

# The Case for Medium-Sized Regional Data Centres

Kate Craig-Wood

Memset Ltd, Surrey University  
25 Fredrick Sanger Rd.  
Guildford, UK  
+44 1483 608012  
kate@memset.com

Prof. Paul Krause

Dept. of Computing  
Surrey University  
Guildford, UK  
+44 1483 689861  
p.krause@surrey.ac.uk

Nick Craig-Wood

Memset Ltd.  
25 Fredrick Sanger Rd.  
Guildford, UK  
+44 1483 608013  
nick@memset.com

**Abstract**—Cloud computing is widely associated with major capital investment in mega data centres, housing expensive blade servers and storage area networks. In this paper we argue that a modular approach to building local or regional data centres using commodity hardware and open source hardware can produce a cost effective solution that better addresses the goals of cloud computing, and provides a scalable architecture that meets the service requirements of a high quality data centre.

In support of this goal, we provide data that supports three research hypotheses:

1. that central processor unit (CPU) resources are not normally limiting;
2. that disk I/O transactions (TPS) are more often limiting, but this can be mitigated by maximizing the TPS-CPU ratio;
3. that customer CPU loads are generally static and small.

Our results indicate that the modular, commodity hardware based architecture is near optimal. This is a very significant result, as it opens the door to alternative business models for the provision of data centres that significantly reduce the need for major up-front capital investment.

**Keywords**—Cloud computing; Virtualisation; Data centres; Utility computing; Infrastructure as a Service.

## 1. BACKGROUND AND RATIONALE

A key factor in the dramatic increase in interest in Cloud Computing has been the offering of Cloud services by major keystones in the digital economy: Amazon, Google, Microsoft, Sun. Other less publicly known providers, such as Rackspace, have joined this space. However, there is a widespread perception that “Cloud” means a flexible rental scheme for compute resource from a single mega-data centre. This view is well articulated in the Berkeley take on Cloud Computing where it is argued that:

“the advantages of the economy of scale and statistical multiplexing may ultimately lead to a handful of Cloud Computing providers who can amortize the cost of their large datacenters over the products of many “datacenter-less” companies.” (Armbrust et al, 2009)

One of the arguments for this is by analogy with the semiconductor industry’s move to a small number of high cost fabrication lines supporting a large number of “fab-less” semiconductor innovators. However, there is a key distinction in the form of business relationship in the two cases. In the semiconductor industry, the relationship between requestor (fab-less semiconductor innovator) and supplier (mega fabrication line) is asynchronous. The customer can, within limits, absorb an unexpected delay on the supplier side. In addition, the customer could mitigate this risk by contracting more than one supplier: there is a standard “interface” between

designer and manufacturer. This provides a critically important duality. The supplier can time-multiplex the use of expensive fab lines in order to secure return on investment. In return, the consumers can choose and transfer their relationships between a choice of fab plants in event of loss, or reduction in quality, of service.

In contrast, the relationship between a “datacentre-less” service innovator and a cloud provider is synchronous. As witnessed by the 22-hour downtime of Microsoft Azure on 13/14 March 2009, an interruption or loss of quality on the supply side will immediately and directly impact on the revenue generated by the customer’s service. In addition, the interface between supplier and consumer is typically non-standard. We are not really seeing the resilience of a true cloud. Instead, the consumer sees some important benefits of flexibility and scalability, but has a very tight dependency on their chosen cloud provider – the relationship is highly asymmetric in the degree of empowerment to mitigate risk.

In fact, it turns out that the arguments for economy of scale are largely fallacious. We challenge the vendor-led view that expensive blade servers and/or storage area networks are needed for a virtualised platform. The use of commodity hardware, combined with the use of open source software for the platform can deliver cost effective utility computing.

Armbrust et al (2009) provide the following cost comparison (quoting (Hamilton, 2009)) between a medium-sized data centre and a mega data centre:

Technology	Cost in Medium-sized DC	Cost in Very large DC	Ratio
Network	\$95 per Mbit/sec/month	\$13 per Mbit/sec/month	7.1
Storage	\$2.20 per GByte/month	\$0.40 per GByte/month	5.7
Administration	≈140 Servers / Administrator	>1000 Servers / Administrator	7.1

Our figures from Memset, a successful example of a small-sized data centre, show that the economies of scale can be realised through the use of commodity hardware rather than powerful blade servers.

