Building the future proof datacentre

Architectural considerations for building a service centric datacentre

By Andrew Buss, January 2013

Originally published on

The Register® Biting the hand that feeds IT

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

The last decade has seen massive leaps in performance and capability across all areas of IT. This has enabled organisations of all sizes and types to approach doing business in new ways, or to do things better and more effectively. Despite this, many companies feel that the actual service delivered by IT often falls short of what is needed and expected. The end result is a lack of alignment between the direction and goals of the business and what IT is able to deliver.

Driving better alignment between business and IT is therefore usually at, or at least very near, the top of the list of priorities for CIOs and IT Directors, even if not formally documented as such. When we look at the top requirements that many companies have of their IT departments, it is to be responsive to requests for new or enhanced services, for performance and reliability guarantees to be met, for costs to be contained and more value to be delivered.

While the success of this type of alignment initiative is very dependent on the abilities and drive of people involved, architectural and technology choices and investment in the datacentre can have a real impact on how this goal is realised. In order to meet the requirements of agility, availability and performance, a great deal comes down to effective and automated service provisioning coupled with comprehensive service monitoring and management.

These are principles which underpin the concepts of Private Cloud, so we would expect those IT departments which are better aligned with business objectives to also be leading the way when it comes to the adoption of private cloud and this is indeed the case (Figure 1).

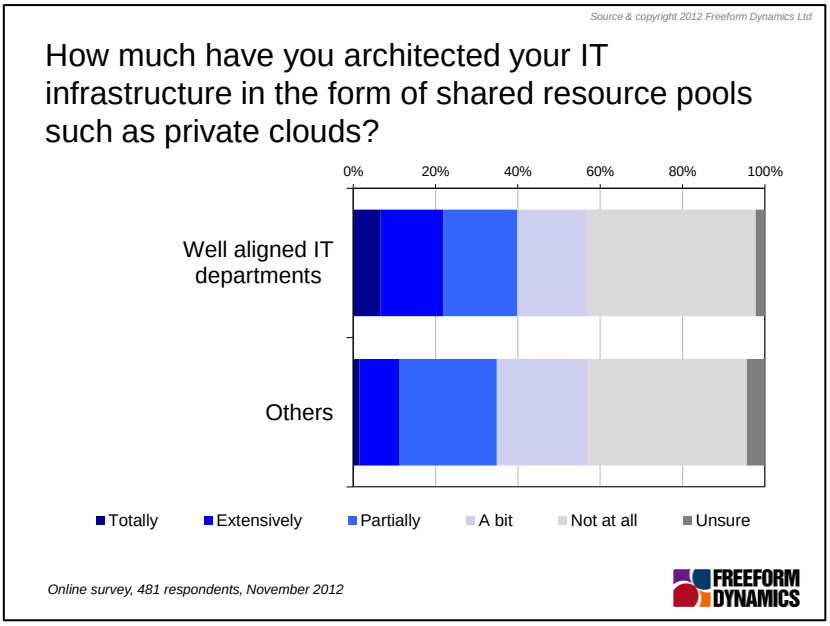

Figure 1

It is not only well aligned IT organisations that grasp the benefits that private cloud can bring. If the constraints of the real world are removed, the concept resonates strongly with many IT managers. However, when it comes to the practicalities of implementing a private cloud many companies are still in the stages of thinking about it rather than moving forward at this point (Figure 2).

