How To Plan A Data Centre

A modern datacentre is vastly different to those of a decade ago that housed huge mainframes. This means that the planning process is different too, though many of the old concerns still apply.

By David Norfolk

A data centre used to be a big building that you put a mainframe computer into, dedicated to servicing the needs of one or more organisations (depending on whether you were running it for your own company or as a bureau).

With the rise of client-server computing in the 1990s, all this changed. Computers became more efficient and no longer needed water-cooling systems and specially strengthened floors. Disk drives which used to be the size of refrigerators now fitted into the palm of a hand. What used to be the data centre moved (if you were lucky) into a locked general-purpose room somewhere in your prestigious HQ or (if you were unlucky) under somebody's desk.

Renaissance

A lot of computing certainly got done in the 1990s, but nonetheless it wasn't a total success story. People noticed that the service levels went down as the data centres closed and that, despite an impressive initial cost saving, a lot of people seemed to be spending a lot of time fiddling with PCs rather than getting on with business work. People started worrying about Total Cost of Ownership rather than just the up-front cost of buying computers, and business leaders started to worry about the security and business continuity threats associated with the reliance of business on IT. In fact, some of the biggest insurance firms and other major players found that they really couldn't afford to close down their data centres and replace their mainframe computers. So the mainframe, despite all the predictions, did not die and we have experienced a renaissance of the data centre, although it now also accommodates Intel-based server farms run with mainframe-like discipline.

According to Richard Dracott of Intel Online Services (**http://www.intelonlineservices.com**) there are two main drivers for the return of the data centre. Firstly economies of scale (lower power, cooling, cost per square foot, support costs). Secondly, far greater control over the processing environment (security controls, redundancy of network, power and cooling services, and closeness to support staff).

The new data centres are somewhat different to the old ones, but they are data centres nonetheless. Dracott associates the key differences with technology improvements and the availability of server farm designs based on commodity Wintel servers. Today's data centres are more robust, in terms of power, cooling and security. They offer a much lower cost per square foot. They are able to take advantage of much better tools and technology. They are more likely to be server farms based on Unix or even Wintel technology than mainframe processing centres (although mainframes are still in use and innovations such as the "virtual Linux server" promise to extend their life). They are able to scale out (adding more servers) as well as up (buying bigger servers).

Modern data centres also have to address a new class of security issues, says Dracott, associated with universal access to systems via the Internet. However, they still share the key characteristics of the old data centres. They are staffed 24 hours a day, seven days a week, and they operate to the same basic premise of maximum availability through managed redundancy

New Mainframes

Some commentators see the whole data centre as the new mainframe, with the servers inside it corresponding to the nodular processors plugged into the old

machines. However there is no one model for the modern data centre and some, such as the one at Legal and General's head office in the south of England, have evolved from older mainframe-oriented data centres through the addition of Windows NT eCommerce systems.

The managers of such data centres seem concerned to maintain the old mainframe-style attitude to service levels and management for the benefit organisation as a whole as new technologies come in. However, looking at some brand new data centres recently, it seems to me that little has changed. According to Mark Davies, UK MD of Integra (a pan-European developer and operator of complex e-business services), its new data centre is providing a modern version of a computer bureau, for clients that don't see running a data centre as a core part of their business.

Integra emphasises its sense of partnership with the customers. Key to its operation is the agreed service level - the client specifies requirements, ITC staff make it happen. They do provide mechanisms for client access to systems but in a separate room and via a keyboard / video / monitor switch to a system console - clients don't have access to the general facility without being accompanied by ITC staff (in fact, there has been very little demand for hands-on access by clients to their systems).

Behind data centre service levels is the concept of the data centre as providing a better managed environment than a firm, or corporate department, could afford to provide for itself. One in which the threats facing a computer installation are kept under control by skilled professionals.

