



# The Evolution of z/VM

## 25 Years of Innovation

Romney White, IBM  
[romneyw@us.ibm.com](mailto:romneyw@us.ibm.com)



# Agenda

- Why z/VM?
- Goals
- z/Architecture Challenges
- Design Approach
- Development Staging Plan
- z/VM Release Chronology
- The Path to the Future



## Why z/VM?

- New Z architecture being introduced with a new machine (z900) to meet client requirements for larger memory
- Backward compatible with ESA/390
- Uncertain whether development would be funded
  - OS/390 (Bob Rogers insisted) needed z/VM and its ViCom derivative in order to be able to deliver its exploitation of z/Architecture



## Goals

- Provide support for OS/390 to test its Z support in a timely manner
- Provide support for Z-capable processors running in Z mode on or immediately after processor general availability, including
  - the ability to run VM in Z mode on the machine
  - the ability to allow operating systems that exploit Z mode such as OS/390 and VM to run as guests, in Z mode on a Z-capable machine
  - the ability to support all current guests as ESA/390 guests (compatibility with prior releases)
- Maintain ability to run in ESA/390 architecture mode



# z/Architecture Challenges

- Fundamental architectural differences
  - 2-page Prefix Area
  - Different fixed storage assignments
    - Old and New PSWs
    - Interruption parameter areas
    - Logout areas
  - 8-byte general-purpose registers
  - Different address translation structure
    - 64-bit addresses
    - Region tables
  - New-format SIE State Descriptor
  - New instructions
  - New instruction formats
  - New I/O interfaces for 64-bit addressing



# Design Approach

- Develop the virtual environment first, wherever possible
  - Provides a VM environment to test new code
  - Gives some reasonable confidence in the quality of a given level of code before it "moves down" to run first level on the real hardware
- Results in a kind of "reverse bootstrap" ordering
  - Do the minimum amount of host support at each level to enable the next level of function at the guest level
  - For any given "function", add the support at the "highest" level
    - E.g., support virtual 64-bit addresses (i.e., allow guests to use them) before exploiting 64-bit guest or host real addresses, so that host support can be debugged in a guest



# Development Staging Plan

Stage	Description	Lines of Code
1	VICOM for OS/390	1,150
2	VICOM for VM	3,300
3	Base Hardware Support	16,716
4	VM Z Guest (<2G)	14,550
5	OS/390 Z Guest - Large Virtual	10,200
6	VM Z Guest - Large Virtual	4,150
7	OS/390 Z Guest - Large Real	6,075
8	VM Z Guest - Large Real	500
9	Host large real toleration	450
10	Host large real exploitation	7,850
11	Additional Infrastructure	6,302
12	Functions not yet assigned to stages	



## Stage 1 – VICOM for OS/390

- Characteristics
  - Code by the VM team for the VICOM environment
  - Changes for limited 64-bit guest virtual address translation functions available in VICOM environment
  - Limited support for new Z formats of guest DAT tables: segment and page tables
- Resulting environment
  - Basic Z support
  - No region support, implying guest virtual storage limited to  $\leq 2\text{G}$  and cannot use region tables
  - No VSIE support - a VM CP guest cannot issue SIE successfully in Z mode
  - Can test some VM Z functions in this environment, but only initialization since no VGuest virtual machines are available to set up and drive CP paths
  - The existence of large registers is simulated by VICOM



## Stage 1 – VICOM for OS/390 ...

- Content
  - Guest operand translation
  - New guest segment and page table formats
  - Architected control blocks
  - Guest operand decode, including for new instruction formats



## Stage 2 – VICOM for VM

- Characteristics
  - Limited VSIE support for Z guests (test a VM Z Basic Hardware Support guest)
  - Limited support for guest Z architecture
    - Allow 64-bit Guest CRs 1, 7, 13 with Format 1 STDs and simulate LCTLG/STCTLG support
    - Interpret RGuest Format 2 State Descriptor and translate to host shadow State Descriptor (Format 1)
  - All addresses (at all levels, host, RGuest and VGuest) limited to < 2G
  - Limited VSIE support for new Z formats of RGuest DAT tables: segment and page tables
  - VSIE shadow table support collapses guest Z DAT tables with host 390 DAT tables
  - Need to support Z mode VGuests (VGuest real and VGuest virtual addresses < 2G)
  - Minimal support for Region tables
- Resulting environment
  - Limited region support but guest virtual storage limited to <= 2G
  - VSIE support - a VM CP guest can issue SIE
  - Can test many VM Z functions in this environment



## Stage 2 – VICOM for VM ...

- Content
  - Support ESAME mode RGuest (RGuest < 2G VGuest real and virtual addresses) - RGuest Format 2 State Description
  - Handle new format segment and page tables in Rguest (Format 1 STDs)
  - Shadow tables: translate RGuest new format DAT tables through host old format DAT tables
  - Support Z VGuest (< 2G VGuest real and virtual addresses)
  - Handle new format segment and page tables in Vguest (Format 1 STDs)
  - RGuest use of region tables
  - VGuest use of region tables



