Planning the Corporate Data Centre

BUTLER COX FOUNDATION

Research Report 55, April 1987

BUTLER COX FOUNDATION

Planning the Corporate Data Centre

Research Report 55, April 1987

Butler Cox & Partners Limited

LONDON
AMSTERDAM MUNICH NEW YORK PARIS

Published by Butler Cox & Partners Limited
Butler Cox House
12 Bloomsbury Square
London WC1A 2LL
England

Availability of reports
Members of the Butler Cox Foundation receive three copies of each report upon publication; additional copies and copies of earlier reports may be purchased by members from Butler Cox.

Printed in Great Britain by Flexiprint Ltd., Lancing, Sussex.

BUTLER COX FOUNDATION

Planning the Corporate Data Centre

Research Report 55, April 1987

Contents

1 Purpose and content of the report	1
Research for this report	2
Research team	2
2 Future of the data centre	3
Centralised computing has a continuing role	3
The role of the data centre is changing	5
Equipment will be changed more frequently	7
Fewer staff will be required	7
The data centre of the future	9
3 Critical factors for planning the data centre	10
Equipment space planning	10
Environment planning	15
Security planning	18
Summary	19
4 Data centre planning guidelines	20
Space requirements will be inaccurate, so plan accordingly	21
Move to a greenfield site wherever possible	22
Plan for flexibility by using a modular design	22
Set a realistic budget	22
Appendix 1: Case histories	24
Appendix 2: Four-step methodology for predicting equipment space requirements	28

BUTLER COX FOUNDATION

Chapter 1

Purpose and contents of the report

A data centre is a concentration of data processing, storage, and communications equipment that serves many users, and that is managed within a secure and professional environment. The processing equipment may be one large computer shared between many users, or it may be several smaller computers each dedicated to a specific group of users. Data centres have evolved from the computer rooms of 15 years ago. The change in name has turned out to be more appropriate than perhaps was first envisaged because, today, the planning and management of data centres is dominated by the space and environmental requirements of data-storage devices, rather than by the computers.

Before commencing the research for this report we asked Foundation members about their data centre planning concerns. The main ones were:

- The likely trends in the performance and space requirements for data-storage devices and processors.
- The impact of distributed computing on the corporate data centre.
- The trend towards unmanned data centres.
- The security measures that could be taken to protect the data centre premises and the equipment and data housed in them.

Security was of greater concern to those organisations that would be particularly vulnerable to the loss of their computer resource because of the nature of their business applications, and/or because of possible terrorist attacks.

Corporate data centres require substantial investment both in the building space they occupy and in their specialised requirements for air conditioning and power supplies. Increasing dependence on information systems has also emphasised the need for security of the premises, and for protection from hazards such as fire or sabotage. Hence planning a new data centre is a significant exercise in which there are substantial risks to be counterbalanced. For example:

- The risk of not providing enough space to allow for additional equipment to be installed in the data centre, so that it is prematurely full and cannot cope with the demands of the business, compared with the risk of providing too much space, so that the data centre premises cost more than they need to.
- The risk of overinvesting in security and environmental safeguards, compared with the risk of consequent financial loss to the organisation, if the safeguards are inadequate.

Furthermore, the environmental requirements of the data centre are complex and involve more than straightforward air conditioning. In particular, air filtration can be a problem. A failure to use the most up-to-date expertise can result in severe (and apparently random) equipment failures. (This particular aspect of data centre planning is discussed more fully on page 16.)

This report has been prepared to provide guidance in planning a new data centre. It is written primarily from the point of view of an organisation that can build a new data centre on a greenfield site. However, most of the research results and advice we give can readily be adapted for situations where existing centres are being extended or refurbished.

The remaining chapters of the report address the following questions:

- What is the future of corporate data centres? Will they continue to exist at all, bearing in mind the trend towards distributed computing? If they do have a future, how will their role change and what equipment will they contain?
- What are the critical factors for planning the future data centre? What will determine the space required? What needs to be

considered in planning the environment within the centre, and how should security be provided?

— How should a data centre be planned? What is the best approach to coping with the uncertainty of future needs?

RESEARCH FOR THIS REPORT

In attempting to answer the above questions we have drawn on the experiences of data centre managers, of the suppliers of equipment used in data centres, and of experts concerned with designing data centres.

Following a thorough review of the available literature on data centre planning, a questionnaire survey was carried out of organisations known to have at least one large data centre. The purpose of the questionnaire was to ask about recent growth trends and future plans in terms of equipment installed and floor space occupied. Fifty-six replies were received from organisations in eight countries.

Finally, interviews were conducted with selected systems managers and with representatives of equipment suppliers. The purpose of the interviews with systems managers was to obtain detailed information about the practical approaches to data centre planning. (The results of some of these interviews are included as case histories in Appendix 1.) The supplier interviews provided useful information about technological trends affecting the equipment most commonly found in corporate data centres.

RESEARCH TEAM

The research for this report was carried out in the second half of 1986 and early 1987 by *Tony Manley* and *Jim O'Connor*. Tony is a principal consultant with Butler Cox. He has more than 25 years experience in the information technology industry and specialises in the development of computer policy and strategy, and the application of computers to business needs. Jim O'Connor is Butler Cox's director of systems consultancy and was formerly director of the UK Foundation. His background includes senior systems management positions in Europe and the United States, and he is the author of two previous Foundation Reports.

In conducting the research for this report, they drew on their own consulting experiences, and those of other Butler Cox specialists.

Chapter 2
Future of the data centre

One of the concerns expressed by some Foundation Members was that corporate data centres might be made obsolete by the trend toward distributed processing and distributed databases. A new centre could become a 'white elephant' - a memorial to a wasted investment in time and money if the need to concentrate computer equipment in one location disappears.

As a result of our research, we believe that, despite the growth of distributed processing, the corporate data centre does have a continuing role. However the role of the data centre is changing, with resultant implications for its contents and staffing requirements.

CENTRALISED COMPUTING HAS A CONTINUING ROLE

At first sight, the least expensive way of installing a large amount of computing power is to install several smaller computers. (The cost of 20 microcomputers, each of 1 mips, is significantly less than a 20-mips mainframe, for example.) However, from our discussions with senior systems managers from a wide variety of organisations that have corporate data centres, we concluded that, although some medium-sized data centres may be replaced by distributed systems or by microcomputing facilities, large corporate data centres will continue to expand in the medium to long term. Indeed, it appears that the growth of distributed computing and the proliferation of microcomputers in larger organisations is causing a parallel, albeit smaller, increase in the demand for central data centre facilities. Evidence for the continuing growth of data centres can be seen in Figure 2.1, which shows the projected value of shipments of computer systems in the United States until 1990, by class of configuration. The most important aspects of these projections are as follows:

— There is 30 per cent overall growth in the annual value of all computer system classes from 1986 to 1990.

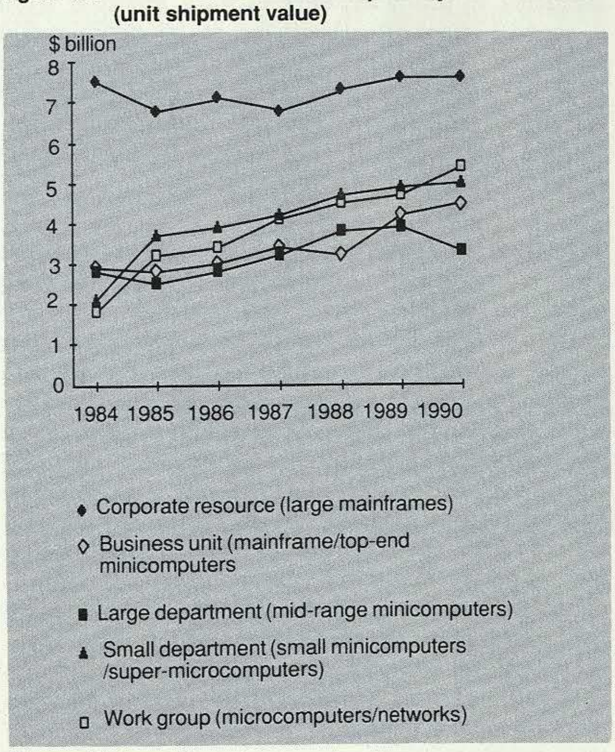
— While the greater part of this growth (40 per cent) is computers that would normally be classified as distributed or personal, there is still expected to be 20 per cent growth in the annual value of systems usually installed in data centres (top-end minicomputers and large mainframes).

There are many applications-related and organisational reasons for the continued demand for central computer facilities and therefore the corporate data centres that house them. The most important of them are discussed below.

APPLICATIONS-RELATED REASONS

Many data processing applications, by their very nature, will always require centrally located processing facilities. Organisations

Figure 2.1 **US business computer systems forecast (unit shipment value)**



(Source: Dataquest)

such as airlines, banks, and insurance companies will probably continue to require fast-response access to large databases, and large organisations in many other industries will continue to have similar requirements. Large centralised systems can support much larger data-storage systems than small computers. Data files could be divided, but the software problems have not yet been solved, and the problems of controlling distributed databases are substantial. Although some minicomputer manufacturers now claim to have effective distributed database software, the experiences of many large organisations have shown that the most cost-effective and least risky method of providing access to large databases is still through the use of a centralised system. Despite the claims of the minicomputer manufacturers, it is still difficult to keep the components of a distributed database in step with each other.

Another applications-related reason for a continuing demand for central computing facilities arises from the growth of personal computing. There are many instances where personal computing facilities have led directly or indirectly to an increased workload at the central data centre. For example, many organisations have found it necessary to create a separate enquiry database in addition to the existing operational database. The enquiry database is used to provide ad hoc access to the data. And many organisations have provided facilities to allow subsets of data selected from a central corporate database to be downloaded from the mainframe to local workstations. In both of these examples, extra central resources for processing and/or storage and/or communications are required in addition to the resources at the personal computing location.

The inertia of the established systems base is also a significant reason for the continuing existence of corporate data centres. Approximately 60 per cent of the total computing power in the average data centre is used for routine production work. Often, half of the routine workload arises from applications that were originally implemented between five and ten years ago. Although most, if not all, of these applications will have been enhanced significantly (and the volume of transactions being processed by them will have increased significantly), they are still being operated in much the same way as they were when they were originally developed. The inevitable result is that changes in data centre operations are always constrained by the operational needs of the existing systems, and the high cost of conversion.

There is yet another applications-related reason for believing that the corporate data centre will continue to exist. The very fact that centralised facilities already exist means that it is much easier to develop new applications to use these facilities. A new distributed computing environment will not be introduced specifically for small applications. Thus the existence of the data centre tends to be self perpetuating.

ORGANISATIONAL REASONS

In the past, changes in organisation structure or the relocation of significant numbers of users often meant that a data centre was closed down or relocated. Today, this does not happen so readily because most users (including development staff) are now linked to the data centre facilities through communications networks. As a consequence, there is now no need for a data centre to be located with the majority of the user community. Thus, organisational changes or staff relocations are now less likely to lead to the closure of a data centre.

Another organisational reason for the continuing existence of data centres concerns the scarcity of skilled senior operations and technical support staff. The shortage of skilled staff means that they are expensive to employ. It therefore makes sense to concentrate the available skills at a central location, rather than to spread them thinly throughout the organisation. Moreover, a large, professionally run data centre staffed by high-quality personnel makes it easier to attract technical experts and provide them with work that is challenging and stimulating. Furthermore, it is possible to provide better opportunities for staff development and career progression.

An additional organisational reason for the continuing role of the corporate data centre arises from the reluctance of some user departments to take on the responsibility for running their own computers. Some departments fight hard to control their own local computing facilities. Others have the responsibility thrust upon them as a result of a senior management decision. Our research discovered instances where departments had either accepted the responsibility for a period of time before asking to be relieved of it, or had refused to accept the responsibility after they had reviewed the costs and problems that would be incurred. In both cases, the responsibility for providing and running the computing resources remained with the central data centre.

