	107,0	680	5.100
Total da Instalação		43.000	347,0	187,0	5.650	15.410

TABELA III.6
PERFIL DE UTILIZAÇÃO MENSAL MÉDIA PARA 1984

Tipo de tarefa		Número de tarefas	Tempo de CPU(horas)	Tempo de E/S(horas)	Cartões lidos (mil)	Linhas impressas (mil)
Compilações	PEQUENA	38.860	50,0	72,0	9.040	10.930
	MÉDIA	1.270	1,4	2,0	240	300
	GRANDE	1.980	5,6	9,3	1.100	1.300
Execuções	PEQUENA	32.400	146,0	22,0	1.170	7.100
	MÉDIA	1.930	31,0	23,0	370	2.500
	GRANDE	1.980	715,0	99,7	65	2.400
	ROTINEIRA	12.200	118,0	223,0	1.420	10.600
Total da Instalação		90.620	1.060,0	451,0	13.405	35.130

CAPÍTULO IV

REQUISITOS DA CONFIGURAÇÃO BÁSICA

Tendo sido determinada, com precisão, a demanda para os próximos anos, fazem-se, neste capítulo, considerações sobre a configuração que atenderá a esta demanda, sem levar em conta, entretanto, qualquer tipo particular de fabricante ou modelo. Este processo é conveniente para estimar o "*hardware*" mínimo, exceção feita à quantidade de memória final, que é muito dependente da estrutura dos sistemas operacionais e dos compiladores, variando muito, de computador para computador, o ponto ótimo para uma determinada carga.

Para determinar a velocidade de processador que atenda a demanda, estimou-se, inicialmente, o número máximo de horas mensais, que se pode extrair de uma CPU, como igual ao tempo em que o computador estiver ligado, menos o tempo consumido pelos "*overheads*", referentes a "*overlays*", "*spooling*", "*time-sharing*", acesso a banco de dados, balanceamento da carga de programas e outros.

As estatísticas da DPD indicam que a máquina não opera, por causa de manutenção, preventiva ou corretiva, ou por causa de falhas ambientais (defeitos no ar condicionado, falta de energia etc) durante 8% do mês, ou seja, 57 horas. Admitindo que os "*overhead*" atinjam 20%, eles consumirão 133 horas, o que leva ao valor teórico máximo de utilização do processador de 530 horas/mês. Este valor só poderá ser alcançado em um ambiente de multiprogramação equilibrado, que permita a utilização total do recurso de processador, por exemplo, a execução sistemática de um programa "*cpu-bound*" juntamente com os demais serviços da instalação.

Tomando os valores mínimos e máximos da demanda esperada, correspondentes as taxas de crescimento de 27% e 50%, pode-se compor um quadro de utilização percentual, em relação a um máximo de uso de 530 horas mensais, para processadores de diversas velocidades (Tabela IV.1).

TABELA IV.1
UTILIZAÇÃO PERCENTUAL DE PROCESSADOR

a) Demanda mínima

ANO	VELOCIDADE DO PROCESSADOR EM MIPS							
	0.5	1.0	1.5	2.0/2x1.0	3.0/2x1.5	6.0	10	15
78	48	-	-	-	-	-	-	-
80	77	39	26	19	12	-	-	-
82	125**	63	42	31	21	10	6	-
84	201**	101	67	50	34	17	10	7

**) Processador saturado (maior que 100%)

-) Processador sub-utilizado.

b) Demanda máxima

ANO	0.5	1.0	1.5	2.0	3.0	6.0	10	15
78	48	-	-	-	-	-	-	-
80	108**	54	36	27	18	-	-	-
82	238**	119**	79	60	39	20	12	-
84	542**	271**	178**	136**	89	45	27	18

Das inúmeras alternativas que podem ser adotadas, levando em consideração os dados da Tabela IV.1, destacam-se duas soluções principais:

- a) Instalação de um sistema com velocidade de 1,5 MIPS e que apresente facilidades de multiprocessamento, permitindo a instalação de um segundo processador por volta de 1984, quando o primeiro estaria alcançando o seu ponto de saturação.

- b) Instalação de um computador próprio com velocidade de 1.0 Mips e garantia de acesso a um outro sistema, bem mais veloz, que fosse capaz de absorver o excesso de demanda.

Devido ao grande número de usuários já existentes e ao crescimento previsto pelo aumento de pessoal, é imprescindível a descentralização dos meios de entrada, utilizando recursos de teleprocessamento. Entretanto, é pouco razoável estender a todas as pessoas que se utilizam do computador, o recurso do "*time-sharing*", pelo grande investimento necessário. Por outro lado, a maior eficiência decorrente do uso dessa ferramenta deve ser preferencialmente oferecida aos projetos prioritários do Instituto, que congregam cerca de 30% dos programadores em atividade. Os demais programadores, envolvidos em outros projetos de pesquisa, em atividades acadêmicas ou originários de entidades em convênio com o INPE, continuariam a se valer do "*batch*".

Desta forma, 30% da carga de desenvolvimento, composta por programas pequenos e médios serão absorvidos pela rede. No entanto, tomando por base a experiência de outras organizações, 10% do volume de entrada e saída destes serviços continuariam utilizando os meios convencionais, isto é, leitora de cartões e impressoras.