Technology	Memset (small DC)
Network	£14 (\$21) per Mbit/sec/month
Storage	£8.83 (\$13.30) per TByte / month
Administration	≈385 Servers / Administrator

The only area where there is a significant economy of scale (only a factor of 2.6) for a mega data centre is in the number of servers that can be maintained by a single administrator.

Furthermore, a significant body of data supports the effectiveness of a data centre architecture based on commodity hardware. Hourly samples of CPU utilization, disk read/write data rate and disk transactions per second (TPS) were measured for 950 individual virtual machines (VMs) on 45 virtual machine hosts (VMHs) in Memset's (see below) real-world deployment. The users and applications of the VMs ranged across all sizes and types of business. The VM platform was managed using open source Xen 3.2, and the VMHs were low-cost single-CPU commodity server hardware.

Analysis of the data shows that CPU availability is very rarely a limiting factor; directly challenging the widely-held view that CPU resources are the main limiter in VM deployments. The paper also shows that TPS, while rarely limiting, are more frequently limiting in the architecture used. This supports the hypothesis that one should aim to keep the TPS-to-CPU ratio high, supporting the practice of using a large number of commodity servers with local disks.

Overall, this paper argues that the move to utility computing is better served by a move to more "agile" local or regional data centres, built in a modular fashion. The data provided in the paper provides strong supporting evidence for this claim.

### 1.1 Memset

Memset, a managed hosting company, started renting virtual servers (Miniserver VM<sup>®</sup>) to consumers and business in late 2002, originally using User Mode Linux, but in 2005 moved to the open source Xen hypervisor (now using version 3.2). This practice has been popularised by a growing acceptance of virtualisation as "industrial strength", and is now more commonly referred to as Cloud Computing or Infrastructure as a Service (IaaS).

Over the last 7.5 years Memset has accumulated a set of best practices for optimal environments for hosting large numbers of virtual machines. However, many of those practices go against widely held beliefs, and this research is intended to analytically test those views.

Despite its leadership in the field of virtual servers, Memset is a relatively small company. They are turning over £2m, profitable, debt-free and growing at roughly 30% per annum. Their share of the approximately £1bn UK hosting market is therefore 0.2%. However, despite their size they are able to offer the highest quality services; voted UK's Best Web host 4 years running by PCPro magazine, and consistently in the top 10 global rankings of hosting company reliability according to the Netcraft survey. They also currently boast the lowest prices for virtual server rental in the UK, compared against the open information on other providers' Web sites. This makes them an ideal test case against the growing view that IaaS can only be delivered at commodity prices from mega-scale data centres. Memset's servers are housed in a 10,000 square foot 2MWatt data centre facility.

### 1.2 Commodity hardware and open source software platform

Accepted wisdom of the Memset systems administrators is that central processor unit (CPU)

resources are not normally limiting (research hypothesis #1), but rather disk I/O transactions per second (TPS) are more often limiting (hypothesis #2). RAM is a statically provisioned resource in the Memset architecture, and is fast-becoming the unit of comparison between VM providers.

Their rationale is that the limiter is normally the speed at which the disk head can physically seek<sup>1</sup>, measured in TPS, rather than CPU which appears to be the general belief, based on conferences and debates attended by the authors.

As a consequence of the beliefs of their systems administrators, the Memset virtual machine (VM) hosting architecture does not follow vendor-led wisdom of expensive blade arrays for huge CPU capacity, with storage on even-more expensive disk arrays accessed over fibre channel. Instead, Memset uses commodity 1U Dell PowerEdge servers, with two disks in a RAID(1) configuration.

