

L74

Penguins Board the Stagecoach for the Linux Frontier: A User Experience with Linux on z Series

Marcy Cortes, Wells Fargo.com

IBM System z Expo

September 17-21, 2007
San Antonio, TX



Disclaimer

- **The information and opinions found herein are for informational purposes only and not necessarily those of Wells Fargo Bank and should not be considered an endorsement.**

Session Topics

- **Who we are**
- **Some history**
- **Our environment (hw and sw)**
- **Our applications**
- **Operational considerations**
- **Futures**
- **Hints and tips**
- **Summary**

About Wells Fargo Bank

- **Diversified financial services company**
- **Retail banking (23 western states), Mortgage in all 50**
- **23 Million customers**
- **~6000 stores (branches) 6800 ATM locations (13,000 ATM)**
- **10 Million Debit cards issued (2nd largest issuer US)**
- **5th largest US bank**
- **158,000 employees**
- **Headquarters in San Francisco, but very decentralized**

About me

- **Team lead for z/VM and Linux on z Series at WF**
- **AVP, Operating Systems Engineer**
- **Enterprise Hosting Services group**
- **At WF since 1994**
- **7 team members now – me plus:**
 - 2 new hires (VM backgrounds)
 - 2 z/OS (ex z/OS) about 1.5 years ago
 - 1 recent (8 mos) transfer from Unix/Linux patching/quality team
 - 1 50/50 VM and Linux (out on disability)
- **Install/support/engineer z/Linux and z/VM builds**

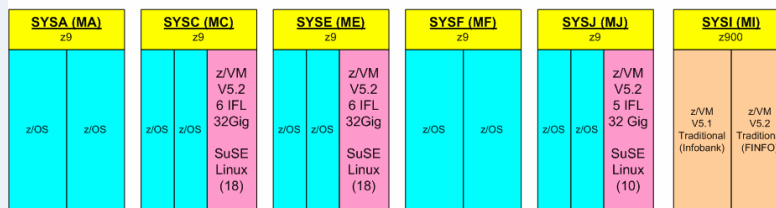
VM at Wells Fargo

- **Traditional VM**
 - 2 z/VM systems on 1 z900
 - Traditional CMS applications – mainly reporting
 - IBI's FOCUS, Connect:Direct, Rexx, Pipes
 - Installed in 1983
 - Participant in VM ESPs since V2.1.0
- **Linux hosting**
 - 4 z/VM systems
 - 100 test and development Linux servers
 - 50 production Linux servers
 - 19 total IFL on z9 EC machines

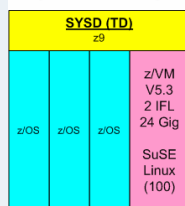
Hardware Environment

- **Production Data Center – Minneapolis MN**
 - 6 systems (5 z9 EC, 1 z900)
 - Over 80,000 MIPs (9300 VM)
 - 17 IFL engines on 3 LPARs
 - 2 engine z900 for traditional z/VM work
- **HA/Test/Dev Data Center – Tempe, AZ**
 - 2 z990s for z/OS HA
 - 1 z9 EC for z/OS and z/Linux test and development
 - 13,000 MIPs
- **BCP (disaster recovery) Data Center in house – Tempe AZ**
 - 5 z9 EC activated by Capacity Backup Up (CBU)
- **Servers – at least 6500 in data centers, 15000 total**

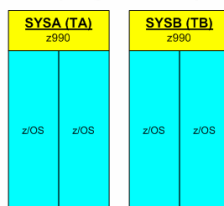
Minneapolis Production



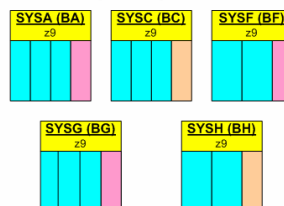
Tempe, AZ (Dev/Test)



Tempe, AZ (High Avail)



Tempe, AZ (BCP)



