

Explorando o Escalonamento no Desempenho de Aplicações Móveis Distribuídas

Adenauer Yamin^{1,3}, Iara Augustin^{2,3}, Jorge Barbosa^{1,3}, Luciano da Silva³, Cláudio Geyer³

¹ Escola de Informática, Universidade Católica de Pelotas (UCPel)
R. Félix da Cunha, 412, Pelotas, RS, Brasil
{adenauer, barbosa}@atlas.ucpel.tche.br

² Departamento de Eletrônica e Computação, Universidade Federal de Santa Maria (UFSM)
Campus Universitário, Santa Maria, RS, Brasil
august@inf.ufsm.br

³ Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS)
Av. Bento Gonçalves, 9500, Porto Alegre, RS, Brasil
{adenauer, august, barbosa, lucc, geyer}@inf.ufrgs.br

Resumo—

Uma nova classe de aplicações está emergindo, as aplicações móveis distribuídas. Estas aplicações são adaptativas por natureza, porque necessitam modificar o comportamento à medida que seu contexto de execução se altera (dados, recursos, serviços). Com este tema, está em desenvolvimento o projeto ISAM (Infra-Estrutura de Suporte às Aplicações Móveis). Este projeto objetiva definir uma arquitetura para o desenvolvimento e a execução desta classe de aplicações, a qual otimiza o desempenho das mobilidades lógica e física, e libera o programador de cuidados com as particularidades do ambiente de execução. A proposta contempla o uso de um novo modelo de desenvolvimento denominado Holoparadigma. Este artigo apresenta a arquitetura ISAM, e se concentra em aspectos do escalonamento de tarefas em sistemas distribuídos heterogêneos com suporte à computação móvel.

Palavras-chave— Escalonamento em Sistemas Distribuídos, Aplicações Móveis Distribuídas e Computação Móvel

Abstract—

A new class of applications is emerging, the distributed mobile applications. These applications are adaptable by nature, because they need to modify their behavior in response to changes in the execution context (data, resources, services). Within this theme, it is in development the ISAM project (Mobile Applications Support Infrastructure), which aims to define an architecture to the development and execution of mobile application. ISAM optimizes the performance of physical and logical mobility, and it releases the programmer of caring about execution environment details. The proposal contemplates the use of a new development model called Holoparadigm. This paper presents the ISAM architecture, and it concentrates on the tasks scheduling issues in distributed heterogeneous systems with mobile computing support.

Keywords— Scheduling in Distributed Systems, Distributed Mobile Applications and Mobile Computing.

I. INTRODUÇÃO

É esperado que os futuros ambientes de execução para aplicações distribuídas contemplem obrigatoriamente suporte, tanto para a mobilidade lógica como física, de processos e/ou recursos. Estes ambientes vão ser caracterizados por nodos de processamento bastante heterogêneos, os quais serão interconectados por redes sujeitas a freqüentes flutuações nos serviços fornecidos. Como consequência deste cenário, a busca de desempenho no software para ambientes móveis é complexa, seus componentes são variáveis no tempo e no espaço em termos de conectividade, portabilidade e mobilidade. Existem, portanto, requerimentos emergindo para uma nova classe de aplicações projetadas especificamente para este ambiente dinâmico. Esta nova classe de aplicações tem sido referenciada na literatura de muitas formas: *environment-aware*, *network-aware*, *resource-aware*, *context-aware*. A característica comum entre elas é a capacidade das aplicações adaptarem sua funcionalidade às condições dos recursos envolvidos nos diferentes momentos da execução.

A idéia de sistemas adaptativos não é nova. Segundo Katz [KAT 94], mobilidade exige adaptabilidade, o que significa que sistemas devem ter consciência da localização e do contexto, e devem tirar vantagem desta informação, estruturando-se de modo distribuído e reconfigurando-se dinamicamente. Davies [DAV 97] acredita que é somente através de um processo de adaptação, fornecendo às aplicações informações gerenciadas sobre trocas na sua infra-estrutura de execução, que se pode operar com desempenho (eficientemente) em um ambiente altamente dinâmico. Porém, modelos, arquiteturas e tecnologias para a computação móvel estão ainda nos seus primeiros estágios de desenvolvimento, e estão somente iniciando a tratar os desafios postos pela mobilidade [PIC 00]. Parece, portanto, ser necessário definir uma nova arquitetura de

sistemas móveis, projetada desde seu início com flexibilidade e adaptabilidade em mente.

Com esta visão, o projeto ISAM (Infra-estrutura de Suporte a Aplicações Móveis), propõe uma arquitetura de software que simplifica a tarefa de implementação de aplicações móveis distribuídas. O objetivo é conceber um ambiente de desenvolvimento e execução no qual todos os componentes, mesmo os básicos, estarão comprometidos com a premissa de elevada adaptabilidade.

Este artigo apresenta uma proposta onde o escalonamento é projetado como a estratégia central de adaptação, e por consequência de aumento de desempenho.

O texto está estruturado da seguinte forma: primeiro, na seção II, discute-se a necessidade da adaptação e seu conceito no ambiente móvel. Segue-se, na seção III, com uma introdução à arquitetura ISAM. Na seção IV, apresenta-se o modelo de escalonamento proposto. Faz-se na seção V uma análise dos trabalhos relacionados, e na seção VI apresenta-se a conclusão geral.

