

ARDINA - Difusão de Informação em Computação Móvel

José Orlando Pereira Rui Oliveira António Sousa Francisco Moura
{jop,rco,als,fsm}@di.uminho.pt

GSD - DI - Universidade do Minho

1 Introdução

Os *palmtops* surgiram, no início desta década, como computadores ultra-portáteis destinados a servir como agenda electrónica para a gestão de informação pessoal. A pretensão natural de alargar a sua utilização a substituto ou complemento do computador pessoal foi na grande maioria dos casos defraudada. Tanto as capacidades de armazenamento e processamento, como as características ergonómicas de um *palmtop* dificilmente viabilizam esta sua utilização.

Actualmente, com a banalização das comunicações pessoais a utilização do *palmtop* assume uma nova perspectiva como dispositivo móvel de acesso a informação por excelência. Por um lado, a capacidade de memória oferecida nas máquinas actuais é suficiente para o armazenamento da informação típica de uma agenda pessoal e de documentos de utilização transitória e por outro, as múltiplas possibilidades de comunicação oferecidas, permitem o acesso eficaz a fontes remotas para a obtenção de informação.

Nesta perspectiva, a sua aceitação e conseqüente ubiquidade, permite actualmente considerar o *palmtop* como um veículo privilegiado para a difusão e troca de informação. No entanto, os protocolos e aplicações tradicionais, demasiado exigentes em termos de recursos de *hardware* e de tipo de ligação, não são os mais adaptados a *pamltops*.

Este artigo apresenta o ARDINA, uma infra-estrutura modular de suporte à difusão de informação por intermédio de *palmtops*. A difusão de informação do ARDINA reside na exploração da habitual operação de sincronização ponto-a-ponto entre o *palmtop* e uma estação de suporte ou entre dois *palmtops*.

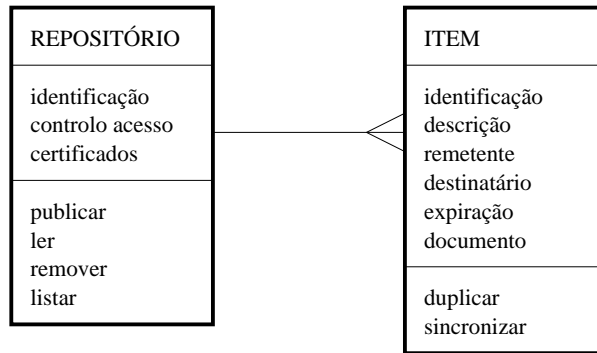


Figura 1: A arquitectura ARDINA.

2 Arquitectura

O conceito principal na arquitectura do ARDINA é o *repositório* de informação. Todos os intervenientes, sejam *palmtops* ou estações de suporte, têm associado um repositório que oferece uma interface uniforme para a sua manipulação. O tipo de informação contida num repositório é arbitrário e apenas condicionado pela possibilidade de ser manipulado pelas máquinas consideradas. À informação publicada num repositório chamamos genericamente *documento*. De um modo geral, a funcionalidade do sistema consiste na publicação e manutenção de documentos em repositórios e sincronização de pares de repositórios.

Conceptualmente, todos os repositórios são idênticos. No entanto, e porque na prática o padrão de utilização dos repositórios assim o exige, distingue-se entre repositórios de carácter *peçoal* ou *comum*. Um repositório pessoal é o que tipicamente se encontra num *palmtop*; o seu utilizador é responsável pela publicação de documentos e manutenção do repositório e tem um papel activo como moderador na sua sincronização. Um repositório comum encontra-se numa estação de suporte e é acessível a um grupo de utilizadores. A publicação de documentos num repositório comum é aberta a um grupo de utilizadores e a aplicações de angariação de informação. A manutenção do repositório é essencialmente automática podendo, ainda assim, também estar a cargo de administradores para efeitos de moderação.

