

Carla Osthoff Barros
 Laboratório de Física Experimental de Altas Energias/DRP
 Centro Brasileiro de Pesquisas Físicas - CBPF/CNPq
 Ministério da Ciência e Tecnologia
 Rua Dr. Xavier Sigaud, 150, 22290 - Rio de Janeiro, RJ

RESUMO

O Centro Brasileiro de Pesquisas Físicas, CBPF, em colaboração com o Fermi National Accelerator Laboratory, FERMILAB, está participando do desenvolvimento de sistemas de multiprocessadores paralelos de baixo custo efetivo capazes de satisfazer não só aos problemas de computação da Física de Altas Energias como também de outras áreas de pesquisa científica.

O CBPF possui no momento um sistema de 21 processadores e está desenvolvendo software básico e aplicativo em colaboração com outras instituições.

ABSTRACT

The Centro Brasileiro de Pesquisas Físicas in collaboration with Fermi National Laboratory is participating on the development of multiprocessors systems which is easy to use and low cost effective to solve many computational problems in High Energy Physics and in others fields.

CBPF in the moment has a system with 21 processors and is developing basic and application software for the ACP in collaboration with others institutes.

1. INTRODUÇÃO

A necessidade computacional em física de altas energias tem crescido tanto que computadores tradicionais não são mais capazes de satisfazer-la. Uma única experiência é capaz de produzir centenas de milhões de eventos onde cada evento necessita para o seu processamento de alguns segundos do tempo de uma cpu. O Fermi National Laboratory criou o Advanced Computer Program para o desenvolvimento de computadores capazes de suprir esta necessidade [2].

O primeiro sistema computacional desenvolvido, o ACP, tem como finalidade satisfazer problemas do tipo "evento orientado". Neste problema um mesmo processo é executado para a análise de cada uma dentre as centenas de milhões de eventos. Este tipo de processo é caracterizado por necessitar de muito tempo de cpu, de realizar poucas operações de "entrada e saída", e de efetuar pouca ou nenhuma interação entre os processadores.

Novos sistemas estão em fase de desenvolvimento possuindo uma arquitetura mais flexível capaz de apresentar alta performance para problemas onde existe a necessidade de uma maior interação entre os processadores.

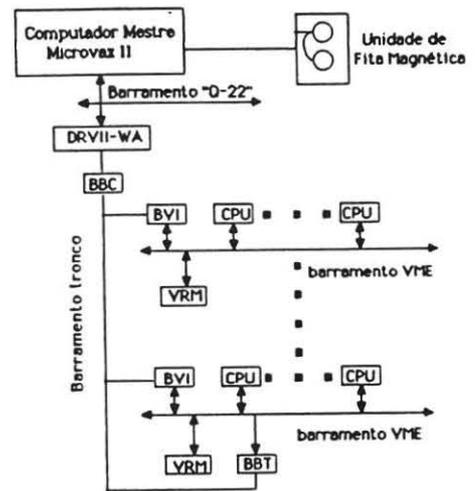
Tal como existe no presente, o sistema ACP [1] foi indicado pela "RESEARCH AND DEVELOPMENT MAGAZINE" para o prêmio R-100, dado aos 100 mais significantes produtos técnicos do ano de 1986.

2. ARQUITETURA DO SISTEMA ACP

O primeiro Sistema ACP desenvolvido pode ser classificado como processador MIMD, "MULTIPLE INSTRUCTION MULTIPLE DATA". Sua arquitetura e

do tipo "árvore", ver figura 1.

SISTEMA ACP
 Configuração Padrão



LEGENDA

BBC - Interface controladora do Barramento Tronco	CPU - Nó de processamento
BVI - Interface do Barramento tronco VME	BBT - Terminal do Barramento tronco
DRVII-WA - Unidade de DMA	VRM - Arbitro do Barramento VME

Figura 1.

Cada processador possui uma unidade de processamento e uma memória local. Cada processador pode acessar as memórias de todos os processadores do sistema excluindo a do hospedeiro. Os pro

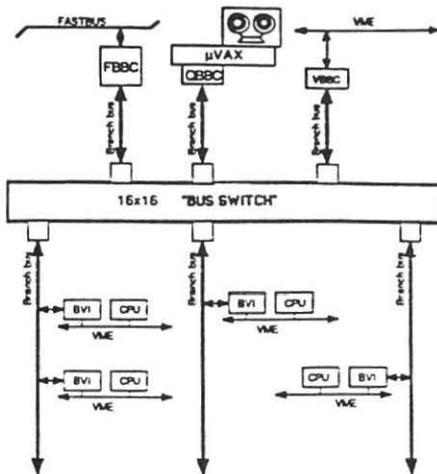
cessadores são interligados via bastidores de padrão VME capazes de suportar até 21 módulos.

