

UNIVERSIDADE DE SÃO PAULO

**Sistema auxiliar de arquivos baseado em disco ótico
WORM
para ambiente computacional distribuído**

**LUIZ ANTONIO FRAMARTINO BEZERRA
REGINA HELENA CARLUCCI SANTANA
MARCOS JOSÉ SANTANA**

No. 21

N O T A S



Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

**Sistema auxiliar de arquivos baseado em disco ótico
WORM
para ambiente computacional distribuído**

**LUIZ ANTONIO FRAMARTINO BEZERRA
REGINA HELENA CARLUCCI SANTANA
MARCOS JOSÉ SANTANA**

No. 21

**NOTAS DO ICMSC
Série Computação**

**São Carlos
Dez./1995**

Sistema Auxiliar de Arquivos Baseado em Disco Ótico WORM Para Ambiente Computacional Distribuído

Luiz Antonio Framartino Bezerra
Dep. de Genética e Matemática Aplicada à Biologia - FMRP-USP
CEP 14049-900 Ribeirão Preto - SP
lafbezer@genbov.fmrp.usp.br

Regina Helena Carlucci Santana
Departamento de Computação - ICMSC-USP
Caixa Postal 668 - 13560-970 São Carlos-SP
rsc@icmssc.sc.usp.br

Marcos José Santana
Departamento de Computação - ICMSC-USP
Caixa Postal 668 - 13560-970 São Carlos-SP
mjs@icmssc.sc.usp.br

Resumo

Este trabalho discute o desenvolvimento e implementação de um sistema auxiliar de arquivos, utilizando disco ótico de tecnologia WORM. O objetivo é proporcionar aos servidores de arquivos de um ambiente computacional distribuído, um mecanismo para transferência de arquivos originalmente armazenados em discos magnéticos, para disco WORM. As idéias apresentadas são baseadas na análise de dados reais, coletados em sistemas em operação diária. O mecanismo de transferência de arquivos proposto é transparente ao usuário e tem como objetivos o melhor aproveitamento da mídia magnética disponível e a melhoria da segurança e confiabilidade do sistema. A técnica básica utilizada envolve replicação de arquivos importantes e a transferência para o disco ótico de arquivos com baixa probabilidade de acesso. O sistema foi projetado integralmente e testado através de implementação na linguagem C.

Abstract

This work discusses the design and implementation of an auxiliar file system, adopting optical disks (WORM). The aim is to implement mechanisms to file servers of distributed computing environments for file transfer from traditional magnetic disks to optical units. The ideas presented are completely based on the analysis of real data collected from existing systems during daily operation. The transfer mechanisms proposed are user-transparent and aim to improve both disk-space availability, security and reliability. The basic techniques adopted include the replication of important files and the transfer of files with low access probability, from magnetic disks to optical units. The system was fully developed and implemented for tests using the C programming language.

Introdução

A rápida evolução dos sistemas computacionais ao longo das últimas 4 décadas tem influenciado de modo significativo quase todas as áreas do conhecimento humano. A evolução trouxe os sistemas computacionais distribuídos onde, na maioria das instalações atuais, têm-se máquinas operando como servidores (prestadores de serviços) e máquinas funcionando como clientes (as estações de trabalho), o que caracteriza o modelo arquitetural das estações de trabalho/servidores [TAN92, COU94, MUL89].

Apesar de toda evolução e toda melhoria alcançada na qualidade do serviço oferecido pelos sistemas computacionais ao longo desses anos, ainda se tem pontos críticos que parecem sobreviver a qualquer tipo de evolução.

Qualquer análise dos sistemas computacionais ao longo das 4 décadas mostrará que um problema (aparentemente eterno) é a utilização do espaço de armazenamento do sistema (por exemplo a mídia magnética). Um terror que sempre ameaça os usuários de sistemas computacionais é a mensagem "disk full" (quantos arquivos já não foram perdidos porque na hora "H" o sistema de arquivos não encontra espaço disponível nos discos para o armazenamento).

A solução que parece prevalecer ao longo das décadas é, ou solicitar aos usuários que façam uma "limpeza" em seus diretórios (gerando algum espaço), ou procurar por fontes de recursos visando substituir o(s) disco(s) do sistema por unidades de maior capacidade, ou ambas (acreditam-se que soluções ditatoriais do tipo o gerente do sistema sai removendo arquivos à revelia dos usuários, já não sejam mais utilizadas!). Uma alternativa às vezes reportada é a utilização de um sistema de arquivamento (tipo arquivo morto) em fitas magnéticas, onde arquivos antigos (por exemplo com mais de 6 meses sem acesso) são removidos do sistema e transferidos para fitas magnéticas. Os usuários são informados pelo gerente do sistema sobre a remoção e recebem as informações necessárias para uma futura recuperação do arquivo (por exemplo, solicitando que a fita "x" seja montada no sistema para que o arquivo "y" seja recuperado).

