



Capacity Planning

Enhancing resource allocation to support future technology and drive the business

A BMC Software whitepaper

www.datamonitor.com

Datamonitor USA

245 Fifth Avenue
4th Floor
New York, NY 10016
USA

t: +1 212 686 7400
f: +1 212 686 2626
e: usinfo@datamonitor.com

Datamonitor Europe

Charles House
108-110 Finchley Road
London NW3 5JJ
United Kingdom

t: +44 20 7675 7000
f: +44 20 7675 7500
e: eurinfo@datamonitor.com

Datamonitor Germany

Kastor & Pollux
Platz der Einheit 1
60327 Frankfurt
Deutschland

t: +49 69 9750 3119
f: +49 69 9750 3320
e: deinfo@datamonitor.com

Datamonitor Asia Pacific

Darling Park
Tower 2, Level 21
201 Sussex Street
Sydney NSW 2000
Australia

t: +61 2 9006 1526
f: +61 2 9006 1559
e: apinfo@datamonitor.com

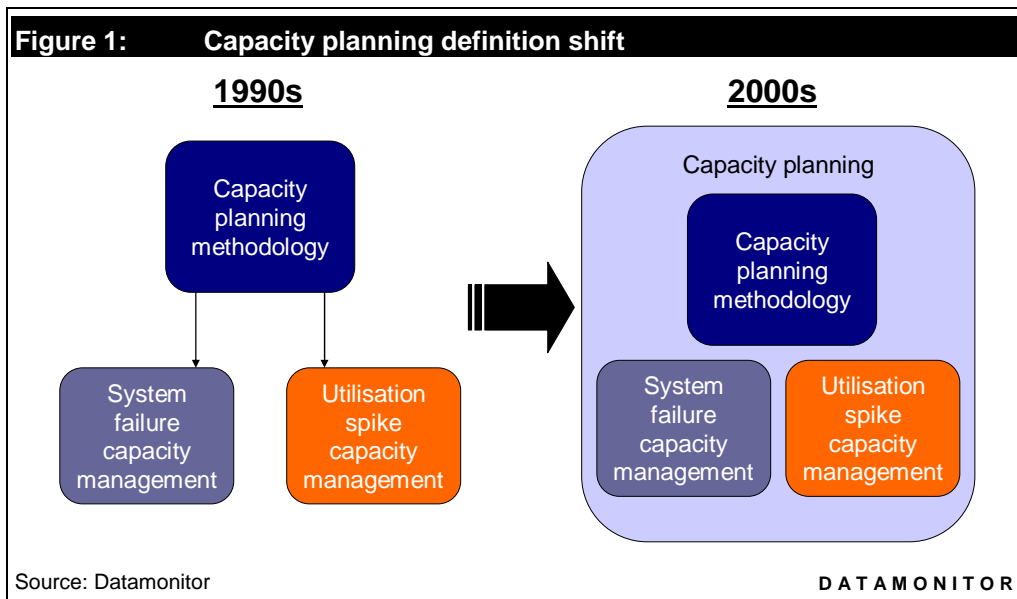
Datamonitor Japan

Wakamatsu Bldg 7F
3-3-6 Nihonbashi-Honcho
Chuo-ku
Tokyo 103-0023
Japan

t: +813 6202 7681
f: +813 5778 7537
e: jpinfo@datamonitor.com

Foreword

Taken at its broadest level, capacity planning can incorporate a range of different elements, including service, business and resource delivery. The IT industry Best Practice ITIL¹, has perhaps the widest-reaching of the commonly accepted definitions of this space. However, this whitepaper deals solely with the resource capacity planning element. The concept of capacity planning as a methodology is to insure against operational failure and excess utilisation is by no means a new phenomenon. Businesses have already been implementing strategies to ensure that the IT department is able to fully support the organisation's operational roles. With increasing pressures being placed on IT managers to minimise their cost bases, capacity planning is undergoing a transitional evolution to what is now becoming known as capacity management. While the term capacity planning, in the methodological sense will not disappear, IT managers are looking to take more immediate action and as such are turning to capacity management solutions. For the purposes of this whitepaper, *capacity planning* refers to this next stage in the evolution of capacity technology: capacity management and the dynamic allocation of resources for both system failure and utilisation spikes (as seen in Figure 1).