Os bastidores VME são por sua vez interligados a um barramento padrão RS485, chamado BRANCH BUS, que é capaz de suportar até 31 bastidores.

Barramentos BRANCH podem também são interligados via um barramento chaveado, BUS SWITCH [4] que consiste de um bastidor capaz de interligar bi-direcionalmente até 16 barramentos BRANCH.

Existem dois módulos compatíveis com o padrão B.SWITCH atualmente em desenvolvimento: os módulos de interface "BUS SWITCH/BRANCH BUS" e os "FLOATING POINT ARRAY PROCESSOR" desenvolvidos para cálculos de "LATTICE GAUGE".[10]

Um barramento "BRANCH" é capaz de realizar transferência de dados em até 20Mbytes. Um sistema acima descrito possuindo 16 "Branch Buses" interligadas pelo "BUS SWITCH" é capaz de realizar transferência de dados em até 160 Mbytes, ver figura 2.



Interligação de diversas raízes utilizando o "BUS SWITCH". Este sistema pode suportar uma taxa de transferência de até 160 Mbytes.

Ref.-4

Figura 2.

Cada processador [5], figura 3, é composto por:

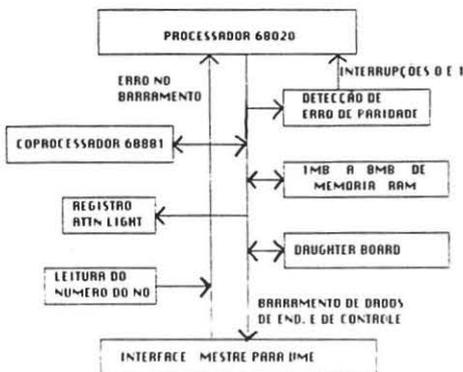


Figura 3.

- um microprocessador comercial de 32bits/16MHZ, MC68020 ou ATT32100, e seu respectivo coprocessador de ponto flutuante,

- 2Mbytes de memória real ram, "Random Access Memory", podendo ser expandida em até 8 Mbytes. A memória realiza ciclos de leitura a 240 ns e ciclos de escrita a 120 ns (0 wait states).

- interface para implementação de uma "daughter board" com 64k de memória não volátil PROM, "Programable Read Only Memory"; interface serial para padrão RS232, e/ou de outras necessidades.

- interface para o barramento VME do tipo "mestre/escravo". No modo escravo a memória do processador pode ser acessada por qualquer módulo no barramento VME. No modo mestre o processador tem acesso direto às memórias locais dos outros processadores do barramento. O sistema vê cada processador como parte de uma memória virtual global de 4 Gigabytes.

Está em fase de desenvolvimento um novo processador utilizando um microprocessador de arquitetura RISC, "Reduced Instruction set Computer", o microprocessador "MIPS". Este processador deverá apresentar uma performance dez vezes superior a performance dos atuais.

O sistema ACP é um sistema de processamento paralelo do tipo mestre/escravo. O microcomputador uVAX II ou o minicomputador VAX 11/780 são utilizados como seu hospedeiro.

O hospedeiro tem a função de gerenciar todo o acesso de "entrada/saída" dos processadores assim como outras necessidades do sistema. Cabe a cada processador apenas a tarefa de processar. Para isto, foi desenvolvido o sistema operacional LUNI, "LITTLE UNIX". O hospedeiro carrega o LUNI e o programa a ser executado pelo processador na fase de inicialização do sistema. O LUNI coloca a processador no estado "READY" esperando o carregamento do evento a ser processado, ver figura 4.



Estados assumidos pelo processador

Figura 4.

Ao receber o evento o processador passa para o estado "RUNNING" e executa o programa que foi carregado pelo hospedeiro. Ao término da exe

ção do programa o processador passa para o estado "DONE" e fica esperando a coleta dos resultados para só então voltar para o estado "READY". Caso ocorra uma exceção, o LUNI coloca o processador no estado "DEAD" e espera uma re inicialização.

Está em fase de desenvolvimento um novo sistema operacional compatível com o sistema operacional "UNIX" para ser carregado nos processadores do sistema ACP. Este sistema operacional tornará viável a substituição do hospedeiro por um dos processadores do sistema; tornando também possível a substituição do sistema "mestre/escravo" por um sistema operacional do tipo distribuído ou fluutuante.

3. SISTEMA OFF-LINE

Para um melhor performance [1], é aconselhável a implementação de um sistema de produção como na figura 5.

