64-bit z/Architecture Overview
Part II: Supervisor Facilities

SHARE
Nashville, TN
Session 2825

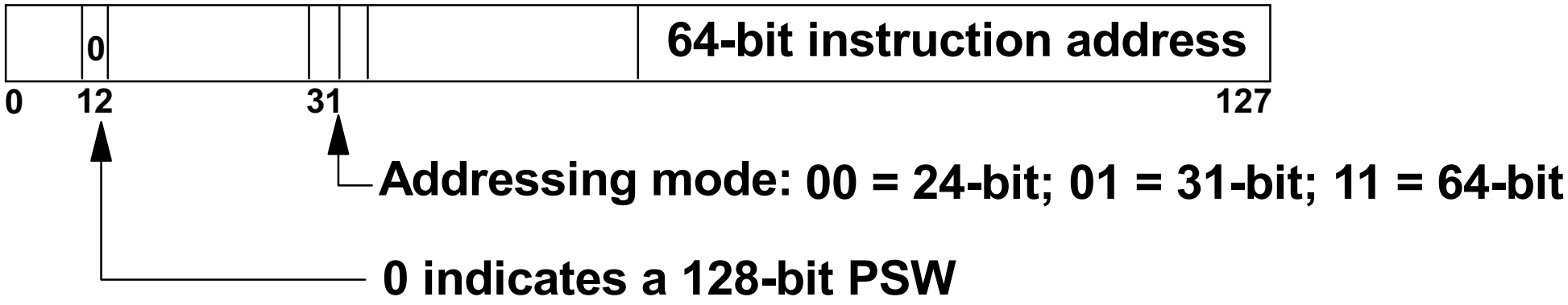
Bob Rogers
IBM Corporation
rrrogers@us.ibm.com
1-845-435-1647

Architecture Elements

- **Part II - Authorized Facilities**
  - The Program Status Word
  - The Address Translation Architecture
  - Switching into z/Architecture
  - Prefixing and PSA
  - Linkage Stack
  - Other changed formats
  - Tracing, Trap and Monitor Call
  - Control Registers
  - Store Facilites List
  - Operand Consistency
  - Dropped Facilities

Program Status Word

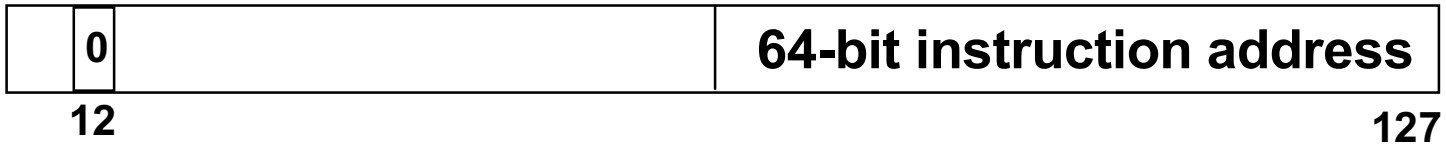
z/Architecture PSW



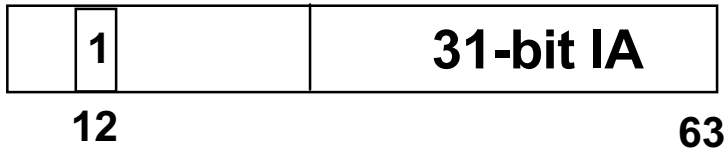
- The PSW is 128-bits with bit 12 set to 0
- LPSW expands its 8-byte operand to 128 bits and sets bit 12 to 0
- LPSWE instruction to set the full 128-bit PSW
- Bit 31 is the Extended Addressing mode bit
- Bit 32 is the Basic Addressing mode bit

The "Apparent" PSW of z/OS

Actual PSW

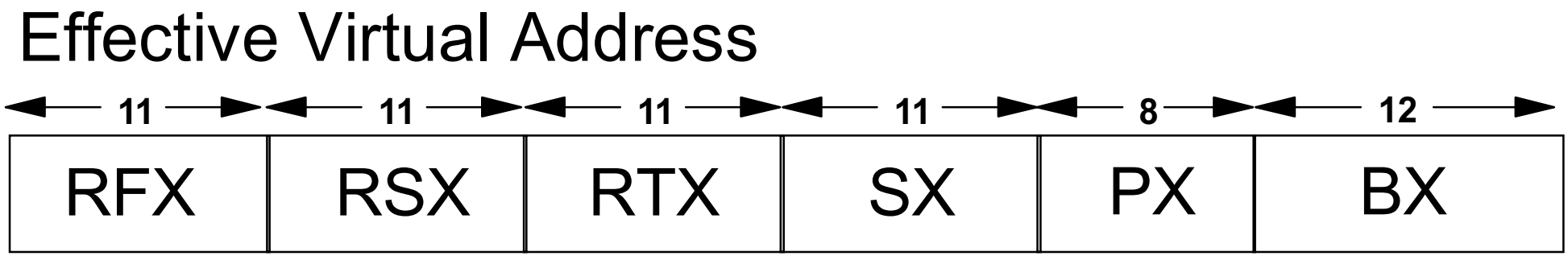


z/OS Apparent PSW



- The z/Architecture PSW is actually 128-bits long.
- The z/OS supervisor remaps it to 64-bits in control blocks, e.g. RB, IHSA, SDWA
  - bits 33-96 are removed and bit 12 set to 1
- LPSW works compatibly for these shortened PSWs