Qualquer uma das soluções acima, apesar de resolver momentaneamente o problema, não condizem com o atual estágio de evolução dos sistemas computacionais. Quando se imagina um sistema computacional moderno, deve-se estar pensando em algo com dezenas, talvez centenas (ou mesmo milhares) de estações de trabalho, com acesso a um número elevado de servidores de arquivos. O que se requer em sistemas desse nível é um mecanismo automático para gerenciamento do problema de espaço em disco, que possa prover o sistema com um uso racional da mídia magnética (assume-se aqui que a mídia padrão dos sistemas computacionais modernos é a magnética).

Outro problema que também persegue os usuários de sistemas computacionais ao longo de décadas é segurança e confiabilidade [MUL89]. Muito tem sido feito nessa área com a introdução de técnicas sofisticadas para se proteger os usuários contra acessos indevidos, bem como perda de informação por danos lógicos ou físicos. Infelizmente, quando mecanismos especiais visando

melhorias na segurança e confiabilidade são introduzidos, inevitavelmente há perdas em desempenho. Por exemplo, quando criptografia é adotada sempre tem-se danos no desempenho. Da mesma forma, se técnicas de replicação de arquivos são utilizadas, o desempenho será penalizado, além de se ter mais blocos dos discos utilizados, podendo agravar ainda mais o problema da falta de espaço.

Análise da utilização do meio de armazenamento de diversos sistemas tem mostrado que muito do espaço utilizado poderia ser economizado, liberando áreas preciosas para os trabalhos em andamento. Por exemplo, análise efetuada no sistema de arquivos do sistema TRICE (um sistema computacional distribuído baseado em rede local, utilizando o modelo estação de trabalho/servidores e com utilização diária básica para desenvolvimento de sistemas, com cerca de 100 estações de trabalho, 500 usuários cadastrados e 10 servidores de arquivos, no Departamento de Eletrônica e Computação da Universidade de Southampton, Inglaterra [SAN90a]) mostra que os arquivos produzidos em um dia de trabalho são parcialmente removidos do sistema em curto espaço de tempo. Após 10 dias, por exemplo, cerca de 50 por cento dos arquivos originalmente criados já foram removidos. Entre duas a quatro semanas apenas 30 por cento dos arquivos originais continuam no sistema. Após um mês, os arquivos que ainda permanecem no sistema dificilmente são modificados ou mesmo removidos [SAN87].

Análise semelhante, conduzida considerando três sistemas de computação em utilização diária na Faculdade de Medicina de Ribeirão Preto - USP, forneceu os dados apresentados nas tabelas 1, 2, 3, e 4 [BEZ93]. Os sistemas analisados compreendem: (A) estação de trabalho Intergraph, (B) estação de trabalho SUN e (C) computador CDC 4330.

Sistema	Número de Usuários	Área Total (blocos)	Área em Uso (blocos)	Área Comprometida (blocos)	%
A	53	236.602	204994	170.244	83,1
B	49	140.638	106.572	84.769	79,5
C	715	558.438	435.727	187.026	42,9

Tab. 1 - Espaço com Arquivos não Acessados a mais de 180 Dias

Classe (blocos)	Número de Arquivos	Área Comprometida	% Total em Uso
(0, 2]	1.893	2.847	1,4
(2, 10]	2.385	12.063	5,9
(10, 50]	1.410	34.733	16,9
(50, 250]	418	47.132	23,0
(250, 500]	70	23.572	11,5
(500, 1000]	48	33.204	16,2
(1000, -)	11	16.693	8,1
Total	6.235	170.244	83,0

Tab. 2 - Espaço com Arquivos não Acessados a mais de 180 Dias - Sistema A

□ Classe (blocos)	Número de Arquivos	Área Comprometida	% Total em Uso
(0, 2]	5.149	6.551	6,2
(2, 10]	2.552	13.330	12,5
(10, 50]	1.005	20.381	19,1
(50, 250]	180	18.203	17,1
(250, 500]	23	7.632	7,2
(500, 1000]	13	9.552	9,0
(1000, -)	7	9.120	8,6
Total	8.929	84.769	79,7

Tab. 3 - Espaço com Arquivos não Acessados a mais de 180 Dias - Sistema B

Classe (blocos)	Número de Arquivos	Área Comprometida	% Total em Uso
(0, 2]	1.203	1.515	0,4
(2, 10]	1.878	10.588	2,4
(10, 50]	1.410	29.964	6,9
(50, 250]	362	37.479	8,6
(250, 500]	39	12.984	3,0
(500, 1000]	74	51.768	11,9
(1000, -)	25	42.728	9,8
Total	4.991	187.026	43,0

Tab. 4 - Espaço com Arquivos não Acessados a mais de 180 Dias - Sistema C

Nos três sistemas analisados, o comprometimento de espaço dos sistemas de arquivos com arquivos não utilizados recentemente é significativo, como pode ser observado pelos valores totais (83,0% para o sistema A, 79,7% para B e 43,0% para o sistema C).

