

INTEGRAÇÃO DE REDES ACADÊMICAS: APLICAÇÕES E GERENCIAMENTO COOPERANTES

Liane Margarida Rockenbach Tarouco
Instituto de Informática - UFRGS
Porto Alegre - RS - BRASIL
Liane@vortex.ufrgs.br

SUMÁRIO

A organização típica das redes acadêmicas e as aplicações mais comuns nelas existentes são comentadas e é mostrada a necessidade de sistemas de informação integrados que apoiem a pesquisa de informações relevantes na rede. Uma proposta de sistema capaz de oferecer este tipo de serviço é também mostrado no presente trabalho.

ABSTRACT

The typical organization and the most common existing applications of academic networks are commented. The need for integrated information systems in support to the search for meaningful information in a network is shown. A proposal of a system for supporting this need, as a service, is also presented.

1. TENDÊNCIA NA ORGANIZAÇÃO DAS REDES ACADÊMICAS

Existe uma tendência internacional em organizar as redes de computação acadêmicas de forma hierárquica em três grandes níveis:

Nível 1: Um "backbone" nacional cujo objetivo é conectar os centros provedores de serviços especiais (tais como supercomputadores) e uma gama de redes regionais. Estas infra-estruturas devem prover um serviço muito confiável e de boa performance.

Nível 2: Redes regionais usualmente organizadas geograficamente em áreas metropolitanas, estaduais ou mesmo de regiões englobando segmentos maiores.

Nível 3: Redes internas nos campi compostas de redes locais provendo serviços para uma entidade que é responsável pela implantação, custeio e operação da mesma.

Segundo [MAN], esta foi a estrutura recomendada para a NSF (National Science Foundation) nos Estados Unidos organizar a rede acadêmica naquele país. A operação e o custeio da rede de primeiro nível deve ser responsabilidade da entidade nacional de apoio à pesquisa, NSF nos Estados Unidos, e a rede norte-americana é denominada NSFNET.

No Brasil, o CNPq está apoiando uma iniciativa similar e a RNP-Rede Nacional de Pesquisa [TAK 89] está sendo implantada para servir de backbone para as redes regionais.

Diversas universidades e centros de pesquisa no Brasil, a exemplo do que acontece nos países mais avançados, já dispõem de redes de nível 3 em funcionamento.

Para o nível 2 a situação é totalmente nova. Diferentemente do que acontece no exterior, no Brasil ainda não existe tradição na busca de soluções regionais. As primeiras redes regionais começam a ser organizadas. Estas redes, em princípio, tem como meta básica prover de forma econômica e eficiente o acesso das redes de nível 3 ao backbone nacional. Os problemas básicos na busca de tal tipo de solução envolvem questões relativas a:

- roteadores a serem usados
- arquitetura da rede
- responsabilidade pela monitoração e rastreamento de problemas
- provimento de recursos para o custeio das linhas
- serviços a serem oferecidos na rede

Alguma autoridade de nível estadual deve arcar com a responsabilidade de catalizar a organização de uma rede regional, congregando as partes interessadas na busca da solução mais apropriada para cada caso. Os estados de São Paulo (FAPESP), Rio de Janeiro (FAPERJ) e Rio Grande do Sul (FAPERGS) foram os primeiros a se mobilizarem para montar redes regionais. A atuação da Secretaria de Ciência e Tecnologia no Rio Grande do Sul, desde cedo, apoiando a criação de uma rede regional, enseja condições para que se consiga resultados relevantes. Neste sentido, o presente trabalho propõe-se a trazer para discussão alguns aspectos relevantes na implantação de uma rede acadêmica, os quais estão sendo objeto de intensos estudos visando levar a uma solução adequada. Trata-se das aplicações e isto na verdade significa os serviços que serão oferecidos na rede e do gerenciamento da mesma.

Considerando que uma rede acadêmica constitui na verdade uma cooperativa em que entidades autônomas concordam em interoperar, oferecendo e usando recursos umas das outras, é preciso buscar alguma forma de integração que por um lado mantenha a operação independente de cada instituição integrante da rede mas por outro lado assegure condições mínimas para o mútuo acesso o que só é possível com algum grau de coordenação.

No caso do Rio Grande do Sul, foi decidido desde o início que a ênfase da rede seria no sentido de apoiar o acesso ao acervo de informações em Ciência e Tecnologia existentes em cada instituição. Por isto, o que se buscou desde o início foi a viabilização de um serviço de recuperação de informações bibliográficas pois já existiam em diversas instituições sistemas on-line oferecendo este tipo de serviço.