II. A NECESSIDADE DA ADAPTAÇÃO

A computação móvel genericamente se refere a um cenário onde todos ou alguns *hosts* que tomam parte na computação são móveis [BAG 98]. Desta definição pode-se derivar diferentes interpretações. Em um extremo, a mobilidade leva em conta as necessidades dos usuários nômades, isto é, usuários que se conectam na rede de localizações arbitrárias e que não ficam permanentemente conectados. Em outro extremo, estão os usuários móveis, os quais retêm a conectividade durante o deslocamento, tipicamente explorando links sem fio. Desta forma, a computação móvel é caracterizada por três propriedades: mobilidade, portabilidade e conectividade [AUG 00, PIC 00]. Para ser portátil, um computador deve ser pequeno, leve e requerer fontes pequenas de energia. Isto significa que este tipo de equipamento apresenta restrições no tamanho de memória, na capacidade de armazenamento, no consumo de energia e na interface do usuário. Os segmentos sem fio da rede levantam outros obstáculos: comunicação intermitente (desconexões freqüentes, bloqueio no caminho do sinal, ruído), restrita (e altamente variável), largura da banda, alta latência e alta taxa de erros. Quando em movimento, o dispositivo móvel freqüentemente altera seu ponto de contato com a rede fixa. Essa natureza dinâmica do deslocamento introduz questões relativas tanto ao endereçamento dos nós, quanto às informações dependentes da localização. Os ambientes móveis requerem mecanismos dinâmicos que façam o mapeamento das unidades móveis na infra-estrutura de comunicação, para permitir a comunicação entre as mesmas.

Portanto, o projeto de aplicações móveis deve levar em conta essas limitações, de modo que as mesmas não percam sua consistência quando um recurso não estiver disponível

[AUG 01]. Para que apresentem um desempenho compatível com a expectativa do usuário, muitos pesquisadores concordam que as arquiteturas para suporte à computação móvel necessitam contemplar capacidade de adaptação às freqüentes e rápidas alterações no ambiente de execução durante o curso do processamento da aplicação [KAT 94].

Isso que nos conduz à indagação: como especificar aplicações móveis distribuídas? A perspectiva do projeto ISAM é oferecer um ambiente de desenvolvimento e execução comprometido em otimizar o desempenho destas aplicações, e que também permita ao desenvolvedor abstrair da camada de aplicação as especificidades dos diversos recursos reais envolvidos no processamento. Os principais componentes deste ambiente são apresentados na seção III, a seguir.

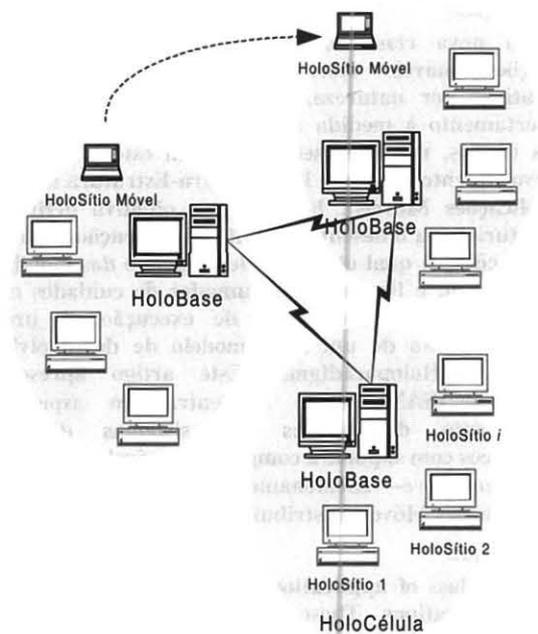


Fig.1 Ambiente de Execução ISAM

III. A ARQUITETURA ISAM

Como mencionado anteriormente, a computação móvel abrange uma faixa de cenários, os quais têm diferentes requerimentos no sistema de suporte. A princípio, uma categorização distingue entre dois cenários:

- (a) infra-estruturado - cenário composto pela presença de uma rede fixa onde alguns *hosts*, referenciados como estações-base, constituem os pontos de acesso para os *hosts* móveis.
- (b) ad-hoc - cenário dinâmico composto somente por *hosts* móveis (sem o suporte dado por uma rede fixa). A topologia resultante é altamente variável,

constituída a partir das interseções das áreas (células) de abrangência dos *hosts* móveis.

Considera-se que, para o desenvolvimento de aplicações distribuídas mais avançadas, é necessário que os *hosts* móveis usufruam a infra-estrutura da rede fixa existente, e possam se beneficiar de ambientes como o oferecido pela Internet. Desta forma, o modelo de rede adotado é o de uma rede móvel infra-estruturada. Este modelo é refletido nos elementos básicos do ambiente de execução do sistema ISAM apresentado na figura 1. Estes elementos são:

- HoloBase - é o ponto inicial de contato do *host* móvel com os serviços ISAM residentes na parte fixa da rede. Possui as funções de identificação, autenticação e de ativação das ações básicas do sistema;
- HoloCélula - denota a área de atuação de uma HoloBase, e é composta pela mesma e por HoloSítios;
- HoloSítio - são os nodos do sistema responsáveis pela execução da aplicação móvel distribuída propriamente dita. Nestes também são processados serviços de gerenciamento ISAM;
- HoloSítioMóvel - são os nodos móveis do sistema, responsáveis por serviços de interface com o usuário, e por algumas funções de monitoramento de recursos ISAM;
- HoloHome - é um ponto de referência único por usuário móvel no âmbito de toda rede. Está associado a um HoloSítio registrado para tal na arquitetura.

