Linux for zSeries
Technical Update

Redbooks

International Technical Support Organization

Agenda

Performance Toolkit for VM

FCP Implementation Guide on Linux for zSeries

What's New for Linux in z/VM 5.1

Redbooks

ibm.com/redbooks

Performance Toolkit for VM

IBM.

Redbooks

International Technical Support Organization

---

ibm.com

Topics - Performance Toolkit

Performance Toolkit for VM overview

Configuration and operation

History, trend, and benchmark data collection

Remote monitoring with Performance Toolkit

Real-time monitoring with Performance Toolkit

Mapping RTM and PRF functions to Performance Toolkit

Changes for FL510

Redbooks

ibm.com/redbooks

Related Publications

Other Publications

- *These publications are relevant as information on Performance Toolkit for VM*

Title	Publication Number
Linux on IBM eServer and S/390: Performance Toolkit for VM	SG24-6059
Linux on IBM eServer and S/390: Performance Measurement and Tuning	SG24-6926
Accounting and Monitoring for z/VM Linux Guest Machines	REDP-3818
z/VM: Performance Toolkit	SC24-6062
z/VM: Performance	SC24-5999

Redbooks

ibm.com/redbooks

Performance Toolkit for VM Overview

z/VM performance analysis and monitoring tool

- Derived from FCON/ESA
- Installed on z/VM 4.4 and higher systems by default
  - Licensed z/VM product
  - Must be enabled before using!

Replacement for:

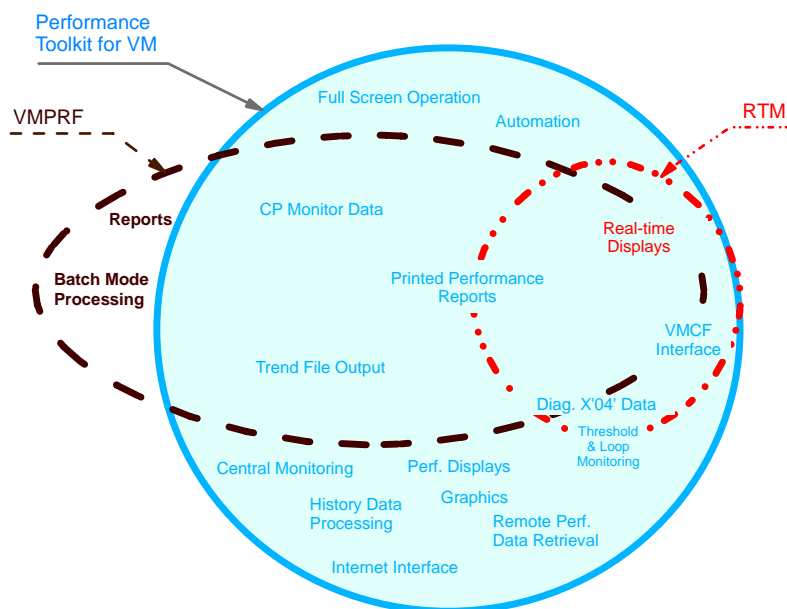- Real Time Monitor (RTM)
- Performance Reporting Facility (PRF)

Note:

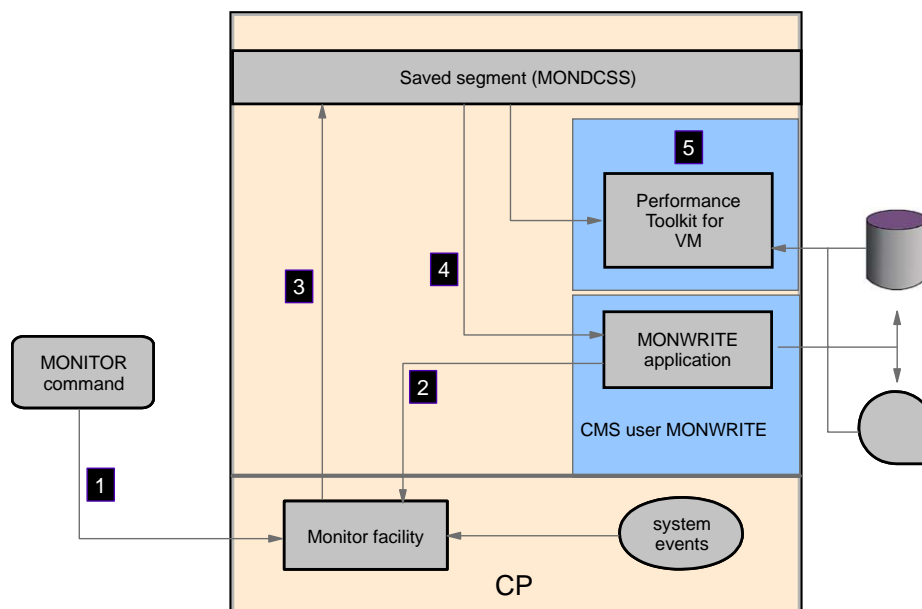- Performance Toolkit replaces RTM and PRF in z/VM 5.1

Redbooks

ibm.com/redbooks

Performance Toolkit Functions

Performance
Toolkit for VM

VMPRF

Full Screen Operation

Automation

RTM

CP Monitor Data

Real-time
Displays

Reports

Printed Performance
Reports

**Batch Mode
Processing**

VMCF
Interface

Trend File Output

Diag. X'04' Data

Threshold
& Loop
Monitoring

Central Monitoring

Perf. Displays

History Data
Processing

Graphics

Remote Perf.
Data Retrieval

Internet Interface

Redbooks

ibm.com/redbooks

Monitor Data Collection

Saved segment (MONDCSS)

5

Performance
Toolkit for
VM

MONITOR
command

3

4

MONWRITE
application

2

CMS user MONWRITE

1

Monitor facility

system
events

CP

Redbooks

ibm.com/redbooks

Monitor Data Collection (cont'd)

1. **CP MONITOR command**

   Controls type, amount of data collected

2. **MONWRITE connects to Monitor facility**

   Establishes data link to CP

3. **Monitor facility stores performance data**

   Uses MONDCSS shared segment

4. **MONWRITE retrieves, processes, stores monitor records**

   Data can be written to disk or tape

5. **Performance Toolkit performs real-time analysis**

   Can also process data written by MONWRITE

Redbooks

Types of Monitor Data

Event Data

- Collected/reported each time a system event occurs
- Data represents system status at the time of the event
- Example:
  - CP MONITOR ENABLE EVENT I/O DEVICE 1234

Sample Data

- Single-sample data
  - Data collected/reported once
  - May be status or accumulated counter values
- High-frequency sample data
  - Data added to counters at end of sampling time
  - Accumulated counters reported at end of interval

Redbooks

Monitor Data Domains

Domain 0 - System   system-wide resource usage (sample)

Domain 1 - Monitor   system configuration (sample and event)

Domain 2 - Scheduler   scheduler queues (event)

Domain 3 - Storage   storage usage (sample and event)

Domain 4 - User   scheduling, I/O, events (sample and event)

Domain 5 - Processor   processor usage (sample and event)

Domain 6 - I/O   I/O requests, interrupts (sample and event)

Domain 7 - Seek   DASD seek operations (event)

Domain 10 - Application application-specific (sample and event)

CP Commands

Controls data selection, collection, reporting
- Requires A or E privilege

MONITOR EVENT
- Establish/modify profile for event data collection

MONITOR SAMPLE
- Establish/modify profile for single and high frequency data collection

MONITOR START/STOP
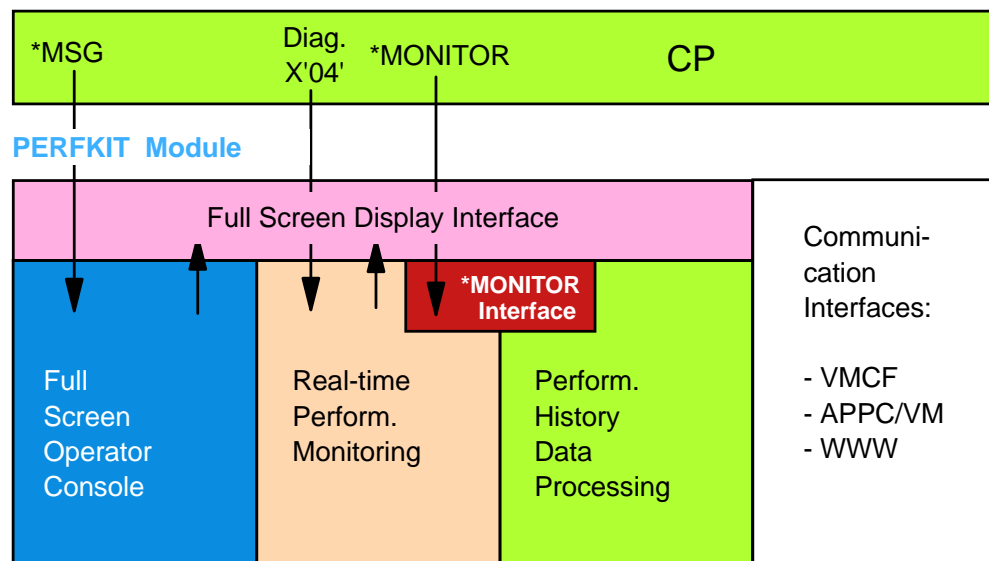- Activates/deactivates event and sample monitoring

QUERY MONITOR
- Display current monitor profile

PERFKIT Module

*MSG	Diag. X'04'	*MONITOR	CP

PERFKIT Module

Full Screen Display Interface

*MONITOR Interface

Full Screen Operator Console	Real-time Perform. Monitoring	Perform. History Data Processing	Communi-cation Interfaces: - VMCF - APPC/VM - WWW

Performance Toolkit Overhead

Generally very low for real-time monitoring

- CPU load for normal collection <<1% with default 60 sec interval
- Typically 0.1% - 0.2% with 1000 users

Your mileage may vary - actual load depends on:

- System being monitored
  - I/O configuration
  - Number of monitored virtual machines
- Number of enabled monitor domains
- Time interval between samples

Factors That Contribute to Overhead

Monitor domains

- Performance Toolkit collects data for all enabled domains
- Can be considerable overhead for:
  - MONITOR ENABLE USER ALL
  - MONITOR ENABLE SEEKS ALL

Data collection frequency

- Sample interval used for collecting data from CP control blocks
- Default interval is 60 secs
- Longer intervals reduce overhead

When to Use Performance Toolkit

Part of normal system operation

- Use discretion when enabling data collection

Real-time monitoring of system performance

- Monitor specific virtual machines or resources

Automation based on specific events

- Use thresholds to invoke REXX execs

History data collection

- Trend analysis and capacity planning

To Summarize:

Performance Toolkit is strategic direction for VM monitoring

- Replaces RTM and VMPRF in z/VM 5.1

Reads data from CP monitor facility

- Can read data from files created by MONWRITE

Analyzes both event and sample data

- CP commands control data collection

Generally low overhead

- Use data collection intelligently

Redbooks

Configuration and Operation

The 4VMPTK10 z/VM user

- Contains Performance Toolkit code / configuration files

The PERFSVM z/VM user

- Default Performance Toolkit user

Invoking the Performance Toolkit

- PERFKIT command
- Modes of operation
  - Basic
  - Redisplay
  - Monitor
- Help facility

Redbooks

4VMPTK40 Minidisks

200

- Test build disk

201

- Production build disk (contains executable code)

1CC

- IBM-supplied sample files

CCC

- Customized control files

29D

- Help files

PERFSVM User Directory Entry

```
USER PERFSVM  PERFSVM 64M 512M ABDEG
  MACHINE XA
  XAUTOLOG AUTOLOG1
  ACCOUNT xxxx
  NAMESAVE MONDCSS
  IUCV *MONITOR MSGLIMIT 255
  IUCV *IDENT FCXRES00 GLOBAL
  IUCV *IDENT FCXSYSTM GLOBAL
  IUCV ALLOW
  SHARE ABS 3%
  IPL CMS
  OPTION QUICKDSP
  CONSOLE 0009 3215
  SPOOL 000C 2540 READER *
  SPOOL 000D 2540 PUNCH A
  SPOOL 000E 1403 A
  LINK MAINT 190 190 RR
  LINK MAINT 19D 19D RR
  LINK MAINT 19E 19E RR
  LINK 4VMPTK40 200 200 RR
  LINK 4VMPTK40 201 201 RR
  LINK 4VMPTK40 1CC 1CC RR
  LINK 4VMPTK40 CCC CCC RR
  LINK 4VMPTK40 29D 29D RR
  MDISK 191 3390 1007 060 440W02  MR READ      WRITE     MULTIPLE
  MDISK 195 3390 1067 060 440W02  MR READ      WRITE     MULTIPLE
```

Directory Entry Options

NAMESAVE MONDCSS

- Enables access to CP monitor data

IUCV *IDENT FCXRES00 GLOBAL
IUCV *IDENT FCXSYSTM GLOBAL
IUCV ALLOW

- Defines resources for remote data retrieval

SHARE ABS 3%
OPTION QUICKDSP

- Ensures PERFSVM is dispatched long enough
- Will not wait in dispatch list

LINK 4VMPTK40 ..

- Access control files and executables

PROFILE EXEC for PERFSVM

```
/* */
'ACC  201 B '        /* Production Disk            @FC068BD*/
'ACC  CCC D '        /* Customized controls disk         */
'ACC  1CC E '        /* Sample controls disk             */
'ACC  29D F '        /* Help files                 @FC089BD*/

/* Once you have PERFKIT enabled and running uncomment the   */
/* following comments                                        */

/* 'CP MONITOR SAMPLE ENABLE PROCESSOR'    */
/* 'CP MONITOR SAMPLE ENABLE STORAGE'      */
/* 'CP MONITOR SAMPLE ENABLE USER ALL'     */
/* 'CP MONITOR SAMPLE ENABLE I/O ALL'      */
/* 'CP MONITOR SAMPLE ENABLE APPLDATA ALL' */
/* 'CP MONITOR EVENT  ENABLE STORAGE'      */
/* 'CP MONITOR EVENT  ENABLE I/O ALL'      */
'PERFKIT'            /* Invoke the PERFKIT module         */

Exit
```
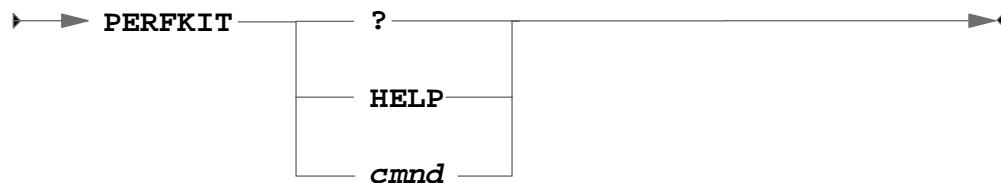
Starting Performance Toolkit

```
   ┌──────> PERFKIT ──┬──── ? ────┬────────────────────────>◄
                      │           │
                      ├─── HELP ──┤
                      │           │
                      └─── cmnd ──┘
```

Configuration Files

FCONX $PROFILE
- Controls Perfomance Toolkit operation

FCONX REPORTS
- Controls reports to produce

FCONX TRENDREC
- Controls trend records to produce

FCONRMT AUTHORIZ
FCONRMT SYSTEMS
- Controls remote data retrieval

FCONX LINUXUSR
- Controls RMF/DDS interface to Linux guests

FCONX $PROFILE

Perfomance toolkit initialization file
- Used to customize operation
- Sample provided on 4VMPTK40 1CC disk
- Modified copy goes on the 4VMPTK40 CCC disk

Contains:
- Performance Toolkit sub-commands
  - Many FCONTROL statements
- CP or CMS commands
- Comments

Executed:
- Immediately when PERFKIT is started
- Commands executed in order

FCONTROL subcommand

Accepts many options (~36)
- ACTMSG   set number of action pending action messages
- BENCHMRK   define device/user to benchmark
- DEFSCRN   create customized display
- MONCOLL   control data collection
- LIMIT   define thresholds
- PFKEY   customize PF key assignments
- PROCESS   specifies actions based on basic mode output
- SEARCH   set command search order
- SETEVENT   define day/time command is executed
- UCLASS   define class of user
- USRLIMIT   set threshold for user resource consumption

Basic Mode Screen Layout

Screen identifer   **Status message area**   **Scroll setting**

```
FCX001              Performance Toolkit for VM      AUTOSCROLL 10
01
02
03  ◄── Line numbers (optional)
04
05
06
07
08
09
10
11    Output area
12
13
14
15
16
17
18
19
20
.... Message area ....
Command ===>  Command line
F1=Help  F2=Redisplay  F3=Quit  F12=Return
```
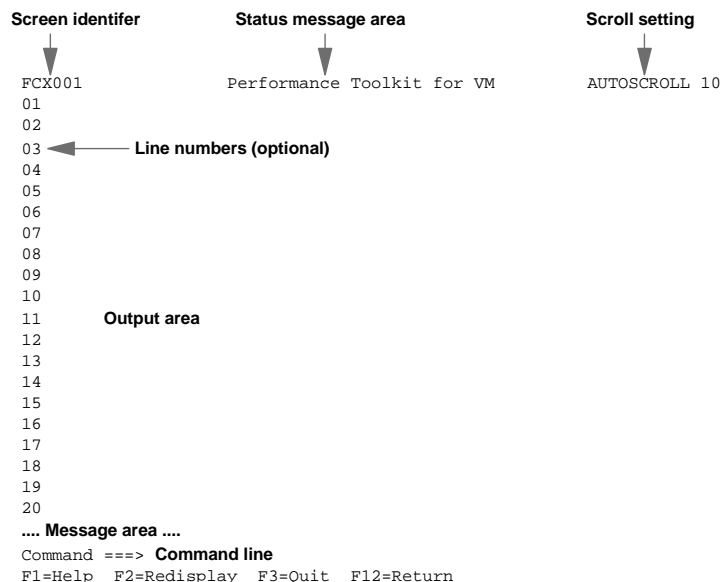
Basic Mode

Command search order
- Set by FC SEARCH command (default is CPCMS)
- May use CP, CMS, or EXEC prefix

Scrolling
- Output displayed if free lines are available
- Hit Enter to freeze current screen

Message handling
- Error messages displayed in Message Area
- Action messages displayed at top of screen (FC ACTMSG)
- DELETE subcommand removes messages
- '=' re-executes previous command
- '?' retrieves previous command

Controlling Scrolling Mechanism

Scrolling is diffent than standard VM console

- Screen is not automatically cleared
- Lines shift upward
- Messages and last action messages remain on screen

FCONTROL SCROLL sets scrolling behavior

- Default is automatic
  - FC SCROLL AUTO 10
- Can be made manual
  - FC SCROLL MANUAL 12

Use color highlighting distinguish message types

- FC COLOR command

Processing Basic Mode Messages

Basic mode output lines can be filtered

- Actions can be taken based on type of message

Use FCONTROL PROCESS command to specify

- Type of message
- Filter conditions
- Action to perform
  - Reroute messages
  - Execute REXX script / CMS module

Route tape-related messages to tape operator (TAPEOP):

FC PROCESS CPMSG * 'TAPE' NODISP RER TAPEOP CPMSG

FC PROCESS CPMSG * 'HCPER2215A' NODISP RER TAPEOP CPMSG

Action Messages

Indicate operator action is required

- Appear at top of screen
  - FC ACTMSG controls the number left pending (default is 10)
- Message is preceeded by '+*nn*'
  - *nn* idicates line number

Action messages can be:

- Performance Toolkit messages
  - Threshold exceeded alerts
  - Exception messages
- CP / CMS messages

Defined using FC PROCESS command

Threshold Monitoring

Set threshold values for key performance variables

- Defined with FCONTROL LIMIT command
- Additional parameters can be applied to threshold values
  - Time period value is exceeded
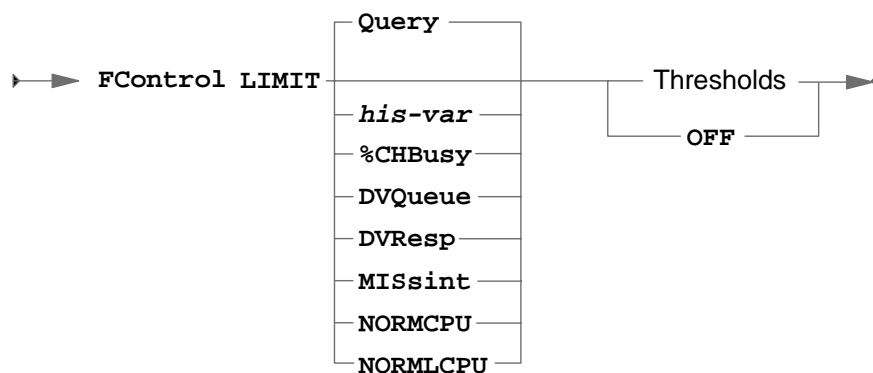 - Minimum number of users in queue (Q1 and Q0 / Q2 / Q3)

Alert messages generated when threshold reached

- Displayed in basic mode
- Alert messages are highlighted, do not scroll sutomatically
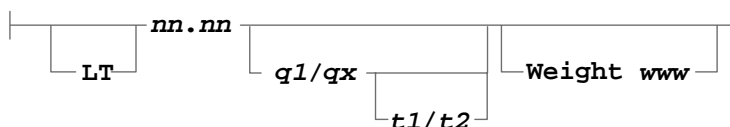 - Use DELETE command to remove messages

FCONTROL LIMIT Syntax

```
                          ┌─ Query ─────┐
                          │             │
  ►──► FControl LIMIT ────┼─ his-var ───┤──── Thresholds ──────►◄
                          ├─ %CHBusy ───┤    └─ OFF ─┘
                          ├─ DVQueue ───┤
                          ├─ DVResp ────┤
                          ├─ MISsint ───┤
                          ├─ NORMCPU ───┤
                          └─ NORMLCPU ──┘
```

Thresholds:

```
  ┌──────────┬─ nn.nn ──┬──────────┬── Weight www ──┐
  │          │          │          │                │
  └─ LT ─┘              ├─ q1/qx ──┤
                        │          │
                        └─ t1/t2 ──┘
```
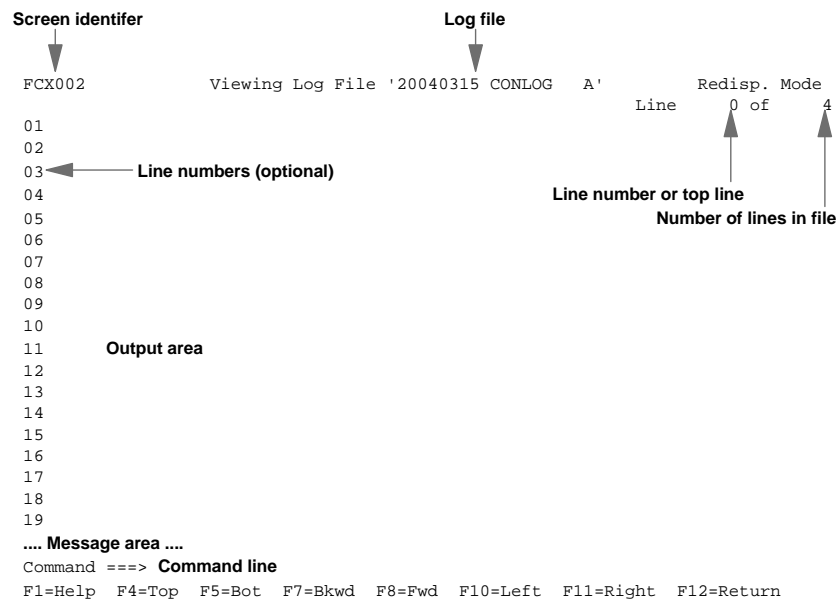

Threshold Alert Messages

```
FCX001   -------------- Press 'ENTER' to continue -------------- Autoscroll 12
+23 08:37:00 FCXPER315A Cl1 time slice 2.002 exceeds limit 1.000 (Q1=01 Qx=02)
+24 08:59:00 FCXUSL317A User LNXSU1 IO/s  377 exceeded threshold 50.0 for 5 min.
+25 09:51:00 FCXUST457A PERFKLA 60% active: may need higher priority
+26 10:12:00 FCXPER315A % I/O wait (sys)  52 exceeds limit  50 (LNXRH1   .5 IO/S)
+27 10:27:00 FCXPER315A Cl1 time slice 2.002 exceeds limit 1.000 (Q1=01 Qx=03)
+28 10:27:00 FCXPER315A % I/O wait (sys)  52 exceeds limit  50 (PERFROG   .5 IO/S)
*29 12:00:00 MSG FROM PERFKLA : IT IS NOW 12:00H
+30 12:32:00 FCXPER315A Cl1 time slice 2.000 exceeds limit 1.000 (Q1=01 Qx=02)
+31 12:32:00 FCXPER315A % I/O wait (sys)  55 exceeds limit  50 (PERFROG   .6 IO/S)
+32 12:39:00 FCXPER315A Cl1 time slice 2.002 exceeds limit 1.000 (Q1=01 Qx=03)
+33 12:50:00 FCXCHA318A Channel 48: 67% busy exceeds limit of 40%
+34 13:03:00 FCXCHA318A Channel 49: 70% busy exceeds limit of 40%
+35 14:28:00 FCXCHA318A Channel 75: 70% busy exceeds limit of 40%
+36 14:39:00 FCXUSL317A User LNXSU2 IO/s  574 exceeded threshold 50.0 for 5 min.
+37 14:49:00 FCXUSL317A User LNXSU1 %CPU 33.9 exceeded threshold 30.0 for 5 min.
+38 14:59:00 FCXUSL317A User LNXSU1 %CPU 56.7 exceeded threshold 30.0 for 15 min.
+38 15:09:00 FCXUSL317A User LNXSU1 %CPU 37.1 exceeded threshold 30.0 for 25 min.




Command ===>
F1=Help  F2=Redisplay  F3=Quit  F12=Return
```

Redisplay Screen Layout

Screen identifer   **Log file**

```
FCX002              Viewing Log File '20040315 CONLOG   A'      Redisp. Mode
                                                             Line      0 of      4
01
02
03  ←  Line numbers (optional)
04                                          Line number or top line
05                                          Number of lines in file
06
07
08
09
10
11      Output area
12
13
14
15
16
17
18
19
.... Message area ....
Command ===>  Command line
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F10=Left  F11=Right  F12=Return
```

Redisplay Mode

Use REDISP command to enter redisplay mode

- Can use PF2 from basic mode
- Initial screen displays last page of console log file
- Console log file: *mmddyy* CONLOG

Navigate using subcommands or PF-keys

- FORWARD (PF8)
- BACKWARD (PF7)
- TOP (PF4)
- BOTTOM (PF5)
- LOCATE (*/text*)

Monitor Mode Screen Layout

```
FCX124        Performance Screen Selection  (FL440 VM63447)   Perf. Monitor

General System Data      I/O Data                History Data (by Time)
1. CPU load and trans.   11. Channel load        31. Graphics selection
2. Storage utilization   12. Control units       32. History data files*
3. Storage subpools      13. I/O device load*    33. Benchmark displays*
4. Priv. operations      14. CP owned disks*      34. Correlation coeff.
5. System counters       15. Cache extend. func.* 35. System summary*
6. CP IUCV services      16. DASD I/O assist      36. Auxiliary storage
7. SPOOL file display*   17. DASD seek distance*  37. CP communications*
8. LPAR data             18. I/O prior. queueing* 38. DASD load
9. Shared segments       19. I/O configuration    39. Minidisk cache*
A. Shared data spaces    1A. I/O config. changes  3A. Paging activity
B. Virt. disks in stor.                           3B. Proc. load & config*
C. Transact. statistics  User Data                3C. Logical part. load
D. Monitor data          21. User resource usage* 3D. Response time (all)*
E. Monitor settings      22. User paging load*    3E. RSK data menu*
F. System settings       23. User wait states*    3F. Scheduler queues
G. System configuration  24. User response time*  3G. Scheduler data
H. VM Resource Manager   25. Resources/transact.* 3H. SFS/BFS logs menu*
                         26. User communication*  3I. System log
I. Exceptions            27. Multitasking users*  3K. TCP/IP data menu*
                         28. User configuration*  3L. User communication
K. User defined data*    29. Linux systems*       3M. User wait states

          Pointers to related or more detailed performance data
          can be found on displays marked with an asterisk (*).
Select performance screen with cursor and hit ENTER
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

ibm.com/redbooks

Monitor Mode

MONITOR command switches to monitor mode

Reports are selectable by:
- Subcommand name
- Option number
- Sensitive selection

If option not highlighted, no monitor data is available
- At least two sample intervals required to accumulate data

Asterisk (*) indicates option contains submenus

MENU command returns to main menu

ibm.com/redbooks

Initial Help Screen

```
PERFKIT TASKS              Task Help Information            line 1 of 27

                          PERFORMANCE TOOLKIT FOR VM

 Move the cursor to the selected item and press ENTER or the PF1 key.


 Introduction  - General introduction for beginners
               - Use in basic command mode for system operation
               - Use in redisplay mode for re-viewing accumulated log
               - Use in perf. monitor mode (selecting performance screens)
               - Use in remote performance monitor mode
               - Use of VMCF data retrieval interface
               - Use of APPC/VM data retrieval interface
               - Use of the WWW data retrieval interface

 Sub-commands  - General (all modes)              (provides a detailed
               - Basic mode sub-commands           list of commands
               - Redisplay mode sub-commands       which are valid
 PF1= Help    2= Top      3= Quit     4= Return    5= Clocate   6= ?
 PF7= Backward 8= Forward  9= PFkeys  10=          11=          12= Cursor

 ====>
                                                    Macro-read 1 File
```

Help Facility

Access the master help menu using the PF1 key
- Use PF3 to exit the menu

Place cursor over the desired topic
- Press Enter to drill down

Context-sensitive help is available
- Move cursor over field heading and press Enter
- Help text is displayed in pop-up window
- May describe more than one field
- Only available for screen with fixed layout

Context Sensitive Help

```
FCX112     CPU 2084  SER 96A3A  Interval 14:41:47 - 14:42:07   Perf. Monitor

 .         ___    .    .     .    .    .    .    .     .
           <----- CPU Load -----> <------ Virtual IO/s ------>
            <-Seconds->   T/V
Userid   %CPU  TCPU  VCPU Ratio Total DASD Avoid Diag98    UR Pg/s  User Status
>System< 3.34  .668  .508  1.3   102  101   141    .0     .0  .0   ---,---,---
+------------------------------------------------------------------------+
|                              Help Text                                 |
|                                                                        |
| %CPU           Percent of total CPU used.                              |
|                This value is based on the utilization of a single      |
|                processor: Values exceeding 100% are possible for virtual MP |
|                users.                                                  |
|                                                                        |
| F12=Return                                                             |
+------------------------------------------------------------------------+
LNXSU3    .21  .041  .027  1.5   .1   .1    .0     .0    .0   .0  ESA,CL3,DIS
LNXSU4    .21  .041  .028  1.5   .0   .0    .0     .0    .0   .0  ESA,CL3,DIS
LNXSU5    .19  .037  .023  1.6   .0   .0    .0     .0    .0   .0  ESA,CL3,DIS
PERFKLA   .12  .024  .023  1.0   .4   .3    .2     .0    .0   .0  ESA,---,DOR
PERFLIV   .11  .022  .021  1.0   .4   .3    .1     .0    .0   .0  ESA,---,DOR
PERFROG   .11  .022  .020  1.1   .7   .6    .3     .0    .0   .0  ESA,CL0,DIS
PERFSVM   .10  .020  .019  1.1   .4   .3    .1     .0    .0   .0  ESA,---,DOR
PERFAPP   .08  .016  .015  1.1   .4   .4    .2     .0    .0   .0  ESA,---,DOR
Select a user for user details or IDLEUSER for a list of idle users
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F10=Left  F11=Right  F12=Return
```

To Summarize:

PERFSVM is default user for Performance Toolkit
- Performance Toolkit installed on 4VMPTK40 minidisks

PERFKIT module invokes Performance Toolkit
- PROFILE EXEC access required disks
- Executes PERFKIT

Modes of operation:
- Basic mode
- Redisplay mode
- Monitor mode

Help facility
- Context sensitive

Performance Data Collection

```
                    ┌─────────────────────┐
                    │    Monitor data     │
                    └─────────────────────┘
                              │
          FCONTROL MONCOLL CPMON
                              ▼
        ┌──────────────────────┐        ┌──────────────────┐
        │ Collection and       │───────▶│  Performance     │
        │ analysis             │        │  displays        │
        └──────────────────────┘        └──────────────────┘
                              │
          FCONTROL MONCOLL REDISP
                              ▼
        ┌──────────────────────┐
        │  Redisplay buffer    │
        └──────────────────────┘
                              │
          FCONTROL MONCOLL PERFLOG
                              ▼
              ┌──────────────────┐
              │  History files   │
              └──────────────────┘
```
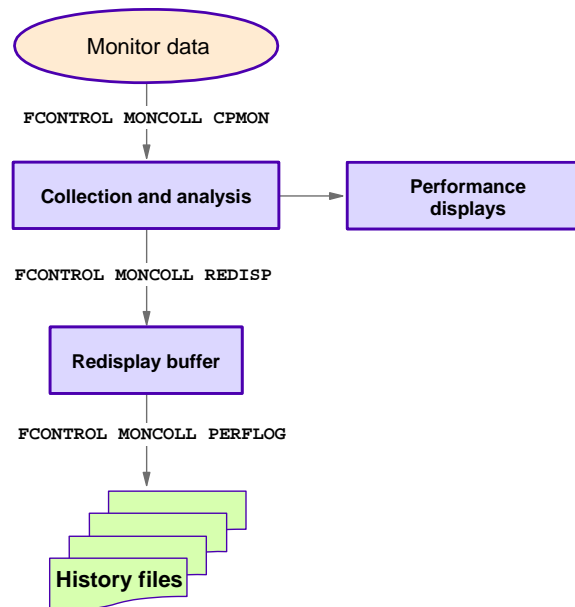
Controlling Data Collection

FC MONCOLL CPMON

- Controls whether data is collected from MONDCSS or disk

FC MONCOLL REDISP

- Sets number of lines used in the redisplay buffer

FC MONCOLL PERLOG

- Controls history data collection

Current vs. Average Data Reporting

Initially set to display current data
- Time interval appears in header field

CURRENT command displays current values
- Displayed data is value since last measurement interval
- Start/end times indicate interval

AVERAGE command displays average values
- Displayed data is average value since start or data reset
- FC MONCOLL RESET command

System Counters - CURRENT values

```
FCX102     CPU 2084  SER 96A3A  Interval 11:12:33 - 11:13:33   Perf. Monitor

Operation                  Count Rate/s  Operation                  Count Rate/s
Real SSCH instructions      206    3.4   Real CSCH instructions        0    .0
Real HSCH instructions        0    .0    El. time slice drops         22    .4
SVC instr. simulated          0    .0    SVC interrupts reflectd       0    .0
SVC 76 reflected              0    .0    Diagnose I/O requests        40    .7
FP external call simul.       0    .0    FP partial executions        72   1.2
Fast-path SIGP simulat.       0    .0    FP simul. of Diag.X'44'       0    .0
FP successful x-lates        76   1.3    CCW chains not FP-elig.       0    .0
Fast-path aborts             12    .2    Total FP xlate attempts      88   1.5
Nr. of SIE executions     19961  332     Nr. of SIE intercepts     19894  331
Entries to enabled wait   18277  304

Storage Management
Subpool FREE requests      2041   34.0   Total FREE requests        2045  34.1
V=R subpool FREE req.         0    .0    Storage fast clears          92   1.5
Avail. list frame req.      179   3.0    Available list empty          0    .0
Demand scan 1st pass          0    .0    Demand scan 2nd pass          0    .0
Demand scan emergency         0    .0    Demand scan not satisf.       0    .0
System stor. pgs taken        0    .0    Shared stor. pgs taken        0    .0
Dispatch lst pgs stolen       0    .0    Eligible lst pgs stolen       0    .0
Pgs from dormant users        0    .0    Pages taken for FREE          0    .0
Fast PGINs from XSTORE         0    .0    Slow PGINs from XSTORE         0    .0
PGOUTs main to XSTORE         0    .0    No XSTORE available           0    .0
XSTORE allocations            0    .0    XSTORE releases               0    .0
Glbl cycl list searched       0    .0    Migr. target time reset       0    .0
Migr thresh buf increas       0    .0    Migr thresh buf lowered       0    .0
Page migr. from dormant       0    .0    Dormant with page migr.       0    .0
```

System Counters - AVERAGE Values

```
FCX102     CPU 2084  SER 96A3A  Interval 08:30:33 - 11:15:33   Perf. Monitor

Operation                Count Rate/s  Operation                Count Rate/s
Real SSCH instructions   48423   4.4   Real CSCH instructions       2    .0
Real HSCH instructions       0    .0   El. time slice drops      4289    .3
SVC instr. simulated        10    .0   SVC interrupts reflectd      0    .0
SVC 76 reflected             0    .0   Diagnose I/O requests    10950   1.0
FP external call simul.      0    .0   FP partial executions    16729   1.5
Fast-path SIGP simulat.      0    .0   FP simul. of Diag.X'44'  74375   6.8
FP successful x-lates    19389   1.7   CCW chains not FP-elig.    142    .0
Fast-path aborts          2868    .2   Total FP xlate attempts  22399   2.0
Nr. of SIE executions   4.87E6   451   Nr. of SIE intercepts   4.77E6   441
Entries to enabled wait 3.28E6   303

Storage Management
Subpool FREE requests   1.56E6   145   Total FREE requests     1.57E6   145
V=R subpool FREE req.         0    .0   Storage fast clears     100113   9.2
Avail. list frame req.   153515  14.2   Available list empty         0    .0
Demand scan 1st pass          0    .0   Demand scan 2nd pass         0    .0
Demand scan emergency         0    .0   Demand scan not satisf.      0    .0
System stor. pgs taken        0    .0   Shared stor. pgs taken       0    .0
Dispatch lst pgs stolen       0    .0   Eligible lst pgs stolen      0    .0
Pgs from dormant users        0    .0   Pages taken for FREE         0    .0
Fast PGINs from XSTORE         0    .0   Slow PGINs from XSTORE       0    .0
PGOUTs main to XSTORE         0    .0   No XSTORE available          0    .0
XSTORE allocations            0    .0   XSTORE releases              0    .0
Glbl cycl list searched       0    .0   Migr. target time reset      0    .0
Migr thresh buf increas       0    .0   Migr thresh buf lowered      0    .0
Page migr. from dormant       0    .0   Dormant with page migr.      0    .0
```
FCXCMD511I AVERAGE data set, active for this display after next update

---

Sorting Reported Data

Most screens allow sorting by column

- Sortable colums indicated by "." above heading
- Current sort order indicated by "_" above heading

To set sort order:

- Move cursor to the column, press Enter
- Use SORT *column* command
  - Specify column name as parameter

Sorting by Disconnect Time

```
FCX108      CPU 2084  SER 96A3A  Interval 12:13:19 - 12:14:19    Perf. Monitor
.   .                      .      .      .    .  ___    .    .     .     .
<-- Device Descr. -->  Mdisk Pa- <-Rate/s-> <------- Time (msec) ------> Req.
Addr Type  Label/ID    Links ths  I/O Avoid Pend Disc Conn Serv Resp CUWt Qued
>> All DASD <<          ....        .1   .1   .2   .1  2.0  2.3  2.3   .0  .00
5092 CTCA  >PVM         ...   1    .0   ...  .0 3000   .1 3000 3000   .0  .00
5090 CTCA  >RSCS        ...   1    .3   ...  .1 2000   .3 2000 2000   .0  .00
1590 3390-3 LX4W02 CP    48   3   1.0   1.5  .2   .3  2.4  2.9  2.9   .0  .00
150C 3390-3 LX150C        0   3    .0    .0  .1   .1   .4   .6   .6   .0  .00
150E 3390-3 LX150E        0   3    .0    .0  .1   .1   .4   .6   .6   .0  .00
```

History Performance Analysis

Performance data can be written to log files

- Use FC MONCOLL PERFLOG ON
  - Time interval to save can be specified

History data files:

- Extended history data   *yyyymmdd* HISTLOG
  - 1 record per monitor inter
- Summary data   ACUM HISLOG
  - 1 record per hour

Up to 3 generations retained

- *yyyymmdd* HISTLOG
- *yyyymmdd* HISTLOG1
- *yyyymmdd* HISTLOG2

Extended Trend Files

Contain both:
- General system load data
- Detailed data for specific:
  - Users, processors, disks, Channels

Similar to VMPRF trend files
- Typically used for capacity planning
- Are not interchangeable with PRF data!

Control files:
- FCONX TRENDREC
- FCONX REPORTS

Redbooks

FCONX TRENDREC

```
*--------------------------------------------------------------*
* Performance Toolkit for VM Trend File Definition. This file  *
* defines the kind of records to be generated when a trend file*
* build request is processed                                   *
*--------------------------------------------------------------*

* General Trend Record Selection
  RECORDS   CHANNEL NSS DSPACES USER DASD SEEKS MTUSER TCPIP RSK

* Specific User Selection
* USERID    userid1 userid2  userid3  userid4  ...
  USERID    OPERATOR RSCS     TCPIP     VTAM     RSKSERV

* Specific DASD Selection (for DASD and SEEKS Records)
* DEVNO     devno1 devno2 devno3 devno4  ...
  DEVNO     0781 0782 0784 0786 0788
```

Redbooks

FCONX REPORTS