Production z/Linux environment

- 3 LPARs on z9 EC machines
- z/VM 5.2
- 32Gig per machine – split 24G central – 8G xstore
- 50 servers
- Applications

DB2 Connect

TMS

FAI

OpsCons

EPS

IDC

WFHM Lien Subordination

WFHM RMD

HHP

PhoneBank

Levanta

ETM

Linux Test/Dev Environment

- z9 EC in Tempe, AZ
- 24G (split 20/4) and just 2 IFL engines – heavily loaded
- z/VM 5.3 (as of 7/21)
- Approx 100 servers – variety of things
 - 50/50 SLES8 vs SLES9x
 - Apache/Tomcat/MySQL
 - DB2 Connect
 - WebSphere ND 5 and V6 (50 of these are WAS)
 - Some Oracle 9i
 - Test and Training servers
 - Bunch of miscellaneous things

z/Linux History

- **Started in December 2001**
- **HW folks wanted to POC before customers asked for it**
- **Moved internal website from old Solaris to SLES7 Apache**
- **Chose SuSE just because they seemed a little ahead of RedHat**
- **Installed as a guest on existing IBM 9672 machine (G5)**
- **Began using VM Guest Lan in Feb 2002 (z/VM 4.2)**
- **Worked well – moved website in about a week**

z/Linux History ... 2002

- **Initial hurdles**
 - Security plan
 - Resistance from unix groups (stealth project)
- **Added server for them to play with**
- **Gave several informational presentations internally**
- **PhoneBank app (call center) added in Fall 2002**
 - MySQL and Perl (6 servers)
 - Backup (disaster recovery) for their Intel servers

z/Linux History ... 2003

- Added several more departmental Apache servers in spring
- Stopped deploying servers mid 2003 pending:
 - Proper HW
 - zSeries – not a G5 9672 box
 - Modern DASD
 - Off of VM production machine
 - Deployment and maintenance product
 - Ourselves out of z/OS group

Fish or cut bait time!!

z/Linux History ... a turning point

- Management buy-in thanks to reorg and IBM influence with upper management
- VM Systems Programming team moved into “Open Systems” group and out of MVS ... and then to “Emerging Technologies”
- Purchased the Levanta tool
- Installed 1st IFL on z900 in California Nov 2003
 - Appeared as part of a “deal”
 - 20 or so servers moved off 9672 to z900

z/Linux History ... 2004

- **Began adding servers again**
 - First Apache
 - Then big things like WebSphere AS and MQ Series
- **Deployed Levanta**
- **DB2 Connect project given OK to go production**
- **TMS Application (WebSphere AS + MQ Series) started**
- **2 additional z990 IFL LPARs deployed in MN data center**
 - Production separate from test at data center level
 - Test/Dev remained in CA and later moved to AZ in 2005
- **Moved prod servers (6!) to MN IFLs**
 - (TMS, DB2Connect, Phonebank)

z/Linux History... 2005

- **Continued to deploy servers**
- **Moved test/dev z/Linux workload from CA to new z990 in Tempe AZ (July)**
 - New DS8000 FICON attached DASD
 - 8 Gig central storage
- **Added 2 more IFL engines to MN (1 each LPAR) for TMS**
- **Added more memory and another IFL engine for test/dev**
- **Made part of SPO (Standard Product Offering) for the “On Demand” initiative (ICOD – Infrastructure Capacity on Demand)**

z/Linux History ... 2006

- **New business applications**
 - EPS
 - OpsCons
 - Mortgage RMD and Lien Subordination
- **Significantly grew production environment to 17 IFL**
- **All z990's upgraded to z9 EC**
- **Added new staff**
 - 2 z/OS sys progs added to the group
- **Built 3rd production LPAR**
- **Began utilizing z/OS to do prod Linux backups**