A crescente disponibilidade de facilidades de comunicação tem deslocado as aplicações da computação móvel de uma perspectiva de uso pessoal para outras mais avançadas de uso corporativo. Exatamente este domínio de aplicações constitui o escopo de interesse da arquitetura ISAM, ilustrada na figura 2. A arquitetura proposta é organizada em camadas com níveis diferenciados de abstração.

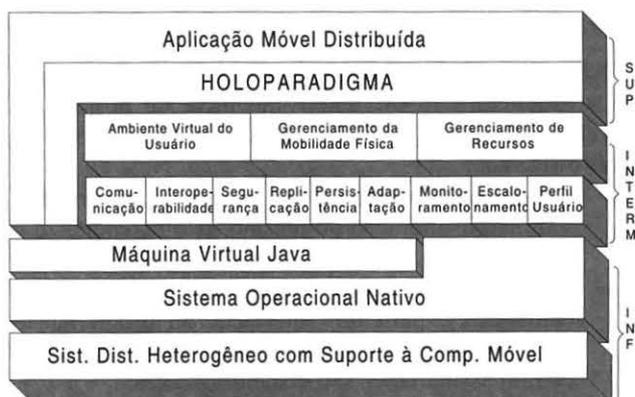


Fig.2 A Arquitetura ISAM

A camada superior (SUP) constitui a aplicação do usuário desenvolvida com a HoloLinguagem [BAR 01a], uma linguagem de programação que integra os paradigmas em lógica, imperativo e orientado a objetos. Além disso, a HoloLinguagem utiliza um modelo de coordenação que suporta invocações implícitas (*blackboard*) e explícitas (mensagens). Este modelo de coordenação é apropriado ao ISAM porque contempla o desacoplamento espacial e temporal da comunicação e sincronização, propriedades importantes para a computação móvel [PIC 99]. A HoloLinguagem suporta ainda concorrência, modularidade, mobilidade e encapsulamento de *blackboards* em tipos abstratos de dados. No Holoparadigma [BAR 01b], a aplicação é modelada com entes (entidade de existência) e símbolos (entidade de informação). Existem dois tipos de entes: elementar e composto. Um ente elementar é organizado em três partes: Interface, Comportamento e História. A interface descreve suas possíveis relações com os demais entes. O comportamento contém ações que implementam sua funcionalidade. Por sua vez, a história é um espaço de armazenamento compartilhado no interior de um ente. Um ente composto possui a mesma organização do ente elementar, no entanto, suporta a co-existência de outros entes na sua composição (entes componentes). Cada ente possui uma história. A história fica encapsulada no ente e, no caso dos entes compostos, é compartilhada pelos entes componentes. Os entes componentes participam do desenvolvimento da história compartilhada e sofrem os reflexos das mudanças históricas. Sendo assim, podem existir vários níveis de encapsulamento da história. Os entes acessam somente a história no seu nível. Como o HoloParadigma foi concebido com a perspectiva de mobilidade tanto lógica quanto física, sua utilização torna-se natural na proposta ISAM. A figura 3 ilustra duas possibilidades de mobilidade dos entes que são de interesse da arquitetura ISAM: a mobilidade física e lógica (A), e a mobilidade física (B).

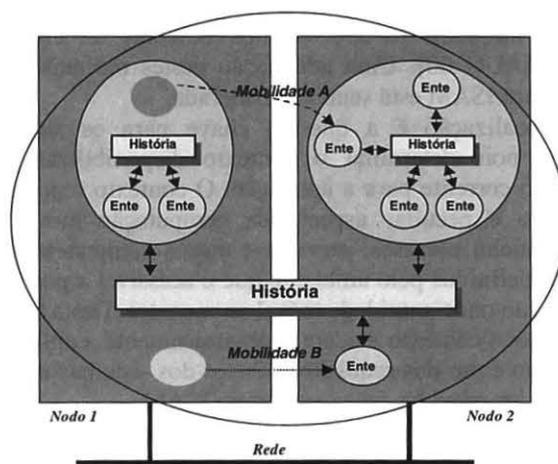


Fig.3 Mobilidade no HoloParadigma

A camada intermediária (INTERM) da figura 2 é o núcleo funcional da arquitetura ISAM, e é fornecida em dois níveis de abstração. O primeiro nível é composto por três módulos de suporte à aplicação: Ambiente Virtual do Usuário, Gerenciamento da Mobilidade Física e Gerenciamento de Recursos.

A. Ambiente Virtual do Usuário (AVU)

Compõe-se dos elementos que integram a interface de interação do usuário móvel com o sistema. O modelo foi projetado para suportar a exploração de aplicações contextualizadas (adaptadas aos recursos, serviços e localização corrente) e individualizadas (adaptadas aos interesses e preferências do usuário móvel). O desafio da adaptabilidade é suportar os usuários em diferentes localizações com diferentes sistemas de interação que demandam diferentes sistemas de apresentação dentro dos limites da mobilidade. Este módulo deve caracterizar, selecionar e apresentar as informações de acordo com as necessidades e o contexto em que o usuário se encontra. Para realizar estas tarefas, o sistema se baseia num modelo de uso, onde as informações sobre o ambiente de trabalho do usuário, preferências, padrões de uso, padrões de movimento físico e hardware do usuário são dinamicamente monitoradas e integram o Perfil do Usuário.