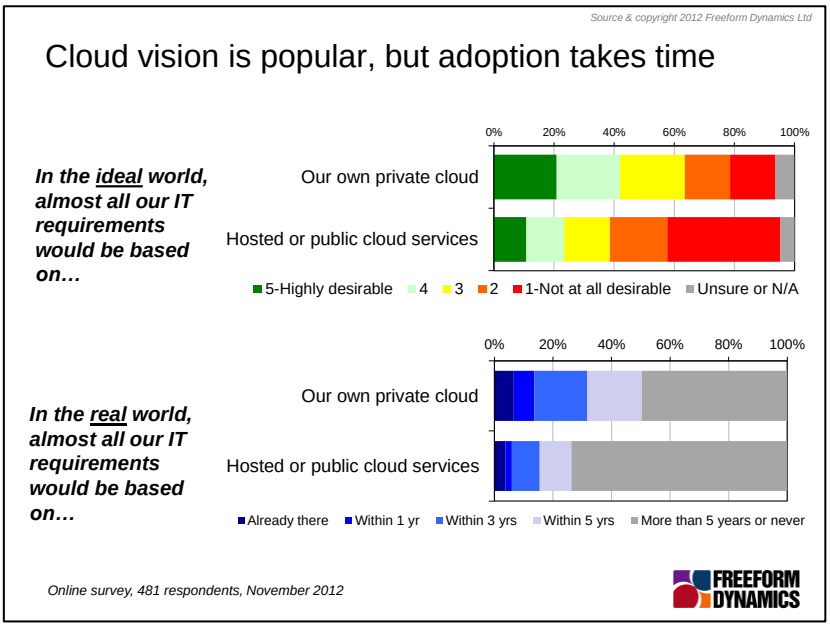

Figure 2

A minority may never be convinced of the benefits of a more dynamic approach to IT, but for most this is an appealing vision and one that delivers tangible benefits for those that have already made the move.

The end result is that the datacentre of the future will be built increasingly around the concepts of private cloud, including automation, orchestration, service level agreements, unified infrastructure and proactive monitoring and management. This may seem pretty obvious when sitting back and thinking about it.

But the way that projects and budget focus the attention on the here and now often means that the bigger picture is by necessity put off for another day - usually many times over. Making the case and getting started is often the hardest part, and to do this it can be useful to consider some of the architectural options that can shape your investment towards better alignment and more dynamic IT.

Evolution rather than big bang

While IT vendors may like to imply that the way to solve many of the problems facing the datacentre infrastructure is to sweep out the old and replace it en-masse, the path to private cloud is definitely not about a big transformational change.

For most IT organisations, it is very much about choosing a path that allows incremental improvements that helps make an initially small - but growing - part of the IT infrastructure more dynamic and 'cloud like' while preserving and even enhancing what is already in place (Figure 3).

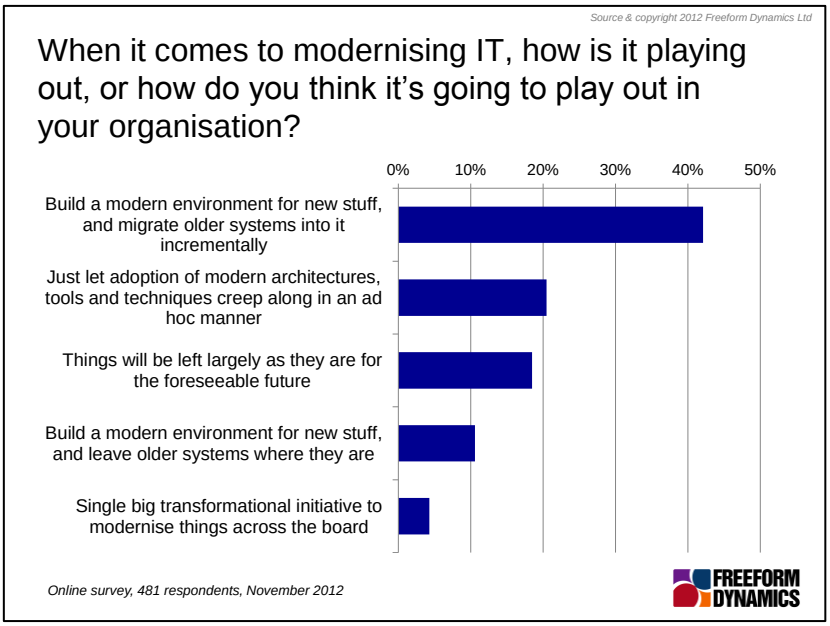

Figure 3

This then raises the question of what to do about existing or legacy workloads and systems. Leaving them where they are can simplify the introduction of modern environments such as private cloud, allowing a focus purely on supporting the new workloads. The end result is typically a smoother introduction with less time, cost and associated risk for the new environment.

Enhance the old, but not at any cost

This however, then leaves the older workloads without the benefits of advances in architecture and management that come with cloud type environments. The result is continuing fragmentation of the IT infrastructure, with workload or service 'islands' emerging that may not function very well together. They may also introduce, in areas such as change management or provisioning, potential bottlenecks that impact on how effective the modern environment is able to be.

The other approach is to try to gradually migrate the older services into the new environment. The advantage here is that these systems can start to benefit from much of the investment in shared services that are being integrated into the new infrastructure, with the potential for improved monitoring, management and flexibility. It is no surprise then that this is by far the most popular approach to modernising IT among the respondents.

