Success with Linux on System z at Nationwide Lessons Learned

Session 7275

Jim Vincent
Nationwide Insurance
This information is for sharing only and not an endorsement by Nationwide Insurance

August 2010

SHARE in Boston

Overview and Disclaimer

Disclaimer:

The content of this presentation is for information only and is not intended to be an endorsement by Nationwide Insurance. Each site is responsible for their own use of the concepts and examples presented.

First, a word from our announcer:

With a few exceptions, this is an overview! Where possible there are technical details you may be able to use. As you frequently hear when anyone asks for recommendations, "IT DEPENDS" is the answer and it applies here too. The information in this session is based on *our* experiences as long-time VM-ers building virtual Linux farms.

Interaction is good! Please ask questions whenever you want. We'll all get the most out of this session that way.

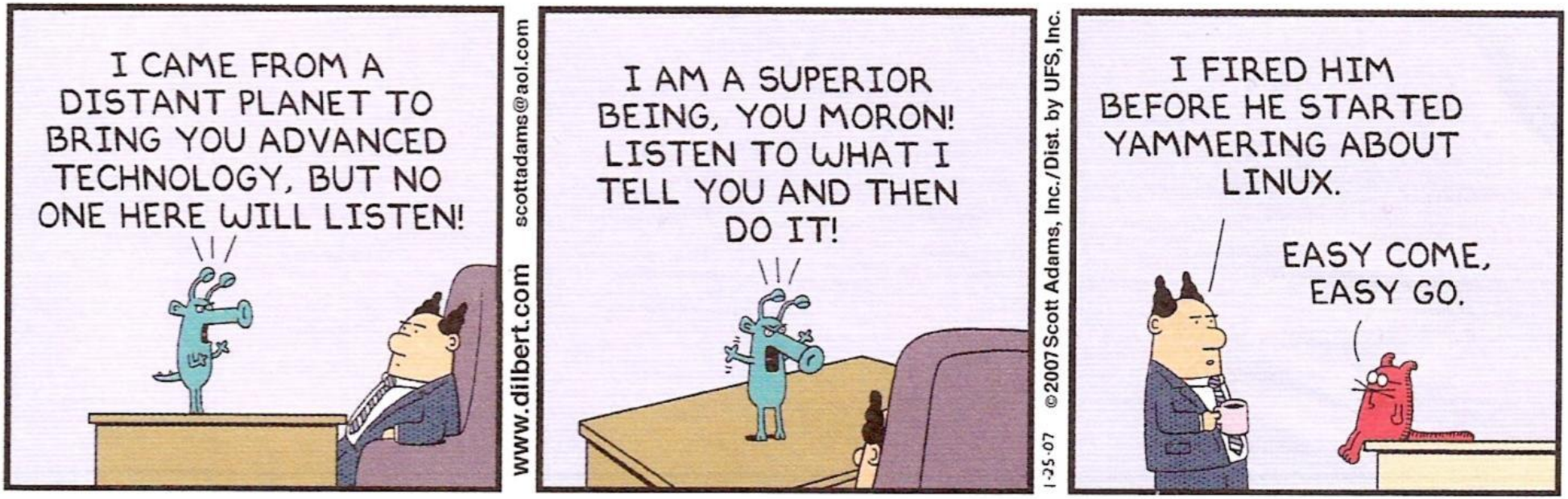
Topics

- *Very* Briefly:
  - Our Linux Decision History
  - Our Environment
- Expectations and Reality
- What's next?
- Conclusions

Our Linux Decision History

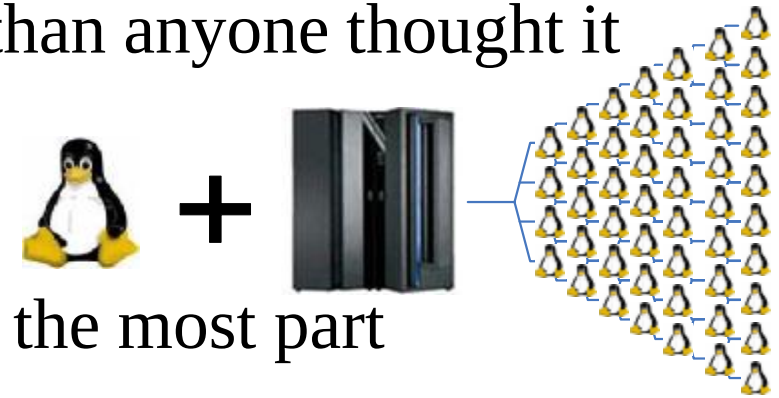
In the beginning, there was darkness…

- 2000 – Marist Distribution (based on Red Hat)
  - First offering of install lab at SHARE
  - Built one in-house to play with
  - Wrote up recommendation to management; Little interest or direction
- 2002 – SUSE 7
  - Basic demo of Apache and Samba
  - Wrote up recommendation to management; Little interest or direction
- 2004 – Red Hat
  - Intel, pSeries and zSeries pilots planned and started
  - zSeries work ceased within two weeks – “out of scope”


