Remote PC Information Network

John Wilson
Dan Macfayden

Coles Book Stores Limited
90 Ronson Drive
Rexdale, Ontario
M9W 1C1

I would like to introduce myself, my name is John Wilson and I am the Systems Development Manager for Coles Book Stores. I am speaking to you this morning along with Dan Macfayden, the Project Leader responsible for our network systems.

This paper is a bi-product of the development of a PC based store information network for a retail book chain with 192 loctions across Canada and a head office in Toronto utilizing two HP3000s (a 68 and a 70)

With 100 locations now up and live, we hope to share with you some of our findings on how to integrate PCs into a distributed DP environment.

THE ENVIRONMENT

I would like you to consider the following corporate environment:

* The existance of multiple business locations spread across a large geographic area.

* A centralized corporate office supporting the business functions common to all locations.

This type of environment is common to many industries because it makes good business sense to exploit the advantages associated with an economy of scale.

Examples - Customer Service Organizations with multiple depots.
- Branch Sales and Distribution Operations located close to the customers.
- Retail specialty chains commonly found in malls across the country.

The centralized business functions are supported by a computing facility running customized business application software for:
- Financials
- Order Processing
- Invoicing
- Shipping

and most other facets of the business to improve both the company's profitability and level of support to its customer base and remote sites.

Traditionally, much of the activity in the remote sites is used to feed the systems at H.O. with only limited information being sent back to the remote sites to support "their" own business activities.

Herein lies both the problem and the opportunity. Access to the corporate information data base is essential to the optimum operation of the remote sites. At the same time, the value of the corporate data is only as good as the accuracy and timeliness under which it is acquired from those same remote sites. There are two basic opportunities available in moving the access to corporate data and business applications out to the remote sites:

1. An online network to provide access to the corporate information and applications.

2. A PC network with remote application software and a subset of the corporate data updated using batch data communications to keep the remote and corporate data bases syncronized.

Both of these alternatives are feasible but up until recently, the online network has usually been the choice of most corporations willing to bear the costs of developing and running a network. However, the availability of low cost micro computers, comprehensive software development tools and solid data communications software has meant that our second alternative is now both feasible and cost justifiable.

When making a decision on which alternative to select, given the environment under consideration, the main consideration is based on a need for information updated minute by minute at the remote sites or will remote information updated on a daily basis be sufficient. The answer to this question will indicate whether a batch integrated distributed environment is applicable or not for your business.

BENEFITS

There are many benefits to be obtained from implementing the integrated PC solution in the remote sites:

- Savings: There is the opportunity to save manhours through the automation of tasks at the remote sites.

- Improved Operations and Profitability: Operations at the remote site can be improved in its efficiency and profitability through access to accurate and timely information at the remote site. (eg. reduced inventory levels, faster reorder cycles, better customer service, improved margin with better costs and price controls, improved cash flow)

- the corporate facilities do not have to be enhanced to support an online network with all of its added complexity and back up requirements.

- There is no single point of failure with the concept of distributed processing.

- Much of the processing requirements are at the remote site with little impact on the corporate resources.

- The data communication costs are substantially less in volume, can be done during non-business hours and do not even require dedicated phone lines at each remote site.

- Tailored application systems can be developed for the remote PCs without the high cost of multiple site licenses from software vendors.

- Onsite service is available across the country for all hardware

Lets first take a look at the overall framework of a distributed data processing environment using PCs. We will then delve further into what the structure of the remote PCs data base should be and the actual components that will make the distributed process work.
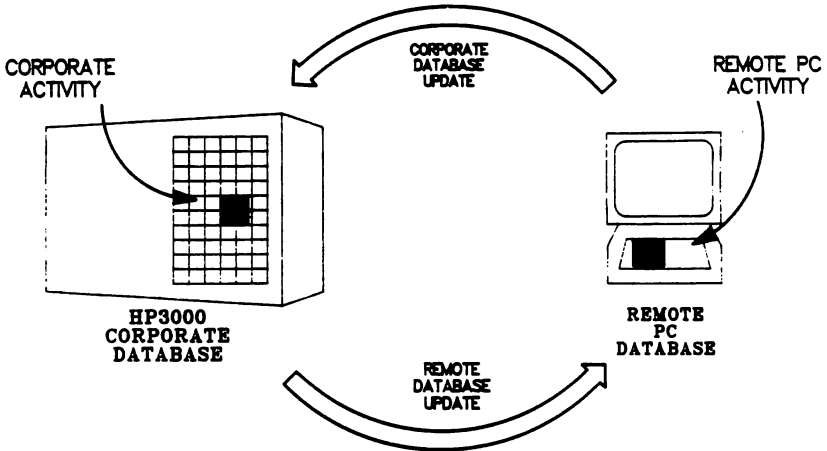