Do ponto de vista de arquitetura de rede, optou-se pela pelo conjunto de protocolos mais usual no ambiente acadêmico, a arquitetura TCP/IP que é a única solução independente de fornecedor disponível em praticamente todos os tipos de sistemas computacionais existentes nas universidades, e tem sido adotada como a alternativa a ser usada a curto prazo.

A arquitetura OSI, padronizada pela ISO, embora mais completa

e que deve ser adotada a médio prazo, ainda não tem implementações para um número suficiente de sistemas computacionais para poder ser adotada a curto prazo.

Soluções baseadas em arquiteturas proprietárias devem ser evitadas por motivos óbvios; a ninguém interessa ficar dependente de um único fornecedor e, portanto, inibido de buscar a melhor solução ao melhor custo que o mercado, em função de uma competição sadia, pode vir a oferecer.

2. OUTROS SERVIÇOS RELEVANTES A OFERECER AOS USUÁRIOS DAS REDES

Os principais serviços usualmente oferecidos nas redes acadêmicas ou de pesquisa incluem **correio eletrônico, transferência de arquivos, acesso remoto a outras máquinas, acesso remoto a bancos de dados, intercâmbio interativo de mensagens em tempo real, conferências computadorizadas.**

Os serviços de **correio eletrônico** já estão disponíveis no país e sendo utilizados por uma comunidade que cresce rapidamente. Inicialmente as conexões foram estabelecidas usando protocolos definidos para a rede BITNET mas atualmente existe uma grande número de locais usando TCP/IP, sendo esta a modalidade preferencial. Alguns locais conseguem intercambiar mensagens usando o protocolo X.400 do CCITT mas esta forma tem sido utilizada apenas em caráter experimental pois não esta disponível em um número razoável de máquinas. Gateways conseguem converter mensagens de qualquer um dos sistemas acima referidos para qualquer um dos outros, embora nem sempre tais conversões estejam sendo possíveis no país. Pode-se concluir que qualquer instituição conseguirá, sem investimento substancial dispor e um serviço de correio eletrônico integrado inclusive a nível internacional. Existe, todavia um investimento de custo um tanto elevado que é o inerente à interconexão ao nó mais próximo para poder então dispor da conectividade ilimitada que a rede acadêmica já possibilita.

A **transferência de arquivos** já é mais limitada mas mesmo assim tem tido um uso crescente a partir da disponibilidade do TCP/IP no país. Usando o protocolo FTP (File Transfer Protocol) integrante da arquitetura INTERNET (na qual o TCP/IP está inserido) é possível transferir arquivos de muitas máquinas na rede internacional e várias delas atuam como servidores de software de livre domínio que pode ser "file-transferido" sem qualquer custo e livremente usado.

O **acesso remoto a outras máquinas** não tem sido muito utilizado pois demanda a disponibilidade de um "usercode" e senha na máquina destinatária. Mas superado este impecilho, mediante acordos bilaterais entre as organizações envolvidas, com o uso da facilidade TELNET da arquitetura INTERNET é possível a qualquer usuário "logar-se" numa máquina remota e trabalhar interativamente usando todos os serviços a que tem direito. Face à limitada velocidade usada nas conexões existentes no país, percebe-se um tempo de resposta muito elevado neste tipo de uso.

O **acesso remoto a bancos de dados** apresenta duas dificuldades: necessidade de dispor de "usercode" na máquina destinatária e conhecimento da linguagem de comando. No caso de bancos de dados destinados a prover informações bibliográficas verifica-se no exterior um maior grau de integração entre sistemas acadêmicos e comerciais que permitem a um usuário de qualquer sistema, formu-

lar consultas que são direcionadas a outro sistema usando protocolos específicos para a área, tal como o ISO 10162/10163 (Search and Retrieval for computer-to-computer information retrieval) referido em [CLI 90]. A unificação da linguagem de acesso foi a meta de um projeto conduzido pelo IBICT e CIN/CNEN resultando em um padrão de linguagem de acesso, denominada LINCE-Linguagem Comum de Acesso a Bases de Dados [IBI 89]. Este trabalho deve resultar no oferecimento de serviços com linguagem de acesso padronizada (LINCE) por um conjunto de provedores (inicialmente IBICT, BIREME, CIN/CNEN, IPT e CENAGRI entre outros). Contudo tais provedores tem a oferta de serviços inicialmente prevista para ser viabilizada através da RENPAC, para assinantes (pagantes). Isto provavelmente exclue uma parcela significativa de potenciais usuários do ambiente acadêmico e de pesquisa, que não conseguirão recursos para custear o acesso RENPAC e os serviços de informação ofertado por estes provedores.

Por outro lado, diversas universidades estão implementando seus próprios serviços de recuperação de informações bibliográficas colocando on-line pelo menos as referências sobre o acervo produzido ou disponível em cada instituição. Tais implementações não necessariamente seguem alguma padronização, usando, ao contrário, os mais variados softwares.