## Stage 3 – Base Hardware Support

- Characteristics
  - Initialization detects presence of Z-capable hardware and does whatever is required to bring up CP in Z mode
  - CP runs with SIGP SARCH to Z mode having been issued
    - CP runs with new segment and page table formats
    - Prefix area is now two pages
    - Host CRs are 64-bits
    - Host CRs 1, 7, 13 contain ASCEs with STDs in them (until Stage 7)
    - Format 2 SIE State Descriptor used for all guests
    - Host general register are now 64- bits wide (only visible through new instructions)
    - New instructions are available to CP code (subject to design constraints)
    - Existing guests (SET MACH XC, XA, ESA) run, but cannot enable Z via SIGP SARCH



## Stage 3 – Base Hardware Support ...

- Resulting Environment
  - A stable host environment that can run on real Z-capable hardware and on which the later stages can be built
  - Somewhat in parallel with Stage 4, since VM Z is not a test environment until it can virtualize the Z architecture to guests, while running on real Z hardware itself



## Stage 3 – Base Hardware Support ...

- Content

- Save and Restore Guest Registers
- 8K Host Prefix Area with new layout
- Allocate an 8K Block on an 8K Boundary
- IPL of CP and Alternate Processor Initialization
- Dual CP Text Deck Generation
- SYSGEN - System Build
- Initial Load Into Storage
- Detect and Enable Z
- Start CPLOAD MODULE
- Bounce Processing
- Storage and CR Initialization
- Alternate Processor Initialization
- Host Virtual Storage Support
- Translation, Guest Storage and Key Manipulation
- Available List Management
- Pageable PGMBK Management
- Paging I/O Management
- Auxiliary Storage Management
- Expanded Storage (XSTORE) Management
- NSS/DCSS Processing
- Segment table size increment changes
- Minidisk Cache
- CP Hard Abend Dump
- Basic CP Dump Tool support
- Host FLIHs
- State Descriptor changes - format 2
- VMDBK restructure
- CP Trace Table
- DIAG 00 and other CP image indicators



## Stage 4 – VM Z Guest

- Characteristics
  - Provides limited Z guest support when running on Z-capable hardware
  - Start to "bootstrap" up by virtualizing the real Z architecture to the guest (in a limited fashion)
  - Something useful to OS/390
- Resulting environment
  - ESA guests can issue SIGP SARCH to enter Z mode
  - All guest real and virtual addresses and storage limited to < 2G
  - Z guest GPRs are now 64 bits (though cannot contain large addresses)
  - If Z guest uses DAT, segment and page tables are new Z format
  - New instructions available to Z guest
    - New instruction simulation (perhaps limited)
    - Guest Operand Decode
    - Guest Operand Translation (new guest segment and page table formats)
  - Debug facilities are available for testing and debugging the guest
    - TRACE (at least new instructions, 64 bit addresses not needed yet)
    - DISPLAY/DUMP/STORE (includes DISPLAY R and I as registers are larger and new instructions are executable)
  - Host VSIE for Z guest - SIE shadow table support collapses guest (full architecture) Z DAT tables with host Z DAT tables
  - Only page 0 of guest prefix area must be resident at SIE entry



## Stage 4 – VM Z Guest

- Content
  - Instruction Simulation
  - Guest operand decode
  - Guest operand translation
  - Simulation of 8K Prefix Storage Area
  - TRACE support for new instructions used by VM
  - DIAGNOSE Codes
  - CP Dump Tool support
  - VSIE - "basic" Z RGuests and VGuests



# z/VM 3.1 General Availability

- GA Date: 03/23/2001
- Originally planned to be VM/ESA 2.5
- Proposed to be z/VM 1.0 but z/VM 3.1 made more sense, especially with dual nature of support



# Release Chronology

- High-level tour
- Focus on Linux-related function



## z/VM V4

- z/VM 4.1 07/20/2001
  - IFL processor
  - Fast CCW Translation for Linux network CCWs
- z/VM 4.2 10/26/2001
  - Asynchronous page fault for Linux
  - z/VM Guest LAN
  - OBSERVER support
  - 64-bit IDAWs for DASD Fast CCW translation
  - Guest RSA-Assist support (AP crypto for Linux)
  - FICON CTC
  - DASD Large volumes



## z/VM V4 ...

- z/VM 4.3 05/31/2002
  - FCP support
  - Hipersockets II - QDIO adapter simulation, broadcast/multicast, IP takeover
  - Linux guest accounting improvements - virtual device activity
  - Linux automated shutdown (SIGNAL SHUTDOWN)
  