THE FRAMEWORK

The concept of developing applications to run on a remote PC is based on taking a subset of the corporate data base as it is applicable to a particular remote site and to put it on a PC. Applications are then written for the PC to

Remote PC Info Network   0048-3

support the business activities at the remote site using the
information on the PC data base.

Any transaction affecting the remote data base must be
accumulated for transmission to H.O. to be used as input to
the production systems and update the corporate data base.
Conversely, any online or batch production activity on the
corporate systems that affects the remote PC data base must
be accumulated for transmission to the remote site.

What is important is the ability to develop a
distributed environment where the remote and corporate
databases cannot go out of sync. Controls must be built
into the database update processes to ensure that the
databases cannot be corrupted by duplicate, lost or bad
data. This means that transactions unsuccessful in updating
a database must be cycled back to the originating location
for the necessary action. The measure of how solid the
systems are at both ends, including the data communications
software, will dictate the number of error situations to be
dealt with. A word of advice to those of you contemplating
development of a distributed processing environment: if
your present corporate systems are not solid don't
extrapolate your problems by spreading your errors around
onto additional machines.

FRAMEWORK

Remote PC Info Network   0048-4

STRUCTURE

Let us consider one more area when looking at the
framework of a distributed environment and that is the
logical structure of the databases on the corporate and
remote system. Corporate systems are usually made up of
"static master file" data such as product, vendor,
department information and secondly, "transaction based
data" such as purchase orders, invoices, etc. Integrated
corporate systems structure the data to support both company
level processing and for production based systems focused at
the remote business activity level (eg. creating purchase
orders or invoices).

The processing requirements for the remote PC only
require those pieces of static master file data and
transaction based data as is necessary for that individual
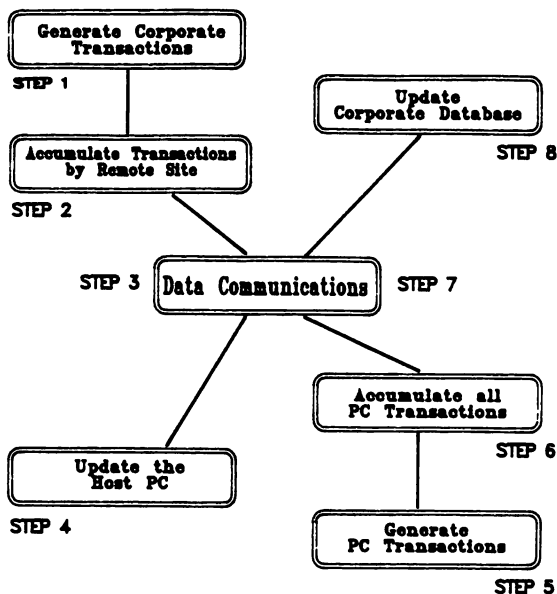remote site to function.

ie. Products carried, related vendors, store orders,
    associated sales history

What we can recommend is that the structure of the
remote data base must reflect the data base structure on the
corporate systems. This gives you a situation which is
straight forward to understand and possible to keep
synchronized. Without the similar structures, it would be
extremely difficult if not impossible to even detect when
the systems are out of sync let alone know how to go about
correcting the situation.