I CAME FROM A DISTANT PLANET TO BRING YOU ADVANCED TECHNOLOGY, BUT NO ONE HERE WILL LISTEN!
www.dilbert.com scottadams@aol.com
I AM A SUPERIOR BEING, YOU MORON! LISTEN TO WHAT I TELL YOU AND THEN DO IT!
1-25-07 © 2007 Scott Adams, Inc./Dist. by UFS, Inc.
I FIRED HIM BEFORE HE STARTED YAMMERING ABOUT LINUX.
EASY COME, EASY GO.

Our Linux Decision History

- 2005 – The fun began!
  - Because of emphasis on virtualization from CIO
- Proof of concept system grew FAST
- Anticipated having more than 1000 servers by now
  - Strangely (?) we don't!
  - But it is actually growing faster than anyone thought it would
  - Over 587 zLinux servers
  - Growing UP instead of OUT for the most part
  - Over **2700** JVMs!
    - Normally those would have been separate servers



Environment

- **Before**: One small LPAR on z900, then two z990s, replaced with two z9s and…
- **Today** – Two z10s dedicated to z/VM and Linux (~30K MIPS)
  - **21** IFL engines for test/dev and **32** IFL engines for production
  - 431GB memory for test, 318GB for production
    - 4 z/VM 5.4 LPARs on each
    - 1 additional test LPAR on development box for sandbox
    - 2 other LPARs on dev box for DR of traditional VM images
    - 9 total zLinux LPARs
    - total of 12 z/VM images to manage, and growing

What were we trying to solve?

- Server Proliferation
  - Space that previously was required to house a few mainframes is now mostly consumed by multitudes of all type of servers, network hardware, other support hardware
    - Sun, HP, multiple brands of Intel
    - Routers and switches
    - SAN, NAS, data warehouse, etc

What were we trying to solve?

- Provisioning
  - Many requirements for stand-alone server
    - Order and obtain hardware – several weeks
    - Physical install
    - Optional external disk subsystem configuration and connection
    - Network configuration and connection
    - OS load
    - Middle-ware load
    - Application load
  - Many hands and significant time
    - Usually would take several weeks (6-8 at least) or more before the customer would get the box

Vision and Expectations

- Physical space and environmental reduction
  - *A z990 IFL can support 10-30 (or more) virtual servers*
    - So a z990 can have up to 32 IFL engines so it *could* replace 300+ servers
    - **But in fact, we had 330+ large servers running on 15 IFLs between two z990s**
  - Significant savings in physical space, power, cooling
- Reduce network complexity
  - Small number of physical network connections (OSAs with VSWITCH) can support all of the virtual servers
    - contrast to every stand-alone server having 2 or more interfaces
- Quicker provisioning
  - Setting up new server can be as fast as your disk copy tool

What we learned… Experts?

"My definition of an expert in any field is a person who knows enough about what's really going on to be scared."
- PJ Plauger

- "Total Experts" - Do they really exist?
  - There are many people with varying levels of experience in specific areas
  - There are few (if any) who know enough about *everything*
  - Make friends with people who have knowledge in:
    - Mainframe disciplines and z/VM
    - Linux
    - Networks
  - Learn as much as possible about all of these areas
    - at least learn how to contact the right person when you need to!

What we learned…
"The z Team"

- Consider joining the Unix/Linux SA and z/VM Mainframe support groups together
- Initially the most challenging aspect of the project
  - I am a recovering "Mainframe Fundamentalist"
- The payoff was more than worth it - tremendous success
  - z/VM guys learned a lot about Linux technology
  - The Unix/Linux guys learned a lot about Mainframe technology and z/VM
  - Brainstorming and strategic planning is phenomenal

What we learned… Applications

- Virtualization is very foreign to a majority of people
  - Add multiple levels of virtualization and it gets worse
- Applications that are “challenging” will likely point outside their code as the problem
  - *Lack of understanding* increases probability of finger pointing
- Hand-hold your application folks!
  - Show them how things work in a virtual environment
  - Give them help in diagnosing **issues** and/or showing that the z environment is *not* the cause
- The best approach: Teach them about the ‘new’ environment *before* they get there!