¹ The IT Infrastructure Library – The world's most widely accepted approach to IT service management

TABLE OF CONTENTS

Foreword	2
Introduction	4
Current IT challenges driving the need for capacity planning	4
Current state of the capacity planning market	6
Examples of major outages and their effects	8
Benefits and drawbacks of capacity planning	9
Benefits of capacity planning	9
Potential drawbacks	11
How to overcome the drawbacks and win buy-in from your organisation	11
Using capacity planning in the real world – Egg PLC	12
Capacity planning offers long-term (LT) benefits as well as short-term (ST)	13
Higher service rates, lower costs and greater insight into resource allocation can be readily achieved (ST)	14
The impact of miscommunication can be minimised (ST)	15
The cost of not considering capacity planning (ST-MT)	15
Supporting the future with capacity planning (LT)	15
Definitions	16
Related readings	16

Introduction

IT is increasingly underpinning multiple elements of businesses around the world. Consequently, IT managers are bearing the accountability for providing predefined services to business units that in turn are responsible for the company's success. Additionally, trying to minimise operational and procurement costs to achieve these goals is a challenging task and, as a result, IT departments are looking to optimise existing infrastructure and ensure maximum usage is leveraged.

Coping with spikes in the utilisation of an application or system needs to be dealt with quickly and effectively to ensure that speed of service is maintained and the business can operate as per usual. Similarly, system failover² plans remain a critical aspect of an IT environment as server downtime can create catastrophic effects for the business as a whole. Both internally and externally facing applications need to be able to function should a primary server go down. Faced with these two dynamics of maximising service and minimising costs, IT managers are increasingly looking to capacity planning solutions.

IT managers have typically provisioned 100% redundancy for system failure, and significantly more for unpredictable application utilisation spikes. Given the nature of usage spikes and the reality that response time and transaction throughput are not directly related to resource utilization rates, using average utilisation to provide backup resources is not sufficient. Instead, companies are forced to look at providing reserve resources based on the peak utilisation rates of applications – typically up to five times the average. Dynamically allocating resources to application needs from a combined reserve resource pool, capacity planning allows IT departments to run a lower number of reserve servers, translating to a significantly lower cost base, and increased level of service, helping to drive the business further.

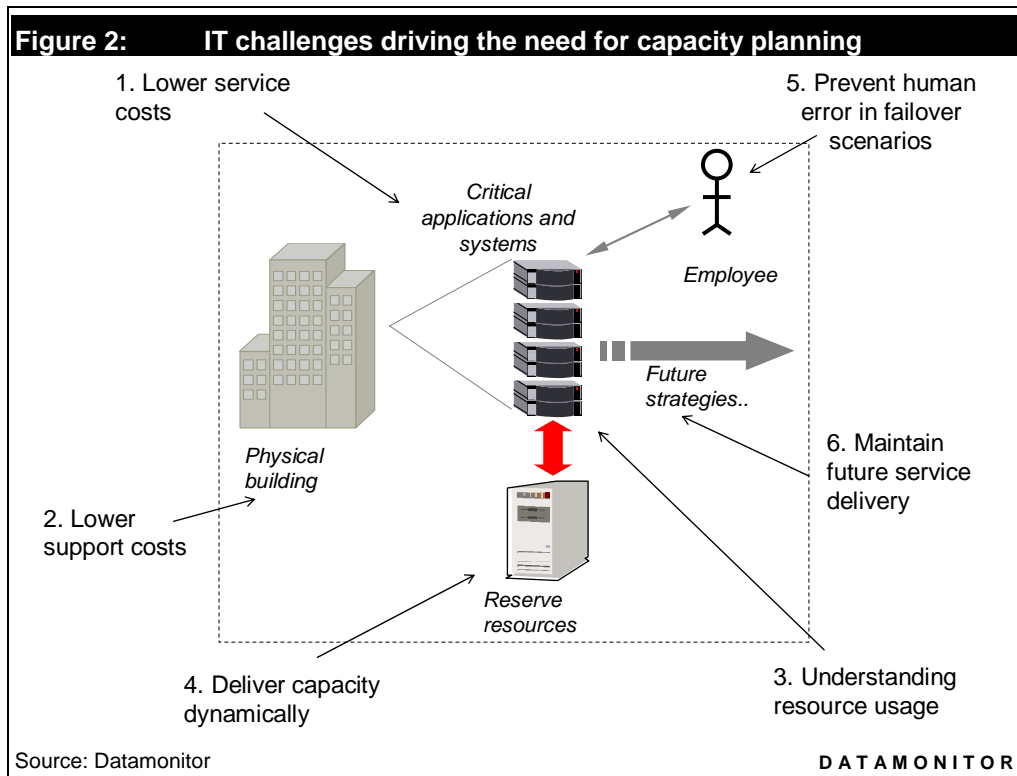
Current IT challenges driving the need for capacity planning