Take Advice

Integra took on professional advice (from Hoare Lea & Partners - **http://www.hoare-lea.com**) for the detailed design to ensure that the new data centre was built to accepted "good practice" standards. Such consultants are expected to be aware of threat levels in different geographical areas and the level of investment appropriate to counter them. It is inappropriate to attempt to eliminate all threats. A more realistic ambition is to mange threats appropriately and cost-effectively.

Expert advice based on practical and historical knowledge of actual contingencies in a given geographical location is essential to building an effective data centre. An accurate threat analysis from the American mid-west is unlikely to translate well to European conditions, for example, although some threats will be common. Natural phenomena like geological fault lines are country-specific, though, and some areas have much more reliable power supplies than others.

When you start to plan a data centre, the high level design can be simple because the detailed formal design depends to a great extent on re-use of a common blueprint - the shape of a new data centre may be different but the equipment rack designs and so on will be the same. Intel also emphasises this approach when talking about

"Reproducing standard designs not only keeps planning and inventory cost down, it increases flexibility and resilience."

its own data centres. Intel treats the design of a data centre with the same meticulous attention to detail as it would give to a chip fabrication plant. In a fab, even the chemicals given off by different paints can affect the product and, while you don't need quite such attention to detail for a data centre, this background stands it in good stead.

Flexibility

Reproducing standard designs not only keeps planning and inventory cost down, it increases flexibility and resilience. If the worst happens and you need to move processing for a given client to another data centre, knowledge that it uses the same basic design and even the same kit and software reduces the possibilities for things to go wrong. Re-use of standard designs also simplifies equipment choice. Obviously, flexibility to accommodate the needs of as many clients as possible is important, but you also want to deploy industry-standard kit and maintain good relationships with a manageable number of suppliers. Consider access to the servers, as a separate console and monitor for each server is unmanageable and you will need a switched arrangement, possibly with a set of floor-to-ceiling screens displaying background information (news, global Internet performance) as well as the console for a particular server.

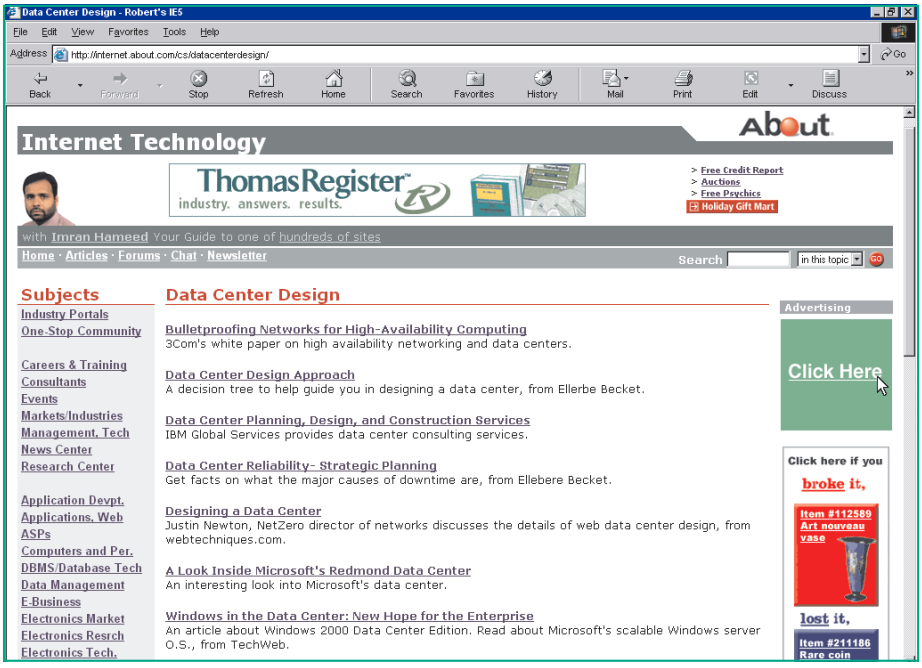
The detailed design of Integra's data centre includes "flood ducting" of cold air flow through the floor and standard rack designs, so that equipment can be moved or added without causing problems with overheating. Integra specialises in Wintel servers but didn't want to install full capacity until it had customers to use it. Prediction of future demand is always tricky but even if radically different solutions become fashionable (an IBM zSeries mainframe running thousands of Linux virtual Web servers is an option that some firms are adopting for Web hosting) they are likely to be generally compatible with Compaq server dimensions these days.