We also encountered situations where distributed minicomputers had been installed at local sites, but had subsequently been moved to the central data processing facility in order to improve reliability or security, or to gain the advantages of scale provided by an established operating environment (see, for example, the clearing bank case history on page 26). This situation had arisen particularly where the decentralised facilities had grown to the point where full-time operating staff were needed.

On the other hand, the decentralisation of computing facilities has sometimes led to the creation of substantial new data centres. Initially, the decentralised computer facilities were not very significant, but, over time, they have often evolved into full-scale data centres in their own right. In some organisations, the individual divisional data centres have become as large as the original central data centre.

THE ROLE OF THE DATA CENTRE IS CHANGING

We believe the role of the corporate data centre is changing fundamentally because the emphasis of the work performed in the centre is changing. Increasingly, it will be focused on operations rather than development, and on managing data storage and communications rather than processing applications and minding peripherals.

MORE OF A DATA UTILITY

The traditional role of the typical corporate data centre was to process a predictable workload. A predictable volume of input was processed sequentially by a well-defined process to produce printed outputs or screens of information. Until recently, the data centre's workload was largely self contained, with every processing step being carefully controlled by the operations staff. This was particularly true for batch-processing systems, but also still applies to many online systems, which continue to use batch-updating procedures. Today, however, an increasing proportion of the data centre's role is to provide online ad hoc enquiry services. Also, the increasing number of online systems in a typical data centre means that significant amounts of data and applications processing are located outside the data centre in the user community. As a consequence the use of the data is no longer always controlled from within the data centre. For example, the use of data stored in local files will be controlled by the users. Thus, users will decide whether (and when) this type of data is transmitted to a central data centre for archiving or for control purposes, and they will decide whether the data can be made available to other users.

The result of these changes is that the data centre has to provide adequate computing resources to meet an increasingly unpredictable workload. The data centre planner has therefore to balance the cost of providing excessive capacity against the risk of not being able to meet exceptional peak demands. In this respect, the provision of computing facilities in a data centre is analogous to the provision of generating capacity by a national electricity supply company. The analogy can be extended further by considering the role of an electricity supply company in constructing and maintaining the distribution network, and in promoting the use of electricity. Data centres will increasingly play an analogous role in the distribution and use of data throughout the organisation.

In Foundation Report 50 (Unlocking the Corporate Data Resource) we identified two different types of data: authenticated data, which is vital to the organisation, and is usually controlled and validated by a data administration function and is made available via a carefully managed database; and unauthenticated data, which is relevant only to local needs, and is not centrally controlled. It is clear from our research that the volume of both types of data will grow dramatically. In particular, new classes of applications with extensive storage requirements both for authenticated and unauthenticated data are now being developed to satisfy increasingly complex user needs. Although much of the processing for future applications may be distributed, the corporate authenticated databases will continue to be managed and maintained centrally because they are the essential lifeblood of the organisation.

LESS RESPONSIBILITY FOR DATA ENTRY

The widespread adoption of online systems has meant that the need for a central data-entry department has all but disappeared. In most organisations, data entry via visual display terminals is now the responsibility of user departments.

Other forms of data-entry technology, such as document reading, are also increasingly being installed in the user community. Many of the organisations that participated in our research told us they expected to make greater use of document scanners to meet the growing demand for capturing text and image information, particularly from documents originating outside the organisation. The rate at which document scanners can input pages of text or images is expected to double in the next five years.

MORE RESPONSIBILITY FOR COMMUNICATIONS

In Foundation Report 54 (Integrated Networks) we highlighted the developments that, eventually, will lead to integrated voice/data networks becoming commonplace. Initially, the trend towards integrated networks will be seen in corporate wide-area networks, spurred on by the emerging ISDN services that will be offered by the PTTs. We believe that in some organisations the responsibility for managing the integrated wide-area network will fall on the systems department, and that the network management role will be located in one of the corporate data centres. Other organisations will choose to continue to operate discrete data networks, and these too will be managed from a corporate data centre. In either case, the corporate information network will become more important, and its management will be a key responsibility of the data centre manager.

Even where separate voice and data networks are maintained, there is a case for housing the PABX in the corporate data centre. As the number of terminals connected to the data centre increases and approaches one per office worker, it will be necessary for cables to run from the data centre to every desk. This, of course, mirrors the wiring requirements for the PABX, and it might make sense to use the same cable ducts for telephone wiring and the data network.

One example of the growing importance of networking is the increasing use of electronic mail by many organisations. The data centre will usually house and control the computer hardware used as the message-switching centre. Thus the data centre will need to install additional storage, processing, and communications equipment to handle the flow of information through the electronic mail network.

LESS RESPONSIBILITY FOR APPLICATIONS PROCESSING

In most large organisations, a growing proportion of the total processing power is being placed in the user community through the increasing use of personal computers and departmental minicomputers. Currently, the bulk of the applications software in the user community supports either an individual or a specific departmental processing need mainly word processing, spreadsheets, and personal job aids. However, an increasing amount of mainstream applications software is being developed to run on processors located in the user community, not on the central mainframe. This is particularly true for transaction-validation processing and for applications that allow users to access data from a copy of the corporate databases and then manipulate it.

In the long term, the shift to the user community of the computing power required for manipulating data will reduce the requirement for central processor power to perform the analytical aspects of most applications. However, this reduction will occur very gradually and will, in the end, be more than compensated for by an increase in workload associated with centralised database management systems as the central mainframe takes on more of a file-server role. The net result will be an increased requirement for mainframe processing power and data storage in the typical data centre.

LESS RESPONSIBILITY FOR PRINTED OUTPUT

For some time now there has been a trend to remove the preparation of printed output from the computer room and place it in a specialised area. More recently, with the introduction of small, inexpensive laser printers, there has been a trend to remove the preparation of printed output from the data centre altogether and place it in the user community. As a result, the needs for manual paper handling and stationery storage in the data centre are becoming less.

For example, one organisation known to us is placing with the user community the responsibility for selecting and formatting the data for printed reports. Instead of writing and running programs to print reports, this organisation extracts the data fields that will be required to produce the reports and dumps the data onto a floppy disc in a format compatible with well-known microcomputer software packages (Lotus 1-2-3 and so on). (When defining the extract, data fields not required immediately, but which might be required in the future, are also included.) Users then manipulate the data on their own personal computers, selecting the required data, sorting it, formatting it, and printing the reports locally. The advantage of this approach is that the user gets exactly what is required without any involvement from systems development staff.

In some organisations, the need for printed output could be reduced even further as synthesised-voice-output technology develops. We foresee voice-synthesis becoming a standard component of applications that access corporate databases, particularly where the user may not have ready access to a screen or printer. Situations where synthesised voice response might be valuable would include enquiries by field sales or maintenance staff, or enquiries from remote construction sites.

MORE OF A CORE SERVICE TO THE ORGANISATION'S OPERATIONS

In summary, we foresee the data centre increasingly becoming both the central store of the data common to different areas of the business, and the organisation's communications centre. It will be the home of the technical infrastructure that supports the information and communication needs of the organisation. As such, it will fulfil a vital role in keeping the business going. Other activities previously housed in the data centre (such as data entry, system development, and operating peripherals) will increasingly be carried out at separate locations, some of which will be in the user community. The changing balance of the responsibilities of the data centre and the user community are shown in Figure 2.2.

In addition, the corporate data centre may also become the home of smaller computers that user departments do not want to manage themselves.

EQUIPMENT WILL BE CHANGED MORE FREQUENTLY

The equipment installed in large data centres is almost continually being changed, partly as new models become available but mainly to accommodate new applications or extensions to existing ones. We believe that, increasingly, data centre equipment will also be changed because it will be perceived as a commodity that can be bought and sold to take advantage of changes in price/performance ratios. Opportunities for doing this already exist within the IBM architecture. Astute managers can swap IBM processors for those provided by the plug-compatible manufacturers (and vice versa) as one supplier leapfrogs another in terms of price/performance. (Bill Cook's presentation at the Gleneagles Foundation conference in 1986 described how Morgan Stanley Bank now perceives processing power as a commodity that can be sold and bought at very short notice.)

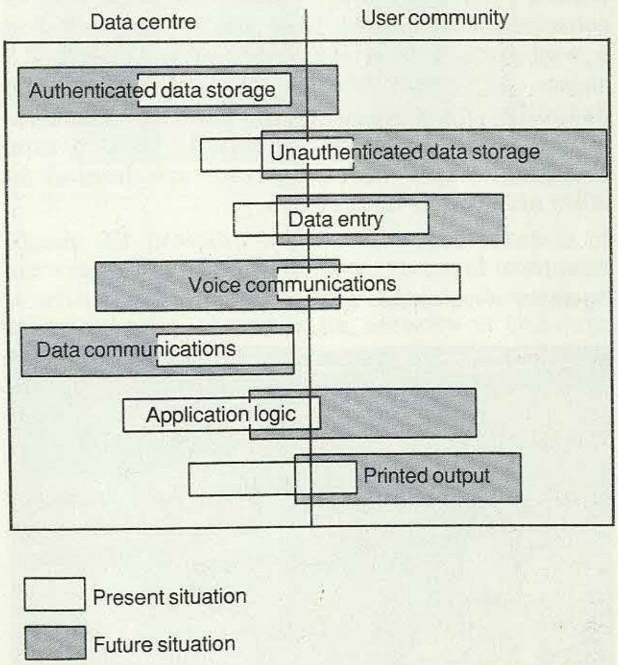
This trend will be reinforced by recent initiatives that are leading to standard interfaces for computer peripherals. The Intelligent Peripheral Interface (IPI) standard drafted by ANSI (American National Standards Institute) will allow peripherals from different manufacturers to be connected to a variety of computer systems. Unisys is already building products that conform with the IPI standards, and the architecture of the new IBM 9370 mid-range system has been designed with IPI in mind.

As a consequence, equipment changes within the data centre will become more frequent as opportunities for 'commodity trading' are added to the need both to replace obsolete equipment and to install new equipment to cater for a larger workload. The implication is that adequate 'spare' space needs to be set aside in the data centre so that the new equipment can be installed and commissioned before the old equipment is removed. In the past, some organisations have not allowed for space that will be used in this way.

FEWER STAFF WILL BE REQUIRED

The staff that work in data centres are specialists: operations staff, systems programmers, and, increasingly, telecommunications staff. Corporate data centres do not necessarily house applications development staff, however.

Figure 2.2 The changing role of the corporate data centre



The size and position of the boxes represents the extent to which the data centre and the user community are responsible for each area.

(Some organisations regard system development as an integral part of the data centre's responsibilities, with the development staff being located at the data centre. The reasons for this are largely historical. Today, system development staff can be considered as just another group of specialised users of the data centre. Like other groups of users, they no longer have to be physically located at the data centre - the clearing bank case history on page 26 is a case in point. For this reason, we have have not considered system development staff as part of the data centre's staff complement.)

Lights-out rooms will be controlled from a data centre bridge

Much of the equipment now housed in a data centre does not need to be continually 'minded' by human operators. Hence, the unmanned or 'lights-out' computer room, containing the 'inert' equipment (processors, fixed disc drives, communications equipment, etc), is now a possibility. The equipment in the lights-out computer room will be controlled from a command centre or 'bridge' (as in a ship's bridge) where computer-room and network-control decisions will be made.

A few organisations have already implemented a data centre bridge and find that it allows them to run complex data centres with a minimum of staff. In particular, the bridge concept allows them to make much better use of highly skilled computer operators' time. The highly skilled staff control the operations from remote consoles located in an area that is furnished in a way similar to a standard office. Figure 2.3 shows a typical data centre bridge. The dwindling number of lower skilled staff required for the more manual tasks (tape changing, paper handling, etc) are located in other areas of the data centre.

Figure 2.3 A typical data centre bridge

Another advantage of the data centre bridge concept is that it minimises the risks of mishandling or damaging (whether deliberately or by accident) the expensive equipment located in the lights-out computer room. People only need to handle the equipment either when it is installed or for maintenance purposes. In addition, by locating the inert equipment in an area that is virtually off limits to operations staff, it is possible to provide much better fire-prevention systems. (Some major organisations have removed fire-prevention systems from computer rooms because of the potential risk to people should the chemicals in the systems be discharged.)