A preocupação com a utilização de espaço em disco com arquivos que simplesmente foram "esquecidos" no sistema é, portanto, um assunto que deve estar sempre em pauta no dia a dia de um sistema computacional. Um ponto criticamente difícil é saber se um dado arquivo foi realmente "esquecido" no sistema, ou se é um arquivo com dados importantes que podem ser necessários, por exemplo, na próxima década (é o caso de dados estatísticos sobre epidemias).

Servidores de arquivamento (archive servers [SAN90]) surgiram como uma alternativa para o gerenciamento automático do espaço de armazenagem de sistemas computacionais distribuídos, bem como apoio à implementação de mecanismos para melhorar a segurança e a confiabilidade dos sistemas.

Um servidor de arquivamento é aqui referido como sendo um servidor de arquivos com alta capacidade de armazenagem (na ordem de vários Gigabytes) e alto grau de segurança (garantindo a consistência da informação ali depositada). Santana [SAN90], propõem a utilização de discos óticos do tipo **WORM** na definição de um servidor de arquivamento para o sistema TRICE, onde o disco ótico é utilizado em conjunto com dois discos magnéticos de tecnologia Winchester. O disco ótico armazena os arquivos que são transferidos do sistema convencional

segundo certos critérios; os discos magnéticos são utilizados para armazenar o diretório correspondente ao servidor de arquivamento.

Com base na argumentação apresentada nesta seção, nos números apresentados em [SAN90, SAN87] e nos resultados observados nos três sistemas analisados (tabelas 1, 2, 3 e 4), este artigo discute o desenvolvimento e implementação de um sistema auxiliar de arquivos, a ser utilizado como um servidor adicional, capaz de oferecer mecanismos automáticos para a transferência de certos arquivos para a mídia ótica (liberando espaço na mídia convencional) e prover o sistema com algumas melhorias em termos de segurança e confiabilidade (por exemplo, arquivos podem ser replicados na unidade ótica que, por suas características - discutidas na seção que segue -, diminuem radicalmente a probabilidade de se ter uma inconsistência no sistema).

Porquê Discos Óticos?

Os sistemas óticos de armazenagem de informação, apesar de serem resultado de uma tecnologia razoavelmente recente, têm sido muito utilizados nas grandes organizações com a finalidade de minimizar a problemática dos arquivos de papel, consulta a microfimes, discos magnéticos e segurança [GAR91, OLI91]. Os discos óticos estão disponíveis em três categorias básicas: **CD-ROM** (compact disc), **WORM** (**Write Once Read Many**) e **Regravável**.

Algumas das vantagens apresentadas pelos discos óticos em relação a outros dispositivos, principalmente em relação aos discos magnéticos são: baixo custo por byte armazenado, grande capacidade de informação por unidade de área de mídia, possibilidade de reprodução do disco e de seu conteúdo em série, possibilidade de remoção da sua unidade acionadora, as informações uma vez escritas no disco não podem mais ser apagadas (WORM), apresentam maior estabilidade e confiabilidade quanto a fatores ambientais tais como temperatura, umidade e principalmente perturbações magnéticas [FUJ84, CHR86, MIK88].

Por outro lado, como desvantagens fundamentais dos discos óticos em relação aos magnéticos têm-se: tempo médio de acesso cerca de seis vezes maior e falta de padronização de protocolo para gravação de informações nos discos **WORMs** e regraváveis.

Considerando-se sistemas de arquivamento, uma informação relevante é o custo do byte armazenado. Isto é, quanto custa para se armazenar um dado arquivo, considerando-se o custo da mídia necessária para armazená-lo? O custo médio de armazenagem de 5 megabytes de informação em um disco rígido removível é de US\$15.00 / Mbyte, em um disco rígido fixo este custo cai para US\$10.00 / Mbyte. Para disco ótico (WORM) este custo é de US\$0.13 / Mbyte e para o CD-ROM é de US\$0.02 / Mbyte, ficando claro a superioridade dos sistemas óticos [OLI91]:

A utilização da mídia ótica do tipo **WORM** tem sido apontada como uma alternativa atrativa para a implementação de sistemas de arquivamento, principalmente devido ao baixo custo de armazenamento (por byte) apresentado,

pela alta capacidade dos discos disponíveis e por sua característica de uma vez escrita uma informação ela não pode mais ser removida. Dessa forma discos **WORM** compreendem uma opção para a implementação de mecanismos que permitam não apenas o gerenciamento da utilização do meio de armazenagem de informação de modo racional, mas também como apoio à implementação de mecanismos de segurança e de alta confiabilidade.

Projeto do Sistema Auxiliar de Arquivos (SAA)

A política proposta por Santana no desenvolvimento do servidor de arquivamento do sistema TRICE [SAN90], é particularmente adequada à utilização de discos óticos do tipo **WORM** e serve como base para o **sistema auxiliar de arquivos** discutido neste artigo.