```
*------------------------------------------------------------------*
* Performance Toolkit for VM Report Selection File                 *
*                                                                  *
* File format:                                                     *
* - Asterisks in column 1 indicate a comment card                  *
*------------------------------------------------------------------*
*-General System Data----------------------------------------------*
  CPU
  STORAGE
  SYSTEM
*-I/O Device Data--------------------------------------------------*
  CACHDBSE
  CACHEXT
  DEVICE
  DEVICE CPOWN
*-User Data--------------------------------------------------------*
  UPAGE     (100  PAGESIZE 29
  USER      (100  SORT %CPU
  USTAT     (100
*-System Load by Time----------------------------------------------*
  AUXLOG
  LPARLOG
*-Benchmarking Logs for Specific I/O Devices and Users-----------*
```

Controlling Report Generation

FCONTROL MONCOLL RESET controls report generation

- FC MONCOLL RESET *hh:mmOption*

Options:

- P   Print reports specified in FCONX REPORTS
- T   Create trend records
- R   Reset counters

Benchmark Data Analysis

Provides detailed analysis for:
- Users
- I/O devices
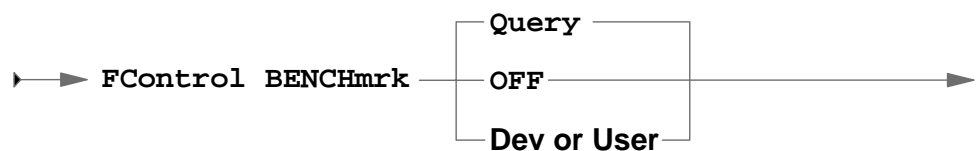
Use FCONTROL BENCHMRK command

Data collected each monitor interval (60 sec by default)
- Logs written to volatile storage by default
- Can be saved to DASD using FCONTROL BENCHMRK command
- Up to 3 generations of log files are retained

FCONTROL BENCHMRK Syntax

```
                              ┌─ Query ──────┐
    ►──► FControl BENCHmrk ───┼─ OFF ────────┼──►◄
                              └─ Dev or User ┘
```

Dev or User

```
    ┌─ DEVice xxxx ─┐   ┌─ OFF ─────────────┐
  ──┤               ├───┤                   ├──
    └─ USEr userid ─┘   └─ FILE hh:mm ── hh:mm ─┘
                              └─── TO ───┘
```
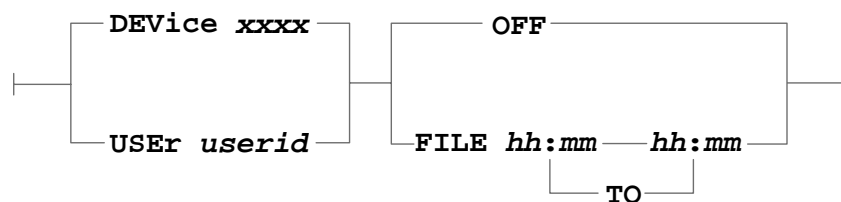
History Data Selection Options

```
FCX124         Performance Screen Selection  (FL440 VM63447)    Perf. Monitor

General System Data       I/O Data              History Data (by Time)
1. CPU load and trans.    11. Channel load      31. Graphics selection
2. Storage utilization    12. Control units     32. History data files*
3. Storage subpools       13. I/O device load*  33. Benchmark displays*
4. Priv. operations       14. CP owned disks*    34. Correlation coeff.
5. System counters        15. Cache extend. func.* 35. System summary*
6. CP IUCV services       16. DASD I/O assist   36. Auxiliary storage
7. SPOOL file display*    17. DASD seek distance* 37. CP communications*
8. LPAR data              18. I/O prior. queueing* 38. DASD load
9. Shared segments        19. I/O configuration 39. Minidisk cache*
A. Shared data spaces     1A. I/O config. changes 3A. Paging activity
B. Virt. disks in stor.                         3B. Proc. load & config*
C. Transact. statistics   User Data             3C. Logical part. load
D. Monitor data           21. User resource usage* 3D. Response time (all)*
E. Monitor settings       22. User paging load* 3E. RSK data menu*
F. System settings        23. User wait states* 3F. Scheduler queues
G. System configuration   24. User response time* 3G. Scheduler data
H. VM Resource Manager    25. Resources/transact.* 3H. SFS/BFS logs menu*
                          26. User communication* 3I. System log
I. Exceptions             27. Multitasking users* 3K. TCP/IP data menu*
                          28. User configuration* 3L. User communication
K. User defined data*     29. Linux systems*    3M. User wait states

         Pointers to related or more detailed performance data
         can be found on displays marked with an asterisk (*).
Select performance screen with cursor and hit ENTER
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

ibm.com/redbooks

Graphics Selection Menu

```
FCX128                    Graphics Selection Menu              Perf. Monitor
General Specifications

Format         : GDDM                 Truncate X-var. at:
Data origin    : FILE 20040317 HISTLOG2 B    Y-var. cumulative (Y/N): N
Graphics type  : DETailed history     Scatter plot (Y/N): Y
Selected period: FROM 14:00
Selected hours :
Selected days  : All days

Variables Selection            Trend Record Selection
Var  Name      Description      Display   Record Details
X  : ACT      - Active users
Y *: CPU      - % total CPU
  o : VIO/S    - Virtual I/O rate
  = :          -
  - :          -


To select graphics
- either key in your choice directly in the fields above
- or enter '?' in the first position of any field for which you
  want additional information. A detailed selection menu will then
  be shown with further explanations.
Hitting ENTER without any changes will start graphics creation.
```

ibm.com/redbooks

Graphics Display Using GDDM

```
LX4-32                                                                    _|□|×|
File  Edit  View  Communication  Actions  Window  Help
FCX196        CPU 2084   SER 96A3A   GDDM Graphic History Display    Perf. Monitor

                      % total CPU        Virtual I/O rate

   6.0 %CPU
   800 IO/s

   3.0 %CPU
   400 IO/s

     0
        14:00        14:30         15:00         15:30         16:00

                  Source data:  20040317 HISTLOG2 B


    F6=Save   F9=Print  F12=Return                    File name  CONMON    Copies 1
MA□    c                                                                    32/062
     Connected to remote server/host 9.12.4.17 using port 23        IBM4076E on
```

History Data Files Selection Screen

```
FCX160        File Selection: All Files with Default Filetypes      Perf. Monitor

        Move cursor to the history data file you are interested in and
        - enter 'G' to select the file for creating graphics
        - just hit ENTER for viewing the data

   .         .       .    _____       .        .      .    _____
 S Filename Filetype FM  Created      S Filename Filetype FM  Created
 . ACUM     HISTSUM  A   2004/03/25   . 20040325 HISTLOG  A   2004/03/25
 . 20040324 HISTLOG1 A   2004/03/24   . 20040324 HISTLOG1 D   2004/03/24
 . 20040324 HISTLOG  D   2004/03/24   . 20040323 HISTLOG1 D   2004/03/23
 . 20040323 HISTLOG2 A   2004/03/23   . 20040322 HISTLOG  D   2004/03/22
 . 20040318 HISTLOG1 D   2004/03/18   . 20040317 HISTLOG2 D   2004/03/17
 . LNXSU1   UCOMMLG1 D   2004/03/12   . LNXSU1   USTATLG1 D   2004/03/12
 . LNXSU1   UPAGELG1 D   2004/03/12   . LNXSU1   USERLOG1 D   2004/03/12
 . 20040312 HISTLOG1 D   2004/03/12   . LNXSU1   UCOMMLG2 D   2004/03/11
 . LNXSU1   USTATLG2 D   2004/03/11   . LNXSU1   UPAGELG2 D   2004/03/11
 . LNXSU1   USERLOG2 D   2004/03/11




Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

Extended Data History Log Screen

```
FCX195              Input File: '20040324 HISTLOG1 A'          Perf. Monitor

Date       Time  <--Interval--->     <-------------------- CPU Load (%) -
yyyy/mm/dd hh:mm El_Time Samples #CPU  CPU  %US  %CP  %EM  %SY %Spin   %WT
2004/03/24 19:51   20.0       1  2.0  3.1  2.2  1.6  1.5   .9    .0    197
2004/03/24 19:51   20.0       1  2.0  3.0  2.1  1.6  1.4   .8    .0    197
2004/03/24 19:51   20.0       1  2.0  2.8  2.0  1.5  1.3   .8    .0    197
2004/03/24 19:52   20.0       1  2.0  3.1  2.2  1.6  1.5   .9    .0    197
2004/03/24 19:52   20.0       1  2.0  3.0  2.1  1.6  1.3   .8    .0    197
2004/03/24 19:52   20.0       1  2.0  2.9  2.1  1.5  1.4   .8    .0    197
2004/03/24 19:53   20.0       1  2.0  3.3  2.5  1.6  1.8   .8    .0    197
2004/03/24 19:53   20.0       1  2.0  3.0  2.1  1.6  1.3   .9    .0    197
2004/03/24 19:53   20.0       1  2.0  2.8  1.9  1.5  1.3   .9    .0    197
2004/03/24 19:54   20.0       1  2.0  2.9  2.1  1.6  1.3   .9    .0    197
2004/03/24 19:54   20.0       1  2.0  3.0  2.1  1.7  1.4   .9    .0    197
2004/03/24 19:54   20.0       1  2.0  2.8  2.0  1.5  1.3   .8    .0    197
2004/03/24 19:55   20.0       1  2.0  2.9  2.1  1.6  1.3   .8    .0    197
2004/03/24 19:55   20.0       1  2.0  3.0  2.1  1.6  1.3   .9    .0    197
2004/03/24 19:55   20.0       1  2.0  2.8  2.0  1.5  1.3   .9    .0    197
2004/03/24 19:56   20.0       1  2.0  3.3  2.4  1.6  1.7   .8    .0    197
2004/03/24 19:56   20.0       1  2.0  3.0  2.2  1.6  1.4   .9    .0    197
2004/03/24 19:56   20.0       1  2.0  2.8  2.0  1.5  1.3   .8    .0    197
2004/03/24 19:57   20.0       1  2.0  3.1  2.2  1.6  1.5   .8    .0    197
2004/03/24 19:57   20.0       1  2.0  3.0  2.1  1.6  1.4   .8    .0    197

                        *** End of File ***


Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F10=Left  F11=Right  F12=Return
```

Benchmark Log Selection Screen

```
FCX173      CPU 2084  SER 96A3A       BENCHMRK Log Data         Perf.
_____      .
  Userid    Log File
S Devnum    Name      Description
. LNXSU1    UCOMMLOG  User IUCV and VMCF communications log
. LNXSU1    UPAGELOG  User paging load log
- LNXSU1    USERLOG   User resource consumption log
. LNXSU1    USTATLOG  User wait state log
. LNXSU3    UCOMMLOG  User IUCV and VMCF communications log
. LNXSU3    UPAGELOG  User paging load log
. LNXSU3    USERLOG   User resource consumption log
. LNXSU3    USTATLOG  User wait state log
. LNXSU4    UCOMMLOG  User IUCV and VMCF communications log
. LNXSU4    UPAGELOG  User paging load log
. LNXSU4    USERLOG   User resource consumption log
. LNXSU4    USTATLOG  User wait state log
. 1510      CACHDLOG  I/O device CU cache data log
. 1510      CACHELOG  Extended function cache data log
. 1510      CPOWNLOG  CPOWNed I/O device log
. 1510      DEVLOG    General I/O device performance log


Select a user or device log with cursor and hit ENTER
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

Correlation Coefficients

```
FCX194              Input File: '20040323 HISTLOG  *'          Perf. Monitor

 Correl. Coefficients for Base Variable CPU (% total CPU    , 1892 Samp)

 Selection: From (not set)    Days : All days
            To   (not set)    Hours: All hours
 Corre-  Variable Variable               Corre-  Variable Variable
 lation  Name     Description            lation  Name     Description
 1.000   %US      % user CPU              .999   %LogLd   % logical CPU
  .997   %EM      % emulation CPU         .941   Max. CPU MaxUser: TotCPU
  .885   CPU/TR   CPU MSEC/TR             .734   Conn     Connected time
  .721   Serv     DASD serv. time         .721   Resp     DASD resp. time
  .708   InvR/s   Inval. request/s        .702   SendQ    User IUCV send Q
  .690   Busy     Ave. DASD busy          .643   %Run     % Running
  .448   DIAG     DIAG instr./s           .444   MDR/S    MDC read req/s
  .444   Avoid/s  I/O avoided rate        .444   MDH/s    MDC read hits/s
  .441   %CP      % supervisor CPU        .441   FrSav    Save area pages
  .420   IO/S     SSCH/RSCH rate          .418   DASD/s   DASD I/O rate
  .415   Disc     Disconnect time         .407   PGIN/s   PAGIN rate/s
  .406   Alo/s    Page alloc. rate        .404   FEx/s    Frame extends/s
 Select new base variable to recompute
 Command ===>
 F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

ibm.com/redbooks

To Summarize:

Performance Toolkit can collect / analyze:
- Real-time
- History
- Trend
- Benchmark

Data can be displayed as:
- CURRENT values
- AVERAGE values

Data can be sorted:
- Look for sort indicators on field headings

ibm.com/redbooks

Remote Performance Monitoring

Performance Toolkit can act as central data collector

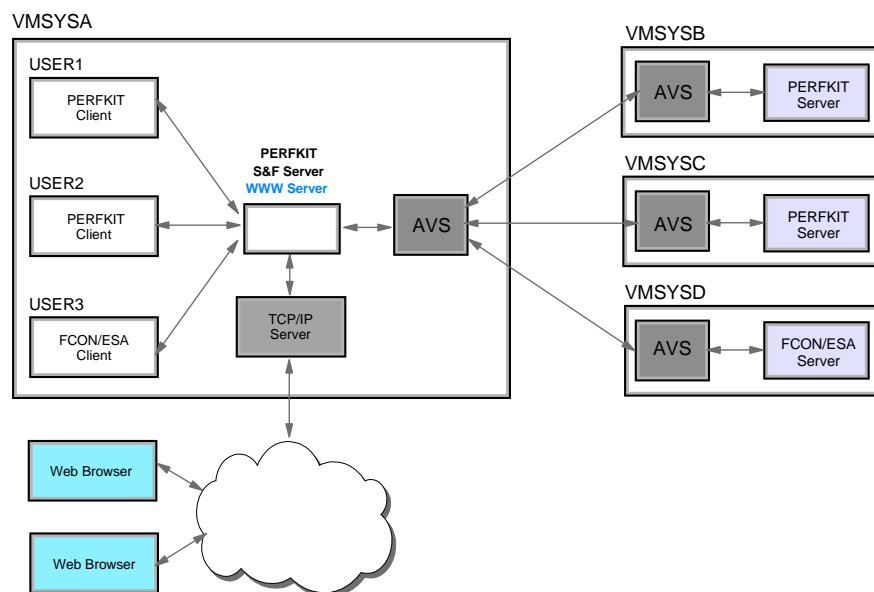- Store and Forward (S&F) server

Remote access methods include:

- APPC/VM
  - Utilizes IUCV resources
  - Uses APPC/VTAM Support (AVS)
- Web interface
  - Relies on APPC/VM
  - Access through standard Web browser
- RMF PM interface
  - Suitable for monitoring Linux guests
  - Access through Web browseror Java application

Remote Monitoring Facilites

Remote Monitoring Configuration Files

FCONRMT AUTHORIZ

- Enables data access and command execution

FCONRMT SYSTEMS

- Defines remote server to monitor

FCONRMT PASSFILE

- Defines client users and passwords
  - Configurable authentication methods

FCONX LINUXUSR

- Defines Linux guests to monitor

ibm.com/redbooks

The FCONX AUTHORIZ File

```
**************************************************************
*Authorization file for local and remote performance data *
*retrieval and command execution                          *
**************************************************************
*Need to setup as a S&F server for APPC connections.       *
**************************************************************
*Node-id User-id  Authorized for
*   |       |        |
NODE1     PERFSVM  S&FSERV
NODE2     PERFSVM  CMD DATA
**************************************************************
*The following allows ANYONE from the specified SYSTEMS   *
*to request DATA from the ID running PERFKIT               *
**************************************************************
NODE1     *         DATA
**************************************************************
* YOU ARE RESONSIBLE FOR WHO YOU ALLOW TO EXECUTE COMMANDS*
*The following ID's are allowed to request DATA and also  *
*execute COMMANDS on the ID running PERFKIT                *
**************************************************************
*Allow me CMD and DATA access
NODE1     MYUSERID CMD DATA
*Allow my backup CMD and DATA access
NODE1     MYBACKUP CMD DATA
*Allow some users DATA and some command access
NODE1     USER1    DATA CPQRY
******************************
```

ibm.com/redbooks

The FCONRMT SYSTEMS File

```
***********************************************************
* System Definition File for Remote Monitoring            *
***********************************************************
*Node-ID PERFKIT-ID VMTYPE    APPEND  Nickname
NODE1    PERFSVM    z/VM_440 N
NODE2    PERFSVM    z/VM_440 N
NODE3    PERFSVM    z/VM_440 N         FCXRES03
NODE4    FCONX      ESA
```

Redbooks **ibm.com**/redbooks

The FCONRMT PASSFILE

```
*User-ID         Password
USER1           PASSWORD1
USER2           PASSWORD2
USER3           PASSWORD3
```

Redbooks **ibm.com**/redbooks

The FCONX LINUXUSR File

```
***********************************************************************
*   INITIALIZAITON FILE WITH IP ADDRESS DEFINITIONS FOR LINUX
*   SYSTEMS THAT MAY BE MONITORED
*
*LINUX ID   IP ADDRESS FOR DDS INTERFACE
LINUX01    111.111.111:8080
LINUX02    222.222.222:8080
LINUX03    111.222.222:8080
LINUXTST   111.111.222:8080
```

APPC/VM Interface to Performance Toolkit

Two modes of operation:
- Retrieval from local PERFSVM machine
- Retrieval from remote PERFSVM machine

Relies on global resource names
- FCXRES00 by default
- FCXSYSTM for remote systems

Activated and accessed using:
- FCONTROL MONCOLL VMCF ON command
- FCONAPPC command
  - For local PERFSVM machine
- FCONAPPC FCXSYSTM command
  - For remote PERFSVM machines

Using Default APPC/VM Resource Name

Ensure resource is defined to PERFSVM user
- IUCV *IDENT FCXRES00 GLOBAL
- IUCV ALLOW

Define authorized users in FCONRMT AUTHORIZ file
- FC RELOAD AUTHORIZ to reload authorization rules

Activate APPC/VM interface in FCONX $PROFILE
- FC MONCOLL VMCF ON
- On startup, look for:
  - `FCXAPP5300 Connected to *IDENT for resource FCXRES00`
- FCONAPPC command
  - Establishes connection to default resource name

Redbooks

Using an Alternate Resource Name

Needed for APPC/VM access to remote systems
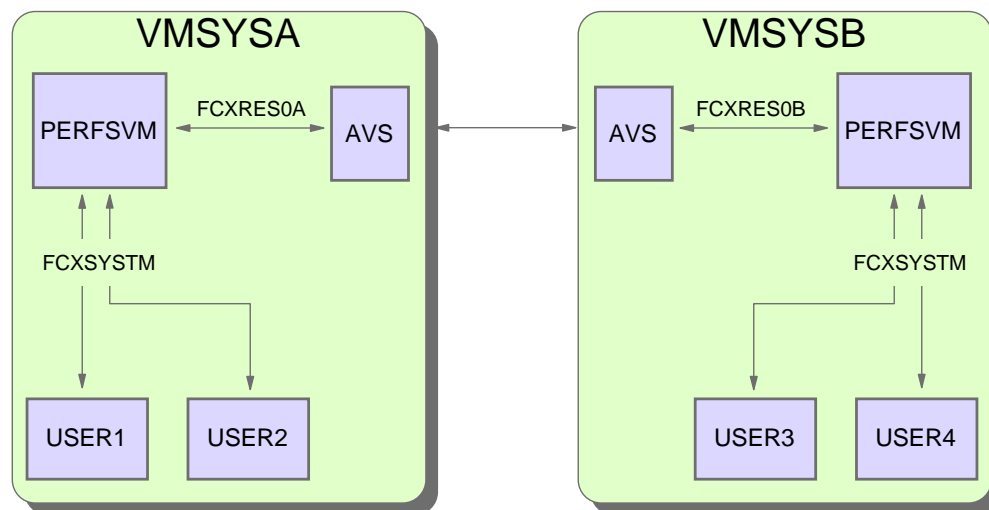- Unique global names are required

Steps:
- Assign unique name to each PERFSVM user directory:
  - IUCV *IDENT FCXRES0A GLOBAL
  - IUCV ALLOW
- Map new name to FCXRES00 in SCOMDIR NAMES file:
  - :NICK.FCXRES00   :LUNAME.*IDENT
                        :TPN.FCXRES0A
                        :SECURITY.SAME
- On startup, look for:
  - `FCXAPP5300 Connected to *IDENT for resource FCXRES0A`
- Use FCONAPPC command to connect

Redbooks

APPC/VM Interface for Remote Access

VMSYSA

PERFSVM ←—— FCXRES0A ——→ AVS ←——→

FCXSYSTM

USER1   USER2

VMSYSB

←——→ AVS ←—— FCXRES0B ——→ PERFSVM

FCXSYSTM

USER3   USER4

APPC/VM for Remote Systems

Remote connections exist between PERFSVM servers

- Users connect to local PERFSVM using FCXSYSTM resource

Steps:

- Use FCXSYSTM resource in PERFSVM user directory entry
  - IUCV *IDENT FCXSYSTM GLOBAL
- Use unique global resource names
  - See Using an Alternate Resource Name
- Define remote systems and authorizations
  - FCONRMT SYSTEMS and FCONRMT AUTHORIZ files
- Issue FCONAPPC FCXSYSTM command

System Selection Menu

```
FCX193      CPU 2084  SER 96A3A    (Intermediate Server)        VMSYSA

                        Systems Accessible via APPC/VM

    System    Resource        System    Resource        System    Resource
    VMSYSA    FCXRES00        VMSYSB    FCXRES0B        VMSYSC    FCXRES0C
    VMITSO    FCXRESBE




 Select the system to be monitored
 Command ===>
 F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

Web Interface to Performance Toolkit

Performance Toolkit screens accessible from a Web browser

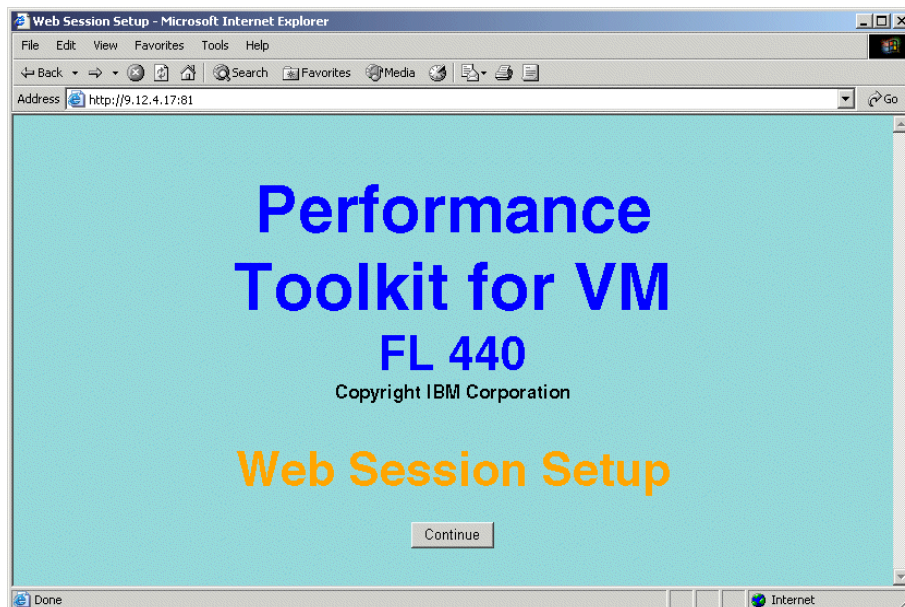- PERFSVM must be configured as Store and Forward server

Features:

- Remote access to Performance Toolkit screens
  - No 3270 connection required
- Graphic displays without GDDM
  - Uses Java applet
- Flexible authentication
  - Standard VM userid / password
  - RACF authentication
  - FCONRMT PASSFILE
  - No authentication

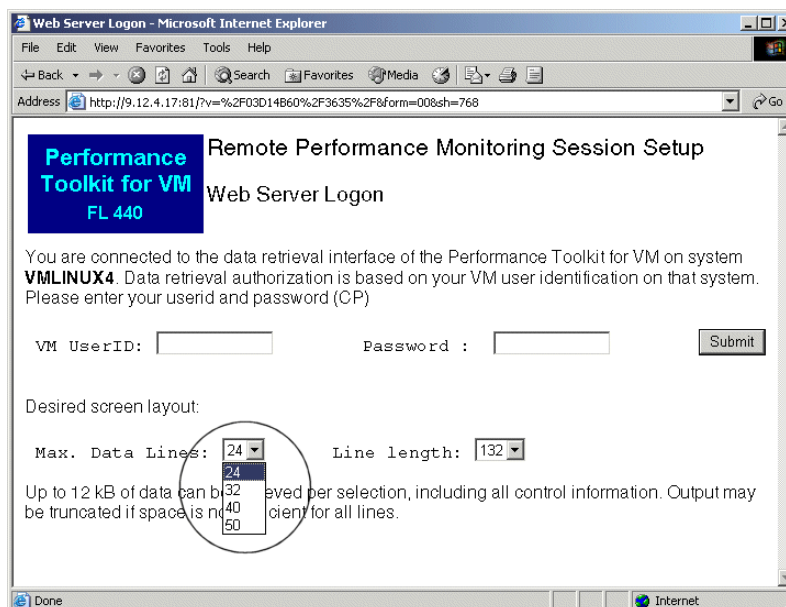# Web Interface Welcome Page

Web Server Logon Page

Web Interface Selection Menu

VMLINUX4 Data Retrieval Session (Performance Toolkit for VM FL440 VM63419) - Microsoft Internet Ex...

File   Edit   View   Favorites   Tools   Help

Back   Search   Favorites   Media

Address http://9.12.4.17:81/03C81780/34C9/VMLINUX4

Performance Toolkit for VM FL 440

Initial Performance Data Selection Menu   (**VMLINUX4**)
Select performance screen

Command   Refresh   Systems   Forw   Help   Auto-Refresh

General System Data
1. CPU load and trans.
2. Storage utilization
3. Storage subpools
4. Priv. operations
5. System counters
6. CP IUCV services
7. SPOOL file display*
8. LPAR data
9. Shared segments
A. Shared data spaces
B. Virt. disks in stor.
C. Transact. statistics
D. Monitor data
E. Monitor settings
F. System settings
G. System configuration
H. VM Resource Manager

I. Exceptions

K. User defined data*

I/O Data
11. Channel load
12. Control units
13. I/O device load*
14. CP owned disks*
15. Cache extend. func.*
16. DASD I/O assist
17. DASD seek distance*
18. I/O prior. queueing*
19. I/O configuration
1A. I/O config. changes

User Data
21. User resource usage*
22. User paging load*
23. User wait states*
24. User response time*
25. Resources/transact.*
26. User communication*
27. Multitasking users*
28. User configuration*
29. Linux systems*

History Data (by Time)
31. Graphics selection
32. History data files*
33. Benchmark displays*
34. Correlation coeff.
35. System summary*
36. Auxiliary storage
37. CP communications*
38. DASD load
39. Minidisk cache*
3A. Paging activity
3B. Proc. load & config*
3C. Logical part. load
3D. Response time (all)*
3E. RSK data menu*
3F. Scheduler queues
3G. Scheduler data
3H. SFS/BFS logs menu*
3I. System log
3K. TCP/IP data menu*
3L. User communication
3M. User wait states

Internet

Graphics Selection Web Page

VMLINUX4 Data Retrieval Session (Performance Toolkit for VM FL440 VM63447) - Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

Back   Search   Favorites   Media

Links IBM Business Transformation   IBM Internal Help   IBM Standard Software Installer   Customize Links   Free Hotmail   IBM Standard Software   Windows Media

Address http://9.12.4.17:8088/0366EFF8/200B/FC/DATA?v=%2F0366EFF8%2F200B%2FB/orm=038f=Line+graphics&s=Validate&i=Storage+++++++++++++++++&t=Detailed

Performance Toolkit for VM FL 440

Graphics Selection Menu   (**VMLINUX4**)

Command   Refresh   Systems   Menu   Return   Help   Auto-Refresh

General Specifications

Output format   Line graphics   Validate   Submit

Data origin   Storage

Graphics type   Detailed graphics (detailed time scale)

Selected period Last measurements

Selected days   All days

Selected hours   All hours

Variable Selection

X-Variable

   Truncate at

Y-Variables   1 CPU   % total CPU
              2 IO/S   SSCH rate/s
              3 VIO/S   Virt. IO/sec
              4 PG/S   Page rate

Cumulative

Done   Internet

Line Graphics in the Web Interface

Enabling the Web Interface

Configure PERFSVM as Store and Forward server

- See APPC Interface to Performance Toolkit

Connect PERFSVM to TCPIP machine

- Choose TCP/IP port number in PROFILE TCPIP:
  - PORT
    - 81 TCP PERFSVM   NOAUTOLOG;

Enable Web interface

- FCONTROL MONCOLL WEBSERV command
  - Specify configured port number

Access Performance Toolkit from Web browser

- http://192.168.100.1:81

Web Interface Authentication

Controlled by IDTEST option of FC MONCOLL WEBSERV

Methods:

- Standard VM userid / password
  - FC MONCOLL WEBSERV ON TCPIP TCPIP 81 IDTEST CP
- RACF
  - FC MONCOLL WEBSERV ON TCPIP TCPIP 81 IDTEST RACF
- FCONRMT PASSFILE
  - FC MONCOLL WEBSERV ON TCPIP TCPIP 81 IDTEST FILE
- None
  - FC MONCOLL WEBSERV ON TCPIP TCPIP 81 IDTEST OFF

Redbooks

RMF PM Interface to Linux Guests

Uses Distributed Data Server (DDS) to collect Linux data

- Runs in Linux guest as server process
- Feeds data to Performance Toolkit
  - Uses HTTP transactions

Available as separate component:

- http://www.ibm.com/servers/eserver/zseries/zos/rmf/rmfhtmls/pmweb/pmlim.htm
- Java client is also available

Be aware:

- Monitoring Linux guests consume CPU cycles
- Reported data is relative to Linux virtual machine (not VM system)

Redbooks

Data Collection Using DDS

PERFSVM

TCP/IP

DDS	DDS	DDS	DDS
Linux1	Linux2	Linux3	Linuxn

Redbooks

Linux Systems Option

VMLINUX4 Data Retrieval Session (Performance Toolkit for VM FL440 VM63406) - Microsoft Internet Ex...

File Edit View Favorites Tools Help

Back → ⊗ ⊗ ⌂ Search Favorites Media ⊗ ⊗ ⊗ ▼ ⊗

Address http://9.12.4.17:8081/03DF9200/1811/VMLINUX4 ▼ Go

Performance Toolkit for VM FL 440

Initial Performance Data Selection Menu (**VMLINUX4**)
Select performance screen

Command Refresh Systems Help ☐ Auto-Refresh

```
General System Data          I/O Data                    History Data (by Time)
1.  CPU load and trans.     11.  Channel load            31.  Graphics selection
2.  Storage utilization     12.  Control units           32.  History data files*
3.  Storage subpools        13.  I/O device load*        33.  Benchmark displays*
4.  Priv. operations        14.  CP owned disks*         34.  Correlation coeff.
5.  System counters         15.  Cache extend. func.*    35.  System summary*
6.  CP IUCV services        16.  DASD I/O assist         36.  Auxiliary storage
7.  SPOOL file display*     17.  DASD seek distance*     37.  CP communications*
8.  LPAR data               18.  I/O prior. queueing*    38.  DASD load
9.  Shared segments         19.  I/O configuration       39.  Minidisk cache*
A.  Shared data spaces      1A.  I/O config. changes     3A.  Paging activity
B.  Virt. disks in stor.                                 3B.  Proc. load & config*
C.  Transact. statistics    User Data                    3C.  Logical part. load
D.  Monitor data            21.  User resource usage*    3D.  Response time (all)*
E.  Monitor settings        22.  User paging load*       3E.  RSK data menu*
F.  System settings         23.  User wait states*       3F.  Scheduler queues
G.  System configuration    24.  User response time*     3G.  Scheduler data
H.  VM Resource Manager     25.  Resources/transact.*    3H.  SFS/BFS logs menu*
                            26.  User communication*     3I.  System log
I.  Exceptions              27.  Multitasking users*     3K.  TCP/IP data menu*
                            28.  User configuration*     3L.  User communication
K.  User defined data*      29.  Linux systems*          3M.  User wait states

          Pointers to related or more detailed performance data
          can be found on displays marked with an asterisk (*).
```

http://9.12.4.17:8081/03DF9200/BEAE/3F Internet

Redbooks

Linux Memory Usage (LXMEM)
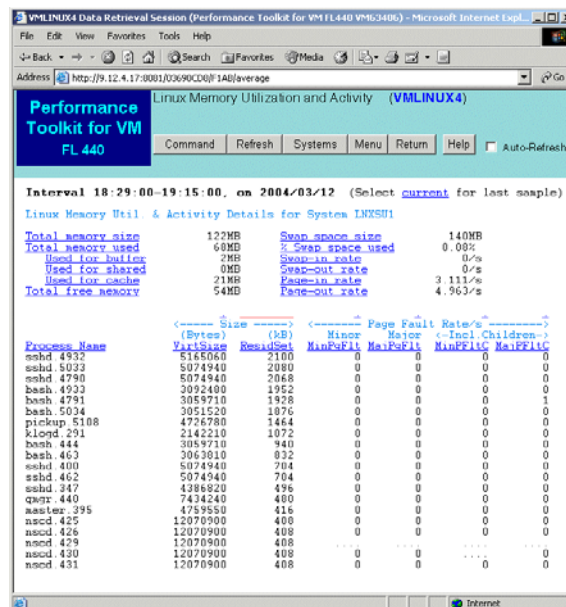
RMF DDS Web Interface
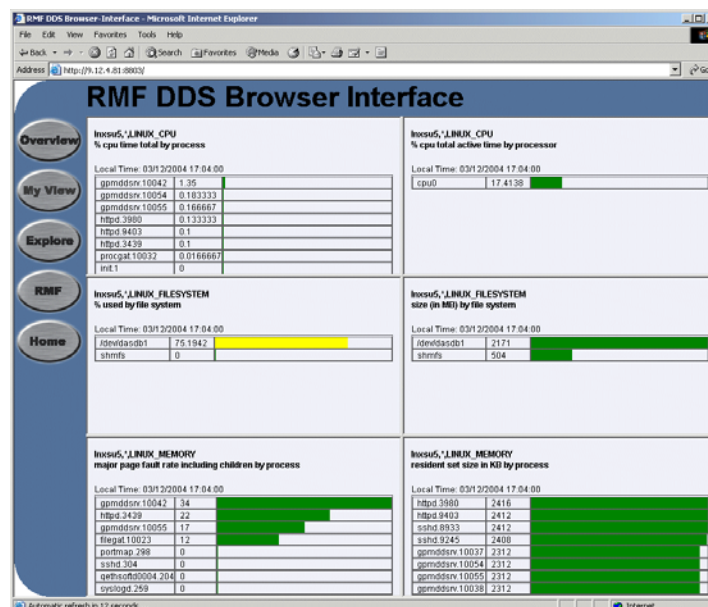
DDS Java Client
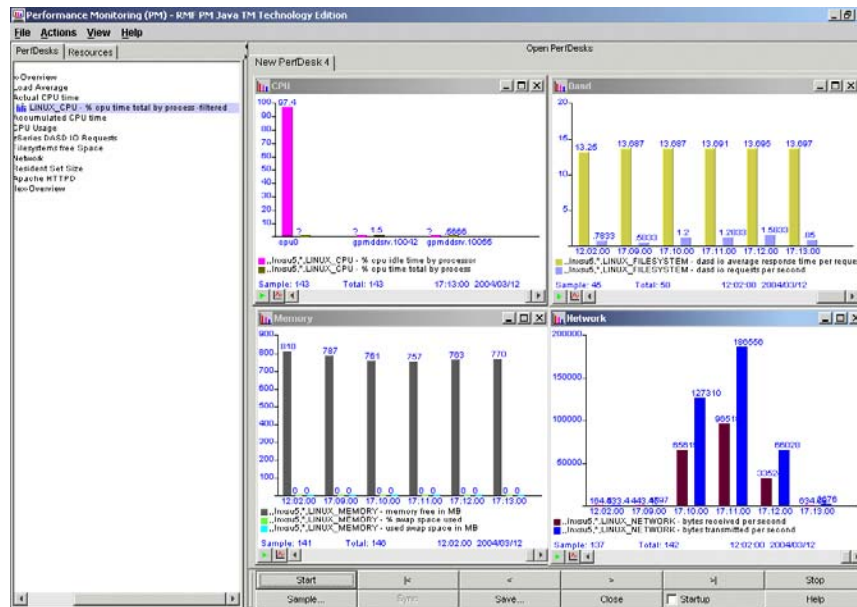
To Summarize:

Performance Toolkit can monitor multiple VM systems

- Relies on APPC/VM interface
  - Default resources defined to PERFSVM user

Web interface allows remote monitoring

- Requires:
  - APPC/VM interface
  - Configuration of Web server component

DDS interface to Linux guests

- Allows integrated view of Linux performance data
- Reported data relative to Linux virtual machine

Some Important Real-time Displays

```
FCX124        Performance Screen Selection  (FL440 VM63447)    Perf. Monitor

General System Data      I/O Data                   History Data (by Time)
1. CPU load and trans.   11. Channel load           31. Graphics selection
2. Storage utilization   12. Control units          32. History data files*
3. Storage subpools      13. I/O device load*       33. Benchmark displays*
4. Priv. operations      14. CP owned disks*         34. Correlation coeff.
5. System counters       15. Cache extend. func.*   35. System summary*
6. CP IUCV services      16. DASD I/O assist        36. Auxiliary storage
7. SPOOL file display*   17. DASD seek distance*    37. CP communications*
8. LPAR data             18. I/O prior. queueing*   38. DASD load
9. Shared segments       19. I/O configuration      39. Minidisk cache*
A. Shared data spaces    1A. I/O config. changes    3A. Paging activity
B. Virt. disks in stor.                             3B. Proc. load & config*
C. Transact. statistics  User Data                  3C. Logical part. load
D. Monitor data          21. User resource usage*   3D. Response time (all)*
E. Monitor settings      22. User paging load*      3E. RSK data menu*
F. System settings       23. User wait states*      3F. Scheduler queues
G. System configuration  24. User response time*    3G. Scheduler data
H. VM Resource Manager   25. Resources/transact.*   3H. SFS/BFS logs menu*
                         26. User communication*    3I. System log
I. Exceptions            27. Multitasking users*    3K. TCP/IP data menu*
                         28. User configuration*    3L. User communication
K. User defined data*    29. Linux systems*         3M. User wait states

            Pointers to related or more detailed performance data
            can be found on displays marked with an asterisk (*).
Select performance screen with cursor and hit ENTER
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

CPU Load Screen - Option 1 (CPU)

```
FCX100      CPU 2084   SER 96A3A  Interval 17:18:19 - 17:18:39    Perf. Monitor

CPU Load                                          Vector Facility   Status or
PROC  %CPU  %CP %EMU  %WT %SYS  %SP %SIC %LOGLD   %VTOT %VEMU REST  ded. User
P00    21    0   21   79    0    0   63   23       not installed    Master
P01    51    0   51   49    0    0   31   65       not installed    Alternate

Total SSCH/RSCH     5/s    Page rate        .0/s   Priv. instruct.   30/s
Virtual I/O rate    3/s    XSTORE paging    .0/s   Diagnose instr.   26/s
Total rel. SHARE   3200    Tot. abs SHARE     6%

Queue Statistics:    Q0    Q1    Q2    Q3    User Status:
VMDBKs in queue       3     1     0     2    # of logged on users      25
VMDBKs loading        0     0     0     0    # of dialled users         0
Eligible VMDBKs             0     0     0    # of active users          8
El. VMDBKs loading          0     0     0    # of in-queue users        6
Tot. WS (pages)    5109   153     0 30984    % in-Q users in PGWAIT     0
Expansion factor           1     1     1    % in-Q users in IOWAIT    30
85% elapsed time  112.2 14.03 112.2 673.3   % elig. (resource wait)    0

Transactions     Q-Disp   trivial   non-trv   User Extremes:

Command ===>
```

CPU Command

Displays per processor utilization

- %CPU   Total CPU load
- %CP   Time spent in supervisor state
- %EMU   Time spent under SIE (vritual CPU time)
- %WT   Time spent in wait state
- %SYS   Time spent executing system services
- %LOGLD Logical load - based on active processor time / wait time

I/O rates

- Virtual I/O rate
- Page / XSTORE rate

Queue statistics

Storage Utilization - Option 2 (STORAGE)

```
FCX103     CPU 2084  SER 96A3A  Interval 12:12:33 - 12:13:33    Perf. Monitor

Main storage utilization:              XSTORE utilization:
Total real storage          3'072MB    Total available              1'024MB
Total available             3'072MB    Att. to virt. machines           0kB
Offline storage frames          0kB    Size of CP partition         1'024MB
SYSGEN storage size         3'072MB    CP XSTORE utilization            0%
.
.
.
V=R user                  ........     Users with MDCACHE inserts       0
                                       MDISK cache read rate          0/s
Paging / spooling activity:            MDISK cache write rate       ...../s
Page moves <2GB for trans.     0/s     MDISK cache read hit rate      0/s
Fast path page-in rate         0/s     MDISK cache read hit ratio     100%
Long path page-in rate         0/s
Long path page-out rate        0/s     VDISKs:
Page read rate                 2/s     System limit (blocks)       Unlim.
Page write rate                0/s     User limit (blocks)         500000
Page read blocking factor       34     Main store page frames          30
Page write blocking factor     ...     Expanded stor. pages             0
.
.
Enter 'FREesub' command for Free Storage Subpool details
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

USER Resources - Option 21 (USER)

```
FCX112     CPU 2084  SER 96A3A  Interval 14:41:47 - 14:42:07    Perf. Monitor
.          ___    .      .      .     .     .     .     .    .     .
           <----- CPU Load -----> <------ Virtual IO/s ------>
           <-Seconds->   T/V
Userid     %CPU  TCPU   VCPU  Ratio Total DASD Avoid Diag98   UR  Pg/s  User Status
>System<   3.34  .668   .508   1.3   102  101   141     .0   .0   .0   ---,---,---
LNXSU1     51.6 10.32  8.855   1.2  2267 2267  2267     .0   .0   .0   ESA,CL3,DIS
LNXRH3     24.0  4.796 4.066   1.2   4.7  4.7  1429     .0   .0   .0   ESA,CL3,DIS
LNXSU2     15.2  3.042 1.511   2.0    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXRH1     14.1  2.825 2.333   1.2  1381 1376  1376     .0   .0   .0   XC, CL0,DIS
TCPIP      13.0  2.598 1.165   2.2    .0   .0    .0     .0   .0   .0   ESA,CL0,DIS
TCPMAINT    .27  .054  .050    1.1   1.2  1.2    .0     .0   .0   .0   ESA,CL3,DIS
DIRMAINT    .23  .046  .033    1.4    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXSU6      .22  .044  .030    1.5    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXRH2      .22  .044  .031    1.4    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXRH4      .22  .044  .031    1.4    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXSU3      .21  .041  .027    1.5    .1   .1    .0     .0   .0   .0   ESA,CL3,DIS
LNXSU4      .21  .041  .028    1.5    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
LNXSU5      .19  .037  .023    1.6    .0   .0    .0     .0   .0   .0   ESA,CL3,DIS
PERFKLA     .12  .024  .023    1.0    .4   .3    .2     .0   .0   .0   ESA,---,DOR
PERFLIV     .11  .022  .021    1.0    .4   .3    .1     .0   .0   .0   ESA,---,DOR
PERFROG     .11  .022  .020    1.1    .7   .6    .3     .0   .0   .0   ESA,CL0,DIS
PERFSVM     .10  .020  .019    1.1    .4   .3    .1     .0   .0   .0   ESA,---,DOR
```

Scheduler Settings

Scheduler settings affect how machines are dispatched

- Default settings give priority to low resource consumers
- Long running, high resource tasks wait longer on average
- Virtual machine classification:
  - Q0   special (do not wait)
  - Q1   interactive
  - Q2   non-interactive
  - Q3   batch

Scheduler parameters:

- STORBUF
  - Projected working set size
- LDUBUF
  - Projected paging usage

System Settings - Option F (SYSSET)

```
FCX154      CPU 2084  SER 96A3A       System   Settings         Perf. Monitor

Initial Scheduler Settings: 2004/03/03 at 17:01:52
DSPSLICE (minor) 5.000 msec.           IABIAS Intensity   90 Percent
Hotshot T-slice  1.999 msec.           IABIAS Duration     2 Minor T-slices
DSPBUF Q1          32767 Openings       STORBUF Q1 Q2 Q3  125 % Main storage
DSPBUF Q1 Q2       32767 Openings       STORBUF Q2 Q3     105 % Main storage
DSPBUF Q1 Q2 Q3    32767 Openings       STORBUF Q3         95 % Main storage
LDUBUF Q1 Q2 Q3     100 % Paging exp.   Max. working set 9999 % Main storage
LDUBUF Q2 Q3         75 % Paging exp.   Loading user        5 Pgrd / T-slice
LDUBUF Q3            60 % Paging exp.   Loading capacity    3 Paging expos.


Changed Scheduler Settings
Date  Time       Changed
..... ........   No changes processed
```

Redbooks

Effect of Virtual Machine Size for Linux

Memory is a shared resource under VM

- Normal to overcommit
  - VM paging ensures guest access to memory

To maximize efficient memory usage:

- Ensure Linux guests defined with 'just enough' virtual machine size

Virtual machine size affects performance!