The latest VMH specification is a R300, with single quad-core 2.5GHz Intel Xeon processor (minimum spec), 2 x 2,000GB 7,200 RPM 3.5" SATA disks software RAID(1) and 6 x 4GB DIMMs (total value approx. £1,200). Faster (eg. 15k RPM 2.5") disks are not employed because they are prohibitively expensive, and are only an option in larger, multi-disk chassis, and overall give a much less cost-effective TPS-to-RAM ratio.<sup>2</sup>

Each VMH runs the Xen hypervisor. Its domains (VMs) run their disk off a local disk array using the Logical Volume Manager on a RAID1 array. Their CPU scheduling parameters are set and their RAM allocations are fixed to make them as independent as possible. CPU is fairly shared among the VMs in proportion to the amount of RAM allocated to them. Disk TPS is shared fairly, but is not shared in proportion to RAM, using the Linux kernel's CFQ ("completely fair scheduler") I/O scheduler.

One oft-cited reason for using a shared storage array for the VM disk images is to facilitate real-time migration of VMs between front end machines. Memset have not found there to be a requirement for this among their customers, however, since those that need resilience simply rent two VMs on different VMHs with fail-over between them. Additionally, the CPU load profiles for most customers appear to be quite static and small (research hypothesis #3), which

<sup>1</sup> A 7,200 RPM disk will rotate 120 times per second. On average, a seek will take half a rotation. Therefore, when using RAID1 which effectively doubles the transaction capacity, the peak TPS is expected to be a little above 480 due to intelligent caching and request sequencing.

<sup>2</sup> Memset has been experimenting with Solid state (flash) drives (SSDs) as an alternative to rotary disk drives since they would appear to be the answer to TPS-bound applications. In their tests, however, they found SSDs performed no better than high-end rotary disk drives in seek-intensive operations. The issue appears to be the large, fixed 128kByte read/write block size of SSDs, meaning that even for a tiny read operation an entire 128kB block has to be read, and that additional bandwidth clogs up the serial interface. Memset anticipates that the issue will be resolved in time, perhaps with placement of SSD on the PCI bus, with better caching or with a reduction in the SSD read/write block size itself.

further reduces the need for live VM migration.

To summarise, Memset use commodity hardware and local VM images for five reasons:

1. It is cheaper.
2. Costs scale in smaller, more manageable increments.
3. Disk TPS are believed to be limiting, thus local RAID(1) should be better.
4. 2x 7.2k disks chassis (vs. 4/6/8 15k disks) gives the most cost-effective TPS-to-RAM ratio.
5. There is no demand for live/hot VM migration.

### 1.3 Data included in study

VM host (VMH) servers are used exclusively for either Linux or Windows VMs, there being 36 and 17 respectively of each type in the study. Due to unreliability with the open source paravirtualised Windows drivers (found during Memset's testing), only the Linux VMH use paravirtualisation. There is, therefore, expected to be an additional performance overhead associated with the virtualisation layer on the Windows hosts compared with the Linux hosts.

Until April 2009 the VMs were provisioned with swap partitions enabled. This greatly exacerbated the disk I/O issues, and it was frequently (and painfully) obvious that the host servers were I/O bound with disks "thrashing". Since then, with the rapid reduction in RAM prices making it more feasible, swap has been disabled on the Linux VM range (not on Windows VMs and VMHs)

This study only looks at the newer VMH deployed since April 2009. To compensate for the lack of swap, and in anticipation of disk I/O being less loaded, Memset increased the RAM in the host servers to 16GBytes. This also increased the number of VMs deployable on each host server.

During the course of this study it was noted that even with the increased number of VMs per host the servers' resources were being significantly under utilized. Therefore Memset increased the RAM in new VMH to 24GByte during the study (mid-January 2010), allowing 50% more VMs to be hosted on each VHM. This study examines the 21 large (24GB) VMHs and 36 small (16GB) VMHs independently where suitable since it is anticipated that they have different usage patterns.

Generally, Memset's VM customers are motivated more by cost savings than over-specification of the server, otherwise they would rent a physically dedicated server from the company. Therefore, this user group should represent a higher workload for a given amount of computer resource than, for example, a virtual machine deployment servicing an internal corporate demand.

## 2. METHODS

### 2.1 Power and bandwidth data collection

From bench testing of server equipment, Memset has also shown that Dell PowerEdge servers power draw only changes appreciably from the

idle state with increased CPU utilization, and not greatly with increased disk activity. A Dell R300 like those used as VMHs (but with less RAM in this case) draw 87 Watts when idle, 131 Watts when at maximum CPU, and 137 Watts when both CPU and disks are at maximum. Those results suggest a 50.6% power draw increase possible due to increased CPU activity and a possible 6.9% increase from increased disk activity.