Para o projeto do sistema auxiliar de arquivos, baseado em disco ótico do tipo **WORM**, devem ser considerados três elementos básicos: **arquivos envolvidos**, **mecanismo de transferência** (do disco magnético para o disco ótico) e a organização geral do **disco ótico WORM**.

Arquivos Envolvidos

Três tipos de arquivos são definidos para o SAA em função dos aspectos de segurança e disponibilidade de espaço do servidor de arquivos do sistema considerado:

- **crítico:** um arquivo é considerado do tipo crítico quando não se admite perda da informação contida no arquivo ou da disponibilidade do arquivo. Nesse caso o serviço auxiliar de arquivos (implementado em um servidor auxiliar de arquivos) provê o sistema de arquivos (convencional) com uma cópia adicional do arquivo, armazenada na mídia ótica, toda vez que o arquivo sofrer alguma modificação (observe que as cópias, ou versões, anteriores, não podem ser fisicamente removidas do disco **WORM**);
- **normal:** um arquivo normal do sistema, que não exige cuidados especiais;
- **morto:** define a possibilidade de se adotarem cópias em disco ótico de algumas classes especiais de arquivos em processo de remoção do sistema convencional de arquivos, tais como fontes, dados, arquivos de usuários especiais, etc.

Como conseqüência da adoção dos três tipos de arquivos acima descritos, deve-se ter no sistema uma categoria de usuários especiais, os quais são autorizados a definir arquivos do tipo crítico (com a conseqüente ação do SAA, replicando o arquivo na mídia ótica).

Mecanismos de Transferência para o SAA

Para os arquivos *críticos* tem-se inicialmente a necessidade de um procedimento disponível aos usuários "especiais", habilitados a indicar arquivos críticos. A alternativa adotada é ter-se no sistema uma primitiva que permite a inclusão de um arquivo na lista de arquivos críticos gerenciada pelo SAA, seguindo um algoritmo simples como segue:

faça

*recebe do usuário habilitado identificação do arquivo crítico
atualiza lista de arquivos críticos do SAA*

fim

Uma vez que o sistema possui uma lista com a identificação dos arquivos críticos, toda vez que há alteração em um arquivo deve-se executar o seguinte algoritmo para verificar se o arquivo modificado pertence ou não à lista de arquivos críticos, devendo nesse caso ser também copiado no disco ótico (no caso de arquivos críticos o mecanismo de transferência é na verdade um mecanismo de criação de cópia de arquivo):

repita

*recebe do servidor de arquivo a identificação do último arquivo alterado "id_arq"
se id_arq consta da lista de arquivos críticos
então*

*faz nova cópia (versão) do arquivo no WORM
atualiza diretório WORM*

fim

Para os *arquivos normais* a transferência para o disco ótico não exige a mesma urgência dos arquivos críticos, podendo ser empregado um escalonamento de transferências que utilize os períodos ociosos do sistema. A transferência de um arquivo normal pode ser ativada segundo dois critérios:

- **longevidade (ou desuso):** a permanência de um arquivo normal no disco magnético é assegurada por um período de *n_dias* sem acesso. O tempo de permanência *n_dias* é um parâmetro definido em função dos dados apresentados na introdução deste artigo. O conteúdo do arquivo, que estiver com seu tempo de permanência no meio magnético vencido, é removido do disco magnético. Uma ligação ("link") entre o diretório original e o diretório de uso do disco ótico (o diretório **WORM**) será criada, possibilitando a reinstalação de uma nova cópia no meio magnético. O algoritmo proposto para esse caso é o seguinte:

faça

*percorre todos os diretórios e localiza os arquivos a serem transferidos
enquanto existir arquivo a transferir*

faça

*cria "link" entre diretório original e o de uso do disco ótico
copia arquivo para o disco ótico
remove conteúdo do arquivo do disco magnético
atualiza diretório do disco ótico*

fim

fim

fim

- **solicitação do usuário:** o usuário habilitado poderá solicitar a transferência de qualquer um de seus arquivos para o disco ótico, bastando para isso informar ao SAA a identificação do seu arquivo. Nesse caso, o SAA acrescentará a identificação do arquivo na lista de arquivos a transferir para futura transferência. O conteúdo do arquivo original permanece no disco magnético, caracterizando uma estratégia de replicação sob demanda:

```

faça
    percorre lista de arquivos a transferir
enquanto existir arquivo a transferir
    faça
        copia arquivo para o disco ótico
        atualiza diretório do disco ótico
    fim
fim
fim