Nevertheless, not every data centre can be built to an absolutely identical plan in exactly the same environment. Location is a vital influence on data centre design, and operation and the shape of the plot available can influence the use of standard designs. However, since a data centre is largely rows and rows of electrical equipment in racks, a simple rectangular building or room is usually all that's needed, although access must be considered. You need to move the kit in and must provide access for fire engines and so on, but at the same time, some restrictions or, at least, strong perimeter fencing, will limit the potential of "ram-raid" style attacks.

Location

When choosing the location for its data centre, Integra balanced access to electrical

"Effective security involves more than configuring a firewall properly. A data centre must have a security policy, based on a threat analysis, as a basis for configuring its firewall and ensuring that firewall security can't be bypassed."



supplies and several sources of high-bandwidth fibre with site cost and the attractiveness of the site for both existing and new staff. Obtaining a suitable calibre of staff shouldn't be underestimated as a factor in siting a data centre.

Armoured fibre-optic cables are fairly resilient but you still need two or even three fibre cables and power supplies coming into the data centre, preferably from different directions and different suppliers.

Another aspect of location is the quality or otherwise of local emergency services. In any case, the requirements and design of a data centre should be discussed with the local police and fire services and the design or implementation modified as a result, if necessary. Insurance should be arranged at the same time.

Once the physical installation and environment of a data centre is under control, security becomes a major concern. One of the key drivers for the return of data centres is the greater control of the processing environment they offer and the corresponding comfort factor they offer with respect to security. According to Intel, its data centre customers tend to consider firewall design as the major item of logical security and value a data centre's ability to keep the firewall design and configuration current against known threats.

But effective security involves more than configuring a firewall properly. A data centre must have a security policy, based on a threat analysis, as a basis for configuring its firewall and ensuring that firewall security can't be bypassed. Standards like ISO 17799 provide a benchmark against which a security policy can be assessed but you can't just pick a policy off the shelf - it must be based on a thorough analysis of likely threats and will provide an underlying basis (whether explicitly referred to or not) for service level agreements with data centre clients.

Integra, for example, considers its security policy to be a vital part of its confidential service level agreements with its clients although it will make a generalised overview available to prospective clients.

A typical client-focused security policy (a data centre will have internal security policies too) will cover physical access control, data security, logical access control and firewall provision. It will enable clients to assess the degree of security provided for their business data and the resilience of their processing environment.

