SHARE
Technology · Connections · Results

Choose the Wrong Architecture and Waste Millions - A Customer Case Study

Mark Post
Novell, Inc.

Wednesday, August 15, 2007
9205

Agenda

- Disclaimer
- Caveats and Emptors
- Basic Scenario
- The First Project Meeting
- Internal Constraints
- High-level System Hardware Overview
- Other Hardware
- Software
- Power and Cooling

Copyright 2007 Mark Post

Agenda (2)

- Floor Space
- Wear and Tear on People
- How Busy Were Those Servers?
- The Mainframe Alternative
- Questions

I'll take questions during the talk unless time gets tight

Copyright 2007 Mark Post

Disclaimer

- My current employer (Novell) had nothing whatsoever to do with the project I will be describing.
- I will not be talking about which companies were involved, for obvious reasons.
- Costs for the mainframe version will be approximate, since IBM doesn't have list prices, per se, for z9 hardware (except IFLs). I was provided with cost figures (at the very high end), from someone who had access to them.
- I ***was*** involved in the project, and still have the scars...

Copyright 2007 Mark Post

Caveats and Emptors

- All hardware, software and maintenance costs are for 3 years.
  - The numbers for the mainframe get even better over 5 years.
- Not all hardware costs for the Intel deployment will be included in the dollar totals. (Don't worry, it's OK.)
- VMWare, although available at the time of the project, wasn't considered a viable option, so no virtualization was done.
- I'll only be talking about the Linux part of the project.
- z/VM will be configured to “over commit” real storage.
  - Standard operating procedure for z/VM shops
- z/VM's Virtual Switch will be used to connect the Linux guests to the network.

Copyright 2007 Mark Post

Caveats and Emptors (2)

- Storage will be configured with some percentage as expanded storage.
- Disk mirroring will be done in the DASD storage array, not by Linux or RAID controller.
- SAN costs are assumed to be equal between the two choices.
- People costs are assumed to be equal, even though managing mainframe Linux systems is less people-intensive.

Copyright 2007 Mark Post

Basic Scenario

- A new client required us to build ~50 Linux systems on Intel-based server class equipment.
- A much smaller number of Windows and Solaris systems were also to be built.
- Due to the client's promises to others, we had 2 months to get everything installed and in production.
- The systems were all going to be remote from where any of the Linux system administrators were located.
- A number of other bidders on the contract refused to commit to the 2-month timeframe.
- The project had upper management “visibility.”

Copyright 2007 Mark Post

The First Project Meeting

- ***No*** hardware had been ordered
  - This included racks, power distribution units, etc.
- The account rep. working with the client refused to let us order only two server configurations – they thought the client would complain about the cost of over-spec'ing.
- Pushing the schedule out was not an option.
- All of the teams working on the project were “leveraged,” i.e., not dedicated to one particular client.
- The network was still being designed.
- The requirements for system builds had not yet been received from the customer.

Copyright 2007 Mark Post

Internal Constraints

- We were supposed to use a commercial system management product to provision and patch midrange systems. This worked best when the system was cabled directly to the provisioning network.
  - We didn't have a place to do that, so other measures were taken to get access to the provisioning network. They weren't terribly reliable/fast.
- All servers had out-of-band remote management cards in them, connected via Ethernet, accessed via SSL-enabled HTTP (port 443).
- Racks could not be fully populated because of data center policies
- Are we having fun yet?

Copyright 2007 Mark Post

Development Systems

- 17 Systems
  - Additional GIGe NICs - 17
  - RAM (total of 146GB)
    - 1 x 2GB
    - 6 x 4GB
    - 5 x 8GB
    - 5 x 16GB
  - Processors (3.33 and 3.4 GHz)
  - 2 x 1 CPUs
  - 5 x 2 CPUs
  - 10 x 4 CPUs
  - Internal 72GB SCSI disks - 10 x 4, 7 x 5 = 75

Copyright 2007 Mark Post

Test Systems

- 7 Systems
  - Additional GIGe NICs - 7
  - RAM (total 96GB)
    - 2 x 8GB
    - 5 x 16GB
  - Processors (3.33 and 3.4 GHz)
    - 5 x 2 CPUs
  - 2 x 4 CPUs
  - Internal 72GB SCSI disks – 7 x 4 = 28