Recent advances in networking and operating systems have also made it easier to implement an operations bridge remote from the data centre. These advances mean that some computer systems no longer require a dedicated operating console — any terminal can be assigned as the central command console, subject to appropriate security requirements. For example, the Inland Revenue in the United Kingdom has one command centre from which the operations of several ICL mainframe installations are controlled. The operators no longer need to be near, or even to know where the mainframes are physically located.

Fewer operators will be required

The changing role of the data centre described earlier in this chapter also means that fewer computer operators will be required for a given workload. In particular, there will be significant reductions in the requirement for less-skilled staff — the main reasons being the transfer of data-entry functions and (to a lesser extent) printed output preparation to user departments (with a consequent reduction in the requirement to load and unload preprinted stationery), and the reduction in the use of conventional magnetic tapes and removable disc drives.

High-speed communications links between data centres are also contributing to a reduction in the number of staff needed to handle peripheral devices. Many organisations (over half in our survey) now have more than one data centre. Often, they use high-speed links to transfer files between data centres so that one centre can act as a backup to the another. This procedure reduces the need to make security tape copies of files and physically transfer them to another site for storage. Security copies of files can be transferred very quickly over the communications link, often without the operations staff needing to be involved at all. (Some systems now transfer transactions to a remote 'shadow' database a few seconds after the operational database has been updated. In this way, the shadow database can automatically take over in the event of a failure at the data centre housing the operational database.)

The need for skilled staff will not reduce

However, there are three areas in which data centre manpower requirements will stay the same, or will possibly increase slightly. The first of these is the continuing need for systems programmers. The increasing complexity of the services provided by the typical data centre, and the continuing inability of the major hardware suppliers to provide easily maintained systems software, means that there is a continuing demand for highly skilled systems programmers to oversee the increasingly complex system software environment found in most data centres.

The second category of staff whose numbers will not decrease in the foreseeable future is network planners and controllers. As more advanced telecommunications facilities become available, and as an organisation's network becomes increasingly complex, there will be an increasing requirement for skilled network staff.

The third area is database administration, which will become an increasingly important support function as the data centre develops its 'data-utility' role. Clearly, one person can look after only a certain amount of data storage. As the amount of data stored online increases, organisations will need to employ additional database administration staff.

Nevertheless, we believe that any increases in the numbers of systems programmers, network staff, and database administrators will be more than offset by reductions in other areas. Overall, there will be fewer staff employed in the data centre of the future.

THE DATA CENTRE OF THE FUTURE

To summarise, our view of the data centre in the medium- to long-term future is that it will be:

— A focal point (or points) for data storage and communications, with the organisation being increasingly dependent on the services and facilities provided by the data centre.

— A secure environment for equipment, with fewer operators than today's data centre, and no systems development staff.

Figure 2.4 The change in content of the data centre

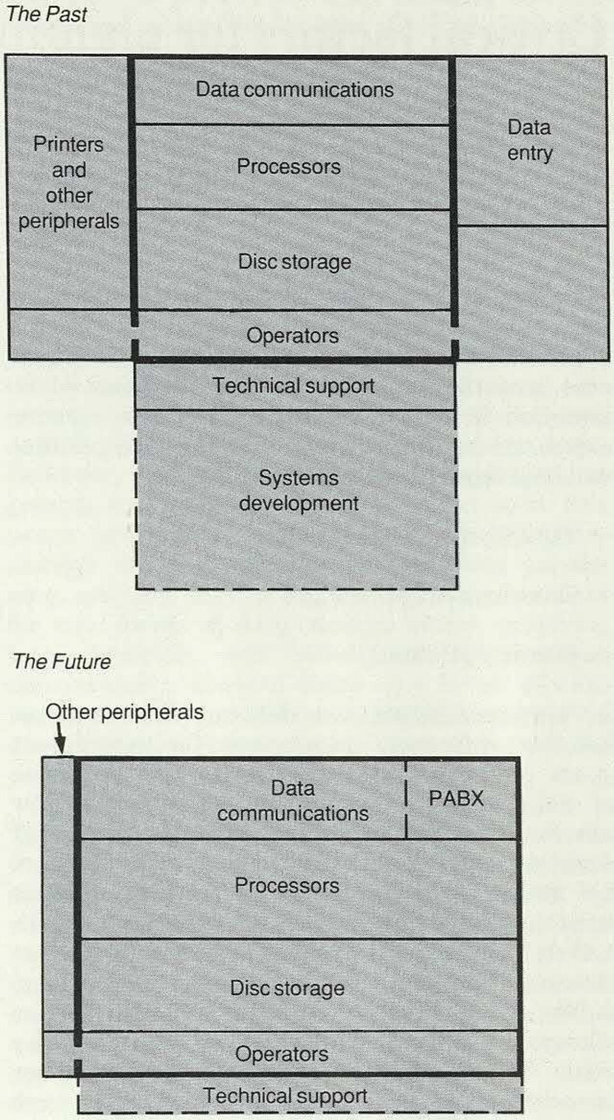


Figure 2.4 provides a pictorial representation of our view of the future data centre, and compares it with the data centre of today. As a result of these changes, the need for security at the data centre will increase because the data centre will be more vital to the day-to-day operation of the organisation.

It follows that the critical factors for planning the data centre will be:

— The space required to house the equipment.

— The physical environment for the equipment, and the remaining staff who will work there.

— The security of the data centre.

Each of these is considered in the next chapter.

Chapter 3
Critical factors for planning the data centre

Planning a new data centre requires a very wide range of inputs not only from the systems specialist, but also from architects, plant engineers, and the suppliers of the equipment to be installed in the data centre. For the reasons explained in the previous chapter, the planner will be primarily concerned with:

— Equipment space planning.

— Environment planning.

— Security planning.

In our researches, we did not find comprehensive reference documents for equipment space planning. We therefore discuss this area of data centre planning in some detail. For environment planning and security planning, several useful guidelines have been published by mainframe suppliers, by advisory bodies such as the CCTA for UK Government installations, and by specialist authors such as Howard Schaeffer (*Data Centre Operations*, published by Prentice Hall) and Robert Halper (*Computer Data Center Design*, published by John Wiley). In these areas, we therefore concentrate on highlighting those aspects that are not yet covered in the published guidelines.

EQUIPMENT SPACE PLANNING

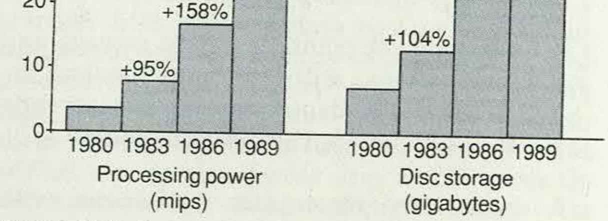
In Chapter 2 we described how data entry, data output devices, and unauthenticated data storage will increasingly be located in the user community. The floor space required for equipment in the data centre will therefore be determined by the amount of processing equipment, authenticated-data-storage equipment, and communications equipment. Of these three types of equipment, authenticated-data-storage devices will require the lion's share of space in the data centre. If an organisation can predict the floor space required for authenticated-data-storage equipment, it will have gone a long way to determining the overall space requirements for its data centre. The second most important factor in planning equipment floor space is the requirement for processing power.

Technological advances in data-storage and processing techniques are reducing the physical size (the 'footprint') of equipment for a given amount of storage or processing power. Nevertheless, the requirement for additional storage and processing power will grow at a rate that means the overall floor-space requirements for both types of equipment will continue to grow.

In the survey that formed part of our research, we asked the respondents about the growth in processing power and disc storage in their data centres. The results for 45 data centres for the period 1980 to 1989 are shown in Figure 3.1, and are summarised below:

— The average installed processing power, measured in millions of instructions per second (mips) or equivalent, grew by 95 per cent between 1980 and 1983, and by 158 per cent between 1983 and 1986. By the end of 1986, the average number of mips installed

Figure 3.1 Increases in processing power and disc storage requirements 1980 to 1989



(Source: Butler Cox survey of data centre managers) was 18.6. When asked to project their growth in processing power from 1986 to 1989, the same organisations predicted a growth of 95 per cent to 36.2 mips.

- The average amount of disc storage (in gigabytes) installed increased by 104 per cent between 1980 and 1983, and by 115 per cent between 1983 and 1986. By 1986, the average number of gigabytes installed was 29.4. The predicted growth between 1986 and 1989 is 77 per cent to 52 gigabytes.

We believe the predicted rates of increase for both categories are somewhat conservative, as do other industry sources, which predict much higher rates of growth. For example, Figure 3.2 shows how IBM believes the requirements for processing power and data storage will grow in large installations between 1984 to 1990. According to this data, large system users can expect a growth in online storage requirements in excess of 170 per cent a year. Whilst IBM's predictions may well contain an element of wishful thinking, we believe that the rate of increase will be considerably higher than that predicted by our survey respondents.

Our research also indicated that, regardless of the effort put into forecasting increases in processing power or data storage, most data centre planners have, in the past, underestimated the actual increases. We believe that the inability to forecast accurately is due to the inherent unpredictability of resource growth, and not to incompetence on the part of data centre planners. This situation is unlikely to change in the foreseeable future.

Figure 3.3 shows that the average data centre floor area in 1980 reported by our survey respondents was 310 square metres. By 1983, the average had grown by 27 per cent to 395 square metres. The growth over the next three years to 1986 was 16 per cent to 460 square metres. Between 1986 and 1989, our survey respondents predicted that the average floor area will grow by a further 17 per cent to 540 square metres.

Figure 3.2 IBM's view of the likely growth of computing resources in large installations

Resource	1984	1990
Processing power (mips)	29	289
Online storage (gigabytes)	102	1,157

Given the tendency in the past to underestimate the growth of processing power and data storage, and the predictions from industry sources referred to above, we believe that the projected growth in Figure 3.3 of 17 per cent over three years is likely to be understated.

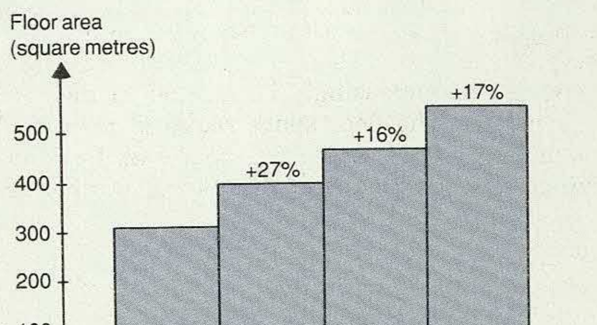
We now examine in more detail the factors that will determine the floor space requirements for the various types of equipment housed in data centres.

DATA-STORAGE DEVICES

The amount of data stored at a typical data centre has grown considerably over the past few years, and it is tempting to think that, soon, organisations will run out of new data to store. However, we believe that the present rate of growth will continue for at least the next five years, and may even increase. Whilst the data-storage demands of conventional data processing systems may level out, the requirements for new forms of data storage (voice, graphics, text, facsimile, and video) will all increase substantially. Each of these new forms of data requires much greater amounts of storage than conventional systems (see Figure 3.4 over the page).

Thus, more data storage capacity will be required for online storage and for archival storage. More floor space will be required for online storage because this will continue to be based on conventional disc technology for the foreseeable future. However, the floor space required for archival storage will be contained as cartridge tape and mass-storage devices replace magnetic tapes, and as optical discs come into use. The technical limitations of optical discs mean that they will not replace conventional discs for online storage.

Figure 3.3 Data centre floor space requirements 1980 to 1989



(Source: Butler Cox survey of data centre managers)

More space for online storage

The requirement for online data-storage devices is the key factor in determining the amount of floor space required in a computer room. The trend is firmly towards the use of fixed-disc units that provide greater storage density, increased reliability and 'hands-off' operation. Most disc units purchased during the past five years have been fixed discs. Even so, disc drives and disc controllers occupy between two and four times as much floor space as the processing units.