What we learned… Management

- **Train your managers** to "speak the environment" correctly
  - "The z" **or** "z/VM never goes down. Period."
  - "…Linux on z/OS*…"
- Multi-layered virtualization is tough enough to grasp
  - Misleading statements can cause a lot of anguish
- If they don't want to learn or (ahem) can't, at least convince them to include you in discussions with others until they do

* z/OS \== Mainframe

What we learned… Chargeback

- We have to deal with chargeback models
- Give some good, realistic thought to how to set charges
  - Basic server charge?
  - Per GB fee?
  - CPU usage? An idle server: $5 a CPU hog: $1000
  - Growth (up and out) all factor in
- Adjust as needed – but not too much
  - Even “refunds” can be a pain to deal with

What we learned… Reality hits

“You’re on top of the wave again, but don’t get cocky!”

-Bob Rogers, when he was knighted at the 35th VM birthday – Aug 2007

- Thought that zLinux would change the perception of the mainframe finally
  - Ha! Don’t count on it.
- Keep looking at the solution(s) from a business perspective
  - The “cool” and “wow” factors don’t always pay the bills (or save money)
  - Remember to do what is best for *your business*
- Watch the “competition” at your shop – they are watching too
  - Even though they may not really “get it”
  - If you stumble, they will be there to help (kick you)

What we learned… Benchmarking/TCO

- Know your competition in your own shop (friend or foe??)
  - Understanding the solutions being compared to zLinux is important
- Keep everyone honest
  - Don't try to compare a brand new fully decked out rack/server with a moderate to heavily loaded z system
- You can twist the numbers any way you want
  - Run one Intel server, one app, one test – then EXTRAPOLATE for dozens of servers – yeah, that's the ticket!
    - Give me a break!!
- Publish and talk the FACTS!
  - Check out IBM's System z Myth vs Truth

What we learned… Hygiene

- Keeping your z/VM systems up-to-date may be a challenge
- As zLinux grows in use/size, be prepared
  - Contractual business requirements for **no outages** or unrealistic outage windows
  - Set up multiple LPARs, or better, multiple z's
  - **H**igh **A**vailability – does it really work correctly?
    - Convince your application areas to agree that HA is important and then make them fix it
- Do your math
  - Number of servers, time to cleanly shutdown, do z/VM maintenance, start up servers & apps, and test
  - Is there enough time to do it all in one shot?