O intercâmbio interativo de mensagens em tempo real é outro serviço pouco utilizado e depende de acesso simultâneo entre os usuários que se comunicam. Em países desenvolvidos, onde os pesquisadores trabalham praticamente todo o tempo em suas estações de trabalho e estas estão conectadas a redes de nível 3, é mais provável que ao tentar conexão com um usuário remoto o mesmo esteja acessível naquele momento. No Brasil, todavia, a realidade é muito diferente e, via de regra, os pesquisadores somente estão on-line durante um período de tempo limitado, quando saem de suas sala e dirigem-se ao local onde estão os terminais ou estações de trabalho utilizadas para acessar suas caixas postais eletrônicas. Isto reduz a probabilidade de encontrar alguém on-line para intercâmbio interativo, sem previa combinação.

As conferências computadorizadas implantadas através de um mecanismo de listas de distribuição permitem que cada contribuição sobre tópicos em discussão numa dada lista seja enviada a todos os usuários inscritos naquela lista. O recebimento efetivo da mensagem contendo aquela manifestação somente ocorrerá quando o destinatário acessar sua caixa postal. Assim, tais conferências não são usualmente realizadas em tempo real, ao contrário, cada usuário recebe as contribuições, tem tempo para lê-las atentamente, pensar a respeito, efetuar alguma pesquisa para referendar sua eventual resposta e elaborá-la sem pressa. Este tipo de serviço tem sido responsável pela maior parte do tráfego de mensagens nas redes acadêmicas. A quantidade e a diversidade das conferências realizadas desta maneira é quase inacreditável.

3. A PROPOSTA PARA UM SERVIÇO DE INFORMAÇÕES INTEGRADO

Face ao que foi acima exposto, percebe-se que existe uma lacuna a ser preenchida: existe necessidade de prover melhor acesso aos usuários da rede acadêmica a serviços de informação. O acesso ao acervo bibliográfico existente em cada instituição é um dos mais importantes serviços que a rede pode viabilizar e o que não está sendo satisfatoriamente atendido.

O acesso à informação contida neste acervo é uma necessidade premente e isto não é uma colocação de validade regional ou inerente apenas ao momento atual. Ao contrário, desde há muito tempo, a comunidade científica percebeu a necessidade de acesso otimizado à informação, conforme destacado por [PEN 88] em seu relato sobre a montagem da rede de bibliotecas na AT&T. Desde que a Bell Laboratories abriu suas portas pela primeira vez, em 1925, foi reconhecida a importância de ter a informação certa no tempo certo, na forma certa. Os cientistas tem sido constantes em sua demanda por informação pronta, adequadamente apresentada. O desafio para os especialistas em informação tem sido sempre o de como alcançar o objetivo e não sobre qual o objetivo a alcançar. Uma biblioteca adequadamente abastecida tem sido tradicionalmente vista como o mecanismo apropriado para atender às necessidades da comunidade científica. Contudo é preciso que seu conteúdo seja adequadamente disseminado na comunidade potencialmente interessada em seu uso. O acesso on-line a bases de dados contendo referências bibliográficas ou mesmo os próprios documentos, como no projeto Mercury da Universidade de Carnegie Mellon [TRO 91], torna-se então a forma básica de atingir esta meta.

A força inicial movendo as bibliotecas na direção das redes adveio do desenvolvimento da tecnologia que permitiu o uso dos computadores em time-sharing. Isto forneceu os meios para disseminar o acesso às informações, via teleprocessamento, para usuários inclusive remotos. Mais tarde houve uma evolução para uma forma distribuída de processamento na qual, mesmo que diferentes computadores fossem usados para gerenciar um segmento do acervo, redes integrando os serviços providos por cada uma, permitiam aos usuários de um serviços acessarem as informações de outros.

Na fase inicial, foram implementados tipicamente serviços de controle de circulação de material. A UNB foi, talvez pioneira no país na implantação e operacionalização deste tipo de aplicação. Posteriormente, surgiram no país sistemas de recuperação de informações bibliográficos de acesso interno ou mesmo público, tal como o sistema de informações jurídicas do Senado Federal ou o serviço oferecido pela BIREME-Biblioteca Regional de Medicina que coleta e organiza informações sobre publicações na área médica tanto a nível nacional, como na América Latina, além de contar com referências sobre publicações a nível internacional importadas de provedores deste tipo de serviços no exterior. Vários serviços públicos de informação foram implantados no Brasil, cada qual usando sua própria linguagem de acesso o que dificultava seu uso pois cada usuário era obrigado a aprender a linguagem de comando de cada sistema.