This knowledge, combined with the observation by Memset systems administrators that there is little diurnal variation in their power requirements, suggests a crude test of CPU utilization. Therefore, the hourly power draw across all servers was recorded, and was compared against the total bandwidth use across all servers as a very crude measure of activity.

Most of Memset's customers are UK-based, and their customers' users are also mostly UK-based. The majority of the applications hosted by Memset are Web sites or Web applications, therefore bandwidth should be a reasonable approximation of average load across the servers. This data includes fully dedicated servers as well as VHM, and also includes older VHM machines.

Both sets of data were already being collected via the Memset infrastructure and stored in their central database, and was extracted with simple SQL queries. If CPU utilization is generally low without limiting little-to-no correlation between bandwidth, power draw is expected (hypothesis 3).

### 2.2 VM resource utilization data collection

For the main research project (testing hypotheses 1 and 2), a set of scripts were written, using Python and some Linux shell, to query the VHM and log resource utilization data hourly.

The Xen hypervisor provides accurate accounting data for CPU usage since it is involved in scheduling the Virtual CPUs. Its user space tools from the master domain were used to collect CPU usage figures (both spot and cumulative).

The Linux kernel on the master domain accounts for the disk traffic to and from the LVM disk array. Use standard Linux userspace tools to account for the disk traffic of domain attached to respective logical disk. This includes total disk reads and writes, spot disk I/O and spot transactions per second (TPS). Table 1 shows the data fields collected.

### 2.3 Data analysis

The VHM were grouped into four classes (Linux 16GB, Linux 24GB, Windows 16GB and Windows 24GB) so that comparisons were with-like. These groups are referred to as the VMH-groups.

Of the variables available, only CPU percentage, TPS, and spot read KBps (R-KB/s) were examined in detail (the core metrics). Memory

utilization profile and VM count was also noted for some data sets.

In total, a little under one gigabyte of data was collected over 8 months. 5.1 million individual VM-hour samples from 1,416 different VMs across 53 VMHs. This volume of data necessitated a programmatic approach to the analysis. Therefore, the data was organized in a SQLite database, and a set of Python programs were designed and written which could break down the data sets and produce the following:

- **Averages:** Simple average of resource utilization for each core metric, within each VMH group.
- **Load bands distribution:** Average time spent at 5 percentile ranges for an individual core metric averaged across all VMH within each group which were at least 75% full (by RAM).
- **Find peaks:** Peak search for each core metric among each VMH-group. Where VMHs share the highest peak, finds one with most counts of peak.
- **Single VMH detailed:** Detailed data for each core metric on each VM for a specified VMH around a specified date-time (usually a peak). One chart per core metric, each showing that metric over time for each VM individually.

Table 1. Data fields collected and their sources.

Field	Description	Source
created	The date and time the info was collected.	
host	The VMH that this info was collected from.	
name	The Miniserver VM name for which this data was collected (hidden in results).	
id	The xen domain id - a number just used for identification purposes	xm list
mem	MBytes of RAM for the VM. This is a fixed allocation without ballooning.	xm list
vcpus	Number of Virtual CPUs assigned to this VM.	xm list
stat	Spot status of the VM - whether it is running or not.	xm list
cpu_time	Total amount of CPU time the VM has used (since last reset)	
cpu_pc	Instantaneous percentage of CPU the VM is using.	xm top
tps	Average disk transactions per section.	iostat
read_kbps	Instantaneous kBytes/s read from the disk.	iostat
write_kbps	Instantaneous kBytes/s read from the disk.	iostat
kb_read	Total kBytes read from the disk.	iostat
kb_written	Total kBytes written to the disk.	iostat

Names of individual VMs were removed for security and customer privacy since host names are discernable from the Memset VM naming scheme.

### 3. RESULTS AND ANALYSIS

#### 3.1 Power and bandwidth

A plot of power draw (instantaneous) and bandwidth (averaged over the previous hour) against time over a period of five consecutive days (Friday to Tuesday) is shown in figure 1.