Although the capacity of discs has increased considerably over the past few years (and will continue to do so for the foreseeable future), the physical size of the cabinets has not changed appreciably. However, in addition to technological improvements (such as thin-film discs and heads) that are improving the data-storage density, suppliers will increasingly be offering repackaged or stackable drives, giving double or more capacity for the same floor space. Early versions of 'double-density' disc drives were, in effect, two single-density drives stacked one on top of the other. This approach enabled the 'dead' vertical space in the data centre to be used, thereby doubling the amount of data that can be stored per square metre of floor space before true double-density technology was available. The suppliers correctly judged that user organisations are more concerned with factors such as reliability and compatibility than with the size or appearance of equipment.

The same technique is likely to be used when the first 'quadruple-density' drives become available in 1988. They will probably be two double-density drives stacked one on top of the other. There are problems with this approach, however. Speaking at a Foundation management briefing in early 1987, the representative of one manufacturer said of a prototype double-density unit "when we found that our unit was too big to fit in the hold of a 747 we knew we had to modify it". The implications for data centre planning are that much greater floor loadings and ceiling heights will be required, and the access routes to the data centre will need to be higher. The alternative would be to assemble the equipment on the customer's premises, which would not be acceptable to many organisations.

Figure 3.5 illustrates a typical 'repackaged' double-density disc drive product and provides data about its physical dimensions and floor loading.

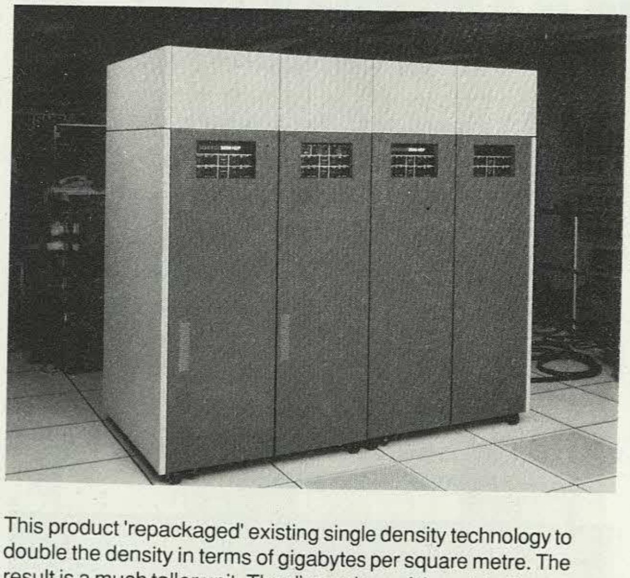
Despite these developments in disc technology the demand for online storage is growing at a faster rate than the increases in data density on the discs, so the floor space required for disc drives will continue to increase. One of the suppliers of IBM plug-compatible disc drives believes that online storage (in terms of gigabytes) in a typical large installation will increase by between 30 per cent and 50 per cent a year for the foreseeable future. The same supplier also estimates that the density of disc

Figure 3.4 Storage required for different types of data

Type of data*	Relative amount of storage required
Structured data or text	1
Free text (500 words)	3
Line drawing	10
Facsimile (monochrome)	30
Spoken words (500 words)	100
Video images (4 images in colour)	300

* Each example refers to what could be represented on one A4 sheet of paper. Most data stored in computer systems today is in the form of structured data or text.

Figure 3.5 Memorex's 3680 HDP double-density disc



This product 'repackaged' existing single density technology to double the density in terms of gigabytes per square metre. The result is a much taller unit. The dimensions of the cabinet are:

- Height: 82" (2,080mm)
- Width: 89" (2,280mm)
- Depth: 42" (1,065mm)

The unit weighs 4,144lbs (2,110kgs)

drives (in terms of gigabytes per square metre of floor space occupied) will increase by about 25 per cent to 30 per cent a year.

Taking both these estimates to be realistic, the likely range of increase in the data centre floor-space required for online storage is shown in Figure 3.6. At the high end (50 per cent increase a year in online gigabytes and 25 per cent increase a year in density), the floor space required for data storage will increase by 20 per cent per year, which means it will more than double after four years, and triple after six years.

Based on this analysis, we believe that most large data centres should be planning to increase the floor space for data storage by about 7 or 8 per cent a year, which corresponds with a 40 per cent increase a year in online storage and a 30 per cent increase per year in disc density. This means that the floor space for data storage requirements will double within ten years.

Cartridge tape and mass storage will replace magnetic tape

We have also noted a trend for databases to be duplicated (on discs) at different sites, for purposes of resilience and backup. This trend means that there is an even greater need for floor space for disc drives. However, it also means that many data centres have been able to reduce the number of magnetic tape drives (and the amount of tape storage), because magnetic tape is used less for archival purposes. Where tape drives are retained, they are increasingly being located outside the main computer room. Eventually, magnetic tape drives will be replaced by cartridge tape files and mass-storage devices that require little operator intervention. (The international insurance company case history on page 25 describes how the space required for magnetic tapes has been reduced in this way.)

Mass-storage devices will continue to occupy a relatively large floor area, however, but their storage density (in terms of gigabytes per square metre of floor space) will be much higher than for fixed disc units. Improvements continue to be made to the technology of cartridge-storage (and other similar) devices, but these developments are likely to be overtaken by developments in optical laser-recording techniques.

Optical discs will not replace conventional discs

Laser-recording techniques can now provide storage capacities of one or more gigabytes on a 12-inch optical disc. The main advantages of optical discs are that they are robust and reliable. Furthermore they can be duplicated easily if many (hundreds or thousands, rather than a few) copies of a database are required. The main drawback at present of optical discs is that they provide essentially a write-once read-only storage medium. For several years, manufacturers have been promising to develop an updatable optical disc, and have claimed that such products will be available"within two years". Sooner or later, an updatable optical disc will be available as a commercial product. However, at the time of writing this report, we can see no clear evidence as to when that will be.

Another disadvantage of optical discs is that the access time is inferior to that of magnetic discs because it is not possible, for engineering reasons, for the read head to 'fly' over an optical disc as quickly as it can over a conventional disc. Furthermore, the bit error rates are higher than with conventional discs, which means that complex error correction procedures are required if data accuracy is a paramount concern. The complex procedures also slow down the data transmission rate. For all these reasons, optical discs will be used primarily for archival storage of images. When transmitting an image, it is not essential that every bit of data is transmitted correctly. A relatively high bit error rate can be tolerated before a significant degradation in the image is perceived. This means that complex error correction procedures are not

Figure 3.6 Growth in floor space required for online storage

% increase pa in online gigabytes	%increase pa in disc density	Resultant % increase in floor space	Number of years before floor space required doubles
25	30	-3.8	Floor space reduces
25	25	0.0	Floor space stays constant
30	30	0.0	Floor space stays constant
30	25	4.0	More than 10 years
40	30	7.7	9
40	25	12.0	6
50	30	15.3	5
50	25	20.0	4

required for image transmission. Thus, optical discs will not be seen as a replacement for online storage, which will continue to be based on conventional magnetic discs.

Nevertheless, optical-disc technology is developing rapidly. As it is perfected, the cost advantages of optical discs, and their huge data storage capacity, will ensure that they are accepted rapidly, particularly for archival storage. When updatable optical discs become available, 'jukeboxes' of optical discs are likely to replace magnetic mass-storage devices.

Developments of this kind could result in the 1990s in significant reductions in floor-space requirements for archival data-storage devices. Recent experiments conducted by the Central Computer and Telecommunications Agency in the United Kingdom have shown that two gigabytes of data can be stored on one double-sided 12-inch (30 cm) optical disc, which results in a considerable space saving compared with current magnetic disc storage. In addition, optical storage media have far less stringent environmental requirements than magnetic media and can therefore operate in a normal office environment.

PROCESSORS

During the past few years, the amount of raw processing power per square metre of floor space occupied has increased by about 50 per cent per major hardware upgrade (in other words, about every three to five years), and we believe that this trend will continue for next few years (see Figure 3.7). This means that the processing power per square metre of floor space is growing at between 10 and 15 per cent a year. However, as Figure 3.1 showed, the demand for processing power has grown at a faster rate (between 25 and 35 per cent a year). Figure 3.8 shows what these growth rates mean for the total floor area occupied by processors. The figure shows that most organisations should probably be planning to increase the floor space for processors by about 13 per cent per year, corresponding with a 30 per cent increase per year in mips, and a 15 per cent increase per year in mips per square metre.

But, mips per square metre is not the only factor governing the amount of floor space required by processors. Manufacturers also need to house auxiliary equipment such as power supplies and cooling equipment within processor cabinets, and they have not been able to reduce the size of these components to the same extent that processor circuits have been reduced. The need to provide maintenance access to the cabinets also means that it is not possible to reduce the size of the cabinets significantly.

Another factor preventing significant reductions in the physical size of processor cabinets is the trend to multiple-processor mainframe computers. In the future, 16 or more processors will be housed in a cabinet of similar size to (or perhaps slightly smaller than) today's single-processor mainframes. In the immediate future, these developments will be made possible by the use of denser chips and by improved packaging and interconnection techniques (such as the use of optical fibres for interconnections to and within the processor), rather than by significant improvements in basic processor technology.

Figure 3.7 Increase in processing power per square metre of floor space

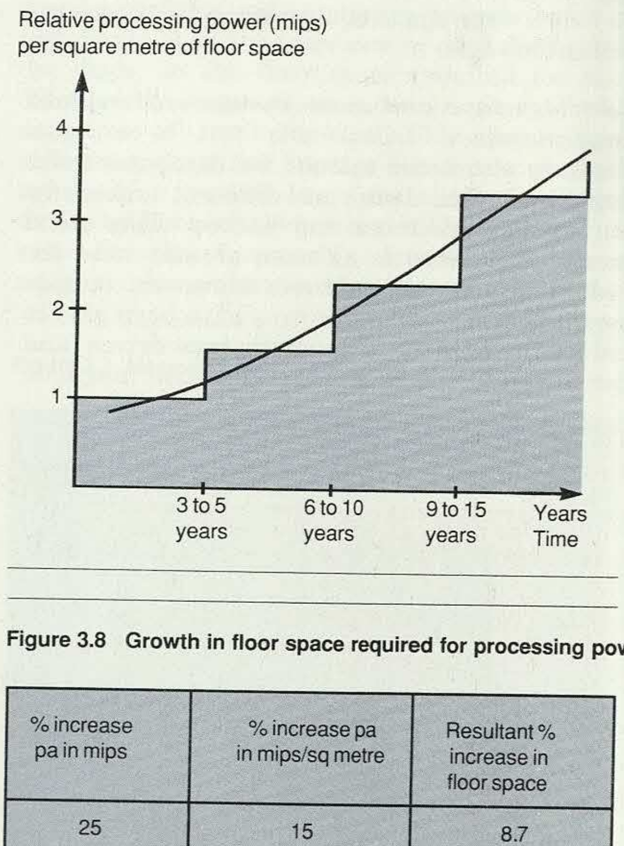


Figure 3.8 Growth in floor space required for processing power

% increase pa in mips	% increase pa in mips/sq metre	Resultant % increase in floor space
25	15	8.7
25	10	13.7
30	15	13.0
30	10	18.1
35	15	17.4
35	10	22.7

There is also a trend towards the greater use of specialised processors, and this will affect the amount of floor space required for processors in the data centre. Most of the suppliers we spoke with during our research said they believed there would be a move towards specialised processors to handle specific tasks. In the same way that front-end communications processors were installed to handle the growing communications processing load, the functions currently performed by the typical mainframe will be subdivided — for example, into a database machine and a transaction processor. Thus, although the individual processor cabinets may be smaller, there will be more of them and, overall, they will require more floor space.

In some organisations, the data centre may house several smaller processors of different types, each dedicated to a particular part of the business. In effect, the processing will be logically decentralised, but the equipment used to run the applications will be managed centrally.