The need for an effective resource provision strategy is critical, regardless of the IT environment. The general strategy employed by IT managers is to match reserve

² The transfer of operation from a failed system to a similar, redundant system to ensure uninterrupted availability.

Capacity Planning

resources to the application's peak needs, however this is generally an expensive exercise that leaves unused capacity.



Now, IT executives are facing increasing pressure from their business level counterparts to provide IT services at a reduced cost while maintaining predefined service levels – in real time. Detailed challenges driving the need for capacity planning solutions include:

- **Ensuring maximum service delivery** through advanced modelling and forecasting of the future impact on resources of changing business conditions or new application implementations;
- **Reducing service costs and lower operational risk** at optimised resource performance levels. Consolidation of resources to meet these lower cost targets has been a key strategy for IT managers, however they are struggling to balance this with the ability to adapt to resource requirement changes quickly and effectively;

Capacity Planning

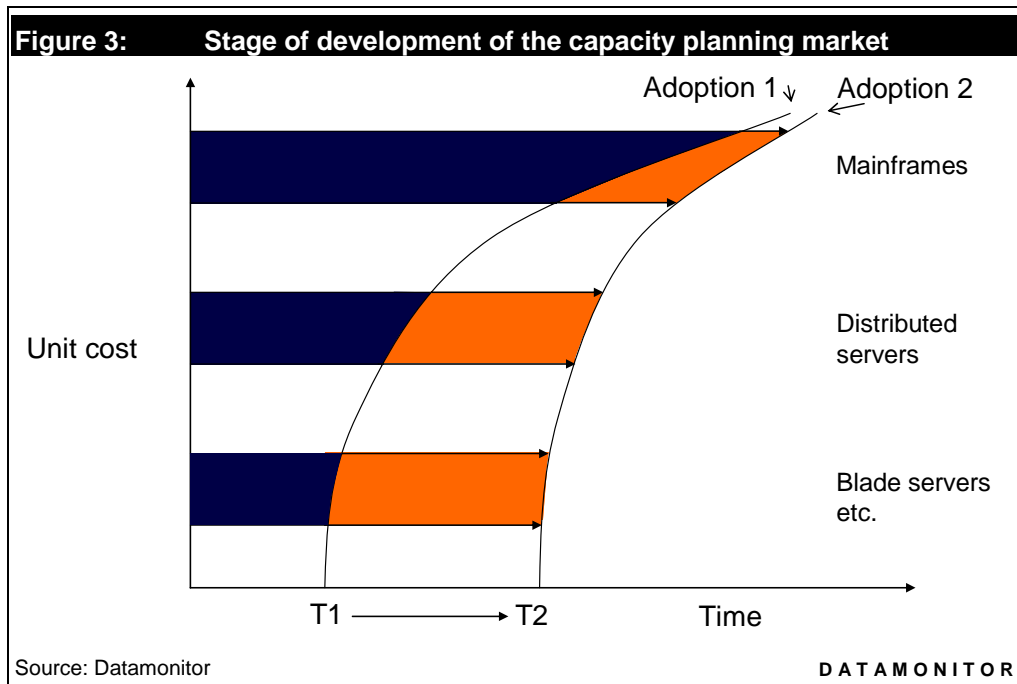
- **Lowering the 'support' costs of IT resources.** Electricity costs, being driven up by wider market pressures, floor space, cooling and management costs are all areas directly linked to the amount of IT infrastructure in place within an organisation;
- **Assessing and quantifying the actual resource requirements of applications** across the business and allocating those resources according to business priority using a policy based architecture;
- **Delivering resource capacity dynamically** in as short a period of time as possible;
- **Maintaining critical business services in failover scenarios** through automation and a reduced risk of down-time caused by human error;
- **Minimising the impact of poor communication** between business units and the IT department with regard to operational activities that may necessitate greater temporary resource allocation.