Copyright 2007 Mark Post

Production Systems

- 27 Systems
  - Additional GIGe NICs - 27
  - RAM (total 254GB)
    - 3 x 2GB
    - 2 x 4GB
    - 14 x 8GB
    - 8 x 16GB
  - Processors (3.33 and 3.4 GHz)
    - 3 x 1 CPU
    - 12 x 2 CPUs
    - 12 x 4 CPUs
  - Internal 72GB SCSI disks – 20 x 4, 7 x 5 = 115

Copyright 2007 Mark Post

Other Hardware

- HBAs for SAN storage – 56
- Fibre cables for SAN – 56
- RAM expansion boards – 37
- External 72GB disks for alternate boot drives – 37
- External disk enclosures - 19
- SCSI cards for external disk access – 37
- Mezzanine riser cards – 37
- Rack units (42u) – 13 (78 sq. ft.)
- Power distribution units – 26

Copyright 2007 Mark Post

Other Hardware (2)

- Switch ports – 204
- Ethernet cables – 204
- SAN switch ports - 56
- Hardware support 24x7 – 27 (+10 for ext. storage)
- Hardware support 13x5 – 24 (+9 for ext. storage)

Not including the costs of switches, routers, etc.

- They were leveraged (shared with other clients)
- I don't know how many were used, what kind they were, etc.

Copyright 2007 Mark Post

Hardware Costs for 3 Years

- $1,212,130.55
  - Doesn't seem too bad for 3 years, does it?

- Total Cost so far:
  $1,212,130.55

Copyright 2007 Mark Post

Software Licensing

- Test and Development
  - 29 Oracle database (per processor)
  - Oracle maintenance 22% of purchase price per year
  - 24 (3-year) Linux, including 9x5 support

- Production
  - 44 Oracle database (per processor)
  - Oracle maintenance 22% of purchase price per year
  - 27 (3-year) Linux, including 24x7 support
  - 2 clustering software

Copyright 2007 Mark Post

Software Costs for 3 Years

- $5,077,789.74
  - Ooh, that's gotta hurt.
  - And it doesn't even include all the costs, because I didn't have access to some of them. (Powerpath, for example.) But, it was the majority of them, so good enough.

- Total Cost so far: $6,289,920.29

Copyright 2007 Mark Post

Power and Cooling at Idle

- 5 x 1 CPU, 4GB RAM, 5 disks
  - 258 watts
  - 881 BTUs/hr
- 4 x 2 CPUs, 2GB RAM, 5 disks
  - 337 watts
  - 1150 BTUs/hr
- 5 x 2 CPUs, 4GB RAM, 5 disks
  - 341 watts
  - 1163 BTUs/hr

Copyright 2007 Mark Post

Power and Cooling at Idle (2)

- 3 x 2 CPUs, 4GB RAM, 4 disks
  - 473 watts
  - 1614 BTUs/hr
- 10 x 2 CPUs, 8GB RAM, 4 disks
  - 476 watts
  - 1622 BTUs/hr
- 12 x 4 CPUs, 8GB RAM, 4 disks
  - 626 watts
  - 2134 BTUs/hr
- 12 x 4 CPUs, 16GB RAM, 4 disks
  - 642 watts
  - 2191 BTUs/hr

Copyright 2007 Mark Post

Power and Cooling Costs for 3 Years

- 25,738 watts
- 87,782 BTU/hr

- At $0.0936 (US average for 2006) for electricity
  - 25,738 * 24 * 365 / 1000 * 0.0936 = $21,103.51 per year
  - For 3 years, $63,310.53
- At 0.7 of power costs for HVAC:
  - $44,317.37

- Total Cost so far: $6,397,548.19

Copyright 2007 Mark Post

Floor Space Costs for 3 Years

- Including service clearance, each rack requires ~17.5 sq. ft.
- At $220 per square foot per year:
  - 13 racks: 17.49 * 13 * 220 = $50,021 per year
    For 3 years = $150,064

Copyright 2007 Mark Post

Midrange Totals

- Hardware - $1,212,130
- Software - $5,077,789
- Power and Cooling - $107,627
- Floor Space - $150,064

- Midrange Total - $6,547,610

Copyright 2007 Mark Post

Wear and Tear on People

- Getting the remote access cards configured so they could be used remotely
- Configuring the RAID arrays using the internal disks
- Firewalls blocking access to the remote access cards
- Firewalls blocking access to the systems
- Wrong system build information
- Wrong network information (IP addresses, network masks, default gateways)
- Getting network connections made in the first place
- Getting correct network connections made