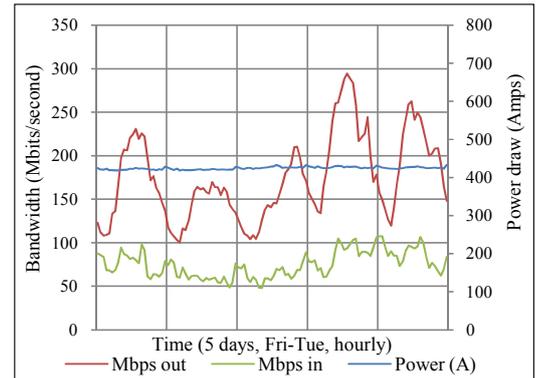


Fig. 1. Bandwidth and power over time

The diurnal bandwidth variance for the measured set of servers was significant, with the outbound data rate dropping below 75 Mbits/second (Mbps) at night and peaking at over 200Mbps during the day.

From figure 1 it is not possible to discern any relationship between activity (bandwidth) and power draw, therefore a second scatter plot of power draw at varying bandwidth levels was rendered (figure 2). The trend line on figure 2 (dotted, red) show the expected upward trend in power consumption with bandwidth, however it is a very small effect, only visible thanks to the large data set, which includes a total of 12 days' readings. Even taking the most extreme samples, the power variance with bandwidth is only 4.1%, suggesting that the CPUs are very lightly loaded and not appreciably moving out of an idle state even with the manifold increase in bandwidth activity.

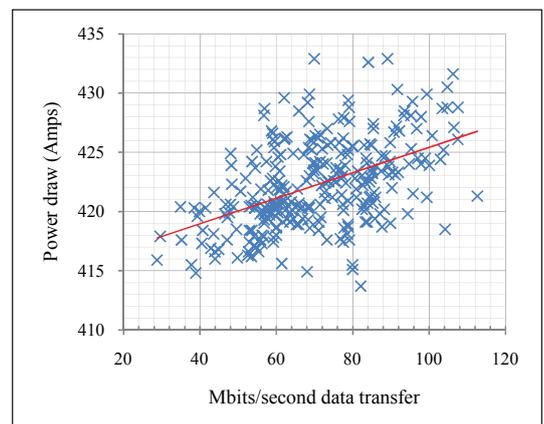


Fig. 2. Power against bandwidth

### 3.2 Averages

Table 2 shows the average utilization of each core metric across the entire VMH set.

Table 2. Averages

	Lin		16GB		Win		24GB	
	Raw	%	Raw	%	Raw	%	Raw	%
Host #	26		14		10		7	
CPU	46.3	11.6%	96.1	24.0%	58.4	14.6%	81.8	20.4%
TPS	46.2	9.2%	63.3	12.7%	-	-	-	-
R-KB/s	196	0.5%	362	0.9%	315	0.8%	210	0.5%
Mem MB	14,969	93.6%	22,262	92.8%	14,264	89.1%	21,483	89.5%
VM #	18.9		21.1		14.8		16.8	

For calculation of CPU percentage utilization, 100% was taken to be 400. For TPS 100% was taken to be 500 (the theoretical maximum) and for read KBytes/sec the observed maximum of 40,000 was used as the 100% figure.

As expected, the average resource utilization is overall higher on 24GB VMH compared with 16GB VMH, which can be attributed to a larger number of VMs on the 24GB hosts.

### 3.3 Peaks

Part of the analysis was to find peak states in the data. As well as locating useful time-slices on peaking VMH for further examination (see 3.6) this also yielded some differences between the Linux and Windows hosts.

#### 3.3.1 Disk I/O statistics under Windows and Linux

There is a clear and large difference in the peak TPS and R-KB/s figures between Windows and Linux servers. This is because the Linux VMH have a paravirtualised driver, but Memset does not use one under Windows. Therefore when streaming from the disk each 512 byte block appears as one transaction, limiting the usefulness of TPS statistics on Windows VMHs. Under Linux, the kernel is more efficient, grouping adjacent blocks together to reduce the number of transactions when streaming data. That efficiency is clearly working through the Xen layer too. However, that does still suggest an unforeseen weakness in our measurement system. Even under Linux

the TPS increased with data read rate, and streaming data results in high TPS readings, making the TPS readings less useful in supporting hypothesis 2.