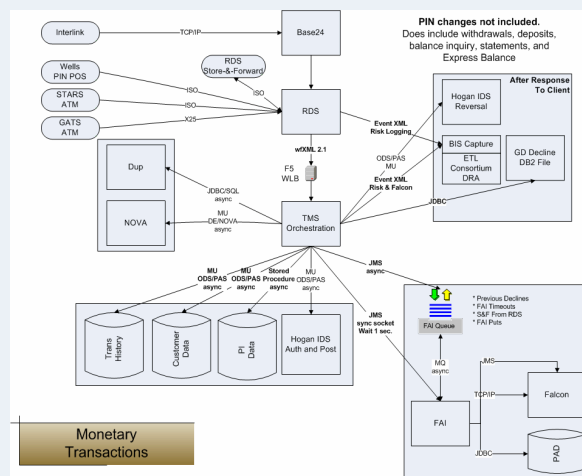
TMS Application

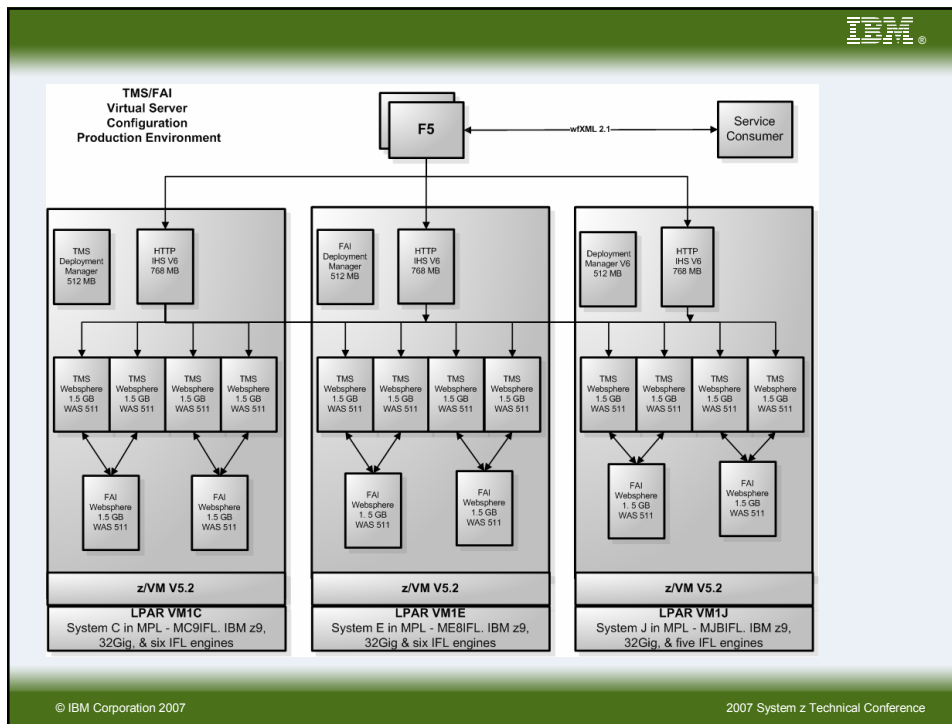
- **SOA enable existing z/OS Systems of Record (SORs)**
- **Present monetary transactions as web services (debits/credits/inquiries)**
- **Foundation for future initiatives - wfXML**
- **Contains FRAUD detection component called FAI**
 - Interfaces with Fair Isaac product FALCON on AIX (via MQ)
 - Rules based real time fraud detection
- **In use by many clients (5 Million trans a day)**
 - Bank Official Check
 - Cash Vault
 - Debit Card Point of Sale
 - ATM

TMS Application.. continued

- In house written Java routine
- WebSphere AppServer (WAS) 5.1.1 / MQ Series 5.3 / SLES 8
- WAS 6.1 and MQ 6 in test/dev now on SLES9 64bit
- Uses JDBC type 4, MU (CICS), and MQ Series
- SSL Option for clients
 - IBM Crypto Express 2 cards in Accelerator mode (C2EXA)
 - IBM HTTP Server with mod_ssl for crypto card
 - Provides nearly equivalent performance to non-SSL
- Posts debits/credits to z/OS SORs
- Runs on WAS Cluster over 3 LPARS load balanced by F5 BigIP
- Critical application with strict performance needs (\$\$\$ flowing through)