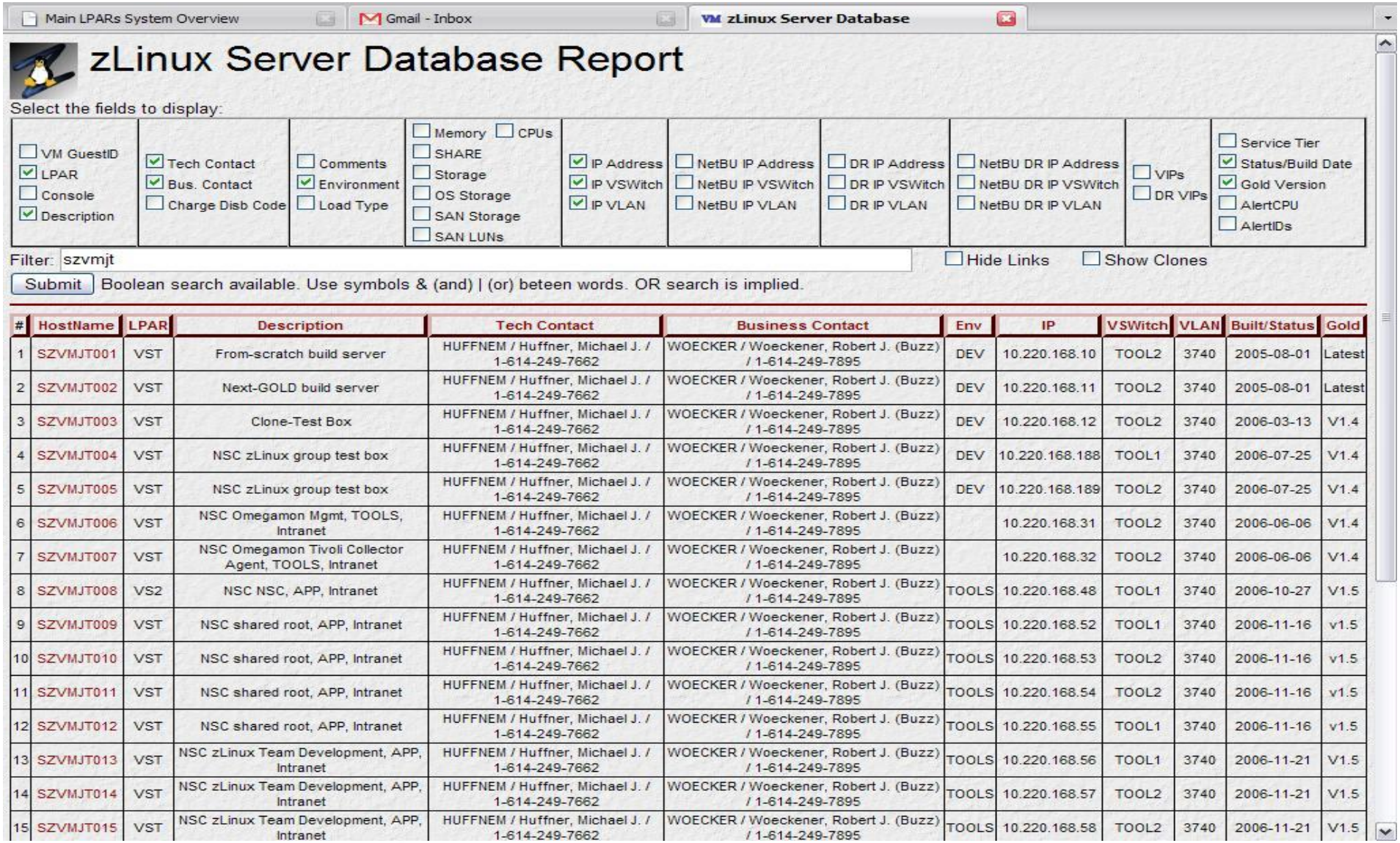
What we learned… ECKD or SAN?

- SAN volumes solved a lot of problems for storage
  - Helped for large storage (DASD) needs
- SAN volumes caused a lot of problems
  - Management of the SAN devices, WWPNs (real and virtual), LUNs, etc can be daunting
    - NPIV can make that even more of a challenge
  - Think Disaster Recovery – are you ready to re-map?
  - EDEV might be useful – still needs a little tweaking
- ECKD may actually be something to look at 'going back' to *Large volumes (> 100G) will be great

The Tool Belt

- As the zLinux project continued, it was fairly obvious that providing some “basic” tools would help everyone
- Access to details about the servers
- Access to easy-to-read CPU charts & information on resource usage by server
- Bottom line – put the data/view into the user’s hands
- You can build them yourselves or look to a vendor

The zLinux Database – details!



Main LPARs System Overview | Gmail - Inbox | zLinux Server Database

zLinux Server Database Report

Select the fields to display:

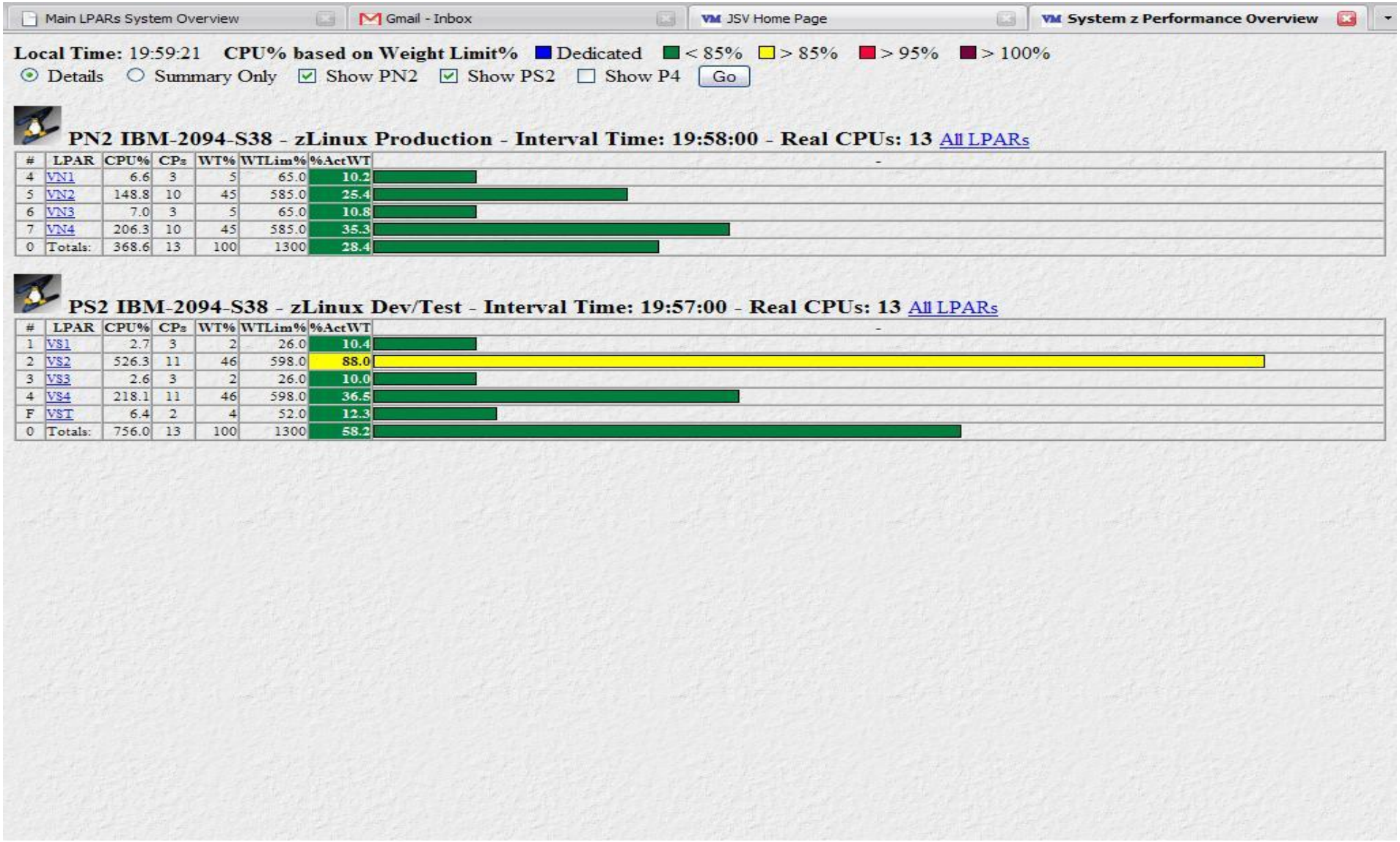
☐ VM GuestID ☑ LPAR ☐ Console ☑ Description	☑ Tech Contact ☑ Bus. Contact ☐ Charge Disb Code	☐ Comments ☑ Environment ☐ Load Type	☐ Memory ☐ CPUs ☐ SHARE ☐ Storage ☐ OS Storage ☐ SAN Storage ☐ SAN LUNs	☑ IP Address ☑ IP VSWitch ☑ IP VLAN	☐ NetBU IP Address ☐ NetBU IP VSWitch ☐ NetBU IP VLAN	☐ DR IP Address ☐ DR IP VSWitch ☐ DR IP VLAN	☐ NetBU DR IP Address ☐ NetBU DR IP VSWitch ☐ NetBU DR IP VLAN	☐ VIPs ☐ DR VIPs	☐ Service Tier ☑ Status/Build Date ☑ Gold Version ☐ AlertCPU ☐ AlertIDs

Filter: szvmjt ☐ Hide Links ☐ Show Clones

Submit Boolean search available. Use symbols & (and) | (or) beteen words. OR search is implied.