```

Para arquivos mortos, a adoção deste modo de transferência está relacionada ao processo de remoção definitiva de arquivos do meio magnético. Normalmente o procedimento de remoção de arquivos de um sistema de arquivos dispõe de parâmetros de controle que podem evitar a remoção acidental de um arquivo. Por exemplo em ambientes UNIX o comando *rm -i arquivo* interroga o usuário se realmente deseja remover o arquivo, sendo possível cancelar-se o comando. Se o parâmetro *-i* não é especificado, a remoção é concluída, a menos que o usuário não tenha permissão para fazê-la. O que se tem no SAA é um novo processo de remoção que permite que arquivos importantes, mas que não estarão em uso diário, como é o caso de arquivos fontes de sistemas já concluídos e de dados estatísticos, entre outros exemplos, sejam removidos do sistema convencional (liberando espaço na mídia magnética) e transferidos para o disco ótico. A transferência do arquivo se dá através do algoritmo que segue:

```

faça
    percorre a lista de arquivos a transferir
enquanto existir arquivo a transferir
    faça
        copia arquivo para o disco ótico
        remove arquivo do disco magnético
        atualiza o diretório do disco ótico
    fim
fim
fim

```

As transferências de arquivos do sistema convencional para o sistema ótico é gerenciada por um **sistema escalador de transferências**, que monitora as listas de arquivos críticos e normais, indicando sempre ao processo **copiador de arquivos** qual o próximo arquivo a ser transferido para o disco ótico e qual o tipo de transferência a ser executada.

Uma vez que um determinado arquivo tenha sido transferido para o disco ótico, pode haver necessidade de sua relocação para o disco magnético. Quando um processo necessita de um arquivo removido para o disco ótico por longevidade, a reinstalação será automática. Isto é possível porque o sistema de arquivos, através de uma nova primitiva **OPEN** (a primitiva padrão do sistema é substituída), é capaz de reconhecer este tipo de arquivo. Todo o processo ocorre transparentemente durante a operação **OPEN**. Quando um usuário habilitado solicitar transferências ao SAA.

Os usuários comuns, sem habilitação para acesso especial ao SAA, não precisam conhecer nenhum detalhe sobre a atuação do SAA. Todos os mecanismos discutidos aqui são completamente transparentes, a menos que solicitados pelos usuários especiais.

Estrutura do Sistema de Arquivos do Disco Ótico

A organização do sistema de arquivos no disco ótico proposta é completamente independente do sistema operacional do sistema considerado, facilitando a adoção do SAA em diferentes ambientes.

Uma vez que o disco ótico **WORM** possibilita acesso aleatório e considerando-se a finalidade da escolha desse tipo de disco como uma medida de segurança para os arquivos, o disco é organizado em três regiões: para **arquivos**, **livre e diretório**. Essas regiões não têm limites pré-estabelecidos a não ser pela capacidade de armazenamento do disco considerado.

A região de arquivos (região 1), criada a partir do início do disco, é constituída por agrupamentos de blocos gravados correspondentes aos arquivos transferidos para o disco ótico. Cada arquivo é constituído de um bloco de cabeçalho seguido pelos blocos de dados do arquivo. A segunda região, livre ou intermediária, compreende os blocos disponíveis para a gravação de novos blocos de arquivos ou diretórios. A região de diretórios (região 3), criada a partir do final do disco, é constituída pelos blocos de diretório. A escolha de uma região para o diretório se deve a dois motivos:

- maior rapidez na localização de um arquivo e reconstituição do diretório em disco magnético;
- replicação (por medida de segurança) das entradas dos arquivos no disco ótico.

Considerando-se que todo arquivo gravado no disco ótico possui um cabeçalho, não haveria necessidade de se ter um diretório gravado na mídia ótica. Porém, a localização de um arquivo no disco ótico dentro da região 1 pode tomar muito tempo. Dessa forma, a existência de uma região para o diretório acelera a localização dos arquivos, enquanto que a adoção do cabeçalho identificador do arquivo, introduz uma redundância que pode auxiliar na recuperação do diretório em caso de falhas no momento de escrita no disco ótico.

O diretório do disco ótico é também mantido no disco magnético, com o propósito de melhorar o desempenho do sistema. A utilização do disco magnético como apoio ao SAA deve-se principalmente à grande diferença no tempo de acesso, conforme já discutido neste artigo.

A estrutura proposta para cada entrada do diretório do disco ótico, bem como para cada entrada do arquivo de diretório geral mantido no disco magnético, contém todas as informações necessárias para permitir o gerenciamento de várias unidades simultâneas de discos óticos no SAA e a correta ligação com a mídia ótica utilizada na gravação do arquivo, mesmo quando o disco for removido do sistema. O SAA controla a instalação e remoção de mídias óticas através da estrutura mantida em disco magnético contendo informações do dispositivo onde está instalado o disco ótico e a identificação do mesmo.

Implementação do Sistema Auxiliar de Arquivos

A implementação de um protótipo deste sistema, visando validar os algoritmos e estrutura propostos, foi realizada em um ambiente computacional baseado no sistema operacional MINIX [TAN87], em linguagem C, incluindo alterações das primitivas **OPEN** e **CLOSE** do sistema de arquivos do MINIX. Diversos programas de suporte e gerenciamento de arquivos transferidos do meio magnético para o disco ótico (e vice-versa) foram também desenvolvidos.

