The next decade of data centre evolution

If you don't know where you're going, any road will take you there

By Dale Vile, January 2013

Originally published on

http://www.theregister.co.uk/

To some, the mere concept of enterprise datacentres still being around in ten years' time is anathema. By then, it is asserted, all enterprise IT should be running in public clouds. The only people who can't see this are insecure box-huggers frightened for their jobs, and IT dinosaurs with no imagination.

For the remainder of this discussion, however, we're going to make the assumption that the world's IT won't be under the total control of cloud providers such as Amazon, Microsoft, Google and Salesforce.com in 2023. That's not to say we won't be using more cloud services, we undoubtedly will, but infrastructure running on premise will probably still be the centre of gravity for IT in the majority of medium and large organisations.

So does that mean datacentres will remain as they are? Almost certainly not, if for no other reason than natural refresh cycles will bring with them new technology and new ways of doing things, even if you have no plan to change things proactively.

Best to think ahead

If things are going to be evolving anyway, though, it makes sense to think ahead and make sure that changes are introduced in a coordinated and optimum manner. If you are dubious about the value of doing this, just take a look back over the last ten years.

While most IT departments have seen x86 server virtualisation enter their worlds, the ones that planned and managed its adoption in a considered manner are in a much better state today than those who just let things happen. If you took a proactive approach, you are probably enjoying the advantages of a more coherent, manageable and cost-effective environment. If you adopted server virtualisation in more of an opportunistic or ad hoc way, there's a good chance you are battling with virtual server sprawl and networking and storage bottlenecks, while still being a slave to a lot of tedious and error-prone manual administration that others have eradicated.

If you look across the data centre computing world, you'll find lots of other examples where new technology has had little or no impact, or even a negative impact, because of uncontrolled adoption: Unmanaged SharePoint installations leading to document and information sprawl, tactical data warehouse initiatives creating yet more disjoints and integration headaches, ill-planned unified communications implementations running into quality of service issues, etc.

Get it right, though, and it's possible to move the game forward with every round of change and investment driving towards better efficiency, improved service levels, greater flexibility and, not least, an easier life for IT managers and professionals.

A forward looking approach also means that investments today are more likely to be laying a firm foundation for the future. Those who took a structured and managed approach to server virtualisation, for example, now find themselves in good position to start looking at advanced workload management and orchestration (a.k.a. private cloud).

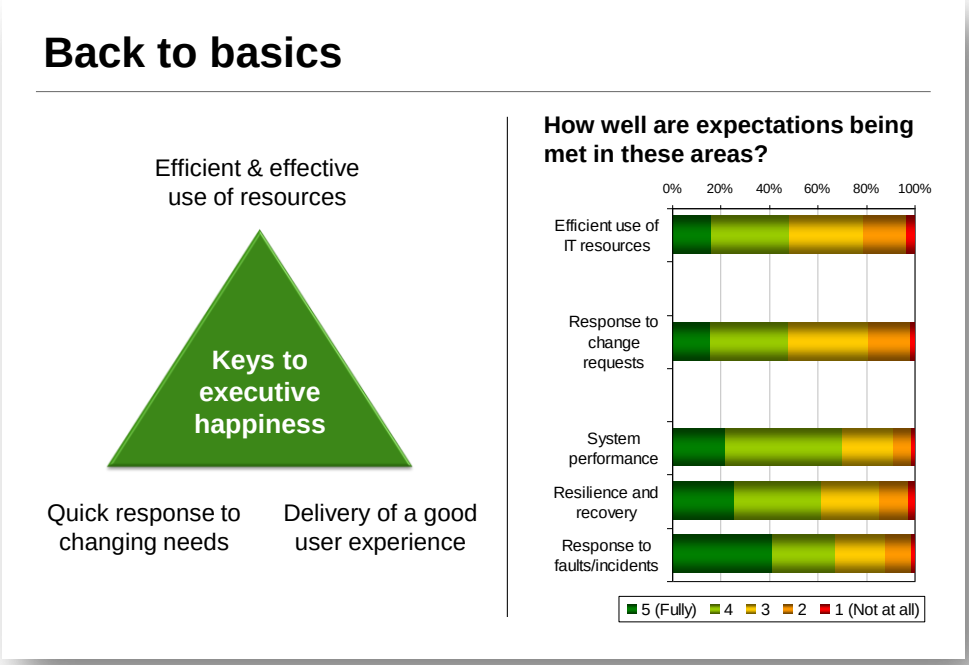
However, it's not just infrastructure and management technology that's evolving; the needs and expectations of users and business stakeholders are too. It's a bit of a cliché, but it really is true that IT departments are generally being asked to deliver more for less each year, and at a faster pace. It's therefore worth taking a minute to consider some of the specifics here.