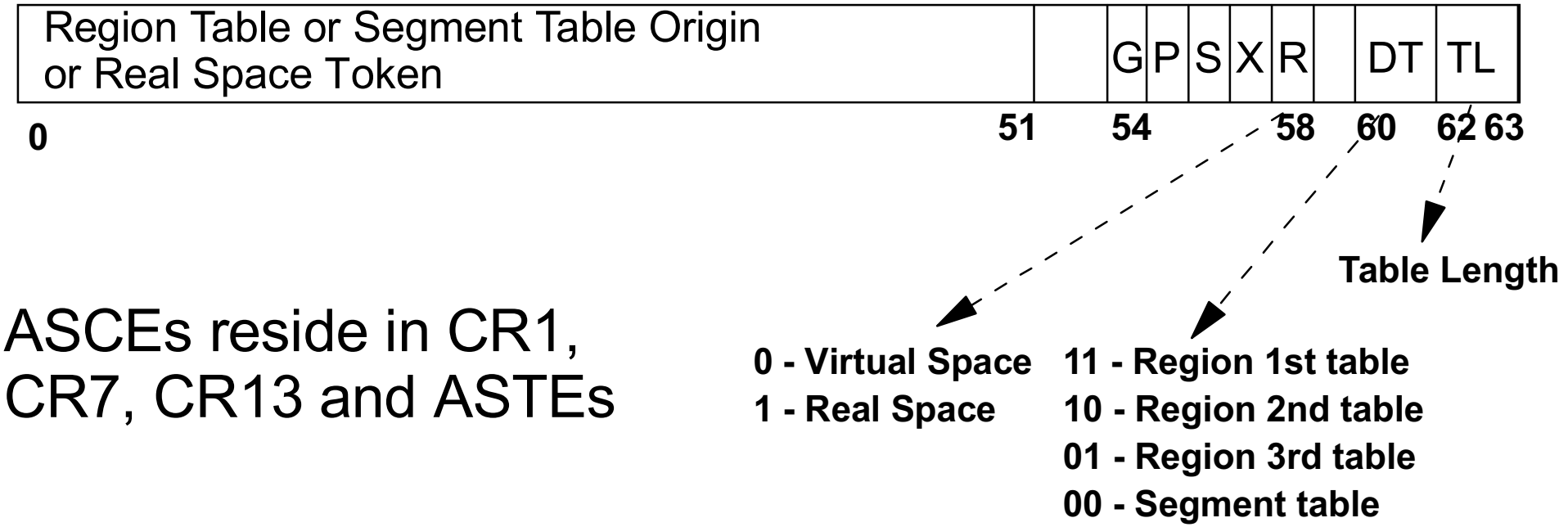
Virtual Address Translation



- Up to five levels of Dynamic Address Translation tables:
  - Region First (indexed by RFX)
  - Region Second (indexed by RSX)
  - Region Third (indexed by RTX)
  - Segment (indexed by SX)
  - Page (indexed by PX)

The Address Space Control Element

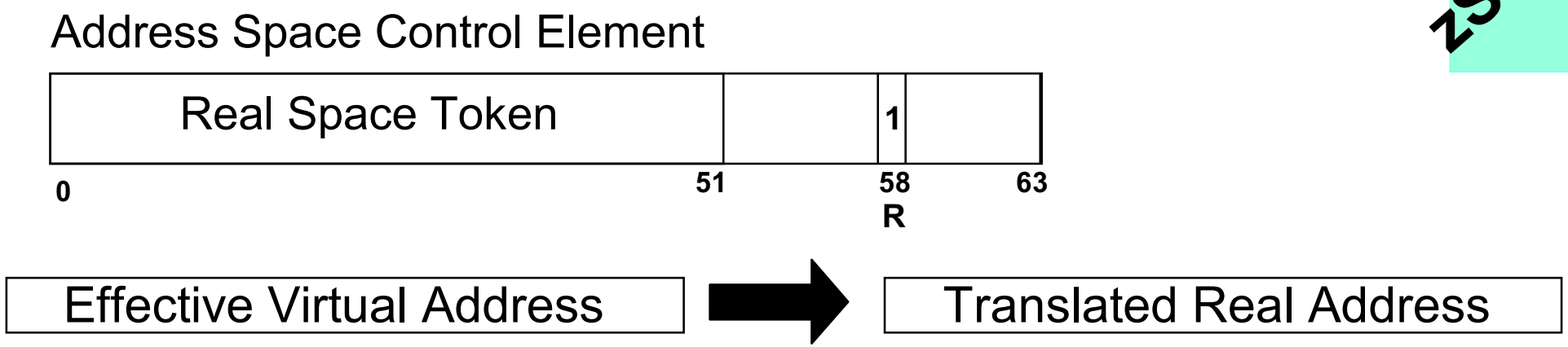
Address Space Control Element (ASCE)



ASCEs reside in CR1, CR7, CR13 and ASTEs

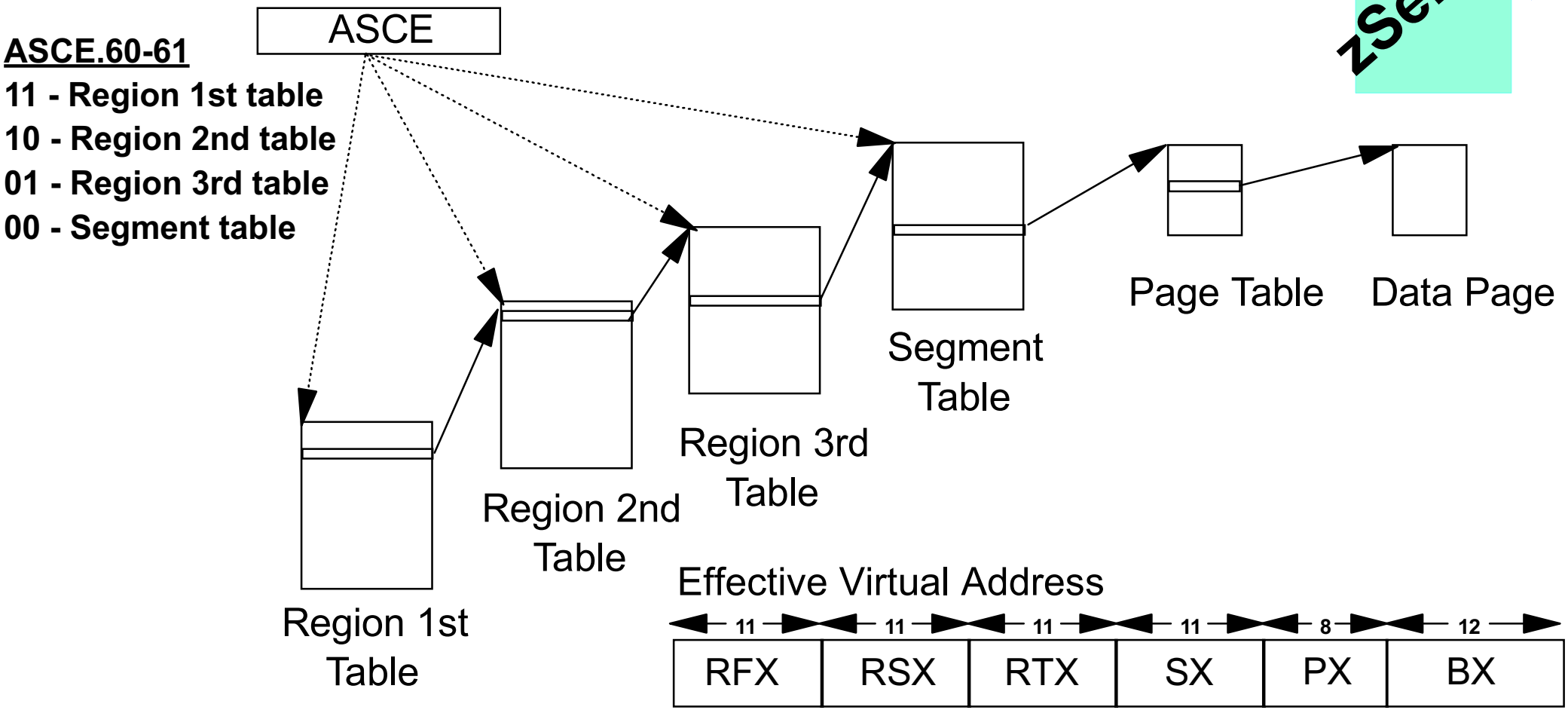
- The ASCE describes an address space.
  - A virtual space described by translation tables, or
  - The Real Space for which virtual addresses are translated to the same identical real address - no translation tables

