Server Consolidation with Linux for zSeries

Scott Courtney
Sine Nomine Associates
scourtney@sinenomine.net

Perceptions About Mainframes in a Distributed World

- System/390 had its day, but…
  - Our new hires don't know how to run it.
  - It's….(gasp)…old!
  - It's expensive.
  - What's an EBCDIC and what's a 3270?
  - It can't run "modern" software
  - It isn't an Intel PC.

Contents

- A Brief History of Linux on zSeries
- Sine Nomine Associates' customer experiences
- Challenges of zSeries Linux deployment

Genesis of a Dream

- VM is nice, but…
  - Our new hires don't know how to run it!
  - It's….(gasp)…old!
  - What's an EBCDIC?
  - Can I run Apache on that thing?
  - It won't run on my PC!

A Brief History of zSeries Linux

How did this happen?
Why is it important?

Genesis of a Dream

- OS/390 is nice, but…
  - Our new hires don't know how to run it!
  - It's….(gasp)…old!
  - It smells of Hollerith!
  - What's an EBCDIC?
  - Gee, it's awfully expensive!
  - It won't run on my PC!

Genesis of a Dream

- UNIX System Services are nice, but…
  - You want *how much* for the C compiler?!
  - It smells of 3270!
  - It's just <sigh> not quite UNIX.
  - What's an EBCDIC?

Linux on S/390 Delivers…

- Not emulation – it's the real McCoy!
- Smells of UNIX
- ASCII and byte streams
- Open Source: Free Speech, Free Beer
- Runs in LPAR, on native metal, or under VM
- S/390 architecture in the mainstream kernel
- Up-to-date kernel, utilities, and applications

Genesis of a Dream

- I'm dreaming of…
  - Using all my UNIX skills on Big Iron
  - Running Apache on my mainframe
  - Leveraging Open Source in my business
  - Getting rid of all those #&$* little servers
  - Gluing my mainframe apps to the intranet
  - Squashing Microsoft like a bug!

…But on VM, it *REALLY* delivers

- Try it risk-free (forever)!
- Test Plan Charlie: Buy one, get 40,000 free
- You call that a LAN? Now *this* is a LAN!
- Which version, which version? Run 'em both!
- It's good to be the root!
- Reliable, scalable, and secure – just like VM!
- Microsoft: You can't touch this!

The Dream Becomes Reality

- Linas Vepstas : Bigfoot
- IBM Boeblingen builds "official/unofficial" port
- Marist College distribution
- May, 2000: "It's official!"
- LinuxWorldExpo 2001: That's "B" as in "Billion"

…But on VM, it *REALLY* Delivers

- Risk-free test and pilot deployment
- Test Plan Charlie: Buy one, get 40,000 free
- Ultrafast internal networking
  - Hipersockets, IUCV, CTC, and VCTC
- Multiple simultaneous environments/versions
- Distributed empowerment, centralized security
- Reliable, scalable, and secure

Virtualizing the Virtual: It's Turtles All the Way Down!

VM can virtualize virtual hardware, *n* levels deep, and can run *thousands* of images at once.

- Test Plan Charlie: 41,400 Linux images in an LPAR on a G5
- Test Plan Omega: 97,943 Linux images on a ZZ7, 12-way@160 MIPS each, 16G RAM
- Thornton: Linux/390 under Hercules under Linux/390 under VM … whoda thunk it?

These specific demonstrations are "lab queens" but the practical value of this capability is very, very real!

Linux Used by…

- IBM
- Oracle (development)
- SAP (develops and deploys on Linux)
- Bynari Systems
- PeopleSoft
- Winnebago
- InfoCrossing
- Siebel
- Lawson
- Veritas
- Reuters
- Rational Software
- Computer Associates
- Telia
- Korean Air

"Why Would We Want to Do This?"

- Server consolidation
- Common operating system across all architectures
- New life for big iron "dinosaurs"
- Integrating platform for OS/390 legacy apps
- Rapid deployment of Open Source technology
- Because we **CAN**!