TMS Architecture





TMS application – why z/Linux

- **Why z/VM vs. Linux (Intel) vs. z/OS**
 - z/OS licensing costs
 - Compatibility with distributed systems
 - Reliability of mainframe hardware
 - Time to market
 - Proximity to z/OS and SORs
 - WAS for z/OS not integrated into our complex z/OS env (TopSecret, many LPARs, lots of non-IBM sw)
- **Java Based**
 - Developers use RedHat/Tomcat on Intel and easily moved to z/Linux with no app changes needed. Did have to set default code set to UFT-8 because of EBCIDIC data sent to z/OS

TMS / IBM Design Review

- Participated in the TMS/Falcon Design Review
- Poughkeepsie, NY – Aug 2005
- 10 WF employees + IBM experts
- Included both WebSphere app architecture as well as infrastructure and platform discussions
- Validated choice of WAS on z/Linux vs. z/OS
- J2EE app could easily move to another platform if necessary
- Identified weaknesses in the environment

TMS Challenges

- Must run high volumes of “fat” transactions (peaking at 120 TPS now, 500 TPS by 2009)
- Transactions that don’t make 2 seconds are “stand-in” by the ATM support systems
- Used more CPU than originally sized for (but requirements changed as well)
- Recoding parts (with IBM help) lowered CPU quite a bit
- Extensive testing is done with Load Runner to evaluate app and SW changes
- On Intel, perhaps inefficiencies would have been discovered much later down the line?

Recent TMS challenges

- **Peak period standins (Friday at 4-6pm)**
- **Extreme CPU “sensitivity”**
 - Other apps using more than normal CPU (batch type job or loop)
 - TMS standins occurred
 - TMS CPU utilization soared across all 3 LPARs
 - Standin’s result in “do overs” causing snowball effect
- **Spiky CPU utilization**
- **TCPIP service machine experiencing too much CPU wait**
- **Sent Velocity a day’s ESAMAP report**
 - Pointed out high fast path number on ESACPUA report
 - Indicated global spin lock contention
- **IBM also got data and came up with same recommendation**
- **Taking virtual CPUs from 2 to 1 on all TMS and FAI servers solved the problem**

WFHM RMD / Lien Subordination Apps

- **Mortgage group**
- **RMS – Relationship Management Dashboard**
- **Web app for small group of mortgage staff**
- **Lien Subordination – web app used to record secondary mortgages**
- **Low volume – 2-3% of a single IFL at peak, no data storage**
- **Would have had to purchase 6 “standard size” Intel servers (test, dev, prod, disaster recovery) at > \$100K**
- **Small ones much more cost effective on z/Linux vs. Intel**

EPS/OpsCons/ASI/IDC/ETM Apps

- All 6 belong to Corporate Operations LoB
- Some parts from Vendor Carreker, some locally written
 - Vendor had no zSeries Linux experience
 - Supported it on z because we asked
 - Came on site, quickly checked it out (needed to do nothing to app)
- J2EE
- Using WAS 6.0 and 6.1
- MQ Series
- Sterling Commerce Connect:Direct Secure+ for file transfer (aka NDM)
- Initially chose z/Linux because of ability to get HW quickly and cheaply for proof of concept
- Corporate LoB continues to choose z/Linux first
- All related to image clearing in order to support Check 21 image presentation