A abordagem básica da implementação compreende um processo gerenciador de transferências (denominado **wgt**), disparado durante a operação de carregamento do sistema operacional, através da execução pela "shell" do arquivo */etc/rc*. O processo instala-se por si só em "background" e tem a função de gerenciar as listas de arquivos a serem copiados para o disco ótico.

As listas de arquivos foram implementadas como subdiretórios de um diretório criado em disco magnético para uso exclusivo do disco ótico. Assim, os arquivos críticos são armazenados no subdiretório *crítico*, os arquivos removidos por longevidade são armazenados no subdiretório *desuso* e os solicitados pelos usuários no subdiretório *cópia*.

O gerenciamento dessas listas (ou subdiretórios) é realizado com a utilização de um arquivo diretório (**worm.dir**), contendo todas as entradas (identificações dos arquivos) das listas.

Componentes do Sistema Auxiliar de Arquivos

O procedimento **CLOSE** foi alterado para que, toda vez que um arquivo for fechado, seja feita uma consulta à lista de arquivos críticos e em caso afirmativo a identificação do arquivo é passada ao SAA através do disparo de um processo copiador (**cp** no MINIX). O copiador encarrega-se de fazer uma cópia do arquivo crítico no próprio disco magnético em um subdiretório (denominado crítico) de uso do disco ótico. Essa cópia permanece nesse subdiretório até que o arquivo seja definitivamente copiado para o disco ótico, sendo então removido do disco magnético (o arquivo original permanece também gravado no disco magnético).

A necessidade de alteração da primitiva **OPEN** está relacionada com os arquivos transferidos para o disco ótico por longevidade. Toda vez que um desses arquivos é requisitado por um processo qualquer, não é possível ao sistema de arquivos do MIMIX atender de imediato à solicitação. Assim, durante a execução da primitiva **OPEN**, a lista de arquivos removidos por longevidade é consultada e se for o caso, a primitiva **OPEN** indica ao SAA que é necessário reinstalar o arquivo na mídia magnética. Isto é feito através do disparo do processo re-instalador de cópias do disco ótico para o disco magnético (denominado **wrc**).

Além das alterações em procedimentos do sistema de arquivos do sistema operacional MINIX, a implementação proposta necessitou de novos procedimentos constituídos dos seguintes módulos básicos:

- **wuser**: autorização de usuários,

- **wgt**: gerenciador de transferência de arquivos para o disco ótico,
- **wrc**: re-instalador de arquivos do disco ótico para o magnético,
- **wdes**: coletor de arquivos não utilizados há mais de *n_dias*,
- **wci**: coletor de arquivos críticos,
- **wcp**: coletor de arquivos normais solicitados por usuários especiais.

Procedimentos Transparentes ao Usuário

O SAA dispõe de procedimentos que são totalmente transparentes ao usuário: **wgt** e **wdes** além das primitivas **OPEN** e **CLOSE** modificadas. Para os usuários especiais (habilitados) estão disponíveis os seguintes comandos: **wci**, **wcp** e **wrc**. Outros comandos de apoio ao usuário estão disponíveis em dois grupos básicos: comandos de uso específico do super-usuário (**wuser**, **wim**, **wmm** e **wdm**) e comandos para usuários especiais (**wdl** e **wlw**).

O comando **wuser** permite ao super-usuário cadastrar um usuário do SAA e tem a seguinte sintaxe: **wuser** *<identificação do usuário>* [*área*]. O parâmetro *área* permite a adoção de uma política de distribuição de área em disco de acordo com a classe do usuário.

O comando **wim** (**wim** *<dispositivo MINIX>* *<tamanho em blocos>* *<identificação para a mídia>*) é utilizado pelo super-usuário para iniciar uma mídia para uso no SAA. No caso da implementação para testes, não estando disponível o disco **WORM**, todo o mecanismo foi simulado utilizando-se disquetes.

O comando **wmm** é utilizado pelo super-usuário toda vez que houver necessidade de montar uma nova mídia (no caso da implementação simulada, um disquete) no SAA. A sintaxe é a seguinte: **wmm** *<dispositivo MINIX>*.

Para desmontar (retirar de uso) uma mídia do SAA, o super-usuário necessita do comando **wdm**, cuja sintaxe é: **wdm** *<dispositivo MINIX>*.

O comando **wdl** permite ao usuário do SAA listar o diretório do disco ótico, contendo todas as entradas de todos os arquivos gravados nos discos óticos. A sintaxe para esse comando é: **wld**.

O comando **wlw** tem função análoga ao **wld**, porém lista somente as entradas no diretório corresponde aos arquivos que estão na mídia ativa. A sintaxe deste comando é: **wlw**.

A figura 1 apresenta um diagrama ilustrando a estrutura e organização do SAA, com o inter-relacionamento de seus elementos e comandos de apoio.