COMPONENTS

The process of keeping the corporate database and the
remote locations' data bases synchronized is comprised of a
number of component activities. These components can be
viewed separately in order to more easily understand and
maintain the integrated processing that is involved:

Remote PC Info Network   0048-5

COMPONENTS

```
        ┌─────────────────────┐
        │ Generate Corporate  │
        │    Transactions     │
        └─────────────────────┘          ┌─────────────────────┐
         STEP 1                           │       Update        │
                                          │ Corporate Database  │
                                          └─────────────────────┘
        ┌─────────────────────┐                        STEP 8
        │ Accumulate Transactions │
        │    by Remote Site   │
        └─────────────────────┘
         STEP 2
                    ┌─────────────────────┐
          STEP 3    │ Data Communications │   STEP 7
                    └─────────────────────┘

                              ┌─────────────────────┐
                              │   Accumulate all    │
                              │  PC Transactions    │
                              └─────────────────────┘
                                            STEP 6
        ┌─────────────────────┐
        │   Update the        │
        │    Host PC          │     ┌─────────────────────┐
        └─────────────────────┘     │     Generate        │
         STEP 4                      │  PC Transactions    │
                                     └─────────────────────┘
                                              STEP 5
```

Step 1   **Generate Corporate Transactions**

Every corporate system, (online or batch), that
alters masterfile or transaction data on the
corporate database that is also carried on the
remote PC database must generate a transaction
that can be used to update the remote databases.

There are two alternatives for the transaction
record format: The first is to simply send all
information about the record to be updated or
secondly, to develop complex techniques that
identify the pieces of data that have been changed
and only send this data. This will obviously have
an impact when considering your data communication
costs.

Remote PC Info Network   0048-6

Step 2   **Accumulate All Transactions by Remote Location**

The transactions that have been generated by the
corporate systems must be accumulated by remote
location and organized by date and time so the
same processing sequence will occur at the remote
sites. It will also be necessary to append
transactions to any remote site's files that have
not transmitted since the last time this
accumulation was run.

On the HP3000 it is advisable to set up separate
groups for each remote site to hold the files that
are specifically for that location.

Step 3   **Data Communications from Head Office to the Remote PC**

There are a number of products available today
that are able to provide error free data
transmissions between an HP3000 and remote PCs.
There are, however, a number of other features of
the communications component that are not only
desirable, but essential, if communications with a
large number of remote PCs is to be successful.
These include:

-       the ability to initiate communications
        without requiring manual intervention at the
        remote locations or at head office for
        unattended after hours transmission

-       an automatic restart capability beginning
        from the last successfully transmitted
        transaction, whenever communications is
        interrupted.

-       file integrity routines such as traditional
        batch header control techniques to ensure the
        correct data is being processed by the proper
        location.

-       the ability to maintain and monitor the
        transmission status for each transmit file,
        coupled with the ability to retransmit
        unsuccessfully transmitted files in the next
        run.

- the availability of network management information for performance and trend reporting as well as daily problem and exception reporting.

Step 4   **Update the Remote PC**

Software on the remote PC will process the Head Office transactions and update the data base. Controls should be built in to insure the integrity of the remote database is not threatened by corrupt incoming transactions or by failure during the transaction processing. The update component can also provide the ability to generate reports to be printed at the start of the next day. New versions of programs can be implemented at this time (if they have been down loaded from the corporate machine). A transaction activity summary should be generated to inform the remote site of the activity that occurred over night and any transactions that could not be processed successfully should be added to an error file to be transmitted back to the corporate site for reivew.

Step 5   **Generate PC Transactions**

All systems on the PC that update master file data or generate dynamic data that impacts corporate systems must generate a logical transaction that can be sent back to the corporate site.

Step 6   **Accumulate All Transactions on the PC**

Before setting up for data communications this component has to create the files containing:

- the new transactions to be processed at head office
- any previously rejected head office transactions
- relevant PC database and disk utilization statistics to be used at head office to manage and support the remote PCs

Step 7   **Data Communications from the Remote PC to Head Office**

This process is the same as the head office to store data communications with the same controls and restart capabilities.

Step 8   **Update the Corproate Database**

The incoming transaction files from the remote loctions must be split up by transaction type and the transactions integrated into the associated corporate production systems.