Em 1986 o IBICT empreendeu um esforço no sentido de encontrar uma solução integradora para os serviços de informações oferecidos no país. Um trabalho realizado na oportunidade [TAR 86] identificou inicialmente uma série de serviços típicos passíveis de serem oferecidos e que demandariam a interconexão dos sistemas. A definição destes serviços foi baseada num trabalho preliminar desenvolvido no Canadá. Conforme definido pelo Task Group on Computer/Communication Protocols for Bibliographic Data Interchange no Canada [NAT 83], o fluxo de transações que flue de/para uma biblioteca típica pode ser representado pelo diagrama a seguir:



FORNECEDORES (material, serviços)

1. Formulários de pedidos, pesquisas a catálogos, pedidos de circulação de periódicos
2. Formulários de compras preenchidos, localização, serviço de circulação
3. Pedidos de informação, de pesquisas, de documentos
4. Informação, disponibilidades, empréstimos etc...
5. Pagamentos, ordens de compra, notificações, pesquisa de cópias catalográficas
6. Material, faturas, cópias catalográficas, etiquetas etc
7. Pesquisa de citações, pedidos de localização, pagamentos
8. Citações, localizações, faturas etc...
9. Empréstimos, pedidos de empréstimos, notificações etc...

Estes serviços podem ser apoiados por sistemas automatizados, conforme discutido em [CLI 90 e TUR 90]. Existem na realidade diversos sistemas implantados no exterior, como por exemplo, o da "Library of Congress" nos Estados Unidos que juntamente com duas outras organizações provedores de serviços de automação de bibliotecas implantaram os protocolos Z39.50 (versão da ANSI para o Search and Retrieval for computer-to-computer information retrieval) referidos em [CLI 90] e o ILL-Interlibrary Loan Service and protocol referido em [TUR 90].

Os serviços numerados de 1 a 4 são providos aos usuários e alguns (pedidos de informações e pesquisas, por exemplo) são oferecidos por provedores de serviços de bancos de dados não bibliográficos também. Estes foram os serviços que o grupo coordenado pelo IBICT optou por implantar de forma padronizada no país, numa primeira fase.

A partir de pesquisa junto a alguns provedores de serviços de bancos de dados oferecidos no Brasil, constatou-se que varios deles já prestam uma parcela de tais serviços em modalidade on-line, embora alguns apenas em uso interno. A modalidade de interação usuário-sistema, varia de caso a caso, embora não substancialmente. Os recursos computacionais disponíveis para o provimento de serviços on-line são um tanto limitados em grande parte dos casos.

O uso de serviços de bancos de dados poderia ser mais amplo,

vesse disponível através de um módulo de interface (terminal T*). Para o encaminhamento de pedidos de serviços a outros provedores, o programa que atende aos requisitos comandados na linguagem nativa (AN-Aplicação Nativa) encaminharia tais solicitações a um programa capaz de prover o interfaceamento (MI-Módulo de Interfaceamento) com os sistemas remotos. O MI, também aceitaria requisitos de serviço comandados usando a linguagem padronizada (linguagem de comando comum). O MI poderia, alternativamente, ser implementado em processador diferente daquele no qual é prestado o serviço de banco de dados.

Seriam usados os protocolos ISO 10162/10162 (Documentation - Search and Retrieve Service Definition and Protocol Specification) e ISO 10160 (ILL-Interlibrary Loan) para implantar a interconexão entre os sistemas. Embora estes protocolos correspondam à camada de aplicação do modelo OSI é possível utilizá-los sobre outras pilhas de protocolo, tal como a do TCP/IP, a exemplo do que é feito na Universidade de Carnegie Mellon [TRO 90].

Por outro lado, considerando-se a tendência atual de utilizar em Teleprocessamento, terminais inteligentes, dever-se-ia usar microcomputadores ou mesmo estações de trabalho para prover o serviço padrão (T*). Em termos de custo, há casos em que um microcomputador com unidade de disco flexível, custa menos do que um terminal. A disponibilidade de inteligência e armazenamento secundário junto ao usuário poderia ser aproveitada para fins de treinamento. Um software de acesso a banco de dados em microcomputador, usando a linguagem de acesso comum seria tornado disponível para este contexto.

Tal software permitiria não apenas treinamento off-line do usuário como também que pesquisas por ele requisitadas, fossem enviadas em meio magnético ou por transferência de arquivos através da rede. Este software também poderia ser utilizado pelo usuário para criar seus bancos de dados pessoais, catalogando os documentos de uso pessoal (publicações lidas ou escritas por ele, etc...) estimulando, deste modo o hábito de manipular informação e documentos de forma organizada e automatizada (apoiada pelo microcomputador).