Comentários Finais

Este artigo apresentou aspectos fundamentais do projeto e implementação de um sistema auxiliar de arquivos, baseado em disco ótico do tipo **WORM**, para utilização em ambientes computacionais distribuídos. A definição do sistema foi integralmente baseada na análise de dados reais, que mostram que o uso racional

da mídia magnética de um sistema de arquivos deve incluir a remoção de alguns arquivos com baixa probabilidade de acesso durante operação diária do sistema. Em adição, a solução proposta, com a utilização de discos óticos do tipo **WORM**, permite a criação de mecanismos que melhoram a segurança e confiabilidade dos sistemas, sem impor um custo elevado em termos de perda de desempenho.

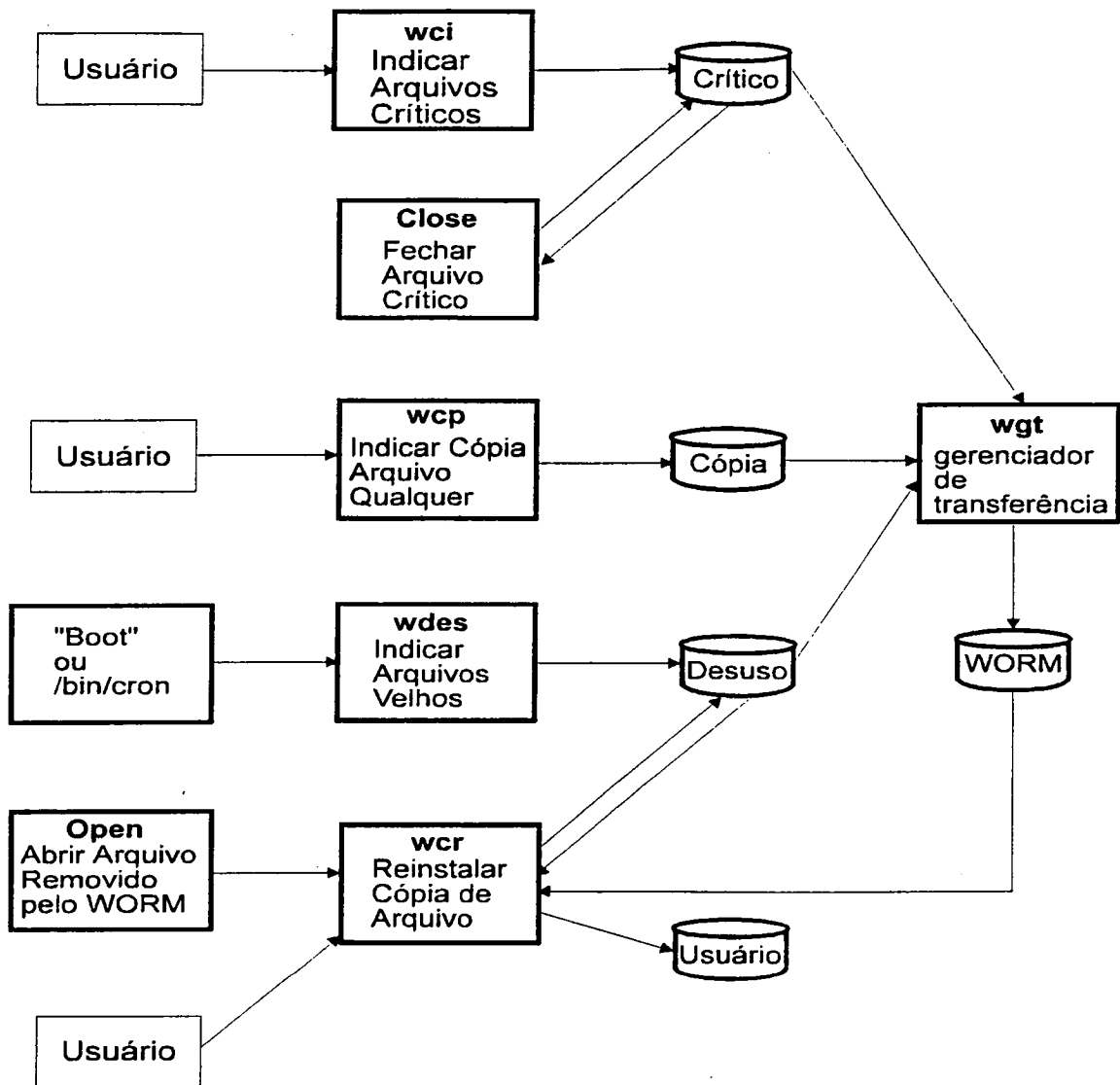


Figura 1 - Estrutura e Organização do SAA - WORM

O sistema proposto SAA pode, facilmente, ser adaptado a diversos ambientes computacionais existentes, bastando para isso que se façam as adaptações pertinentes nos procedimentos dependentes do sistema operacional. A estrutura e organização do disco ótico são totalmente independentes do sistema operacional adotado.