- z/VM 4.4 08/15/2003
  - VSWITCH
  - VLAN support
  - QDIO Adapter Interruption passthrough
  - FCP List-directed IPL
  - Hipersockets broadcast and VLAN support
  - 3270 Integrated Console
  - Virtual FICON CTCA
  - SILENT CP



# z/VM V5

- z/VM 5.1 09/24/2004
  - > 16 CPs
  - Native SCSI
  - OSA 3270
  - VSWITCH Failover
  - Layer 2 LAN
  - Virtual Machine timebomb
  - Hyperswap
  - Getting Started With Linux
  - Collaborative Memory Management
- z/VM 5.2 12/16/2005
  - QDIO V=V passthrough
  - Native SCSI II
  - Guest LAN sniffing
  - 64-bit Diagnose x'250'



## z/VM V5 ...

- z/VM 5.3 06/29/2007
  - IP Takeover
  - Mixed-engine partitions
  - >31 CPUs
  - HyperPAV
  - Link Aggregation
  - Flashcopy V2
  - Guest Program-Directed IPL
  - VSWITCH Usability
  - Native SCSI III
  - Virtual Switch SNMP Agent
  - Asynchronous CP Commands
  - Guest MIDAW support
  - Guest ASCII Console support
  - User Directory COMMAND support
  - San Volume Controller Support
  - \*VMEVENT



## z/VM V5 ...

- z/VM 5.4 09/12/2008
  - Guest FCP Dump
  - Virtual CPU SHARE redistribution
  - Host Dynamic Memory Upgrade
  - OSA Multiport support
  - VSWITCH Management improvements
  - Mixed-Engine Partition enhancements
  - DCSS >2G



## z/VM V6

- z/VM 6.1 10/23/2009
  - OSA Large Receive
  - Page Release Serialization improvement
  - XRC time-stamping
- z/VM 6.2 12/02/2011
  - Single System Image
  - Guest Dynamic Storage Reconfiguration
  - Memory constraint relief
  - VSWITCH Multiple Access Ports
  - SCSI Discovery



## z/VM V6 ...

- z/VM 6.3 07/26/2013
  - HiperDispatch
  - FCP Data Router
  - Guest Hipersocket Completion Queue support
  - Memory scalability
  - IPL NSS retention (Linux dump)
  - VSWITCH Virtual Edge Port Aggregation support
  - VSWITCH Recovery Stall Prevention
  - Live Guest Relocation Support for port-based VSWITCHes
  - Multiple Subchannel Sets support



## z/VM V6 ...

- z/VM 6.4 11/11/2016
  - Guest Enhanced DAT (1M pages)
  - Surplus CPU distribution
  - HyperPAV Paging scalability
  - SCSI Management Queries, RAS improvements, Flash Systems support
  - Memory Management Serialization contention relief
  - Dynamic SMT
  - > 1TB Memory
  - EAV Minidisk
  - Multi-VSWITCH Link Aggregation
  - Host Page Management Assist 2
  - XIV Concurrent I/O support



## z/VM V7

- z/VM 7.1 09/21/2018
  - 80 Logical CPUs
  - VSWITCH priority queueing
  - Consolidated Linux Boot Loader support
  - Memory reclamation
  
- z/VM 7.2 09/18/2020
  - EDEVICE path management
  - 1-END HyperPAV Minidisk
  - Improved Live Guest Relocation for Shared Crypto guests
  - Enhanced VSWITCH Bridge Port support
  - Automatic STANDBY Memory
  - 4 TB Memory



## z/VM V7 ...

- z/VM 7.3 09/16/2022
  - 8-member SSI
  - Guest Secure Boot
  - Large Guest LOGOFF Time mitigation
  - 2 TB guest virtual memory
  - SCSI List-directed IPL Alternate Paths
- z/VM 7.4 09/20/2024
  - Control-Only Crypto Access
  - EQDIO VSWITCH
  - QUERY SCPID
- z/VM 7.5 09/18/2025
  - HYPERSWAP for Guest HyperPAV



# The Path to the Future

- z/VM-centric community started in June 2018
  - Meeting about once a month via telephone and web conferencing
  - Clients, ISVs, IBMers
  - User research
  - Propose, define, and prioritize new project proposals
  - Sponsor User Recruitment and Playbacks
  - Other collaboration via Community
- Membership Requirements
  - IBM Z Feedback Program Agreement (FPA)
  - Regular participation
  - Sponsor User for at least one project per year
  - Homework assignments
- Additional details:
  - [http://www.vm.ibm.com/sponsor\\_user/zvm\\_council.html](http://www.vm.ibm.com/sponsor_user/zvm_council.html)
  - Contact: Kerry Wilson – [kerryw@us.ibm.com](mailto:kerryw@us.ibm.com)