- Smaller virtual machine size = smaller working set size
  - WSS factors into scheduling
- Smaller virtual machines more likely to use memory < 2GB
  - I/O performed in buffers below 2GB

Redbooks

User Resources for 1024MB Guest

```
FCX115     CPU 2084  SER 96A3A  Interval 17:34:30 - 17:34:31   Perf. Monitor

Detailed data for user LNXSU1
Total CPU  :  73.6%   Storage def. :   1024MB   Page fault rate:    .0/s
Superv. CPU :   .0%   Resident <2GB:      859   Page read rate :    .0/s
Emulat. CPU : 73.6%   Resident >2GB:    28536   Page write rate:    .0/s
VF total   : ....%    Proj. WSET   :    29085   Pgs moved >2GB>:    .0/s
VF overhead : ....%   Reserved pgs :        0   Main > XSTORE  :    .0/s
VF emulation: ....%   Locked pages :       10   XSTORE > main  :    .0/s
VF load rate: ..../s  XSTORE dedic.:      0MB   XSTORE > DASD  :    .0/s
I/O rate   :   .0/s   XSTORE pages :        0   SPOOL pg reads :    .0/s
DASD IO rate:  .0/s   DASD slots   :        1   SPOOL pg writes:    .0/s
UR I/O rate :  .0/s   IUCV X-fer/s :      .0/s  MDC insert rate:    .0/s
Diag. X'98' :  .0/s   Share        :      100   MDC I/O avoided:    .0/s
*BLOCKIO   :   .0/s   Max. share   :      ...

#I/O active :    0    Active   : 93%    PSW wait : 67%    I/O act. :  3%
Stacked blk :    ..   Page wait :  0%   CF wait  :  1%    Eligible :  0%
Stat.: ESA,P02,RNBL   I/O wait  :  0%   Sim. wait:  0%    Runnable : 33%

Proc.  %CPU %CP %EM   %VECT %VOHD %VEMU VLD/S  IO/S   Status
Enter 'STOrage Display' for storage details
Command ===>
```

User Resources for 128MB Guest

```
FCX115     CPU 2084  SER 96A3A  Interval 14:17:34 - 14:17:36   Perf. Monitor

Detailed data for user LNXSU1
Total CPU  :  69.3%   Storage def. :    128MB   Page fault rate:    .0/s
Superv. CPU :   .0%   Resident <2GB:    28346   Page read rate :    .0/s
Emulat. CPU : 69.3%   Resident >2GB:     2556   Page write rate:    .0/s
VF total   : ....%    Proj. WSET   :    30892   Pgs moved >2GB>:    .0/s
VF overhead : ....%   Reserved pgs :        0   Main > XSTORE  :    .0/s
VF emulation: ....%   Locked pages :       10   XSTORE > main  :    .0/s
VF load rate: ..../s  XSTORE dedic.:      0MB   XSTORE > DASD  :    .0/s
I/O rate   :   .0/s   XSTORE pages :        0   SPOOL pg reads :    .0/s
DASD IO rate:  .0/s   DASD slots   :        0   SPOOL pg writes:    .0/s
UR I/O rate :  .0/s   IUCV X-fer/s :      .0/s  MDC insert rate:    .0/s
Diag. X'98' :  .0/s   Share        :      100   MDC I/O avoided:    .0/s
*BLOCKIO   :   .0/s   Max. share   :      ...

#I/O active :    0    Active   : 97%    PSW wait : 67%    I/O act. :  6%
Stacked blk :    ..   Page wait :  0%   CF wait  :  0%    Eligible :  0%
Stat.: ESA,P02,RNBL   I/O wait  :  2%   Sim. wait:  0%    Runnable : 31%

Proc.  %CPU %CP %EM   %VECT %VOHD %VEMU VLD/S  IO/S   Status
Enter 'STOrage Display' for storage details
Command ===>
```

Timer Patch and Scheduling

By default, Linux wakes every 30 ms
- Updates system jiffies counter

This disrupts VM scheduling
- Virtual machine wakes up to do nothing (essentially)
- Scheduler 'rates' virtual machine lower
  - Looks like long-runing task (Q3)

Net effect:
- Less responsive Linux guests over time

Redbooks

Effect of Timer Patch
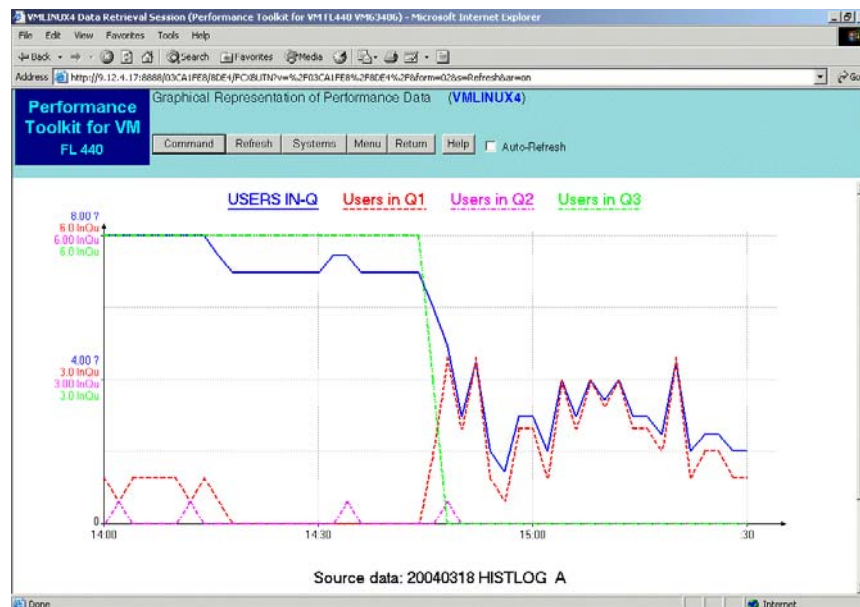
Redbooks

The QUICKDSP Option

Special virtual machine classification
- Do not wait on eligible list
- Go immediately to dispatch list

Fast mechanism to define 'high-priority' virtual machine(s)
- SET QUICKDSP ON

Use sparingly!
- Reserve for virtual machines that require immediate dispatching

Redbooks **ibm.com**/redbooks

QUICKDSP Users Always in Q0

```
CPU Load                                    Vector Facility    Status or
PROC   %CPU   %CP %EMU  %WT %SYS  %SP %SIC %LOGLD   %VTOT %VEMU  REST  ded. User
P00     81     0  81   19    0    0   14   100      not installed   Master
P01     81     0  81   19    0    0   28   100      not installed   Alternate

Total SSCH/RSCH     7/s      Page rate      3.1/s    Priv. instruct.    24/s
Virtual I/O rate    4/s      XSTORE paging   .0/s    Diagnose instr.    18/s
Total rel. SHARE  3500       Tot. abs SHARE    0%

<-- Device Descr. --> Mdisk Pa- <-Rate/s-> <------- Time (msec) -------> Req. <Percent>  SEEK Recov <-Throttle->
Addr Type   Label/ID  Links ths  I/O Avoid Pend Disc Conn Serv Resp CUWt Qued Busy READ  Cyls SSCH Set/s Dly/s
150D 3390-3 LX150D        2   3  1.4   .0   .2   .6  2.2  3.0  3.0   .0  .00    0    0 1172    0  ...   ...
1510 3390-3 LX4RES CP   183   3  1.0   .2   .2   .2  4.3  4.7  4.7   .0  .00    0    7  172    0  ...   ...
1590 3390-3 LX4W02 CP    49   3   .7   .4   .2   .2  2.3  2.7  2.7   .0  .00    0    0   82    0  ...   ...
1512 3390-3 LX1512        2   3   .3   .0   .2   .2  2.0  2.4  2.4   .0  .00    0    0  129    0  ...   ...

CHPID Chan-Group   <%Busy>     <----- Channel %Busy Distribution 13:29:47-14:57:27 ------>
(Hex) Descr Qual   Cur Ave    0-10 11-20 21-30 31-40 41-50 51-60 61-70 71-80 81-90 91-100
 7F   ESCON   00     4   9      76    8   13    3    0    0    0    0    0    0
 79   ESCON   00     3   4     100    0    0    0    0    0    0    0    0    0

Userid    %CPU  TCPU  VCPU Ratio Total DASD Avoid Diag98   UR Pg/s  User Status  Logged Active Pages SPg/s  MDC/s
Share Users
LNXSU2   80.3 16.06 16.06  1.0   .1   .1   .0    .0   .0   .0  ESA,CL0,DISP     0      0     0  .00    .0  100
LNXSU3   47.8 9.557 9.553  1.0   .1   .1   .0    .0   .0  1.0  ESA,CL0,DISP     0      0     0  .00    .0  100
LNXSU1   33.2 6.629 6.626  1.0  1.3  1.3   .0    .0   .0   .0  ESA,CL0,DISP     0      0     0  .00    .0  100
LNXSU4    .21  .041  .024  1.7   .2   .2   .0    .0   .0  2.1  ESA,CL0,DISP     0      0     0  .00    .0  100
FCX148        CPU 2084  SER 96A3A  Interval 14:57:27 - 14:57:47    Print on
```

Redbooks **ibm.com**/redbooks

Effect of SHARE Settings

SHARE determines position in eligible / dispatch lists
- Have an effect when only when CPU is overcommitted
- May be better alternative to QUICKDSP

Absolute share
- Guarantee minumum processor resources

Relative share
- Assigned processor after absolute shares are satisfied

Recommendations:
- Use relative shares for most users
- Save absolute share for critical virtual machines

ibm.com/redbooks

Using the Default SHARE Settings

```
   <----- CPU Load ----->       <-User Time-> <--Spool-->    MDC
                  <-Seconds->          <--Minutes--> Total  Rate Insert
   Userid    %CPU  TCPU  VCPU   User Status  Logged Active Pages SPg/s  MDC/s Share
   >System<  4.40  .881  .877   ---,---,----    .0     .0    .0    .0     .0   ---
   LNXSU1    52.8 10.56 10.55   ESA,CL3,DISP     0      0     0   .00     .0   100
   LNXSU2    52.5 10.50 10.49   ESA,CL3,DISP     0      0     0   .00     .0   100
   LNXSU3    51.3 10.26 10.25   ESA,CL3,DISP     0      0     0   .00     .0   100
   LNXRH4     .31  .062  .046   ESA,CL3,DISP     0      0     0   .00     .0   100
   PERFROG    .07  .013  .011   ESA,---,DORM     0      0     1   .05     .0 3.0%A
   PERFSVM    .05  .009  .009   ESA,---,DORM     0      0     0   .00     .0 3.0%A
   PERFAPP    .02  .003  .002   ESA,---,DORM     0      0     0   .00     .0 3.0%A
   PERFBER    .02  .004  .003   ESA,---,DORM     0      0     0   .00     .0 3.0%A
   BLDNUC     .01  .001  .001   EME,---,DORM     0      0     0   .00     .0   100
   RSCS       .01  .001  .000   ESA,---,DORM     0      0     0   .00     .0   100
   TCPIP      .01  .002  .001   ESA,CL0,DISP     0      0     0   .00     .0  3000
   TCPIP2     .01  .001  .000   ESA,---,DORM     0      0     0   .00     .0  3000
   BROYOLE      0     0     0   ESA,---,DORM     0      0     0     0      0   100
   CMS1         0     0     0   ESA,---,DORM     0      0     0     0      0   100
   DATAMOVE     0     0     0   ESA,---,DORM     0      0     0     0      0   100
   DIRMAINT     0     0     0   ESA,---,DORM     0      0     0     0      0   100
```

ibm.com/redbooks

Effect of Increasing Relative SHARE

```
<----- CPU Load ----->        <-User Time-> <--Spool-->    MDC
             <-Seconds->                 <--Minutes--> Total  Rate Insert
  Userid   %CPU  TCPU  VCPU  User Status  Logged Active Pages SPg/s MDC/sShare
>System<   4.46  .892  .887  ---,---,----    .0    .0    .1    .0    .0
LNXSU1      136 27.11 27.09  ESA,CL3,DISP     0     0     0   .00    .0  1000
LNXSU2     12.1 2.415 2.409  ESA,CL3,DISP     0     0     0   .00    .0   100
LNXSU3     10.9 2.171 2.164  ESA,CL3,DISP     0     0     0   .00    .0   100
LNXRH4      .25  .050  .035  ESA,CL3,DISP     0     0     0   .00    .0   100
VMRTM       .15  .029  .007  ESA,---,DORM     0     0     0   .00    .0 3.0%A
PERFROG     .06  .011  .010  ESA,---,DORM     0     0     2   .10    .0 3.0%A
BLDNUC      .04  .008  .006  EME,CL1,DISP     0     0     0   .00    .0   100
PERFSVM     .04  .008  .008  ESA,---,DORM     0     0     0   .00    .0 3.0%A
PERFAPP     .02  .003  .002  ESA,---,DORM     0     0     0   .00    .0 3.0%A
RSCS        .01  .001  .000  ESA,CL2,DISP     0     0     0   .00    .0   100
TCPIP       .01  .002  .001  ESA,---,DORM     0     0     0   .00    .0  3000
BROYOLE       0     0     0  ESA,---,DORM     0     0     0     0     0   100
CMS1          0     0     0  ESA,---,DORM     0     0     0     0     0   100
DATAMOVE      0     0     0  ESA,---,DORM     0     0     0     0     0   100
DIRMAINT      0     0     0  ESA,---,DORM     0     0     0     0     0   100
```

Redbooks **ibm.com**/redbooks

To Summarize:

Most effective tuning for Linux guests:

- Keep virtual machine size small!

Consider tuning VM scheduler for:

- STORBUF   overcommit memory
- LDUBUF   overcommit paging

To give priority to specific Linux guests:

- Increase SHARE
  - RELATIVE is preferred
- Possibly assign QUICKDSP
  - Do not overuse!

Redbooks **ibm.com**/redbooks

Equivalent RTM Functions

Performance Toolkit provides most RTM functions

- See:
  - http://www.vm.ibm.com/related/perfkit/pkitrtm.html   RTM-to-PERFKIT
  - http://www.vm.ibm.com/related/perfkit/pkitprf.html   PRF-to-PERFKIT

Some corresponding displays:

- DISPLAY GENERAL   GENERAL (user defined)
- DISPLAY SLOG LAST   REDHIST
- QUERY ENVIRON   SYSCONF
- DISPLAY USER   USER
- DISPLAY IDLE   IDLEUSER
- DISPLAY I/O   I/O

RTM General Display

```
+------------------------------------------------------------------------+
| z/VM   CPU2084 SERIAL 196A3A    3G DATE 03/11/04 START 10:21:00 END 10:21:30 |
|          *                                                             |
| <USERID> %CPU %CP %EM ISEC PAG  WSS  RES   UR PGES SHARE VMSIZE TYP,CHR,STAT |
| LNXSU5    1.4 .05 1.3  .50 .00  75K  75K   .0   1   100     1G VUS,DSC,DISP |
| SYSTEM    .41 .41 .00  .00 .00   0   55K   .0  571 .....     2G SYS,        |
| LNXSU1    .18 .08 .10  .06 .00  31K  31K   .0   0   100   128M VMS,DSC,DISP |
|                                                                        |
| <--- DEVICE ---> <----- DEVICE  RDEV  DATA ------> <-- MEASUREMENT FACILITY ->|
|                *                                                       |
| DEV TYPE VOLSER IOREQST SEC  %Q %ER R %LK  LNK PA %UT  ACC  FPT  DCT   CN %CN|
|1590 3390 LX4W02      35   1 .00 .00   .00   56  3 .35    3    0    0    2 .26|
|154C 3390 LX154C      17   0 .00 .00   .00    2  3 .54    9    0    4    5 .29|
|5090 CTCA              7   0 .00 .00   .00    0  1  46 2.0S    0 2.0S    0 .00|
|158C 3390 LX4SP1       5   0 .00 .00   .00    0  3 .01    1    0    0    1 .01|
|                                                                        |
| <---------- CPU STATISTICS --------> <-- VECTOR ---> <STORAGE><XSTORE> |
|  NC %CPU %US %EM %WT %SY %SP  XSI %SC NV %VT %OT RSTR %ST PSEC %XS XSEC  TTM|
|-> 2  2.5 .26 1.8 198 .41 .00  580  96  0   0   0    0  13    0   2    0 2.006|
|<-.. 2.6 .28 1.8 197 .45 .00  609  96 ..   0   0    0  13    0   2    0 2.766|
+----------------------<-- 04 LOG ACTIONS INDICATED -->----------------------+
```

User-defined General Display

```
*C DEFSCRN GENERAL LINE  2 TO  2 COPY USER     FROM 4
*C DEFSCRN GENERAL LINE  3 TO  6 COPY USER     FROM 6
*C DEFSCRN GENERAL LINE  7 TO  7 COPY DEVICE   FROM 2
*C DEFSCRN GENERAL LINE  8 TO 11 COPY DEVICE   FROM 5
*C DEFSCRN GENERAL LINE 12 TO 19 COPY CPU      FROM 1
```

Note:
- Defined in default FCONX $PROFILE
- Remove comments to enable

Performance Toolkit General Display

```
FCX148     CPU 2084  SER 96A3A  Interval 10:21:39 - 10:21:59    Remote Data

Userid    %CPU  TCPU  VCPU Ratio Total DASD Avoid Diag98   UR Pg/s  User Status
LNXSU5    1.47  .293  .281  1.0   .5   .5   .0    .0   .0   .0  ESA,CL3,DIS
LNXSU1     .18  .036  .019  1.9   .0   .0   .0    .0   .0   .0  ESA,CL3,DIS
PERFLIV    .05  .009  .008  1.1   .8   .7   .3    .0   .0   .0  ESA,---,DOR
PERFROG    .05  .009  .008  1.1  1.4  1.3   .5    .0   .0   .0  ESA,---,DOR
<-- Device Descr. --> Mdisk Pa- <-Rate/s-> <------- Time (msec) -------> Req.
1590 3390-3 LX4W02 CP    56   3  1.9  1.4   .2   .2  2.2  2.6  2.6   .0  .00
154C 3390-3 LX154C        2   3   .6   .0   .2   .2  2.1  2.5  2.5   .0  .00
5090 CTCA   >RSCS       ...   1   .3  ...   .1 2000   .3 2000 2000   .0  .00
150C 3390-3 LX150C        0   3   .1   .0   .2   .0   .4   .6   .6   .0  .00
CPU Load                                   Vector Facility   Status or
PROC  %CPU  %CP %EMU  %WT %SYS  %SP %SIC %LOGLD  %VTOT %VEMU  REST  ded. User
P00    1    0    1   99   0    0   97    1      not installed     Master
P01    1    0    1   99   0    0   98    1      not installed     Alternate

Total SSCH/RSCH    6/s    Page rate      .0/s   Priv. instruct.   41/s
Virtual I/O rate   8/s    XSTORE paging  .0/s   Diagnose instr.   39/s
Total rel. SHARE  3300    Tot. abs SHARE  0%
```

RTM SLOG Display

```
TOD z/VM   CPU2084 SERIAL 196A3A   3G DATE 03/11/04 START 10:18:24 END 11:28:30
H:M NC %CPU %US %EM %WT %SY %SP  XSI %SC NV %VT %OT RSTR %ST PSEC %XS XSEC   TTM
1120 2  2.4 .25 1.6 198 .46 .00  563  98  0   0   0    0  15    0   2    0 5.002
1120 2  2.7 .25 2.0 197 .42 .00  560  98  0   0   0    0  13    0   2    0 4.002
1121 2  2.4 .26 1.6 198 .48 .00  563  97  0   0   0    0  15    0   2    0 5.049
1121 2  2.8 .25 2.1 197 .42 .00  569  97  0   0   0    0  13    0   2    0 3.435
1122 2  3.0 .27 2.2 197 .52 .00  565  97  0   0   0    0  15    0   2    0 6.702
1122 2  3.2 .25 2.5 197 .47 .00  560  98  0   0   0    0  13    0   2    0 6.670
1123 2  2.8 .26 2.1 197 .48 .00  565  97  0   0   0    0  15    0   2    0 8.670
1123 2  2.7 .25 2.0 197 .42 .00  555  98  0   0   0    0  13    0   2    0 5.088
1124 2  2.8 .25 2.1 197 .46 .00  565  97  0   0   0    0  15    0   2    0 4.961
1124 2  2.7 .25 2.0 197 .42 .00  558  98  0   0   0    0  13    0   2    0 5.502
1125 2  2.3 .25 1.6 198 .46 .00  560  97  0   0   0    0  15    0   2    0 5.002
1125 2  2.7 .25 2.0 197 .41 .00  552  98  0   0   0    0  13    0   2    0 4.802
1126 2  2.3 .25 1.6 198 .46 .00  563  98  0   0   0    0  15    0   2    0 4.961
1126 2  2.8 .25 2.1 197 .42 .00  565  97  0   0   0    0  14    0   2    0 1.434
1127 2  2.5 .27 1.7 198 .46 .00  577  98  0   0   0    0  15    0   2    0 2.761
1127 2  3.2 .31 2.4 197 .48 .00  610  97  0   0   0    0  14    0   2    0 9.336
1128 2  2.8 .27 2.0 197 .47 .00  575  98  0   0   0    0  15    3   2    0 2.604
1128 2  2.3 .25 1.7 198 .42 .00  555  98  0   0   0    0  13    0   2    0 5.502
AVG  .. 2.7 .26 2.0 197 .46 .00  573  97 ..   0   0    0  13    0   2    0 4.042
```

Redbooks ibm.com/redbooks

Performance Toolkit REDHIST Display

```
FCX195              Input File: '20040311 HISTLOG  A'         Perf. Monitor

Date       Time  <--Interval--->      <-------------------- CPU Load (%) -
yyyy/mm/dd hh:mm El_Time Samples #CPU  CPU  %US  %CP  %EM  %SY %Spin   %WT
2004/03/11 11:20   20.0       1  2.0   2.5  2.1   .7  1.8   .4   .0    198
2004/03/11 11:20   20.0       1  2.0   2.3  1.9   .6  1.7   .5   .0    198
2004/03/11 11:21   20.0       1  2.0   2.9  2.4   .7  2.1   .4   .0    197
2004/03/11 11:21   20.0       1  2.0   2.6  2.1   .8  1.8   .5   .0    197
2004/03/11 11:21   20.0       1  2.0   3.1  2.7   .6  2.5   .5   .0    197
2004/03/11 11:22   20.0       1  2.0   3.5  3.0   .8  2.7   .5   .0    197
2004/03/11 11:22   20.0       1  2.0   2.6  2.1   .8  1.8   .5   .0    197
2004/03/11 11:22   20.0       1  2.0   3.0  2.6   .6  2.4   .4   .0    197
2004/03/11 11:23   20.0       1  2.0   2.7  2.3   .7  2.0   .4   .0    197
2004/03/11 11:23   20.0       1  2.0   3.2  2.8   .7  2.5   .4   .0    197
2004/03/11 11:23   20.0       1  2.0   2.4  1.9   .6  1.7   .4   .0    198
2004/03/11 11:24   20.0       1  2.0   2.7  2.3   .7  2.0   .4   .0    197
2004/03/11 11:24   20.0       1  2.0   2.5  2.1   .7  1.8   .4   .0    198
2004/03/11 11:24   20.0       1  2.0   2.3  1.9   .6  1.7   .4   .0    198
2004/03/11 11:25   20.0       1  2.0   2.7  2.3   .7  2.0   .4   .0    197
2004/03/11 11:25   20.0       1  2.0   2.5  2.0   .7  1.7   .4   .0    198
2004/03/11 11:25   20.0       1  2.0   2.3  1.9   .6  1.7   .4   .0    198
2004/03/11 11:26   20.0       1  2.0   3.0  2.6   .7  2.3   .4   .0    197
2004/03/11 11:26   20.0       1  2.0   2.5  2.1   .7  1.8   .4   .0    198
2004/03/11 11:26   20.0       1  2.0   2.4  1.9   .6  1.7   .4   .0    198
                       *** End of File ***
Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F10=Left  F11=Right  F12=Return
```

Redbooks ibm.com/redbooks

RTM Environment Display

```
+------------------------------------------------------------------------------+
| z/VM   CPU2084 SERIAL 196A3A    3G DATE 03/11/04 START 17:07:07 END 17:07:37 |
+------------------------------------------------------------------------------+
 z/VM Version 4 Release 4.0, service level 0401 (64-bit)
 Generated at 02/18/04 09:53:36 EST
 IPL at 03/08/04 09:36:44 EST

 CMS LEVEL 20, Service Level 401

 z/VM RTM Function Level 4.1.0

 Cryptographic Facility Available - No

 IBM               2084-                    <-------- CPU -------------------->
                   Cap Adj  Total  Config Standby  Reserve  Dedicated   Shared
 PROCESSOR                    10      5       0        5
 LPAR    25 A19     400        2      2       0        0         0         2
```

Performance Toolkit SYSCONF Display

```
FCX180      CPU 2084  SER 96A3A        System   Config.          Remote Data

System History for VMLINUX4 (running in LPAR)
CP Level               z/VM Version 4.4.0, Service Level 0401 (64-bit)
Last Termination Time  ........
Last Termination Code  ........
Last IPL Time          2004/03/08 at 09:36:44
Time Zone Offset       -05:00:00 From Greenwich Mean Time
Checkpoint Volume      LX4RES
Warmstart Volume       LX4RES

Initial Status on 2004/03/09 at 11:46, Processor 2084-
Real Proc.: Cap  1968, Total 10, Conf  5, Stby  0, Resvd 5
Log. Proc.: CAF   400, Total  2, Conf  2, Stby  0, Resvd  0, Ded  0, Shrd  2

<----- Processor ----->  Crypto Facility  <------- Vector Facility ------->
Num  Serial-Nr Status    ID     Status    Installed Connected Operational
 0    196A3A    Online    ..     Offline      No        No         No
 1    196A3A    Online    ..     Offline      No        No         No
S/370 guests not supported

Changed Processor Status
Date  Time       Changed
..... ........    No status changes received
```

RTM User Display

```
z/VM   CPU2084 SERIAL 196A3A   3G DATE 03/12/04 START 08:48:52 END 08:49:22
        *
  <USERID> %CPU %CP %EM ISEC PAG  WSS  RES   UR PGES SHARE VMSIZE TYP,CHR,STAT

   LNXSU5  .81 .05 .75  1.0 .00  76K  76K  .0    0  100    1G VUS,DSC,DISP
   SYSTEM  .47 .47 .00  .00 .00   0   55K  .0  570 .....   2G SYS,
   LNXSU1  .19 .07 .11  .06 .00  31K  31K  .0    0  100  128M VMS,DSC,DISP
   VMRTM   .10 .08 .02  .00 .00  831  852  .0    0  3%A   32M VUS,QDS,SIMW
   PERFLIV .05 .00 .04  .46 .00 2724 2731  .0    0  3%A   64M VUX,IAB,IDLE
   PERFROG .05 .00 .05  .60 .00 2856 2864  .0  195  3%A   64M VUX,IAB,IDLE
   PERFSVM .05 .00 .04  .56 .00 2560 2570  .0 1139  3%A   64M VUX,IAB,IDLE
   PERFKLA .05 .00 .04  .33 .00 2770 2777  .0 1120  3%A   64M VUX,IAB,IDLE
   PERFBER .01 .00 .01  .73 .00 1417 1425  .0    0  3%A   64M VUX,IAB,IDLE
   TCPIP   .01 .00 .00  .00 .00 1338 1788  .0 2028 3000   32M VUX,QDS,DISP
   RSCS    .00 .00 .00  .23 .00  223  225  .0 1157  100   16M VUS,IAB,IDLE
```

ibm.com/redbooks

Performance Toolkit USER Display

```
FCX112     CPU 2084  SER 96A3A  Interval 08:49:39 - 08:49:59    Remote Data
.       ____   .     .     .     .    .     .      .    .    .
         <----- CPU Load -----> <------ Virtual IO/s ------>
           <-Seconds->   T/V
Userid   %CPU  TCPU  VCPU Ratio Total DASD Avoid Diag98   UR Pg/s  User Status
>System<  .07  .015  .012  1.2   .1   .1    .0    .0     .0  .0  ---,---,---
LNXSU1    .83  .166  .150  1.1   .0   .0    .0    .0     .0  .0  ESA,CL3,DIS
LNXSU5    .71  .142  .131  1.1   .5   .5    .0    .0     .0  .0  ESA,CL3,DIS
VMRTM     .16  .032  .007  4.6   .0   .0    .0    .0     .0  .0  ESA,---,DOR
PERFKLA   .04  .008  .007  1.1   .3   .3    .1    .0     .0  .0  ESA,---,DOR
PERFLIV   .04  .008  .008  1.0   .4   .3    .1    .0     .0  .0  ESA,---,DOR
PERFROG   .04  .008  .007  1.1   .3   .3    .1    .0     .0  .0  ESA,---,DOR
PERFSVM   .04  .008  .007  1.1   .3   .3    .1    .0     .0  .0  ESA,---,DOR
PERFBER   .02  .003  .002  1.5   .3   .3    .1    .0     .0  .0  ESA,---,DOR
TCPIP     .02  .003  .001  3.0   .0   .0    .0    .0     .0  .0  ESA,CL0,DIS
```

ibm.com/redbooks

RTM Idle User Display

```
<>z/VM   CPU2084 SERIAL 196A3A    3G DATE 03/12/04 START 08:00:17 END 09:22:22<>

USERID->  IMIN  USERID->  IMIN  USERID->  IMIN  USERID->  IMIN  USERID->  IMIN
BROYOLE     27  DATAMOVE     1  DIRMAINT     1  DISKACNT  1031  EGELER    1385
EREP       444  GCS       1385  HAIMO     1329  OPERATOR  1385  OPERSYMP  1385
OP1        211  VMSERVR   1385  VMSERVS   1385  VMSERVU   1385
```

Redbooks ibm.com/redbooks

Performance Toolkit IDLEUSER Display

```
FCX238     CPU 2084  SER 96A3A      Status   09:26:39        Remote Data
_____      .   _____      .   _____      .   _____      .   _____      .
            Min            Min            Min            Min            Min
Userid     Idle  Userid   Idle  Userid   Idle  Userid   Idle  Userid   Idle
BROYOLE      31  DATAMOVE    0  DIRMAINT    0  DISKACNT 1035  EGELER   4180
EREP        448  GCS      4180  HAIMO    1333  OPERATOR  216  OPERSYMP 4180
OP1         216  PVM         0  VMSERVR  4180  VMSERVS  4180  VMSERVU  4180
```

Redbooks ibm.com/redbooks

RTM I/O Display

```
<>z/VM   CPU2084 SERIAL 196A3A   3G DATE 03/12/04 START 09:39:52 END 09:40:22<>
 <--- DEVICE ---> <----- DEVICE  RDEV  DATA ------> <-- MEASUREMENT FACILITY ->
                *
 DEV TYPE VOLSER IOREQST SEC  %Q %ER R %LK   LNK PA %UT   ACC  FPT  DCT    CN %CN
1590 3390 LX4W02      63   2 .00 .00   .00    55  3 .51     2    0    0     2 .48
154C 3390 LX154C      36   1 .00 .00   .00     2  3 .50     4    0    0     3 .40
5090 CTCA              7   0 .00 .00   .00     0  1  46   2.0S    0 2.0S     0 .00
150D 3390 LX150D       6   0 .00 .00   .00     2  3 .02     1    0    0     1 .02
150C 3390 LX150C       4   0 .00 .00   .00     0  3 .00     0    0    0     0 .00
150E 3390 LX150E       4   0 .00 .00   .00     1  3 .00     0    0    0     0 .00
150F 3390 LX150F       4   0 .00 .00   .00     0  3 .00     0    0    0     0 .00
1510 3390 LX4RES       4   0 .00 .00   .00   151  3 .00     0    0    0     0 .00
```

Performance Toolkit I/O Display

```
 FCX108      CPU 2084  SER 96A3A  Interval 09:40:39 - 09:40:59     Remote Data
 .   .                  .       ___     .     .     .     .     .     .     .
 <-- Device Descr. --> Mdisk Pa- <-Rate/s-> <------- Time (msec) ------> Req.
 Addr Type   Label/ID  Links ths  I/O Avoid Pend Disc Conn Serv Resp CUWt Qued
 >> All DASD <<        ....        .2    .0   .2   .1  1.3  1.6  1.6   .0  .00
 1590 3390-3 LX4W02 CP   55   3   1.5    .9   .2   .2  2.3  2.7  2.7   .0  .00
 154C 3390-3 LX154C       2   3    .6    .0   .2   .1  4.0  4.3  4.3   .0  .00
 5090 CTCA   >RSCS     ...   1    .3   ...   .1 2000   .3 2000 2000   .0  .00
 150C 3390-3 LX150C       0   3    .1    .0   .2   .0   .4   .6   .6   .0  .00
 150D 3390-3 LX150D       2   3    .1    .0   .2   .0   .4   .6   .6   .0  .00
 150E 3390-3 LX150E       1   3    .1    .0   .1   .0   .4   .5   .5   .0  .00
 150F 3390-3 LX150F       0   3    .1    .0   .2   .0   .4   .6   .6   .0  .00
 1510 3390-3 LX4RES CP  151   3    .1    .0   .2   .0   .4   .6   .6   .0  .00
```

Equivalent VMPRF Functions

VMPRF processes CP MONITOR data (MONWRITE)

- Can be automatically run on daily basis
- Creates:
  - Printed reports
  - machine-readable data (trend and summary records)

Performance Toolkit provides similar functions

- MONSCAN *fn ft fm*
  - Step through intervals using NEXTSAMP
  - Some reports rely on CP control blocks
- TRNDSCAN *fn ft fm*
  - Processes large amounts of data
  - Suitable for capacity planning

ibm.com/redbooks

Monitor Mode During MONSCAN

```
FCX124        Performance Screen Selection  (FL440 VM63447)   Monitor Scan

General System Data      I/O Data                History Data (by Time)
1. CPU load and trans.   11. Channel load        31. Graphics selection
2. Storage utilization   12. Control units       32. History data files*
3. Storage subpools      13. I/O device load*    33. Benchmark displays*
4. Priv. operations      14. CP owned disks*      34. Correlation coeff.
5. System counters       15. Cache extend. func.* 35. System summary*
6. CP IUCV services      16. DASD I/O assist      36. Auxiliary storage
7. SPOOL file display*   17. DASD seek distance*  37. CP communications*
8. LPAR data             18. I/O prior. queueing* 38. DASD load
9. Shared segments       19. I/O configuration    39. Minidisk cache*
A. Shared data spaces    1A. I/O config. changes  3A. Paging activity
B. Virt. disks in stor.                           3B. Proc. load & config*
C. Transact. statistics  User Data                3C. Logical part. load
D. Monitor data          21. User resource usage* 3D. Response time (all)*
E. Monitor settings      22. User paging load*    3E. RSK data menu*
F. System settings       23. User wait states*    3F. Scheduler queues
G. System configuration  24. User response time*  3G. Scheduler data
H. VM Resource Manager   25. Resources/transact.* 3H. SFS/BFS logs menu*
                         26. User communication*  3I. System log
I. Exceptions            27. Multitasking users*  3K. TCP/IP data menu*
                         28. User configuration*  3L. User communication
K. User defined data*    29. Linux systems*       3M. User wait states

        Pointers to related or more detailed performance data
        can be found on displays marked with an asterisk (*).
```

ibm.com/redbooks

Performance Toolkit SYSSUMLG Display

```
FCX225    Data for 2004/03/22  Interval 14:17:47 - 14:50:27   Monitor Scan

          <------- CPU --------> <Vec> <--Users--> <---I/O---> <Stg> <-Paging--
          <--Ratio-->                    SSCH  DASD Users <-Rate/s--
Interval   Pct          Cap-  On-   Pct  Log-       +RSCH Resp   in PGIN+ Read
End Time   Busy  T/V   ture  line  Busy  ged Activ   /s   msec Elist PGOUT Writ
>>Mean>>   27.5  1.02 .9873  2.0   .0   27   12    7.4  7.6   .8  47.8  28.
14:28:47     .5  2.93 .5317  2.0   .0   27   11    5.3  2.3   .0   .0    .
14:29:07   71.2  1.09 .9187  2.0   .0   27   12    5.9  4.0   .0  4233   .
14:29:27   80.8  1.03 .9684  2.0   .0   27   13    5.9  2.1   .0  64.3   .
14:29:47   80.7  1.02 .9834  2.0   .0   27   11    5.3  1.6   .0  17.5   .
14:30:07   69.4  1.01 .9951  2.0   .0   27   13    6.0  3.0   .0  37.7   .
14:30:27   80.6  1.00 .9968  2.0   .0   27   13    5.1  2.1   .0  29.0   .
14:30:47   81.5  1.00 .9967  2.0   .0   27   13    7.6  2.5   .0  15.2   .
14:31:07   81.2  1.01 .9956  2.0   .0   27   14    8.2  4.1   .0 130.5   .
14:31:27   81.6  1.00 .9968  2.0   .0   27   14    7.4  2.5   .0   3.0   .
14:31:47   11.5  1.04 .9741  2.0   .0   27   13    6.0  1.9   .0  37.9   .
14:32:07     .7  2.17 .6294  2.0   .0   27   14    7.1  2.3   .0  19.7   .
14:32:27     .7  2.17 .6117  2.0   .0   27   14    5.3  1.8   .0  11.7   .
14:32:47     .7  2.21 .6054  2.0   .0   27   13    6.0  2.8   .0   9.7   .
14:33:07     .7  2.07 .6541  2.0   .0   27   13    6.5  1.9   .0   8.6   .
14:33:27     .7  2.09 .6423  2.0   .0   27   13    6.3  2.6   .0   4.5   .
14:33:47     .7  2.15 .6335  2.0   .0   27   13    6.5  3.6   .0    .9   .
14:34:07     .7  2.13 .6360  2.0   .0   27   14    5.7  2.1   .0    .4   .
14:35:26    1.0  3.58 .5270  2.0   .0   27   16   45.6 17.5   .0  15.1 602.
14:35:27    2.0  1.67 .7250  2.0   .0   27   11   98.0  2.2   .0    .0  20.
14:35:47     .7  2.25 .6209  2.0   .0   27   14    6.8  2.4   .0    .0   3.
14:36:07     .8  1.92 .6898  2.0   .0   27   15   17.9  4.5   .0    .0  14.
14:36:27   63.3  1.01 .9958  2.0   .0   27   12   24.1  2.8   .0    .0  58.

Command ===>
F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F10=Left  F11=Right  F12=Return
```