B. Gerenciamento da Mobilidade Física (GMF)

Este gerenciamento dá suporte ao movimento dos dispositivos móveis entre diferentes HoloCélulas, mantendo a execução durante o deslocamento. Para tal, deve redirecionar as referências dos recursos e serviços alocados, ligando recursos e serviços equivalentes na nova localização. Se isto não for possível, mantém as referências para os recursos originais, agora remotos. O GMF também deve dar suporte às operações desconectadas. A desconexão no ambiente móvel é mais uma regra do que uma exceção, e tem sido tratada extensivamente no âmbito do acesso aos dados. Mecanismos de *caching* e replicação otimista são a solução amplamente adotada nos sistemas de arquivos móveis [AUG 00]. Uma adequação destes mecanismos à arquitetura ISAM está sendo considerada.

A localização é a questão chave para os sistemas móveis, pois determina o contexto disponibilizado no momento corrente para a aplicação. O contexto representa um novo e peculiar aspecto da computação móvel. O mesmo inclui recursos, serviços e outros componentes do sistema definidos pelo ambiente que é acessível a partir da localização onde a unidade móvel se encontra. Desta forma, identificar o contexto e adaptar dinamicamente a aplicação ao mesmo é um dos requisitos básicos dos sistemas móveis distribuídos providos pela arquitetura ISAM.

C. Gerenciamento de Recursos (GR)

A adaptação é ativada quando determinado nível de alteração no contexto de execução da aplicação é detectado. Para tal, o sistema provê serviços de monitoramento dos recursos de interesse da aplicação e do escalonador. Na arquitetura ISAM o escalonamento é um componente central do GR. As informações necessárias podem ser agrupadas em categorias: estáticas (poder da CPU, memória instalada, versão do sistema operacional, etc.), dinâmicas (bateria, latência, ocupação processador/memória, etc.) e previsões futuras. Além disso, como as aplicações podem depender de diferentes tipos de informações, a arquitetura ISAM deve gerenciar dados, que podem variar desde a topologia de um subconjunto de nós da rede (grafos estruturados) a valores simples (escalares). Para a adaptação é exigida do módulo de monitoramento e gerenciamento dos recursos uma infra-estrutura sofisticada, que permita tratar não somente um recurso isoladamente, mas também a relação entre eles. Considerar a relação entre recursos é um dos principais aspectos deste módulo.

No segundo nível da camada intermediária estão os serviços básicos do ambiente de execução ISAM. Os serviços pertinentes ao escopo deste artigo, serão discutidos na seção IV.

Por sua vez, a camada inferior (INF) da figura 2 é composta dos sistemas de infra-estrutura distribuída pré-existent, tais como sistemas de rede móvel, sistemas operacionais nativos e a Máquina Virtual Java.

IV. O ESCALONAMENTO NO ISAM

De forma análoga às aplicações tradicionais de alto desempenho, uma aplicação móvel deve oferecer resultados o mais rapidamente possível ao seu usuário. Isto é uma exigência implícita à situação de mobilidade do usuário: equipamento com pouca autonomia (operado a bateria), custo de uso da rede móvel, inserção no tempo/espço do contexto da tomada de decisão (e.g. reuniões com clientes), etc. Algumas decisões pertinentes ao escalonamento na arquitetura ISAM são abordadas a seguir.

A. Características Gerais

O *framework* que fornece o suporte de escalonamento para a arquitetura ISAM tem como meta de projeto ser flexível e extensível [YAM 01]. Suas características mais significativas são:

- ♦ sua operação ocorre sobre o sistema operacional, e sem exigir alteração do mesmo. Isto potencializa a portabilidade;
- ♦ pode suportar tanto execuções paralelas como distribuídas. Para tal, interfaces de programação para comunicação interprocessos, tanto síncronas quanto assíncronas, são disponibilizadas;

- ◆ não está comprometido com uma heurística de escalonamento em particular. Ao contrário, disponibiliza facilidades para que novas heurísticas sejam implementadas;
- ◆ a heurística a ser utilizada é selecionada e/ou contextualizada por usuário e aplicação;
- ◆ os componentes que tomam decisão são replicados, e são capazes de atividades autônomas e assíncronas;
- ◆ as metas de escalonamento são perseguidas em escopos. Cada componente que toma decisão escala serviços no seu domínio;
- ◆ uso intensivo de registro histórico como auxiliar na tomada de decisão.

Algumas dessas características são típicas de propostas de balanceamento de carga difusas voltadas para aplicações com elevada dinamicidade de execução [COR 99].

No que diz respeito às estratégias para maximização do desempenho, o escalonamento no ISAM utiliza as seguintes:

- ✓ balanceamento de carga nos nodos responsáveis pelo processamento;
- ✓ localização dos recursos (software e hardware) mais próximos (reduzir custo de comunicação);
- ✓ emprego de replicação de serviços e de dados;
- ✓ disponibilização antecipada, por usuário, da demanda de componentes das aplicações e dos dados;
- ✓ otimização no volume de comunicações, utilizando transferências de contextos e componentes de aplicação personalizadas por usuário;
- ✓ monitoração da comunicação praticada pelos componentes das aplicações em execução, com intuito de otimizar aspectos de mapeamento.

O emprego destes procedimentos fica potencializado pela possível alternância do ponto de conexão dos HoloSítios móveis no contexto da rede, comportamento este inerente à computação móvel.