In the longer term, beyond the mid-1990s, there may be more radical changes in processor technology, as research into optical processors and 'fifth-generation' computers results in commercial products. As yet, it is too early to speculate about the impact these developments will have on data centre space requirements. For the foreseeable future, these impacts will remain beyond the typical data centre planning horizon.

COMMUNICATIONS EQUIPMENT

Like other electronic equipment, communications equipment has become smaller over the last ten years. Since the mid-1970s, a 2400-baud modem has reduced from the size of a typewriter, to a rack-mounted card, and single-chip modems are now available for incorporation into terminals. In the early 1970s, a 9600-baud modem required a box measuring 43x 26.5x 58.5 cm. Today, the equivalent modem requires a box measuring 38x9x35.5 cm, which means the physical volume of the modem has reduced by more than 80 per cent.

Similar trends have affected the size of voice communications equipment. In the mid 1970s, a 400-line PABX would have required a room of 34 square metres. Today, a 500-line digital PABX requires only 17 square metres — a reduction of 50 per cent. Although the number of telephones installed in the typical office has increased slightly during the past ten years, this has not usually been sufficient to offset the reduction in space requirements caused by technological developments. Over the next ten years we expect to see further reductions in the size of voice-switching equipment. There will also be some reduction in the space required for distribution frames because of the increasing use of wideband circuits (often ISDN primary-rate bearers), rather than multipair cables, for:

— Intersite links.

— Site-to-public network links.

— Links between the central equipment and distributed 'multiplexors' located in wiring closets.(Future PABX technology was discussed in Foundation Report 54 — Integrated Networks.)

However, the growth in the numbers of terminals, networked personal computers, and departmental and distributed processors has often outrun improvements in communications technology, leading to an overall increase in the space required for data communications equipment. Over the next five years, the growth in the number of terminals and workstations will, in many organisations, come to an end as the ratio of terminals-to-staff approaches one-to-one. At the same time, the introduction of ISDN and local area networks will favour the use of multiplex connections, further reducing space requirements. (A recently announced LAN-to-mainframe gateway can support several hundred concurrent calls, for example.)

These trends may be offset to some extent by increases in the functionality of communications equipment rather than in the numbers of devices, with a consequent increase in space requirements. For voice, the extra functions could include voice-mail and voice-response systems, whilst for data they might include electronic mail and protocol conversion. Many of these extra functions are, however, closer in nature to data processing and office systems than to basic communications. As such, they will probably be provided by workstations or the central computing complex, not the communications system.

The net result is that the space required for communications equipment is likely to decrease over the coming ten years, although it may increase in the short-term in some organisations.

ENVIRONMENT PLANNING

It is not our purpose to include here detailed information on the physical and environmental planning of a data centre. Most large organisations have a wealth of expertise in this field, and there are supplier and other reference books readily available to guide planners. However, our research did identify several aspects of environment planning that may not be so well known and these are discussed below.

ELECTRICAL SYSTEMS

A data centre cannot operate without a reliable and stable source of electrical power. The electrical power requirements of a data centre differ significantly from those for ordinary office buildings, both in terms of the electrical load and the quality of the supply that is needed. Average loads in a computer room can range from 25 to 50 watts of electrical power per square foot of floor area (far greater than in conventional offices). Figure 3.9 lists the typical power requirements for different areas of the data centre. Furthermore, the fluctuations in commercial power supplies, which are acceptable for ordinary industrial and commercial uses, are not acceptable for computing equipment. Data centres require so-called 'clean-line' electrical supplies. The electrical power supply arrangements in a data centre also need to provide the flexibility both to relocate existing equipment and to add new equipment. Other difficulties concern the different power requirements (in terms of voltage and phasing) of different equipment.

Figure 3.9 Data centre electrical power requirements

Area	Power requirements (watts per square foot)
Mainframe computer room	25 to 50
Printer room	25
Command centre	20
Telecommunications	25
Tape/disc library	5
Input/output	5 to 7
Storage	2
Programming	5 to 7
Data entry	7
Administration	5 to 7
Equipment maintenance*	

* Depends upon equipment requirements

(Source: David Oxman, P.E. of Butlan and Oxman, Engineers)

For all these reasons, the provision of electrical power for a data centre must be considered at the earliest possible stage in the planning process.

Many of the problems associated with data centre power requirements have been resolved in recent years by the development of pre-packaged power units, connected to the computer equipment via flexible cables under the raised floor. The main advantage of this type of equipment is that it provides a factory-assembled and tested power source (including transformers) that electrically isolates the equipment from the building power supply. The shielded transformers and surge suppressors in the power units provide a reliable and stable power supply suitable for computer equipment.

Most data centres also require a standby power supply that will be used in the event of the mains supply failing. In most cases, a few seconds (or minutes) downtime whilst the standby supply is switched on will be acceptable. However, if even a momentary break in the power supply is unacceptable it may be necessary to install uninterruptable power supply (UPS) equipment. The cost of such equipment is very high and in most cases can be justified only where the continuous operation of the data centre is vital to the functioning of the organisation.

AIR-CONDITIONING SYSTEMS

The introduction of 'lights-out' computer rooms reduces the problems of environmental control systems. Absence of staff movement means that the clean air flow is uninterrupted so that the equipment can be kept at a constant temperature. Furthermore, dust particles are not introduced into the controlled atmosphere. It has been shown that even relatively minor movements of operators introduce measurable impurities into the air that circulates through the installation. For example, brushing a hand across the sleeve of a 'clean-room' uniform can increase the number of particles in the atmosphere by between 50 and 200 per cent. The mere presence of people disturbs the air flow and creates areas of turbulence that pull in contaminated air from outside.

Conventional air-conditioning systems keep the air at a constant temperature and remove any solid particles from the air. They do not filter out specific gases, however, and disc-drive manufacturers have been able to trace the cause of apparently random failures to the presence of hydrocarbon gases in the computer-room atmosphere. In one case, the air intake for the air-conditioning system was positioned above the loading bay for the organisation's delivery trucks. Diesel fumes were being sucked into the air-conditioning system, and hydrocarbon gases found their way into the disc drives. Once the air flow was changed, the disc drive failures stopped occurring.

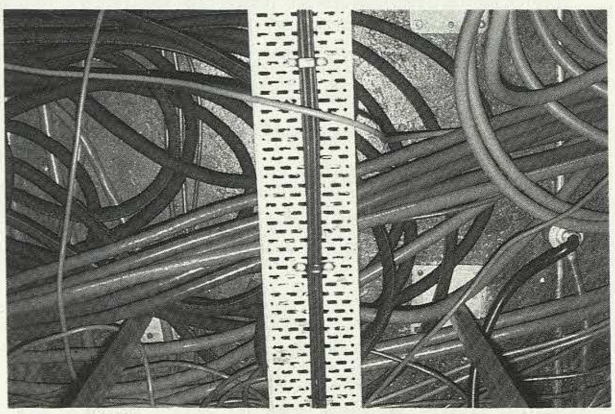
Disc drive problems have also been traced to dust particles being lifted from the void under the raised floor in the computer room. If the concrete floor has not been sealed properly during construction, fine concrete dust may be blown into the units, bypassing the filters that normally extract 99.99 per cent of particles.

Environmental problems such as these are extremely difficult to track down, and require very special expertise. The difficulty is that the cause of equipment failures is often not recognised as being due to an environmental problem. Worldwide, there are very few experts on this specialised field - the only one known to us is 'Storm' Larkins, who works for the UK-based Winton Industries.

FLOORING

We found false floor heights in large computer rooms varying from 23 cm (which is too small) to 120 cm (which is over generous). The majority were between 45 cm and 90 cm. One common problem with false floors in computer rooms is that the smart floor panelling hides a tangled web of cables that have accumulated over the years (Figure 3.10 is a photograph of a typical installation). Many suppliers are in the habit of leaving old cables in place when new equipment is installed because it is time-consuming and difficult to remove them. The suppliers are usually under pressure to install new equipment as quickly as possible, and removing the old cables would slow them down. The resultant large accumulation of redundant underfloor cables can lead to problems with future equipment changes, air flows, and fire hazards. It can also mean that a new wiring scheme that could solve the problem (optical fibres, for example) could not be installed without shutting down the complete installation because all of the old cables would have to be removed to make space for the new cables.

Figure 3.10 Underfloor wiring can easily get out of control



We therefore advise data centre managers to ensure that the old cables are removed when new equipment is installed, even if this means that it will take slightly longer to install the equipment.

The loading on raised floors can also be a problem. Several Foundation members were obliged to reinforce their floor during 1986 to accommodate their supplier's latest processor, whose point weight exceeded all previous loadings. Access paths to the installation area had also to be reinforced. (One organisation forgot to do this, and the new equipment fell through the floor.)

CEILING HEIGHTS

Earlier in this chapter, we described some of the techniques now being used by disc manufacturers to increase storage density per square metre (stackable drives and taller units). This means that data centre planners may have to allow for greater ceiling heights both within the data centre and in the access routes to it. (Figure 3.5 showed the dimensions of one such disc unit.)

FIRE PREVENTION

In a survey of Foundation members, more than half admitted that a fire had occurred in their data centre, although all of the fires were small and localised, and had not affected the operations of the computer systems. About half of the fires had occurred outside the computer room (in the mail room, electricity supply area, and so forth). The remainder of the fires were caused by minor electrical faults in hardware units.

All of the organisations had installed automatic fire-detection systems and some form of extinguishing system. Opinions varied widely on the effectiveness of different types of systems (water, carbon dioxide, Halon, and foam), and on the amount of damage that each type could cause to equipment and staff. In several cases, Halon and carbon dioxide systems had been dismantled because of fears about health and safety hazards.

However, the 'lights-out' concept reduces these fears because the room is usually unoccupied. Water sprinkler or gas flooding systems, which are automatically activated when heat or smoke is detected, can be installed in such rooms in the knowledge that there is less chance of people being present if the system is activated.

STAFF FACILITIES

Environmental planning for the data centre should take account of the people requirements as well as those for the equipment. It is easy to become so concerned with the environmental conditions required by the equipment that will be installed in the data centre, and forget that the staff who will work there have very different requirements.

In many data centres, however, it is no longer necessary to provide accommodation for resident maintenance engineers. Improvements in machine reliability and the introduction of remote diagnosis are removing the need for on-site engineers in all but the largest, most changeable installations. In addition, the introduction of a 'lights-out' computer room removes the need for special rest rooms where the operators can recuperate from the uncongenial environment of the computer room, with its chilly atmosphere and constant hum of equipment. Instead the operators and network control staff are located in the data-centre 'bridge', which is furnished in the same way as a normal office.

SECURITY PLANNING

We mentioned in Chapter 1 that data centre security was one of the key concerns of Foundation members. Clearly, the extent to which an organisation is prepared to go in ensuring security depends on the nature of its business, its dependence on computer systems, and on its local situation, which determines how likely it is to experience an event (whether natural or man-made) that could lead to a breakdown at the data centre.

The first step in security planning is to analyse the risks to the organisation arising from a loss of data processing or telecommunications services. There is a wealth of information readily available on this subject (see, for example, Foundation Reports 33 — Managing Operational Computer Systems — and 51 — Threats to Computer Systems), and we do not elaborate further on it here.

Once the risk analysis has been carried out, the security measures for minimising the risks must be considered and costed. Usually, there are two costs to consider: the cost of installing the security measure and the cost of operating it. Figure 3.11 (opposite) shows some of the security measures that can be considered for the site, for the services to the site, for the building, and for the computer room. For each measure, the level of security provided is ranked as high, medium, or low, as are the installation cost and operating cost. Once the cost of each of the security measures has been ascertained for the particular data centre environment, the planners can provide senior management with recommendations for the appropriate level of security, setting the costs of the measures against the potential costs of the risks.