Mais uma vez, pelo fato de não se possuir dados próprios, recorre-se à experiência de outras entidades que já se utilizam de terminais remotos, para determinar o tempo de duração de uma sessão típica de desenvolvimento de programas. É estimado um tempo médio de 30 minutos para a edição, compilação e execução de um programa, resultando num máximo de 48 sessões diárias por terminal. No entanto, um terminal não será utilizado mais de 12 horas por dia, pela própria natureza do serviço, reduzindo o número de sessões diárias para 24, ou seja 48 tarefas (compilação ou execução).

Relacionando os valores obtidos, na Tabela III.5, com a demanda de utilização, configura-se uma rede inicial para 1980 de 12 terminais. De modo análogo, utilizando a Tabela III.6, determina-se um total de 24 terminais para 1984, sempre mantendo a relação de 30% da

carga originando-se nos terminais. Havendo recursos e interesse em descentralizar mais ainda a entrada de serviço, outros terminais podem ser agregados à rede; por exemplo, para ter 70% da carga oriunda de terminais, em 1984, necessitar-se-á de 55 terminais.

A demanda de cartões e linhas impressas, foi obtida levando em consideração a utilização da rede dimensionada e os valores dispostos nas Tabelas III.5 e III.6, e é apresentada na Tabela IV.2. Conclui-se que, para 1980, duas leitoras rápidas de cartões e duas impressoras rápidas serão suficientes. Já em 1984 o sistema deverá sofrer o acréscimo de uma leitora de cartões e de uma impressora, para atender à demanda dos programas grandes e rotineiros.

TABELA IV.2
DEMANDA DE LEITURA DE CARTÕES E LINHAS IMPRESSAS

ANO	Cartões por minuto		Linhas por minuto	
	Tarefas de desenvolvimento	Outras tarefas	Tarefas de desenvolvimento	Outras tarefas
1980	240	35	470	210
1984	545	90	960	490

A existência de "time-sharing" exige, por si, só uma utilização intensa de disco magnético, uma vez que os programas e arquivos devem residir neste meio. Atualmente é permitido aos usuários armazenar programas e arquivos em disco de uma forma controlada, em virtude da pequena capacidade existente (60 Mbytes). Esta forma de utilização, que atinge 20 Mbytes, representa uma parcela muito pequena da que realmente seria necessária. O restante das informações são armazenadas em fita magnética, ou algumas vezes em cartão, sendo transferidas para disco antes de sua utilização e removidos logo após o uso, causando um aumento considerável do "overhead" do sistema.

As dificuldades seriam ainda maiores, se as aplicações não fossem orientadas, quase que exclusivamente, para o uso de fitas magnéticas, mesmo quando a utilização de disco é mais aconselhável. Assim, as aplicações de consultas, que requerem um grande uso de disco e que tem sido evitadas, com a criação da rede de teleprocessamento, tornar-se-ão inadiáveis. Nesta classe podem citar-se os sistemas de recuperação de informação clássicos e os bancos de dados científicos. Estes últimos são mantidos em fita magnética, em virtude do grande volume de informação, permanecendo em disco apenas conjuntos de dados selecionados e as estruturas de indexação existente. Na Tabela IV.3 são apresentadas as previsões para discos magnéticos, relativas aos anos de 1980 e 1984.

TABELA IV.3
DEMANDA DE DISCO MAGNÉTICO

Tipo de uso		Demanda em MBytes	
		1980	1984
1	"Software" básico (sistema operacional, compiladores, utilitários, "spooling", "rollin-rollout").	50	50
2	Arquivos de trabalho	100	200
3	Programas e arquivos de usuários.	120	250
4	Aplicações "on-line"	30	100
	TOTAL:	300	600

Havendo disponibilidade de memória de massa, o uso de fitas magnéticas diminui substancialmente, sendo até mesmo admissível a permanência das quatro unidades atualmente em operação.

Por outro lado, a aquisição, em caracter permanente, de grandes massas de dados científicos traz consigo o problema de armazenamento dos volumes que os contem, cuja solução natural é o aumento da densidade de gravação, com a consequente compactação. Outro aspecto do mesmo problema é a compatibilidade entre os diversos equipamentos que o INPE opera, e que atualmente usam apenas a densidade de 800 bpi, compatibilidade essa que deve ser mantida.

Enfim, tem-se aqui, a exemplo do processador central, duas possibilidades:

- a) Caso as unidades atuais possam permanecer em atividade, acrêscimo de quatro unidades de 1600 bpi, ou, preferencialmente, acrêscimo de quatro unidades de 6200 bpi.
- b) Caso as unidades atuais devam ser retiradas de serviço, adição de quatro unidades de densidade dupla 800/1600 bpi, ou, preferencialmente, adição de quatro unidades de densidade tripla 800/1600/6200 bpi.

Resumindo, as características básicas da configuração que atende à demanda prevista para os próximos anos são as seguintes:

- Processador central com velocidade entre 1,0 e 1,5 Mips.
- Memória central, com ciclo de acesso abaixo de 1 μ seg, e capacidade aproximada de 1,5 MB.
- Possibilidade de instalação de um segundo processador central, transformando o sistema em dual.
- Capacidade de teleprocessamento, com possibilidade de ligação de grande número de terminais de diferentes características.
- Discos magnéticos com capacidade inicial de 300 MB.
- Fitas magnéticas com as densidades 800, 1600 e/ou 6200 bpi.
- Duas impressoras de pelo menos 1000 linhas/minuto.
- Duas leitoras de cartões de pelo menos 1000 cartões/minuto.
- Doze terminais de vídeo.