On the surface, this approach makes a lot of sense. But if taken too far can have unintended or negative consequences. The ultimate aim is to be able to move workloads where it can be done

quite easily in terms of compatibility and performance, and where the effort involved does not result in having to compromise the implementation or functionality of the new, modern environment.

The ideal is not to think about forcing older workloads into the new system at any cost. Instead, a pragmatic approach is needed, with an understanding of the end goal of a successful migration. An assessment period would examine the feasibility of migration, and should any issues arise the decision can be made to either invest in migration so long as it does not interfere with the new environment, or to leave the workload as it is.

Investing in storage and network virtualisation

One of the critical underpinnings of the datacentre of the future will be the ability for the different elements of the infrastructure to be able to actively participate in any required reconfiguration due to changing or migrating workloads. To do this, servers, storage and networking need to be free of their physical or topological constraints in order to be configured on the fly by management policies or automated provisioning tools.

Virtualisation is one of the technology foundations for enabling this flexibility, and it has been broadly adopted for servers by many organisations. More recently, there has been a move to bring the advances seen in server provisioning and management to both storage and networking. While storage has seen a pickup in adoption, virtualisation in both remains some way behind that of servers (Figure 4).

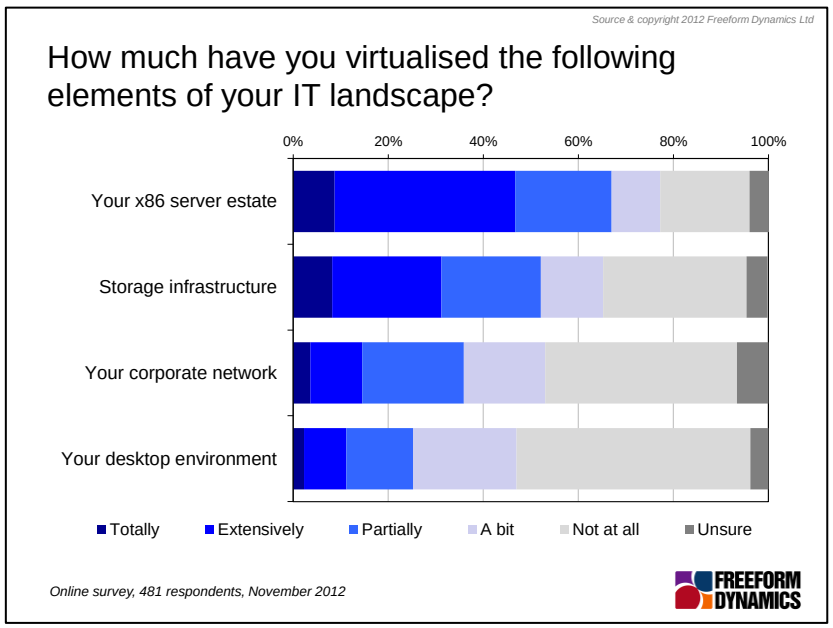

Figure 4

It may not seem to be a pressing priority to bring networking and storage up to the same level of virtualisation capability as the server estate. For many IT departments there is very likely still a lot of optimisation left to do with respect to servers and workloads never mind adding storage and networking into the mix.

But workloads are becoming increasingly virtualised, and server consolidation is fast approaching its limits in many companies. In order to drive further improvements in IT delivery, facilities such as automated workload migration are becoming more widely adopted.

For workload migration to be successful, both storage and networking need to be flexible enough to reconfigure quickly and easily when change is requested and virtualisation can greatly help with this. With storage and networking typically having a lifespan of five years or more once deployed, thinking of this requirement now and building it into current and future investment plans can help to stave off an inflexible infrastructure in the future.

Integrated stacks need to be interoperable too

Earlier in this workshop we touched on [LINK to Defragging the DC workshop piece] the concept of integrated stacks. This is where server, storage and networking from a single vendor, or close partnership of vendors, are pre-integrated and designed to work together as a whole stack. Many IT architects are wary of this approach, and view it as 'incomplete' or a 'lock in', preferring to choose the best vendors for each layer.

Much of this scepticism has been justified in the past, but vendors have been developing or acquiring the portfolio and capabilities to make this more workable. These 'best of need' integrated systems usually have good-enough overall performance at a much-reduced cost and with greater simplicity compared to the optimised performance and increased cost of using 'best of breed' components that need to be integrated. When it comes to jump-starting a move to private cloud, in many cases an integrated stack may prove attractive.