Copyright 2007 Mark Post

Wear and Tear on People (2)

- Getting “final” go ahead to actually build a particular set of systems on Friday afternoons
- System builds and hardening taking hours instead of minutes due to unreliability of provisioning software workaround
- Hardware failures
- Verifying servers were assembled with the correct combination of hardware
- Confusion about what servers were named what, and what rack they were in. (Inventory control.)
- Having to rebuild systems multiple times due to client requested changes

Copyright 2007 Mark Post

Wear and Tear on People (3)

- Assembling servers
- Racking servers
- Labeling and relabeling servers
- Moving parts around between servers
- Running and re-running Ethernet cables
- Testing Ethernet connections to switches, etc.
- Configuring and re-configuring network switches
- Running and re-running fibre cables for SAN
- Testing fibre connections
- Configuring and re-configuring firewalls

Copyright 2007 Mark Post

Just How Busy Were Those Systems?

- The operating systems were taking up about 16GB of the 72GB (x2 for mirroring) disk drives
  - This was only because we drastically over allocated space
  - Swap partitions took up another 4GB
- Once the systems were in production, I looked at several 24-hour periods during the week.
  - One system was running around 10-15%
  - The other 50 systems were running < 1%
  - None of them were doing any paging because of the large RAM sizes.
  - Almost all of them had no page space usage.

Copyright 2007 Mark Post

The Mainframe Alternative

- z9 EC (2094-S18)
  - 1 CP (Oddity due to IBM Pricing Policy)
  - 2 IFLs
  - 256GB RAM (16x16GB cards, requires 2 books)
  - 2 FICON Cards = 8 Ports/Channels
  - 2 FCP Cards = 8 Ports/Channels
  - 2 OSA-Express2 1000Base-T

- DASD (DS8100)
  - 4 TB ECKD

  There was also a lower-cost DS6800 option, which I chose not to take.

Copyright 2007 Mark Post

The Mainframe Alternative

- Other Hardware
  - Hardware support 24x7 – 2 (The z9 and DASD)
  - Ethernet Switch Ports – 4
  - Ethernet cables – 4
  - Fiber cables for FICON DASD – 8
  - Fiber cables for SCSI over FCP - 8

- Again, this isn't including switches, routers, etc.

Copyright 2007 Mark Post

The Mainframe Alternative

- The totally unrealistic “reference price” hardware costs for 3 years, including maintenance
  $5,295,092
- The more realistic price:
  $3,575,096

- Again, this isn't including switches, routers, etc.

  The IBM business partner who helped me didn't want to send me the reference price:
  “I really hate to give you these prices.. They are so out of line!”

Copyright 2007 Mark Post

The Mainframe Alternative

Software licensing, including support and maintenance

- 2 – z/VM
- 2 – Linux
- 2 - Oracle
- 2 - Dirmaint
- 2 - Performance Tool Kit (substitute your own favorite here)

- Software costs for 3 years:
  $309,080

Copyright 2007 Mark Post

The Mainframe Alternative

- Power
  - 6,300 * 24 * 365 / 1000 * 0.0936 = $5,165.60
  - Power for 3 years = $15,496.80
- Cooling
  - $15,496.80 * 0.7 = $10,847.76 for cooling over 3 years
- Floor Space Costs
  - z9 BC: 58.7 * 220 = $12,914 * 3 years = $38,742

Copyright 2007 Mark Post

The Mainframe Alternative

Category	Unrealistic	More Realistic
• Hardware	- $5,295,092	$3,575,096
• Software	- $309,080	$309,080
• Power and Cooling	- $26,345	$26,345
• Floor Space	- $38,742	$38,742
• Mainframe Total	- $5,669,259	$3,949,263
• Midrange Total	- $6,547,610	

Copyright 2007 Mark Post

The Mainframe Alternative – 5 Years

Category	Unrealistic	More Realistic
• Hardware	- $5,555,492	$3,835,496
• Software	- $458,960	$458,960
• Power and Cooling	- $43,908	$43,908
• Floor Space	- $64,570	$64,570
• Mainframe Total	- $6,122,930	$4,402,934
• Midrange Total	- $8,204,091	

Copyright 2007 Mark Post

Questions?

Copyright 2007 Mark Post