The Real Space



- When ASCE.58 is 1, the Real Space is designated.
- Address translation for the Real Space is idempotent, that is, the effective virtual address is translated to itself as a real address.
- Addresses in the Real Space are virtual address, they just translate to the identical real address.
- The Real Space Token can be any value which is not identical to the RTO or STO in any other ASCE.

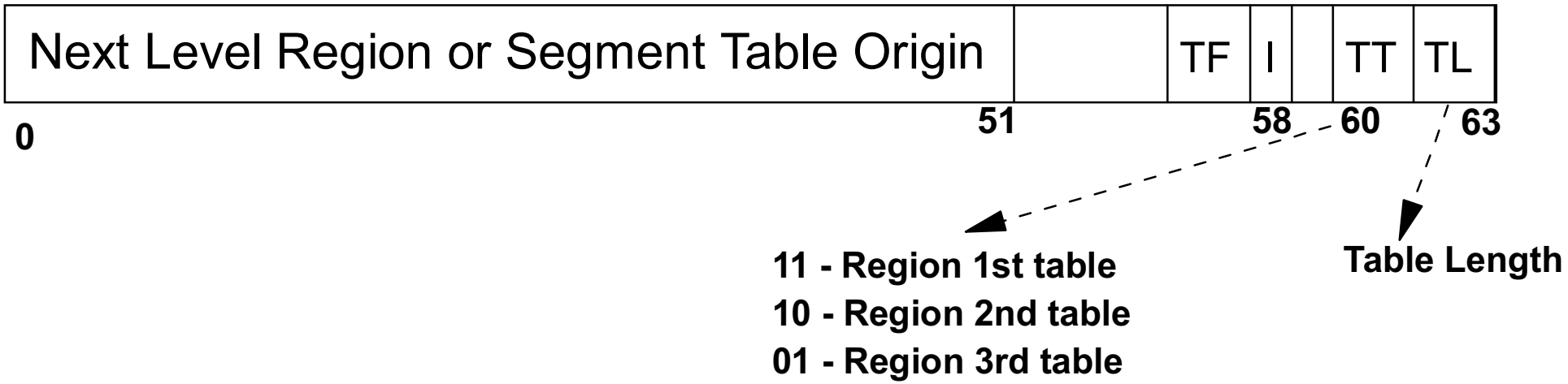
Virtual Address Spaces



- Translation can start at R1T, R2T, R3T or SGT
- The starting point of the translation is designated in the Address Space Control Element (ASCE.60-61)

The Region Table Entry

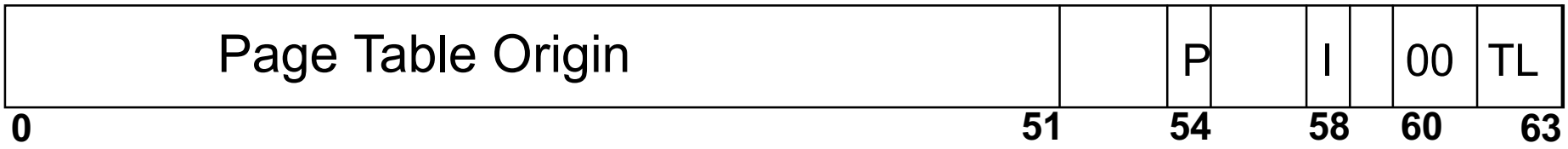
Region Table Entry



- The Region First, Second and Third Table Entries all have the same format.

The Segment Table Entry

Segment Table Entry



- The Segment Table Entries have a format very similar to the Region Table Entries.
- The Table Type (TT), bits 60-61, is 00 for Segment Table Entries.
- A Page Protection (P) bit, bit 54, is ORed with the P-bit in Page Table Entires and effectively makes the entire segment page protected.

The Page Table Entry

Page Table Entry

- The Page Table Entry is like the ESA/390 format but with the real frame address extended on the left to support 64-bit addresses

New Translation Exception Codes

- ASCE Type Exception - x'38'
  - Virtual Address > 2^{31}-1 and ASCE.60-61 = 00
  - Virtual Address > 2^{42}-1 and ASCE.60-61 = 01
  - Virtual Address > 2^{53}-1 and ASCE.60-61 = 10
- Region First Translation Exception - x'39'
  - I = 1 in Region First Table entry indexed by RFX.
- Region Second Translation Exception - x'3A'
  - I = 1 in Region Second Table entry indexed by RSX.
- Region Third Translation Exception - x'3B'
  - I = 1 in Region Third Table entry indexed by RTX.