Current state of the capacity planning market

The popularity of capacity planning solutions has been steadily growing, however the technology is generally being implemented at different 'infrastructure levels'. Mainframe systems have been utilising capacity planning in multiple industries at a fairly widespread level. The inherent cost of provisioning additional mainframe capacity on a 100-500% reserve ratio has been far too great for many IT departments to realistically justify. Consequently, using pools of resources shared between multiple or virtualised and/or partitioned mainframe systems has been a popular strategy to ensure continued timely availability of systems in the event of an outage or spike in utilisation. This trend has continued to non-mainframe systems such as UNIX servers, as the benefits realised at the mainframe level are demanded by senior business management in greater quantities. While capacity planning is not yet fully mature at this level, its adoption is rapidly increasing. At the lowest level of servers, populated by blade and industry-standard rack mount servers and the like, there has not been widespread uptake of capacity planning, however Datamonitor believes this is about to change as IT departments are forced to squeeze out all remaining excess cost yet maintain the provision of service levels. In Figure 3, this is depicted by the shifting lines in adoption (Adoption 1 → Adoption 2) as the market moves forward. Adoption of capacity planning solutions will grow quickest in the lower cost server markets (e.g. blade and rack servers) while distributed servers (e.g. UNIX) will also

Capacity Planning

continue, but at a slower rate. The mainframe market will grow at the slowest rate of all three.



Businesses in different industries are using capacity planning to underpin different strategic initiatives. For cost sensitive industries such as retail, where the concept of capacity planning in various non-IT disciplines is well documented, companies are using capacity planning primarily to lower overall IT infrastructure costs. Business models within these companies have been developed to effectively push out any excess costs, and in this light, capacity planning is a key and relatively untapped element of this strategy.

At the other end of the spectrum are businesses looking to capacity planning to support cutting edge IT deployments with maximum uptime to ensure greater revenue opportunities. Institutions such as financial services companies are implementing strategies to further drive their businesses on the back of new technological innovations. The emergence and popularity of online banking is a testament to the 'pro-IT' outlook of financial services companies. To drive this innovation, banks, investment houses and other such companies need to ensure maximum uptime, yet keep overhead costs down. Similarly, the heavy marketing of new technological features combined with their often cyclical utilisation rates has seen capacity planning enable the business to cope with spikes in usage. By avoiding the excess

Capacity Planning

provisioning of reserve resources, the overall cost of technical innovation is further optimised.

Capacity planning is by no means restricted to these industries however. Governments, manufacturers, broadcasters, health institutions and educational facilities are all using capacity planning to maintain and improve service levels while at the same time lower cost bases.