B. Uso de Adaptação Dinâmica de Interesse do Usuário

À medida que o usuário interage com o sistema, seu comportamento é monitorado e seu perfil é definido (figura 4). Esta informação é utilizada pelo ambiente de escalonamento para tomadas de decisão em diferentes situações, como pode ser observado na subseção B.3.

Nesta proposta, o escalonador emprega uma abordagem estocástica com aprendizado por reforço, na qual são construídas correlações estatísticas entre o usuário, o comportamento das suas aplicações e o ambiente de execução. Trabalho neste mesmo sentido, voltado para o escalonamento de aplicações paralelas em multiprocessadores, pode ser encontrado em [ZOM 98]. Um dos principais objetivos de dotar o escalonamento com uma estratégia de aprendizado por reforço é viabilizar uma

instanciação otimizada e antecipada de recursos nas HoloCélulas.

B.1 Do Ponto de Vista da Instanciação Otimizada

A premissa é buscar um modelo WYNIWYG (*What You Need Is What You Get*) [KON 00]. O escalonador atua como um configurador automático da arquitetura ISAM. Particularmente no suporte à aplicação do usuário, este carrega nos HoloSítios (móveis ou fixos) um conjunto mínimo de componentes que garantam a sua execução.

B.2 Do Ponto de Vista da Instanciação Antecipada

O processo de instanciação começa no momento em que o usuário efetiva sua autenticação na HoloBase, antes do mesmo solicitar a execução de aplicações. Esta instanciação também pode ocorrer com uma antecipação ainda maior, tendo por referência uma expectativa de roteiro de mobilidade já consolidada.

A rede de suporte à arquitetura ISAM é do tipo *wide-area*, e emprega a Internet na maioria dos segmentos. Isto introduz uma elevada heterogeneidade no tocante à velocidade das conexões entre as HoloCélulas. Por sua vez, as aplicações são de natureza distribuída; deste modo, antecipar o tráfego na parte estruturada da rede (com conexão física) é uma opção da arquitetura proposta para aumentar o desempenho da aplicação, e conseqüentemente reduzir o tempo de espera/conexão do usuário do segmento de rede com suporte à mobilidade (conexão sem fio).

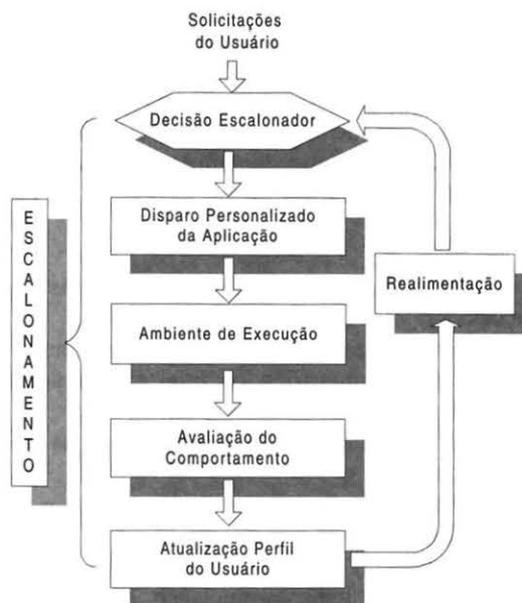


Fig 4. O Aprendizado no Escalonamento ISAM

B.3 Exemplo da Atuação Adaptativa dos Escalonadores

Para exemplificar a atuação do escalonador como uma estratégia de adaptação, descreve-se algumas situações

relativas ao Ambiente Virtual do Usuário (AVU). A instalação do AVU é uma atribuição do escalonador. Os critérios que ele irá utilizar são provenientes do Perfil do Usuário, seja para fazer uma alocação tendo o balanceamento de carga na HoloCélula como meta, seja considerando critérios de afinidade da aplicação com algum HoloSítio.

Na localização dos dados, para instalar o AVU, o escalonador irá verificar qual alternativa de origem apresenta menor custo de comunicação. Os dados para compor o perfil do usuário poderão vir da sua HoloHome, ou da última HoloCélula que o atendeu. Verificada ser a segunda opção o melhor caminho, solicitará da respectiva HoloBase a confirmação se a mesma ainda tem ativo sob sua tutela o AVU. Caso este exista, será feita a coleta de dados a partir do HoloSítio que o contém; caso contrário, os dados virão da HoloHome do usuário. A informação de qual foi a última HoloBase que atendeu determinado usuário, fica registrada no equipamento móvel.

O AVU será atualizado sempre que for detectada uma mudança no perfil de comportamento do usuário. A atualização será feita tanto no registro primário do perfil (na HoloBase), como nas cópias que estejam distribuídas na Arquitetura. O tempo que o AVU será mantido no HoloSítio após a desconexão também é determinado a partir do Perfil do Usuário, especificamente de seu padrão de uso.

C. Organização Distribuída do Escalonamento

O modelo de escalonamento adotado utiliza uma organização fisicamente distribuída e cooperativa [CAS 88] conforme a figura 5. A proposta está baseada em dois escalonadores: um denominado EscWAN, com atuação entre as HoloCélulas, outro denominado EscEnte, que atua localmente nas HoloCélulas.