Load Real Address - LRA, LRAG

- Load Real Address (LRA) in 24/31-bit mode works as in ESA/390 except when the real address does not fit in 31 bits.
  - if the translation is available and the real address does not fit in 31 bits, there is a special operation exception (x'13')
  - if the segement or page table entry is invalid, and the real address of the entry does not fit in 31-bits, the condition code is set to 3, rather than 1 or 2, respectively.
- LRA in 64-bit mode sets a 64-bit real address in R_1.
- Load Real Address (LRAG) always sets a 64-bit real address in R_1, regardless of addressing mode (i.e. it is a non-modal instruction).

Store Real Address - STRAG

STRAG $D_1(B_1),D_2(B_2)$

- STRAG takes two virtual address operands.
  - The first addresses a doubleword in storage where the 64-bit real address is to be stored
  - The second is the virtual address which is to be translated
- STRAG does not set the Condition Code.
  - If the translation is not available, the processor raises a translation exception (PIC x'10', x'11', etc.)
  - Therefore, it cannot be used to determine if a page is currently backed by a real frame.
  - Use TPROT before Store Real Address if you are paranoid. It provides the same non-guarantee as LRA does.
- STRAG is intended for building channel programs which access real storage above 2GB.

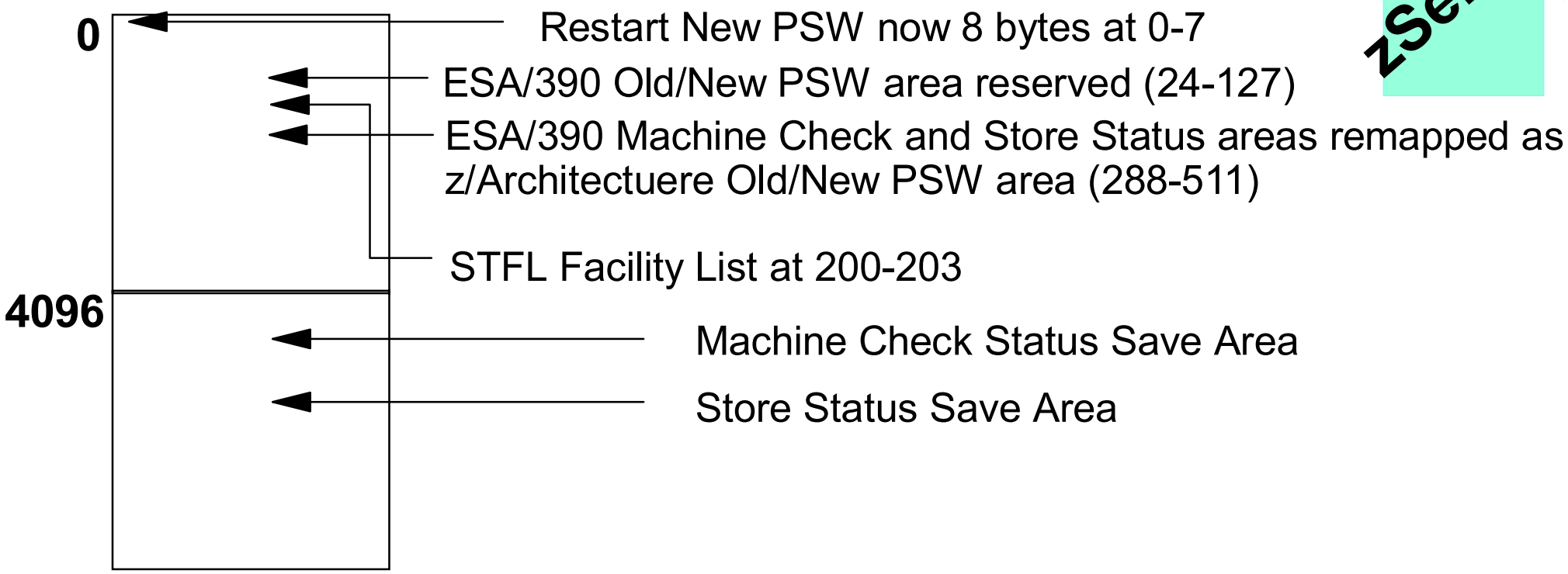
Access Register Translation (ART)

- Access Register Translation is logically unchanged from ESA/390
- There are minor adjustments due to the use of an ASCE rather than STD and to the different mapping of the Address Space Second Table Entry (ASTE)

Switching into z/Architecture

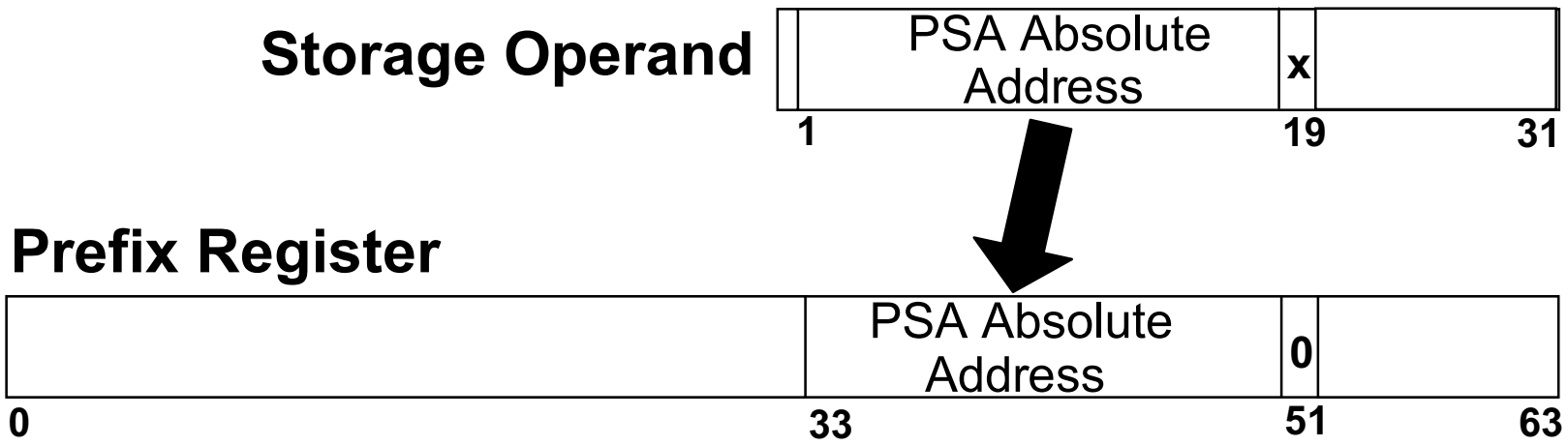
- A z/Architecture system initializes to ESA/390 mode during Initial Program Load.
- Later, an operating systems causes a switch to z/Architecture mode by issuing a SIGP using a new Set-Architecture order.
- This SIGP functions only if all other CPUs in the configuration are in manual or check-stop state, i.e. only one running CPU.
- It is also possible to SIGP back to ESA/390.

