

経済経営研究所における計算機環境について

神戸大学 経済経営研究所

安田 豊

はじめに

経済経営研究所(以後単に研究所)は設立以来過去半世紀以上の歴史を有しています。計算機導入の歴史も古く、1970年には HITAC10 を導入して統計データ処理を行なっています。それ以来、継続的に計算機設備の更新をおこない、常に実用性の高いデータベースサービスを利用者に提供してきました。

本稿では、研究所の計算機環境について、特にネットワークサービスとデータベースサービスを重点的に紹介します。研究所における計算機環境の構築と運用は、そのほとんどを機械計算室(以後単に機械室)というセクションが担当しています。この機械室が提供するサービスは、主としてネットワークサービスとデータベースサービスの二つに分かれます。そのネットワークサービスとは、一般的にインターネットで行なわれている電子メールや WWW などのサービスや、プリンタ共有、ディスク共有などにあたります。そしてデータベースサービスとは、機械室が独自に開発したデータベースシステムを用いて行なっています。これは研究所が購入しているデータを、利用者がより簡単にアクセスできるようにするものです。

研究スタッフには一人一台の PC が割り当てられており、これが各種のサービスを受けるための窓口端末となります。これら全てのサービスと設備には、機械室スタッフによる継続的なサポートとメーカーによる保守が含まれています。

現有の設備は 1996 年に導入された汎用機とワークステーション、PC 混在環境ですが、2000 年の設備更改によってより先進的な計算機環境を構築する予定です。

1. ネットワークサービス

研究所のネットワークサービスは、電子メールなどインターネットで一般に行なわれている開かれたネットワークサービスと、研究所スタッフのための資源共有を目的とした所内向けサービスの二種類に大別されます。利用者数が少なく、管理者も少人数なので、Sun Workstation 一台でほとんど全てのサーバを兼ねさせています。つまりサーバの能力よりも全体のサービス品質の方が重要なのです。

設備の性能やその数量については <http://www.rieb.kobe-u.ac.jp/> 以下を見ていただければわかるでしょうから、以下では具体的にどのようなサービスが行なわれているのか示します。

1. 開かれたサービス

(a) 電子メール

研究所のスタッフなら誰でも foo@rieb.kobe-u.ac.jp というメールアドレス (foo 部分は重複しない範囲で希望するユーザ名)の発行を受けることができます。メイリングリストも必要に応じて作成しています。

最近の研究者は遠隔地の共同執筆者と、Microsoft Word ファイルを添付したメールを往復させながら論文執筆をすることも多く、メイルプールは常に大きめに確保してこれをサポートしています。

また、電子メールはいまや所内の事務連絡に不可欠のものとなっており、全研究スタッフ、全事務スタッフ宛などの同報通信が毎日のように行なわれています。

メイルユーティリティには Microsoft Outlook を使うようにしていますが、一般的な SMTP/POP プロトコルを利用しているので、互換性のあるものなら何でも使えます。

(b) WWW

<http://www.rieb.kobe-u.ac.jp/>が研究所 WWW サイトのトップページです。希望する利用者は個人ページを持つこともできます。最も頻繁に更新されるページはおそらく研究会情報でしょう。日々追加される研究会のスケジュールなどを、研究助成室のスタッフが更新しています。また、研究所の図書掛や附属経営分析文献センターのスタッフは、WWW 上でワーキングペーパー検索システムや所蔵雑誌の目次情報などを公開しています。これらのサービスのために何らかのプログラミングが必要な場合は、ほとんど機械室が行ないます。所内向けサービスもありますので、そちらも御覧下さい。

(c) anonymous ftp

<ftp.rieb.kobe-u.ac.jp> で、anonymous ftp サービスを提供しています。幾らかの限定されたスタッフには ftp ディレクトリの書き込み権を与えており、自分が必要になったファイルをネットワークから取得した場合は、ここに置いて再利用するように心がけています。

2. 所内向けサービス

(a) WWW

WWW は本来情報を公開するためのサービスですが、適切なアクセス制限を設定することによって、限定されたスタッフだけがアクセスできる領域を設けることができます。この機能を用いると、ローカルな情報提供サービスが実現できます。例えば助成金の情報などが随時更新・提供されています。

また、後に示すデータベースサービスは、この限定アクセスの WWW ページを利用して提供しています。

(b) ダイアルアップ接続サービス

ネットワークサービスの利用価値を高めるにつれて、それは研究者にとって必要不可欠なものとなるでしょう。すると今度は研究所ネットワークサービスが受けられないところでは研究が進まないという状況になってきます。つまり今の研究者はいつでもどこでもネットワークサービスが必要なのです。これを実現するためにダイアルアップ接続のための設備を導入し、自宅や出張先からでも研究所ネットワークサービスが受けられるようにしています。これを利用して先に示した所内向け WWW サービスも利用できます。

機器としては Ascend 社の MAX1800 を採用して PPP による TCP/IP 接続を行なっています。ユーザ認証には Radius を利用しています。