This type of analysis has led some organisations to build new data centres that have security features similar to those found at military installations. One example (a major clearing bank) is described in the case history on page 26. The security measures that can be taken include:

- Placing the data centre in a subterranean structure that is isolated from the outside world.
- Constructing the data centre as a dome shape, or with sloping exterior walls, so that the energy from a bomb blast will be deflected upwards away from the facility.
- Creating an earthen wall around a single-storey data centre so that the energy from an external bomb blast will be deflected.
- Constructing the data centre as a concrete block within a concrete block so that more than one blast will be required to breach both walls.
- Positioning the data centre so that its external walls are at least 50 metres from public roads so that car or lorry bombs are not likely to have a serious impact.

Figure 3.11 Checklist of data centre security measures

Security measure	Security level	Installation cost	Operating cost
The Site			
Away from population centres	High	Low	Medium
Large enough for the building to be at least 50 metres from the perimeter at all points	High	High	Low
Free from flooding, landslips, earth movements	High	Medium	Low
Away from flight paths and major roads	High	Low	Low
Not Overlooked	High	Medium	Low
The Services to the Site			
Single road access	High	Low	Low
Underground power cables from two separate substations	High	High	Low
Telephone lines from two separate exchanges	High	Medium	Medium
Adequate water supply	Medium	Low	Low
The Building			
Low rise	High	High	Low
Entirely shielded from view and blast by surrounding mound	High	High	Low
Double perimeter fence	Medium	Low	Low
Car park well away from building	High	Medium	Low
Vehicle entrance to building area restricted to delivery vehicles	Medium	Low	Medium
Single entrance for all personnel	High	Low	Low
Security staff manning entrance at all times	High	Medium	Medium
Full surveillance equipment in entrance security room (CCTV screens, fault and fire detection alarms, etc.)	High	High	Medium
The Computer Room			
Card and keypad access control	Medium	Medium	Low
Heat detectors	High	Medium	Low
Smoke detectors	High	Medium	Low
Automatic Halon flooding	High	High	Low
Manually initiated Halon flooding	Medium	High	Low
Water sprinklers, automatic	Medium	Medium	High

SUMMARY

In this chapter we have considered the three critical factors that need to be considered when planning a data centre. We now need to consider how these critical factors may be brought together with the other specialist and financial factors. This is the subject of the next chapter, which provides guidelines for planning the data centre.

Chapter 4

Data centre planning guidelines

The cost of designing and constructing a new corporate data centre is enormous. Yet our research has shown that many data centres outgrow their accommodation well before the date anticipated when the original plans were drawn up. The result is crowded working conditions, overheating because the air conditioning cannot cope with the additional equipment, and difficulties in updating equipment and applications. The cost of subsequent renovations (excluding equipment costs) can exceed $1,500 per square metre, compared with $325 to $645 for general office space.

During our research we were told of many reasons for running out of space in the data centre earlier than was planned. Some of these reasons were within the overall responsibilities of the systems function and could have been anticipated with a little more forethought. Examples included:

- Inaccurate projections of the growth of existing applications.
- Insufficient allowance made for new applications.
- Insufficient allowance for the take-up of new technologies.
- Insufficient contingency allowances.
- Inability to persuade senior management to allocate funds for capital expenditure to provide 'spare' space.

Other reasons were outside the direct responsibilities of the systems function, and would have been more difficult to anticipate even though they had a significant impact on the data centre. For example:

- Rationalisation of the information systems function, which led to the closure of several data centres, with the workload being transferred to a central data centre.
- An unforeseen takeover of another organisation, whose data processing workload was transferred to the existing data centre.
- Launching of new products or services requiring computer support, and which were not foreseen when the data centre was originally planned.

As an example of how difficult it is to predict accurately the space that will be required, one of the case histories describes how the same organisation (an international insurance company) both overestimated the space required for stationery storage and underestimated the space required for communications equipment.

We believe that a systematic approach to data centre planning should be able to take account of most of these variables. Our research did not discover an all-embracing methodology, however. Instead, each organisation uses a different mix of criteria in its approach to data centre planning. Appendix 1 contains selected case histories that illustrate the various approaches described during the research. For example, some organisations calculate the anticipated workload by estimating the future number of terminals and extrapolating from the computing power and storage requirements needed to support existing terminals. The drawback of this approach is that no attempt is made to take account of significant changes in the processing-power or data-storage requirements of individual applications.

We have analysed the approaches currently being used, selecting the best practices from each. From these, we have been able to construct a set of guidelines for data centre planning, and present these in this final chapter. We have concluded that there are four key guidelines to data centre planning:

- Recognise that the space requirements estimates will be inaccurate, and plan accordingly.
- Where possible, move to a greenfield site.

— Design the premises on a modular basis that can provide flexible accommodation.

— Set a realistic budget.

SPACE REQUIREMENTS WILL BE INACCURATE, SO PLAN ACCORDINGLY

Data centre planners, and general management, must recognise that, for the reasons given earlier in this report, it is not possible to predict accurately the equipment space requirements for a new data centre. Plans for a new data centre should therefore not be based on a single estimate of equipment space requirements. Instead the upper and lower bounds need to be forecast and a risk analysis undertaken to determine the likelihood and penalties of providing insufficient space, and of providing too much space that remains unused. (The case histories in Appendix 1 provide examples of organisations that have suffered at both of these extremes.)

In the end, the initial equipment space allowed for will be a trade-off between the cost of providing too much space and the penalties of providing too little space. The potential cost to the business of providing too little space can be very large if the data centre cannot accommodate new equipment required by changed business needs. There should also be contingency plans to cater for the situation where either too much or too little space has been allowed for.

Clearly, the overall space requirements of the data centre must allow for the staff who will work there as well as for the equipment. However, the uncertainty in space forecasting is not critically dependent on staff space.

We propose that a four-step methodology be used to predict the equipment space requirements for the future data centre and to conduct the risk analysis:

— Extrapolate recent equipment-resource trends into the future.

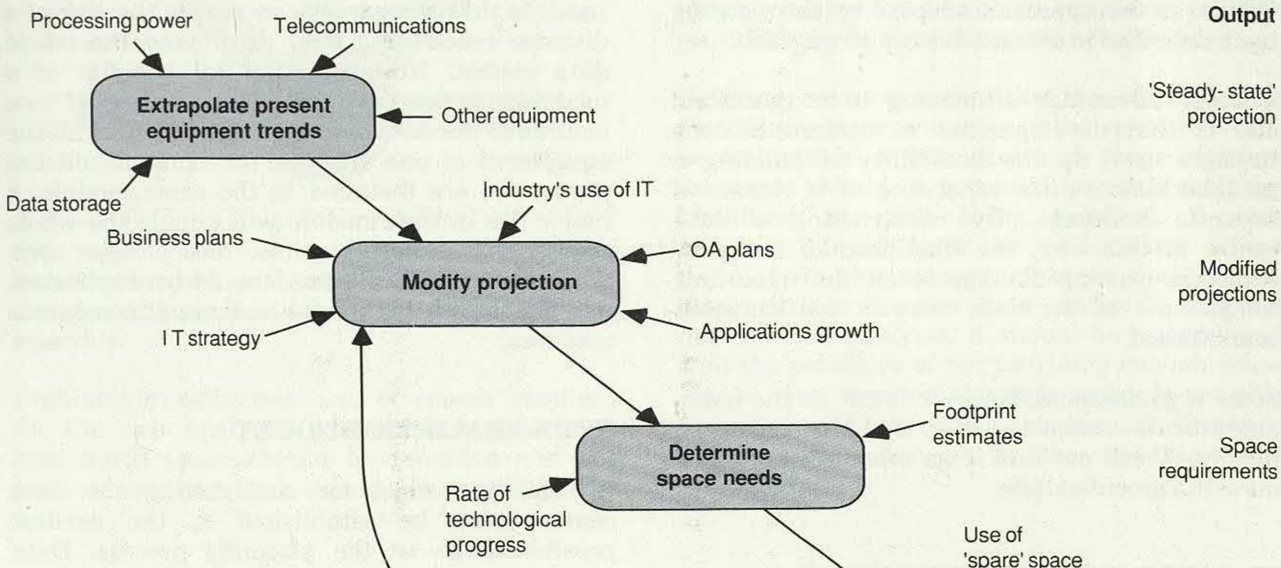
— Modify the projection to take account of possible developments in the use of information technology.

— Translate the modified projections into space requirements.

— Assess the risks of not providing the maximum space.

The steps of this methodology are described in detail in Appendix 2 and are summarised in Figure 4.1. In essence, the methodology begins

Figure 4.1 Four-step methodology for equipment space planning



by estimating the resource requirements (processing power, data storage, and communications) on the assumption that present trends will continue (the steady-state projection). This projection is then modified to take account of possible developments that could occur in the use of information technology, and a range of possible resource requirements is created. These requirements are then translated into a range of possible space requirements. Finally, a risk analysis is carried out to determine the risks of not providing enough space to meet all of the possible (but less likely) requirements.

MOVE TO A GREENFIELD SITE WHEREVER POSSIBLE

Organisations should consider the option of moving their data centre to a greenfield site rather than assume that it must be located at one of the organisation's existing major sites. Many of the technical and staffing trends described in this report indicate that there will be fewer reasons for locating the data centre close to the user departments or even to the system development departments. There may also be direct economic advantages if the existing premises are in an area where the cost of space is high and if the new data centre can be located in a low-cost area. Furthermore, it will be possible to provide better security at a greenfield site because the security measures can be designed in right from the beginning. (This was the approach adopted by the clearing bank described in the case history on page 26.)

Another advantage of moving to a greenfield site is that developments in communications linkages open up the possibility of building a modular data centre consisting of a cluster of separate buildings. By constructing a data centre in this way, the likelihood of a fire or explosion causing damage to all the important components of the data centre's configuration is minimised.

Some organisations, however (such as the pharmaceuticals company described in the case histories), will not find it an economic option to move to a greenfield site.

PLAN FOR FLEXIBILITY BY USING A MODULAR DESIGN

The ideal way to gain flexibility for the future is to build a new data centre on a site with plenty of space around it to allow for expansion. Unfortunately for most organisations this is not a practical proposition. The alternative, therefore, is to consider whether the equipment currently located in the data centre must continue to be concentrated within one small area. In the past, technical constraints have meant that the different components of computer systems must be located close to each other. The basic problem was that to achieve high speed data-transfer rates between computer processors and storage devices, the cables had to be short.

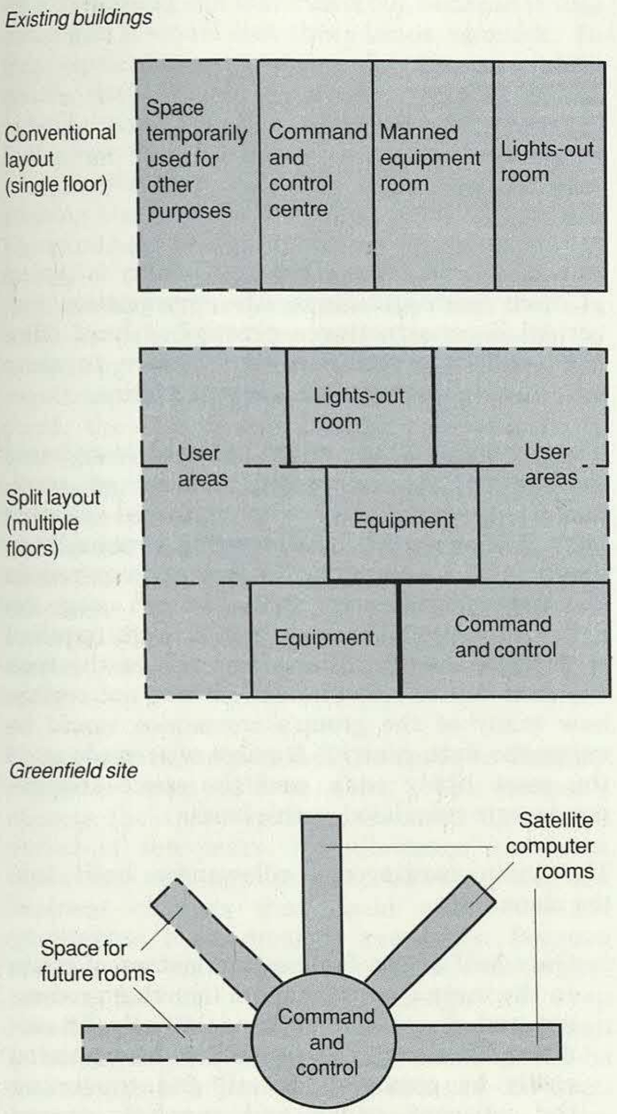
Many of the technical constraints still remain today, although some manufacturers are now providing optical fibre channels between equipment. Optical fibres channels allow data to be transferred at high speeds over greater distances. The implication is that it will be possible to adopt a modular design for the data centre by distributing it among different rooms within a site. (Technically it would be possible to create a data centre that was distributed over several sites, but this would not be economic in the immediate future). Thus, one approach to flexible planning would be to identify suitable space for expansion within the site that can be used for other purposes until it is required. (This approach is being used by the first organisation described in the case histories on page 24.) Possible arrangements for reserving space that will initially not be used, or will be used initially for other purposes, are depicted opposite in Figure 4.2.