The Prefix Area



- 2 4K-frames on an 8K boundary
- Low Address Protection applies to 0-511 and 4096-4607
- MCH and Store Status information defined in 2nd frame
- New/Old PSWs moved and lengthened to 16 bytes
- other miscellaneous changes: Translation Exception Address, Monitor Code, Per Address, etc.

Setting the Prefix Register



- Prefix Register
  - 64 bits wide but only bits 33-50 are used, $PX._{0\text{-}33}$ must be 0
  - Set Prefix sets $PX._{33\text{-}50}$ from bits 1-18 of a word in storage
  - SIGP Set-Architecture to z/Architecature sets bit 51 {19} to zero.

Addressing Mode Switching with PC

- $PC_{stacking}$ can switch to and from 64-bit addessing mode.
  - A G-bit is defined in the ETE to specify whether the PC routine is entered in 64-bit addressing mode.
- PC_{basic} cannot switch to or from 64-bit addessing mode.
  - The ETE G-bit is compared with PSW.31. If they are not equal, an exception is raised.
  - PC_{basic} can be issued in 64-bit addressing mode, but only if the G-bit in the ETE is on.

The Linkage Stack

- The full 64-bit registers are always saved on the stack by $PC_{stacking}$ and BAKR
- PR restores full 64-bit registers
- Extract Status (ESTA) is only ***mostly*** compatable.
- Linkage Stack Entry Formats
  - Changed format for status entries, headers and trailers

Extracting status from the linkage stack

- Extract Status (ESTA)
  - ESTA codes 0, 2 and 3 work as in ESA/390.
  - ESTA code 1 is compatible if the instruction address in the saved PSW is $<2^{31}$. It sets the low-order half of R_1 from the first word of the saved PSW; $R_{1+1\cdot 32\{0\}}$ from saved $PSW._{32}$ and $R_{1+1\cdot 33\text{-}63\{1\text{-}31\}}$ from saved $PSW._{33\text{-}62}$
  - Otherwise, ESTA code 1 sets the low-order half of R_1 from the first word of the saved PSW; $R_{1+1\cdot 32\{0\}}$ from saved $PSW._{32}$; $R_{1+1\cdot 33\text{-}62\{1\text{-}30\}}$ from saved $PSW._{33\text{-}62}$; and sets $R_{1+1\cdot 63\{31\}}$ to 1 as an indication that the instruction address is not true.
  - An ESTA code 4 is added to extract the full PSW into an even-odd pair of 64-bit registers.

Extracting registers from the stack

- Extract Registers - EREG and EREGG
  - EREG extracts the low-order halves of registers saved in the linkage stack and inserts them into the low-order halves of the registers
  - The new EREGG instruction extracts the entire saved 64-bit register values.

Other Changed Formats

- ASN-Second-Table Entry (ASTE)
- Entry-Table Entry (ETE)
- Status Information (stored by Store Status)