(c) ディスク共有

ディスク共有サービスを利用すると、サーバマシンのディスクを研究室の PC の増設ディスクのように扱うことができます。例えば研究助成室のスタッフが WWW ページを更新する時に、この機能を利用しています。つまり研究助成室のパソコンからは、WWW ディレクトリの一部が見えており、WWW ページ作成ソフト(Adobe PageMill)を利用してファイルを直接更新するのです。これによって Unix のコマンドや ftp 転送の方法をおぼえる必要もなく、誰でも安全簡単に WWW ページの更新ができます。

他にも研究会のドキュメントを、担当教官と担当助手の間で共有ディスクを経由して編集したり、すでに示した研究所の図書掛や附属経営分析文献センターのサービス データ更新などでも活用しています。

技術的には NFS を利用しています。Samba による SMB も試したのですが、NFS Maestro という PC NFS クライアントソフトの出来が良かったので、こちらを採用しました。

(d) プリンタ共有

両面印刷が出来る PostScript プリンタや、OHP シートに非常にきれいに出力が可能なカラープリンタなどは、それぞれ高価なうえに利用頻度が低く、個人で買うのは難しいものです。そのような場合はネットワーク対応のプリンタを研究所全体で購入するなどして設備を共用しています。技術的には lpr を採用しています。先にあげた PC NFS ソフトが lpr クライアントも含んでいましたので、それを使っています。

ざっと見ていただいたとおり、研究所のネットワークサービスはまるでミニ・プロバイダのようです。そういう視点で見た場合の研究所サービスの最大の特徴は、そのきめの細かさでしょう。例えば相談を受けて、出張時のメール転送処置などを管理者側で行なうこともよくあります。利用者数が少なく個別対応が可能なことから、そうした細かな対応がある程度行なえるのです。

ところでサーバマシンに telnet login して Unix システムやサーバ上のアプリケーションを利用することはあまり想定していません。例えば SPSS の利用など、CPU 資源を要求するような仕事は手元の PC か総合情報処理センターのサーバマシンを使う方が効率が良いからです。最近で

はサーバマシンより最新の PC の方が CPU 能力が高い場合が多いのでこうした作業を集中ホストで処理する意味がなくなりつつあります。

そしてこれらのサービスは、トラブルがない限り 24 時間 365 日の連続運転でユーザに提供されています。もちろんユーザも連続運転を望んでいますから、トラブル対策には力を入れています。研究所では、全てのサーバ群は基本的な保守契約を結んでおり、故障があっても短時間で必ず復旧できる体制をとっています。ユーザデータの保護に関しては、サーバのディスクを RAID 5 ディスクアレイで構成することによって故障に備えています。万が一のデータ喪失に対応するために DLT テープドライブを用いたバックアップを 毎晩自動的にとっています。全てのサーバ機器には無停電電源装置が付いており、停電時には自動的にサーバを シャットダウンさせ、復旧時には自動的に全てのサービスが立ち上がります。管理者が一人なのでこうした自動運転は必須です。

また、ネットワークサービスではありませんが、高速プリンタのような共用した方が効率の良い設備については、機械室の共有スペースに設置して有効利用をはかっています。CDR、イメージスキャナ、OCR ソフトなどです。

2. データベースサービス

2.1 RIEB データベースシステム

経済・経営データは多くの会社から販売されていますが、そのフォーマットやメディアはまちまちで、そのままでは統一的に扱う手段がありません。多くの研究者は慣れないコンピュータ用語が並んだデータ仕様書を前に、苦労して自分が必要なデータを抽出していることと思います。研究所では需要の多いデータについて、利用者が統一されたインタフェイスと フォーマットでアクセスできるようにしています。このシステムは機械室が開発し、継続的な更新作業も行なっています。これが研究所のデータベースサービスです。

このサービスは 1970 年代から、常に研究所内で開発したシステムによって行なわれてきました。今も SECRETARY という汎用機上で開発されたシステムによるデータベースサービスが提供されています。これは簡易なコマンドでさまざまなデータを取り出すことが出来るものです。

現在では汎用機から Unix + PC 環境への移行のために、SECRETARY にかわる新しいシステム (以後 RIEB データベースシステム)を開発・運用しています。そこでは一切のコマンド操作を廃し、WWW インタフェイスを操作するだけで各種のデータを取り出すことができます。例として多国籍企業データの検索インタフェイスをこちらに示しておきます。

多国籍企業データベース
 (特定企業抜き出し済み)

資料：(表は TAB区切りになっています)
 [項目一覧] [国名一覧] [業種コード表] [親企業コード表, (指定62社コード表)]

出力する項目を指定して下さい。
 項目名に隣りチェックボックス(☐)にチェックをすると、条件指定が有効になります。文字列を指定する場合は“天津伊勢丹(有)”などのように半角の引用符で囲んでください。
 最後のチェックボックス(☑)のチェックを外すと、その項目は出力から除外されます。