Inescapable pressures and demands

At the highest level, it's possible to break out business level expectations of data centre computing into three main areas:

1. Efficient and effective use of resources (assets, people, external services, etc.)
2. Quick response to changing needs (new requirements, additional capacity, etc.)
3. Delivery of a good user experience (systems performance, availability, etc.)

These are summarised on the following figure, which also shows (based on a recent reader study) that IT departments are often perceived to fall short in these areas:



And things aren't going to be getting any easier. Continuous change and the increasing pace of change at a business level comes out strongly when we interview senior business managers, as does the degree to which business processes are becoming ever more dependent on IT through greater automation of business operations and direct interaction with customers and trading partners over the internet

Meanwhile, with everyone from management consultants, through investment analysts, to politicians and the mainstream media talking about cloud computing, the 'C' word has now made it into executive vocabulary. Even though business people often can't articulate the significance of

cloud computing in any precise, accurate or meaningful way, it's entry onto the scene has pushed the question of IT sourcing further up the business agenda.

The hosted cloud opportunity and challenge

Over the past few years, we have seen an explosion in the number and variety of cloud services available on the market, from basic hosted infrastructure at one end, to full-blown business applications at the other. The promises that typically accompany these services are now widely known – no up-front costs, fast access to new capability, less IT infrastructure to worry about, increased ongoing flexibility, and so on.

Less well publicised are the challenges that can arise when you start to make broader and more extensive use of cloud. As the number and type of services consumed proliferates, ensuring adequate integration between offerings from different providers, and between critical cloud services and internal systems, can become a problem. Related to this, end-to-end service level assurance often becomes more difficult, as does troubleshooting across systems and services boundaries, protecting information, assuring compliance, and, not least, monitoring and controlling costs.

To be clear, most of the problems are not to do with individual cloud services themselves (assuming you do your due diligence on providers), it's more about making sure everything works together safely and cost effectively. And in this respect, the ease with which end user departments, workgroups and even individual employees can adopt cloud services, while never thinking about integration, interoperability or information related requirements, already represents a challenge for some organisations.

So how does all this impact the data centre?

The unavoidable conclusion when you consider the dynamics we have been discussing is that 'business as usual' from a data centre perspective is not going to cut it over the coming years.

According to research, the typical data centre is a pretty fragmented and disjointed environment. Traditionally, both infrastructure and toolsets have been acquired off the back of application related investments. The 'stack based' approach to procurement, with each application typically pulling through a specific set of hardware and platform software with it, has led to the accumulation of multiple architectures over the last two or three decades, even multiple generations of each architecture in many cases.

Just think about how many versions of each hardware component, operating system and database management system you have, let alone how even the same software and hardware is configured in different instances to support specific application needs.

And is there one place you look or one team you can ask to get an accurate picture of what's in your infrastructure and how it all fits together? Probably not, and according to feedback from readers a combination of disparate tools and processes, together with demarcation of management activity between server, storage, networking, and other discrete teams, gets in the way of both efficiency and effectiveness.