Refer to Principles of Operation for details

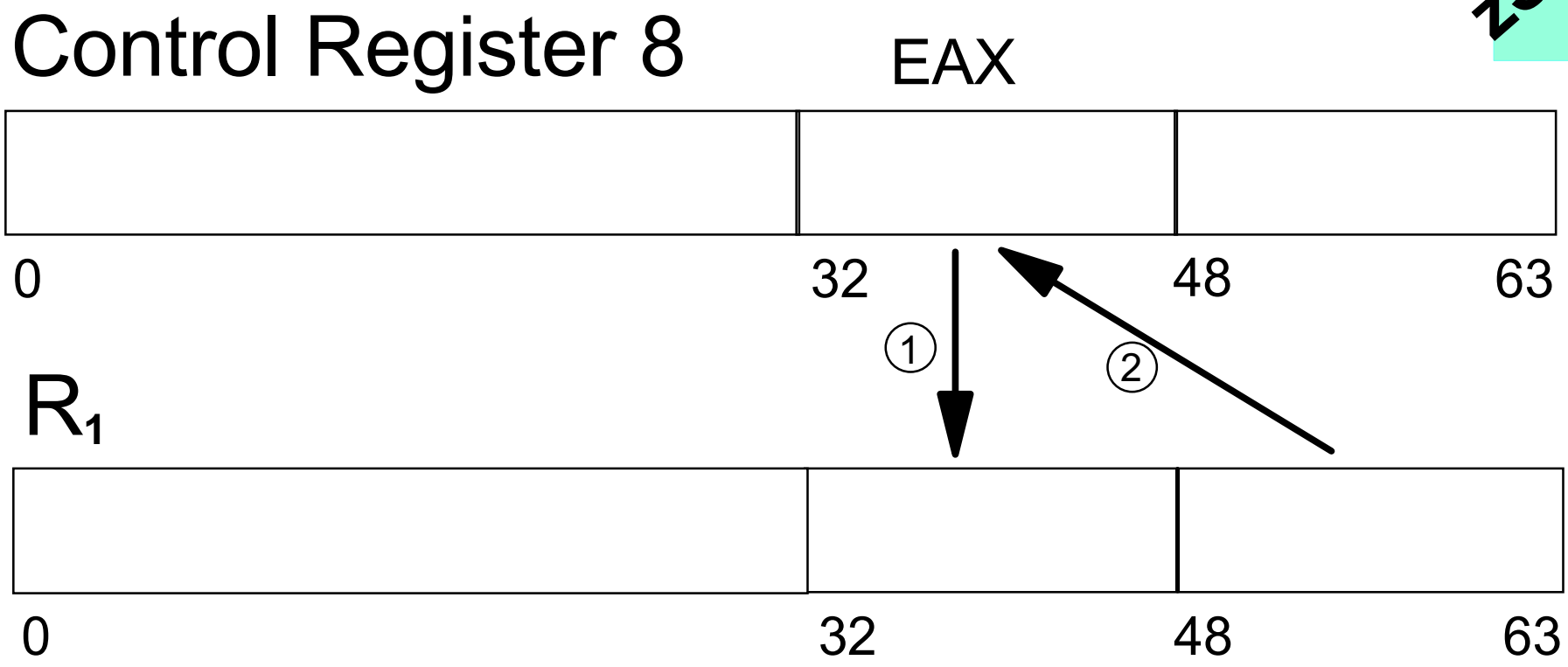
Load and Store Control Registers

- LCTL and STCTL operate only on the low-order 32 bits of the control registers
- LCTLG and STCTG operate on the full 64-bit control registers
- All Control Registers are implemented as 64 bits
  - CRs 1, 7, 10, 11, 13 and 15 are treated as 64-bit address
  - CR 12 (Trace) is a 62-bit address w/ 2 flag bits
  - For other CRs, bits 0-31 are currently unassigned

Control Registers

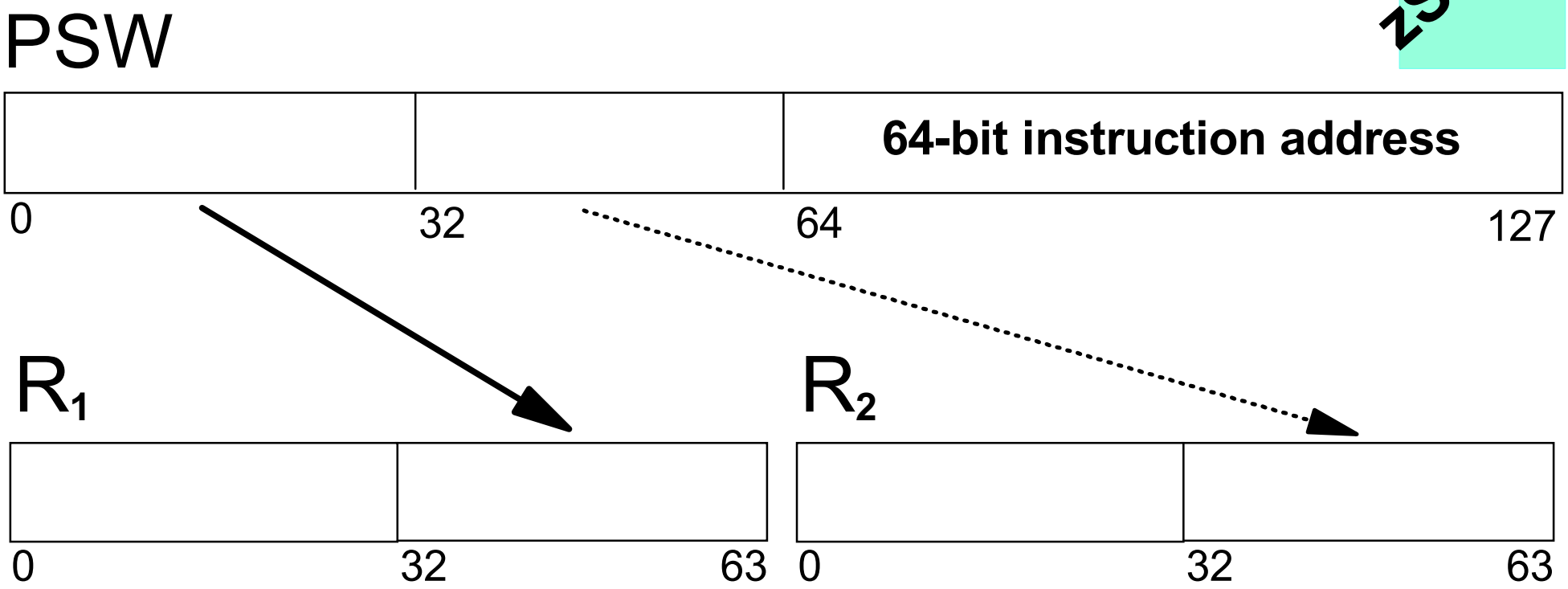
0		8	
1	Primary ASCE	9	
2		A	PER Range Start
3		B	PER Range End
4		C	Trace Entry Address
5		D	Home ASCE
6		F	
7	Secondary ASCE	F	Linkage Stack Designator

Extract and Set Extended Authority



The current EAX is extracted into bits 32-47 of the R_1 , then the EAX is set from bits 48-63 of R_1.

Extract PSW (EPSW)



The first word of the PSW is extracted into the low-order half of R_1. Also, if R_2 is not 0, the second word of the PSW is extracted into the low-order half of R_2.

Priveledged 64-bit Analog Instructions

- Load Control (LCTLG)
- Load Using Real Address (LURAG)
- Store Control (STCTG)
- Store Using Real Address (STURG)
- Trace (TRACG)