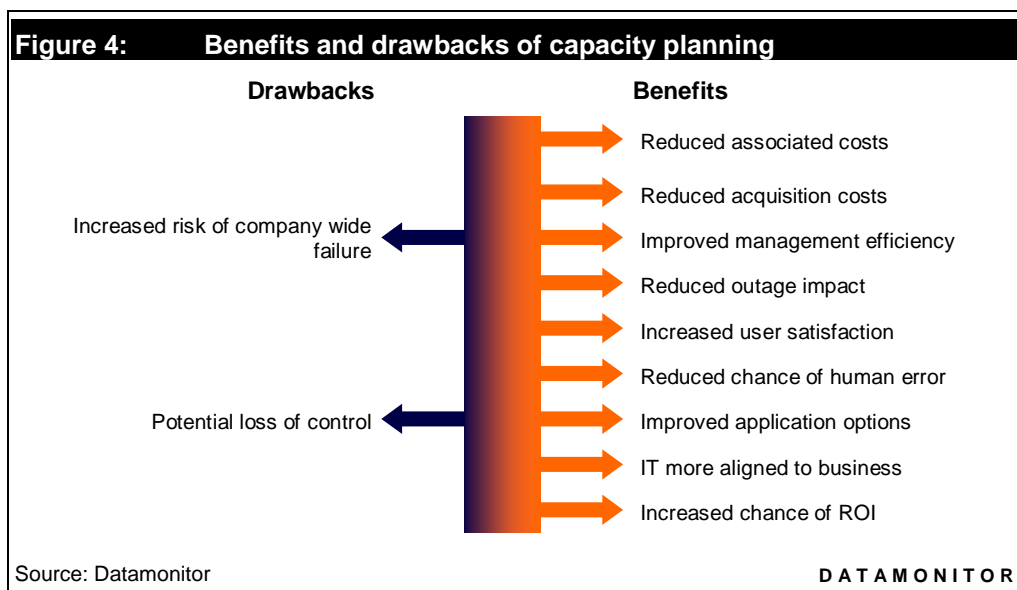
Examples of major outages and their effects

In December 2004, a **major UK retailer** advertised a promotion whereby customers purchasing items for Christmas through its website would receive a 20% discount on all goods creating a massive spike in usage. The site was completely inaccessible for two hours with reports of slow loading occurring for several hours before and after the outage. The problem stemmed from a lack of communication between the business unit responsible for the promotion and the IT department who, with no prior notice, were unable to provide additional resources fast enough. Consequently, customers chose to use competitor websites that were not suffering from similar problems at the time. This outage came at a critical time for the retailer as the lead up to Christmas is one of the most important periods for retail – thereby seeing the company miss out on significant revenues and associated profits. A capacity planning solution would be able to automatically allocate the resources necessary to the website and associated supporting back-end systems to ensure maximum uptime.

The **New York Stock Exchange (NYSE)** experienced an outage in June 2005 as one of its systems responsible for distributing market data and operating computer trading systems suffered from a fault. The system and its backup were swamped with error messages resulting in the exchange closing four minutes early. On an average day, the NYSE trades some 1.46bn shares valued at \$46.1bn. Four minutes in a 6.5 hour day equates to approximately \$470m – a major impact for one of the world's busiest exchanges. This outage directly affected the stock exchange's primary revenue stream leading to an unbudgeted shortfall in profits. The role of the IT department within the business as a whole shows that not using capacity planning can lead to direct business pains. Had a capacity planning solution been implemented, the system would have been able to reallocate sufficient resources from a pool of redundancy servers to deal with the overflow of error messages experienced by the system. Downtime would have been minimised, and the impact of the outage would be decreased.

Benefits and drawbacks of capacity planning

Like any technology, implementing a capacity planning solution brings with it a number of benefits and drawbacks. Ensuring minimal risk in implementation projects is critical to achieving the desired outcome from a new system. By balancing and being aware of the potential drawbacks of capacity planning solutions, IT managers will have a better chance of achieving buy-in and meeting return on investment (ROI) targets.