Redbooks

Changes for FL510

New installation VM user

- 5VMPTK10

Support for VM PRF

- New parameters for PERFKIT command
  - BATCH   run in batch mode
  - VMPRF   migrate PRF control files

Use of APPLDATA

- Reduced overhead for Linux guests

Montoring for FCP-attached SCSI

- New monitor records to support SCSI disks

Redbooks

The 5VMPTK10 Directory Entry

```
USER 5VMPTK10 XXXXXXXX 32M 32M EG
    IPL CMS
    MACHINE XA
    CONSOLE 0009 3215
    SPOOL 000C 2540 READER *
    SPOOL 000D 2540 PUNCH A
    SPOOL 000E 1403 A
    LINK MAINT 0190 0190 RR
    LINK MAINT 019E 019E RR
    LINK MAINT 019D 019D RR
    LINK MAINT 051D 051D RR
    LINK MAINT 05E5 05E5 RR
    ACCOUNT xxxxxxxx
    MDISK 0191 3390 ssss 12 vvvvvv MR     1
    MDISK 02A2 3390 ssss 2  vvvvvv MR     2
    MDISK 02C2 3390 ssss 2  vvvvvv MR     3
    MDISK 02D2 3390 ssss 50 vvvvvv MR     4
    MDISK 02C4 3390 ssss 2  vvvvvv MR     5
    MDISK 0200 3390 ssss 10 vvvvvv MR     6
    MDISK 02A6 3390 ssss 2  vvvvvv MR     7
    MDISK 0201 3390 ssss 10 vvvvvv MR     8
    MDISK 02B2 3390 ssss 9  vvvvvv MR     9
    MDISK 01CC 3390 ssss 1  vvvvvv MR    10
    MDISK 029D 3390 ssss 8  vvvvvv MR    11
```

The 5VMPTK10 Directory Entry for FL510

```
USER PERFSVM  PERFSVM 64M 512M ABDEG
 MACHINE XA
 XAUTOLOG AUTOLOG1
 ACCOUNT xxxx
 NAMESAVE MONDCSS
 IUCV *MONITOR MSGLIMIT 255
 IUCV *IDENT FCXRES00 GLOBAL
 IUCV *IDENT FCXSYSTM GLOBAL
 IUCV ALLOW
 SHARE ABS 3%
 IPL CMS
 OPTION QUICKDSP
 CONSOLE 0009 3215
 SPOOL 000C 2540 READER *
 SPOOL 000D 2540 PUNCH A
 SPOOL 000E 1403 A
 LINK MAINT 190 190 RR
 LINK MAINT 19D 19D RR
 LINK MAINT 19E 19E RR
 LINK 5VMPTK10 200 200 RR
 LINK 5VMPTK10 201 201 RR
 LINK 5VMPTK10 1CC 1CC RR
 LINK 5VMPTK10 29D 29D RR
 MDISK 191 3390 1007 060 440W02  MR READ     WRITE    MULTIPLE
 MDISK 195 3390 1067 060 440W02  MR READ     WRITE    MULTIPLE
```

Support for VMPRF

Batch report generation

- PERFKIT BATCH *profn proft profm* DISK *monfn monft monfm*
  - *profn*   same form as FCONX $PROFILE
  - *monfn*   monitor data to process
- New R_S option on FC MONCOLL RESET command
  - Defines start time for summary record generation

VMPRF to Performance Toolkit migration

- PERFKIT VMPRF *masfn masft masfm* DISK *monfn monft monfm*
  - *masfn*   VMPRF MASTER file
  - *monfn*   monitor data to process
- Creates:
  - FCONX FCXEQUIV   equivalent FCONX REPORTS
  - RUNFILE   FC MONCOLL RESET commands

Redbooks

PERFKIT VMPRF Generated RUNFILE

```
FC MONCOLL RESET 00:00:00R_P 23:59:59P (MERGE
FC MONCOLL RESET 00:00:00R_T 08:00:00T 16:00:00T (MERGE
FC MONCOLL RESET 00:00:00R_S 01:00:00S 02:00:00S 03:00:00S 04:00:00S
FC MONCOLL RESET 05:00:00S 06:00:00 07:00:00S 08:00:00S 09:00:00S
FC MONCOLL RESET 10:00:00S 11:00:00S 12:00:00S 13:00:00S 14:00:00
FC MONCOLL RESET 15:00:00S 16:00:00S 17:00:00S 18:00:00S 19:00:00S
FC MONCOLL RESET 20:00:00S 21:00:00T 22:00:00S 23:00:00S 23:59:59S (MERGE
```

Redbooks

Support for Linux APPLDATA

Linux guests can create monitor data

- Reduced monitoring overhead
  - RMF PM interface not required (still supported)

Affects:

- LXCPU
- LXMEM
- LXNETWRK

One line per monitored Linux guest

- Detailed monitoring available from RMF PM

Redbooks

Linux System Selection for FL510

```
FCX242      CPU 2064  SER 51524        Linux Displays        Perf.


  Linux screens selection
S Display     Description
. LINUX       RMF PM system selection menu
. LXCPU       Summary CPU activity display
. LXMEM       Summary memory util. & activity display
. LXNETWRK    Summary network activity display







 Select performance screen with cursor and hit ENTER
 Command ===>
 F1=Help  F4=Top  F5=Bot  F7=Bkwd  F8=Fwd  F12=Return
```

Redbooks

LXCPU Display for FL510

```
FCX243  Run 2004/03/25 14:53:34          LXCPU
                                         Linux CPU Utilization Summary
From 2004/03/25 14:51:48
2004/03/25 14:52:26
For    39 Secs 00:00:39                  This is a performance report for system XYZ
_____

_____       .      .      .      .      .      .      .      .      .
             <------------------ Total CPU -------------------->
Linux    Virt <--------------- Utilization (%) ----------------->
Userid   CPUs TotCPU  User Kernel  Nice   IRQ SoftIRQ IOWait  Idle
>System<  2.0 103.4 102.5   1.0    .0   ...     ...    ...  96.6
LNXSU1      2 103.4 102.5   1.0    .0   ...     ...    ...  96.6
```

LXMEM Display for FL510

```
FCX244  Run 2004/03/25 14:53:43          LXMEM
Page    1
                                         Linux Memory Util. and Activity Summary
From 2004/03/25 14:51:48
VMLINUX4
To   2004/03/25 14:52:26
CPU 2084      SN  96A3A
For    39 Secs 00:00:39                  This is a performance report for system XYZ
z/VM   V.4.4.0 SLU 0401
_____

_____       .      .      .      .      .      .      .
             <------------ Memory Allocation (MB) ------------->
Linux        <--- Main ---> <--- High --->       Buffers  Cache
Userid   M_Total %MUsed H_Total %HUsed Shared /CaFree  Used
>System<  503.1  43.5    .0    .0    .0   8.5   83.3
LNXSU1    503.1  43.5    .0    .0    .0   8.5   83.3
```

Support for SCSI Disk Monitoring

New SCSI command

- Displays information on all SCSI devices
  - device number
  - number of bytes per block
  - transfer rate (KB / sec)
  - number of transfers
  - number of seeks
  - number of blocks read
  - number of blocks written
  - device utilization (busy time over interval)

New SCSILOG *dev* command

SCSI Disk Screen

```
FCX248  Run 2004/03/25 14:53:50        SCSI
Page    1
                                       SCSI DASD Activity Summary
From 2004/03/25 14:51:48
VMLINUX4
To   2004/03/25 14:52:26
CPU 2084      SN  96A3A
For    39 Secs 00:00:39               This is a performance report for system XYZ
_____
_____  .    .      .       .       .        .       .
       Bytes  <--------- Activity/sec --------->  Utiliz
Device   per          Trans          <--Blocks--->   ation
Number  block  KByte   fers  Seeks   Read Written    Rate
0BCD     512     0    357.6 .0000   129.9  423.0    .0000
0BCE     512     0    360.2 .0000   125.4  411.3    .0000
```

IBM.

FCP Implementation Guide on Linux for zSeries

Redbooks

International Technical Support Organization

FCP Implementation Topics

FCP overview

- Terminology

FCP and SCSI addressing

- Defining FCP devices
- Mapping FCP devices to SCSI subsytem

FCP support for Linux on zSeries

- Installation on / IPL from SCSI disk
  - SLES8 / SLES 9 / RHEL 3

Multipathing with FCP-attached SCSI

- Logical Volume Manager (LVM)
- Enterprise Volume Managment System (EVMS)
- Multiple devices admin (mdadm)

Redbooks

ibm.com/redbooks

Related Publications

Other Publications

- *These publications are relevant as information on Linux for zSeries FCP*

Title	Publication Number
Linux for zSeries: Fibre Channel Protocol Implementation Guide	SG24-6344
Building SuSE SLES8 Systems Under z/VM	REDP-3687
Getting Started With zSeries Fibre Channel Protocol	REDP-0205
IBM TotalStorage Enterprise Storage Server Model 800	SG24-6424
Introduction to Storage Area Networks	SG24-5470
Device Drivers and Installation Commands	LNUX-1313

Redbooks

ibm.com/redbooks

FCP Topologies

point-to-point

arbitrated loop
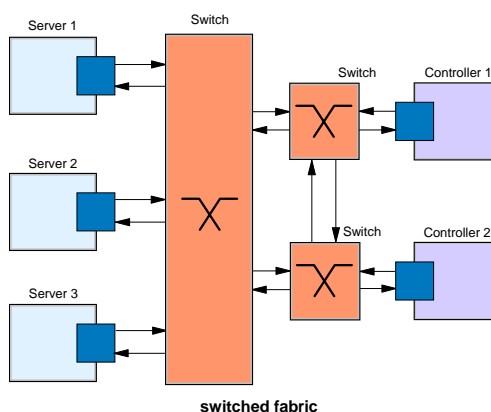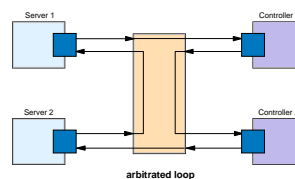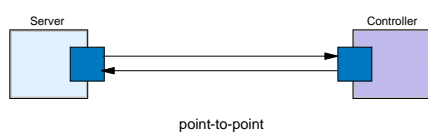
switched fabric

Redbooks

ibm.com/redbooks

FCP Terminology

Node

- Fabric endpoint
  - Server node
  - Storage node

Port

- Hardware interface to node
  - N_Port   Node port connects to node
  - E_Port   Expansion port connects switches
  - F_Port   Fabric port connects to N_port (not loop enabled)
  - FL_Port Fabric loop port connects to N_port (loop enabled)
  - G_Port   Generic port has no roll in fabric
  - L_Port   Loop port connects nodes in a loop
  - NL_Port Node loop port in arbitrated loop
- Ports connect to fabric using fiber optic cables

Redbooks

FC Port Types

Redbooks

Fibre Channel Link

FC port   FC port

Tx   Outbound   Outbound   Tx

Rx   Inbound   Inbound   Rx
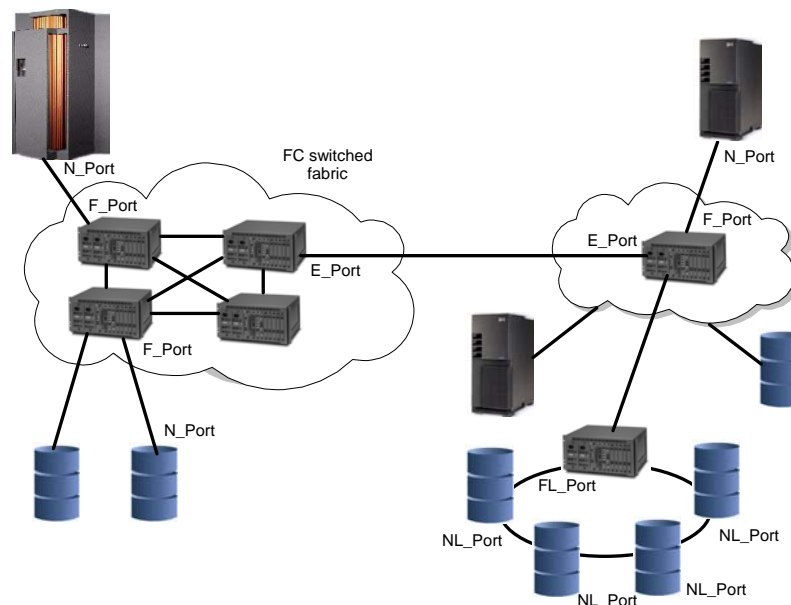
Fibre Channel link
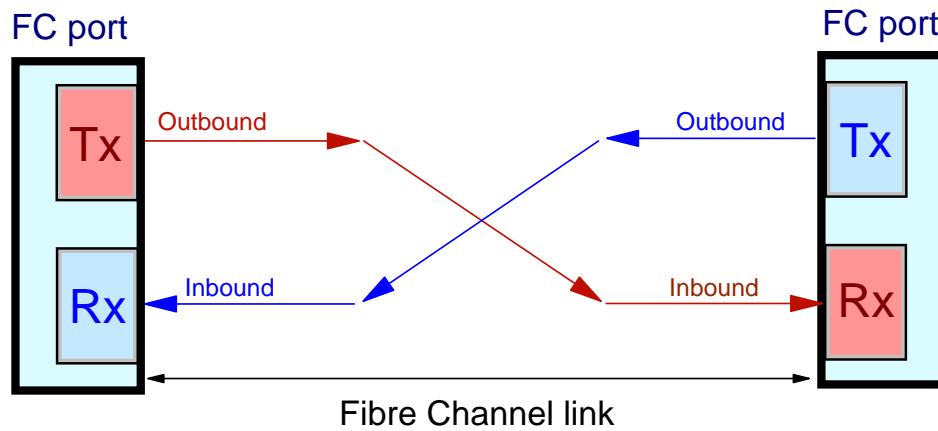
FCP and SCSI Addressing

Accessing FCP devices in a fabric
- Worldwide names

Accessing SCSI devices in Linux
- Linux SCSI subsystem

Defining FCP channel to zSeries
- Hardware configuration

Mapping FCP address to SCSI address
- zfcp device driver
- /proc/scsi interface

Worldwide Names

64-bit address

- Uniquely identifies device
  - Similar to ethernet MAC address
- Assigned by device manufacturer
- Defined by IEEE standard

Types:

- Worldwide node name (WWNN)
- Worldwide port name (WWPN)

Redbooks

World Wide Names

Server WWNN — Port WWPN — Port WWPN — Switch WWNN — Port WWPN — Port WWPN — Controller WWNN

Redbooks

WWNN on a z990

SCZP901 • State Active • Keystrokes remote
Keystrokes Session Services Help
Analyze Channel Information

Channel type:	Fibre FC	Hardware Type:	0B
		Hardware Sub Type:	00
		2Byte Control Unit Link Addr defined:	No
Partition ID:	1		
MIF Image ID:	1	Absolute address:	00000000
Channel mode:	Shared	Absolute address:	62C0E000
CHPARM:	00		
CSS.CHPID:	0.84		
PCHID:	03A0		
Switch number:	62	CVC CCC threshold:	5
Switch number valid:	1	IFCC threshold:	4
		Channel link address:	621513
State:	Online	Temp error threshold:	04
Status:	Operating	Suppress:	0000000000000000
Image chnl state:	Online	SAP affinity	00
Image chnl status:	Operating		
Error code:	00		
Ber inbound:	0	Connection Rate	FICON 1.75 at 2Gb
Ber outbound:	0		
Node type:	Self	Node type:	Attached
Node status:	Valid	Node status:	Valid
Flag/Parm:	10000784	Flag/Parm:	00200A15
Type/Model:	002084-A08	Type/Model:	006064-001
MFG:	IBM	MFG:	MCD
Plant:	02	Plant:	01
Seq. number:	000000026A3A	Seq. number:	0000000119D2
Tag:	8084	Tag:	0011
Worldwide Node name	5005076400C26A3A	Worldwide Node name	1000080088A0BC01
Worldwide Port name	5005076401003E74	Worldwide Port name	2015080088A0BC01

Redbooks ibm.com/redbooks

SCSI Addresses

Addressing scheme

- Adapter number   host
- Channel number   bus
- ID number   target
- Logical Unit Number   LUN

Device names

- /dev/sda   refers to the entire SCSI disk
  - Block device (major number 8)
- /dev/sda1   refers to partition 1 on /dev/sda
  - Up to 15 partitions / disk

Redbooks ibm.com/redbooks

zSeries FCP Addresses

Device number

- Devices assigned to FCP channel
- Defined in IOCDS
- Assigned to VM guest
  - DEDICATE or ATTACH

WWPN

- WWPN of adapter on storage node

SAN device LUN

- Specific LUN to access on storage node

Redbooks

Defining an FCP Channel

```
Session A - [24 x 80]
File  Edit  View  Communication  Actions  Window  Help

 _    Goto  Filter  Backup  Query  Help
 -                       View Channel Path Definition                    --
 C

      Processor ID . . . . : ZAPHOD      IBM AUSTRALIA zSeries
 S    Configuration mode : LPAR

 P    Channel path ID . . . . : 2B
 C    Channel path type . . . : FCP
      Operation mode . . . . . : SHR
      Managed  . . . . . . . . : No   I/O Cluster  . . :
 /
 v    Description . . . . . . : OpenFCP testing

 _    Dynamic entry switch ID  :                                          _
 _    Entry switch ID  . . . . :                                          _
 _    Entry port . . . . . . . :                                          _
 _
 _    ENTER to continue.                                                  _
 _     F1=Help    F2=Split   F3=Exit    F9=Swap   F12=Cancel              _
 _
     F1=Help      F2=Split    F3=Exit     F4=Prompt    F5=Reset    F7=Backward
     F8=Forward   F9=Swap    F10=Actions  F11=Add     F12=Cancel   F13=Instruct
     F20=Right    F22=Command
 MA    a                                                            01/
     Connected to remote server/host 9.190.207.87 using port 23
```

Redbooks

IOCDS Definition for FCP Channels

```
CHPID PATH=(60),SHARED,
PARTITION=((LINUXVM,LINUXVM2),(LINUXVM,LINUXVM2)),TYPE=FCP
CNTLUNIT CUNUMBR=60FC,PATH=(60),UNIT=FCP
IODEVICE ADDRESS=(1700,064),CUNUMBR=(60FC),UNIT=FCP
CHPID PATH=(61),SHARED,
PARTITION=((LINUXVM,LINUXVM2),(LINUXVM,LINUXVM2)),TYPE=FCP
CNTLUNIT CUNUMBR=61FC,PATH=(61),UNIT=FCP
IODEVICE ADDRESS=(1600,064),CUNUMBR=(61FC),UNIT=FCP
```

Redbooks

FCP Map Entry

Target number

Linux LUN

```
0x1625  0x00000007:0x5005076300cd9589  0x00000009:0x5305000000000000
```

Device number

WWPN

SAN device LUN
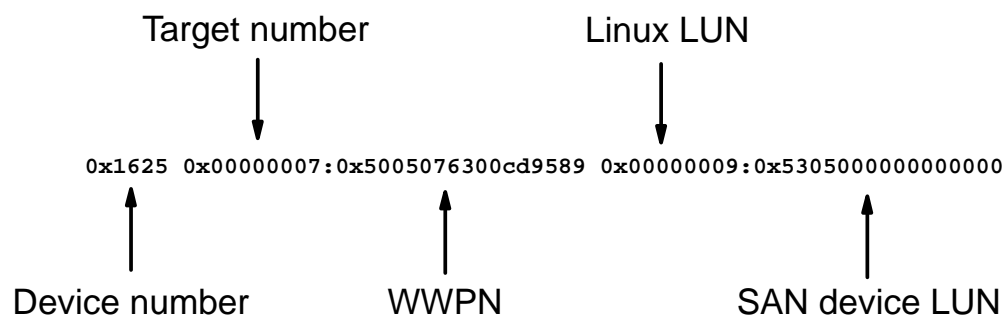
Redbooks

Mapping zSeries FCP Addresses to SCSI

Device number
- Use FCP device number assigned to guest

Target number
- Assigned by Linux administrator

WWPN
- Query switch to find SAN adapter WWPN

Linux LUN
- Assigned by Linux administrator

SAN device LUN
- LUN to access

Device Numbers

2 FCP channels per adapter
- 240 devices / channel on z800 or z900
- 480 devices / channel on z890 or z990

Each VM guest must use different device number

Same device numbers can be used in different LPARs
- Separate units created internally

4 digit hexadecimal number
- Reserved device numbers:
  - FC, FD, FE, and FF

Target Numbers

Target number 0 cannot be used

- This addresses the FCP adapter
- Normally start at 1
  - Increase sequentially

Use different target numbers for each LUN in a WWPN

- Same target number may be used on different WWPN (device number)

All LUNs on a single WWPN must use same target number

- Accessed by the same device number

WWPN

Specifies WWPN of storage device

- 16 bytes in length
- Typically specified in hexadecimal

To find number:

- Query switch
- Query storage device

Linux LUN

Assigned by you!

Typically:
- Begins at 0
- Increments by 1

SAN Device LUN

Use LUN number from storage device
- Assigned by storage controller

Must be 16 hexadecimal characters
- LUN assigned by storage node may be only 4 digits
- Use these 4 characters
  - Fill remaining 12 with 0's

FCP Mapping Syntax Rules

Numbers are C language format:

- 0x*nnnn*   hexadecimal
- 0*nnnn*   octal
- other   decimal

Comments may appear in files

- Leading # (up to end of line is ignored)

Continuation character (\) valid for SUSE

- Not recognized ion RHEL 3 version of /etc/modules.conf!

The zfcp Device Driver

Accepts FCP mapping as parameter on loading:

- insmod zfcp map="0x1617 0x01:0x200300A0B812106F 0X0:000C000000000000"

When loaded:

- Uses map parameter to detect LUNs
  - Attaches LUN to SCSI subsystem
- SCSI host assigned based on device number
  - SCSI host 0 is first device number
- SCSI bus number is always 0
- SCSI target ID assigned based on map value
- SCSI LUN number is next sequential value
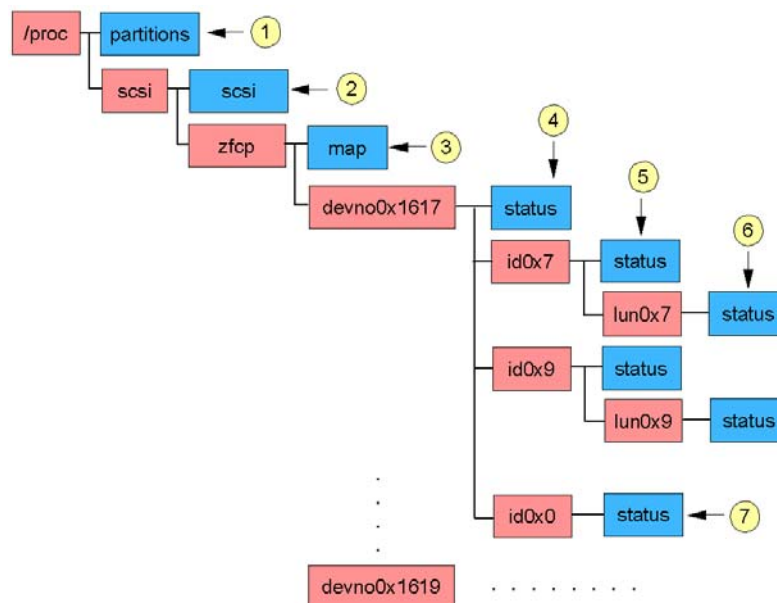
The /proc/scsi Interface

The /proc Interface

/proc/partitions
Lists all partitions (SCSI + ECKD)

/proc/scsi/scsi
Lists information on attached SCSI devices

/proc/scsi/zfcp/map
Reports configured FCP mapping

/proc/scsi/zfcp/devno0x1617/status
Reports status of FCP device 0x1617

/proc/scsi/zfcp/devno0x1617/id0x7/status
Reports status of all connected LUNs (device 0x1617, SCSI id 0x7)

/proc/scsi/zfcp/devno0x1617/id0x7/lun0x7/status
Reports status of accessed LUN

/proc/scsi/zfcp/devno0x1617/status
Reports status of FCP adapter

Customize Activation Profiles
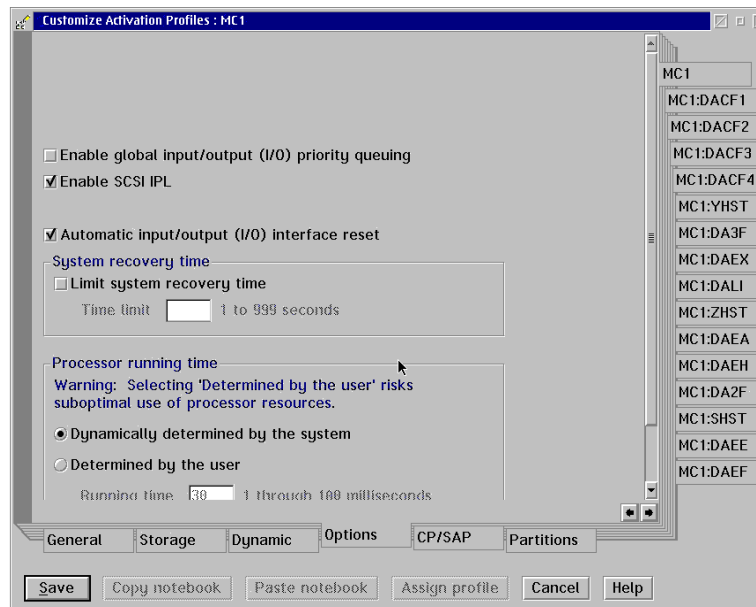
To Summarize:

Switched fabric is the only supported topology fo zSeries!
- FCP devices accessed using FCP channel / adapter
- Must enable SCSI IPL feature in activation profile

Worldwide names uniqely identify FCP devices
- WWNN / WWPN

FCP devices accessed through SCSI subsystem
- zfcp is the FCP device driver
- Mapping parameters used to map FCP device to SCSI subsystem

IPL Linux Guest From FCP-attached SCSI

Define FCP device to guest

- DEDICATE or ATTACH device

Install Linux on SCSI disk

- SUSE
  - SLES8 with Service Pack 3
  - SLES9
- Red Hat
  - With Update 2

Provide WWPN and LUN to machine loader

- SET LOADDEV command

IPL from SCSI device

Redbooks

Using the DEDICATE statement

```
USER LNXSU5 LNXSU5 128M 1G G
    INCLUDE IBMDFLT
    IPL 190 PARM AUTOCR
    MACHINE XA
    DEDICATE 7104 7138
    DEDICATE 7105 7139
    DEDICATE 7106 713A
    DEDICATE 1605 1605
    LINK TCPIP 0592 0592 RR
    MDISK 0191 3390 175 25 440U1R
```

Redbooks

Using the ATTACH command

```
Q V FCP
FCP  1605 ON FCP   1605 CHPID 61 SUBCHANNEL = 0018
     1605 QDIO-ELIGIBLE       QIOASSIST NOT AVAILABLE
Ready; T=0.01/0.01 12:31:51


ATT 1615 LNXSU5
FCP 1615 ATTACHED TO LNXSU5 1615
Ready; T=0.01/0.01 12:30:49
```

Using the SET LOADDEV Command

```
SET LOADDEV PORTNAME 50050763 00CD9589 LUN 53030000 00000000

Q LOADDEV
PORTNAME 50050763 00CD9589    LUN  53030000 00000000    BOOTPROG 0
BR_LBA    00000000 00000000
```

The LOADDEV Directory Entry

```
DIRM LOADDEV PORTNAME 5005076300CD9589
DVHXMT1191I Your LOADDEV request has been sent for processing.
Ready; T=0.04/0.05 09:41:48
 DVHREQ2288I Your LOADDEV request for LNXSU5 at * has been accepted.
 DVHBIU3450I The source for directory entry LNXSU5 has been updated.
 DVHBIU3424I The next ONLINE will take place immediately.
 DVHBIU3428I Changes made to directory entry LNXSU5 have been placed
 DVHBIU3428I online.
 DVHREQ2289I Your LOADDEV request for LNXSU5 at * has completed; with RC =
 DVHREQ2289I 0.


DIRM LOADDEV LUN 5303000000000000
DVHXMT1191I Your LOADDEV request has been sent for processing.
Ready; T=0.04/0.05 09:42:14
 DVHREQ2288I Your LOADDEV request for LNXSU5 at * has been accepted.
 DVHBIU3450I The source for directory entry LNXSU5 has been updated.
 DVHBIU3424I The next ONLINE will take place immediately.
DVHBIU3428I Changes made to directory entry LNXSU5 have been placed
 DVHBIU3428I online.
 DVHREQ2289I Your LOADDEV request for LNXSU5 at * has completed; with RC =
 DVHREQ2289I 0.
```

Redbooks

Installation with SLES8

SCSI support requires SP3!
- Base SLES8 zipl version does not support FCP

Problem:
- Cannot install SP3 directly
  - Install base SLES8
  - Must use YaST to upgrade to SP3
- YaST in SLES8 is not FCP-aware
  - Cannot detect / configure FCP devices

Solution:
- Install on ECKD, then copy installation to SCSI
- Use custom approach with YaST

Redbooks

Customized SLES8 YaST Installation

Manually load device drivers

- Uses SSH installation
- Done early in process (before YaST)
- Similar to fix for un-partitioned ECKD DASD

Partition the SCSI disk

- Use fdisk command

Upgrade to SP3

- Additional last step in process
- Uses secondary YaST session
  - Inside a chroot jail
  - Before primary YaST session completes!

SP3 version of zipl completes installation

Redbooks

Manually Load Device Drivers

Choose SSH installation terminal type

Load FCP and SCSI stack modules

- zfcp / qdio
- scsi_mod / sd_mod

Provide device mapping to zfcp driver

- Use:
  - map parameter on insmod command
  - /proc/scsi/zfcp/add_map interface

Redbooks

The fcpstart Script

```
#! /bin/sh
rmmod zfcp
modprobe qdio
modprobe scsi_mod
insmod zfcp map="\
0x1605 0x1:0x5005076300cd9589 0x0:0x5303000000000000"
modprobe sd_mod
```

ibm.com/redbooks

Manually Loading Device Drivers

```
# modprobe qdio

# modprobe scsi_mod

# insmod zfcp
Using /lib/modules/2.4.19-3suse-SMP/kernel/drivers/s390/scsi/zfcp.o

# echo "0x1605 0x1:0x5005076300cd9589 0x0:0x5303000000000000" >
> /proc/scsi/zfcp/add_map

# echo "scsi add-single-device 0 0 1 0" > /proc/scsi/scsi

# modprobe sd_mod
```

ibm.com/redbooks

Verify Device Configuration

```
# cat /proc/scsi/zfcp/map
0x1606 0x00000001:0x5005076300cd9589 0x00000000:0x5303000000000000


# cat /proc/scsi/scsi
Attached devices:
Host: scsi0 Channel: 00 Id: 01 Lun: 00
  Vendor: IBM      Model: 2105800       Rev: .104
  Type:   Direct-Access               ANSI SCSI revision: 03
```

Partition the SCSI Disk

```
# fdisk /dev/sda
Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-9535, default 1): 1
Last cylinder or +size or +sizeM or +sizeK (1-9535, default 9535): 9535

Command (m for help): p

Disk /dev/sda1: 64 heads, 32 sectors, 9535 cylinders
Units = cylinders of 2048 * 512 bytes

     Device Boot     Start       End     Blocks   Id  System
/dev/sda1              1        7250    7423984   83  Linux
/dev/sda2           7251        9536    2340864   83  Linux

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

Start Installation Using YaST

Execute yast command from SSH session

Things to note:

- YaST requires dasd module
  - Provide dummy parameters (**dasd=404** for instance)
  - No DASD need be defined to the guest
- Swap partition can be:
  - ECKD DASD
  - FCP-attached SCSI
  - VDISK
- If SCSI swap partition:
  - Map the device
  - Partition the disk

Provide Dummy DASD Configuration
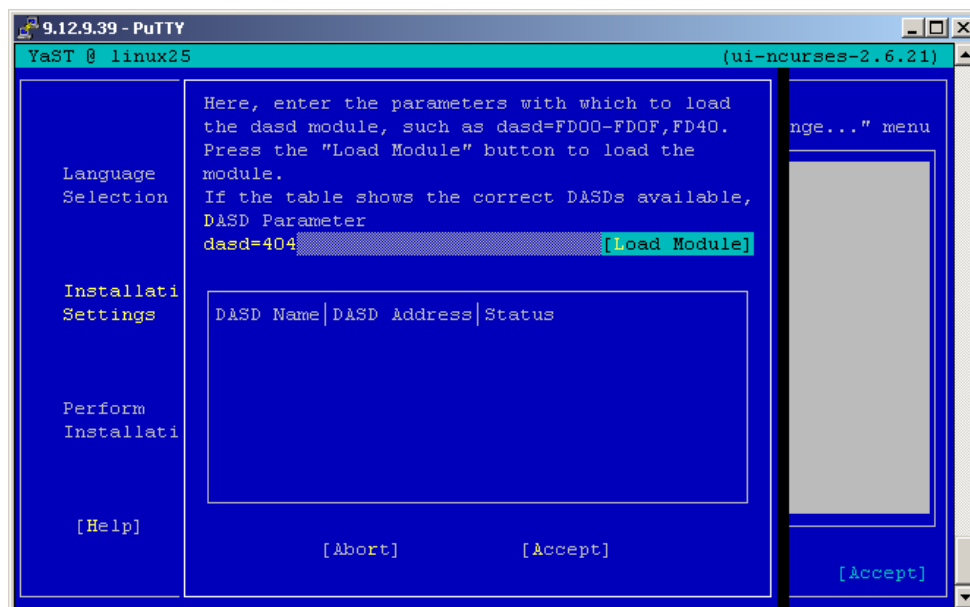
SLES8 Basic Installation

Installation proceeds normally until zipl execution
- SLES8 version of zipl is not FCP-aware
- Installation fails when zipl is executed

At this point:
- Open another SSH session
  - Use this session to apply SP3
  - Operates in chroot jail
- DO NOT click **Continue**!
  - Keep original YaST session open

Return to original YaST session after SP3 upgrade

ibm.com/redbooks

YaST Error Dialog Due to zipl
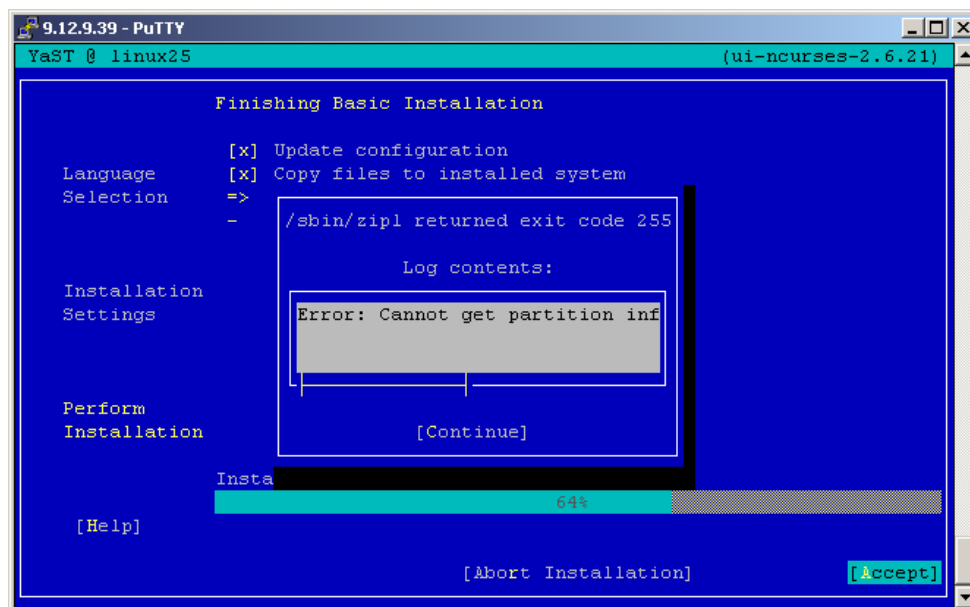
ibm.com/redbooks

Apply Service Pack 3 and Complete Install

In another SSH session:
- Execute YaST in chroot jail
- Select **System Update** to apply SP3
  - Be sure to point to SP3 source location

When update is complete:
- Close secondary SSH session
- Return to primary YaST session

From primary YaST session:
- Complete installation
- SP3 zipl version executes cleanly

Redbooks

ibm.com/redbooks

Execute YaST From chroot Jail

```
# mount
rootfs on / type rootfs (rw)
shmfs on / type shm (rw)
devpts on /dev/pts type devpts (rw)
virtual-proc-filesystem on /proc type proc (rw)
/dev/loop0 on /mounts/instsys type cramfs (ro)
/dev/sda1 on /mnt type ext3 (rw)
proc on /mnt/proc type proc (rw)

# chroot /mnt
# ls
.    bin   dev   home   lost+found   opt    root   srv   usr
..   boot  etc   lib    mnt          proc   sbin   tmp   var


# yast
```

Redbooks

ibm.com/redbooks

SP3 System Update From chroot Jail

```
9.12.9.39 - PuTTY                                         _ □ ×
YaST @ linux25                                (ui-ncurses-2.6.21) ▲

          Software Update


     The update
process recognizesto          Installed version:
packages that must            SuSE SLES 8
be updated. For some          Update version:
packages, the user            SuSE SLES 8 Service-Pack 3
must decide (withl            Update base packages:
the button "Detailed          no
selection") if the            Packages to update:
package should be             152
updated.                      Packages to check manually:
                              0

                                 [Detailed selection...]




