InfoWorld
Tony Bishop
The Real-Time Enterprise

Datacenter in a Box: Delivering Next-Generation Datacenters Today

One of the biggest barriers to innovation today, is dealing with current datacenter complexity. This forces the largest IT investment dollars to be focused on keeping the lights on and containing infrastructure sprawl. The inability to focus a majority of time and investments on innovating and differentiating the business thru IT causes continued missed expectations and disappointment within the business user community.

It is essential to understand that past design choices resulted in complexity, waste, performance barriers and cost models that don't work for the customer or best in class distinction. The lack of transparency of what has been done in the past and how datacenter build-out continues to be a "don't change" mindset – results in continued misalignment with business needs. This prevents business agility and reduces shareholder value.

In our experience, datacenter infrastructure design has to become a holistic footprint of infrastructure (network, storage, compute, memory, disk, I/O, housing container, power, cabling) that is efficient and effective in design, deployment and provisioning. This footprint concept has to be thought of as a composition of service delivery units that can be consumed in a variable manner – on demand consumption and on-demand re-configuration.

Practically, entire datacenters cannot be converted overnight. Additionally, even if organizations adopt build an optimal footprint/design, it requires a bridge strategy to intersect with legacy design & infrastructure. The bridge strategy needs to begin with understanding the traditional datacenter design limitations. Below is a synopsis of design limitations that we have found in the past:

Horizontal Cooling – this approach contradicts natural physics. If heat rises, then optimal cooling strategies should naturally leverage this. The cause/effect of not doing so, results in tremendous performance impacts (user experience of the business suffers) and waste in terms of real estate space, additional power consumption and additional cooling infrastructure (ROE suffers).

AC Power – power to datacenter infrastructure in many cases goes thru a power distribution unit or PDU that transforms feeds into traditional plug connections into servers, network and storage equipment. This results in waste in terms of power loss (not a good thing with the global energy issues today besides additional costs) and potential instability when swings of voltage and current get out of time with each other (user experience can suffer as infrastructure stability can become erratic).

Proximity of Devices – traditional datacenter infrastructure design does not consider proximity of the various devices that comprise a service unit that delivers processing to users. This results in significant performance impacts (user experience of the business suffers) whereby compute, memory, I/O fabric, disk, storage, connectivity to external feeds are provided in terms of layout not in terms of service delivery. Performance can be impacted by 30X due to this approach. Moreover, this creates waste in terms of unnecessary network traffic congestion and bandwidth usage (ROE suffers).

One Size Fits All – most datacenter infrastructures and typical vendor strategies are built around a perceived "standardized" footprint. The problem is that this is designed typically bottom up with little or no correlation to the workflows, workloads, information, content and connectivity requirements of the business and its competitive needs. These disconnects results in performance, unnecessary costs, waste and results in agility issues for both the business and IT.

Layout – the typical datacenter layout incorporates homogenous pooling of various classes of resources. Servers by multiple classes are typically in multiple pools; storage by file or block are in different pools in their own area of the datacenter; network load-balancers, network switches, network routers are pooled/deployed across various areas of the datacenter. This approach is sound in terms of physical floor organization, but is not designed for optimal workload throughput nor time to provision. The average provisioning cycle in datacenters with this kind of layout are in terms of weeks or months versus the minutes or days needed to provision, troubleshoot or perform to meet the needs of the business.

Approach
A top down approach to mapping business "processing demands" with relationship and dependency of "processing resources" correlated against "infrastructure qualities" to visualize, capture and isolate how to enable the build a datacenter infrastructure that ensures the optimal service & information is delivered in the best manner possible in terms of performance, costs, availability and efficiency to help differentiate the business.

To bridge the legacy, we used instrumentation & forensic analysis to systematically and objectively decompose cause/effect, usage of infrastructure, business dependencies of the legacy infrastructure design, and mapped these findings with the associated impact of on our quality of experience provided to the business.

The Strategy – *build a "Datacenter in a Box*:

There are proven architectures, technologies and operating models that can resolve the design limitations outlined above. Strategically oriented organizations have, can and will start to orient future designs and deployments of datacenter infrastructure that will align and differentiate their business.

Below are the three building blocks that can help an organization create an aligned datacenter strategy. This strategy creates a plug and play approach of creating self-contained datacenters within a datacenter – or another way to term this is a datacenter in a box.

Optimized Footprint – a self contained fabric of logical building blocks housed in a single footprint and interconnected with high-speed fabrics between each container. The building footprint container includes: multi-core processors, optimized memory configurations, multiple I/O fabrics (10GigE, Ethernet, Fiber-channel, Infiniband), optimal commodity disk (solid sate and spinning), simplified storage with tiers of provisioning, processing appliances, integration appliances, connectivity appliances, integrated switches, networking accelerators – all housed in vertically cooled containers with DC feeds, simplified cabling and a virtualized stack from the network to the transaction.

Fit-for-Purpose℠ Design – business driven workload and consumption behaviors (with factors of wall clock, calendar, special events, geography and channels) need to be encapsulated in service contracts with qualities of experience associated to performance, priority, costs and efficiencies. This then needs to be translated into a mapping and orchestration model of workload matched with qualities of infrastructure service units that are then translated into the optimized footprint configurations. Predictability for datacenter managers comes in the form of leveraging intuitive infrastructure patterns that correspond to the handful of processing workload types to be able to plan the rollout of optimized footprints.

Virtual, Portable and On-Demand – with the innovation and creation of transforming shipping containers into self-contained datacenters, organizations can radically impact cost, time to provision and accelerate the ability to deliver new datacenters on demand. These containers provide the cooling, ventilation, power supply, UPS', and can house optimized footprints of compute, storage, network, I/O fabrics, etc… in a density manner that requires 1/3 of the traditional datacenter floor space, 1/5 the time to build at up to 80% less costs. Moreover, this strategy transforms the real estate equation/limitation of traditional datacenters by being able to be deployed outside, in a secure warehouse, in a secure parking garage, on an 18 wheeler.

The bottom line is, to remain competitive in a global economy, to provide differentiation for your business and to drive break through innovation in terms of performance costs and efficiencies, firms must start to adopt a new form of datacenter infrastructure design.

Posted by Tony Bishop on March 21, 2008 (http://weblog.infoworld.com/real-time-enterprise/archives/2008/03/datacenter_in_a.html?source=rss)