#### 3.3.2 Effects of paravirtualisation

There is also an apparent difference between the CPU peaks under Linux and Windows, with Windows VMH peaking at only 330.7%, well under the 400% theoretical maximum. Some of the difference can be seen in dom0's CPU utilization. The Linux hosts CPU utilization peaked at 399.7%.

Under Windows the device model (disk controller, ethernet driver emulation etc) requires some dom0 CPU, hence the 26.4% and 16.4% CPU utilization (6.6% and 4.1% of total) by dom0 in those instances. Because the Linux VMs are paravirtualised and there is no device model almost 100% of CPU is available to the VMs, with dom0 only using 0.1% of total CPU capacity. The ~50% CPU unaccounted for under Windows is believed to be the Xen Hypervisor performing the interprocessor communications between device model and VM kernels. This would not show in dom0's CPU utilization, since it is just another domain as far as Xen is concerned. The unaccounted for CPU under Windows is being directly used in the hypervisor.

### 3.4 Load band distribution

Figures 3 shows the average time spent at each 5-percentile CPU load bands for each VMH group.

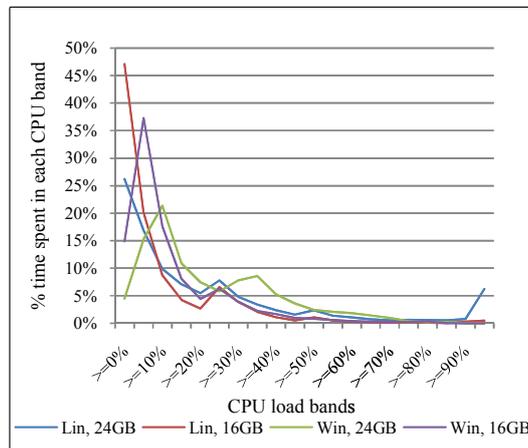


Fig. 3. % of time spent at each 5% load band.

The CPU load curve can be considered accurate, since the peak is definitive (400%), but the effects of paravirtualisation (see 3.3.2) should be kept in mind since this in-effect reduces the maximum CPU utilization under Windows. The sub-peak at 25% is symptomatic of VMs or processes maxing out individual cores on the host machines. This effect was marginal, however on the CPU-peak 24GB Linux VMH (figure 4), there was a small increase in time spent at higher CPU loads. In total, the 24GB Linux VMHs spent 10.1% of their time above 90% CPU load, whereas the

16GB Linux VMH spent only 1.3% of their time above 90%.

The CPU load profile of hosts running Windows VMs shows them to be generally a little more loaded, which would fit with the lack of paravirtualisation. The sub-peak at low CPU utilization for Windows VMH could also be indicating a generally higher background CPU utilization of Windows operating systems. A sub-peak at  $\geq 95\%$  utilization for 24GB Linux VMH suggests that with the increased RAM, CPU is occasionally now becoming a limiter.

### 3.5 Detailed examination of peak states

Figures 4 and 5 show a selection of peak CPU states in detail, with each line representing an individual VM. In each case the peak being observed occurs at the time mid point, with the time range spanning 4.5 days before and after the peak moment. The different vertical scales on Linux (400%, figure 4) and Windows (100%, figure 5) should be noted. Both charts show 24GB VMH peaks, but the ones for the 16GB hosts were similar.

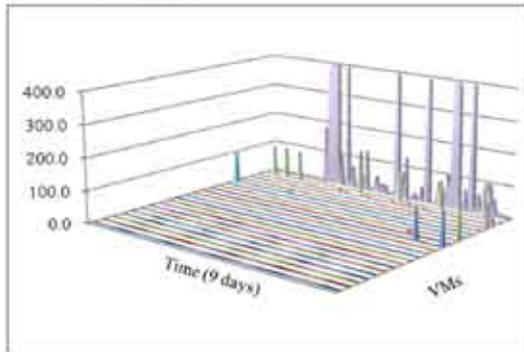


Fig. 4. CPU peak, Linux 24GB

For Linux VMH (figure 4 was typical of others examined) the CPU peak can be attributed to the activities of just one VM. The peak-finding algorithm was designed to find the most intensive (longest lasting) peaks even where many VMHs reached the same level, suggesting that (for the Linux deployment) coincidental VM CPU peaks are vanishingly rare. Those charts also demonstrate the general low CPU activity levels among the majority of VMs.