2.1 Estrutura

A estrutura do ARDINA é representada na Figura 1. Um repositório ARDINA é um dicionário de *itens* que encapsulam documentos. A estrutura e conteúdo desta informação são opacos não sendo por conseguinte manipulados pelas operações do repositório. Um repositório tem como atributos a sua identificação, uma lista de controlo de acessos e outra de certificados de autenticação. Oferece como operações a publicação, leitura, remoção e listagem de itens (publicar, ler, remover e listar, respectivamente).

Um item é definido por um identificador, uma descrição (textual e arbitrária), a identificação do seu remetente e destinatário, uma data de validade, um documento, e por duas operações básicas: duplicar e sincronizar. A especificação assume que o identificador de cada item é único. Compete portanto às concretizações garantir esta unicidade ou gerir a possibilidade de conflitos.

O item é uma estrutura abstracta com o objectivo de ser reificada de forma a acomodar características particulares dos documentos necessárias ao seu processamento e sincronização. Na Tabela 1 são apresentadas tres reificações básicas de item.

2.2 Funcionalidade

A funcionalidade de um repositório resume-se às operações básicas de manipulação de um dicionário (Figura 1). Qualquer uma destas operações poderá estar condicionada à autenticação da invocação. As operações de um item deverão permitir a obtenção de um clone seu, e um item como resultado da sua sincronização com outro item similar.

A sincronização é uma operação realizada sobre pares de repositórios e tem duas fases. Primeiro, os repositórios são sincronizados assegurando-se que ambos têm o mesmo conjunto de itens.¹ Sempre que um item exista apenas num dos repositório, é criado um duplicado do item no outro repositório. Na segunda fase da sincronização, cada item de um repositório é sincronizado com o item correspondente (*i.e.*, com o mesmo identificador) no outro repositório através da operação de sincronizar de um dos itens. O item resultante desta sincronização substitui os itens com o mesmo identificador no repositório destinatário.

Como exemplo, os itens constantes seleccionados em cada um dos repositórios são se possível preenchidos com os documentos correspondentes, se estes se encontram no outro repositório.

O acesso a fontes de informação parametrizada efectua-se em vários passos. Durante uma primeira sincronização a clonagem da fonte de informação produz um item parametrizado. Numa segunda sincronização depois de preenchidos os parâmetros, este item regressa ao repositório inicial onde a clonagem o transforma num item constante ainda sem o documento, notificando ao mesmo tempo a fonte correspondente. Finalmente, quando o documento estiver disponível no repositório correspondente à fonte, pode ser copiado para o seu destino na mesma ou numa posterior sincronização.

3 Concretização

A concretização actual do ARDINA utiliza os *palmtops 3COM PALM* e a plataforma JAVA na estação de suporte. Para a difusão de informação explora-se a operação de *sincroni-*

¹Nesta fase poderá ou não ser considerada a informação sobre o destinatário de cada item.

	<i>Constante</i>	<i>Fonte parametrizada</i>	<i>Item parametrizado</i>
<i>Descrição</i>	Encapsula informação constante num formato adequado a ser armazenado e manipulado no <i>palmtop</i> .	Encapsula uma fonte de informação parametrizada.	Encapsula os parâmetros a serem submetidos a uma fonte de informação.
<i>Identificador</i>	Dependente dos dados.	Arbitrário.	Dependente da identificação da fonte e dos parâmetros.
<i>Dados</i>	Sinal indicando se o documento é desejado no repositório correspondente.	Dependente da concretização.	Uma referência à fonte, dependente da concretização.
<i>Clonagem</i>	Uma cópia do item sem o documento.	Um item parametrizado referente a este, com valores por omissão.	Caso no repositório destino exista a fonte correspondente, transforma-se num constante ainda sem o documento, do qual é produzida uma cópia para esse repositório. Caso contrário, uma cópia exacta dos parâmetros.
<i>Sincronização</i>	Caso o item esteja assinalado como desejado e o outro contenha o documento, este é copiados.	Ignorada.	Ignorada.