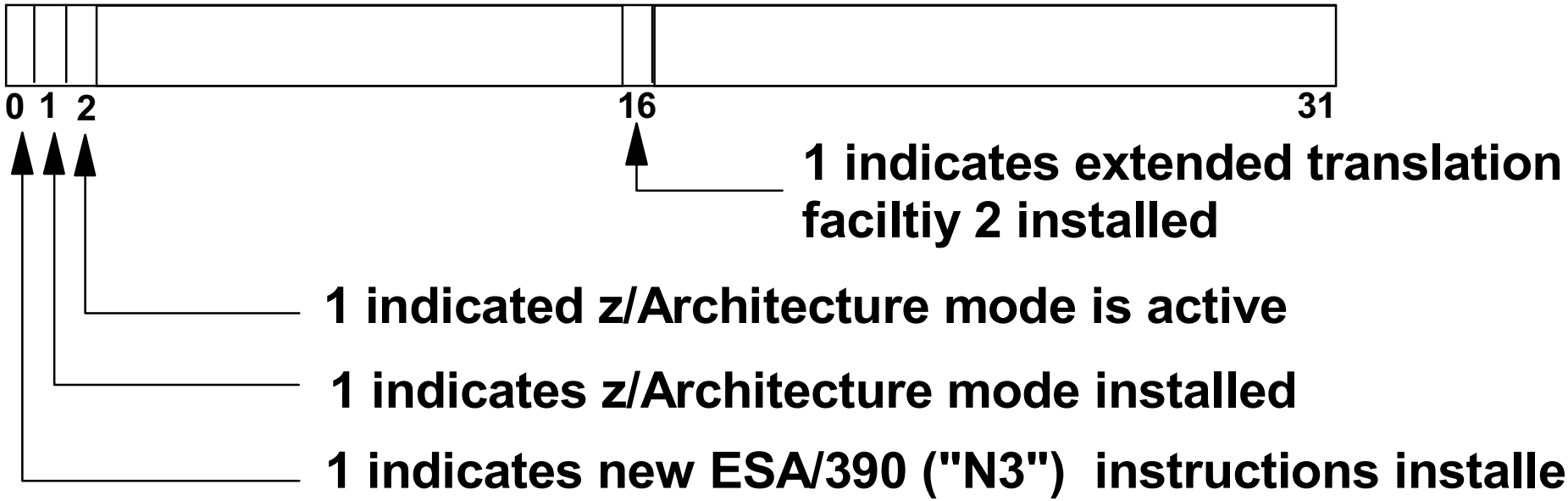
Priviledged Modal Instructions

- Move to Primary (MVCP)
- Move to Secondary (MVCS)
- Move with Key (MVCK)

The length register (R_1) containing the "true length" is 32 bits in 24/31-bit mode and 64 bits in 64-bit mode

Store Facilities List

- Store Facilities List (STFL) stores a bit mask describing installed facilities at location 200. It is a privildged operation.



*STFL added to ESA/390 (i.e. it is an "N3" instruction)

Channel Programming

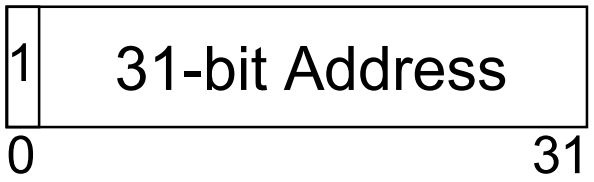
- Channel programs using either Format 0 or Format 1 CCWs are supported.
- Channel program using either Format 1 (32-bit) or Format 2 (64-bit) IDAWs are supported.
- Only a single IDAW type can be used in a channel program - either F1 or F2
- F1 IDAWs have a "span" of 2K
- F2 IDAWs can have a "span" of either 2K or 4K.
- CCWs and IDAWs must reside in absolute storage below 16MB or 2GB, depending upon CCW format.

The Trace Facility

- A new type of tracing which traces addressing mode switches into and out of 64-bit mode.There are 2 new types of trace entries (each with 3 subtypes): mode switch and mode switching branch.
- A new subtype is defined for branch trace entries when the target address is $>2^{31}$.
- A new explicit trace entry is defined for TRACG which stores full 64-bit registers into the entry.

Branch Tracing in 64-bit mode

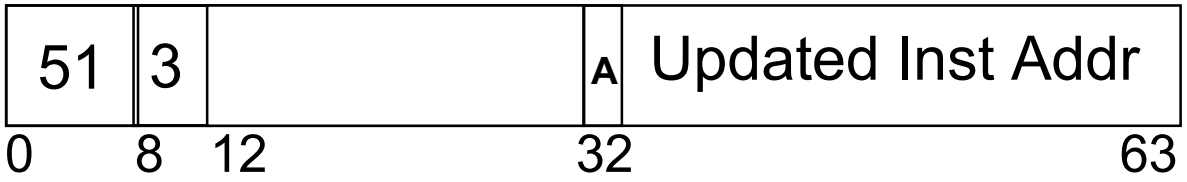
- When bits 0-32 of the branch target address are all 0, the existing 31-bit mode branch trace entry is used.



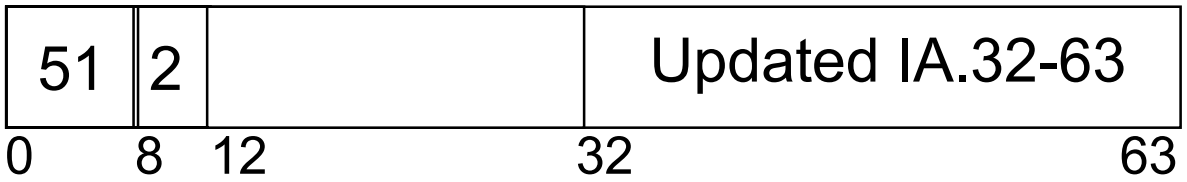
- A third Branch Trace Entry format is defined for cases when bits 0-32 of the branch target address are not all 0.

Mode Switch Tracing

- Three new trace entry formats are defined to record switches between 24/31-bit mode and 64-bit mode. The include the address of the instruction after the mode switching instruction (SAM, BASSM, BSM, PC, PR, RP)