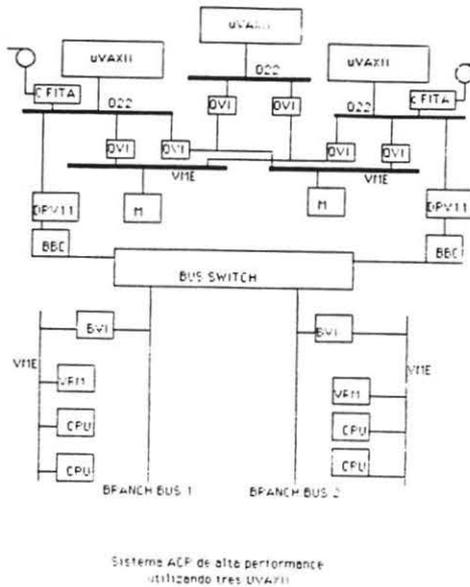


Figura 5

Nesta figura vemos um primeiro UVAX II dedicado apenas a leitura e ao carregamento dos eventos nos processadores. Um segundo uVAX II é dedicado a coleta dos dados do sistema ACP e a escrita dos dados em fita. Um terceiro uVAX II é dedicado ao gerenciamento dos dois, assim como cálculo de estatísticas, etc.

Podemos citar como exemplo de medida de performance de um sistema ACP off-line a implementação do programa de análise de dados da experiência de Fotoprodução de Charm, E-691, que teve a participação do CBPF. Para a análise dos eventos desta experiência foi utilizado um sistema ACP de 53 processadores capaz de processar em 1 mes 400 fitas de eventos; mesma quantidade de eventos processados durante 7 meses de utiliza

ção de 25% de capacidade do FERMILAB COMPUTER CENTER. A capacidade para esta análise era composta de 1 CYBER 850 e 4 CYBERS 175.

4. SISTEMA EM TEMPO REAL

A figura 6 apresenta o sistema de aquisição de dados da experiência E-769 na qual o CBPF é uma das instituições colaboradoras [6]. Dois módulos foram desenvolvidos para a implementação deste sistema: um controlador de bastidor CAMAC SCC ou "SMART CRATE CONTROLER" e o RBUFF ou "Read Buffer". O SCC realiza a coleta dos eventos que chegam no bastidor CAMAC segundo um programa pré-carregado pelo minicomputador VAX 11/780. O SCC irá transmitir os eventos para o RBUFF via uma interface paralela. Os RBUFFS são módulos de padrão VME e apresentam características de "buffer" de leitura. Os eventos são coletados dos "RBUFFS" pelos Nós de processamento. Após o processamento, os eventos são enviados para gravação em fita utilizando para isto um módulo padrão VME controlador de fita, o CIPRICO. O gerenciamento da coleta dos eventos e o processamento são realizados por um dos processadores do sistema ACP. A monitoração do sistema é realizada pelo VAX 11/780.

Tal sistema coletou uma média de 400 eventos por segundo de aproximadamente 4Kbytes cada, produzindo em 8 meses de coleta de dados mais de 10.000 fitas. A análise dos eventos desta experiência necessitaria de 12 anos de processamento no sistema atual do FERMILAB COMPUTER CENTER, mas será realizada em aproximadamente 12 meses em um sistema ACP "OFFLINE" de aproximadamente 100 processadores.

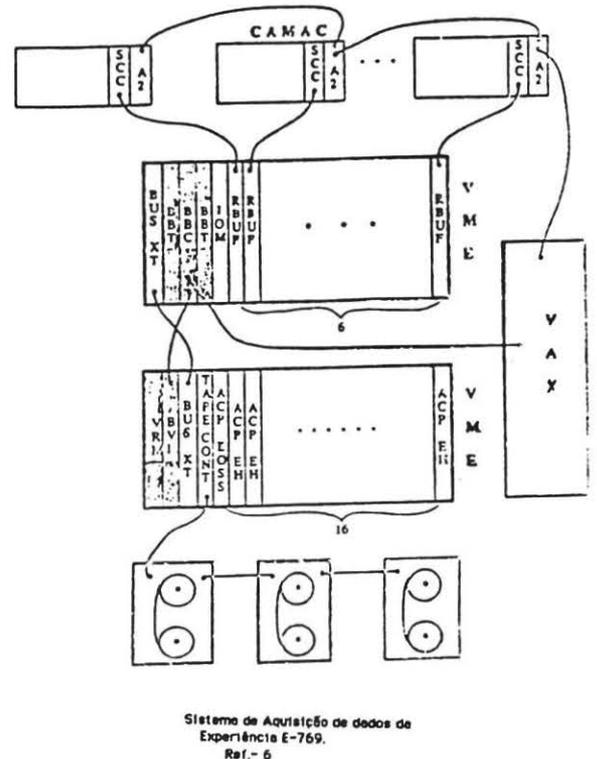


Figura 6.