Tabela 1: Tipos de item na arquitectura básica.

zar o *palmtop* com o *desktop* que ocorre regularmente para o carregamento explícito de informação ou para realizar cópias de segurança.

Do ponto de vista do utilizador comum, que pretende obter informação actualizando o repositório pessoal no seu *palmtop*, a interacção com o sistema implica a sincronização com um ponto de acesso associado a um repositório comum. Este ponto de acesso à estação de suporte que contém o repositório é uma ligação série, uma ligação por infra-vermelhos ou uma ligação através de uma rede celular, deixando como tal em aberto a possibilidade de acomodar diferentes cenários de aplicação.

Depois de sincronizado com o repositório comum, o *palmtop* contém uma lista actualizada dos itens disponíveis, que o utilizador, já desligado do ponto de acesso, pode examinar e seleccionar. Numa próxima sincronização, os documentos correspondentes aos itens seleccionados são sincronizados no *palmtop* para posterior utilização. Geralmente, cada sincronização efectua ambas as operações, ou seja, actualiza o conteúdo dos repositórios e sincroniza os itens, conforme apropriado.

Em termos da arquitectura proposta, esta concretização implica compromissos bastante diferentes nos componentes destinados aos *palmtops* e à estação de suporte. No primeiro caso, os limitados recursos disponíveis e o tipo de utilização a que se destina, levam a que se obtenha uma concretização particular de uma configuração da arquitectura, tendo no entanto em consideração um conjunto mínimo de características para permitir uma fácil extensão.

Em contraste, na estação de suporte não só se realiza directamente a modularidade da arquitectura como se tira partido dessa modularidade para oferecer um conjunto versátil e poderoso de funcionalidades e ferramentas de integração com fontes de informação.

3.1 Geral

Todos os documentos contidos nos itens são ficheiros PALMOS suportados pelo sistema operativo dos 3COM PALM. Actualmente, são contemplados pela concretização os itens constantes e queries. Estes tipos de item reduzem manifestamente a complexidade de sincronização evitando a reconciliação de itens concurrentemente mutáveis.

O identificador de cada item é calculado aplicando um algoritmo de *hashing* ao documento. Deste modo obtém-se um identificador curto, adaptado à manipulação em *palmtops* e permitindo que o repositório possa ser sincronizado com vários outros sem necessidade de guardar estado adicional. Embora, esta solução não garanta a identificação unívoca dos itens no sistema, dado o habitual curto período de existência dos itens nos repositórios e a utilização de um algoritmo de *hashing* seguro, a probabilidade de tal acontecer é mínima e o risco é compensado pela sua simplicidade. É ainda de salientar que uma solução que forneça identificadores únicos não dispensa necessariamente a identificação de itens pelo seu conteúdo para efeitos de identificação de redundância .



Figura 2: ARDINA num *palm*top 3COM PALM.

O controlo de acesso aos repositórios assenta num mecanismo simples de nome e palavra-chave, em que a confiança é depositada na estação de suporte em detrimento dos *palm*tops.

3.2 No *palm*top

No *palm*top, o ARDINA é uma aplicação monolítica com uma interface gráfica que permite inspeccionar e manipular os itens do repositório (Figura 2). Nomeadamente, permite requisitar os documentos no itens constantes e preencher os parâmetros dos outros. Através de uma interface de programação simples, é aberta a possibilidade de outras aplicações gerirem itens, possivelmente de outros tipos não suportados pela concretização básica, ocultos no mesmo repositório de modo a estarem disponíveis apenas durante o processo de sincronização. A visualização e manipulação dos documentos contidos nos itens é exterior ao ARDINA e está a cargo de aplicações específicas.

Esta aplicação ocupa 20Kb e oferece uma interface gráfica simples e integrada no ambiente do *palm*top para ser utilizada pelos utilizadores finais, consumidores da informação.