However good an integrated solution may be though, circumstances change and relying on a single stack can be a big bet to take. For this reason, any stack procured should be open enough to integrate with third party equipment, software and tools without an unreasonable amount of effort, and be appropriately supported by the vendor should such a need arise. If the solution is rigid and inflexible, it can seem like a return to the early days of computing with closed and proprietary systems.

A variety of standardised building blocks enable more flexible service delivery

When it comes to private cloud and the datacentre of the future, it is often claimed that everything should be simplified and standardised so that there is a single, homogenous set of building blocks with which to construct all the services the business needs.

While simplification and standardisation are good ambitions, as is a common set of building blocks, it may be counter-productive to attempt to force this approach across the board. For starters, workloads tend to have widely different requirements in terms of performance, reliability, cost and security to mention just a few.

Trying to run them all from a single set of building blocks can be done, but it can be tough to achieve in practice. In order to make the infrastructure more adaptable to demand, it may be better to have a set of building blocks optimised for some of the most common classes of workload requirements.

This may be done fairly simply along the lines of having good, better and best building blocks, where a variety of ever more capable – and expensive – equipment is used within a single private cloud to add choice and flexibility to the arsenal.

A good solution may be single socket blades for VDI workloads or simple server applications; a better platform could be dual-socket blades supporting more memory for consolidated virtual workloads, email servers or entry-level databases; a best option may be quad-socket servers supporting large amounts of memory for demanding workloads such as enterprise OLTP systems or business support and analytics.

Optimised clouds

In many situations, especially with a complex multi-vendor infrastructure supporting many different business units or customers, trying to build all the requirements into a single infrastructure or private cloud stack may ultimately prove to be unworkable given the conflicting goals and limits of manpower and budget.

This may require the development of optimised solutions that have been tuned for certain classes of workload or infrastructure service types, such as high performance computing, business analytics or

transactional databases. This tuning would typically involve hardware, software and high-level management, resulting in mostly self-contained private cloud 'islands'.

This approach of deploying multiple cloud 'islands' is not as pure or clean a vision of private cloud as many would like. But it can work well, particularly where there is a fairly clean separation between the various applications or services such as delivering end-user facing client application workloads compared to hosting the ERP backend or Exchange server facilities.

Building out into the public cloud

When it comes to actually delivering IT services, the default choice for most IT departments today and for the foreseeable future is to run them from their own private infrastructure. This is quite natural given the history of IT and the development of datacentres.

But there is a growing realisation that building out a datacentre to cover every workload and eventuality may be costing a lot of extra time, money and effort. This is leading many companies to look at the use of public cloud services to augment their own on-premises capabilities.

Software as a Service (SaaS) is typically bought as a complete service and only ever runs in the public cloud. Integration with internal systems is typically done through high level APIs, and all the underlying management and maintenance of the SaaS application is performed by the service provider.

Infrastructure as a Service (SaaS), on the other hand, is typically a raw mix of server, storage and networking resources that needs to be managed. To make matters more complex, each provider tends to have their own unique approach. This may have been acceptable when only a few workloads – usually test and development - were run in IaaS environments. But the desire to move production workloads in volume from the internal private cloud into the public cloud – and back again - requires compatibility with your internal management environment.

It is possible to architect your internal cloud infrastructure to match a specific suppliers own public-facing environment in order to achieve the consistency, but this does limit the choice of equipment and service provider with a long term impact on flexibility and service evolution.

A more sustainable approach may be to ensure that the internal architecture is developed in a way that enables your management and orchestration tools to recognise multiple different IaaS providers' environments. The tools would handle the translation between your private cloud and the different IaaS offerings, while maintaining and enforcing critical elements such as Service Level Agreements (SLAs), security, service monitoring and billing. One way to think of it is that the IaaS environment is just another managed resource pool that can be utilised to deliver services.

Orchestration and automation are the long-term keys to effective service delivery

Whichever way we look at it, the datacentre of the future is becoming less dependent on physical systems and individual applications, and becoming more of a service delivery hub that pulls together a variety of different services both internally and externally.

Underpinning this is a shift to thinking about IT as end-to-end services and this means bringing the overall systems and service management capabilities up to speed as this is often under-invested and highly fragmented. It doesn't matter how advanced the infrastructure or how shiny the new servers are if things do not gel well together.

If there is one thing that will prepare the datacentre of today to be the future proof datacentre of tomorrow, it is recognising that management, and in particular automation and orchestration, will be the glue that holds it all together. In ten years' time the servers and software of today will probably be unrecognisable. But the management policies and service level agreements may well be very familiar to those who put them in place today.

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.