Benefits of capacity planning

There are numerous benefits to be experienced through the use of a capacity planning solution including those that apply directly to cost levels and those that impact business performance across the company. The cost cutting impact of capacity planning solutions is felt in the following areas:

- **Reduced infrastructure costs** including licensing, maintenance and support costs. By running fewer systems, businesses can reduce the outlay required for software as well as hardware. Support costs such as electricity, cooling and floor space are also reduced as less of each is required for a smaller, combined pool of redundancy systems;
- **Reduced cost of acquiring new systems** to maintain a previous initiative of providing one-to-one redundancy for applications. Many businesses will be able to stall the majority of short-term acquisitions as they reallocate and

Capacity Planning

optimise existing infrastructure using existing excess capacity instead. Medium and long term acquisitions will also decrease as the company purchases resources for backup at a lower rate than before. This is particularly the case when looking at capacity provided for utilisation spikes which is sometimes provisioned at a ratio of five times average resource usage;

- **Improved management efficiencies** are gained as monitoring and control of reserve resource systems becomes increasingly automated;
- **Reduced risk of over-utilisation impacts** such as brownouts and outages. As there is a lower chance of redundancy capacity being overdrawn, heavy loads applied to one application or system will not cause a failing device to create a negative effect on a power or network grid.

While cost management will be of key interest to those businesses looking to reduce IT investment, capacity solutions will also have multiple positive effects on the business as a whole, such as:

- Both internal, and perhaps more importantly, external users will **experience increased satisfaction** from interacting with capacity planning systems as downtime decreases, leading to greater operational efficiency and its resultant revenues;
- Part of a greater uptime will be due to a **reduction in human error** when allocating resources in the event of a spike in usage. By automating the process, allocation can occur based on predefined policies or rules that ensure no gaps in availability are experienced;
- Existing and future **applications will benefit from improved resource support** allowing them to run more efficiently and opening the door for larger suites that may have been too costly based on previous reserve resource provisioning requirements;
- **IT will become more aligned to business needs** as the department support more applications better to drive business growth, and cut overall costs. The company will become more agile as it tries to adapt to changing business and market conditions helping to directly contribute towards revenue and profit generation;

Capacity Planning

- An **increased return on investment (ROI)** will be delivered as the IT department provides a more efficient service at a lower cost level. Again, this is particularly the case for spikes in utilisation as reserve resource levels are drastically cut from 4-500% of average usage.

Potential drawbacks

The market for capacity planning has not taken off as quickly as some may have expected, as there have been some minor drawbacks to the solution.

These include:

- **Increased risk of a company wide systems failure** requiring more resources than a reduced systems pool can provide – a traditional rate of redundancy (often 100% for failure, 500% for utilisation spikes) doesn't have this problem when deployed dynamically, but is significantly more expensive. IT managers need to assess the chance of a major catastrophe such as a fire, flood or terrorist attack knocking out an entire facility;
- **A potential loss of control in allocating resources.** Depending on how the capacity planning solution is configured, this is similar to other automated solutions, however at a company level will likely have minimal impact and also reduces the chances of a human error cause problem.

How to overcome the drawbacks and win buy-in from your organisation

Capacity planning solutions meet all the challenges faced by the business discussed previously in this report. Achieving buy-in within the organisation will be dependant on the ability to demonstrate a workable balance between the benefits gained and the risks associated with such a strategy. The risk of an enterprise wide usage spike is unlikely given the architecture adopted by most modern companies. As such, a loss of control is the only other issue created by capacity planning. This should be stressed as a benefit rather than a problem as it removes an element of human error associated with manual provisioning of resources in high usage scenarios.

Greatest resonance with top-level management will be achieved through demonstrating a lower infrastructure cost profile achieved through better utilisation of existing overcapacity into the future, and lower acquisition costs due to decreased redundancy capacity requirements. ROI – commonly the benchmark with which

Capacity Planning

well managed and planned IT department will directly affect the performance of other business units, particularly those involved with customer facing applications, should generate significant interest in a capacity planning solution.

Capacity planning can be applied to two individual areas of reserve resource provisioning: utilisation spikes and system failure. Obviously, the provisioning of reserve resources for spikes is significantly more expensive than for system failure. Datamonitor believes that the evolution of the market with respects to capacity planning will be based upon the reduction of reserve resources for spikes in usage. This will provide the highest ROI as the amount of excess capacity is culled at a greater rate than for system failure. In seeking buy-in from your organisation it will be important to stress both of these.