Dependendo do porte do usuário, uma configuração mais robusta (microcomputador com disco Winchester), poderia permitir até mesmo ampliar o oferecimento do serviço a uma gama um pouco maior de usuários. Cada usuário prepararia sua pesquisa no micro, criando uma consulta sintaticamente correta (verificada localmente). Esta consulta seria então transmitida para o provedor do serviço de recuperação de informações, diretamente numa comunicação processo a processo, usando um protocolo padronizado (tal como o Z39.50 referido em [CLI 90]). Em função do perfil do usuário, catalogado e possivelmente modificado por ele próprio, on-line, arquivos com registros recuperados na base de dados poderiam ser enviados num formato padrão, tal como o ISO 2709 (Documentation - Format for bibliographic information interchange on magnetic tape)

Uma forma alternativa para o envio de dados de/para o Módulo de Interfaceamento seria utilizando um sistema de mensagens baseado em computador, ao invés do acesso on-line interativo e da transferência de arquivos. Nesta alternativa, toda a interação seria transportada por mensagens que, uma vez recebidas, seriam enfileiradas até que seu atendimento fosse possível. Este atendimento poderia ser feito por uma pessoa, conforme sugerido em [TAR

software foi recentemente instalado na UFRGS e esta sendo testado no âmbito do projeto REDURGS. Ele roda sobre serviços X.25 e sobre TCP/IP (em redes locais ou de longa distância).

Visando dar início ao desenvolvimento coerente de esforços no sentido de tornar operacionais todos estes serviços, não apenas no estado mas num âmbito mais amplo, está sendo iniciado um projeto que busca a implantação integrada de todos estes serviços. Este projeto será desenvolvido de forma conjunta por equipes de várias universidades : UFRGS, UFSC, PUC/RS e CEFET/PR. O projeto é denominado KNOWBOT e está sendo submetido ao CNPq para obtenção de suporte ao seu desenvolvimento. O projeto será resumidamente descrito a seguir.

4. PROJETO KNOWBOT

Knowbots são robots especializados em procurar informações digitalmente armazenada buscando compor um relatório com toda as informações relevantes de modo a atender uma determinada solicitação do usuário. A informação recuperada é organizada de modo a compor um resultado sem redundâncias. Os knowbots constituirão elemento fundamental de apoio aos pesquisadores auxiliando-os na pesquisa bibliográfica que costuma consumir uma parcela importante de seu tempo com produtividade baixa devido à falta de dados sobre a localização das informações desejadas.

Trata-se de uma nova área de pesquisa que começa a ser desenvolvida em alguns países e que pode ser adaptada e testada em nosso meio. O sistema ora proposto deverá operar em consonância com o que existe ou deva existir no contexto da RNP- Rede Nacional de Pesquisa [TAK 89] do projeto Sistema Público de Acesso a Bases de Dados do IBICT [IBI 89].

Neste sentido, o projeto KNOWBOT, pretende realizar uma implementação experimental interativa, de um conjunto de soluções capazes de propiciar um nível mínimo de suporte à pesquisa bibliográfica e apropriação de informações para os pesquisadores. O resultado final do sistema a ser projetado e implementado será o oferecimento de serviços de um KNOWBOT.

Inicialmente, prevê-se que o sistema será composto de vários módulos, tal como ilustrado na figura seguinte:

83] ou mesmo por um processo.

No primeiro caso, uma pessoa seria designada em cada instituição a auscultar periodicamente sua caixa postal e atenderia à solicitações recebidas, mediante consulta a arquivos manuais ou automatizados (acessando o MI via um terminal T*), gerando uma mensagem com o resultado das consultas solicitadas.

Na biblioteca setorial do Instituto de Informática da UFRGS, uma das bibliotecárias já presta um serviço similar, atendendo a consultas encaminhadas via correio eletrônico. Para atender às consultas ela acessa tanto o sistema on-line já implementado na UFRGS (SABI) quanto o arquivo não disponíveis on-line (listagens) ou publicações contendo referências bibliográficas. Ela pode, também, acessar bases de dados externas à Universidade para recuperar informações para algum usuário da biblioteca. Nesta modalidade, o usuário não precisa conhecer qualquer mecanismo de acesso a bases de dados internas ou externas e pode formular suas consultas usando linguagem natural pois está interagindo com um intermediário humano.