Questions

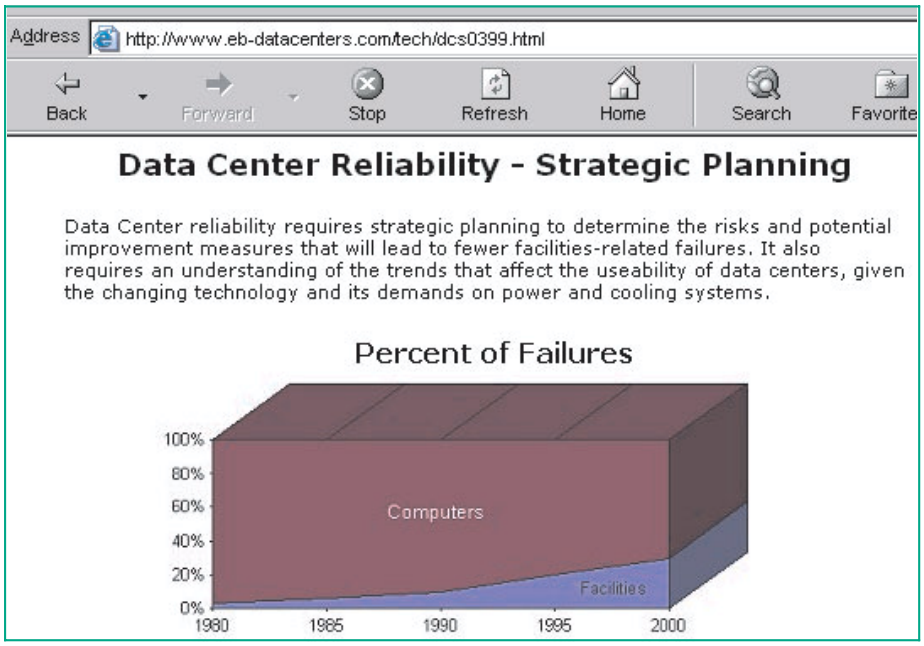
A practical policy document shown to data centre clients might include points such as: who is allowed into sensitive parts of the building? What are those sensitive parts? Are clients allowed in and on what terms? Are all visitors escorted at all times? Are all operations staff allowed access to all of the building? Is a badge scheme in operation? Are audit records of access kept? What sort of lock technology is used and who has access to the keys? Is a zoned alarmed system installed and who manages it? What callout/monitoring provisions are implemented? Is equipment in a temperature-controlled environment behind locked doors and secure fencing at all times? Is encryption available for data travelling outside the data centre? What encryption technology is available? How is console access to systems managed? How are system administrators controlled and are "super admin" accounts forbidden? What audit trails are maintained for when staff access client data? How are passwords managed (as regards password expiry, history and checks for the use of real words)? What firewall technology is deployed and how is it managed?

"You need to structure the testing so that you are confident that all the elements of your contingency provisions work reliably before you test whole system failover."

It is particularly important that the human aspects of security aren't overlooked. Simply not advertising a data centre with a large logo outside the gate can reduce risk, and a company will often arrange to hire security guards from external specialists, in which case such things as vetting of CVs and informal testing of practical security should be written into the contract.

Testing

A vital part of giving customers confidence in data centre service levels is testing of the facilities in advance. Even reusing a general plan doesn't guarantee success and Integra, for example, found during a simulated emergency that they couldn't get at the cylinders holding gas for fire control in order to refill them. Not a very serious issue, perhaps, as two fires breaking out within a short period isn't very likely, but



not something you wish to find out in the heat of an actual incident. If you find out about the problem in advance it can be fixed reasonably cheaply and easily (in this case, by adding a walkway).

However, testing is itself a threat to service levels. Apocryphal tales abound of CEOs testing their new data centre's resilience by flipping the main power breaker. This is a very foolish test - it doesn't tell you much if the backup systems cope (except that maybe the main power breaker should be protected in some way) but the consequences of the test failing could be dire.

Tests must be designed to include contingency plans to minimise the consequences of failure and this may limit the testing of some of the extreme threats. Essentially, you need to structure the testing so that you are confident that all the elements of your contingency provisions work reliably before you test whole system failover.

There is an issue as to whether you inform customers about contingency testing. Ideally, carefully phased testing means that there is minimal risk of impacting service levels. You should publish a method statement regarding recovery after testing, detailing your contingency plans, maximum allowable service interruption and so on. If you're within this, there is no reason to specifically warn customers of testing and doing so might impact their confidence. However, if the consequences of a failed test might be significant then customers should be warned and the data centre should assist with customers' contingency planning.

"A data centre looks like an example of putting all your eggs in one basket, but it is more robust than a collection of servers scattered throughout an organisation."

Health And Safety

One aspect of data centre implementation that might be overlooked is health and safety. These are major issues today, as although most data centres will aim at automated "lights out" operation, there are always some people around. Fire protection systems, for example, can flood a room with inert gas in seconds - what if there are people in the room at the time?