An additional advantage of this modular approach is that companies can reduce the risk of a disaster (such as a fire) destroying the whole data centre. However, the full benefits of a modular design will not be obtained if one module of the data centre is used to house all the equipment of one type. If, for example, all the processors are installed in the same module, a major fire in that module will cripple the whole centre. Wherever possible, every key component of the installation should be duplicated, and the duplicates should be housed in separate modules.

SET A REALISTIC BUDGET

A realistic budget for constructing the data centre must be established at the earliest possible stage in the planning process. Data centre construction costs (excluding computer equipment) can be two or three times higher than for normal construction projects, and many organisations have badly underestimated the cost of constructing a new data centre. The result is that the 'ideal' centre is scaled down to conform to the original underestimated budget. One useful budgeting approach is for

Figure 4.2 Data centre layouts to allow for contingency and expansion



the project team to provide a preliminary construction cost estimate at the beginning of the planning process and to update it at regular intervals.

Preliminary estimates are, of course required for the risk analyses performed to determine how much space should be provided and the level of security that should be provided. The general management of the organisation will want to be involved in making the decisions that result from the analyses. Decisions about the trade-off between risks and the costs of eliminating those risks is a business management responsibility, especially when the potential consequences, and the costs of preventing them, are so high.

REPORT CONCLUSION

Planning, designing, and constructing a new data centre is a substantial undertaking, and requires considerable investment. Our research has shown that it is easy to make mistakes, and that the effects of those mistakes will be evident for many years.

In this report, we have established that there will be a continuing need for corporate data centres and we have identified the critical factors for planning a data centre (space planning, environment planning, and security planning). Security planning is much easier if the data centre is being constructed at a greenfield site because the security measures can be designed in right from the beginning. Environment planning is not just concerned with conventional air conditioning; computer equipment can be disrupted by very small levels of impurities in the atmosphere. The trend towards 'lights-out' computer rooms makes it easier to provide the controlled atmosphere required by computer equipment and to provide powerful fire extinguishing systems.

However, we have shown that it is impossible to provide accurate estimates of the space that will be required in a data centre. A range of space estimates should be prepared, together with the likelihood of each being required. A risk analysis should then be undertaken before deciding on the amount of space that will be designed into the new data centre. In carrying out the risk analysis, it should be remembered that the penalties of not providing enough space can be very severe if the data centre is not able to cope with increased equipment demands required by changes in the needs of the business.

Appendix 1
Case histories

In this appendix we present as case histories the experiences of six organisations. The cases have been selected to illustrate the range of approaches to data centre planning adopted by different organisations. They are presented anonymously because some of the organisations requested that their identities should not be revealed.

A GOVERNMENT-CONTROLLED INSURANCE INSTITUTE

This organisation is in the process of relocating its data centre to a high-rise building and is planning to distribute the data centre over several floors. However, it has not made a systematic attempt to predict how its computing workload will grow over the next few years.

The main computer room will house three central processors and two communications processors. In addition, 18 smaller rooms, two on each of nine floors in the high-rise building, have been allocated for data centre use. Not all of these rooms will be used initially, however, so a considerable contingency allowance has been built into the data centre plans. The rationale is that this approach will allow for gradual expansion into the smaller rooms as the workload grows. It will also provide greater resilience for the data centre because it is less likely that the entire building will sustain damage in the event of a catastrophe.

The anticipated workload was calculated by estimating the future number of terminals and extrapolating from the computing power and storage requirements needed to support the existing terminals. In calculating the floor-space requirements, the general reduction of equipment size was taken into account. No attempt was made to predict the growth in individual applications, however.

AN INTERNATIONAL CONSTRUCTION COMPANY

For many organisations, the opportunity to plan a new data centre arises because of a move to a new head office. This case history is typical of many such situations. The organisation concerned is constructing a prestigious head office that will incorporate a data centre to serve several of the group's operating companies.

The estimates of the space required were based mainly on the estimated number of workstations that will need to be supported (starting from a base of 600, and growing eventually to 2,800). The computing resources required in the data centre were estimated by using the average processing power and storage required to support a workstation. However, at the time the data centre was planned, it was not certain how many of the group's companies would be using the data centre. A guess was made as to the most likely ones, and the space requirements were calculated on this basis.

The main contingency allowances built into the plans are:

— Only half of the 300 square metres of space in the main computer room (housing processors and discs) will be used initially. In addition, this room can (in theory) be expanded easily by removing a wall and converting the adjacent paper and supplies storage area. This would expand the computer room by a further 25 per cent; the storage area would need to be relocated, however.

— The 150 square metre room housing the operators will be fully occupied right from the start. However, it has a removable partition, which would allow it to be expanded by 50 per cent. The additional space that would be used will initially house the data control office. Expanding the operators' room would require the data control office to be relocated.

A MAJOR PHARMACEUTICALS COMPANY

This organisation is building a new computer room at its existing site. The main requirements are to permit an expansion of the computing services provided and to improve the security of the computing facilities. The alternative of constructing a new data centre on a greenfield site was ruled out because it was estimated it would cost three times as much. In this particular case, space for the new data centre could readily be made available at the existing site, and the alternative site was in fact a few hundred metres away in the suburbs of a major European city. The cost of purchasing the land for the alternative site would therefore have been prohibitively expensive.

Again, a thorough, systematic approach was not used to calculate the anticipated growth in computing requirements and the consequent requirements for data centre floor space. Instead, the size of the existing computer room was taken as the base line, and future growth in processing power and storage was extrapolated from the past growth. However, an allowance was made for reductions in equipment size when calculating the floor space required. For example, the company expects eventually to replace 16 IBM 3350 disc drives with two IBM 3380 drives, with a consequent reduction in floor space.

To begin with, only half of the space in the new computer room will be used. The company expects the spare space to be used up over a period of ten years. No allowance has been made for new applications, however, nor for business changes that could affect the requirements for computing resources. Because of this, we would be surprised if this organisation did not have to redesign (or even rebuild) its data centre much earlier than anticipated. The new computer room cannot be extended into adjacent areas because the ducting in the ceiling in these areas leaves insufficient vertical space for a false floor.

A MAJOR INTERNATIONAL INSURANCE COMPANY

This case history is presented in more detail than the others because it illustrates many of the factors that have to be taken into account when planning a data centre. In particular, it illustrates how difficult it is to estimate accurately the space required.

As an international insurance company, this organisation has a massive transaction-processing load. To accommodate this load, in 1979 it constructed a very large computer hall (1,100 square metres) inside a purpose-built building. When the hall first opened, it looked ridiculously empty, with only 25 per cent of the space utilised. Five years later the hall was full. The growth in equipment was due partly to the increased processing power (the original two IBM 370/158s had been replaced by an IBM 3081K and an IBM 3084Q), but mainly to the growth of online storage devices. The disc capacity grew at 40 per cent a year, and rapidly filled the hall.

Eighty-five per cent of the space in the hall is now used for production equipment, with the remaining 15 per cent reserved for assembling and testing new equipment prior to final installation. This requirement was not anticipated when the computer hall was planned in the late 1970s, and is caused by the need to test online and distributed systems.

One factor that has helped to prevent even more space being required has been the replacement of magnetic tape drives by cartridge equipment, which reduces the floor space by a factor of two-and-a-half for a given storage capacity. The use of cartridges has also helped to reduce the space required in the tape library (which now houses 17,000 tapes and 8,000 cartridges). A cartridge storage rack is only two-thirds as deep as a tape storage rack. Moreover, twice as many cartridge racks can fit into the vertical space, and 10 cartridges can be stored in the horizontal space required by six tapes.

By contrast with the main computer hall, the space allocated to the stationery store has proved to be 50 per cent more than required. When the data centre was originally planned, there were abnormally long lead times for paper deliveries, and the stationery storage space was planned on the assumption that this situation would continue. On the other hand, the space allowed for telecommunications equipment was seriously underestimated. The telecommunications equipment room now looks more like a store room than an operating area.

As the space in the main computer hall was used up, this insurance company opened a second data centre, which was located in the same area of the city. A mainframe (currently an IBM 3090/200) was installed in the second centre, mainly for development work. However, the company realised that major disasters (a water supply failure, or the area of the city being cordoned off because of a terrorist siege, for example) could disrupt both data centres, so it has decided to develop another data centre 65 kilometres away from the main site. The new centre will comprise 3,150 square metres, of which 790 square metres will be used as a computer hall. The computer hall will house only processors and disc drives; tape drives, printers, etc will be housed elsewhere in the data centre.

Capacity planning for the new centre is being carried out by a specialist team comprising a manager (who has other responsibilities as well) and five full-time staff. The planning process is broadly as follows:

— Users are given their actual usage figures and are asked to predict the likely growth in existing applications.

— Systems development staff provide forecasts of the processing loads and storage requirements of new systems likely to be developed.

— The capacity-planning team adds contingency allowances, reviews the total demand during peak periods, reviews the demand for overnight work against the elapsed time available, and forecasts the capacity requirements.

Experience has shown that the forecasts provided by system development staff tend to be overestimated by about 20 per cent because of delays in implementing new systems. The company's original plan for the new site was based on the estimated capacity requirements up to 1990. Once the forecasts had been scaled down to account for the usual optimism of system development plans, it was realised that the new building should be able to meet the capacity requirements up to 1992.

A MAJOR CLEARING BANK

This case history concerns a major clearing bank, and illustrates the lengths to which some organisations are prepared to go to provide security for their data centre. The bank operates 24 data centres of varying sizes and varying levels of security. In 1976, as part of its computer development programme, the bank decided to construct a large, ultra-secure, central data centre facility. A greenfield site was chosen, because the bank's management was aware that it would be exceedingly difficult to provide the level of security required at an existing site — really good security has to be built into the site and buildings right from the beginning.

The design life of the facility is 30 years, from 1980 to 2010. The total complex provides 18,600 square metres of space, of which 2,200 square metres was originally allocated for housing computer equipment. Total investment in the project was about $50 million.

The complex has been designed literally as a fortress. It is situated on the top of a hill, away from urban development, and is approached by a single access road. The two main computer halls, each of 930 square metres, are buried underground. They are in a concrete inner shell within the main building, whose outer wall is of two-foot-thick concrete. Below the halls are the environment-control areas and above and around them the office areas. The two halls have a 30-metre gap between them and are serviced independently for power, refrigeration, water, and network facilities. The building itself is self-sufficient in water, fuel, oil, and power, and can continue to operate for three weeks if these services were cut off. The building is surrounded by a double perimeter fence 50 metres away from the building, with only one entrance beyond which no cars are allowed to pass. This entrance is guarded 24 hours a day by security staff, and delivery vehicles must also enter and leave through it.

Having entered the grounds, there is a second manned security control at the entrance to the building. There is a third security control room (at the entrance to the computer area) equipped with closed-circuit television screens, a computer controlled door, a foam-control panel for fire prevention, and two printers for monitoring information. Finally, in the main equipment hall, there is a control room where the status of cooling, heating, air conditioning, diesel generators, electrical services, oil services, water services, the fire prevention system, the lifts, and computer room water leakage can be monitored.