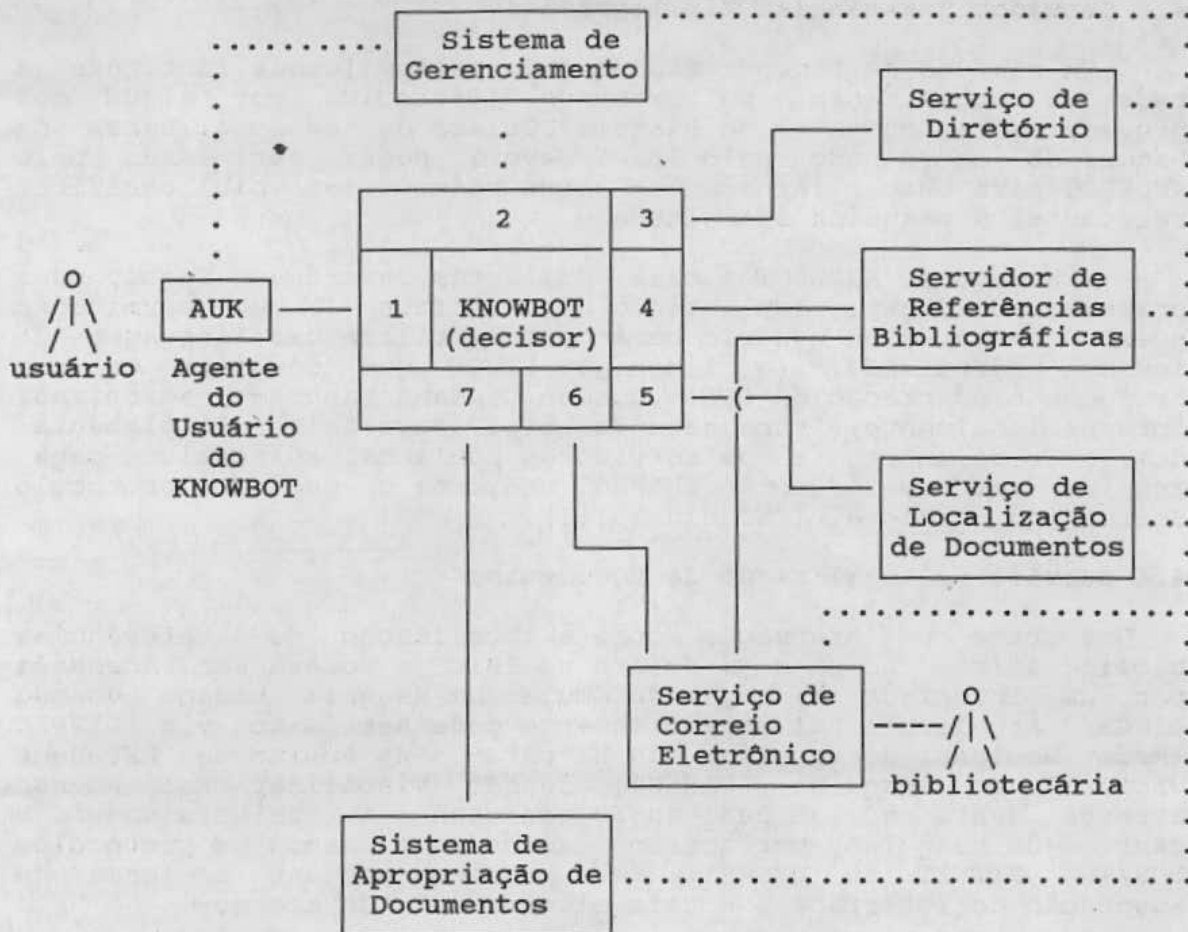
No segundo caso, o pedido de pesquisa seria encaminhado como corpo de uma mensagem destinada a um processo servidor no computador destino, o qual processaria a consulta e remeteria a resposta como corpo de uma mensagem destinada ao originador da consulta.

Destaque-se que esta última modalidade de acesso já está em uso no Brasil. Na FAPESP, um antigo sistema on-line de informações sobre status de processos pode atualmente ser também acessado através de mensagem que os interessados compõem de acordo com certas regras e remetem a um endereço específico. Uma modalidade similar de solicitação de informações já foi também implementada pelo IBICT mas não teve seu uso disseminado devido a problemas com o computador no qual o processo provedor foi implementado.

Adicionalmente à operacionalização de alguma modalidade de acesso a bases de informações bibliográficas é preciso implementar um serviço de diretórios que indique a existência e a localização na rede das diversas fontes/pessoas a serem acessadas na busca de informações. A implementação de um serviço de diretórios deve ser pautada pela orientação atual a nível internacional que implica na utilização das normas definidas nas recomendações da série X.500 do CCITT, tal como descrito em [TAR 91].

A vantagem de um serviço de diretórios deste tipo é que ele pode ser implementado de forma distribuída, com cada instituição criando e mantendo a base de dados com as informações sobre seus próprios recursos e as consultas cruzadas (de alguém de uma instituição concernente a recursos em outra instituição) podem ser atendidas mediante replicação local dos dados de algumas outras instituições ou mediante re-encaminhamento da consulta ao provedor provável ou ainda mediante mera informação sobre para onde o usuário deve dirigir uma outra consulta para poder ser atendido.