#	HostName	LPAR	Description	Tech Contact	Business Contact	Env	IP	VSWitch	VLAN	Built/Status	Gold
1	SZVMJT001	VST	From-scratch build server	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	DEV	10.220.168.10	TOOL2	3740	2005-08-01	Latest
2	SZVMJT002	VST	Next-GOLD build server	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	DEV	10.220.168.11	TOOL2	3740	2005-08-01	Latest
3	SZVMJT003	VST	Clone-Test Box	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	DEV	10.220.168.12	TOOL2	3740	2006-03-13	V1.4
4	SZVMJT004	VST	NSC zLinux group test box	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	DEV	10.220.168.188	TOOL1	3740	2006-07-25	V1.4
5	SZVMJT005	VST	NSC zLinux group test box	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	DEV	10.220.168.189	TOOL2	3740	2006-07-25	V1.4
6	SZVMJT006	VST	NSC Omegamon Mgmt, TOOLS, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895		10.220.168.31	TOOL2	3740	2006-06-06	V1.4
7	SZVMJT007	VST	NSC Omegamon Tivoli Collector Agent, TOOLS, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895		10.220.168.32	TOOL2	3740	2006-06-06	V1.4
8	SZVMJT008	VS2	NSC NSC, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.48	TOOL1	3740	2006-10-27	V1.5
9	SZVMJT009	VST	NSC shared root, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.52	TOOL1	3740	2006-11-16	v1.5
10	SZVMJT010	VST	NSC shared root, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.53	TOOL2	3740	2006-11-16	v1.5
11	SZVMJT011	VST	NSC shared root, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.54	TOOL2	3740	2006-11-16	v1.5
12	SZVMJT012	VST	NSC shared root, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.55	TOOL1	3740	2006-11-16	v1.5
13	SZVMJT013	VST	NSC zLinux Team Development, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.56	TOOL1	3740	2006-11-21	V1.5
14	SZVMJT014	VST	NSC zLinux Team Development, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.57	TOOL2	3740	2006-11-21	V1.5
15	SZVMJT015	VST	NSC zLinux Team Development, APP, Intranet	HUFFNEM / Huffner, Michael J. / 1-614-249-7662	WOECKER / Woeckener, Robert J. (Buzz) / 1-614-249-7895	TOOLS	10.220.168.58	TOOL2	3740	2006-11-21	V1.5

CPU overview



The big picture

Main LPARs System Overview - Local Time 19:58:19

VN2
z/VM CPU 2094 SERIAL 0FBFCD DATE 08/01/07 START 19:58:00 END 19:59:00

AVGPROC-015% 10 MODEL-2094 SERIAL-FBFCD
CPU-148% USER-143% PROB-135% SYST-5% OVHD-8% IDLE-849% PAGING-13/SEC

INQUEUE-00101 Q0-00003 Q1-00023 Q2-00012 Q3-00063
ELIG-00000 LOADING-00000

Host/uid:	CPU%	IO/Sec	Pg/Sec	Description
NZVMAS740	27.6	4.9	0.4	P&C NorthStar, App, PROD
NZVMAS807	10.5	3.2	2.2	NSC Tops, Was/MQ App, PROD
NZVMAS804	6.9	1.9	0.2	NSC Nationwide.Com, WAS App, PROD
NZVMAS738	6.4	1.3	0.6	P&C Horizon Claims, App, PROD
NZVMAS728	6.3	1.3	1.2	P&C Shared, WAS, PROD
NZVMAS734	5.6	1.7	0.5	P&C Northstar, App, PROD
NZVMDS702	4.7	3.7	0.1	P&C Agent Gateway, DB, PROD
NZVMAS702	4.0	2.0	0.1	P&C Vin / Vms Lookup / Nol, WAS App, PROD
NZVMAS718	3.9	1.3	1.0	P&C MarketLink, App, PROD
NZVMAS828	3.6	1.8	3.4	NSC Assoc Resiliency, WAS, PROD
NZVMWS824	3.5	2.3	0	NSC Assoc Resiliency, IHS, PROD

28.0%

VN4
z/VM CPU 2094 SERIAL 04BFCD DATE 08/01/07 START 19:58:01 END 19:59:01

AVGPROC-019% 10 MODEL-2094 SERIAL-4BFCD
CPU-194% USER-185% PROB-172% SYST-9% OVHD-12% IDLE-802% PAGING-29/SEC

INQUEUE-00105 Q0-00003 Q1-00028 Q2-00015 Q3-00058
ELIG-00000 LOADING-00000

Host/uid:	CPU%	IO/Sec	Pg/Sec	Description
NZVMAS741	23.3	46.2	0.2	P&C NorthStar, App MQ, PROD
NZVMAS813	18.4	12.0	0	Gates TOPS, WAS/MQ/Rpt, PROD
NZVMAS739	12.7	2.8	2.0	P&C NorthStar, App, PROD
NZVMAS808	8.5	2.3	0	NSC Tops, Was/MQ App, PROD
NZVMAS733	8.1	4.4	0.7	P&C Northstar, App, PROD
NZVMAS727	8.0	1.3	0.5	P&C Shared, WAS, PROD
NZVMAS833	7.1	0.8	0.4	NSC LMS, App, PROD
NZVMAS805	6.3	1.8	4.4	NSC Nationwide.Com, WAS App, PROD
NZVMAS729	6.1	5.4	0.5	P&C Shared, WAS, PROD
NZVMDS701	5.6	11.1	0.3	P&C Agent Gateway, DB, PROD
NZVMAS737	4.7	1.7	1.3	P&C Horizon Claims, App, PROD