Sine Nomine Associates' Customer Experiences

- Customer 1: Insurance Industry
- Customer 2: Financial Services
- Customer 3: Telecommunications
- Customer 4: Higher Education

This is a partial listing.

Linux/390: Enterprise-Ready

- CERT Advisories equivalent to commercial UNIX vendors
- NSA "Trusted Linux" and Bastille Linux
- Peer review of source code enhances security
- Journaling file systems and LVM
- IBM Java Technology with JIT compiler
- DB2, Domino, MQ Series … middleware is here
- 85% of US colleges teaching Linux courses

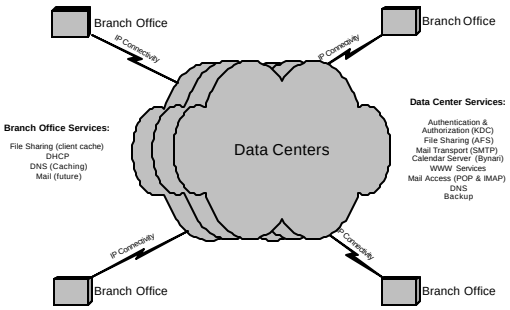
Customer 1: Insurance Industry

- Project to replace:
  - Windows with Linux (get off the Microsoft upgrade train)
  - MS services with Open Source equivalents (add flexibility, control)
  - Core Intel hardware with IBM mainframe (for reliability, scalability, manageability, disaster recovery)

Customer 1: Project Goals

- Horizontally scalable:
  - Modular distributed file and print servers
  - Incoming/outgoing email capability
- Capable of supporting 50,000+ users
- Straightforward to maintain and expand
- Low licensing costs

Customer 1: High-Level Overview



Customer 1: Project Tactics

- Samba Open Source file and print services replace Windows file/print servers (and WINS)
- Bynari Insight Server replaces Microsoft Exchange
  - Compatible calendaring
  - Mail delivery via POP and IMAP
- Open Source dhcpd replaces Microsoft DHCP
- Apache replaces IIS

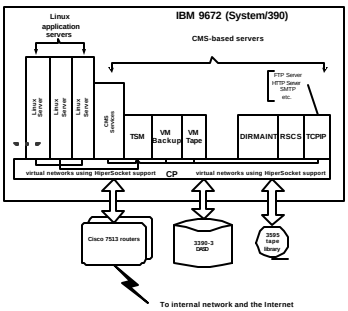
Customer 1: Architecture

- Uses “hub and spokes” model
- Data Center is “hub”
  - Provides all centralized services
  - Security, file sharing, mail, calendar, directory
- Branch offices are “spokes”
  - Provide local services
  - File sharing (cache), DNS (cache), DHCP

Customer 1: Infrastructure

- Authentication & authorization: MIT Kerberos
- Distributed file system: OpenAFS
- Mail/calendar server Bynari Insight Server
- Directory services: OpenLDAP
- Domain Name Service: ISC BIND

Customer 1: Data Center



Customer 1: Scalability Tactics

- Centralized (mainframe) services provided by virtual network of z/VM guests
  - Connected via HiperSockets guest LAN
- Within data center
  - Add more guests for additional task load
  - Round-robin DNS balances load between guests
- At branch offices
  - Add more spoke machines for additional offices
  - Also if needed to spread load at a single office

Customer 1: Disaster Recovery

- Mainframe facilities enable on-the-fly backups
- Can recover entire system (all servers) from single set of tapes
- Can ***verify*** DR capability at "hotsite"
  - Practice makes perfect…!