Destaque-se também que existe software de livre domínio capaz de oferecer um serviço de diretórios coerente com a recomendação X.500. Trata-se do ISODE [ROS 89] que pode ser "file-transferido" pela rede e que roda em ambiente UNIX propiciando uma pilha OSI (um tanto simplificada) e algumas aplicações (QUIPU-serviço de diretório e FTAM-File Transfer Access and Management). Este



4.1 AUK-Agente do Usuário do KNOWBOT

Usando o modelo de sistema distribuído não simétrico (cliente servidor), o processo AUK tem por finalidade apoiar o usuário humano na criação de uma consulta. Isto implica prover um interface ergonômico para facilitar sua interação e evitar que construa consultas erradas. Para isto deverá utilizar, possivelmente, um dicionário local para validar as solicitações do usuário e gerar uma consulta sintaticamente correta e compactada a qual será enviada para o KNOWBOT para atendimento. O interface 1 receberá tais mensagens, de acordo com um protocolo projetado especificamente para tal situação.

4.2 Serviço de Diretório

Um serviço de diretório, de acordo com a recomendação X.500 do CCITT é constituído por um ou mais ASD-Agentes do Sistema de Diretório que gerem um conjunto distribuído de informações sobre os recursos da rede e sua localização. Isto inclui os servidores de informação a serem consultados pelo KNOWBOT.

Portanto o KNOWBOT terá um interface (3) para emular um AUD-Agente do Usuário do Serviço de Diretórios e interagir com um ASD usando o protocolo definido na X.500.

4.3 Servidor Referências Bibliográficas

Um sistema de recuperação de informações locais (internas à rede da organização) ou externos (oferecido por algum dos provedores integrantes do Sistema Público de Acesso a Bases de Dados do coordenado pelo IBICT deverá poder ser usado pelo KNOWBOT para buscar informações sobre referências bibliográficas relevantes à pesquisa solicitada.

O acesso do KNOWBOT a tais servidores deverá ser feito, numa primeira bordagem, implantando um interface (4) que permita o acesso emulando um usuário humano, que utiliza uma linguagem de acesso padronizada, (a linguagem LINCE que consiste em uma tradução e adaptação da CCL - Common Command Language padronizada internacionalmente). Numa segunda etapa, deverão ser implementados, na rede interna e nos servidores públicos, aplicativos capazes de interoperar com o KNOWBOT mediante o uso do protocolo Search and Retrieval (ISO DIS 10163).

4.4 Serviço de Localização de Documentos

Uma base com informação sobre a localização de referências bibliográficas no país já existe no IBICT e poderá ser acessada por um interface (5) capaz de emular um usuário humano usando LINCE. Atualmente tal acesso somente pode ser feito via RENPAC (Rede Nacional de Pacotes) mas há estudos no âmbito da RNP-Rede Nacional de Pesquisa [TAK 89] visando viabilizar tal acesso através deste rede, à qual estão conectadas as universidades e centros de pesquisa, por acessos dedicados e usando os protocolos TCP/IP, BITNET ou DECNET. Será preciso estudar a forma de adaptação do interface 5 a tais alternativas de acesso.

4.5 Correio Eletrônico

O Sistema KNOWBOT deve ser capaz de gerar mensagens dirigidas às bibliotecárias responsáveis pelo acervo que necessita acessar, usando um serviço normal de correio eletrônico (interface 6) disponível na Rede Nacional de Pesquisa.

As mensagens intercambiadas com elas corresponderão às interações previstas no protocolo ILL-Interlibrary Loan (ISO DIS 10161) mas traduzidas para um formato compreensível pelas pessoas que as receberão. O KNOWBOT também deverá ser capaz de analisar a respostas e enquadrá-las em uma das possíveis respostas previstas neste protocolo. Caso não consiga, a mensagem pode ser redirecionada para um operador especializado, juntamente com as mensagens precedentes que parecerem estar relacionadas para que este possa analisar a situação e dar uma solução.

4.6 Sistema de Apropriação de Documento

Este sub-sistema deverá prover um contexto que permita o acesso ao conteúdo dos documentos. Um scanner deverá ser usado para apropriação de imagens de documentos. Tais imagens serão também analisadas (via um sistema OCR-Optical Character Recognition) e um processo de indexação semi-automático (possivelmente completado por um operador especializado) permitirá alimentar uma base de dados que, além de conter as referências bibliográficas do documento apropriado, conterá também dados sobre a localização da imagem do documento. assim,

quando é recebido um pedido de documento, o sistema primeiro verifica se o mesmo já não foi apropriado anteriormente. se for necessário apropriá-lo, o sistema pede ao operador que aproprie o texto desejado, analisa-o, indexa-o, guarda uma cópia do mesmo e envia outra para o KNOWBOT, usando um mecanismo de transferência de arquivos e estruturando-o de acordo com o Formato para Transferência de Informação Bibliográfica (ISO 2709).