If already overstretched operations teams are going keep up with continually escalating and changing demands, it's going to be necessary to remove complexity, increase the level of automation, and start to develop and manage the data centre in a much more holistic and inclusive manner.

The vision most often put forward to enable this is of the data centre becoming a coherent and dynamic virtualised environment, based on an architecture which allows internal and external services to be blended effectively. Service levels and responsiveness are then achieved via a unified management approach that works coherently across all important domains, i.e. servers, networking, storage, applications and cloud services.

But the Devil is in the detail

Of course such high level statements are easy to make, but hard to act upon. One of the biggest questions we hear back from our research is how to move things forward towards a more coherent vision or goal, while at the same time keeping up with all of the day to day stuff.

Sure, if you are starting with a green field site, putting your data centre together based on hybrid cloud and unified management principles with an integrated multi-disciplinary team looking after it is the way to go. You can even make sure that you have the right blend of cloud services in the mix, delivered by responsible, open and reliable service providers that are 'IT department friendly'. For most people, though, it's a case of setting an overall direction, then moving towards the creation of a more modern, flexible and efficient environment in a step wise manner.

This generally starts with the formation of a new team made up of some of your best server, networking and storage specialists, then using a discrete application requirement as an opportunity to start laying the right foundations.

The basic idea, and one that seems to work well for many, is to create a modestly scoped initiative that aims to get everything relevant in place in a joined up manner before scaling up. This provides an opportunity for the unified team to get to grips with new technology, tools and techniques, establish the right set of processes, and, not least, to figure out how different disciplines are going to work together effectively. The aim is to avoid the problems of trying to run before you can walk.

Once the initial beach head of future-proof goodness is established, it's a case of broadening the scope of the new environment in a controlled manner. This comes down to prioritising, scoping and phasing subsequent activity to deal with the accumulated problems of the past.

As a next step, some choose to focus on migrating key applications that would benefit from a more flexible, responsive and efficient execution and management environment. Others elect to get stuck into sorting out the 'long tail' of small footprint applications that clutter a lot of data centres, consolidating as much as possible onto a shared private cloud infrastructure. Where you start and how you move forward will depend on what's important to you – improved service levels and rapid change management for dynamic core systems, for example, versus cost savings and easier administration in relation to the broader Windows and Linux server estates.

Along the way, decisions can be made about which applications and data might be candidates to run via some kind of hosted cloud model, and equally, which systems should be left as they are (e.g. static legacy applications) or developed along their own evolutionary path (e.g. the mainframe, your HPC environment, etc).

It's a marathon, not a sprint, and IT can't go it alone

In terms of timescales, unless you have a data centre that is totally dysfunctional, the evolution we have been talking about is likely to play out over years rather than months. Over the coming few weeks, we'll be drilling down on some of the important aspects of data centre evolution including key technology developments, emerging best practices, and some of the inevitable political issues that IT departments are likely to encounter, if they haven't seen them already.

On that note, we'll leave you with one final thought for now.

We need to stop thinking about the data centre as a facility in which things are housed, and start regarding it as a shared corporate resource and a notional hub for coordinating the safe and effective use of internal and external services.

This translates to a significant mind-set and cultural shift for many organisations, with ramifications in terms of IT governance and funding, and even the fundamental relationship between the IT department and business stakeholders. The upshot is that any long term plan or vision for the data centre cannot be developed unilaterally. It's as much a business matter as it is an IT one.

About Freeform Dynamics

Freeform Dynamics is a research and analysis firm. We track and report on the business impact of developments in the IT and communications sectors.

As part of this, we use an innovative research methodology to gather feedback directly from those involved in IT strategy, planning, procurement and implementation. Our output is therefore grounded in real-world practicality for use by mainstream business and IT professionals.

For further information or to subscribe to the Freeform Dynamics free research service, please visit www.freeformdynamics.com or contact us via info@freeformdynamics.com.