C.1 Escalonador EscWAN

Fica localizado no nodo HoloBase, e tem as seguintes atribuições:

- definir, entre as HoloCélulas que tenham alguns recursos compartilhados, qual oferece acesso com menor custo de comunicação;
- decidir quando e onde replicar serviços e/ou base de dados;
- decidir quando e para onde migrar a base de dados;
- instanciar o Ambiente Virtual do Usuário nos HoloSítios. Esta instanciação é feita sob duas óticas: (i) balanceamento de carga - neste caso é escolhido o nodo menos carregado, (ii) aspectos de afinidade da aplicação - exigência de memória, bases de dados, etc.;
- escalonar componentes da aplicação provenientes de outras HoloCélulas, utilizando também tanto

critérios de balanceamento de carga quanto de afinidade.

Pelas suas atribuições, além da consideração de custos de comunicação e balanceamento de carga, o escalonador EscWAN atua de forma intensiva sobre aspectos de replicação e migração. Uma discussão neste sentido pode ser encontrada nos trabalhos [FER 00, AHM 98].

C.2 Escalonador EscEnte

Existente em todos os HoloSítios, tem como principal atribuição efetuar o mapeamento dos componentes da aplicação nos HoloSítios da HoloCélula. Os critérios utilizados também são balanceamento de carga e critérios funcionais de afinidade.

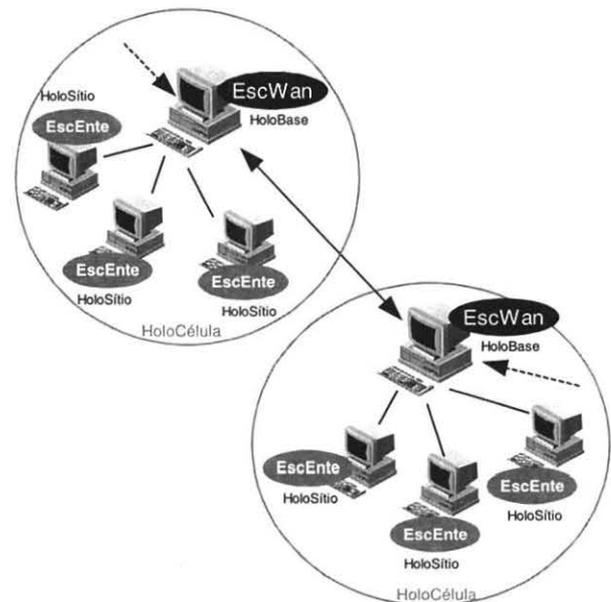


Fig. 5. A Organização do Escalonamento na Arquitetura ISAM

V. TRABALHOS RELACIONADOS

Uma quantidade substancial de pesquisa no campo da computação móvel é devotada a tornar as aplicações adaptativas e cômicas dos recursos [AND 00, BAG 98, NOB 00]. O foco das soluções é variável, tendo uma concentração em técnicas de adaptação dos tipos de dados à variação nos recursos da rede (largura de banda, em especial) [ANG 98, BAG 98, NOB 00, RAN 97, WEL 98]. Em geral, esses sistemas usam processos intermediários entre o cliente móvel e o servidor, na forma de *proxy* ou agente, os quais executam algum tipo de filtro que modifica a estrutura/quantidade de dados antes de serem transmitidos na rede sem fio. Outra estratégia de adaptação muito usada é a migração, de *thread* [RAN 97], de *proxy* [AHM 95] ou de agente [GRA 97], onde a decisão de migrar para um ponto específico é da aplicação. Diferentemente destes sistemas, ISAM utiliza o conceito de escalonamento como

uma estratégia central de adaptação. No ISAM o escalonador pode ser visto como um gerente geral distribuído, que negocia com a aplicação as decisões de adaptação. O escalonador, com base nas informações de alteração de contexto e nas políticas adotadas para a aplicação, pode deliberar ações de adaptação. Por sua vez, a aplicação pode requisitar a intervenção do escalonador.

Em geral, pesquisas em aplicações móveis adaptativas podem ser agrupadas em quatro categorias: (i) monitoramento de recursos, em especial recursos da rede [ANG 98, BHA 98, DEW 97, WEL 98]; (ii) aplicações móveis específicas [JIN 98, NOB 00]; (iii) toolkits para o desenvolvimento de aplicações [BAY 97, BLA 98, BOL 98, KUN 99, MAS 01, RAN 97, TAU 96] (iv) protocolos de suporte de comunicação, em particular variantes do TCP/IP [KID 98]. Esses sistemas tratam de aspectos específicos do ambiente móvel. Por outro lado, ISAM propõe abordar as questões introduzidas pela mobilidade de forma integrada, oferecendo a linguagem para o desenvolvimento das aplicações e o sistema de execução que monitora o contexto e fornece mecanismos de adaptação às alterações contextuais.

Em sistemas altamente distribuídos com suporte à computação móvel, o conceito de escalonamento como estratégia de adaptação não foi abordado pela comunidade científica, no que é de nosso conhecimento. A simplicidade desta idéia contrasta com a complexidade de sua implementação em um ambiente altamente dinâmico. Esta circunstância exige um alto grau de adaptação tanto da aplicação móvel quanto da própria plataforma de suporte à execução.