         [Back]               [Abort]                   [Next]  ▼
```
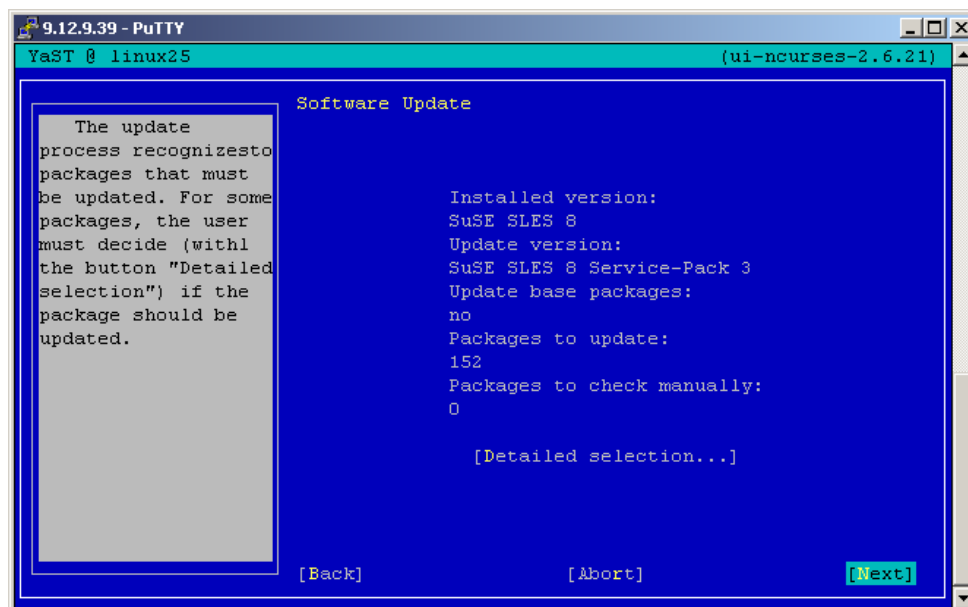
Finish Install From Primary YaST Session

```
9.12.9.39 - PuTTY                                         _ □ ×
YaST @ linux25                                (ui-ncurses-2.6.21) ▲

          Finishing Basic Installation

                  [x] Update configuration
     Language     [x] Copy files to installed system
     Selection    =>  Install boot manager
                  -   Prepare system for initial boot


     Installation
     Settings        Your system will now be shut down.
                     After shutdown, reload the system
                     with the load address of your root DASD.
                              [OK]

     Perform
     Installation


                  Installing boot manager...
                              64%

     [Help]

                     [Abort Installation]            [Accept]  ▼
```
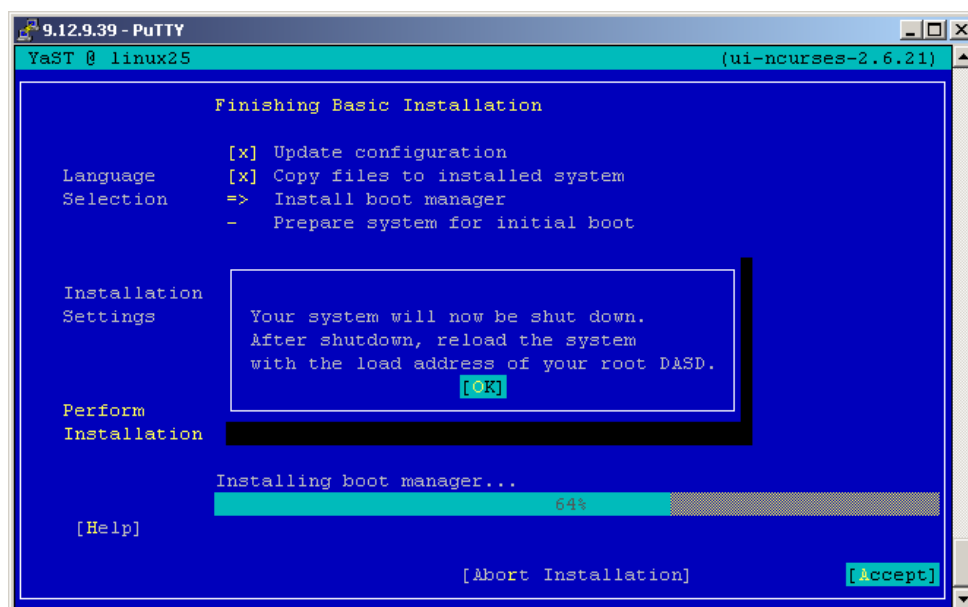
IPL From SCSI Disk

Provide WWPN and LUN to machine loader:

```
SET LOADDEV PORTNAME 50050763 00cd9589 LUN 53030000 00000000
```

IPL the FCP device:

```
I 1605
HCPLDI2816I Acquiring the machine loader from the processor
controller.
HCPLDI2817I Load completed from the processor controller.
HCPLDI2817I Now starting machine loader version 0001.
MLOEVL012I: Machine loader up and running (version 0.13).
MLOPDM003I: Machine loader finished, moving data to final storage
location.
Linux version 2.4.21-83-default (root@s390z04) (gcc version 3.2.2) #1
SMP
```

Redbooks

To Summarize

SLES8 SP3 is required for IPL from SCSI

- Cannot directly install SP3
  - Must upgrade existing SLES8 base installation

Modified installation can install / upgrade SLES on SCSI

- Start installation
- Load zfcp driver
- Install SLES8
  - Upgrade to SP3 *before* completing installation!
- Use SP3 version of zipl to complete install

Redbooks

Copy Existing SLES8 Installation to FCP

Copy SLES8 from ECKD DASD to FCP-attached SCSI disk

- Existing installation must have SP3 applied
- FCP device is defined to Linux guest

Steps to perform from Linux:

- Load SCSI device drivers
- Prepare SCSI disk
- Copy SLES8 installation to SCSI disk
- Make the SCSI partition bootable
- IPL from FCP device

Redbooks

Load SCSI Device Drivers

```
# lsmod
Module                  Size  Used by     Not tainted
qeth                  163376  1
ipv6                  329028  -1  [qeth]
key                    41840  0   [ipv6]
qdio                   35576  1   [qeth]
lvm-mod                70632  0   (autoclean)
dasd_eckd_mod          56676  2
dasd_mod               51092  3   [dasd_eckd_mod]
ext3                   94560  1
jbd                    55076  1   [ext3]

# modprobe scsi_mod

# insmod zfcp
Using /lib/modules/2.4.21-83-default/kernel/drivers/s390/scsi/zfcp.o

# echo "0x1617 0x02:0x200300a0b812106f 0x00:0x0009000000000000" >
  /proc/scsi/zfcp/add_map

# echo "scsi add-single-device 0 0 2 0" > /proc/scsi/scsi

# modprobe sd_mod
```

Redbooks

The SCSI Disk is Now Seen by Linux

```
# cat /proc/partitions
major minor  #blocks  name   ...
   8     0  16777216 sda    ...
  94     0   2259360 dasda  ...
  94     1   2259264 dasda1 ...
  94     4    144000 dasdb  ...
  94     5    143904 dasdb1 ...
```

Prepare the SCSI Disk

```
# fdisk /dev/sda
Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-9535, default 1): 1
Last cylinder or +size or +sizeM or +sizeK (1-9535, default 9535): 9535

Command (m for help): p

Disk /dev/sda1: 64 heads, 32 sectors, 9535 cylinders
Units = cylinders of 2048 * 512 bytes

     Device Boot    Start      End    Blocks   Id  System
/dev/sda1             1       7250   7423984   83  Linux
/dev/sda2          7251       9536   2340864   83  Linux

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

Create a Filesystem on the Partition

```
# mke2fs -j /dev/sda1
mke2fs 1.28 (31-Aug-2002)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
1154176 inodes, 2303996 blocks
115199 blocks (5.00%) reserved for the super user
First data block=0
71 block groups
32768 blocks per group, 32768 fragments per group
16256 inodes per group
Superblock backups stored on blocks:
        32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 24 mounts or
180 days, whichever comes first.  Use tune2fs -c or -i to override.
```

Copy SLES8 installation to SCSI disk

```
# mount /dev/sda1 /mnt


# mkdir /mnt/mnt /mnt/proc /mnt/tmp /mnt/sys


# cp -r --no-dereference --preserve=all --target-directory=/mnt \
/bin /dev /home /lib /lib64 /sbin /var \
/boot /etc /root /usr /opt /misc
cp: cannot stat `/lib64': No such file or directory
cp: cannot stat `/misc': No such file or directory


# chmod 1777 /mnt/tmp
```

Make the SCSI Partition Bootable

SCSI disk now has SLES8 installed

- /dev/sda1 is mounted on /mnt

To make it bootable:

- Change /mnt/etc/fstab
- Provide FCP mapping in /mnt/etc/zfcp.conf
- Configure /mnt/etc/zipl.conf
- Create initial ramdisk on /mnt/boot
- Execute zipl in /mnt chroot jail

Modified Configuration Files

```
# cat /mnt/etc/fstab
/dev/sda1               /               ext3        defaults        1 1
/dev/dasdb1     swap            swap        pri=42          0 0
devpts          /dev/pts        devpts      mode=0620,gid=5 0 0
proc            /proc           proc        defaults        0 0

# cat /mnt/etc/zfcp.conf
0x1617 0x02:0x200300a0b812106f 0x00:0x0009000000000000
# cat /mnt/etc/zipl.conf
# Generated by YaST2
[defaultboot]
default=ipl

[ipl]
target=/boot/zipl
image=/boot/kernel/image
ramdisk=/boot/initrd
parameters="dasd=201-202 root=/dev/sda1"

[dumpdasd]
target=/boot/zipl
dumpto=/dev/dasd??

[dumptape]
target=/boot/zipl
dumpto=/dev/rtibm0
```

Create Initial Ramdisk on /mnt/boot

```
# chroot /mnt

# mk_initrd
using "/dev/sda1" as root device (mounted on "/" as "ext3")

Found ECKD dasd, adding dasd eckd discipline!

Note: If you want to add ECKD dasd support for later mkinitrd
calls where possibly no ECKD dasd is found, add dasd_eckd_mod
to INITRD_MODULES in /etc/sysconfig/kernel

creating initrd "/boot/initrd" for kernel "/boot/kernel/image"
(version 2.4.21-83-default) (s390)

 - insmod qdio                (kernel/drivers/s390/qdio.o)
 - insmod scsi_mod            (kernel/drivers/scsi/scsi_mod.o)
 - insmod zfcp map="
0x1617 0x02:0x200300a0b812106f 0x00:0x0009000000000000"
(kernel/drivers/s390/scsi/zfcp.o)
 - insmod jbd                 (kernel/fs/jbd/jbd.o)
 - insmod ext3                (kernel/fs/ext3/ext3.o)
 - insmod dasd_mod dasd=$dasd (kernel/drivers/s390/block/dasd_mod.o)
 - insmod dasd_eckd_mod       (kernel/drivers/s390/block/dasd_eckd_mod.o)
 - insmod sd_mod              (kernel/drivers/scsi/sd_mod.o)
 - zfcp support

Run zipl now to update the IPL record!
```

ibm.com/redbooks

Run zipl in the /mnt chroot Jail

```
# zipl -V
Using config file '/etc/zipl.conf'
Target device information
  Device.........................: 08:00
  Partition......................: 08:01
  Type...........................: disk partition
  Disk layout....................: SCSI
  Geometry - heads...............: 64
  Geometry - sectors.............: 32
  Geometry - cylinders...........: 16384
  Geometry - start...............: 32
  File system block size.........: 4096
  Physical block size............: 512
  Device size in physical blocks..: 18431968
Building bootmap '/boot/zipl/bootmap'
Adding IPL section 'ipl' (default)
  kernel image......: /boot/kernel/image at 0x10000
  kernel parmline...: 'dasd=201-202 root=/dev/sda1' at 0x1000
  initial ramdisk...: /boot/initrd at 0x800000
Preparing boot device: 08:00Detected SCSI PCBIOS disk layout.
Writing SCSI master boot record.
Syncing disks...
Done.
```

ibm.com/redbooks

To Summarize

Copying SLES8 to SCSI is always an option
- Must upgrade to SP3 first

Steps:
- Load zcp driver
- Prepare SCSI disk
  - Parition, format, create fileystem
- Copy root filesystem to SCSI
- Modify configuration files
- Create initial ramdisk
- Run zipl

Installing on RHEL 3

RHEL 3 Update 2 is required for SCSI IPL
- Update 2 can be installed directly

SCSI enabled zipl is not distributed with U2
- Modified Anaconda installation is not possible

Must copy existing installation to SCSI disk
- Similar to SLES8 SP3 copy procedure
- Must build SCSI enabled zipl

The scsi_start script

```
# cat scsi_start
modprobe scsi_mod
map="0x1604 0x1:0x5005076300cd9589 0x0:0x5308000000000000"
modprobe zfcp map="$map"
modprobe sd_mod
cat /proc/scsi/scsi


# ./scsi_start
Attached devices:
Host: scsi0 Channel: 00 Id: 01 Lun: 00
  Vendor: IBM       Model: 2105800          Rev: .104
  Type:   Direct-Access                     ANSI SCSI revision: 03


# ls_mod
Module                  Size   Used by    Not tainted
sd_mod                  15220   0  (unused)
zfcp                    291956  0  (unused)
scsi_mod                125984  2  [sd_mod zfcp]
qdio                    40796   1  [qeth]
```

Redbooks

Create SCSI Device Nodes

```
# mknod -m 660 /dev/sda b 8 0

# mknod -m 660 /dev/sda1 b 8 1

# ls -l /dev/sd*
brw-rw----    1 root     root        8,    0 Apr 29 17:43 /dev/sda
brw-rw----    1 root     root        8,    1 Apr 29 17:43 /dev/sda1
```

Redbooks

Partition the SCSI Disk

```
# fdisk /dev/sda
Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-9535, default 1): 1
Last cylinder or +size or +sizeM or +sizeK (1-9535, default 9535): 9535

Command (m for help): p

Disk /dev/sda1: 64 heads, 32 sectors, 9535 cylinders
Units = cylinders of 2048 * 512 bytes

     Device Boot     Start        End     Blocks    Id  System
/dev/sda1               1       7250    7423984    83  Linux
/dev/sda2            7251       9536    2340864    83  Linux

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

SCSI Partition is Now Available

```
# cat /proc/partitions
major minor  #blocks  name

   8     0  9765632 sda
   8     1  9764848 sda1
  94     0  2259360 dasda
  94     1  2259264 dasda1
  94     4   144000 dasdb
  94     5   143904 dasdb1
```

Create Filesystem on SCSI Disk

```
# mke2fs -j /dev/sda1
mke2fs 1.32 (09-Nov-2002)
Filesystem label=
OS type: Linux
Block size=1024 (log=0)
Fragment size=1024 (log=0)
1220608 inodes, 9764848 blocks
488242 blocks (5.00%) reserved for the super user
First data block=1
1192 block groups
8192 blocks per group, 8192 fragments per group
1024 inodes per group
Superblock backups stored on blocks:
        8193, 24577, 40961, 57345, 73729, 204801, 221185, 401409, 663553,
        1024001, 1990657, 2809857, 5120001, 5971969
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done
This filesystem will be automatically checked every 22 mounts or
180 days, whichever comes first.  Use tune2fs -c or -i to override.
```

Redbooks

Copy ECKD Installation to SCSI Disk

```
# mkdir /mnt/disk


# mount /dev/sda1 /mnt/disk


# mkdir /mnt/disk/mnt /mnt/disk/proc /mnt/disk/tmp


# cp -v -r --no-dereference --preserve=all
--target-directory=/mnt/disk /bin /boot /dev /etc /home
/initrd /lib /misc /opt /root /sbin /usr /var
```

Redbooks

Customize Configuration Files

```
# cat /mnt/disk/boot/parmfile
root=/dev/sda1 ro


# cat /mnt/disk/etc/fstab
/dev/sda1/        /                    ext3    defaults        1 1
none             /dev/pts        devpts  gid=5,mode=620  0 0
none             /proc           proc    defaults        0 0
/dev/dasdb1      swap            swap    defaults        0 0


# cat /mnt/disk/etc/modules.conf
alias hsi0 qeth
options dasd_mod dasd=201,202
options scsi_mod max_scsi_luns=50
options zfcp 'map="0x1604 0x1:0x5005076300cd9589 0x0:0x5308000000000000"'
```

Create SCSI Initial Ramdisk

```
# chroot /mnt/disk


# cd boot


# cp initrd-2.4.21-11.EL.img initrd-2.4.21-11.EL.img.old


# rm initrd-2.4.21-11.EL.img
rm: remove regular file `initrd-2.4.21-11.EL.img'? y


# mkinitrd -v --with=scsi_mod --with=zfcp --with=sd_mod
initrd-2.4.21-11.EL.img 2.4.21-11.EL


# exit
```

Build SCSI-enabled zipl

```
# tar xvzpBf s390-tools-1\[1\].2.4-june2003.tar.gz

# cd ~/s390-tools-1.2.4/
# make
.
.
.
***************************************************************************
* ERROR: 'extern/linux-2.4.19.tar.bz2' missing
* ERROR: 'extern/busybox-0.60.5.tar.bz2' missing
* ERROR: 'extern/e2fsprogs-1.32.tar.gz' missing
* Call 'get_files.sh' to get the required tarballs from the internet!
***************************************************************************

# cd zfcpdump/

# ./get_files.sh

# cd ~/s390-tools-1.2.4/

# make
```

Execute SCSI-enabled zipl

```
# cd /mnt/disk/boot

# /root/s390-tools-1.2.4/zipl/src/zipl -V -t . -i vmlinuz-2.4.21-11.EL -p parmfile -r
initrd-2.4.21-11.EL.img
Target device information
  Device..........................: 08:00
  Partition.......................: 08:01
  Device name.....................: sda
  Type............................: disk partition
  Disk layout.....................: SCSI
  Geometry - heads................: 64
  Geometry - sectors..............: 32
  Geometry - cylinders............: 9536
  Geometry - start................: 32
  File system block size..........: 1024
  Physical block size.............: 512
  Device size in physical blocks..: 19529696
Building bootmap './bootmap'
Adding IPL section
  kernel image......: vmlinuz-2.4.21-11.EL at 0x10000
  kernel parmline...: 'root=/dev/sda1 ' at 0x1000
  initial ramdisk...: initrd-2.4.21-11.EL.img at 0x800000
Preparing boot device: sda.
Detected SCSI PCBIOS disk layout.
Writing SCSI master boot record.
Syncing disks...
done.
```

IPL SCSI Disk

```
SET LOADDEV PORTNAME 50050763 00cd9589 LUN 53080000 00000000
Ready; T=0.01/0.01 20:01:12


Q LOADDEV
PORTNAME 50050763 00CD9589    LUN  53080000 00000000     BOOTPROG 0
BR_LBA   00000000 00000000
Ready; T=0.01/0.01 20:02:01


I 1604
HCPLDI2816I Acquiring the machine loader from the processor controller.
HCPLDI2817I Load completed from the processor controller.
HCPLDI2817I Now starting machine loader version 0001.
MLOEVL012I: Machine loader up and running (version 0.13).
MLOPDM003I: Machine loader finished, moving data to final storage
location.
Linux version 2.4.21-11.EL (bhcompile@spade.z900.redhat.com) (gcc version
3.2.3
20030502 (Red Hat Linux 3.2.3-30)) #1 SMP Mon Mar 8 23:19:00 EST 2004
We are running under VM (31 bit mode)
```

Redbooks

ibm.com/redbooks

Adding SCSI Disks to a Running System

FCP devices automatically mapped to SCSI *only* when:

- Parameters supplied to zfcp on load
  - Using insmod map="..." parameters

Manually mapped devices must be added to SCSI subsystem

- Using /proc/scsi/scsi interface:
  - echo "scsi add-single-device *host bus target lun*" >/proc/scsi/scsi

Things to be aware for RHEL 3 for /etc/modules.conf

- No continuations lines
- Add:
  - options scsi_max_luns=50

Redbooks

ibm.com/redbooks

Map New LUNs

```
# cat /proc/scsi/zfcp/map
0x1604 0x00000001:0x5005076300cd9589 0x00000000:0x5307000000000000

# echo "0x1604 0x1:0x5005076300cd9589 0x1:0x5306000000000000" >
/proc/scsi/zfcp/add_map

# echo "scsi add-single-device 0 0 1 1" > /proc/scsi/scsi

# cat /proc/scsi/scsi
Attached devices:
Host: scsi0 Channel: 00 Id: 01 Lun: 00
  Vendor: IBM       Model: 2105800          Rev: .104
  Type:   Direct-Access                     ANSI SCSI revision: 03
Host: scsi0 Channel: 00 Id: 01 Lun: 01
  Vendor: IBM       Model: 2105800          Rev: .104
  Type:   Direct-Access                     ANSI SCSI revision: 03)

# cat /proc/partitions
major minor    #blocks name
    8     0    9765632 sda
    8     1    9764848 sda1
    8    16    9765632 sdb
    8    17    9764848 sdb1
```

Update the /etc/modules.conf File

```
# cat /etc/modules.conf
alias hsi0 qeth
options scsi_mod max_scsi_luns=50
options zfcp 'map="0x1604 0x1:0x5005076300cd9589 0x0:0x5307000000000000;
0x1604 0x1:0x5005076300cd9589 0x1:0x5306000000000000; 0x1604
0x1:0x5005076300cd9589 0x2:0x5308000000000000"'
```

Note:
  Map parameters must appear on a single line!

Prepare SCSI Disk for Next Boot

```
# cd /boot

# mkinitrd -v --with=scsi_mod --with=zfcp --with=sd_mod
initrd-2.4.21-11.EL.img 2.4.21-11.EL
.
.
Loading module scsi_mod with options max_scsi_luns=50
Loading module qdio
Loading module zfcp with options 'map="0x1604 0x1:0x5005076300cd9589
0x0:0x5307000000000000; 0x1604 0x1:0x5005076300cd9589
0x1:0x5306000000000000; 0x1604 0x1:0x5005076300cd9589
0x2:0x5308000000000000"'
Loading module sd_mod
.
.

# zipl -V -t . -i vmlinuz-2.4.21-11.EL -p parmfile -r initrd-2.4.21-11.EL.img
```

Add LUNs Without Using /etc/modules.conf

Adding LUNs can be error prone in RHEL 3

- No continuation lines make /etc/modules.conf hard to edit
- Manually adding to SCSI subsystem is tedious

Better to have automation

- Map FCP devices
- Add to SCSI subsystem

Define LUN Mapping

```
# cat /root/map.conf
0x1614 0x00000002:0x200200a0b812106f 0x00000000:0x000d000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000001:0x000e000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000002:0x000f000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000003:0x0010000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000004:0x0011000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000005:0x0012000000000000
0x1614 0x00000002:0x200200a0b812106f 0x00000006:0x0013000000000000


# cat /root/map.conf > /proc/scsi/zfcp/add_map
```

The add_scsi.sh Script

```
# cat add_scsi.sh
echo "scsi add-single-device 1 0 2 0" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 1" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 2" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 3" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 4" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 5" > /proc/scsi/scsi
echo "scsi add-single-device 1 0 2 6" > /proc/scsi/scsi
```

Define LUNs to SCSI stack

```
# /root/add_scsi.sh


# cat /proc/scsi/scsi
Attached devices:
Host: scsi0 Channel: 00 Id: 01 Lun: 00
  Vendor: IBM      Model: 2105800         Rev: .104
  Type:   Direct-Access                  ANSI SCSI revision: 03
Host: scsi0 Channel: 00 Id: 01 Lun: 01
  Vendor: IBM      Model: 2105800         Rev: .104
  Type:   Direct-Access                  ANSI SCSI revision: 03
Host: scsi0 Channel: 00 Id: 01 Lun: 02
  Vendor: IBM      Model: 2105800         Rev: .104
  Type:   Direct-Access                  ANSI SCSI revision: 03
.
.
.
```

Redbooks

ibm.com/redbooks

The **/root/mount_disks.sh** Script

```
# cat /root/mount_disks.sh
mount -t ext3 /dev/sdb1 /mnt/id01lun01
mount -t ext3 /dev/sdc1 /mnt/id01lun02
mount -t ext3 /dev/sdd1 /mnt/id02lun00
mount -t ext3 /dev/sde1 /mnt/id02lun01
mount -t ext3 /dev/sdf1 /mnt/id02lun02
mount -t ext3 /dev/sdg1 /mnt/id02lun03
mount -t ext3 /dev/sdh1 /mnt/id02lun04
mount -t ext3 /dev/sdi1 /mnt/id02lun05
mount -t ext3 /dev/sdj1 /mnt/id02lun06
```

Redbooks

ibm.com/redbooks

Automation for System Initialization

```
# cat rc.local
#!/bin/sh
#
# This script will be executed *after* all the other init scripts.
# You can put your own initialization stuff in here if you don't
# want to do the full Sys V style init stuff.

touch /var/lock/subsys/local

cat /root/map.conf > /proc/scsi/zfcp/add_map
. /root/add_scsi.sh
. /root/mount_disks.sh
```

Redbooks

To Summarize:

RHEL 3 Update 2 is required for SCSI IPL
- SCSI enabled zipl is not distributed with U2

SCSI installation procedure:
- Install on ECKD DASD
- Copy to SCSI
- Build zipl

Automation can simplify process
- Mapping FCP devices
- Define devices to SCSI subsystem

Redbooks

Installing on SCSI for SLES9

SLES9 uses 2.6-based kernel
- Introduces sysfs filesystems
  - Incorporates proc, devfs, devpty features
- FCP mapping becomes much simpler!

YaST for SLES9 has native SCSI support
- No need for customized YaST install
- Do not need to copy existing installation

ibm.com/redbooks
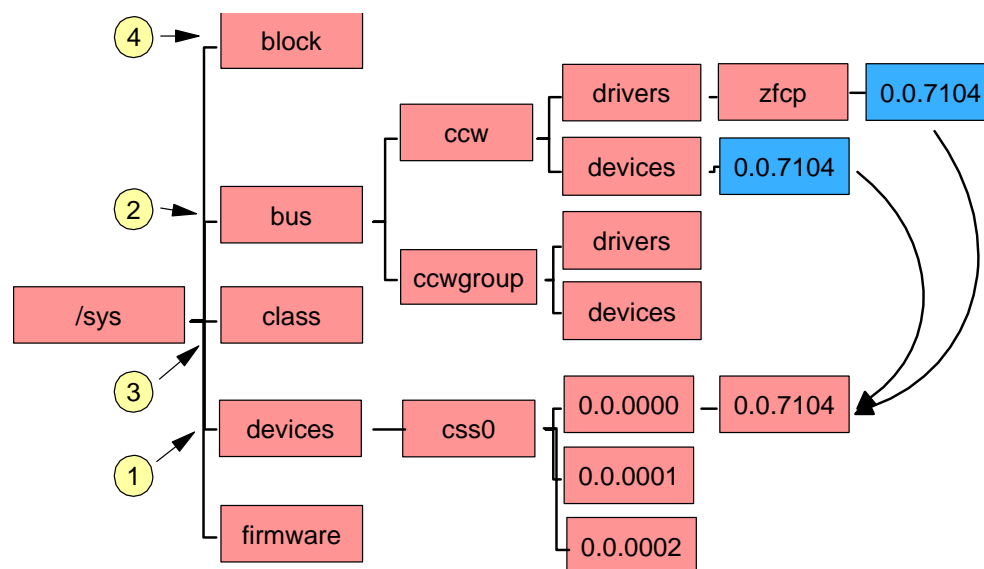
The /sysfs Filesystem

ibm.com/redbooks

The /sys Filesystem

/sys/devices/cssnnn

- Subchannels detected by Linux kernel

/sys/bus/css

- Devices and drivers associated with subchannels

/sysfs/class

- Groups similar devices (TTYs, tape drives, network devices, etc)

/sysfs/block

- Block devices on the system

OSA sysfs Device Configuration

```
First Hipersocket Channels that were detected:
Device Addresses CHPID(s)
          5000 63
          5001 63
          5002 63
          6000 64
          6001 64
          6002 64
          7104 ed
          7105 ed
          7106 ed
Possible configuration: read: 0x5000, write: 0x5001, data: 0x5002
Enter the device addresses for the qeth module, e.g.
'0.0.5000,0.0.5001,0.0.5002
  '
(0.0.5000,0.0.5001,0.0.5002): 0.0.7104,0.0.7105,0.0.7106
Device 0.0.7104 configured
qeth: Device 0.0.7104/0.0.7105/0.0.7106 is a HiperSockets card (level:
D3GF)
```
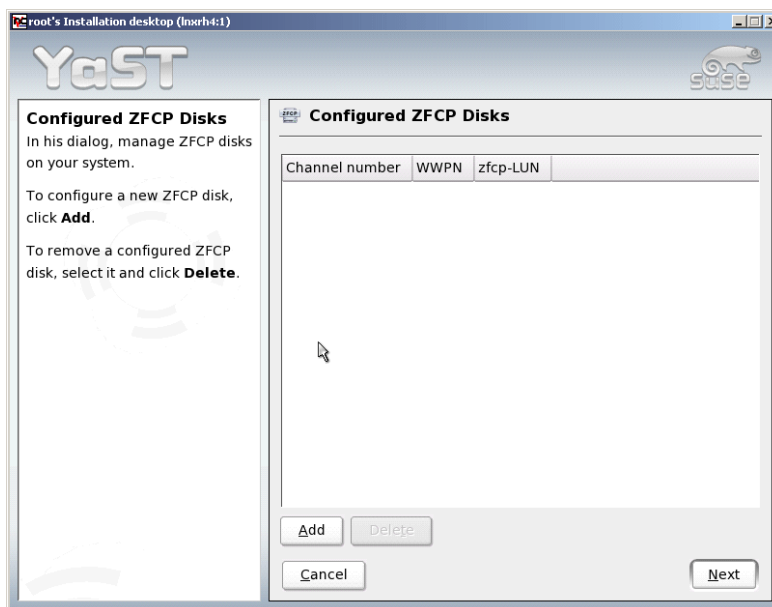
YaST - Configure Disk Controller
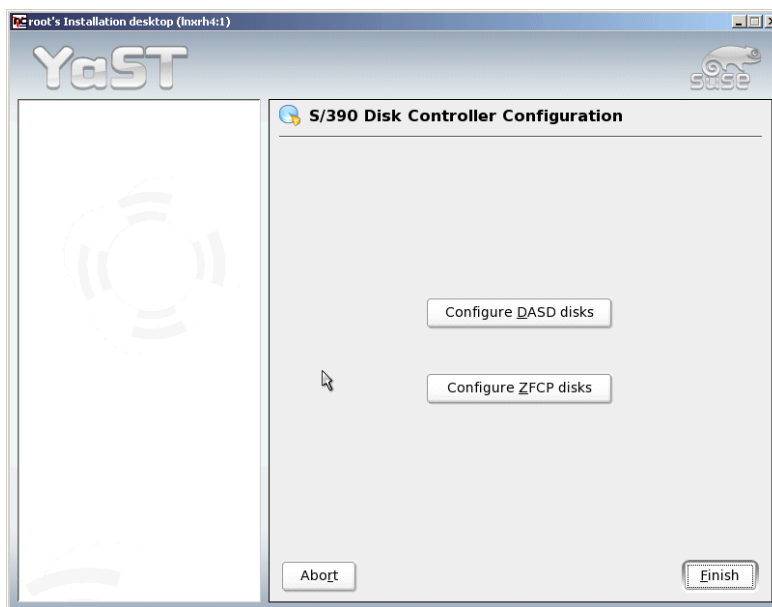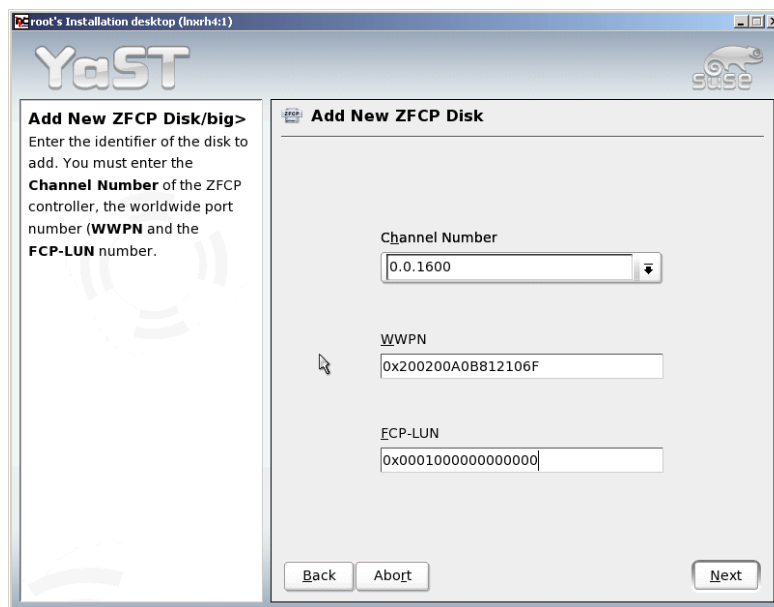
root's Installation desktop (lnxrh4:1)

YaST

S/390 Disk Controller Configuration

Configure DASD disks

Configure ZFCP disks

Abort   Finish

YaST - Configure ZFCP disks

root's Installation desktop (lnxrh4:1)

YaST

Configured ZFCP Disks
In his dialog, manage ZFCP disks on your system.

To configure a new ZFCP disk, click **Add**.

To remove a configured ZFCP disk, select it and click **Delete**.

Configured ZFCP Disks

Channel number	WWPN	zfcp-LUN

Add   Delete

Cancel   Next

YaST - Add New ZFCP Disk

Add New ZFCP Disk/big>
Enter the identifier of the disk to add. You must enter the **Channel Number** of the ZFCP controller, the worldwide port number (**WWPN** and the **FCP-LUN** number.

Add New ZFCP Disk
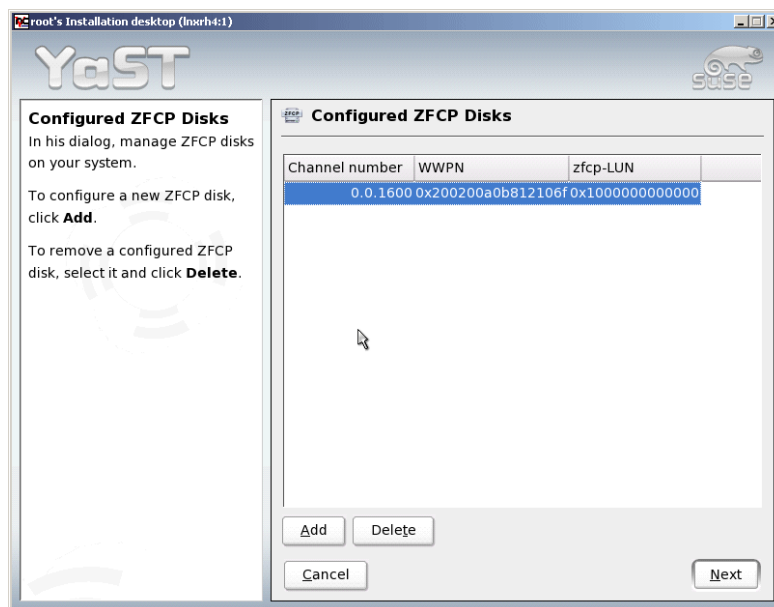
Channel Number
```
0.0.1600
```

WWPN
```
0x200200A0B812106F
```

FCP-LUN
```
0x0001000000000000
```

Back   Abort   Next

ibm.com/redbooks

YaST - FCP Disk Added

Configured ZFCP Disks
In his dialog, manage ZFCP disks on your system.

To configure a new ZFCP disk, click **Add**.

To remove a configured ZFCP disk, select it and click **Delete**.
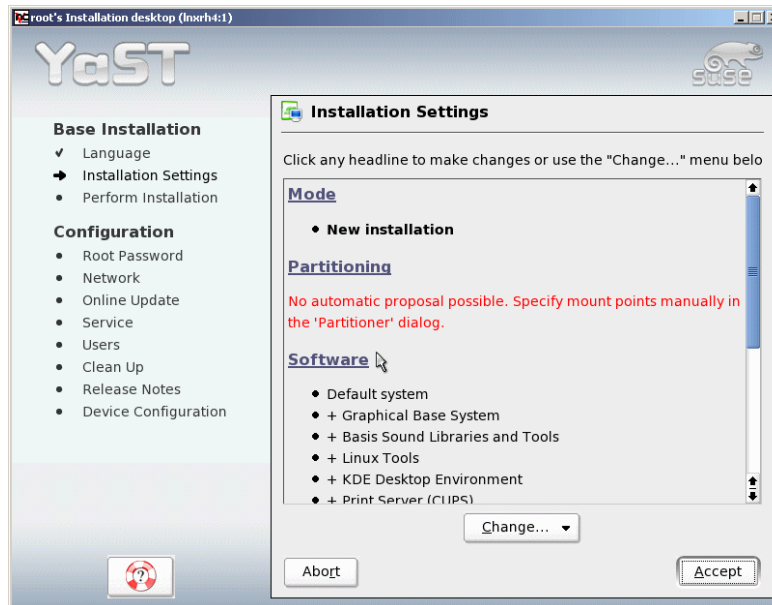
Configured ZFCP Disks

Channel number	WWPN	zfcp-LUN
0.0.1600	0x200200a0b812106f	0x1000000000000000

Add   Delete

Cancel   Next

ibm.com/redbooks

YaST - Choose Partitioning

root's Installation desktop (lnxrh4:1)

YaST

suse

Base Installation
✔ Language
➡ Installation Settings
• Perform Installation

Configuration
• Root Password
• Network
• Online Update
• Service
• Users
• Clean Up
• Release Notes
• Device Configuration

Installation Settings

Click any headline to make changes or use the "Change..." menu belo

Mode
• **New installation**

Partitioning
No automatic proposal possible. Specify mount points manually in the 'Partitioner' dialog.

Software
• Default system
• + Graphical Base System
• + Basis Sound Libraries and Tools
• + Linux Tools
• + KDE Desktop Environment
• + Print Server (CUPS)

Change... ▼

Abort   Accept

YaST - Expert Partitioner

root's Installation desktop (lnxrh4:1)

YaST

suse

For a root file system on SCSI disks, add a /boot partition on DASD to use for IPL.

The table to the right shows the current partitions on all your hard disks. **Nothing will be written to your hard disk** until you confirm the entire installation in the last installation dialog. Until that point, you can safely abort the installation.

Hard disks are designated like this:

/dev/dasda First DASD disk /dev/dasdb Second DASD disk /dev/dasdc Third DASD disk

- or -

/dev/sda First SCSI disk /dev/sdb

Expert Partitioner

Device	Id	Size	F	Type	Mount	Start	End	Us
/dev/sda	()	16.0 GB		1742-900		0	16383	
/dev/sda1		16.0 GB		Linux native		0	16383	

Create   Edit   Delete   dasdfmt

LVM...   EVMS...   RAID... ▼   Expert.. ▼

Back   Abort   Next

YaST - Set Partition Properties

Expert Partitioner

For a root file system on SCSI disks, add a /boot partition on DASD to use for IPL.

The table to the right shows the current partitions on all your hard disks. **Nothing will be written to your hard disk** until you confirm the entire installation in the last installation dialog. Until that point, you can safely abort the installation.

Hard disks are designated like this:

/dev/dasda First DASD disk /dev/dasdb Second DASD disk /dev/dasdc Third DASD disk

- or -

/dev/sda First SCSI disk /dev/sdb

Edit Existing Partition/dev/sda1

On already-existing partitions, you can change everything except the start and size of the partition.

Format
- Do not format
- Format

File system
Reiser

Options

Type of partition: Linux native
Start cylinder: 0
End cylinder: 16383

Fstab Options

Mount Point
/

OK   Cancel

Create   Edit   Delete   dasdfmt

LVM...   EVMS...   RAID... ▾   Expert... ▾

Back   Abort   Next

ibm.com/redbooks

YaST - Partition Configured

For a root file system on SCSI disks, add a /boot partition on DASD to use for IPL.

The table to the right shows the current partitions on all your hard disks. **Nothing will be written to your hard disk** until you confirm the entire installation in the last installation dialog. Until that point, you can safely abort the installation.

Hard disks are designated like this:

/dev/dasda First DASD disk /dev/dasdb Second DASD disk /dev/dasdc Third DASD disk

- or -

/dev/sda First SCSI disk /dev/sdb

Expert Partitioner

Device	Id	Size	F	Type	Mount	Start	E
/dev/sda	()	16.0 GB		1742-900		0	1
/dev/sda1		16.0 GB	F	Linux native (Reiser)	/	0	1

Create   Edit   Delete   dasdfmt

LVM...   EVMS...   RAID... ▾   Expert... ▾

Back   Abort   Next

ibm.com/redbooks

YaST - No Swap Partition

root's Installation desktop (lnxrh4:1)

Expert Partitioner

For a root file system on SCSI disks, add a /boot partition on DASD to use for IPL.

The table to the right shows the current p...
your hard dis...
be written t...
disk until you...
entire installa...
installation di...
point, you ca...
installation.

Hard disks a...
like this:

/dev/dasda First DASD disk /dev/dasdb Second DASD disk /dev/dasdc Third DASD disk

- or -

/dev/sda First SCSI disk /dev/sdb

Device	Id	Size	F	Type	Mount	Start	E
						0	1
						0	1

You have not assigned a swap partition. There is nothing wrong with that, but in most cases it is highly recommended to create and assign a swap partition. Swap partitions on your system are listed in the main window with the type "Linux Swap". An assigned swap partition has the mount point "swap". You can assign more than one swap partition, if desired.

Do you want to change this?

Yes   No

Create   Edit   Delete   dasdfmt

LVM...   EVMS...   RAID...   Expert..

Back   Abort   Next

Redbooks

YaST - Start Installation

root's Installation desktop (lnxrh4:1)

Installation Settings

Base Installation
✔ Language
➔ Installation S...
• Perform Insta...

Configuration
• Root Passwor...
• Network
• Online Update
• Service
• Users
• Clean Up
• Release Note...
• Device Config...

Click any headline to make changes or use the "Change..." menu belo...

Warning:
YaST2 has obtained all the information
required to install SUSE Linux.
The installation will be carried out according
to settings made in the previous dialogs.
To commit the installation and all choices made
so far, choose "Yes". Choose "No" to return
to the previous dialog.

Start installation?

Yes, install   No

+ KDE Desktop Environment
+ Print Server (CUPS)

Change...

Abort   Accept

Redbooks

YaST Admin Console

ibm.com/redbooks

Admin Console - Hardware

ibm.com/redbooks

ZFCP Disks Panel

Add new ZFCP Disk Panel

Configured ZFCP Disks Panel
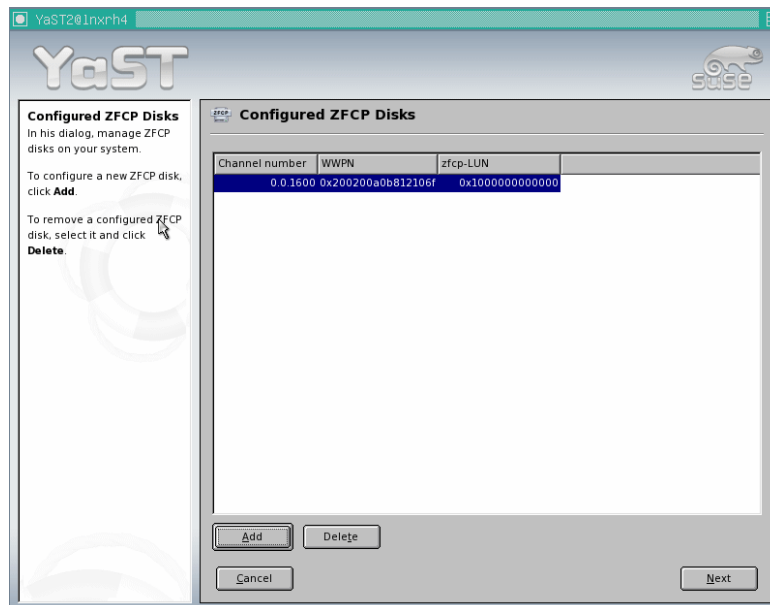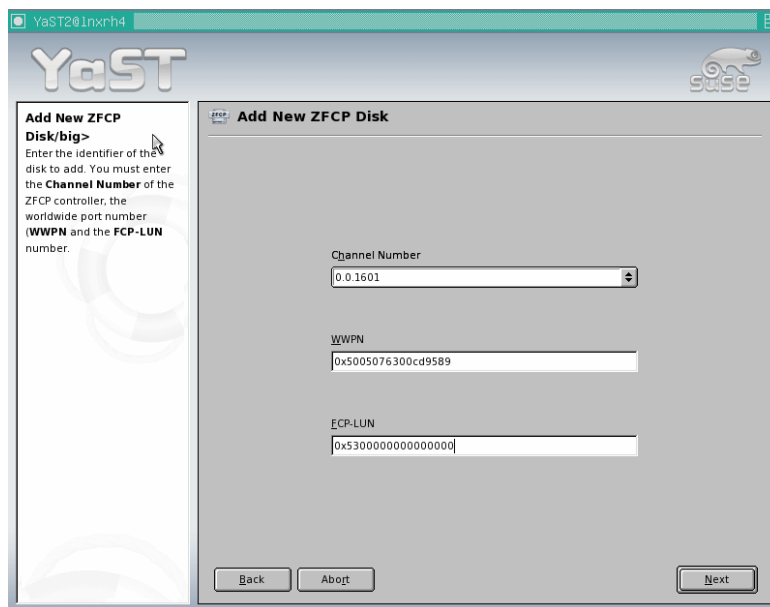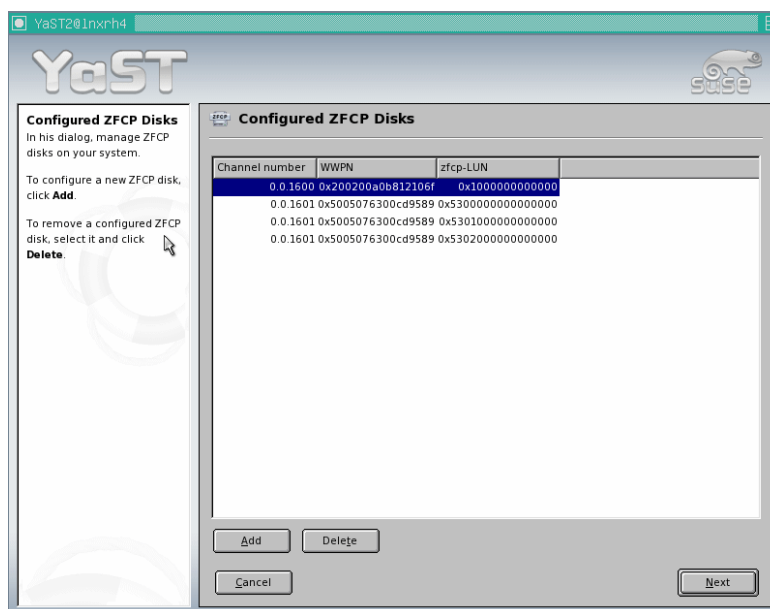
Messages From dmesg