Any transactions that can not be successfully processed should be reported at head office for investigative and corrective action, and the remote loation should be informed via the next transmission.
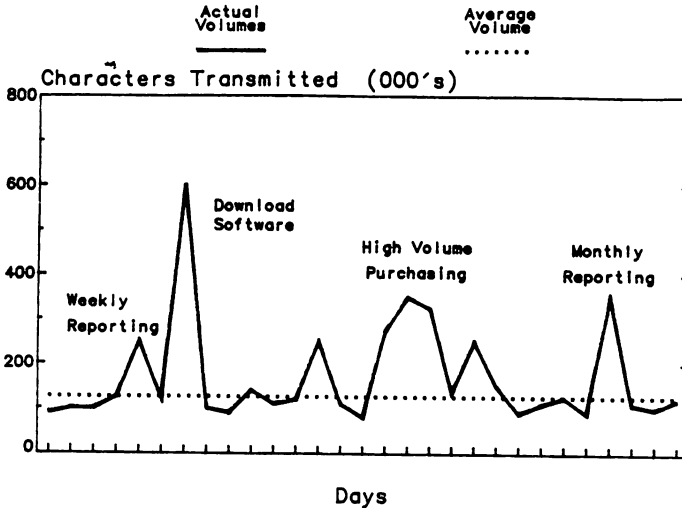

DATA COMM SCHEDULING

There are two basic alternatives for driving the communications process--either the central HP3000 or the remote PCs can initiate or drive the communications process. Software products exist to support either of these alternatives.

In our opinion, PC initiated communications is viable in a multiple remote-location scenario if the volume of transactions per location is small or predictable enough for you to schedule the incoming calls sufficiently far apart to avoid contention. It is also viable if you can afford to add enough incoming communication lines to avoid contention. But recognize that when the remote sites are scheduling the upcoming calls there are certain factors that you cannot control which can cause contention:

- the corporate HP3000 may be unavailable
- the initial data communication attempt may fail, forcing an unscheduled re-dial
- increased transaction volumes can require longer connect times than planned

So, if your corporate processing resources are limited, or your transaction volumes are volatile, we recommend using the corporate HP3000 to poll the remote PCs.

Remote PC Info Network   0048-9

DATA COMM SCHEDULING

Actual Volumes

Average Volume

Characters Transmitted (000's)

Download Software

High Volume Purchasing

Monthly Reporting

Weekly Reporting

Days

This allows you to handle planned and unplanned transmission variances such as the planned release of software revisions or an unplanned increase in transaction volumes. From head office you can also reschedule any redial attempts later in your transmission window whenever a communication link has been unsuccessful without causing contention with other sites.

The priority of communications with specific remote PCs can also be parameterized and controlled. Indeed, it is much easier to alter the entire polling pattern (based on volume changes, business activity changes, special circumstances, etc.) when you have control of the data comm at head office.

Once you have purchased your hardware and your network has been developed and is now operational, the greatest cost you will incur will be your data communication cost. A centralized polling facility with consolidated network performance reporting gives you the best opportunity to control your data comm costs.

MAINTAINING A DISTRIBUTED PC NETWORK

Now that you have your network up and functional with nightly data communication processes to maintain the remote and corporate systems, you must be prepared to support the remote sites.

For most companies, technical support is only available at the central site due to its high cost. Not only is technical expertise unavailable at the remote sites, but in a PC environment the level of computer expertise will usually be limited to the use of business application software only. This makes it unwise to rely on personnel at the remote sites for any computer maintenance related activity that might usually be found in a corporate mini or mainframe environment. Remote personnel cannot be relied upon to perform computer-problem diagnostics as part of their job responsibilities.

We have addressed this limitation of a lack of expertise at the remote sites by employing a strategy that provides for centralized network support, that can resolve any PC or network problem from the central location. In order to achieve this objective we have developed the following tools and procedures:

1. A PC software product that allows access to the remote PC by corporate support personnel. This provides access to MS-DOS level commands and diagnostics. Software problems at the remote location can also be investigated by running versions of the same application programs on a corporate PC while utilizing the remote PC's data.

2. Tools similar to the HP3000 Image database utilities are used to remotely rebuild a PC database or dataset as well as perform other necessary database management functions. It is worth emphasizing that these remote utilities must have a sufficient level of data communications integrity to insure that corruption of the remote data is not caused by the investigation process.

3. A facility is used at the corporate site to download program revisions and special corrective programs to a single site or to all sites as part of the daily transmit process. Instructions and manual updates can also be sent with any program revisions.