A implementação do SAA baseou-se em muitos dos princípios discutidos em [SAN90], mas difere substancialmente em termos de implementação, uma vez que o sistema proposto para o TRICE é fortemente amarrado à organização do sistema de arquivos adotado naquele sistema, que possui diversos servidores de arquivos

organizados hierarquicamente (o servidor de arquivamento está inserido em um desses níveis).

A implementação do SAA efetuada em ambiente MINIX serviu para validar os mecanismos e algoritmos propostos. Os resultados foram completamente bem sucedidos, viabilizando a implementação do sistema em novos ambientes e com a mídia ótica fisicamente presente (não simulada como nos testes). Implementação em ambiente UNIX está dentro dos planos em andamento no grupo de Sistemas Distribuídos e Programação Concorrente do Departamento de Ciências de Computação e Estatística do ICMSC-USP.

Referências

- BEZ93** Bezerra, L.A.F.. *Utilização de Disco Ótico (WORM) como Medida de Segurança de Arquivos em Redes Locais de Computadores*, Dissertação de Mestrado, ICMSC-USP, 1993.
- CHR86** Christodoulakis, S. and Faloutsos, C.. *Design and Performance Considerations for an Optical Disk-Based, Multimedia Object Server*. IEEE Computer, pp. 45-56. Dec. 1986.
- COU94** Coulouris, G.; Dollimore, J. and Kindberg, T.. *Distributed Systems Concepts and Design*, Addison-Wesley, 1994.
- FUJ84** Fujitani, L.. *Lase Optical Disk: The Coming Revolution on-Line Storage*. *Comm. ACM*, 27(6), pp. 546-554, Jun.1984.
- GAR91** Garfinkel, S.. *Designing a Wite-Once File System*. *Dr. Dobb's Journal*, pp. 78-86, Jan. 1991.
- MIK88** Mikro. *Compact Disk com Superfície Orgânica*. *Eletrônica, Informática, Automação e Telecomunicações*. 8, pp.40-51, Set. 1988.
- MUL89** Mullender, S.. *Distributed Systems*, ACM Press, Addison-Wesley Publishing Company, 1989.
- OLI91** Oliveira, E.T.. *O Laser Assume a Vanguarda*. *IPESI Negócios & Informática*, pp. 24-29, Out/Nov. 1991.
- SAN87** Santana, M.J.. *Analysis of the TRICE Filestore Statistical Data*. Thecnical Report. Department of Electonics and Computer Science. University of Southampton, 1987.
- SAN90** Santana, M.J.. *An Advanced Filestore Architecture for Multiple-LAN Distributed Computing System*. PhD Thesis, Department of Electonics and Computer Science. University of Soutampton, U.K., 1990.
- SAN90a** Santana, R.H.C.; Santana, M.J. and Zaluska, E.J.. *TRICE - a Transparent Multi-LAN Distributed Computing System*, *XVI Conferência Latino-Americana de Informática*, Assunção-Paraguai, Set. 1990.
- TAN87** Tanenbaum, A.S.. *Operating Systems: Design and Implementation*. Englewood Cliffs, Prentice-Hall, 1987.
- TAN92** Tanenbaum, A.S.. *Modern Operating Systems*. Englewood Cliffs, Prentice-Hall, 1992.

NOTAS DO ICMSC

SÉRIE COMPUTAÇÃO

- 020/95 NUNES, M.G.V.; HASEGAWA, R. - PROTEMA: intelligent tutoring systems for mathematics.
- 019/95 OLIVEIRA, M.C.; TURINE, M.A.S.; MASIERO, P.C. - A statechart - based model for hypertext.
- 018/95 PIMENTEL, M.G.C. - Alternative operations for browsing hypertext.
- 017/94 ROMEIRO, N.M.L.; CASTELO FILHO, A. - Análise Comparativa de Métodos Numéricos de equações algebrico-diferenciais.
- 016/94 MAGALHÃES, A.L.C.C.; SIQUEIRA, M.F.; OLIVEIRA, M.C.F. - Operadores de Euler na modelagem por fronteira: conceito, aplicação, estudos de casos.
- 015/94 ODA, C.S.; MOREIRA, E.S. - ASNMP graphical network monitor with automatic topology discovery.
- 014/94 FELIPE, L.S.G.; FRANCO, N.M.B. - Sobre a ordem de convergência para as equações integrais de volterra de segunda espécie tipo Abel com soluções não suaves.
- 013/94 PIMENTEL, M.G.C. - A framework for user-hypertext interaction.
- 012/94 TURINE, M.A.S.; MENDES, M.D.C.; NUNES, M.G.V. - TEGRAM: a geometry tutoring system based on Tangram.
- 011/94 SPOLON, R.; SPOLON, R.; SANTANA, M.J.; SANTANA, R.H.C. - Desenvolvimento de um gerador de aplicação para simulação de sistemas discretos.