```
scsi3 : zfcp
zfcp: The adapter 0.0.1601 reported the following characteristics:
WWNN 0x5005076400c10ecb, WWPN 0x50050764014013e2, S_ID 0x00660a00,
adapter version 0x2, LIC version 0x27, FC link speed 2 Gb/s
zfcp: Switched fabric fibrechannel network detected at adapter 0.0.1601.
zfcp: The remote port 0x0000000000000000 via adapter 0.0.1601 was opened,
it's port handle is 0x25f
zfcp: The remote port 0x5005076300cd9589 via adapter 0.0.1601 was opened,
it's port handle is 0x260
  Vendor: IBM        Model: 2105800          Rev: .104
  Type:   Direct-Access                      ANSI SCSI revision: 03
SCSI device sde: 19531264 512-byte hdwr sectors (10000 MB)
SCSI device sde: drive cache: write back
 sde: sde1
Attached scsi disk sde at scsi3, channel 0, id 1, lun 0
Attached scsi generic sg0 at scsi0, channel 0, id 1, lun 0,  type 0
Attached scsi generic sg1 at scsi3, channel 0, id 1, lun 0,  type 0
  Vendor: IBM        Model: 2105800          Rev: .104
  Type:   Direct-Access                      ANSI SCSI revision: 03
```

Manually Adding Disks

Hardware configuration directory:

- /etc/sysconfig/hardware

FCP configuration file:

- hwcfg-zfcp-bus-ccw-0.0.*nnnn*
  - Configuration file for FCP device *nnnn*
- Contains configuration values for the device
  - Startup scripts echo values to /sys filesystem

Checking for FCP Devices

```
# ls /sys/bus/ccw/drivers/zfcp/
.          0.0.1601  loglevel_cio     loglevel_fc     loglevel_qdio
..         0.0.1700  loglevel_config  loglevel_fsf    loglevel_scsi
0.0.1600   0.0.1701  loglevel_erp     loglevel_other  version
```

Configuration for FCP device 1601

```
# cat /etc/sysconfig/hardware/hwcfg-zfcp-bus-ccw-0.0.1601
#!/bin/sh
#
# hwcfg-zfcp-bus-ccw-0.0.1601
#
# Configuration for the zfcp adapter at CCW ID 0.0.1601
#

STARTMODE="auto"
MODULE="zfcp"
MODULE_OPTIONS=""
MODULE_UNLOAD="yes"

# Scripts to be called for the various events.
# If called manually the event is set to 'up'.
SCRIPTUP="hwup-ccw"
SCRIPTUP_ccw="hwup-ccw"
SCRIPTUP_scsi_host="hwup-zfcp"
SCRIPTDOWN="hwdown-scsi"
SCRIPTDOWN_scsi="hwdown-zfcp"

# Configured zfcp disks
ZFCP_LUNS="
0x5005076300cd9589:0x5300000000000000
0x5005076300cd9589:0x5301000000000000
0x5005076300cd9589:0x5302000000000000"
```

Redbooks

ibm.com/redbooks

Bringing FCP Device Online

```
# echo 1 > /sys/bus/ccw/drivers/zfcp/0.0.1601/online
```

Redbooks

ibm.com/redbooks

Adding LUNs Behind New WWPN

```
# ls /sys/bus/ccw/drivers/zfcp/0.0.1601
.                   cutype            fc_topology       online         status
..                  detach_state      hardware_version  port_add       wwnn
0x5005076300cd9589  devtype           host1             port_remove    wwpn
availability        failed            in_recovery       s_id
card_version        fc_link_speed     lic_version       scsi_host_no
cmb_enable          fc_service_class  nameserver        serial_number

# echo 0x200200a0b812106f > /sys/bus/ccw/drivers/zfcp/0.0.1601/port_add

# ls
.                   cutype            hardware_version  port_remove
..                  detach_state      host1             s_id
0x200200a0b812106f  devtype           in_recovery       scsi_host_no
0x5005076300cd9589  failed            lic_version       serial_number
availability        fc_link_speed     nameserver        status
card_version        fc_service_class  online            wwnn
cmb_enable          fc_topology       port_add          wwpn
```

Adding LUNs Behind New WWPN

```
# ls /sys/bus/ccw/drivers/zfcp/0.0.1601/0x200200a0b812106f
.   0x0002000000000000  detach_state  in_recovery  status     unit_remove
..  d_id                failed        scsi_id      unit_add   wwnn

# echo 0x0003000000000000 >
/sys/bus/ccw/drivers/zfcp/0.0.1601/0x200200a0b812106f/unit_add

# ls /sys/bus/ccw/drivers/zfcp/0.0.1601/0x200200a0b812106f
.   0x0002000000000000  d_id          failed       scsi_id  unit_add     wwnn
..  0x0003000000000000  detach_state  in_recovery  status   unit_remove
```

To Summarize:

SLES9 is SUS first 2.6 kernel distribution

- /sys filesystem simplifes dev, proc, udev filesystems

SLES9 supports installation to SCSI!

- No need to:
  - Modify installation process
  - Copy filesystems

Easy to configure SCSI devices

- /etc/sysconfig/harware configuration file
- /sys/bus/ccw/drivers/zfcp interface

Redbooks

Multipath FCP Access

ESCON / FICON

- Multiple connections to device handled transparently
  - Linux sees a single devie

FCP

- Each path to LUN appears as independent device
- Requires software support:
  - SLES8   Logical Volume Manager (LVM)
  - RHEL 3   Multiple Device Administrator (mdadm)
  - SLES9   Enterprise Volume Management System (EVMS)

Redbooks

FICON Multipathing

z990

Linux

| / | /usr | /var | /home | /opt |

/dev/dasda /dev/dasdc /dev/dasde
/dev/dasdb /dev/dasdd

| 7002 | 7102 | 7202 | 7302 | 7402 |

z/Series I/O subsystem

FICON

| CU 7000 | CU 7100 | CU 7200 | CU 7300 | CU 7400 |

unit 02 unit 02 unit 02 unit 02 unit 02

Filesystem sits directly on the device

Single Device node to access a volume

Single Device address to access a volume

I/O subsystem handles selecting from multiple paths

4 Paths to each device

FCP Multipathing

z990

Linux

| / | /usr | /var | /home | /opt |

/dev/v01 /dev/v02 /dev/v03 /dev/v04 /dev/v05

/dev/sda /dev/sdb /dev/sdc /dev/sdd /dev/sde
/dev/sdf /dev/sdg /dev/sdh /dev/sdi /dev/sdj
/dev/sdk /dev/sdl /dev/sdm /dev/sdn /dev/sdo
/dev/sdp /dev/sdq /dev/sdr /dev/sds /dev/sdt

| 6007 | 6107 | 6207 | 6307 |
| CU 6000 | CU 6100 | CU 6200 | CU 6300 |

z/Series I/O subsystem

FCP

FCP Switch

| WWPN 5005076300CD9589 | WWPN 5005076300C19589 | WWPN 5005076300C39589 | WWPN 5005076300C59589 |

lun 5002 lun 5102 lun 5202 lun 5302 lun 5402

Filesystem sits on LVM or RAID device

LVM or software RAID handles selecting from multiple paths

One Device node per path per LUN

Device addresses give access to FC Network

4 Paths to each device

Fully interconnected inside the storage system

RHEL 3 Multipath Implementation

mdadm provides redundancy only:
- Do not provide load balancing!
- Automatically detects active path failure
  - Brings secondary online
- If failed path comes back online:
  - mdadm detects it
  - Makes it secodary path

mdadm is verbose!
- Consider disconnecting VM console

RHEL 3 requires mdadm enabled initial ramdisk

Creating an mdadm Initial Ramdisk

```
# mkinitrd -v --with=scsi_mod --with=zfcp --with=sd_mod --with=dasd_mod --with=dasd_eckd_mod
--with=jbd --with=ext3 --with=qdio --with=raid0 initrd-2.4.21-11.EL.scsi.img 2.4.21-11.EL
Looking for deps of module ide-disk
Looking for deps of module dasd_mod
.
.
.
Using modules:  ./kernel/drivers/s390/block/dasd_mod.o ./kernel/drivers/s390/block/dasd_eckd_mod.o
./kernel/drivers/s390/block/dasd_fba_mod.o ./kernel/fs/jbd/jbd.o ./kernel/fs/ext3/ext3.o
./kernel/drivers/scsi/scsi_mod.o ./kernel/drivers/s390/qdio.o ./kernel/drivers/s390/scsi/zfcp.o
./kernel/drivers/scsi/sd_mod.o ./kernel/drivers/md/raid0.o
Using loopback device /dev/loop0
/sbin/nash -> /tmp/initrd.tZjk8Y/bin/nash
/sbin/insmod.static -> /tmp/initrd.tZjk8Y/bin/insmod
`/lib/modules/2.4.21-11.EL/./kernel/drivers/s390/block/dasd_mod.o' ->
`/tmp/initrd.tZjk8Y/lib/dasd_mod.o'
.
.
.Loading module dasd_mod with options dasd=201-202
Loading module dasd_eckd_mod
Loading module dasd_fba_mod
Loading module jbd
Loading module ext3
Loading module scsi_mod with options max_scsi_luns=50
Loading module qdio
Loading module zfcp with options 'map="0x1601 0x01:0x5005076300cd9589 0x00:0x5300000000000000;0x1601
0x01:0x5005076300cd9589 0x01:0x5301000000000000;0x1601 0x01:0x5005076300cd9589
0x02:0x5302000000000000;0x1701 0x01:0x5005076300c19589 0x00:0x5300000000000000;0x1701
0x01:0x5005076300c19589 0x01:0x5301000000000000;0x1701 0x01:0x5005076300c19589
0x02:0x5302000000000000"'
Loading module sd_mod
Loading module raid0
```

Point to Initial Ramdisk

```
# cat /etc/zipl.conf
[defaultboot]
default=linux
target=/boot
[linux]
        image=/boot/vmlinuz-2.4.21-11.EL
        ramdisk=/boot/initrd-2.4.21-11.scsi.EL
        parameters="root=LABEL=/"


# zipl
Using config file '/etc/zipl.conf'
Building bootmap '/boot//bootmap'
Building menu 'rh-automatic-menu'
Adding #1: IPL section 'linux' (default)
Preparing boot device: dasda (0201).
Done.
```

Redbooks

ibm.com/redbooks

Configuring mdadm

After restarting Linux guest:

- Create SCSI device nodes:
  ```
  mknod /dev/sda  b 8 0
  mknod /dev/sda1 b 8 1
  ```
  .
  .
  .
- Create RAID partition:
  ```
  fdisk /dev/sda
  ```
  - Choose Linux raid auto (fd) partition type
- Create /etc/mdadm.conf

Redbooks

ibm.com/redbooks

The /etc/mdadm.conf File

```
# cat /etc/mdadm.conf
DEVICE /dev/sd[abcdef]1 /dev/md1 /dev/md2 /dev/md3
ARRAY /dev/md1 devices=/dev/sda1,/dev/sdd1
ARRAY /dev/md2 devices=/dev/sdb1,/dev/sde1
ARRAY /dev/md3 devices=/dev/sdc1,/dev/sdf1
ARRAY /dev/md0 devices=/dev/md1,/dev/md2,/dev/md3


# cat /proc/mdstat
Personalities : [raid0]
read_ahead not set
Event: 0
unused devices: <none>
```

Creating Device Arrays

```
# mdadm -C /dev/md1 --level=multipath --raid-devices=2 /dev/sda1 /dev/sdd1
mdadm: /dev/sda1 appears to contain a reiserfs file system
    size = 9764800K
mdadm: /dev/sdd1 appears to contain a reiserfs file system
    size = 9764800K
Continue creating array? yes
mdadm: array /dev/md1 started.
.
.
.

# mdadm -C /dev/md0 --level=raid0 --raid-devices=3 /dev/md1 /dev/md2
/dev/md3
mdadm: /dev/md1 appears to contain a reiserfs file system
    size = 9764800K
mdadm: /dev/md2 appears to contain a reiserfs file system
    size = 9764848K
mdadm: /dev/md3 appears to contain an ext2fs file system
    size=9764848K  mtime=Wed Dec 31 19:00:00 1969
Continue creating array? yes
mdadm: array /dev/md0 started.
```

Create Filesystem on RAID0 Device

```
# mke2fs -j /dev/md0
mke2fs 1.32 (09-Nov-2002)
Filesystem label=
OS type: Linux
Block size=1024 (log=0)
.
.
.


# mount /dev/md0 /mnt
# df -h
Filesystem              Size  Used Avail Use% Mounted on
/dev/dasda1             2.2G  1.4G  726M  65% /
/dev/md0                 28G  8.1M   27G   1% /mnt
```

Enable RAID Array at Startup

```
# cat /etc/rc.d/rc.sysinit
.
.
.
if [  ! -x /etc/mdadm.conf ]; then
       echo "Starting mdadm RAID devices."
       /sbin/mdadm -As
fi
.
.
.
# cat /etc/fstab
LABEL=/      /                    ext3    defaults      1 1
none         /dev/pts             devpts  gid=5,mode=620  0 0
none         /proc                proc    defaults      0 0
/dev/dasdb1  swap                 swap    defaults      0 0
/dev/md0     /mnt                 ext3    defaults      1 2
```

Show RAID Status

```
# mdadm --detail /dev/md1
/dev/md1:
            Version : 00.90.00
      Creation Time : Tue May  4 18:05:27 2004
         Raid Level : multipath
         Array Size : 9764736 (9.31 GiB 9.100 GB)
       Raid Devices : 1
      Total Devices : 2
    Preferred Minor : 1
        Persistence : Superblock is persistent

        Update Time : Wed May  5 16:05:45 2004
              State : dirty, no-errors
     Active Devices : 1
    Working Devices : 2
     Failed Devices : 0
      Spare Devices : 1


        Number   Major   Minor   RaidDevice State
           0       8        1        0      active sync   /dev/sda1
           1       8       49        1      spare         /dev/sdd1
               UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
             Events : 0.32
```

Device Status After /dev/sda1 Failure

```
# mdadm --detail /dev/md1
/dev/md1:
            Version : 00.90.00
      Creation Time : Tue May  4 18:05:27 2004
         Raid Level : multipath
         Array Size : 9764736 (9.31 GiB 9.100 GB)
       Raid Devices : 1
      Total Devices : 2
    Preferred Minor : 1
        Persistence : Superblock is persistent

        Update Time : Wed May  5 16:21:32 2004
              State : dirty, no-errors
     Active Devices : 1
    Working Devices : 1
     Failed Devices : 1
      Spare Devices : 0


        Number   Major   Minor   RaidDevice State
           0       8       49        0      active sync   /dev/sdd1
           1       8        1        1      faulty        /dev/sda1
               UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
             Events : 0.33
```

Path to /dev/sda1 Recovery

```
# mdadm --detail /dev/md1
/dev/md1:
        Version : 00.90.00
  Creation Time : Tue May  4 18:05:27 2004
     Raid Level : multipath
     Array Size : 9764736 (9.31 GiB 9.100 GB)
   Raid Devices : 1
  Total Devices : 2
Preferred Minor : 1
    Persistence : Superblock is persistent

    Update Time : Wed May  5 16:27:08 2004
          State : dirty, no-errors
 Active Devices : 1
Working Devices : 2
 Failed Devices : 0
  Spare Devices : 1


    Number   Major   Minor   RaidDevice State
       0       8      49        0      active sync   /dev/sdd1
       1       8       1        1      spare         /dev/sda1
           UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
         Events : 0.35
```

Forcing Path Failure

```
# mdadm /dev/md1 -f /dev/sdd1
mdadm: set /dev/sdd1 faulty in /dev/md1

# mdadm --detail /dev/md1
/dev/md1:
        Version : 00.90.00
  Creation Time : Tue May  4 18:05:27 2004
     Raid Level : multipath
     Array Size : 9764736 (9.31 GiB 9.100 GB)
   Raid Devices : 1
  Total Devices : 2
Preferred Minor : 1
    Persistence : Superblock is persistent

    Update Time : Wed May  5 16:29:11 2004
          State : dirty, no-errors
 Active Devices : 1
Working Devices : 2
 Failed Devices : 0
  Spare Devices : 1

    Number   Major   Minor   RaidDevice State
       0       8       1        0      active sync   /dev/sda1
       1       8      49        1      spare         /dev/sdd1
           UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
         Events : 0.37
```

Removing Spare Paths

```
# mdadm /dev/md1 -r /dev/sdd1
mdadm: hot removed /dev/sdd1

# mdadm --detail /dev/md1
/dev/md1:
          Version : 00.90.00
    Creation Time : Tue May  4 18:05:27 2004
       Raid Level : multipath
       Array Size : 9764736 (9.31 GiB 9.100 GB)
     Raid Devices : 1
    Total Devices : 1
Preferred Minor : 1
      Persistence : Superblock is persistent

      Update Time : Wed May  5 16:31:05 2004
            State : dirty, no-errors
    Active Devices : 1
   Working Devices : 1
    Failed Devices : 0
     Spare Devices : 0

      Number   Major   Minor   RaidDevice State
         0       8       1        0        active sync   /dev/sda1
             UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
            Events : 0.38
```

Adding Spare Paths

```
# mdadm /dev/md1 -a /dev/sdd1
mdadm: hot added /dev/sdd1

# mdadm --detail /dev/md1
/dev/md1:
          Version : 00.90.00
    Creation Time : Tue May  4 18:05:27 2004
       Raid Level : multipath
       Array Size : 9764736 (9.31 GiB 9.100 GB)
     Raid Devices : 1
    Total Devices : 2
Preferred Minor : 1
      Persistence : Superblock is persistent

      Update Time : Wed May  5 16:32:56 2004
            State : dirty, no-errors
    Active Devices : 1
   Working Devices : 2
    Failed Devices : 0
     Spare Devices : 1

      Number   Major   Minor   RaidDevice State
         0       8       1        0        active sync   /dev/sda1
         1       8      49        1        spare         /dev/sdd1
             UUID : f0b960a7:4c5ee77b:f439e129:92d7887e
            Events : 0.39
```

To Summarize:

RHEL 3 dos not have SCSI nodes defined by default

- Create nodes using mknod command

Partition RAID devices as "Linux raid auto"

- Code fd in fdisk

Define arrays in /etc/mdadm.conf

Manage arrays using mdadm command

- --details   for detailed status
- -C   to create
- -f   to force a path
- -r   to remove a path
- -a   to add a path back

Redbooks

SLES8 Multipath Implementation

Multipathing handled by LVM in SLES8

- LUNs defined as physical volumes
- Physical volumes are added to Volume Group
- Extra paths added using pvpath command

LVM enables both:

- Redundancy
- Load balancing

Redbooks

The /etc/zfcp.conf File

```
# cat /etc/zfcp.conf
0x1602 0x01:0x200200a0b812106f 0x00:0x0001000000000000
0x1602 0x01:0x200200a0b812106f 0x01:0x0002000000000000
0x1602 0x01:0x200200a0b812106f 0x02:0x0003000000000000
0x1702 0x01:0x200300a0b812106f 0x00:0x0001000000000000
0x1702 0x01:0x200300a0b812106f 0x01:0x0002000000000000
0x1702 0x01:0x200300a0b812106f 0x02:0x0003000000000000
```

Note:
  2 FCP devices - 0x1602 / 0x1702
  3 LUNs

Managing Devices With LVM

Steps:
- Create physical volumes
- Create volume group
- Create logical volume
- Enable extra paths to LUNs

To make extra paths persistent
- Use pvpathsave save
- Add pvpathrestore at system initialization

Create Physical Volumes / Volume Group

```
# pvcreate /dev/sd[abc]1
pvcreate -- physical volume "/dev/sda1" successfully created
pvcreate -- physical volume "/dev/sdb1" successfully created
pvcreate -- physical volume "/dev/sdc1" successfully created


# vgcreate vg01 /dev/sd[abc]1
vgcreate -- INFO: using default physical extent size 4 MB
vgcreate -- INFO: maximum logical volume size is 255.99 Gigabyte
vgcreate -- doing automatic backup of volume group "vg01"
vgcreate -- volume group "vg01" successfully created and activated
```

Redbooks

ibm.com/redbooks

Define Logical Volume

```
# vgdisplay vg01
--- Volume group ---
VG Name                vg01
VG Access              read/write
VG Status              available/resizable
VG #                   0
MAX LV                 256
Cur LV                 0
Open LV                0
MAX LV Size            255.99 GB
Max PV                 256
Cur PV                 3
Act PV                 3
VG Size                47.98 GB
PE Size                4 MB
Total PE               12282
Alloc PE / Size        0 / 0
Free  PE / Size        12282 / 47.98 GB
VG UUID                hvka5k-cKvz-0X1J-P6Go-qWPn-z5z0-WoS98k

# lvcreate -i 3 -l 12282 -n lvol01 vg01
lvcreate -- INFO: using default stripe size 16 KB
lvcreate -- doing automatic backup of "vg01"
lvcreate -- logical volume "/dev/vg01/lvol01" successfully created
```

Redbooks

ibm.com/redbooks

Query LVM Volumes

```
# cat /proc/lvm/global
LVM module LVM version 1.0.5+(mp-v6c)(22/07/2002)

Total:  1 VG  3 PVs  1 LV (0 LVs open)
Global: 202877 bytes malloced   IOP version: 10   4:01:15 active

VG:  vg01  [3 PV, 1 LV/0 open]  PE Size: 4096 KB
  Usage [KB/PE]: 50307072 /12282 total  50307072 /12282 used  0 /0 free
  PVs: [AA] sda1                  16769024 /4094   16769024 /4094        0 /0
       +-- sdd1
       [AA] sdb1                  16769024 /4094   16769024 /4094        0 /0
       +-- sde1
       [AA] sdc1                  16769024 /4094   16769024 /4094        0 /0
       +-- sdf1
    LV:  [AWDS3 ] lvol01                     50307072 /12282    close
```

Enable Extra Paths to LUNs

```
# pvpath -q -p1 -ey vg01 1
vg01: setting state of path #1 of PV#1 to enabled
Physical volume /dev/sda1 of vg01 has 2 paths:
      Device  Weight Failed Pending State
 #  0:   8:1      0      0       0 enabled
 #  1:   8:49     0      0       0 enabled

# pvpath -q -p1 -ey vg01 2
vg01: setting state of path #1 of PV#2 to enabled
Physical volume /dev/sdb1 of vg01 has 2 paths:
      Device  Weight Failed Pending State
 #  0:   8:17     0      0       0 enabled
 #  1:   8:65     0      0       0 enabled

# pvpath -q -p1 -ey vg01 3
vg01: setting state of path #1 of PV#3 to enabled
Physical volume /dev/sdc1 of vg01 has 2 paths:
      Device  Weight Failed Pending State
 #  0:   8:33     0      0       0 enabled
 #  1:   8:81     0      0       0 enabled
```

Save Path Configuration

```
# pvpathsave


# cat /etc/pvpath.cfg
VG:vg01:disabled
PV:1:2:/dev/sda1
PATH:0:8-1:0:enabled
PATH:1:8-49:0:enabled
PV:2:2:/dev/sdb1
PATH:0:8-17:0:enabled
PATH:1:8-65:0:enabled
PV:3:2:/dev/sdc1
PATH:0:8-33:0:enabled
PATH:1:8-81:0:enabled
```

Bring Paths Online at System Boot

```
# cat /etc/rc.d/boot.local
#! /bin/sh
#
# Copyright (c) 2002 SuSE Linux AG Nuernberg, Germany.  All rights
reserved.
#
# Author: Werner Fink <werner@suse.de>, 1996
#         Burchard Steinbild <feedback@suse.de>, 1996
#
# /etc/init.d/boot.local
#
# script with local commands to be executed from init on system startup
#
# Here you should add things, that should happen directly after booting
# before we're going to the first run level.
#
pvpathrestore
```

Create Filesystem on Logical Volume

```
# mke2fs -j /dev/vg01/lvol01
mke2fs 1.28 (31-Aug-2002)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
.
.
.
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 34 mounts or
180 days, whichever comes first.  Use tune2fs -c or -i to override.

# cat /etc/fstab
/dev/dasda1             /                   ext3        defaults            1 1
/dev/dasdb1             swap                swap        pri=42              0 0
devpts                  /dev/pts            devpts      mode=0620,gid=5     0 0
proc                    /proc               proc        defaults            0 0
/dev/vg01/lvol01        /mnt                ext3        defaults            1 2
```

Managing LVM RAID Arrays

```
# pvpath -qa
Physical volume /dev/sda1 of vg01 has 2 paths:
       Device  Weight Failed Pending State
 #  0:   8:1        0      0        0 enabled
 #  1:   8:49       0      0        0 enabled
Physical volume /dev/sdb1 of vg01 has 2 paths:
       Device  Weight Failed Pending State
 #  0:   8:17       0      0        0 enabled
 #  1:   8:65       0      0        0 enabled
Physical volume /dev/sdc1 of vg01 has 2 paths:
       Device  Weight Failed Pending State
 #  0:   8:33       0      0        0 enabled
 #  1:   8:81       0      0        0 enabled
```

Enabling and Disabling Paths

```
# pvpath -p1 -en vg01 1
vg01: setting state of path #1 of PV#1 to disabled


# pvpath -p1 -ey vg01 1
vg01: setting state of path #1 of PV#1 to enabled
```

SLES9 Multipath Implementation

Multipathing handled by EVMS in SLES9
- Works in conjunction with 2.6 kernel device mapper

Multiple device (MD) multipath array:
- Uses each path to LUN
- RAID array uses the MD multipath devices

Hardware Configuration

```
# cat /etc/sysconfig/hardware/hwcfg-zfcp-bus-ccw-0.0.1601
#!/bin/sh
#
# hwcfg-zfcp-bus-ccw-0.0.1601
#
# Configuration for the zfcp adapter at CCW ID 0.0.1601
#
STARTMODE="auto"
MODULE="zfcp"
MODULE_OPTIONS=""
MODULE_UNLOAD="yes"
# Scripts to be called for the various events.
# If called manually the event is set to 'up'.
SCRIPTUP="hwup-ccw"
SCRIPTUP_ccw="hwup-ccw"
SCRIPTUP_scsi_host="hwup-zfcp"
SCRIPTDOWN="hwdown-scsi"
SCRIPTDOWN_scsi="hwdown-zfcp"
# Configured zfcp disks
ZFCP_LUNS="
0x5005076300cd9589:0x5300000000000000
0x5005076300cd9589:0x5301000000000000
0x5005076300cd9589:0x5302000000000000"
```

Configuration for 1701 Device

```
# cat hwcfg-zfcp-bus-ccw-0.0.1701
#!/bin/sh
#
# hwcfg-zfcp-bus-ccw-0.0.1701
#
# Configuration for the zfcp adapter at CCW ID 0.0.1701
#
STARTMODE="auto"
MODULE="zfcp"
MODULE_OPTIONS=""
MODULE_UNLOAD="yes"
# Scripts to be called for the various events.
# If called manually the event is set to 'up'.
SCRIPTUP="hwup-ccw"
SCRIPTUP_ccw="hwup-ccw"
SCRIPTUP_scsi_host="hwup-zfcp"
SCRIPTDOWN="hwdown-scsi"
SCRIPTDOWN_scsi="hwdown-zfcp"
# Configured zfcp disks
ZFCP_LUNS="
0x5005076300c19589:0x5300000000000000
0x5005076300c19589:0x5301000000000000
0x5005076300c19589:0x5302000000000000"
```

Vary FCP Devices Online

```
# ls /sys/bus/ccw/drivers/zfcp/0.0.160
.               cutype         fc_service_class  online        serial_number
..              detach_state   fc_topology       port_add      status
availability    devtype        hardware_version  port_remove   wwnn
card_version    failed         in_recovery       s_id          wwpn
cmb_enable      fc_link_speed  lic_version       scsi_host_no

# echo 1 > /sys/bus/ccw/drivers/zfcp/0.0.1601/online
# ls /sys/bus/ccw/drivers/zfcp/0.0.1601
.                   cutype         fc_topology       online        status
..                  detach_state   hardware_version  port_add      wwnn
0x5005076300cd9589  devtype        host1             port_remove   wwpn
availability        failed         in_recovery       s_id
card_version        fc_link_speed  lic_version       scsi_host_no
cmb_enable          fc_service_class  nameserver     serial_number

# echo 1 > /sys/bus/ccw/drivers/zfcp/0.0.1701/online
# ls /sys/bus/ccw/drivers/zfcp/0.0.1701
.                   cutype         fc_topology       online        status
..                  detach_state   hardware_version  port_add      wwnn
0x5005076300c19589  devtype        host2             port_remove   wwpn
availability        failed         in_recovery       s_id
card_version        fc_link_speed  lic_version       scsi_host_no
cmb_enable          fc_service_class  nameserver     serial_number
```

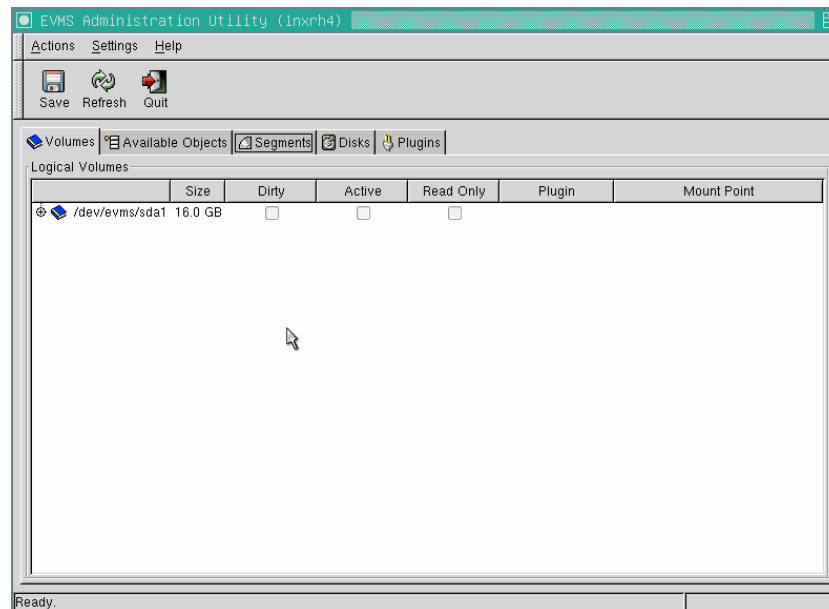**Redbooks**

ibm.com/redbooks

Vary Online Log Messages