EO Image Conversion

- Custom Java code on WAS 6.1
- Used to prepare images from acquisitions for loading into main image repository
- Images come in on NAS or SNAP appliances
- Move several hundred Gig to z/Linux, ran conversion
- Use last summer and now in use again for 3 acquisitions going on
- Saved purchase of a \$200K server for this purpose
- Short time period - can use off peak cycles – done on test/dev IFLs

z/Linux virtualization Advantages

- **Test server needs multiply (1 app can be 16+ servers)**
 - Unit testing
 - QA testing
 - Enterprise testing
 - Regression testing
 - SW version testing (aka WAS 6 and sles9x)
- **Reliable common driver code for all virtual servers**
- **HW platform changes/upgrades/CPU addition/microcode all possible without major disruption to Linux**
- **Every virtual server benefits from HW upgrades/ tech refreshes / HW currency – no waiting for the 3 (4) year upgrade cycle**
- **No cables!**
- **No flaky memory cards, NIC mismatches, CPU failures, real HW multipathing**
- **Significant power / floor space / cooling savings**
- **Decommissioned server resources just returned to pool (approx \$800 to dispose of old server)**

Operational items

- **Our support model**
- **Server Requests**
- **ITIL/ITSM CMDB**
- **Deployment/Provisioning/Patching**
- **Networking**
- **Security**
- **Performance Tuning / Capacity Planning**
- **Monitoring**
- **Backup / Disaster Recovery**

Our Support Model

- **Utilize existing Unix infrastructure where possible**
 - Monitoring
 - Security
- **Utilize existing mainframe HW folks**
- **Utilize existing mainframe Operations for VM side**
- **z/VM and z/Linux group**
 - Build servers
 - Engineer builds
 - Level 3 support for Linux
 - Build and maintain the 6 z/VM systems
 - Patches all z/Linux servers (o/s and related products)

Server Request Process

- **New application requests funding (Incremental / base)**
- **Sizing is prepared using sizing spreadsheet**
- **PTP (Project Tracking Portal) request is opened**
- **Security Risk Assessment submitted by requestor**
- **After security approval, IP admin will update DNS**
- **Servers is built by my group**
- **Turned over to developers (test) or operations (prod)**
- **Operational Readiness is completed (includes CMDB and BCP updates)**
- **Costs – no formal studies here (it's coming) but one app did sizings from both Intel and z (\$297K vs \$137K) for their app**

ITIL CMDB and z/Linux

- **Configuration Management Database (part of ITIL/ITSM)**
- **All virtual servers and real LPARs and real CPUs are in DB**
- **Includes for each server**
 - Support group names for on call
 - Host/LPAR relationship
 - Various info harvested from server scans (Linux level, sw installed, etc)
 - Applications using that server

Provisioning and Patching

- **Currently use Levanta product**
 - Builds new servers based on templates
 - Patches servers
 - Provides significant DASD sharing (savings)
- **Levanta on z Series is unsupported starting 10/31**
- **Investigating potential replacements**
 - IBM Director + z/VM Center + TPM
 - Mainstar
 - Roll your own

Virtual Networking

- Using primarily Guest LAN (we started this in VM 4.2!)
- Some servers on Layer 3 VSWITCH in Tempe test env
- No use of Hipersockets at this time
 - Network transmission time is negligible at the moment for our work load (small transaction sizes)
 - Which LPAR? VM on 3, z/OS on 10 over 5 boxes - would thwart some of z/OS HA workload balancing (F5)
- Dynamic routing via MPROUTE OSPF broadcasts Guest LAN to rest of network (makes BCP easy)
- VSWITCH move in production is coming (probably soon)
 - Save significant CPU (maybe ½ an IFL per LPAR)
 - More OSA and network switch ports required to implement (layer 2)

Security Authentication

- UnixSecure – WF developed authentication model + single signon
 - Quest VAS product
 - Implements kerberos and LDAP on Unix/Linux systems for true single signon “trusted realm” in Active Directory
 - Use PAM modules
 - Users authorized by groups
 - First login creates home directories and populates
 - WF modified OpenSSH (WFSSH) & Putty
 - Provides logged file transfers and validations of transferred files
 - Provides GSS-API for Kerberos protocol support (work with VAS)
 - Microsoft Active Directory
 - Provides LDAP directory
 - VAS schema extensions