4. In the event of a disaster where the hard disk is lost,
    software is available to recover and rebuild the
    systems for a remote site. This can be done from a
    local backup, or from the corporate database which
    contains the subset of data that constitutes the remote
    site's database. This is one of the biggest benefits
    of keeping the remote and corporate databases in sync.

5. Escalation routines are in place to insure problems are
    not left unresolved. Included in these routines is an
    adequate hardware service contract for each remote
    site.

6. Network support reporting is used to:

    - monitor database and disk capacities for the
      remote sites
    - provide historical trend analysis by transaction
      type
    - provide daily operational and management reporting
      of network data communications activity and costs

    We have found that reporting on an exception basis for
    the network support staff is most useful in focusing
    attention on possible problems.

7. A messaging facility is used to keep the remote sites
    informed whenever potential problems are discovered or
    solutions are introduced as well as to introduce new
   processses, complete with the manual updates.

INSIGHT

    Based on our experience at Coles, we would like to
relate the following "bleeding edge" lessons to anyone who
is considering using PCs in a distributed DP environment.

1. Ask yourself how up-to-date the information on your
    remote PCs has to be. If you really need minute-by-
    minute updates then the process we have described may
    not be appropriate.

2. If your corporate systems are not in place and stable,
    or if you anticipate dynamic changes in your corporate
    systems, then you should probably defer the development
    of the PC network until the work is completed.

Moving from a centralized DP environment to a distributed approach with many locations cannot be done overnight. Your corporate systems must be able to support both remote locations on the existing process and also the integration of the distributed PCs as they are implemented. Also recognize that business does not stand still and changes will be required as new business opportunities dictate. There will be a need to maintain and evolve both the new and the old systems during your conversion process.

3. Prototype your remote PC application before developing the interfaces with the corporate systems. This allows the flexibility of fine tuning the PC design before developing the interfaces to the corporate systems.

4. If the number of remote sites is low, then you may not require the level of automation we have described for data communications, remote support, network management, etc. Twenty local sites is about the maximum that can be implemented in a "manual" mode. More sites or a greater geographical spread will require automation of many of your processes.

5. Set up separate project teams and the necessary project controls in the areas shown in the following diagram. Don't short change on the requires planning effort.

PROJECT ACTIVITIES

---
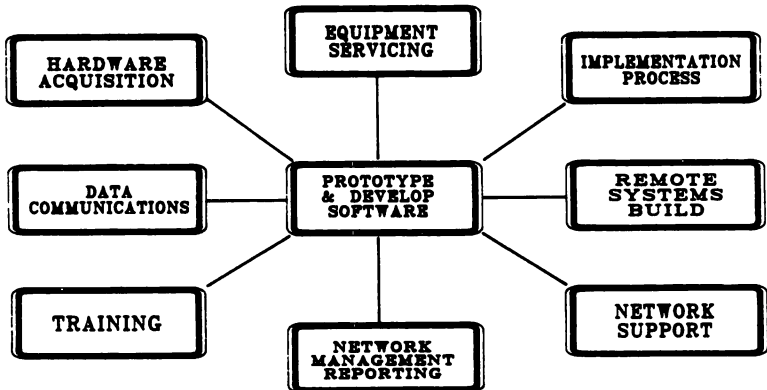
6. Buy or develop the best data comm software you can find. It's an important cornerstone to your network.

7. Be prepared to develop your own in-house technical expertise on both the HP3000 and the PCs to create and maintain your network.

8. Stagger the remote site implementations to give yourself opportunities to learn, revise, and improve before your support level becomes to high. For instance, bring up 1 site, then 5, then 20, then 100, with time in between each group to evaluate and make appropriate changes.

SUMMARY

The advent of micro computers has given us the ability to put processing power into the hands of remote PC users. At the same time, MIS has retained the responsibility for the corporate information resource. The issue that continually faces those of us in MIS is how do we provide the access and update capability of the corporate data base to the remote PC user without threatening the integrity of the corporate data base. The answer of course is that only by using the same development methodology and processing controls and standards as we use in developing corporate systems can the full functionality of corporate data access be passed down to the remote PC user. I stress the use of controls and standards because the hardware, software development tools and data communiations software already exist to do the job. All that is needed is the ability to apply these technologies within a constant framework. This, I think, we have done with the development of the Coles Store Information Network.