```
scsi1 : zfcp
zfcp: The adapter 0.0.1601 reported the following characteristics:
WWNN 0x5005076400c10ecb, WWPN 0x50050764014013e2, S_ID 0x00660a00,
adapter version 0x2, LIC version 0x27, FC link speed 2 Gb/s
zfcp: Switched fabric fibrechannel network detected at adapter 0.0.1601.
zfcp: The remote port 0x0000000000000000 via adapter 0.0.1601 was opened, it's port
handle is 0x29f
zfcp: The remote port 0x5005076300cd9589 via adapter 0.0.1601 was opened, it's port
handle is 0x2a0
  Vendor: IBM       Model: 2105800          Rev: .104
  Type:   Direct-Access                     ANSI SCSI revision: 03
SCSI device sdb: 19531264 512-byte hdwr sectors (10000 MB)
SCSI device sdb: drive cache: write back
 sdb: unknown partition table
Attached scsi disk sdb at scsi1, channel 0, id 1, lun 0
  Vendor: IBM       Model: 2105800          Rev: .104
  Type:   Direct-Access                     ANSI SCSI revision: 03
.
.
.
```
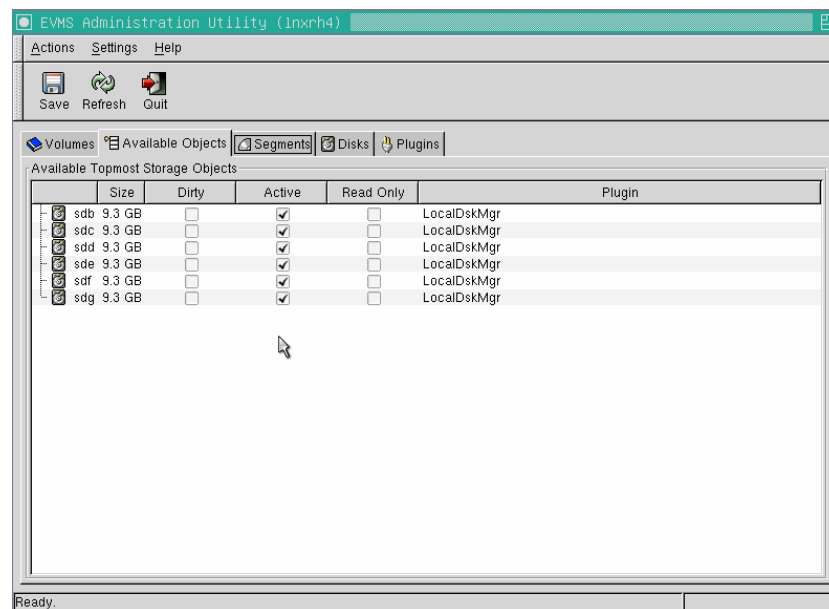
Redbooks

ibm.com/redbooks

Start EVMS Management Tool

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Logical Volumes

	Size	Dirty	Active	Read Only	Plugin	Mount Point
/dev/evms/sda1	16.0 GB	☐	☐	☐		

Ready.

Locate Available Objects

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Available Topmost Storage Objects

	Size	Dirty	Active	Read Only	Plugin
sdb	9.3 GB	☐	☑	☐	LocalDskMgr
sdc	9.3 GB	☐	☑	☐	LocalDskMgr
sdd	9.3 GB	☐	☑	☐	LocalDskMgr
sde	9.3 GB	☐	☑	☐	LocalDskMgr
sdf	9.3 GB	☐	☑	☐	LocalDskMgr
sdg	9.3 GB	☐	☑	☐	LocalDskMgr

Ready.

Locate Segments

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Disk Segments

	Size	Dirty	Active	Read Only	Type	Offset
DosSegMgr						
sda_mbr	16 KB	☐	☐	☐	Meta Data	0
sda1	16.0 GB	☐	☐	☐	Data	32

Ready.

ibm.com/redbooks

Add Segment Manager

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Create
Delete
Expand
Shrink   Objects | Segments | Disks | Plugins
Add   ►   Feature to Volume...
Remove   Segment Manager to Storage Object...   Plugin
Modify   LocalDskMgr
View   ☑   LocalDskMgr
Convert   ☑   LocalDskMgr
Replace...   ☑   LocalDskMgr
                      ☑   LocalDskMgr
Other   ☑   LocalDskMgr
File System
Save...   Ctl+S
Quit   Ctl+Q

Allow a storage object to be subdivided by an assigned segment manager

ibm.com/redbooks

Select DOS Segment Manager
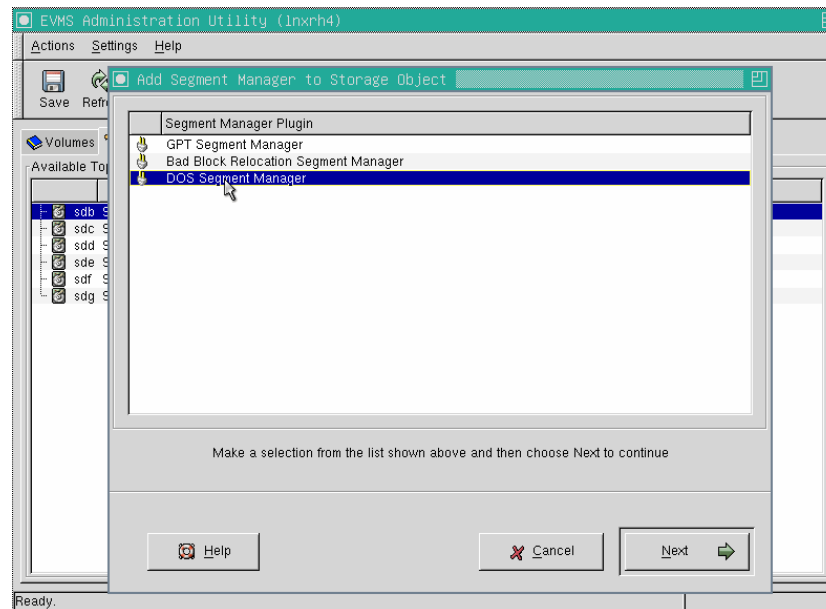
Select the sdb Disk to Manage
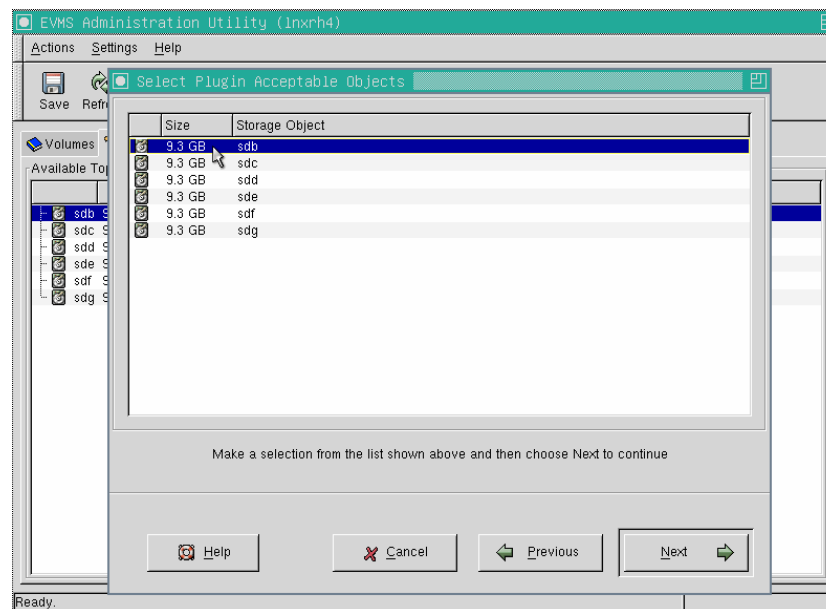
Select Linux Disk Type
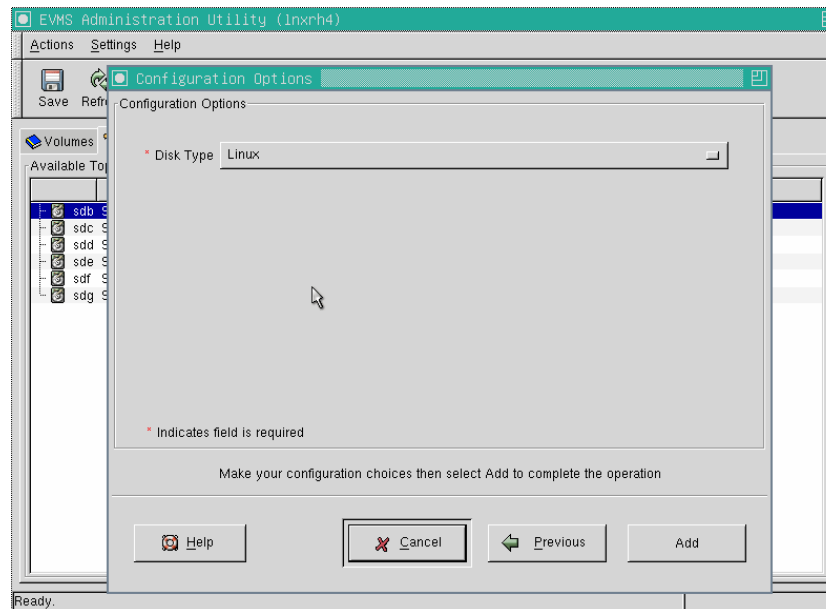
ibm.com/redbooks

Segment Manager Added to sdb
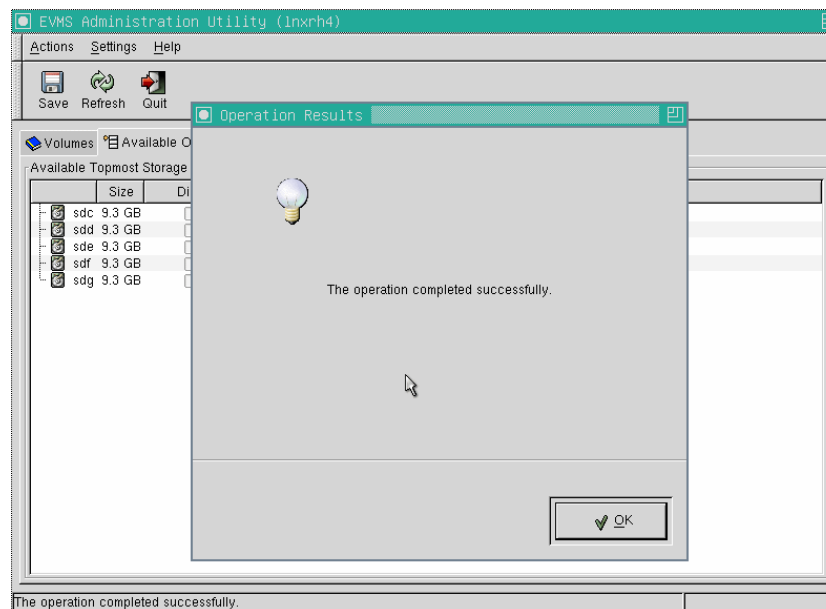
ibm.com/redbooks

Device sdb No Longer Available

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Available Topmost Storage Objects

	Size	Dirty	Active	Read Only	Plugin
sdc	9.3 GB	☐	✔	☐	LocalDskMgr
sdd	9.3 GB	☐	✔	☐	LocalDskMgr
sde	9.3 GB	☐	✔	☐	LocalDskMgr
sdf	9.3 GB	☐	✔	☐	LocalDskMgr
sdg	9.3 GB	☐	✔	☐	LocalDskMgr

The operation completed successfully.
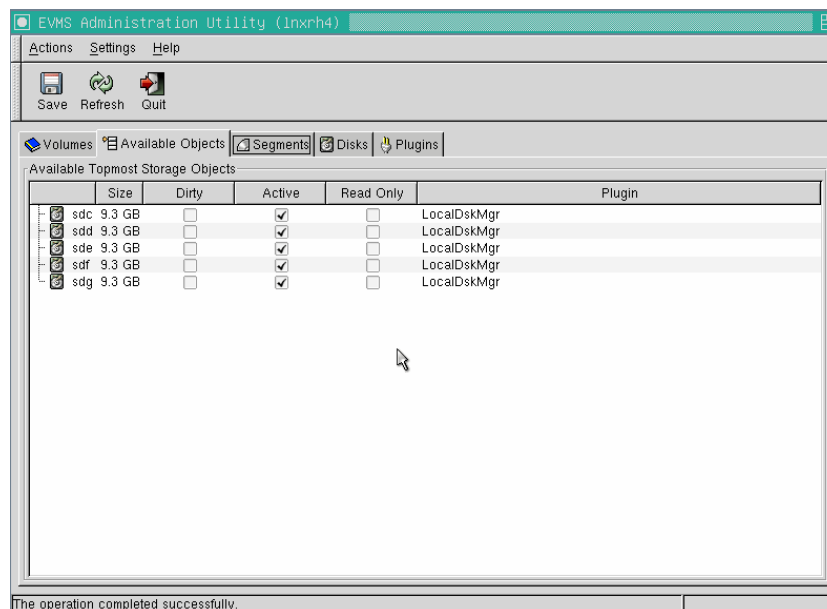
Redbooks

Device sdb Has No Partitions

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Disk Segments

	Size	Dirty	Active	Read Only	Type	Offset
DosSegMgr						
sda_mbr	16 KB	☐	☐	☐	Meta Data	0
sda1	16.0 GB	☐	☐	☐	Data	32
sdb_freespace1	9.3 GB	☐	☐	☐	Free Space	32
sdb_mbr	16 KB	✔	☐	☐	Meta Data	0

The operation completed successfully.
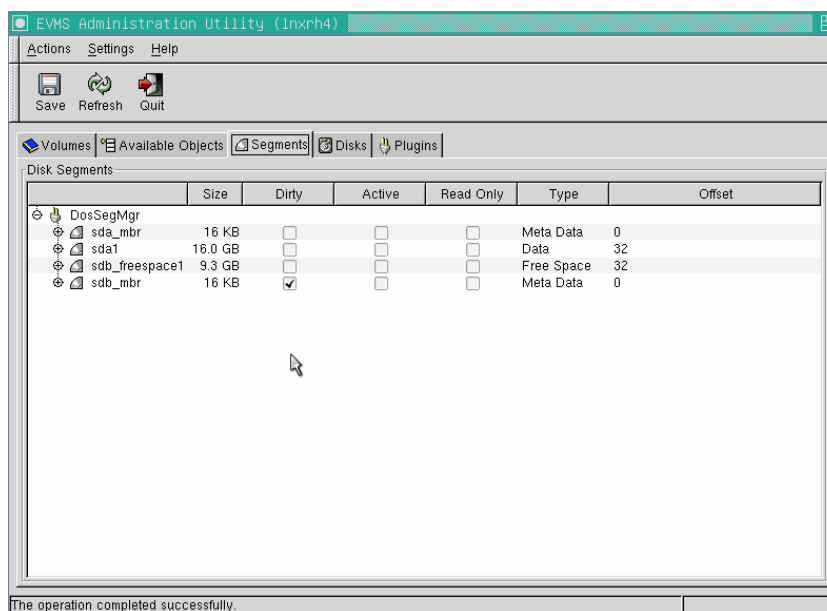
Redbooks

Add Segment Manager to All Disks
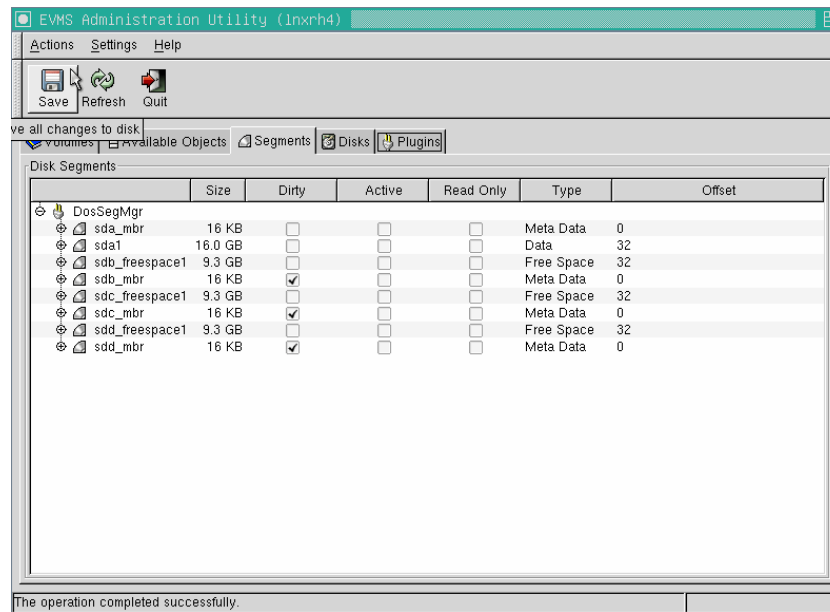
Commit EVMS Changes to Disk
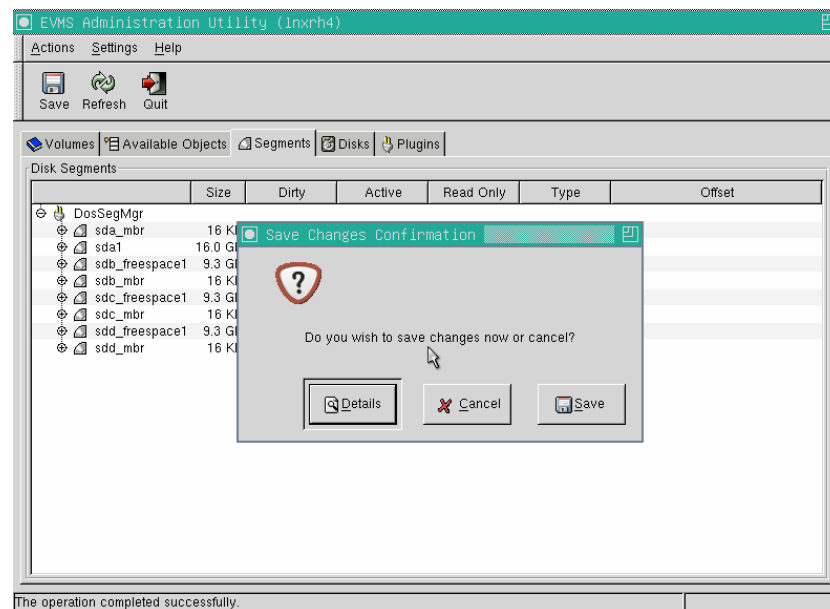
Create EVMS Segment on Disk

EVMS Administration Utility (lnxrh4)

Actions　Settings　Help

Save　Refresh　Quit

Volumes　Available Objects　Segments　Disks　Plugins

Disk Segments

	Size	Dirty	Active	Read Only	Type	Offset
DosSegMgr						
sda_mbr	16 KB	☐	☐	☐	Meta Data	0
sda1	16.0 GB	☐	☐	☐	Data	32
sdb_freespace1	9.3 GB	☐	☐	☐	Free Space	32
sdb_mbr	16	☐	☐	☐	Meta Data	0
sdc_freespace1	9.3	☐	☐	☐	Free Space	32
sdc_mbr	16 KB	☐	☐	☐	Meta Data	0
sdd_freespace1	9.3 GB	☐	☐	☐	Free Space	32
sdd_mbr	16 KB	☐	☐	☐	Meta Data	0

Display details...
Create Segment...

All changes were saved successfully.
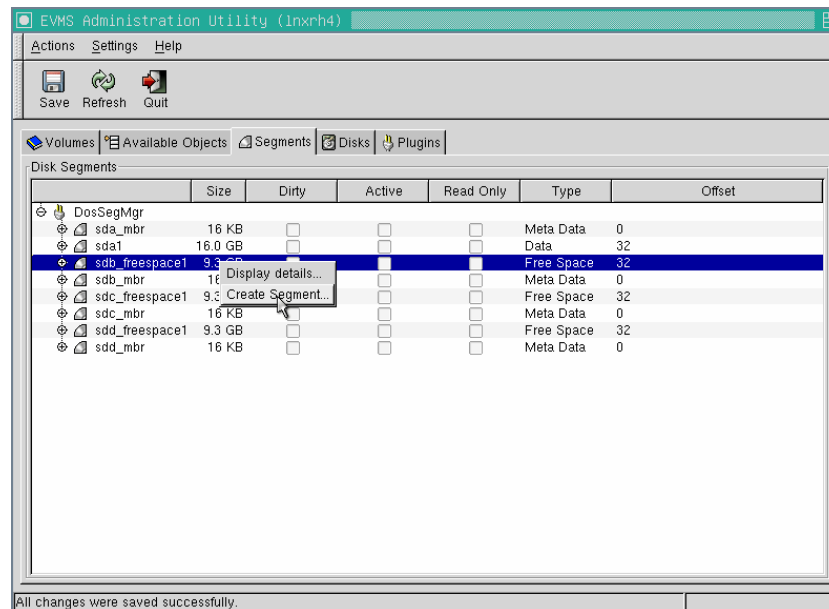
Redbooks

Provide Segment Parameters

EVMS Administration Utility (lnxrh4)

Actions　Settings　Help

Save　Refr

Volumes

Disk Segmen

DosS
sd
sd
sd
sd
sd
sd
sd
sd

Create Segment Configuration Options

Configuration Options

* Size (MB)　9535.98

* Offset (sectors)　0

Partition Type　Linux

* Partition Type Id　0x83

* Bootable　◇ Yes　◆ No

* Primary Partition　◆ Yes　◇ No

* Indicates field is required

Make your configuration choices then select Create to complete the operation

Help　　　Cancel　　　Create

All changes were saved successfully.

Redbooks

Segment Created

EVMS Administration Utility (lnxrh4)

Actions Settings Help

Save Refresh Quit

Volumes Available

Disk Segments

Operation Completed Successfully

The operation completed successfully producing the objects listed below:

	Size	Type	Name	Offset
	9.3 GB	Data	sdb1	

DosSegMgr
- sda_mbr
- sda1
- sdb_mbr
- sdb1
- sdb_freespac
- sdc_mbr
- sdc_freespac
- sdd_mbr
- sdd_freespac

OK

The operation completed successfully.

Create All Segments and Save
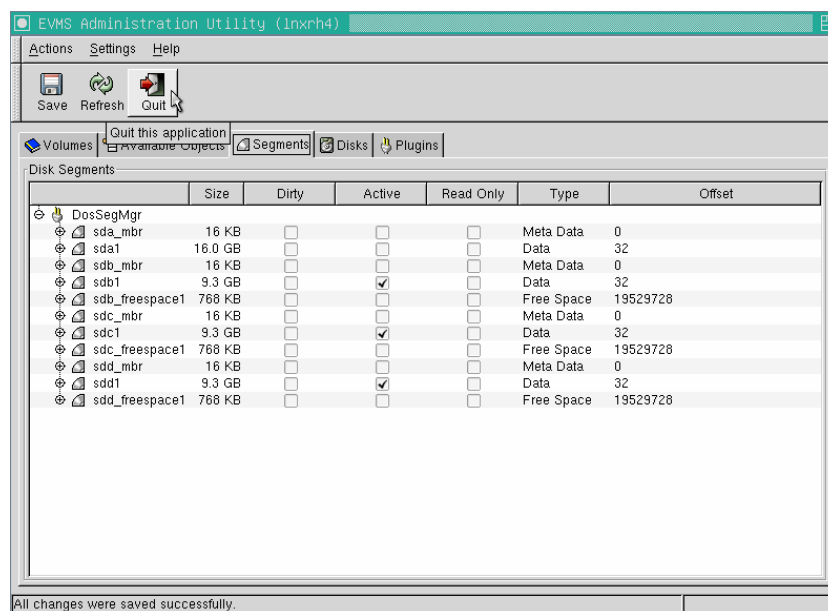
EVMS Administration Utility (lnxrh4)

Actions Settings Help

Save Refresh Quit

Quit this application

Volumes Available Objects Segments Disks Plugins

Disk Segments

	Size	Dirty	Active	Read Only	Type	Offset
DosSegMgr						
sda_mbr	16 KB	☐	☐	☐	Meta Data	0
sda1	16.0 GB	☐	☐	☐	Data	32
sdb_mbr	16 KB	☐	☐	☐	Meta Data	0
sdb1	9.3 GB	☐	☑	☐	Data	32
sdb_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdc_mbr	16 KB	☐	☐	☐	Meta Data	0
sdc1	9.3 GB	☐	☑	☐	Data	32
sdc_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdd_mbr	16 KB	☐	☐	☐	Meta Data	0
sdd1	9.3 GB	☐	☑	☐	Data	32
sdd_freespace1	768 KB	☐	☐	☐	Free Space	19529728

All changes were saved successfully.

Exit and Restart EVMS
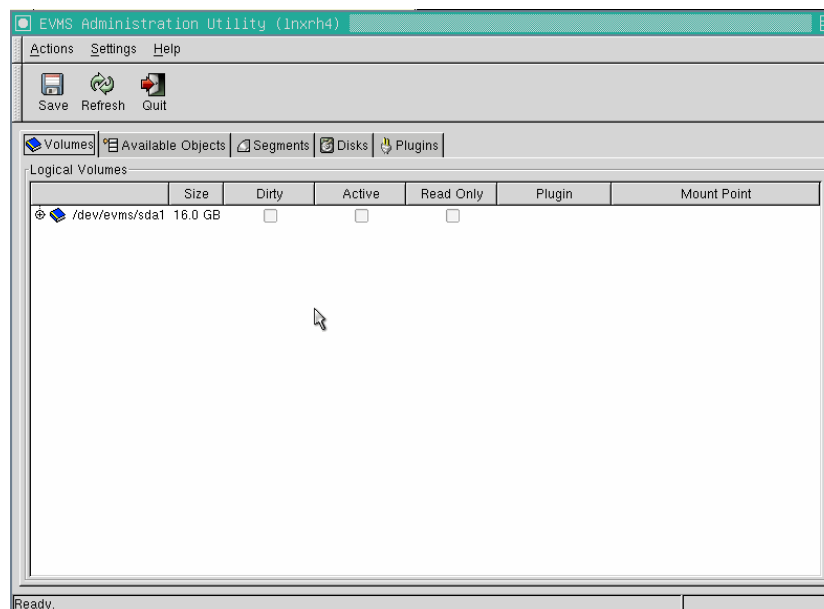
New Available Objects
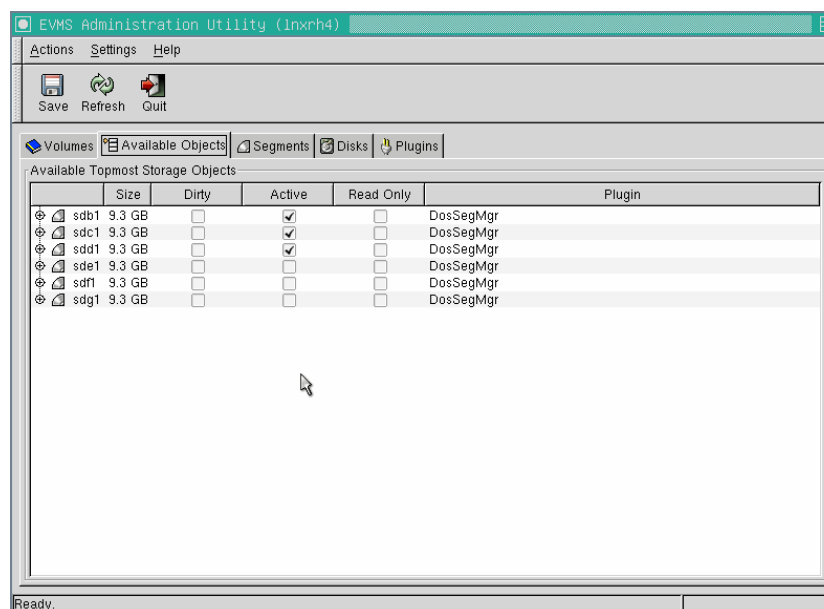
Available EVMS Segments

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save   Refresh   Quit

Volumes | Available Objects | Segments | Disks | Plugins

Disk Segments

	Size	Dirty	Active	Read Only	Type	Offset
DosSegMgr						
sda_mbr	16 KB	☐	☐	☐	Meta Data	0
sda1	16.0 GB	☐	☐	☐	Data	32
sdb_mbr	16 KB	☐	☐	☐	Meta Data	0
sdb1	9.3 GB	☐	☑	☐	Data	32
sdb_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdc_mbr	16 KB	☐	☐	☐	Meta Data	0
sdc1	9.3 GB	☐	☑	☐	Data	32
sdc_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdd_mbr	16 KB	☐	☐	☐	Meta Data	0
sdd1	9.3 GB	☐	☑	☐	Data	32
sdd_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sde_mbr	16 KB	☐	☐	☐	Meta Data	0
sde1	9.3 GB	☐	☐	☐	Data	32
sde_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdf_mbr	16 KB	☐	☐	☐	Meta Data	0
sdf1	9.3 GB	☐	☐	☐	Data	32
sdf_freespace1	768 KB	☐	☐	☐	Free Space	19529728
sdg_mbr	16 KB	☐	☐	☐	Meta Data	0
sdg1	9.3 GB	☐	☐	☐	Data	32
sdg_freespace1	768 KB	☐	☐	☐	Free Space	19529728

Ready.

Redbooks

Create EVMS Region

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Create	▶	EVMS Volume...
Delete		Compatibility Volume...
Expand		Feature Object...
Shrink		Region...
Add		Container...
Remove		Segment...
Modify		
View		
Convert		
Replace...		
Other		
File System		
Save...	Ctl+S	
Quit	Ctl+Q	

| Disks | Plugins |

	Read Only	Plugin
☑	☐	DosSegMgr
☑	☐	DosSegMgr
☑	☐	DosSegMgr
☐	☐	DosSegMgr
☐	☐	DosSegMgr
☐	☐	DosSegMgr

Create a storage region

Redbooks

Select Region Manager

ibm.com/redbooks

Assign Region Manager to Paths
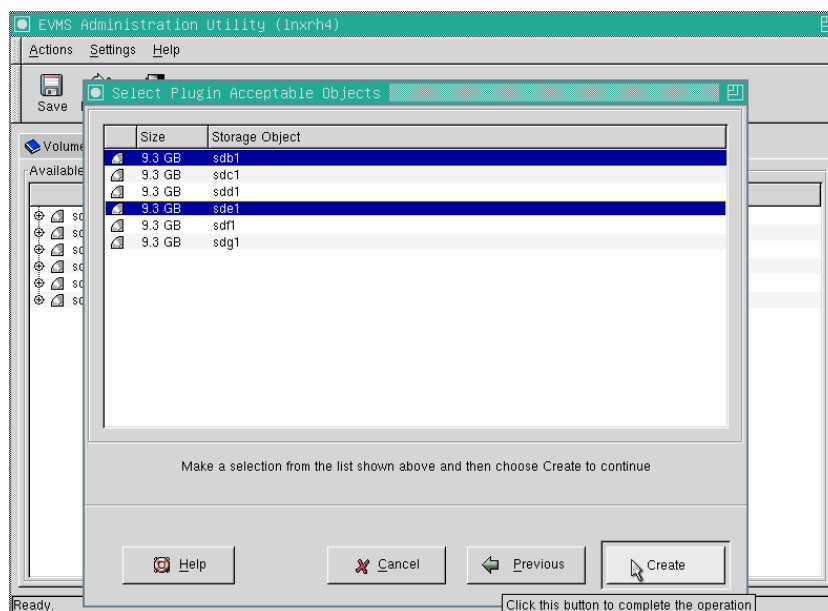
ibm.com/redbooks

Create Multipath Device to LUN
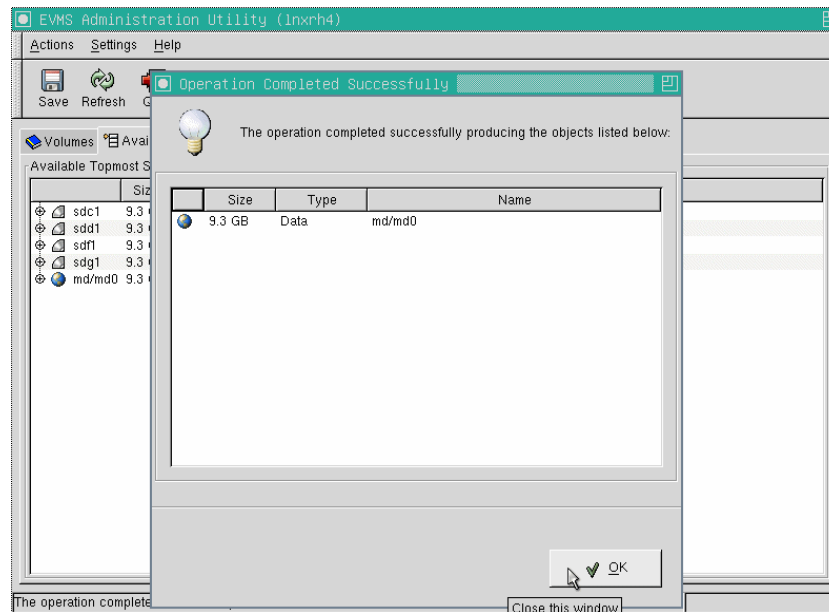
ibm.com/redbooks

Complete for Each Multipath LUN
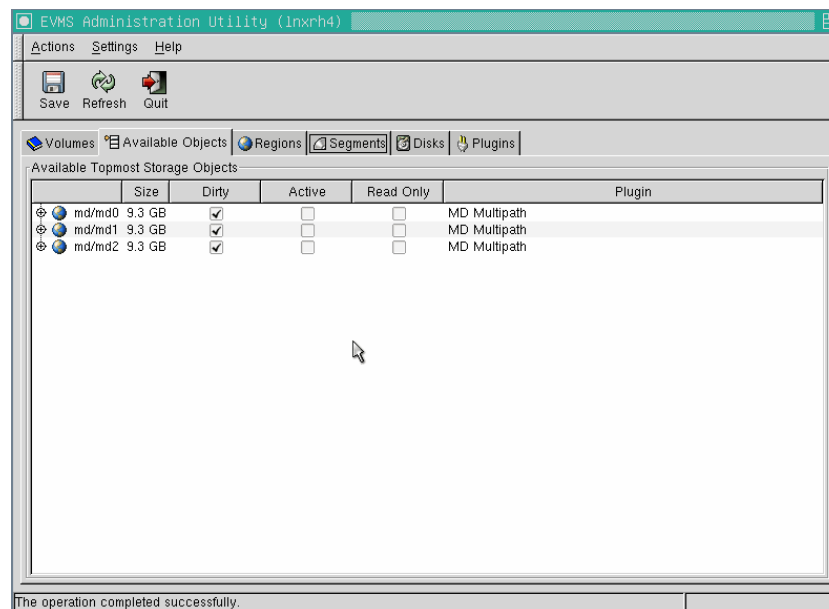
ibm.com/redbooks

Define RAID0 Array of Multipath Regions
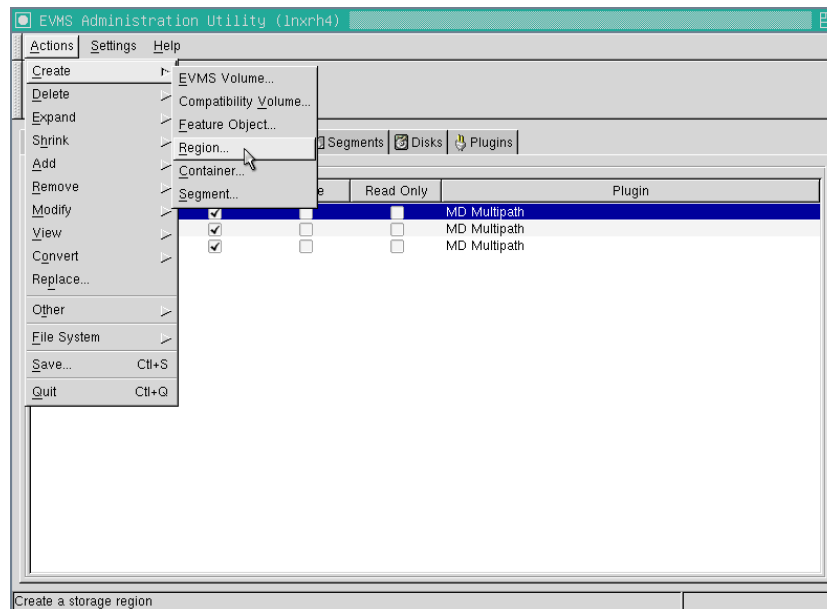
Create MD RAID0 Region Manager
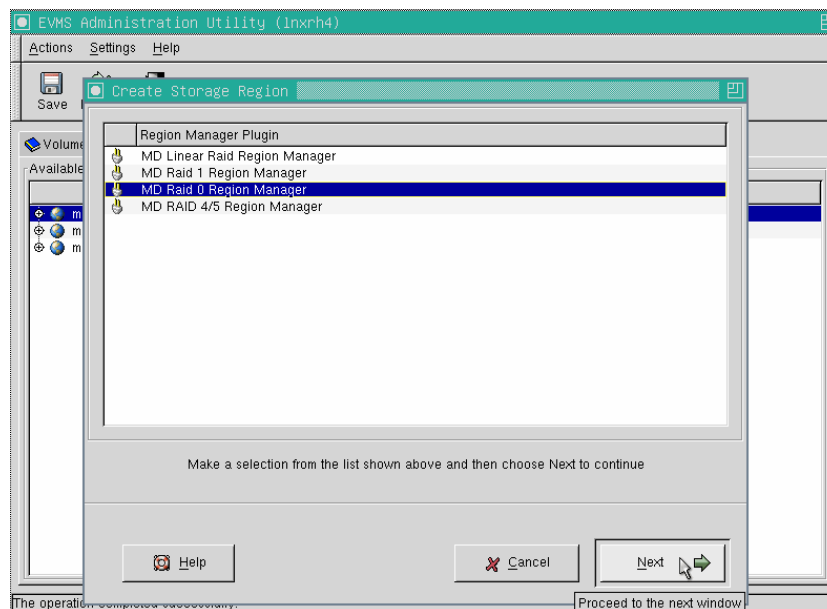
Select MD Multipath Devices
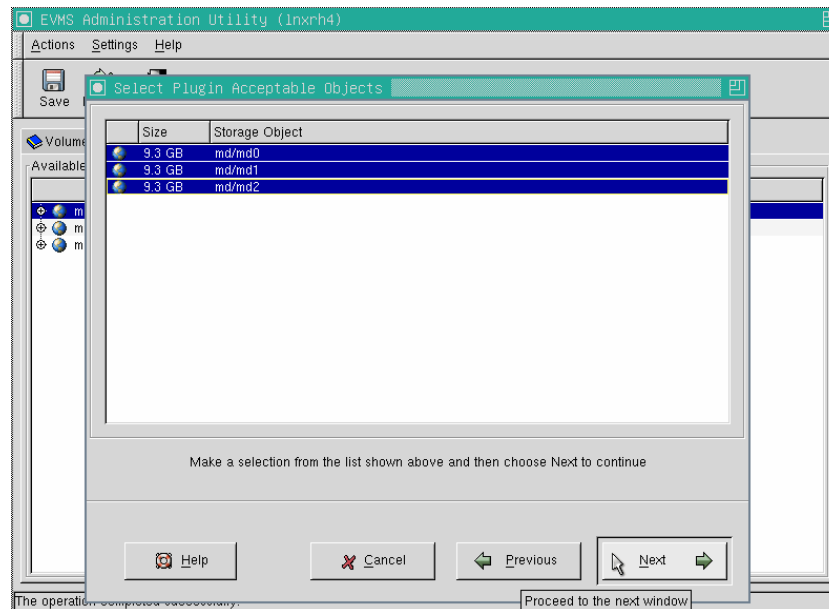
EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save

Volume

Available

Select Plugin Acceptable Objects

Size	Storage Object
9.3 GB	md/md0
9.3 GB	md/md1
9.3 GB	md/md2

Make a selection from the list shown above and then choose Next to continue

Help   Cancel   Previous   Next

Proceed to the next window

The operation completed successfully.

Select Chunk Size

EVMS Administration Utility (lnxrh4)

Actions   Settings   Help

Save

Volume

Available

Configuration Options

Configuration Options

* Chunk Size: (KB)   32

* Indicates field is required

Make your configuration choices then select Create to complete the operation

Help   Cancel   Previous   Create

The operation completed successfully.
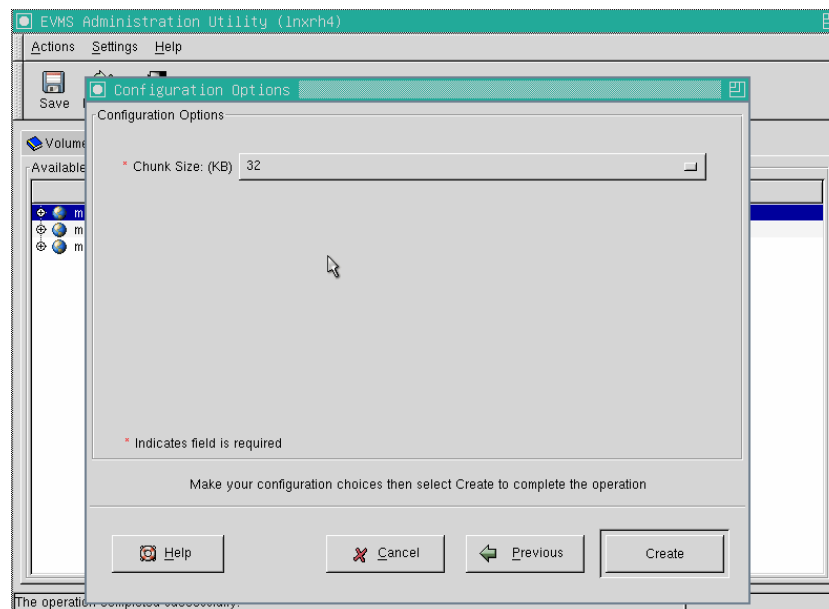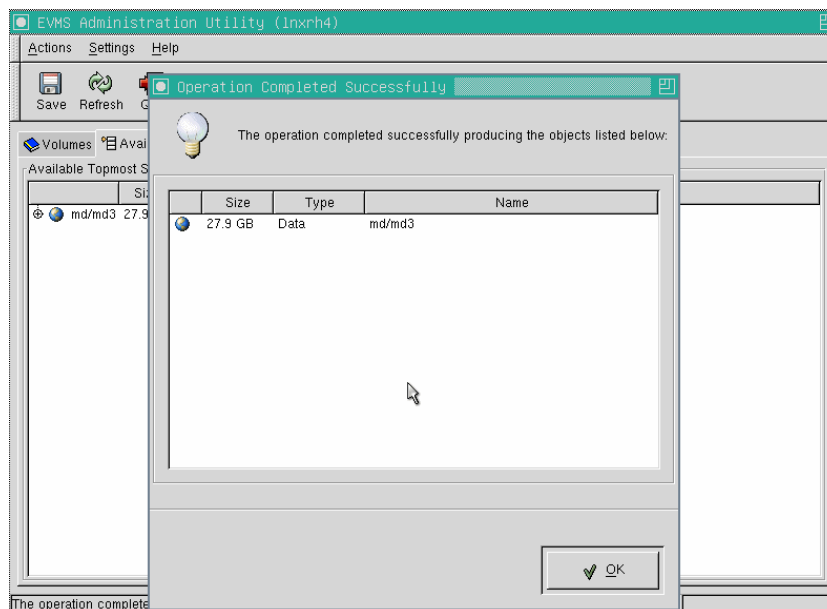
RAID0 Array Created

Commit Changes to Disk
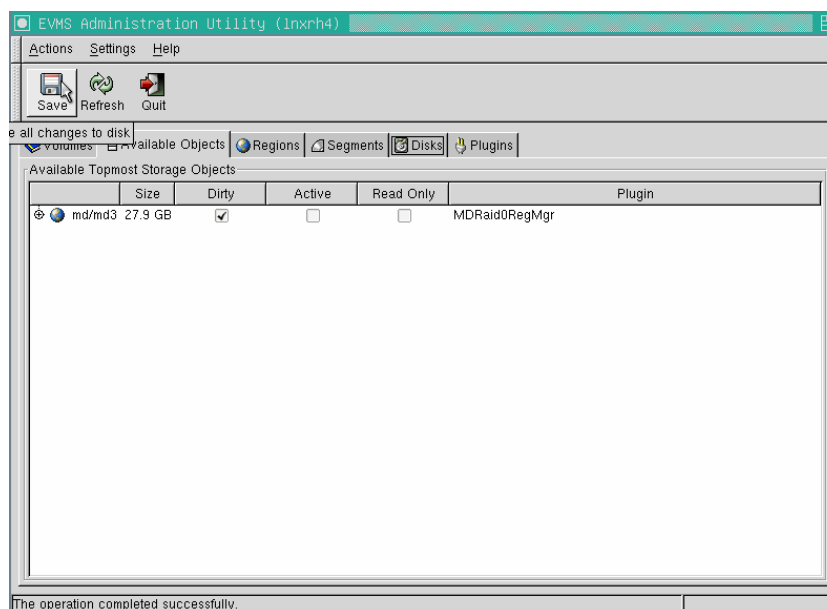
Create Volume Within RAID0 Region
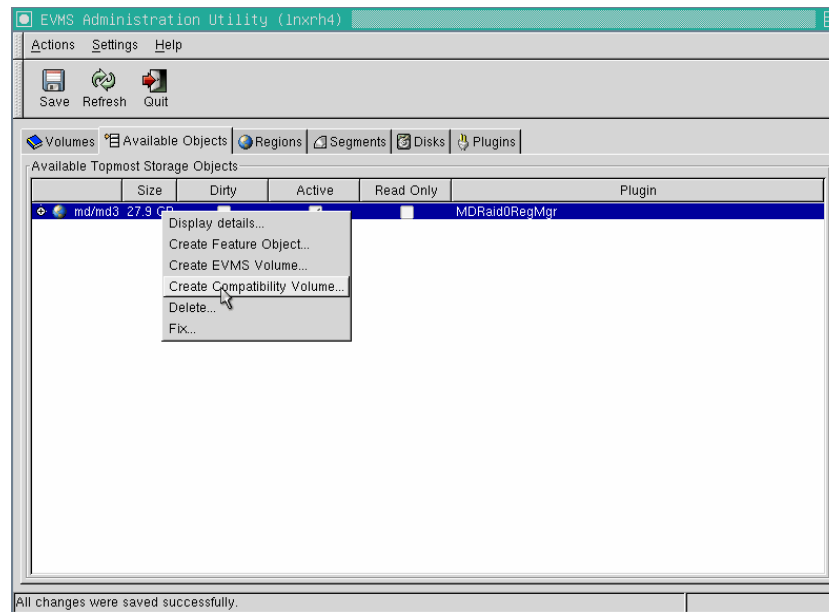
ibm.com/redbooks

Choose Storage Region
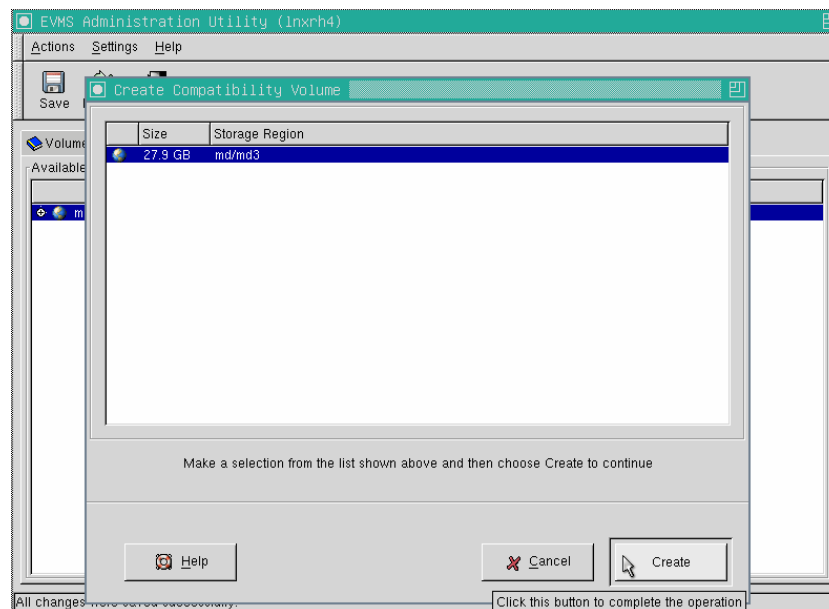
ibm.com/redbooks

Save Changes
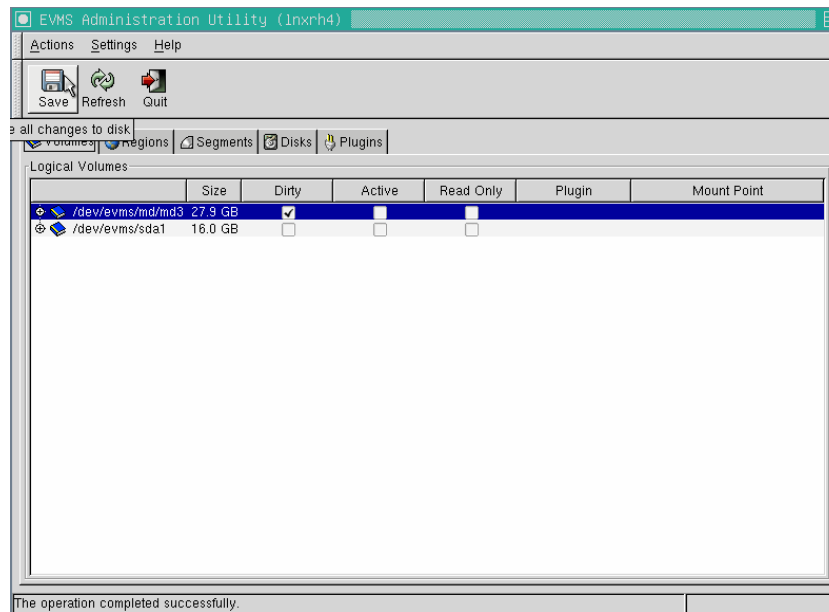
ibm.com/redbooks

Create Filesystem on Volume
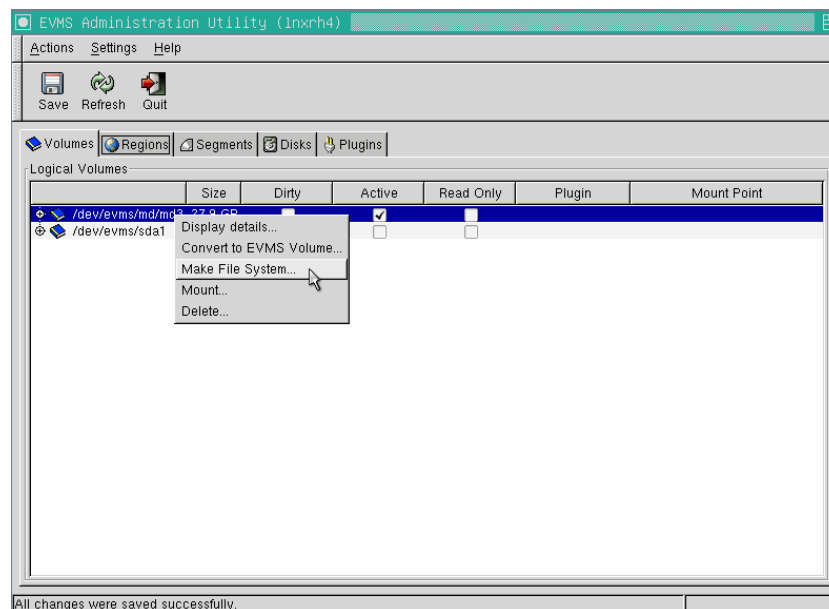
ibm.com/redbooks

Select Filesystem Type
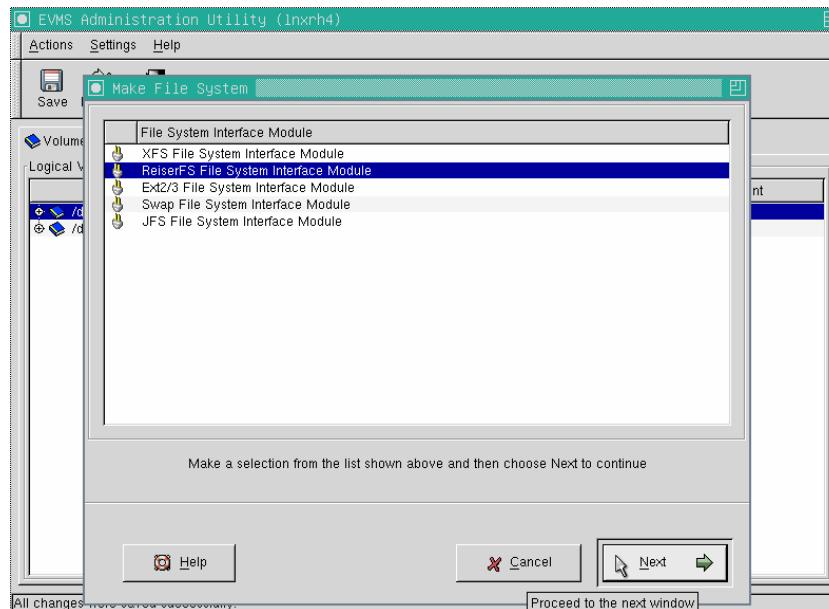
Add Volume Label if Desired
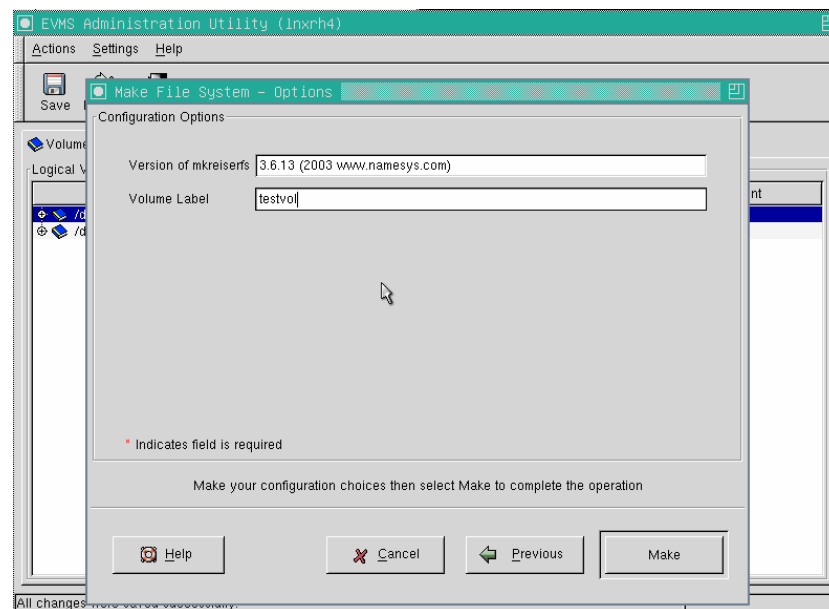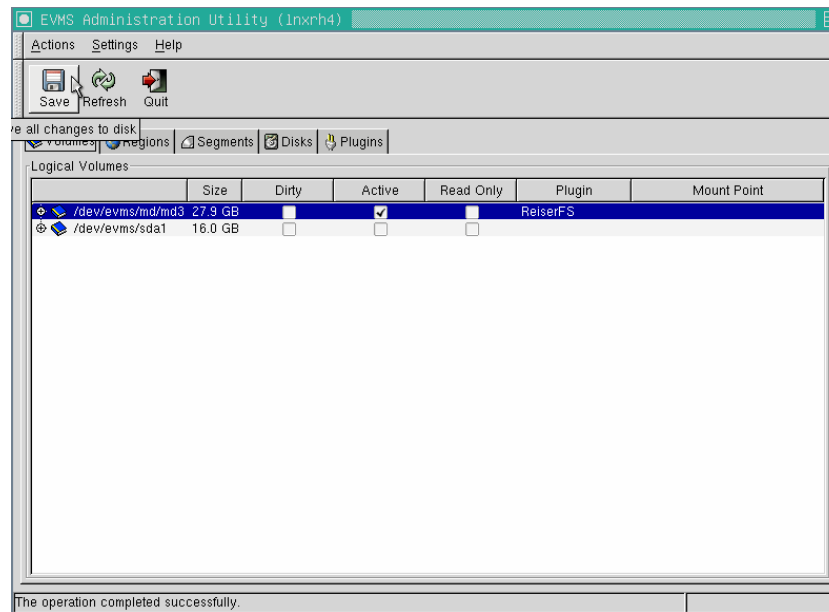
Filesystem Created

Add to /etc/fstab and Start EVMS at Boot