Adicionalmente, este sub-sistema poderá obter uma cópia dos documentos de fornecedores comerciais ou acadêmicos (os próprios autores), no país ou no exterior e aplicar sobre esta cópia procedimentos similares aos usados com as imagens de documentos (analisar, indexar, guardar e remeter ao KNOWBOT os solicitados).

A comunicação entre o KNOWBOT e este sub-sistema deverá ser estabelecida usando o protocolo ILL-Interlibrary Loan [CLI 90].

4.7 Gerenciamento do Sistema

Cada componente do sistema deverá ser dotado de capacidade de auto-gerenciamento. Isto significa que deverá observar seu próprio funcionamento bem como as interações com seu pares e registrar dados relativos às ações e eventos inerentes à sua atividade. Algumas informações serão repassadas para um gerenciador global, outras serão mantidas apenas em escopo local (filtros determinarão as informações a serem propagadas).

As informações serão repassadas pelo KNOWBOT (interface 2 e ligações representadas por linhas pontilhadas) e pelos demais sub-sistemas usando possivelmente o protocolo SNMP previsto na arquitetura INTERNET (TCP/IP) [ROS 91].

5 CONCLUSÕES

O que se pretendeu mostrar neste trabalho foi um apanhado de esforços dirigidos à organização dos serviços a serem oferecidos nas redes acadêmicas. Neste sentido, foram apresentados os principais serviços e os padrões aplicáveis em cada caso, bem como uma forma de integrá-los e gerenciar o funcionamento da rede como um todo.

REFERÊNCIAS BIBLIOGRÁFICAS

- [IBI 89] IBICT. Sistema Público de Acesso a Bases de Dados. PADCT II. Sub-Programa de Informação em Ciência e Tecnologia. IBICT. 1989
- [MAN] MANDELBAUM, R & MANDELBAUM, P. The strategic future of the middle-level networks. University of Rochester
- [BUC 80] Buchinski, Edwin & ISLAM, M. The Context of Interconnection for a nation-wide bibliographic network. National Library of Canada., Ottawa, June 1980
- [CLE 82] CLEMENT, H. The resource network: major issues facing canadian libraries. National Library of Canada., Ottawa, September 1982
- [NAT 83] National Library of Canada. Task Group on Computer/Communication Protocols for Bibliographic Data Interchange- Interim Report, Ottawa, October 1983.
- [TAR 83] TAROUCO, Liane. Sistemas de Mensagens por Computador para Intercambio de Informação e Experiência. In: Encontro do Grupo de Trabalho sobre Interconexão de Sistemas, 3. Patrocinio UNESCO, Manilla - Filipinas - 10-20 Outubro, 1983.
- [PEN 88] PENNIMAN, David. The Library Network at AT&T. In: Sci-Tech Library Networks Within Organizations. Haworth Press, New York, 1988
- [TAR 86] TAROUCO, Liane. Definição de um Sistema Público para Acesso à Base de Dados. Estudo realizado para o IBICT. Porto Alegre, 1986.
- [ROS 89] ROSE, Marshall. The ISO Development Environment. The Wollongong Group, Palo Alto-CA, 1989
- [TAK 89] TAKAHASHI, Tadao. RNP - Rede Nacional de Pesquisa: Plano de Atividades. SECT/PR, Novembro 1989.
- [WIL 89] WILKENS, Maria & TAROUCO, Liane M.R. O Modelo OSI na UFRGS. Informática SUCEsu 89 - XXII Congresso Nacional de Informática. 18-22 setembro 89. São Paulo.
- [CLI 90] CLIFFORD, Lynch. Information Retrieval as a Network Application. Library HI TECH, Volume 8 Number 4 Issue 32, 1990
- [TUR 90] TURNER, Fay. The Interlibrary Loan Protocol: An OSI Solution to ILL Messaging. Library HI TECH, Vol.8 N.4 Issue 32, 1990
- [TRO 90] TROLL, Denise. Library Information System II- Progress Report and Technical Plan. Mercury Technical Report Series, Number Three. Carnegie Mellon University, Pittsburgh, 1990.
- [TAR 91] TAROUCO, Liane. Serviços de Diretórios. Relatório Técnico elaborado para a RNP, Março, 1991