年度	<input checked="" type="checkbox"/>	=	▼	1997	<input checked="" type="checkbox"/>
国コード	<input checked="" type="checkbox"/>	=	▼	304	<input checked="" type="checkbox"/>
社名英語	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
業種名	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
事業内容	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
資本金額	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
資本金単位	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
資本金通貨	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
従業員合計	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
(無し)	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
(無し)	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>
(無し)	<input type="checkbox"/>	=	▼		<input checked="" type="checkbox"/>

日本側出資企業項目を含める (最大16社)
 701200 日本側出資企業指定 (463100など6桁数字)

現地側出資企業項目を含める (最大4社)

検索結果を：
 画面に表示
 メールで送る
 E-mailアドレス：

出力行数制限： 先頭から 100行 先頭から 1000行 制限無し

Query Clear

図 1: 多国籍企業データ WWW インタフェイス

この WWW インタフェイスを用いて各種の条件設定、項目選択を指定したのち、Query ボタンをクリックすることで、即座に希望のデータを抽出することができます。条件設定に必要な項目内容の説明やコード表も、この WWW ページから得られます。

抽出結果は WWW ブラウザ上で見ることもできますが、これはもっぱらプレビューのためにあります。ここで抽出結果を見ながら条件設定を調整し、最適な条件が設定できれば出力データを電子メールの添付ファイルとして送信するか 共用ディスクへ書き込むのか指定して抽出操作を再実行し結果を受けとります。

出力データはTAB区切りのテキスト形式によるMicrosoft Excelドキュメントで、研究所スタッフのPCではメール送信と共用ディスク書き込みの、いずれの方法をとった場合ではExcelドキュメントのアイコンとして見えます。ユーザはこれをダブルクリックするだけで、簡単にデータの分析処理に移行できるのです。

2-2 コンセプト

従来的な経済・経営データベースシステムではサーバ側にデータ分析機能が含まれている場合があります。しかし最近のユーザは分析作業を普段利用している手元のPCの分析ツールで行なうことが多く、PCの劇的な性能向上もあって、サーバ側でデータ分析をすることがなくなりました。この状況の変化に対応して、RIEBデータベースシステムではデータ分析の機能をサーバ側に持ちません。膨大なデータのなかから必要な部分を取りだし、ユーザにとって扱いやすい形式で提供する機能だけを、トータルにわかりやすく簡単な操作方法で提供します。これがRIEBデータベースシステムのコンセプトです。

現在はTAB区切りのテキスト形式によるExcelフォーマットだけしか対応していませんが、利用者の希望によってフォーマットの種類を増やしていくことも可能です。

また最近ではデータがCD-ROMなどで大量に提供されることもあり、そのような場合は検索エンジンにOracleを用いて対応しています。Oracleは非常にスケーラビリティに優れた製品であり、今後のデータ量増加にも対応していけると考えています。

2.3 利用可能なデータベース

現在のところ以下のデータベースが利用可能になっています。

1. 多国籍企業データベース
2. OECD ITCS (International Trade by Commodities Statistics)
3. 日経総合貿易ファイル
4. DRI BASIC Economics (米国マクロ経済データ)
5. IMF EIS (Economic Information System) データ
 - (a) IFS (International Financial Statistics)
 - (b) DOTS (Direction of Trade Statistics)
 - (c) BOPSY (Balance of Payments Statistics Yearbook)
 - (d) GFSY (Government Finance Statistics Yearbook)

3.環境構築にあたって

ネットワークサイトとしての研究所は、スタッフ総勢でも 50 人を切る小さなサイトです。そこでの計算機システムの構築と運用は、総合情報処理センターが提供しているシステムやサービスとは、おのずから異なったものになります。つまりサイトの規模、ユーザの利用目的などに応じて、提供されるサービスはそれぞれ異なったものになるわけです。これは「これだけ揃えば計算機環境は完成」というような、お決まりのスタイルがないことを意味しています。

研究所でも、使っている計算機設備こそワークステーションと PC という恐らくほとんどの部局の計算機環境と共通のものですが、そこで行なわれているサービスや、それを通して実現したいこと、その方法などについては違っている部分があると思います。つまり計算機環境の構築とは、計算機という汎用パーツをもちいて各組織に必要なものをデザインし、実現することでもあります。

本稿では研究所における計算機環境について、特に主要な二つのサービスについての概要を説明しました。研究所スタッフの研究活動において、計算機とネットワークは今後ますます重要なものになっていくでしょう。そしてこれらのサービスのを支えるネットワークシステムのバックボーンは、もちろん KHAN を活用しています。今後の KHAN 及び神戸大学の各ネットワークサービスの発展に期待しています。

i 複数のディスクドライブを用いてひとつのディスクを仮想的に構成し、一本のディスクが壊れてもデータを失わずに運用するための技術

ii 多国籍企業データベースについて <http://www.rieb.kobe-u.ac.jp/doccenter/index-j.html> を御覧下さい。