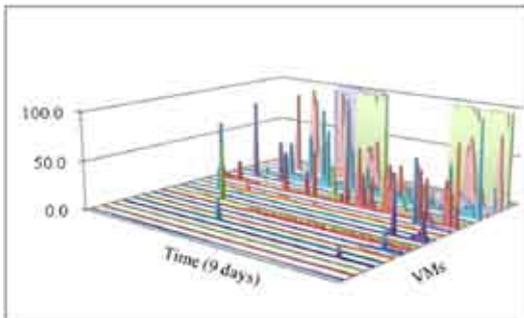


Fig. 5. CPU peak, Win 24GB

The Windows VMH CPU peak shown in figure 5 was also typical of other Windows host peaks, and is the result of a coincidence of individual

VMs maxing out a single CPU-core (of which there are four available). This is partially an artifact of the Memset architecture since it was only towards the very end of the study that they augmented their systems to allow Windows VMs to have more than one virtual CPU core.

## 4. CONCLUSION

### 4.1 Hypothesis #1: CPU is rarely limiting

According to the averages and load band distribution charts, CPU is indeed rarely limiting, confirming Hypothesis #1. Further, the observed CPU peaks were due to only one VM in the case of Linux VM and due to only a few coinciding VM CPU peaks in the case of Windows VMs. The peaks were sufficiently rare that even when in a limited state, only the few VMs bursting were affected; the vast majority of VMs used inconsequential amounts of CPU resource.

The small increase in time at high CPU loads for the 24GB Linux VMH compared with the 16GB Linux VMH suggests that the ratio of 24GB RAM to one quad-core Intel Xeon processor is near optimal. At no time were VMs deprived of provisioned resources (the peaks were as a result of VM CPU bursting beyond guaranteed limits), but a higher ratio could result in a restriction on the helpful ability to burst CPU.

### 4.2 Hypothesis #2: Disk TPS is more limiting

The data collected was an insufficiently accurate representation of what is believed to be the limiting factor (disk head seeks) due to data streaming causing increased apparent transactions and masking seek-bound states. This research has, however, enabled enhancements to the measurement systems to be designed, which should enable assessment of the limiting effects of disk TPS in future work. However, the very low average TPS does suggest that, as with CPU, TPS is at most rarely limiting for Memset's architecture. This also suggests that Memset's approach of maximizing TPS-to-CPU ratio is valid since they have largely avoided fully loading the VMHs, even when provisioned with maximum RAM.

With an architecture that was based on storage area networks (SAN) rather than local disks one would expect TPS to be more often limiting, necessitating a disproportionate investment in very high-performance SAN equipment, further supporting the case for a commodity platform such as Memset's.

### 4.3 Hypothesis #3: Most customer CPU loads are fairly static and small

The very small size of changes in power draw with activity levels (as measured by bandwidth) shown in figure 2 strongly supports the research hypothesis that most of the CPU loads among Memset's user base are fairly static and small. Additionally:

- We see low average CPU utilizations, and load distributions towards the low end of the

spectrum;

- The observed CPU load peaks are largely attributable to only a small number of busy VMs.

This supports hypothesis #1 since there is little need for the bulk of VM users to have the ability to migrate to other VMHs in real time in order to balance CPU load across the VMH grid.

#### 4.4 RAM

While this paper predominantly examines CPU and TPS as limiters in virtual machine deployments, it is worth remembering that the key resource considered when VMs are provisioned is normally RAM. RAM is fast, becoming the unit of comparison of VMs from different providers. Now it accounts for at least half of the cost of the VMH in Memset's architecture, and therefore presumably some others.

RAM is also an additional power drain, and since it is constantly being refreshed regardless of utilization level it is expected that the power drain changes little in relation to utilization (unlike CPU). Therefore it should be considered a resource to use efficiently, perhaps even more so than CPU which does at least partially auto-regulate its power consumption with activity levels. At present varying RAM availability in response to demand is difficult since operating systems (especially Linux) tend to efficiently make use of any spare RAM, giving misleading results for RAM utilization.