It is essential that a company's routine health and safety procedures are extended for the special circumstances of the data centre. It is easier and more efficient, as well as far safer, to design and maintain a safe environment continuously rather than to scurry around tidying up before an inspection. There is also a business benefit from adherence to good practice. Poor practice sends messages to the staff about their importance in the scheme of things and can seriously reduce morale but, in any case, accidents are an avoidable risk to service levels. Poor health and safety practice also has security implications because if procedures aren't defined, starting a small fire or some such may result in sufficient chaos for security defences to be penetrated without this being noticed.

Safety requirements will sometimes mandate simple good practices - such as a requirement that no one is allowed to work in the building on their own. Sometimes procedures will be more complex, such as those for fire control. A gas mixture that extinguishes flame can still contain enough oxygen to be breathable, although the sudden expansion of gas into a room can cause damage and you should have procedures governing the switching off of such automated systems when people are in a room (and for ensuring that they're switched on again when it is empty). I personally know of one case where a fire protection system test resulted in all the data centre's windows being blown out by the expanding gas.

Monitor

Once a data centre is built it must be monitored for performance. Since people are using a data centre, in part, because it increases their confidence in their IT operations, the critical metrics for a data centre's overall success are: the number of service level interruptions (complete interruptions are relatively easy to measure and control but impaired service is relative to the particular SLA in effect); the number of physical security incidents (also relatively easy to measure and control); and the number of logical security incidents (much more difficult to control and successful incidents may be hard to detect).

As a business, a data centre should typically be monitoring: capacity (space) utilization and efficiency (storage density relative to power used; the speed of new solution deployment; performance relative to agreed service levels; new customer acquisition; existing customer retention; return on investment; profitability. The critical success factors include: reuse of standard infrastructure and building designs; pro-active collection of service-oriented metrics that can give clients confidence in the service levels achieved; adherence to external standards where appropriate; and benchmarking of service levels against industry norms where possible; effective contingency plans including damage control and the management of press coverage after any incidents.

Customers may well latch onto very specific indicators for the professionalism of an installation, such as firewall intrusion filtering and virus control, and it is important that firewall configurations and anti virus packages are not only kept up to date but seen to be up to date.

Conclusion

Data centres have a strong future because they allow an organisation to manage the service levels seen by organisational users of the data centre effectively. This applies equally to data centres providing, essentially, a modern computer bureau or ASP service to clients that don't have the resources to manage IT for themselves, or don't see this as a core competence, or data centres that consolidate an in-house IT service into one place.

A data centre looks like an example of putting all your eggs in one basket, but it is more robust than a collection of servers scattered throughout an organisation and it is actually easer to plan for disaster recovery if you know where everything is.

"Data centres have a strong future because they allow an organisation to manage the service levels seen by organisational users of the data centre effectively."

PCNA

New Reviews from Tech Support Alert

Anti-Trojan Software Reviews

A detailed review of six of the best anti trojan software programs. Two products were impressive with a clear gap between these and other contenders in their ability to detect and remove dangerous modern trojans.

Inkjet Printer Cartridge Suppliers

Everyone gets inundated by hundreds of ads for inkjet printer cartridges, all claiming to be the cheapest or best. But which vendor do you believe? Our editors decided to put them to the test by anonymously buying printer cartridges and testing them in our office inkjet printers. Many suppliers disappointed but we came up with several web sites that offer good quality cheap inkjet cartridges with impressive customer service.

Windows Backup Software

In this review we looked at 18 different backup software products for home or SOHO use. In the end we could only recommend six though only two were good enough to get our “Editor’s Choice” award

The 46 Best Freeware Programs

There are many free utilities that perform as well or better than expensive commercial products. Our Editor Ian Richards picks out his selection of the very best freeware programs and he comes up with some real gems.