Para a próxima experiência, o sistema de aquisição pretende desenvolver novos métodos capazes de possibilitar a coleta de 4800 eventos por segundo em uma primeira fase e de 20000 em uma segunda fase.

5. SOFTWARE

O "software" existente fornece suporte para problemas do tipo evento orientado [3]. Onde o usuário deve dividir o seu programa em dois programas distintos: o primeiro programa é processado no hospedeiro e realiza as operações de entrada e saída, estatísticas e etc. O segundo programa é executado nos processadores, e realiza o cálculo propriamente dito. Foram desenvolvidas subrotinas em FORTRAN, fáceis de serem implementadas, e que realizam toda a transmissão de eventos entre o uVAX II e os processadores. Como pode ser observado na figura 7. A linguagem FORTRAN foi escolhida por ser a linguagem padrão na comunidade científica.

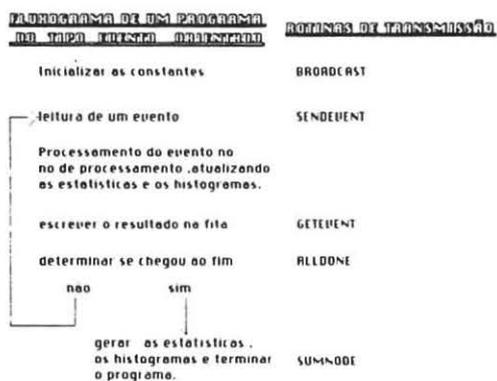


Figura 6.

O software desenvolvido para o ACP suporta toda a classe de problemas que requerem pouca ou nenhuma intercomunicação entre processadores. Nesta classe podemos incluir muitas aplicações de cálculo em redes, cálculo de integrais de muitas dimensões reconstrução da trajetória de partículas, bem como para problemas fora da física tais como cálculo de matrizes, processamento de sinais, etc.

Existe desenvolvido um simulador do sistema ACP que pode implementado em máquinas de ambiente VMS.

6. PARTICIPAÇÃO DO CBPF

O CBPF iniciou sua participação no projeto ACP durante o desenvolvimento do primeiro sistema ACP de 120 processadores, com a participação da engenheira Carla Barros.

A seguir, o CBPF enviou os engenheiros Bruno Schulze e Carla Barros para participarem da implementação do sistema de aquisição de dados da experiência E-769 e para realizar durante o mesmo período a construção de um sistema ACP "OFF-LINE" de 80 processadores para a experiên-

cia CDF, e um sistema ACP de 13 processadores para o CBPF. Este sistema está em funcionamento no CBPF e foi subsequentemente acrescido de mais 8 processadores.

Atualmente o CBPF mantém o engenheiro Mariano Miranda participando do desenvolvimento de um novo sistema operacional compatível com o sistema UNIX.

O CBPF está abrindo programas para o desenvolvimento de software em colaboração com outras instituições tais como Laboratório Nacional de Computação Científica, LNCC [7], Coordenação de Programas de Pós-graduação de Engenharia, COPPE/Universidade Federal do Rio de Janeiro [9] e Instituto de Física Teórica, IFT [8].

REFERÊNCIAS

- [1] Nash, T. et al., "The ACP Multiprocessor System at FERMILAB", Fermilab Conf.86/132.
- [2] Nash, T. et al., "The Fermilab Advanced Computer Program Multimicroprocessor Project", Proc.Conf.Computing in High Energy Physics, Amsterdam, June de 1985 (North-Holland, Amsterdam, 1986) and references therein.
- [3] Nash, T. et al., "Software for Event Oriented Processing on Multiprocessor Systems, Processor and software research for High-Energy Physics, Guanajuato, México (1974) p. 175.
- [4] Bus Switch Specification, BSS. R.ATAC/ACP/FERMILAB.
- [5] ACP 68020 CPU Module User's Manual. R. ATAC/ACP/FERMILAB.
- [6] E-769 Acquisition System. C.Gay/Toronto University.
- [7] Projeto MSPARALLEL. Projeto de colaboração entre LNCC/CNPq, Universidade Simon Bolivar, IM/UFRJ e CBPF/CNPq.
- [8] A colaboração com o IFT visa a pesquisa em algoritmos paralelos e aplicações à física e matemática.
- [9] A colaboração com a COPPE está desenvolvendo trabalhos de pesquisas em conjunto com os departamento de Sistemas de Computação, de Engenharia Elétrica, e Engenharia Civil.
- [10] Schulze, B. e Valois, R., "Segunda Geração de Processadores Paralelos do Advanced Computer Program.