However, hot-(un)plugging VM RAM is possible under Xen, and following this research Memset intends to explore the practicalities of such an approach in order to maximize efficiency. It is expected that, at present, any liberated resources could only be employed in disk-I/O-light activities (e.g. a compute grid) since TPS would otherwise still become a limiter.

#### 4.5 Smaller, more modular data centres, using commodity hardware

Overall, the results clearly demonstrate the validity of the Memset VMH architecture. Following the increase to 24GB, which was as a direct result of this study, one can also argue that their approach is now near-optimal, as demonstrated by the occasional maxing out of resources; if limits were never reached that could be considered wasteful, but to very occasionally "bump" them suggests a good balance.

Further, the facts that Memset, a relatively small but also very financially sound company, has managed to achieve such industry-leading quality of service and value for money with this approach strongly supports the assertion that using commodity hardware for utility / cloud computing platforms is sensible. When combined with other benefits of this approach, in particular the reduced capital expenditure and modular cost scaling, the validity of the widely accepted wisdom that VM platforms should consist of blade arrays and SANs is seriously questioned.

The findings also seriously question the widely held belief that commoditised cloud computing services can only be delivered from mega-scale data centres by huge corporations able to make investments in the hundreds of millions of dollars range. This is important news for smaller, regional data centres since it demonstrates that they can compete with the international computing utilities (such as Amazon EC2) on price. Indeed, localised data centres can provide better quality of service to customers in the region thanks to cultural similarities and reduced connectivity latency.

## REFERENCES

- [1] Armbrust M., Fox A., Griffith R., Joseph A.D., Katz R.H., Konwinski A., Lee G., Patterson D.A., Rabkin A., Stoica I. and Zaharia M. 2009. Above the Clouds: A Berkeley View of Cloud Computing, Technical Report No. UCB/EECS-2009-28, University of California at Berkeley.
- [2] Hamilton, J. Cooperative Expendable Micro-Slice Servers (CEMS). In: Conference on Innovative Data Systems Research (CIDR '09) (January 2009)



#### Kate Craig-Wood

Kate is a multi-award-winning British technology entrepreneur, co-founder and Managing Director of Memset, and a nationally renowned expert-advocate of green ICT and Cloud Computing. She is an officer of BCS's influential Data Centre Specialist Group and is the youngest-ever main board member of Intellect UK, Britain's high-tech trade association. Kate also chairs Intellect's Climate Change Group, and currently co-leads the technical architecture strand of the Cabinet Office's G-Cloud and App Store

project. Kate is also undertaking a collaborative PhD in Cloud Computing with Surrey University.

Her company, Memset, is a nationally leading managed hosting and Cloud Computing provider. Memset is Britain's first Carbon Neutral ISP, and has pioneered virtual machine technology since 2002.



#### Paul Krause

Paul Krause is a Professor of Software Engineering. Originally a Physicist, he has over 20 years research experience both in Industry and Academia as a Software Engineer. Most recently, he was involved in the FP6 Digital Business Ecosystems IP as leader of the test automation work package. The University of Surrey also took on the role of developing the transaction manager for the DBE project. This role continues in the FP6 Project OPAALS, in which Paul Krause leads the workpackage on P2P networks and

transaction management for open B2B collaborations of SMEs. In addition, he is coordinator for the Interactive Computing domain of OPAALS.



#### Nick Craig-Wood

After receiving his Masters Degree in Natural Sciences at Cambridge University, Nick took his deep mathematical & IT skills to the communications sector and gained specialisms in cryptography and compression algorithms. He has always been an active member of the Linux & Open Source communities and is an expert in Linux systems administration and network communications as well as being an unusually highly skilled developer in many programming languages. He honed his expertise in the Web Hosting industry during his time as CTO of Easyspace Ltd. before moving on to co-found Memset in 2002, where he is technical director.

Today, Memset, is a nationally leading managed hosting and Cloud Computing provider. Memset is Britain's first Carbon Neutral ISP, and has pioneered virtual machine technology since 2002 under Nick's technical guidance.