Customer 1: Reliability, Availability, Serviceability

- Initially: round-robin DNS
  - Provides fairly high availability
  - Crude load-balancing, dual-pathed routers guard against network failure
- Ultimately: multiple discrete implementations in multiple data centers
  - Full failover capability
- AFS volume replication provides file storage availability
- IBM hardware MTBF: 50+ years

Customer 2: Financial Services

- Feasibility study / test implementation
- Existing application newly in production
- Current platform is UNIX and NT/IIS
- Web-based 4-tier implementation in Java
- BEA: WebLogic Server
- Task: Port application, evaluate performance

Customer 1: Cost Control

- Open Source packages used as much as possible
- Very little commercial software:
  - z/VM 4.2
  - Bynari Insight Server
    - Includes Open Source packages: Exim, OpenLDAP, Cyrus IMAP/POP server
  - Optional backup from z/VM side
- Predictable, controllable software license costs

Customer 2: Application Port

- Three people, six hours, one working app!
- Zero source code changes (config files only)
- Virtual network setup for app-level clustering
- BEA: WebLogic Server is supported on zSeries Linux
- One z/VM instance, two zSeries Linux instances, four IFL processors – all added to an existing mainframe with zero downtime

Customer 2: Performance Testing

- Gold standard: Match their existing production system's response time
- zSeries with 1 CPU exceeded performance target by approximately 3X
- zSeries with 2 CPUs was too fast for their test environment to saturate, but ran at least 6.5X their normal production load
- External NT systems and LAN were the bottleneck

Overview of ISP/IDC Requirements

- Internet Service Providers (ISP)s and Internet-oriented Data Centers (IDCs) have similar requirements:
  - standard open-source applications (sendmail, bind, UCB POP3, UW IMAP, WUFTPD, INN, etc)
  - primarily Unix-based environment
  - IP-centric (some Novell, some NETBIOS)

Customer 2: Test Notes

- Over 30 hours of intense high-load testing, zero failures of zSeries hardware, z/VM, or Linux
- Even at saturation load on one CPU, no software failures or crashes
- No application or BEA tuning for Linux platform (same parameters as on UNIX)
- Added and removed processors dynamically without rebooting virtual machines

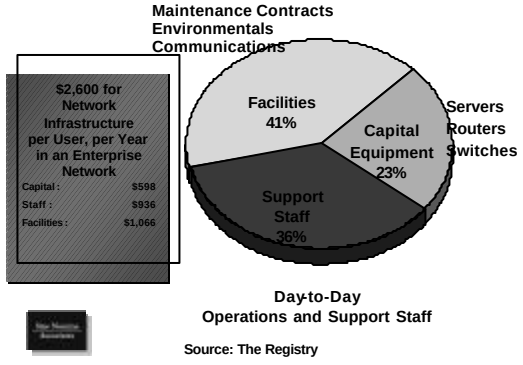
Overview of ISP/IDC Requirements

- Primary differentiator is scalability and TCO:
  - IDC requires substantially larger scalability (average 5000+ systems for industrial scale)
  - Target TCO computation for traditional solution: $1500/sq ft/month
    - Total operational cost, including staff, environmentals, operation and management software, etc.

Customer 3: Telecommunications

- New market initiative into ISP/IDC by existing telecomm carrier
- Needed massive scalability
- Multi-step pilot and feasibility testing
- Rollout to production
- Initial expectation of 250 customers
- Rapid growth after initial rollout

Network Cost of Ownership

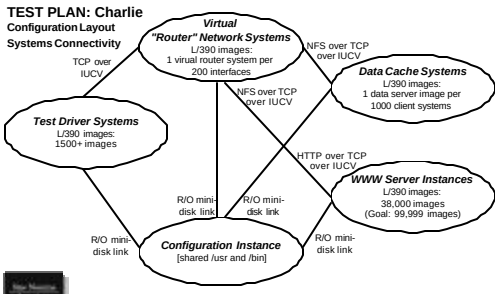


Source: The Registry

Overview of ISP/IDC Requirements

- Secondary differentiator is time to market (TTM):
  - average for discrete machines = 7 days from payment to delivery
  - high-volume sources (Exodus, AboveNet) avg 4-5 days to delivery
- Most business ISP/IDC customers expect dedicated servers to guarantee SLAs

Test Plan: Charlie



System Count: Discrete Solution

- Estimating 2 Sun UE2 class systems for DNS; 1 Sun UE1000 system for INN due to I/O requirements
  - System requirement replicated for each customer.
  - Implies 2 RU per UE2; 4 RU per UE1000 + disk array (2-4 RU)
  - 3 systems per customer: 750 machines!