Often, organisations have to overcome multiple variations on the challenges above at once. Egg PLC is a good example of how such hurdles can be successfully overcome.

Using capacity planning in the real world – Egg PLC

Egg PLC, a successful financial services company offering products such as current accounts, loans and credit cards, operates in a primarily online environment. The performance of its applications and the availability of resources are critical to its business model as it touts its 24 hour access. Outages for the company are an unacceptable situation as they can directly affect customer numbers and reputation. Due to having no branches, if the website suffers from an outage, business with customers will simply halt. Similarly, a slow running website will result in customer frustration and contribute to a poor reputation. The company had previously been running a freeware based application resource management system to manage 50 of its UNIX servers running behind its website. The relative immaturity of this solution meant that Egg was unable to retrieve any statistics about the site's usage past a simple binary status – up or down. With the bank experiencing strong customer growth, now measured in the millions, the availability of resources and the management thereof is becoming increasingly critical to the company's successful operation. A new method of monitoring and allocating resources was needed.

BMC Software was chosen as the vendor that offered solutions best suited to Egg's needs. Standardising on BMC Software's PATROL product portfolio, Egg started with an initial capacity planning pilot using 25 UNIX servers. On the back of significant benefits achieved, the company has now expanded the scope of the project to its entire server stack, comprised of 200 UNIX servers and 150 Microsoft Windows servers. Egg is now able to receive resource usage alerts that are fed through to the

Capacity Planning

company's central command centre where decisions from a predetermined list can be made. According to which piece of software or hardware the alert is generated by, the action taken is decided by the appropriate taskforce. To correct issues, staff can resolve the problem themselves, run preconfigured scripts or call a support team.

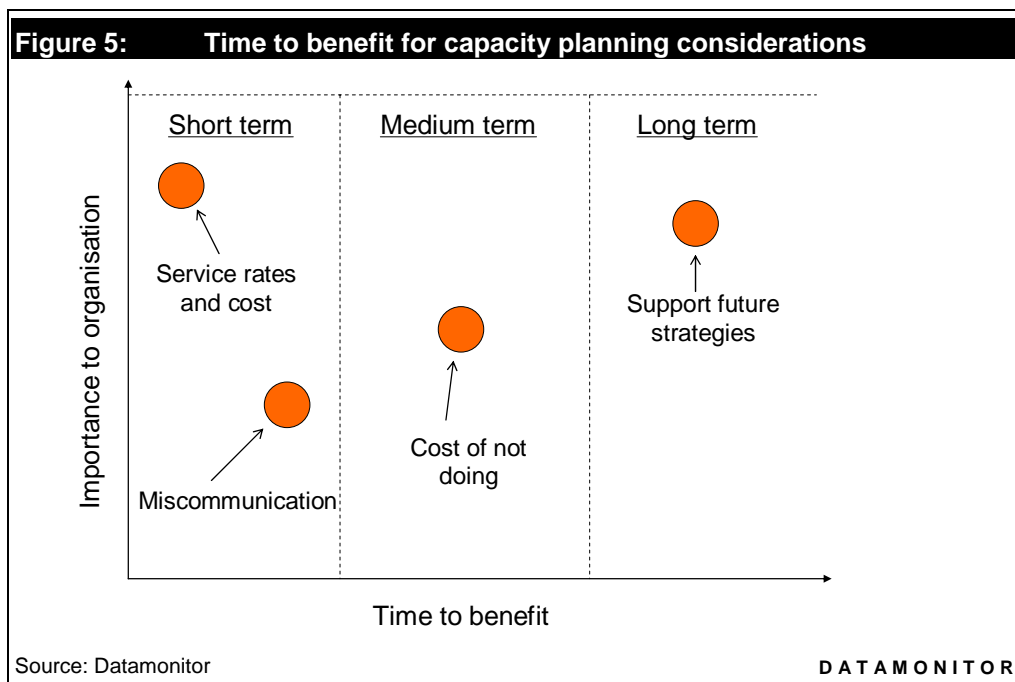
Greater analysis and reporting has also been facilitated by BMC Software's solution which allows staff to drill down to individual reports about machine status including information on operating system, system usage and application availability. Advanced performance analysis is now significantly easier to view, enabling Egg to identify trends in resource usage, and display key indicators in a number of different formats. This information further helps Egg to assess its resource requirements for the future, along with the impact that business decisions such as promotions will have. As a result of capacity planning, the availability of Egg's website has risen from 87% to 97% with reaction time to alerts in some instances cut by more than 10 minutes.