VS2
z/VM CPU 2094 SERIAL 02BFBD DATE 08/01/07 START 19:56:00 END 19:57:00

AVGPROC-047% 11 MODEL-2094 SERIAL-2BFBD
CPU-518% USER-507% PROB-488% SYST-11% OVHD-19% IDLE-538% PAGING-65/SEC

INQUEUE-00232 Q0-00004 Q1-00076 Q2-00020 Q3-00132
ELIG-00000 LOADING-00000

Host/uid:	CPU%	IO/Sec	Pg/Sec	Description
SZVMDS001	101.4	6641.0	2.4	P&C Agent Gateway Data, PT
SZVMAS551	98.8	1.9	0	NSC Tops, WAS ND App, DEV
SZVMAS430	95.5	1.0	1.1	P&C NWBH, App, DEV
SZVMAS050	30.4	0.8	0.8	P&C Agent Gateway WAS App, IT
SZVMAS028	7.9	1.0	0.5	P&C, App, PT
SZVMAS026	5.8	3.3	0.6	P&C Customer Endpoint, App, PT
SZVMAS504	5.6	1.4	2.1	NSC Nationwide.com, App, ST
SZVMAS032	5.3	1.4	0.6	P&C Northstar, App, PT
SZVMAS044	4.7	1.9	0.7	P&C NorthStar, App, PT
SZVMDS002	4.7	7.8	1.4	P&C Agent Gateway Data, PT
SZVMAS550	4.4 C	1.2	0	13:49 NSC Nationwide.Com WAS App, IT

VS4
z/VM CPU 2094 SERIAL 04BFBD DATE 08/01/07 START 19:56:00 END 19:57:00

AVGPROC-019% 11 MODEL-2094 SERIAL-4BFBD
CPU-214% USER-206% PROB-198% SYST-7% OVHD-8% IDLE-881% PAGING-41/SEC

INQUEUE-00223 Q0-00004 Q1-00073 Q2-00029 Q3-00117
ELIG-00000 LOADING-00000

Host/uid:	CPU%	IO/Sec	Pg/Sec	Description
SZVMAS100	30.0	0.7	0.1	P&C Agent Gateway WAS App, ST
SZVMAS110	11.7 C	5.9	2.5	10:03 P&C Agent Gateway, WAS, ST
SZVMAS021	6.0	1.8	0.6	P&C MarketLink, App, PT
SZVMAS031	5.3	1.4	3.4	P&C Northstar, App, PT
SZVMAS043	5.3	1.4	0.4	P&C NorthStar, App, PT
SZVMDS007	5.3	5.9	1.2	P&C Shared, DB, PT
SZVMDS100	4.5	6.2	0.1	P&C Agent Gateway DB, ST
SZVMAS027	4.4	1.6	3.5	P&C Customer Endpoint, App, PT
SZVMAS029	4.2	1.0	0.4	P&C, App, PT
SZVMDS003	3.8	3.4	0.7	P&C SRS/ARAS, WAS, DB, PT
SZVMDS104	3.8	2.6	0.2	P&C Vin / Vms Lookup / Nol, DB, ST

58.2%

So, where are we now?