In the environment-control area there are five generators, a fuel-oil store, an uninterrupted power supply room, a transformer room, a high voltage switch room, a medium voltage switch room, a refrigeration plant room, and a fresh air vent plant room.

If all these precautions should fail, all the key work can be processed at a remote backup site. This site has sufficient processors, controllers, and disc storage to provide a minimum level of systems resilience.

The original design concepts have proved to be sound, because the facility has been able to accommodate the inevitable changes that occur over a period of time. For example the complex was originally planned to house development staff, but these staff have now been moved out into the user community and their space is being converted into additional computer areas. The conversion can be done at minimum cost because all the ducting for connections to the main computer halls was already in place.

It has also proved possible to accommodate additional equipment not foreseen when the facility was originally planned. Some departments originally installed their own computers, but these have now been moved into the data centre because:

— Users now recognise the need to integrate different systems.

— Users appreciate the need to make the optimum use of operating expertise.

— Users have become more aware of the need to provide resilient and secure data processing facilities.

Between 1981 and 1987, the processing capacity installed in one of the computer halls will have increased five-fold (from 10 mips to 50 mips), and the disc storage capacity will have increased from 21 gigabytes to 192 gigabytes. Overall, the floor space used in this computer hall has doubled from 465 square metres to 930 square metres.

The constant upgrading of the equipment installed in the computer halls requires a continuous planning process. To help it in planning the layout of the equipment, the bank uses a software package that charts the optimum layout, taking account of which equipment must be adjacent, and the service clearance areas required for each piece of equipment. The output from this package is produced on a plotter and takes the form of a map of the computer hall.

In summary, the bank has constructed an extremely secure data centre, and is successfully combining the inflexibility of building structures with the flexibility required by the evolving requirements of the business. The cost of providing this level of security and degree of flexibility is, of course, enormous, and very few organisations will be able to justify a data centre of this scale and sophistication.

A COMPUTER SERVICES ORGANISATION

This case history concerns a computer services organisation that still has more than half of the space in its data centre empty after five years. It illustrates the need for a flexible, modular approach. The organisation has been providing computing services for government administration for 35 years. In 1969, it moved to a new building, and by 1980 was using two computer rooms (one housing the processors, and the other housing the peripherals) totalling 560 square metres. At this stage, a new computer room of 1,250 square metres was built alongside the existing rooms. One difference compared with the original rooms is that the operators and the laser printers are located in a separate area of the computer room.

Services for the new room are provided from a separate building, known as the energy centre, which houses:

— Four diesel generators.

— Halon tanks for the fire-prevention system.

— Air conditioning equipment.

— Water supplies.

Five years after it was opened, only 40 per cent of the floor space in the new computer room is used. At first sight, it appears that the size of the room and its associated energy centre are the result of significant overestimates of the capacity that would be required. However, when the new room was planned, several possible circumstances were identified that could significantly increase the workload of the data centre. For example:

— New government legislation could require a large increase in processing and storage requirements.

— Another government data centre could be closed down and its workload transferred to this centre.

— The government might decide to offer commercial bureau services to third parties.

So far these developments have not materialised, and the centre is left with substantial unused space and has an over abundance of environmental facilities. If a modular construction approach had been adopted when building the new computer room, the spare space could have been used for other purposes until it was needed to house computer equipment.

Appendix 2

Four-step methodology for predicting equipment space requirements

Plans for a new data centre should be based on a range of possible space requirements that are then analysed to assess the risks involved in not providing enough space to meet all of the possible resource requirements. Butler Cox has devised a four-step methodology for generating the range of space requirements and carrying out the risk analysis. This appendix sets out the activities to be carried out in each of the steps.

STEP 1: EXTRAPOLATE RECENT TRENDS INTO THE FUTURE

The first step is to examine the trends in processing power, data storage, telecommunications, and other equipment over the past three to five years, and to extrapolate them into the future (say for the next ten years). This calculation provides the broad general growth pattern that would occur (the 'steady-state' projection) if no significant changes occurred in the requirements for information systems.

STEP 2: MODIFY THE PROJECTION TO TAKE ACCOUNT OF POSSIBLE DEVELOPMENTS

The next step is to modify the steady-state projection to take account of possible developments that could occur to impact significantly the use of information technology. We have identified five areas that should be considered (business development plans, your industry's use of information technology, your own information technology strategy, the workload created by office automation, and the growth in current applications.

Considerable judgement will be required in estimating the impact of each of these areas, and it will usually be necessary to develop optimistic and pessimistic estimates, with varying levels of certainty attached to each. Considering each of these areas minimises the chances of overlooking a requirement that could significantly change the amount of equipment required in the data centre.

REVIEW BUSINESS DEVELOPMENT PLANS

The organisation's business development plans should be reviewed to ascertain if there are any significant changes in the pipeline that would increase the demand for information systems. In particular, an attempt should be made to evaluate the impact of 'unforeseen' changes (new business ventures, takeovers, closing of product lines, and so on). This will be difficult (if not impossible) to do, but it is unforeseen changes in the basic structure of the organisation that cause the largest inaccuracies in predicting future growth of data centre facilities. As much information as possible should therefore be extracted from every possible source in an attempt to identify as many of the variables as possible. The more that top management can be involved in this exercise, the more likely it is that the possible variables will be identified.

REVIEW YOUR INDUSTRY'S USE OF INFORMATION TECHNOLOGY

In many cases, the use of information technology by other organisations in your industry can provide useful pointers to possible developments in your organisation. You should assess the potential effect on the data centre if your organisation were to follow suit.

REVIEW YOUR OWN INFORMATION TECHNOLOGY STRATEGY

In an ideal world, planning the future data centre would follow on as a natural consequence of the overall information technology strategy. Regrettably, it is often necessary to plan the data centre before the overall strategy is defined, or before it is updated to take account of impending changes. Nevertheless, it is crucial to review the organisation's long-term plans (five to ten years) for using information technology, and to set out the assumptions on which the data centre plans are based.

System development staff are often asked to estimate the future operational workload for systems yet to be developed. Some capacity planners have a natural tendency to scale down these estimates to account for the usual optimism of system development plans. Their experience tells them that new systems often do not become operational until sometime after the date originally predicted by development staff. However, any significant improvements in system development productivity, through the use of modern system building tools, for example, could mean that development timescales were improved and that new systems were delivered earlier than the capacity planners had assumed.

ESTIMATE THE WORKLOAD CREATED BY OFFICE AUTOMATION

In some organisations, the systems department is not responsible for office automation plans. Wherever possible, these plans should be studied to determine the potential additional workload at the data centre caused by the need to interconnect various office automation facilities with each other and with mainstream computing facilities. Even where the systems department is responsible for office automation, it is easy to underestimate the potential workload that will be created by office automation. Figure 3.4 (on page 12) showed that the types of data more likely to be associated with office automation (text, facsimile, video, and so forth) require proportionally much greater amounts of storage than conventional data or structured text.

In addition, most office automation facilities have, so far, been 'standalone' within a department. In the future, part of the data centre's role will be to act as a control centre or interconnection point for linking the office automation applications of a variety of departments, and for linking office automation systems with data processing systems. The implications of this type of development are explored in Neil Farmer's Position Paper (Hybrid Information Systems), which was published by the Foundation early in 1987.

ESTIMATE THE GROWTH IN CURRENT APPLICATIONS

For many data centres, the growth in processing power, data storage, and telecommunications traffic will be determined largely by the growth in the current applications base. Projections of the growth of existing applications should be sought both from the systems staff responsible for the applications and from user managers. It can be instructive to compare previous growth forecasts with the actual growth, and to attempt to analyse the reasons for any significant variances. By doing this it may be possible to isolate factors specific to your organisation that might otherwise be ignored.

STEP 3: TRANSLATE THE MODIFIED PROJECTIONS INTO SPACE REQUIREMENTS

At the end of the second step, it will be possible to make an informed estimate of the ranges of computing resources that will be required in the data centre over the next seven to ten years. Before these resource requirements can be translated into space requirements, it will be necessary to make some assumptions about the rate of technological progress, and the effect it will have on floor-space requirements for various types of equipment. (We provided some general indications in Chapter 2.)

Having made the technological assumptions, the resource requirements can then be translated into equipment requirements, and hence into an estimate of the space required to house the equipment. It is important to remember to allow space for commissioning new equipment before it is permanently installed in the data centre. During the commissioning period, the data centre will have to house both the new equipment and the equipment it will replace.

At this stage of the planning process, many organisations add in an almost arbitrary contingency allowance. A common rule-of-thumb seems to be to open a new data centre with the equipment occupying between one-third and one-half of the available space. None of the organisations we spoke to had a rationale for doing this, however. Nor could their planning process distinguish between the space reserved for expansion, and that allowed for contingency.

We believe that a better approach is to prepare a range of space estimates, beginning with the space required for the certain equipment requirements. This will provide the minimum equipment space required in the data centre. Further estimates should be prepared based on varying assumptions about the possible developments identified in Step 2. An estimate should also be made of the likelihood of these higher amounts of space being required.

STEP 4: ASSESS THE RISKS OF NOT PROVIDING THE MAXIMUM SPACE

The data centre planner should then use the output from Step 3 as the input to a risk-analysis exercise that will provide the organisation's business managers with information about the risks they will take by not building in 'spare' space at the beginning. For example, the planner might say "if the information centre grows at 75 per cent a year, and not at 30 per cent as assumed, we will have to expand the disc area of the data centre in 1991 at a likely cost of $750,000; however the cost of building in the additional space today is only $350,000." It is then up to business management to decide whether to take the risk of spending $350,000 now to provide space that may not be required. In presenting this type of argument, it is always worthwhile pointing out that the construction costs of providing contingency space is usually far lower than adding that space at a later date.

Sometimes, building a data centre to provide the maximum possible space that may be required can result initially in only a small percentage of the space being used. Several organisations have been prepared to accept the considerable additional costs that the provision of such a large amount of contingency entails (see, for example, the computer services organisation case history on page 27). Sometimes, the additional (and initially unused) space is allocated temporarily for other purposes until it is needed for computing or telecommunications equipment. One organisation, for example, was using the extra space to provide third parties with a 'hot' standby facility on a commercial basis. In this way, the cost of providing the extra space was offset by revenue from the third parties.

Butler Cox & Partners Limited
Butler Cox House, 12 Bloomsbury Square,
London WC1A 2LL, England
☎ (01) 831 0101, Telex 8813717 BUTCOX G
Fax (01) 831 6250

Benelux
Butler Cox BV
Burg Hogguerstraat 791
1064 EB Amsterdam
☎ (020) 139955, Telex 12289 BUCOX NL
Fax (020) 131157

France
Butler Cox SARL
Tour Akzo, 164 Rue Ambroise Croizat,
93204 St Denis-Cedex 1, France
☎ (161) 48.20.61.64, Fax (161) 48.20.72.58

Germany (FR)
Butler Cox Deutschland Ltd.
Richard-Wagner-Str. 13
8000 München 2
☎ (089) 5 23 40 01, Fax (089) 5 23 35 15

United States of America
Butler Cox Inc.
150 East 58th Street, New York, NY 10155, USA
☎ (212) 486 1760 Fax (212) 319 6368

Australia
Mr J Cooper
Consultants (Computer and Financial) plc Australia
Level 5, 303 Pitt Street, Sydney 2000, Australia
☎ (02) 2870400, Fax (02) 2870450

Italy
SISDO
20123 Milano – Via Caradosso 7 – Italy
☎ (02) 498 4651, Telex 350309 SISBDA I

The Nordic Region
Statskonsult AB
Stortorget 9, S-21122 Malmo, Sweden
☎ (040) 1030 40, Telex 12754 SINTABS

Spain
Mr Sidney M Perera
Rosalía de Castro, 84-2°D, 20835 Madrid, Spain
☎ (91) 723 0995