Capacity planning offers long-term (LT) benefits as well as short-term (ST)

Benefits and costs of capacity planning that will have an impact in both the short-term and the long-term include:

- Higher **service rates**, lower **costs** and greater insight into resource allocation can be readily achieved;
- Minimising the impact of **miscommunication**;
- An inherent **cost** of not considering capacity planning;
- **Supporting future** technology led strategies.

Capacity Planning



Higher service rates, lower costs and greater insight into resource allocation can be readily achieved (ST)

Capacity planning as a solution to optimise resources is already widely used in the mainframe server market. Its evolution down towards cheaper and more commoditised systems such as blade servers is inevitable as businesses try to squeeze out every last cost possible. Ensuring maximum uptime of business critical systems such as customer relationship management (CRM), enterprise resource planning (ERP) and websites can be the competitive differentiation of a company in its market. As IT managers are being asked to provide more services more effectively using lower levels of investment, capacity planning solutions provide the perfect answer. The ability to drill down on resource utilisation rates across the company right down to individual system levels will provide many companies with a far greater insight than they have ever experienced before. Being able to gain access to this data in real-time and acting on the information will help to streamline IT processes and ensure the maximum impact of future IT initiatives. Those companies that have opted to implement a capacity planning solution such as BMC Software's PATROL, have immediately felt the impact of their decision, and are now running a more optimised IT environment.

Capacity Planning

The impact of miscommunication can be minimised (ST)

Although communication among business units would seem common sense to many, within companies of all sizes there are often serious lapses. Given the fact that business is being conducted at a much faster rate than ever before, employees are sometimes making business critical decisions without the key input of their peers. Traditionally seen across business units (such as the case of the UK retailer mentioned previously) such as marketing and IT, poor communication – particularly in scenarios with limited resources – can cripple a business. Using a capacity planning solution, those business units that make decisions on the assumption that IT will automatically be able to cope, will have little impact on the company's operation. IT resources will be dynamically allocated, and the dangers of poor communication will be minimised.

The cost of not considering capacity planning (ST-MT)

Capacity planning provides an excellent method to reduce overall IT costs and enhance IT service levels. Rarely does a technology come along that can help align the IT department to the business so effectively. Not looking at capacity planning solutions could prove to be a costly mistake if there is a spike in utilisation or systems failure. If reserve resources are not provisioned for effectively, service levels will fall and the company could start to see a direct impact on revenues and profits. Accountability for scenarios such as this will inevitably fall on the person responsible for the IT department. CIOs, IT directors and similar executives need to minimise the impact of utilisation spikes, systems failure and other associated potential problems. Capacity planning not only provides the means to do this, but can provide a lower overall cost base for the IT division. In Datamonitor's opinion, IT managers should at least consider capacity planning as a way to enhance their business unit's performance and increase the importance of the IT department's role within the organisation.

Supporting the future with capacity planning (LT)

One of the core trends set to emerge in the capacity planning market is the use of the technology to underpin future technological deployments. Service orientated architectures (SOAs), real-time data access initiatives, customer relationship management and industry specific technology such as radio frequency identification (RFID) will all be helped by capacity planning. Technologies such as these can prove

Capacity Planning

to be highly irregular in their utilisation patterns, and as such a capacity planning solution could prove the perfect starting point for multiple future strategies.

Definitions

- ERP Enterprise resource planning;
- CRM Customer relationship management;
- ROI Return on investment;
- Failover The transfer of operation from a failed system to a similar, redundant system to ensure uninterrupted availability.

Related readings

BMC Software <http://www.bmc.com>

ITIL <http://www.itil.co.uk>