Na perspectiva de utilizar a Internet como infraestrutura para aplicações altamente distribuídas, têm surgido diversos trabalhos. Sistemas como o Condor [LIT 88] são voltados para aplicações de alto desempenho em *clusters* de estações de trabalho. Diferentemente do ISAM, utilizam um mecanismo central para disparar processos. O projeto Globus [FOS 98] disponibiliza uma "grade de recursos computacionais" [FOS 99] integrando equipamentos heterogêneos em um único sistema. De forma similar à proposta ISAM, ele contempla uma estrutura escalável e distribuída para o gerenciamento de recursos. Apesar de conter módulo específico para o controle de aplicações (GEM – *Globus Executable Management Service*), a atual versão trata as aplicações como um único executável, ao invés de uma coleção de componentes que podem ser parcial e dinamicamente instanciados como na arquitetura ISAM.

Por sua vez, sistemas como Globe [STE 99], Legion [GRI 97], e WebOs [VAH 98], apesar de suportarem diferentes níveis de configuração, não consideram a adaptabilidade e a configuração automática do ambiente de execução como uma questão central.

VI. CONCLUSÃO

A adaptação é o mais importante requisito para as aplicações móveis atingirem o grau de desempenho que atenda as expectativas dos usuários. A adaptação pode ser realizada em diversos níveis: rede, sistema e aplicação. Muitas técnicas, tais como *prefetching* e *caching*, filtragem, compressão e migração, são empregadas com este propósito. Diferentemente, a proposta da arquitetura ISAM é usar o escalonamento como uma estratégia de adaptação implícita e explícita para as aplicações móveis distribuídas. O conceito de adaptação é desenvolvido como uma negociação entre a aplicação e a arquitetura ISAM. O escalonamento toma decisões baseadas em informações sobre o ambiente de execução, coletadas e mantidas pela arquitetura e pela política de adaptação adotada pela aplicação. O ambiente de execução ISAM compõe-se do perfil do usuário móvel (preferências, padrões de uso e movimento), e do contexto (serviços e recursos) determinados pela localização corrente.

BIBLIOGRAFIA

- [AHM 95] AHMAD, Tahir; et al. The DIANA Approach to Mobile Computing. In *Mobile Computing*: Kluwer Academic Press, 1995.
- [AHM 98] AHMAD, I; KWOK, Y. On Exploiting Task Duplication in Parallel Program Scheduling. *IEEE Transactions on Parallel and Distributed Systems*. New York, v.9, n.9. 1998.
- [AND 00] ANDRÉ, Françoise; SEGARRA, Maria-Teresa. A Generic Approach to Satisfy Adaptability Needs in Mobile Environments. In: 33rd ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCE. *Proceedings...* Maui, Hawaii, USA. 2000.
- [ANG 98] ANGIN, Oguz; et al. The Mobeware Toolkit: Programmable Support for Adaptive Mobile Networking. *IEEE Personal Communications Magazine*. Special Issue on Adapting to Network and Client Variability. Aug. 1998.
- [AUG 00] AUGUSTIN, Iara. Acesso aos Dados no Contexto da Computação Móvel. PPGC/UFRGS. Porto Alegre. Dez. 2000 (Exame de Qualificação).
- [AUG 01] AUGUSTIN, Iara; GEYER, Cláudio. Sistemas Móveis Distribuídos: Análise de Suas Questões. Maio. 2001 (Submetido a publicação).
- [BAG 98] BAGGIO, Aline. System Support for Transparency and Network-aware Adaptation in Mobile Environments. In: ACM SYMPOSIUM ON APPLIED COMPUTING - Special Track on Mobile Comp. Systems and Applications. *Proceedings...* Atlanta, USA. Feb. 1998.
- [BAR 01a] BARBOSA, Jorge L. V.; Geyer, Cláudio F.R. Uma Linguagem Multiparadigma Orientada ao Desenvolvimento de Software Distribuído. V SIMPÓSIO BRASILEIRO DE LINGUAGENS DE PROGRAMAÇÃO (SBLP). *Anais*. Maio. 2001.
- [BAR 01b] BARBOSA, Jorge L. V.; Geyer, Cláudio F.R. Integrating Logic Blackboards and Multiple Paradigm