24/31-bit to 64-bit
A=0 is 24, **A**=1 is 31

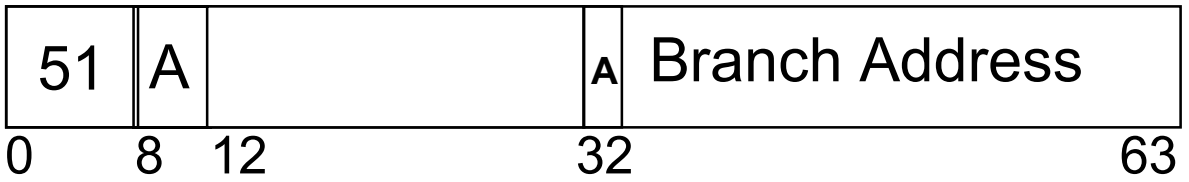


64-bit to 24/31-bit
when UIA.0-31 = 0

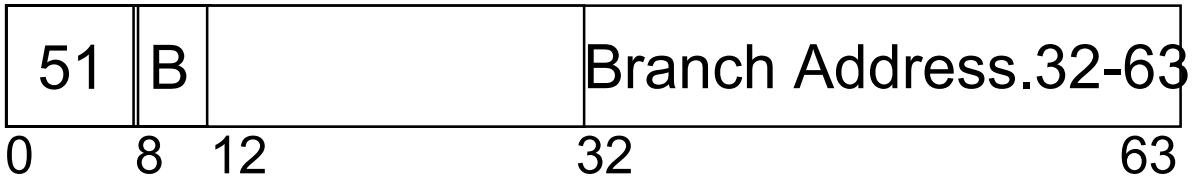
64-bit to 24/31-bit when UIA.0-31 in not 0

Mode Switching Branch Tracing

- Three new trace entry formats are defined to record branches (BASSM and RP) on which a mode switch between 24/31-bit mode and 64-bit mode takes place.The include the address of the target of the branch.



64-bit to 24- or 31-bit
A=0 for 24, **A**=1 for 31



24- or 31-bit to 64-bit
when BA.0-31 = 0

24/31-bit to 64-bit when BA.0-31 in not 0

Mode and Branch Tracing

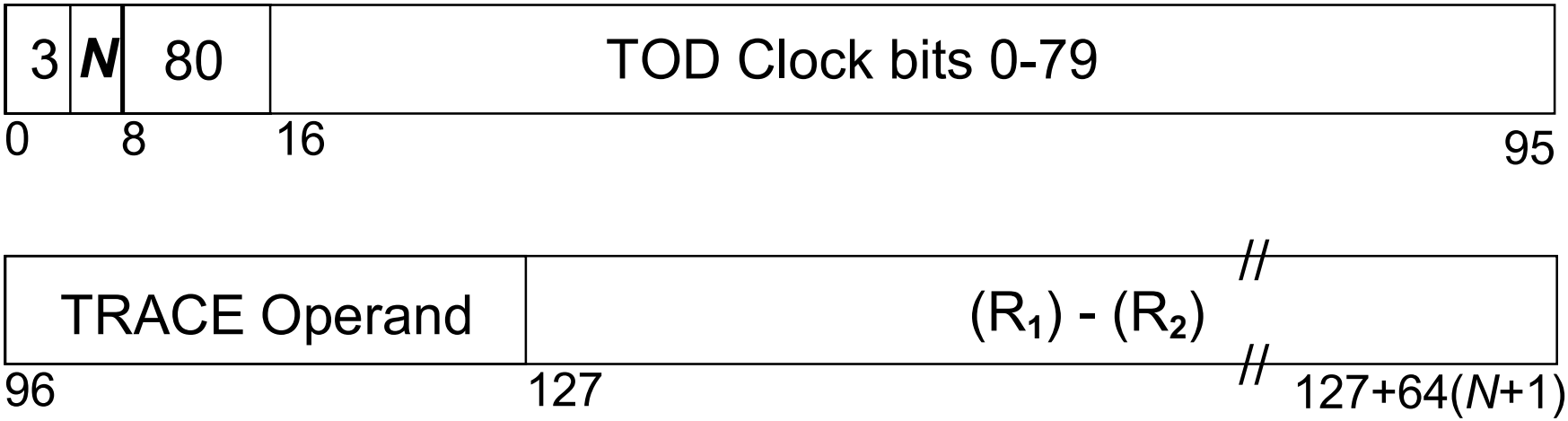
zSeries

Instruction	Branch	Mode	Both
BAKR	B		B
BALR	B		B
BASR	B		B
BASSM	B	MS	B OR MSB[1]
BSA	B		B
BGS	B		B
BSM		MS	MS
PC		MS	MS
PR		MS	MS
RP	B	MS	B OR MSB[1]
SAM 24/31/64		MS	MS
TRAP2/4	B		B

[1]Branch entry if switch is 24/31<->24/31, Mode Switching Branch entry if 24/31<->64

Explicit Tracing

- The TRACG instruction produces a trace entry which differs from that produced by TRACE in that it includes more bits of the TOD Clock and stores the full 64-bit registers into the entry



The Trap Facility

- Compatibly supports debugging programs which are not aware of z/Architecture.
- New Trap Parameter Area flags: one for 64-bit GPRs and one for the128-bit PSW.
  - When both flags are 0, the processor stores compatible status information
  - When one or both flags are 1, the processor stores 64-bit GPRs and/or 128-bit PSW.
- Trap routine always gets control in 31-bit addressing mode.
- Trap areas must reside below 2GB.

Monitor Call (MC)

- The format of the MC instruction is unchanged, as is the monitoring class mechanism.
- The monitor code is now 64 bits and is stored in a doubleword at location 176 (decimal).
- The monitor code value is formed as the address designated by the B_1 and D_1 fields of the instruction.
  - in 24-bit mode, bits 0-39 are zero
  - in 31-bit mode, bits 0-32 are zero
  - in 64-bit mode, a full 64-bit value is stored

Operand Consistency

- Doubleword consistency is provided where it is provided by ESA/390
  - LPSWE and interruption processing provide only doubleword consistency
- New instructions provide quadword consistency
  - Load Pair from Quadword [LPQ] - loads an even/odd pair from a quadword in storage
  - Store Pair from Quadword [STPQ] - stores an even/odd pair to a quadword in storage
  - Compare Double and Swap (64) [CDSG] - performs compare double and swap on 64-bit even/odd pair and a quadword in storage
  - PLO with 128-bit operands - see *Principles of Operation*

Dropped Facilities

- Asynchronous Paging (SPOUT and TAPS)
- Asynchronous Data Mover
- "Valid in Expanded Storage"
  - MVPG only moves from central to central
  - IESBE is dropped
- Program Call Fast
- Address-Space-Function Control CR0.{15}.
  - Compared to ESA/390, the bit is percieved as always 1
- Vector Facility