3.3 Na estação de suporte

Na estação de suporte, foram definidas interfaces em Java correspondentes às diversas entidades da arquitectura bem como concretizado também em Java o algoritmo de sincronização. Este actua sobre quaisquer repositórios através da interface correspondente. Os repositórios no *palm*top são acedidos usando um *proxy*, que faz a ponte com a concretização no *palm*top.

O repositório na estação de suporte é concretizado sobre um sistema de ficheiros mantendo uma directoria separada para cada item. Para os itens constantes, esta directoria contém o ficheiro com o documento correspondente em formato PALMOS, além de um ficheiro de texto com os restantes atributos. Como tal, as operações normais sobre sistema de ficheiros podem ser utilizadas como interface de publicação e manutenção do repositório.

São também disponibilizadas um conjunto ferramentas gráficas e de linha de comando para facilitar a publicação e moderação de documentos e manutenção dos repositórios. Uma vez que funcionam ao nível da interface podem operar sobre quaisquer repositórios, incluindo sobre os pessoais.

A sincronização entre um repositório pessoal (*palmtop*) e um comum (estação de suporte) obedece à descrição da Secção 2.2. Na primeira fase, os repositórios são sincronizados com itens “ocos”, *i.e.* sem documentos. A segunda fase, que corresponde à sincronização de itens com o mesmo identificador, depende do tipo do item como descrito na arquitectura.

A eliminação de itens cujo prazo de validade tenha expirado é efectuada durante a sincronização. As interfaces de utilizador de manipulação de repositórios, com é o caso do cliente no *palmtop*, permitem esconder entradas até que expirem, de modo a simplificar a apresentação.

4 Aplicações

A utilização do ARDINA como plataforma de desenvolvimento de aplicações está intimamente relacionada com a concretização do ARDINA na estação de suporte.

Nomeadamente, é possível manter no repositório itens com diferentes concretizações colocando o código Java correspondente na directoria em questão. Esta funcionalidade é tipicamente utilizada para a publicação automática (ou não-interactiva). Esta oportunidade foi até agora utilizada para:

- efectuar a conversão de ficheiros para formato adequado ao *palmtop*;
- executar programas externos para a angariação de informação, à semelhança do funcionamento da CGI em servidores HTTP;
- efectuar consultas a uma base de dados relacional através de uma interface JDBC.

De modo a auxiliar a concretização de módulos de acesso a dados, a infra-estrutura de serviços do ARDINA é capaz de converter automaticamente estruturas de dados JAVA para a representação adequada aos *palmtops*.

5 Resumo

O ARDINA tem como principal objectivo ser uma ferramenta simples mas poderosa para a integração de fontes de informação e sua difusão integrada no procedimento habitual de sincronização. A difusão e partilha de informação evolutiva, *i.e.* passível de ser alterada e as novas versões redifundidas, conduz rapidamente à necessidade de coordenação e reconciliação das diversas versões dessa informação. Este processo é difícil, senão

mesmo impossível, de generalizar e requer procedimentos de controlo de divergência e reconciliação específicos para cada tipo de dados. Através da reificação da estrutura item, o ARDINA permite a extensão de qualquer concretização de forma a suportar o tratamento de tipos de documentos específicos.

Com o objectivo de obter uma rápida avaliação dos conceitos, a actual concretização do ARDINA considera soluções simplificadas para alguns dos aspectos da arquitectura que requerem, no futuro, maior elaboração. O esforço na concretização do ARDINA concentra-se actualmente no alargamento dos tipos de informação disponibilizada (nomeadamente o acesso genérico a bases de dados, *Web Clipping*, e calendarização de grupo) bem como o tratamento de informação evolutiva. Paralelamente, está a ser desenvolvida a sincronização directa entre repositórios pessoais que de momento é apenas possível através de um repositório comum.