- for Distributed Software Development. Proceedings of Intr. Conference on Parallel and Dist. Processing Techniques and Applications (PDPTA). June. 2001.
- [BAY 97] BAYDERE, Sebnem. MaROS: a Framework for Application Development on Mobile Hosts. INTERNATIONAL CONFERENCE ON DISTRIBUTED AND PARALLEL SYSTEMS (EURO-PDS'97). *Proceedings...* 1997.
- [BHA 98] BHARGHAVAN, Vaduvur; et al. The Timely Adaptive Resource Management Architecture. *IEEE Personal Communications Magazine*. v.5, n. 4. 1998.
- [BLA 98] BLAIR, G. The Role of Open Implementation and Reflection in Supporting Mobile Applications. DATABASE AND EXPERT SYSTEMS APPLICATIONS (DEXA). *Proceedings...* 1998.
- [BOL 98] BOLLIGER, J.; GROSS, T. A Framework-based Approach to the Development of Network-aware Applications. *IEEE Transactions on Software Engineering*. v.24. 1998.
- [CAS 88] CASAVANT, Thomas L.; KUHL, Jon G. A taxonomy of scheduling in general-purpose distributed computing systems. *IEEE Transactions on Software Engineering*, New York, v. 14, n. 2. 1988.
- [COR 99a] CORRADI, Antonio; LETIZIA, Leonardi; ZAMBONELLI, Franco. Diffuse Load-Balancing Policies for Dynamic Applications. *IEEE Concurrency*. New York, v7, n.1. 1999.
- [DAV 97] DAVIES, N; et al. Limbo: a Tuple Space Based Platform for Adaptive Mobile Applications. INTERNATIONAL CONFERENCE ON OPEN DISTRIBUTED PROCESSING / DISTRIBUTED PLATFORMS (ICODP/ICDP'97). *Proceedings...* Toronto, Canada. May. 1997.
- [DEW 97] DeWITT, Tony; et al. ReMoS: A Resource Monitoring System for Network-aware Applications. Technical Report. Carnegie Mellon University. Dec. 1997. Disp. em <http://www.cs.cmu.edu>
- [FER 00] FERRARI, D. Nice. Um modelo de Replicação em ambientes que suportam mobilidade. PPGC/UFRGS. 2000. (Dissertação de Mestrado).
- [FOS 98] FOSTER, I; KESSELMAN, C. The Globus Project: A Status Report. In Proceedings of the IPPS/SPDP – Heterogeneous Computing Workshop. 1988.
- [FOS 99] FOSTER, I; KESSELMAN; C, Editors. *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers. San Francisco. 1999.
- [GRA 97] GRAY, Robert; KOTZ, David; NOG, Saurab; RUS, Daniela; CYBENKO, George. Mobile Agents for Mobile Computing. Proceeding of 2nd International Symposium on Parallel Algorithms/Architectures Synthesis. Japan. Mar. 1997.
- [GRI 97] GRIMSHAW, A; et al. The Legion Vision of a World-Wide Virtual Computer. *Communications of the ACM*. New York, v.40, n.1. 1997.
- [JIN 98] JING, Jin; HUFF, Karen. Adaptation for Mobile Workflow Applications. Proceedings of Workshop on Modeling and Simulation in Wireless Systems. Montreal, Canada. Jul. 1998.
- [KAT 94] KATZ, R.H. Adaptation and Mobility in Wireless Information Systems. *IEEE Personal Communications*. vol.1, n.1, p.6-17. 1994.
- [KID 98] KIDSTON, David; BLACK, J.P.; KUNZ, Thomas; NIDD, Michael E. Comma, a Communication Manager for Mobile Applications. Proceedings of Wireless '98. Canada. 1998.
- [KON 00] KON, Fábio et al. 2K: Distributed Operating System for Dynamic Heterogeneous Environments. Proceedings of the Ninth IEEE International Symposium on High Performance Distributed Computing - HPDC'00. Pennsylvania, USA. 2000.
- [KUN 99] KUNZ, Thomas; BLACK, J.P. An Architecture for Adaptive Mobile Applications. Proceedings 11th International Conference on Wireless Communications. Alberta, Canada. Jul. 1999.
- [LIT 88] LITZKOW, M. et al. Condor – A Hunter of Idle Workstations. In Proceedings of th 8th International Conference of Distributed Computing Systems. 1988.
- [MAS 01] MASCOLO, C.; CAPRA, L. XMIDDLE - A Middleware for AdHoc Networks. Research Notes UCL-CS-00/54. University College London. Jan,2001.
- [NOB 00] NOBLE, Brian. System Support for Mobile, Adaptive Applications. *IEEE Personal Computing Systems*. v.7,n.1,p. 44-9, Feb. 2000.
- [PIC 00] PICCO, Gian Pietro; MURPHY, Amy L.; ROMAN, Gruia-Catalin. A Software Engineering Perspective on Mobility. In *Future of Software Engineering: A.C.W. Finkelstein Editor*, ACM Press. 2000.
- [PIC 99] PICCO, Gian Pietro; MURPHY, Amy L.; ROMAN, Gruia-Catalin. LIME: Linda Meets Mobility. Proceedings of 21 International Conference on Software Engineering (ICSE'99). Los Angeles, USA. May. 1999.
- [RAN 97] RANGANATHAN, M.; ACHARYA, A.; SALTZ, J. Sumatra: a Language for Resource-aware Mobile Programs. In *Mobile Objects Systems: Towards the Programmable Internet*: Springer-Verlag Publisher, Serie LNCS. v.1222. Apr. 1997.
- [STE 99] STEEN, M. van; et al. Globe: A Wide-Area Distributed System. *IEEE Concurrency*. New York, v.7, n.1. 1999.
- [TAU 96] TAUBER, Joshua A. *Issues in Building Mobile-aware Applications with the Rover Toolkit*. Computer Science and Engineering. Massachusetts Institute of Technology. May. 1996 (Master Thesis).
- [VAH 98] VAHDAT, T. et al. WebOS: Operating System Services for Wide Area Applications. In Proceedings of the Seventh Symposium on High Performance Distributed Computing. 1998.
- [WEL 98] WELLING, Girish; BADRINATH, B.R. Na Architecture for Exporting Environment Awareness to Mobile Computing Applications. *IEEE Transactions on Software Engineering*. v. 24, n.5. 1998.
- [YAM 01] YAMIN, Adenauer. Escalonamento em Sistemas Paralelos e Distribuídos. ERAD 2001, Gramado, RS. SBC/UFRGS/PUCRS. Jan. 2001.
- [ZOM 98] ZOMAYA, Albert; CLEMENTS, Matthew; OLARIU, Stephan. A Framework for Reinforcement-Based Scheduling Parallel Processor System. *IEEE Transactions on Parallel and Distributed Systems*. New York, v.9, n.3. 1998.