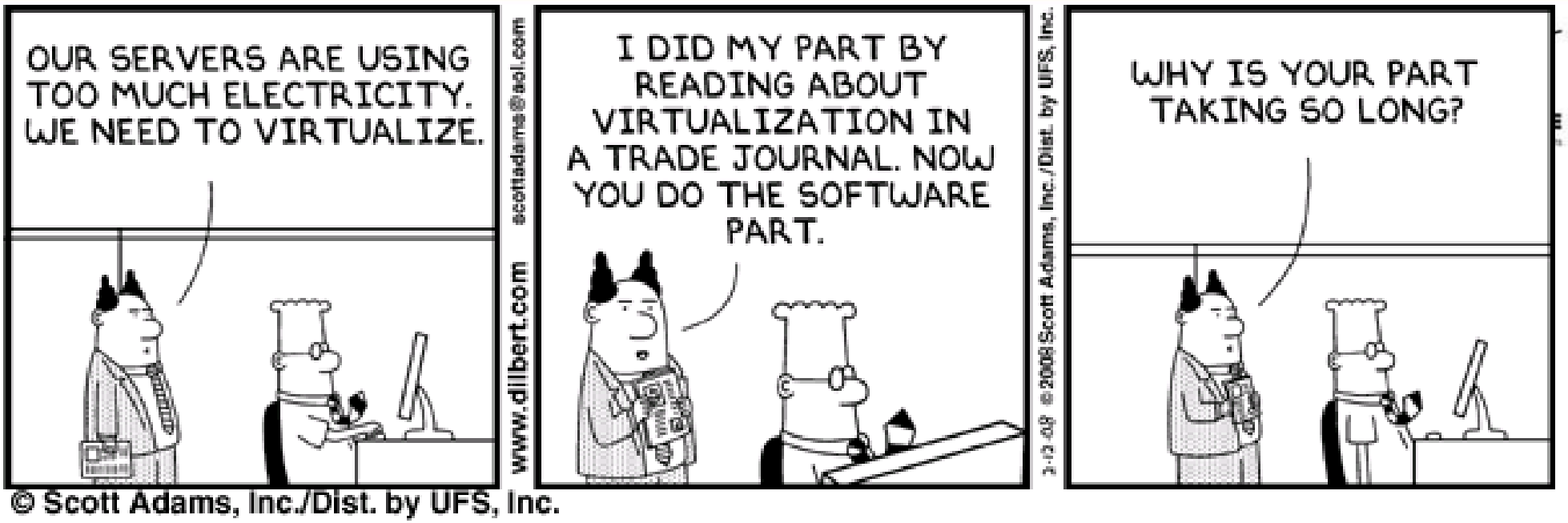
- zLinux **Total** Cost of Ownership is *far* lower, provides faster roll-out (provisioning) and more services (DR) are included than any other platform alternative
- Dozens of live production applications
  - http://www.nationwide.com – the web front door to Nationwide Insurance.
  - More production applications always in progress
- Forecasting indicates zLinux growth to continue
- The zLinux project *has* and *is still* saving real money
- It is still proven that it is the Right Thing to be doing for the business

Where are we going?

- Cooperative Memory Management (CMM)
  - Often Linux servers are built with too much memory
  - CMM helps recovery the “fluff” to use for real things
- *To R/O or not R/O, that is the question!* (Sharing)
  - Is sharing Linux file systems a Good Thing still?
    - Could be for WAS, DB2, MQ or other services
  - DCSS (share segments) for code/data/apps
- True HA
  - Working very closely with application areas to beef up total availability processes
- About 50 other roadmap items on our list!

Conclusions

- Linux virtualization on zSeries can and does:
  - Reduce cost, Reduce complexity, Accelerate provisioning, etc
- Not every workload is suited to Linux on zSeries
  - But you have to try it for yourself
  - What wasn't good even a year ago may fit nicely now
- Things are changing – still!
- With zLinux, working with z/VM is a COOL place to be again!


OUR SERVERS ARE USING TOO MUCH ELECTRICITY. WE NEED TO VIRTUALIZE.
I DID MY PART BY READING ABOUT VIRTUALIZATION IN A TRADE JOURNAL. NOW YOU DO THE SOFTWARE PART.
WHY IS YOUR PART TAKING SO LONG?
scottadams@aol.com
www.dilbert.com
© Scott Adams, Inc./Dist. by UFS, Inc.

I HIRED A CONSULTANT TO HELP WITH OUR VIRTUALIZATION PROJECT BECAUSE I DON'T TRUST EMPLOYEES WITH ANYTHING IMPORTANT.
I WILL DO THE HEAVY THINKING WHILE EACH OF YOU PERFORMS YOUR USUAL DUTIES AS OBSTACLES TO PROGRESS.
YOU SAID THIS IS MY PROJECT!
I'LL LET HIM UNPLUG SOMETHING.
scottadams@aol.com
www.dilbert.com
© Scott Adams, Inc./Dist. by UFS, Inc.

Contact Info

Light travels faster than sound, that's why people seem bright until you hear them...

Jim Vincent
Sr. z/VM Systems Programmer

Phone: (614) 249-5547
Internet: James.Vincent@nationwide.com