Security Monitoring and Root Auth

- **Symark Powerbroker**
 - Agent Still not available Linux under zSeries
 - Used by Linux and Unix throughout the bank
 - Provides centralized root delegation authority
 - Use SUDO for now (security plan exception)
- **Symantec ESM agent**
 - Agent not available for Linux under zSeries
 - Used by Linux and Unix throughout the bank
 - Provides security monitoring (like a centrally managed tripwire)
 - Use simple SECHECK for now (another security plan exception)

Performance Tuning / Cap Planning

- **Provided by team of 2 on Capacity Planning side (relatively new at z/VM)**
- **Velocity ESAMON/MAP/TCP/WEB**
- **Velocity MICS interface to create a file for sending to z/OS for long term charting/trending**
- **Velocity alerts when prod systems exceed CPU targets**
- **Velocity normalizes Linux data to VM data**
 - DO not try TOP on busy WAS / MQ servers – so many processes that Top chases its tail and eats more CPU
 - CPU numbers read from within Linux are not accurate anyway (until we get SLES 10 SP1 installed)
- **Many things are not intuitive – get education for your staff**

Monitoring

- For VM, VM:Operator routes all messages to CommandPost server via IP socket
- For WAS, CA Wily Introscope is used by the WAS admin and app developers
- For general Linux, HP OVO is used
- For MQ Series, QPasa from MQ Software is used

Disaster Recover (BCP) and backup

- Backup Minneapolis prod VM system volumes with CA VM:Backup HIDRO over channel extended 3590 ATL tapes in Tempe
- Backup Linux volumes while servers are down (each LPAR done separately for availability reasons) with z/OS DFDSS to mirrored VTS (to Tempe)
- z/OS storage mgmt team restores DFDSS from VTS in Tempe every Monday morning
- z/VM team restores VM system volumes weekly as well
- For a disaster test, change IPs of restored Linux guests... for a real disaster, keep IPs and Guest LAN moves to Tempe via dynamic routing
- BCP tests done at least 6 times per year now
- File level backups can be done at app request via Netbackup agent

Challenges (opportunities?!)

- **Buy in accomplished in *some* business areas**
 - Politics and incorrect information continue to persist
 - My management putting together pitch to upper mgmt
 - \$ savings to be realized (power / cooling / sw)
 - Preallocate funds for true Capacity on Demand, esp for development
 - (you can turn on IFL easily, but who is paying the bill?)
 - Push other business lines by making this the default for new apps in triple A process (application architectural assessment)
 - Includes additional education on virtualization to many different functional areas
- **Thus far, new workload only, server consolidation has not been on the radar screen (probably won't be until end of server life cycle and tech refresh is needed)**

TO DOs!

- **VSWITCH implementation**
- **Finish SLES 8 to SLES 9x moves**
 - Requires replacement servers
 - Lots of issues with how that work is paid for
- **SAN DASD**
 - In proof of concept now in test system
 - New apps want big big file systems (example 2.5 TB)
- **Possible 4th and 5th LPARs in production (horizontal vs. vertical growth “issue”)**
- **Levanta replacement**
- **z/VM 5.3 upgrades in production (going very well in test so we will do this aggressively)**
- **Upgrade key apps (TMS and FAI to move along to SLES9x/WAS 6)**
 - Benchmarks showed up to 30% CPU savings
 - MQ 5.3 goes unsupported 9/30
- **Evaluate DASD mirroring options**
- **DB2 9 recently certified, several apps wanting to evaluate / test**

Other useful stuff

- Use the Linux-390 and IBMVM listserve lists – nice helpful folks
- See the IBM redbooks site for getting started info
- z/VM is very reliable
- z/Linux is too (very little use of service contract)
- Most servers don't need much resources (esp. things like Apache and test/dev is very unpredictable)
- Developers might have to make different choices
 - Bad code will shine like the sun in a shared environment
- Leverage processes and SW and organizations you already have in place for Unix
- Try it! But realize success probably won't happen until upper management buys in

Contact info

- My info
 - Marcy.D.Cortes@wellsfargo.com
 - 415-243-6343