Lessons Learned

- VM is critical to large scale Linux for System/390 scalability
  - 15 LPARs does not offer sufficient cost/benefit to make the case for Linux on S/390 iron
  - Loss of VM resource management and error recovery substantially complicates system management
- Lack of VM on other platforms is a major differentiator in favor of zSeries

Support Infrastructure: Discrete Solution

- 3 physical LAN ports
- 1/3 of a rack
- VLAN configuration
- Cabling and cabling management
- IP address allocation & routing policy
- Tivoli management agent license
- Tivoli TSM backup client license
- etc., etc., etc.

Customer 3 Outcome

- Customer is now creating between 15 and 30 virtual systems per day on a new 9672
- Clients of the service are pleased with the uptime and low cost
- Virtual system deployment completely automated (integrated into WWW front-end and back-end business systems)

Customer 3: Why?

- TCO for traditional solution: $1500 per square foot per month
- Averages:
  - 3500-7000 discrete systems
  - 15,000-20,000 square feet
  - 3500-7000 network cables and LAN ports at $150/port
  - 3500-7000 power cables
  - Time to market: 4-7 days

Political Challenges

- Challenge: How to sell Linux and Open Source idea to senior management?
- zSeries Linux answers:
  - Deploy alongside existing mainframe software, without interruption to production
  - Small project first, often infrastructural in nature
  - z/VM is key to flexibility of pilot environment

Customer 4: Higher Education

- Small college with cutting-edge CompSci program
- Students get z/VM user accounts with Linux installed
- System administrators can replicate a template for a new student account in seconds
- Student gets "root" privilege on his/her Linux instance, but is still just an ordinary user to z/VM
- 85% of colleges now teaching Linux, but this college is doing so with miniscule incremental cost per student
- Scenario is applicable also to web hosting providers who cater to hundreds or thousands of small companies

Political Challenges (cont'd)

- zSeries Linux answers (continued):
  - Personal case history: Pairing a VM guru with a Linux wiz
  - Empowerment of open systems managers via access to larger-scale resources, new data sources
  - New levels of integration between legacy data and new intranet/internet/extranet applications

Challenges for zSeries Linux Deployment

- Political Challenges
- Technical Challenges
- Project Management Challenges

Technical Challenges

- Backup/Recovery Solutions
  - Can be done within Linux, but not always best to do so
- Performance Monitoring Instrumentation
- Configuration Management
- Security Management
- Software Replication
  - Shared read-only filesystem is one option, but more flexible options are needed

Project Planning Challenges

- What type of project first?
- Implementation planning
  - Who is involved, and at what point in time?
  - Need collaboration between mainframe and UNIX/Linux personnel
  - System automation tools from z/VM environment applied to Linux instances
  - Vertical and horizontal scalability planning

Contact Information

Scott D. Courtney
Senior Engineer
Sine Nomine Associates
43596 Blacksmith Square
Ashburn VA 20147-4606
scourtney@sinenomine.net
http://sinenomine.net/

Resources on the Web

- Marist Distribution
  www.linux390.org
- Linux/390 Project Home Page
  www.linuxvm.org
- IBM Linux zSeries Home
  www.ibm.com/servers/eserver/zseries/os/linux/
- IBM VM Linux Resources
  www.vm.ibm.com/linux

More Resources on the Web

- Official Linux Home Page
  www.linux.org
- IBM Linux Community Development System
  www.ibm.com/servers/eserver/zseries/os/linux/lcds/
- "Dream Machine" Article Online
  www.linuxplanet.com/linuxplanet/reports/1532/
- Sine Nomine Associates
  www.sinenomine.net
  Look in the "Useful Links" for more S/390 Linux resources.