```
# chkconfig boot.evms on


# cat /etc/fstab
/dev/sda1                /                    reiserfs   acl,user_xattr     1 1
devpts                   /dev/pts             devpts     mode=0620,gid=5    0 0
proc                     /proc                proc       defaults           0 0
sysfs                    /sys                 sysfs      noauto             0 0
/dev/evms/md/md3         /mnt                 reiserfs   acl,user_xattr     1 2
```

To Summarize:

2.6 kernel device mapper simplifies multipathing
- EVMS offers easy to use graphic interface

With EVMS:
- Add Segment Manager to partition
- Create:
  - Segment
  - Region
  - Multipath device
  - RAID array
  - Volume

Enable EVMS at system initialization
- Add voume to /etc/fstab

FCP Security Topics

LUN masking
- Restricts LUN access to specific groups of WWPNs
- Defined in the storage controller

Zoning
- Enable or disable communications bewteen nodes in fabric
- Hard zoning
  - Uses switch identifier and port numbers
- Soft zoning
  - Uses WWPNs
- Defined at the switch
  - Applies to entire fabric

LUN masking

ibm.com

Disk Storage

Server1
WWPN1
WWPN2

LUN 0
LUN 1
LUN 2

LUN 0 = WWPN1 + WWPN6
LUN 1 = WWPN2
LUN 2 = WWPN3 + WWPN4
.
.
LUN n = WWPN3 + WWPN5

LUN n

Fabric

Server2
WWPN3
WWPN4

Server3
WWPN5
WWPN6

LUN masking policy resides
in the storage controller
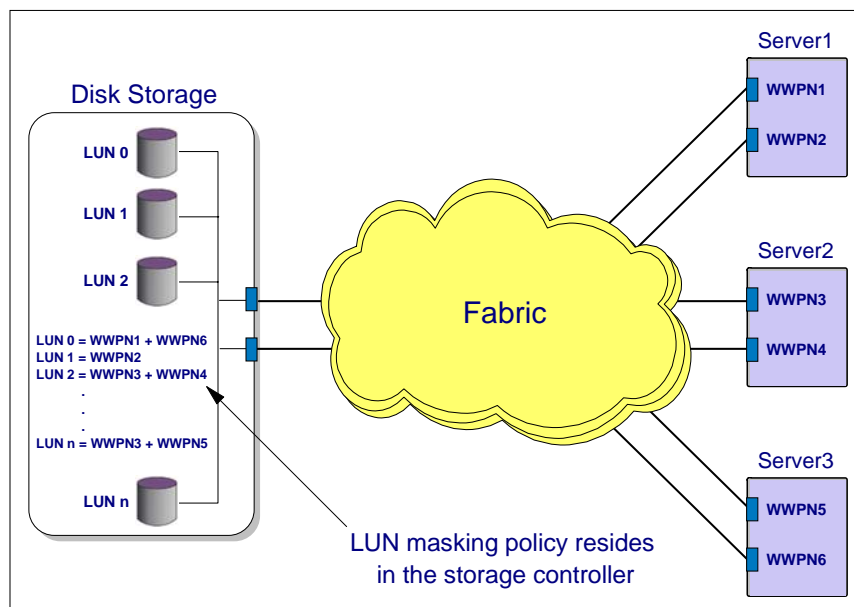
Redbooks

ibm.com/redbooks

© Copyright IBM Corp. 2004. All rights reserved.

Zoning
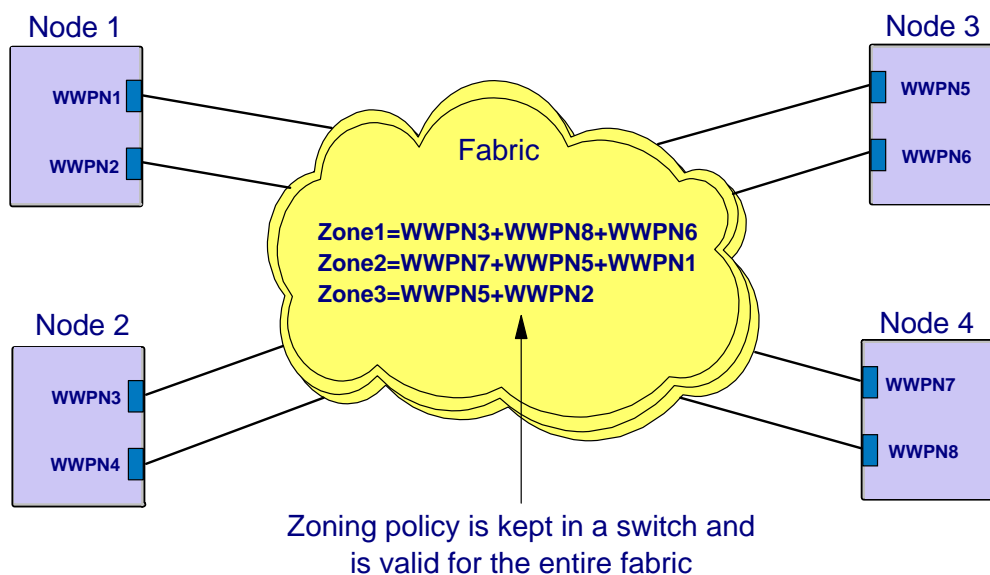
ibm.com

Node 1
WWPN1
WWPN2

Node 3
WWPN5
WWPN6

Fabric

Zone1=WWPN3+WWPN8+WWPN6
Zone2=WWPN7+WWPN5+WWPN1
Zone3=WWPN5+WWPN2

Node 2
WWPN3
WWPN4

Node 4
WWPN7
WWPN8

Zoning policy is kept in a switch and
is valid for the entire fabric

Redbooks

ibm.com/redbooks

© Copyright IBM Corp. 2004. All rights reserved.

325-326

What's New for Linux in z/VM 5.1

Redbooks

International Technical Support Organization

---

ibm.com

Topics for z/VM 5.1

Installing z/VM 5.1 from DVD

Native SCSI support

VSWITCH enhanced authorization

HyperSwap function

Redbooks

ibm.com/redbooks

Related Publications

Other Publications

- *These publications are relevant as information on Performance Toolkit for VM*

Title	Publication Number
Linux on IBM eServer and S/390: Performance Toolkit for VM	SG24-6059
Linux on IBM eServer and S/390: Performance Measurement and Tuning	SG24-6926
Accounting and Monitoring for z/VM Linux Guest Machines	REDP-3818
z/VM: Performance Toolkit	SC24-6062
z/VM: Performance	SC24-5999

Redbooks

ibm.com/redbooks

Installing VM 5.1 From DVD

z/VM is now available on DVD!

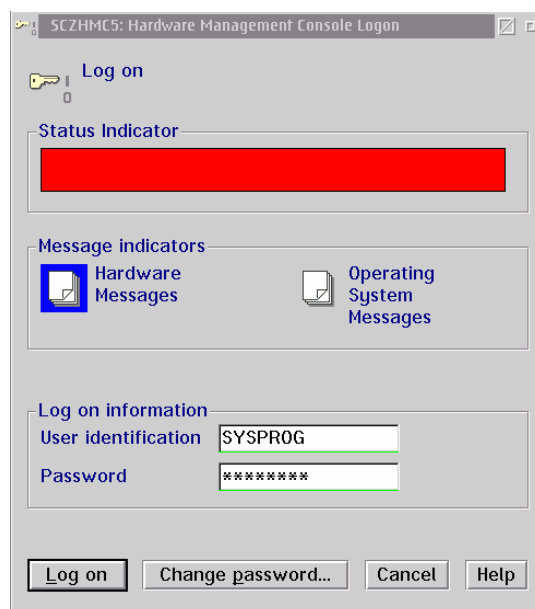- Simplifies procedure
  - No tapes to mount
- Uses INSTDVD dialog

Requires:

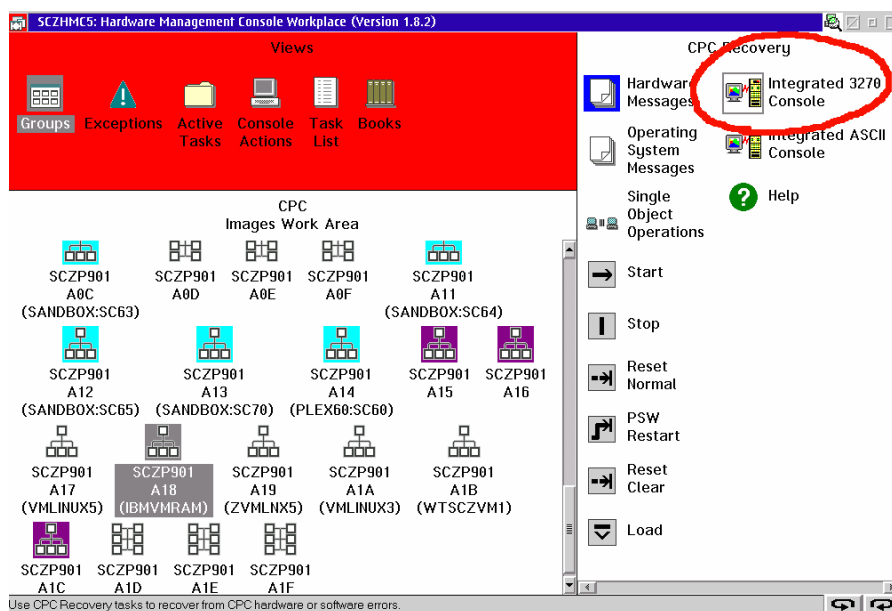- Hardware Management Console V1.8 or higher

Redbooks

ibm.com/redbooks

Logon to HMC

SCZHMC5: Hardware Management Console Logon

Log on

Status Indicator

Message indicators
- Hardware Messages
- Operating System Messages

Log on information

User identification SYSPROG

Password ********

[Log on] [Change password...] [Cancel] [Help]

Redbooks

Select Integrated 3270 Console

SCZHMC5: Hardware Management Console Workplace (Version 1.8.2)

Views

Groups | Exceptions | Active Tasks | Console Actions | Task List | Books

CPC Recovery

- Hardware Messages
- Integrated 3270 Console
- Operating System Messages
- Integrated ASCII Console
- Single Object Operations
- Help
- Start
- Stop
- Reset Normal
- PSW Restart
- Reset Clear
- Load

CPC Images Work Area

SCZP901 A0C (SANDBOX:SC63) | SCZP901 A0D | SCZP901 A0E | SCZP901 A0F | SCZP901 A11 (SANDBOX:SC64)

SCZP901 A12 (SANDBOX:SC65) | SCZP901 A13 (SANDBOX:SC70) | SCZP901 A14 (PLEX60:SC60) | SCZP901 A15 | SCZP901 A16

SCZP901 A17 (VMLINUX5) | SCZP901 A18 (IBMVMRAM) | SCZP901 A19 (ZVMLNX5) | SCZP901 A1A (VMLINUX3) | SCZP901 A1B (WTSCZVM1)

SCZP901 A1C | SCZP901 A1D | SCZP901 A1E | SCZP901 A1F

Use CPC Recovery tasks to recover from CPC hardware or software errors.

Redbooks

Select Single Object Operations
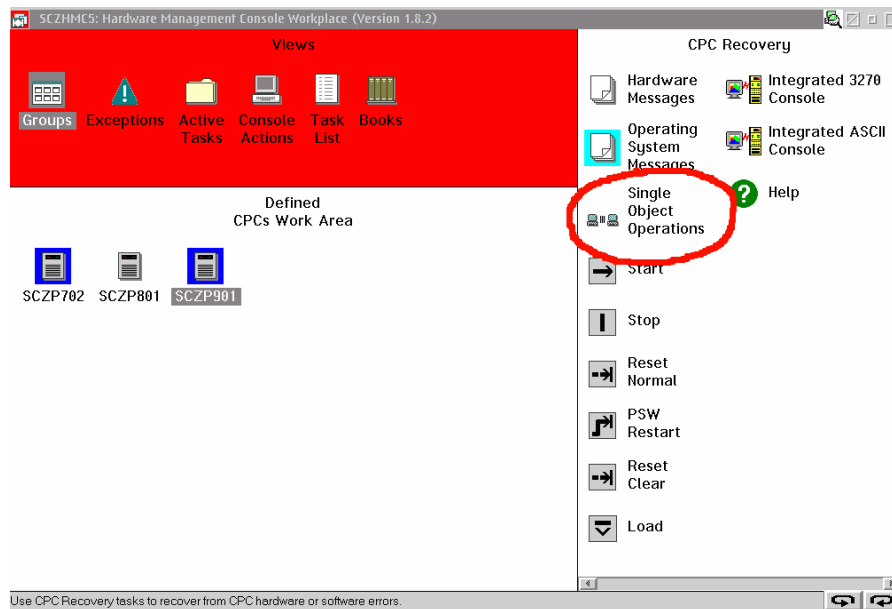
ibm.com/redbooks

Confirm Selection

Single Object Operations Task Confirmation

About to establish a session with a single CPC console.

Do you want to continue with this task?

Object names

SCZP901

[Yes] [No] [Help]
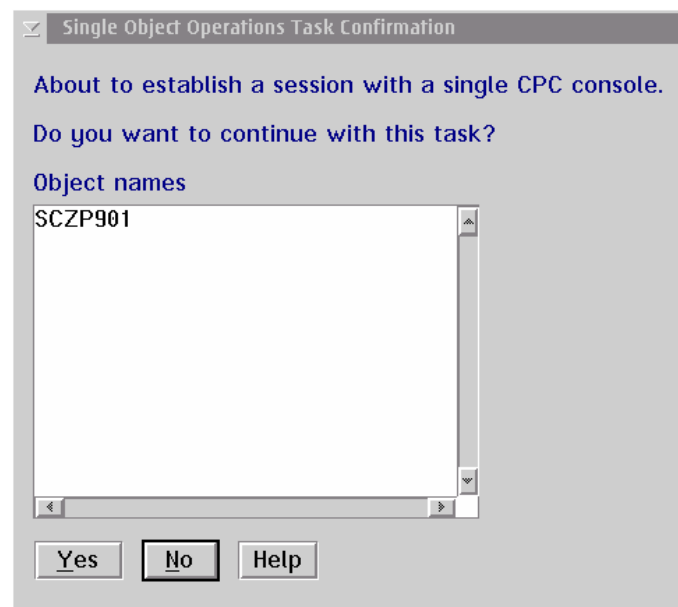
ibm.com/redbooks

Select DVD Drive For Installation
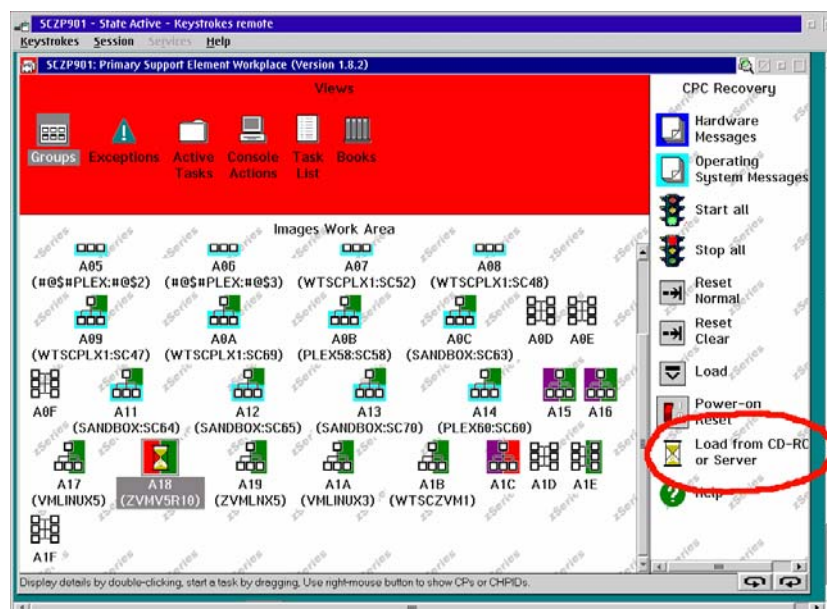
ibm.com/redbooks

Select the Installation Media

ibm.com/redbooks

Confirm Installation Parameters
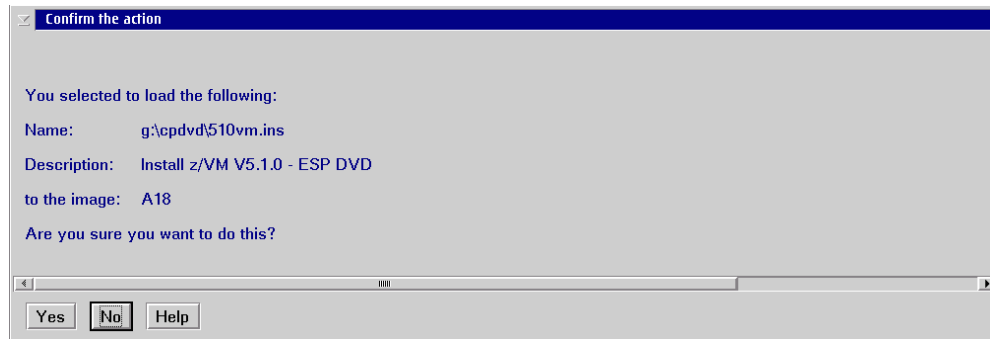
Confirm the action

You selected to load the following:

Name: g:\cpdvd\510vm.ins

Description: Install z/VM V5.1.0 - ESP DVD

to the image: A18

Are you sure you want to do this?

| Yes | No | Help |

Redbooks

ibm.com/redbooks

Image Load Complete

Load from CD-ROM or Server Progress

Function duration time: 00:05:00

Elapsed time: 00:03:38

Load from CD-ROM or server for image A18   completed successfully

| OK | Help |
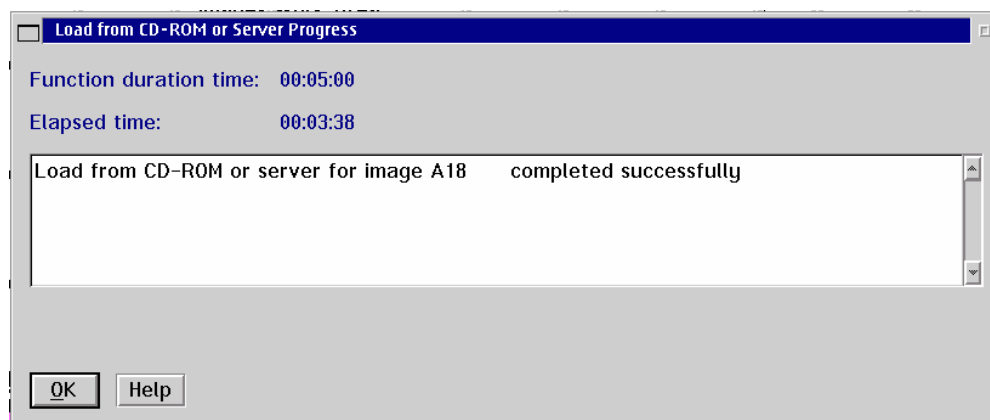
Redbooks

ibm.com/redbooks

IPL From Ramdisk

```
Integrated 3270 Console for SCZP901:A18
File   Keys   Help

13:41:53 z/VM  V5 R1.0  SERVICE LEVEL 0000 (64-BIT)
13:41:53 SYSTEM NUCLEUS CREATED ON 2004-05-14 AT 14:43:33, LOADED FROM $RAMD$
13:41:53
13:41:53 ****************************************************************
13:41:53 * LICENSED MATERIALS - PROPERTY OF IBM*                        *
13:41:53 *                                                              *
13:41:53 * 5741-A05 (C) COPYRIGHT IBM CORP. 1983, 2004. ALL RIGHTS      *
13:41:53 * RESERVED. US GOVERNMENT USERS RESTRICTED RIGHTS - USE,       *
13:41:53 * DUPLICATION OR DISCLOSURE RESTRICTED BY GSA ADP SCHEDULE     *
13:41:53 * CONTRACT WITH IBM CORP.                                      *
13:41:53 *                                                              *
13:41:53 * * TRADEMARK OF INTERNATIONAL BUSINESS MACHINES.              *
13:41:53 ****************************************************************
13:41:53
13:41:53 HCPZCO6718I Using parm disk 1 on volume $RAMD$ (device FFFF).
13:41:53 HCPZCO6718I Parm disk resides on blocks 18000 through 52992.
13:41:53 The directory on volume $RAMD$ at address FFFF has been brought online.
13:41:53 HCPWRS2512I Spooling initialization is complete.
13:41:53 No dump unit - Dump function is SET OFF
13:41:53 HCPAAU2700I System gateway IBMVMRAM identified.
13:41:55 z/VM Version 5 Release 1.0, Service Level 0000 (64-bit),
13:41:55 built on IBM Virtualization Technology
13:41:55 There is no logmsg data
13:41:55 FILES:   NO RDR,   NO PRT,   NO PUN
13:41:55 LOGON AT 13:41:55 EDT TUESDAY 08/10/04
13:41:55 SYSG       LOGON  AS  MAINT   USERS = 1
13:41:55 HCPIOP952I 3G system storage
13:41:55 FILES: 0000001 RDR, 0000001 PRT,      NO PUN
13:41:55 HCPCRC8082I Accounting records are accumulating for userid OPERACCT.
13:41:55 HCPCRC8082I EREP records are accumulating for userid OPEREREP.
DMSIND2015W Unable to access the Y-disk. Filemode Y (19E) not accessed
DMSWSP327I The installation saved segment could not be loaded
z/VM V5.1.0    2004-05-28 13:32
DMSDCS1083E Saved segment CMSPIPES does not exist
DMSDCS1083E Saved segment CMSPIPES does not exist
DMSDCS1083E Saved segment CMSVMLIB does not exist
Ready; T=0.01/0.02 13:41:55

                                                  RUNNING   IBMVMRAM
                                                       42/001
```

ibm.com/redbooks

The INSTPLAN Dialog

```
Integrated 3270 Console for SCZP901:A18
File   Keys   Help
                    *** z/VM INSTALLATION PLANNING ***

Mark the product(s) selected to be installed into the VMSYS filepool with an
"F" and those selected to be installed to minidisks with an "M"

Install To   Product    Install To   Product    Install To   Product
----------   --------    ----------   --------    ----------   --------
     M       VM              M        RSCS            M        TCPIP
     M       OSA             M        ICKDSF          M        DIRM
     M       RACF            M        PERFTK          M        VMHCD


Place a nonblank character in front of the System Default Language you would
like for your system.

   x AMENG        _ UCENG        _ KANJI       _ GERMAN

Place a nonblank character in front of the DASD model onto which your
z/VM system will be loaded.  Only one model may be selected.

   x 3390 Mod 3         _ 3390 Mod 9




    PF1 = HELP   PF3/PF12 = QUIT   PF5 = Process   ENTER = Refresh
                                                         22/029
```
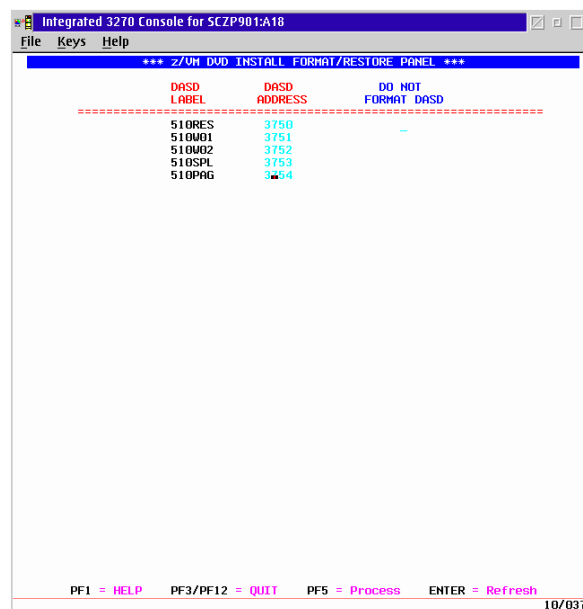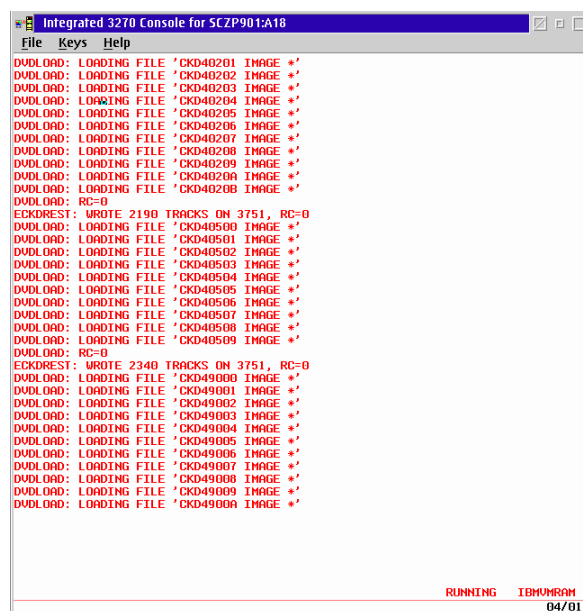
ibm.com/redbooks

The INSTDVD Dialog

```
Integrated 3270 Console for SCZP901:A18
File   Keys   Help

              *** z/VM DVD INSTALL FORMAT/RESTORE PANEL ***

                   DASD          DASD           DO NOT
                   LABEL         ADDRESS        FORMAT DASD
              =================================================
                   510RES        3750              _
                   510W01        3751
                   510W02        3752
                   510SPL        3753
                   510PAG        3754



        PF1 = HELP     PF3/PF12 = QUIT     PF5 = Process     ENTER = Refresh
                                                                    10/037
```

ibm.com/redbooks

Installation Messages

```
Integrated 3270 Console for SCZP901:A18
File   Keys   Help
DVDLOAD: LOADING FILE 'CKD40201 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40202 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40203 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40204 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40205 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40206 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40207 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40208 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40209 IMAGE *'
DVDLOAD: LOADING FILE 'CKD4020A IMAGE *'
DVDLOAD: LOADING FILE 'CKD4020B IMAGE *'
DVDLOAD: RC=0
ECKDREST: WROTE 2190 TRACKS ON 3751, RC=0
DVDLOAD: LOADING FILE 'CKD40500 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40501 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40502 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40503 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40504 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40505 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40506 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40507 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40508 IMAGE *'
DVDLOAD: LOADING FILE 'CKD40509 IMAGE *'
DVDLOAD: RC=0
ECKDREST: WROTE 2340 TRACKS ON 3751, RC=0
DVDLOAD: LOADING FILE 'CKD49000 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49001 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49002 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49003 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49004 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49005 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49006 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49007 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49008 IMAGE *'
DVDLOAD: LOADING FILE 'CKD49009 IMAGE *'
DVDLOAD: LOADING FILE 'CKD4900A IMAGE *'



                                              RUNNING   IBMVMRAM
                                                          04/013
```

ibm.com/redbooks

Load From the 510RES Volume

Stand-alone Program Loader

z/VM IPL Messages

```
Integrated 3270 Console for SCZP901:A18
File  Keys  Help
10:30:15 z/VM  V5 R1.0  SERVICE LEVEL 0000 (64-BIT)
10:30:15 SYSTEM NUCLEUS CREATED ON 2004-05-11 AT 13:54:04, LOADED FROM 510RES
10:30:15
10:30:15 ****************************************************************
10:30:15 * LICENSED MATERIALS - PROPERTY OF IBM*                        *
10:30:15 *                                                             *
10:30:15 * 5741-A05 (C) COPYRIGHT IBM CORP. 1983, 2004. ALL RIGHTS     *
10:30:15 * RESERVED. US GOVERNMENT USERS RESTRICTED RIGHTS - USE,      *
10:30:15 * DUPLICATION OR DISCLOSURE RESTRICTED BY GSA ADP SCHEDULE    *
10:30:15 * CONTRACT WITH IBM CORP.                                     *
10:30:15 *                                                             *
10:30:15 * * TRADEMARK OF INTERNATIONAL BUSINESS MACHINES.             *
10:30:15 ****************************************************************
10:30:15
10:30:15 HCPZCO6718I Using parm disk 1 on volume 510RES (device 3750).
10:30:15 HCPZCO6718I Parm disk resides on cylinders 39 through 83.
10:30:15 Start ((Warm|Force|COLD|CLEAN) (DRain) (DIsable)  (NODIRect)
10:30:15       (NOAUTOlog)) or (SHUTDOWN)
10:30:34 COLD DRAIN NOAUTOLOG
10:30:34 NOW 10:30:34 EDT THURSDAY 2004-08-12
10:30:34 Change TOD clock (Yes|No)
10:30:37 NO
10:30:37 The directory on volume 510RES at address 3750 has been brought online.
10:30:40 HCPWRS2513I
10:30:40 HCPWRS2513I Spool files available        31
10:30:40 HCPWRS2513I
10:30:40 HCPWRS2513I Spool files on offline volumes       NONE
10:30:40 HCPWRS2513I Spool files with I/O errors          NONE
10:30:40 HCPWRS2513I Spool files with control errors      NONE
10:30:40 HCPWRS2513I Spool files to be discarded            2
10:30:40 HCPWRS2513I
10:30:40 HCPWRS2513I Total files to be deleted              2
10:30:40 HCPWRS2511A
10:30:40 HCPWRS2511A Spool files will be deleted because of COLD start.
10:30:40 HCPWRS2511A No files have been deleted yet.
10:30:40 HCPWRS2511A To continue COLD start and delete files, enter GO.
10:30:40 HCPWRS2511A To stop COLD start without deleting files, enter STOP.
10:30:55 GO
10:31:00 HCPWRS2512I Spooling initialization is complete.
10:31:00 DASD 3753 dump unit CP IPL pages 14101
10:31:00 HCPAAU2700I System gateway ZVMV5R10 identified.

                                                HOLDING   ZVMV5R10
                                                          42/001
```

ibm.com/redbooks

Native SCSI Disk Support

SCSI disks used as emulated FBA disks

Requires:
- FCP device number
- Target WWPN
- LUN

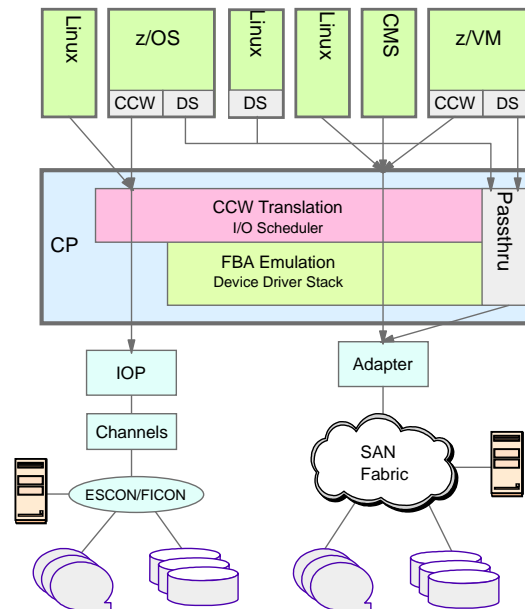**z/VM supports:**
- Individual disks up to 381GB in size
- Directory, paging, spooling must be in first 64GB

ibm.com/redbooks

SCSI Architecture

Install z/VM on SCSI Disk

Specify FBA emulated disks
- INSTPLAN FBA

Define emulated device
- SET EDEVICE command

Attach device
- VARY ON
- ATTACH

Install from media
- INSTDVD

IPL system from HMC

INSTPLAN FBA

```
Integrated 3270 Console for SCZP901:A18
File  Keys  Help
                    *** z/VM INSTALLATION PLANNING ***

   Mark the product(s) selected to be installed into the VMSYS filepool with an
   "F" and those selected to be installed to minidisks with an "M"

   Install To    Product    Install To    Product    Install To    Product
   ----------    -------    ----------    -------    ----------    -------
        M          VM           M          RSCS          M          TCPIP
        M          OSA          M          ICKDSF        M          DIRM
        M          RACF         M          PERFTK        M          VMHCD


   Place a nonblank character in front of the System Default Language you would
   like for your system.

        X AMENG        _ UCENG        _ KANJI        _ GERMAN

   Place a nonblank character in front of the DASD model onto which your
   z/VM system will be loaded.  Only one model may be selected.

        X FBA DASD






      PF1 = HELP   PF3/PF12 = QUIT   PF5 = Process   ENTER = Refresh
                                                               08/008
```

ibm.com/redbooks

SET EDEVICE

```
Integrated 3270 Console for SCZP901:A18
File  Keys  Help
set edevice 5302 type fba attr 2105 fcp_dev b002 WWPN 5005076300c19589 LUN 53020
00000000000
11:47:00 EDEV 5302 was created.
Ready; T=0.01/0.03 11:47:00
set edevice 5303 type fba attr 2105 fcp_dev b003 WWPN 5005076300c19589 LUN 53030
00000000000
11:47:14 EDEV 5303 was created.
Ready; T=0.01/0.03 11:47:14
set edevice 5304 type fba attr 2105 fcp_dev b004 WWPN 5005076300c19589 LUN 53040
00000000000
11:47:32 EDEV 5304 was created.
Ready; T=0.01/0.02 11:47:32
set edevice 5305 type fba attr 2105 fcp_dev b005 WWPN 5005076300c19589 LUN 53050
00000000000
11:47:46 EDEV 5305 was created.
Ready; T=0.01/0.03 11:47:46
set edevice 5306 type fba attr 2105 fcp_dev b006 WWPN 5005076300c19589 LUN 53060
00000000000
11:48:02 EDEV 5306 was created.
Ready; T=0.01/0.03 11:48:02
set edevice 5307 type fba attr 2105 fcp_dev b007 WWPN 5005076300c19589 LUN 53070
00000000000
11:48:15 EDEV 5307 was created.
Ready; T=0.01/0.03 11:48:15
set edevice 5308 type fba attr 2105 fcp_dev b008 WWPN 5005076300c19589 LUN 53080
00000000000
11:49:51 EDEV 5308 was created.
Ready; T=0.01/0.03 11:49:51
vary online 5300-5308
11:50:12 5300 varied online
11:50:14 5301 varied online
11:50:16 5302 varied online
11:50:18 5303 varied online
11:50:20 5304 varied online
11:50:22 5305 varied online
11:50:24 5306 varied online
11:50:26 5307 varied online
11:50:28 5308 varied online
11:50:28 9 device(s) specified; 9 device(s) successfully varied online
Ready; T=0.01/0.01 11:50:28

                                                    RUNNING   IBMVMRAM
                                                               42/001
```

ibm.com/redbooks

Vary On, Attach Device, Install Media

```
VARY ON b002

ATTACH b002 *

INSTDVD
```

IPL z/VM From SCSI - HMC Display

ibm.com

Load

CPC: SCZP901

Image: A18

Load type: ○ Normal ○ Clear ● SCSI ○ SCSI dump

☐ Store status

Load address	b000
Load parameter	sysg
Time-out value	060 60 to 600 seconds
World wide port name	5005076300C19589
Logical unit number	5300000000000000
Boot program selector	0
Boot record logical block address	00000000000000C8
OS specific load parameters	

[OK]   [Reset]   [Cancel]   [Help]

VSWITCH Enhanced Authorization

Connecting to VSWITCH requires authorization

- Standard CP granted authority
  - SET VSWITCH command
  - MODIFY VSWITCH statement
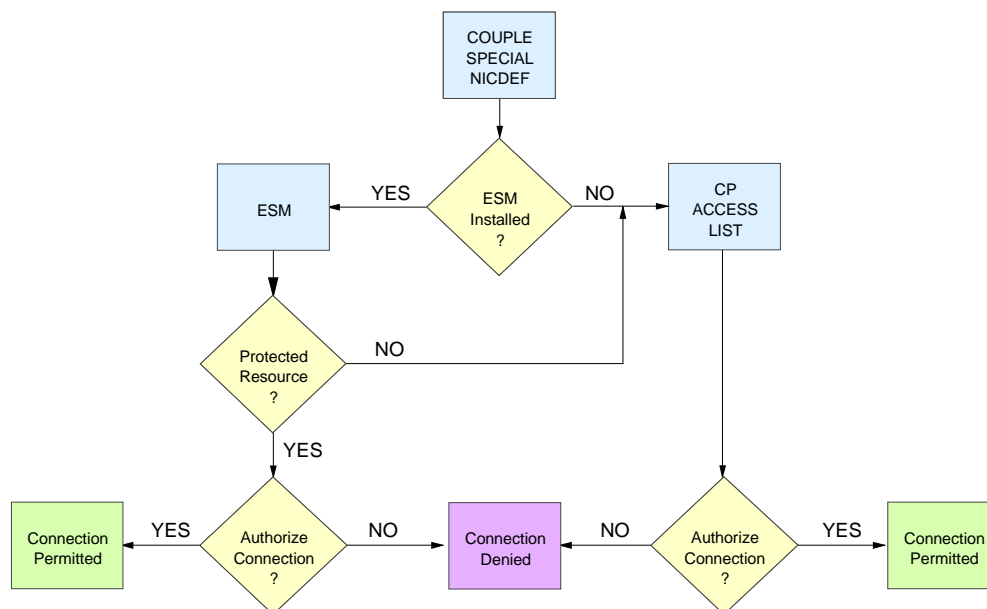- Authorization granted by ESM
  - RACF

Using RACF:

- Define profile
- Grant / revoke access to switch
- Activate profile

VSWITCH Authorization Logic

SET VSWITCH command

```
                                              .-VLAN--ANY--------.
>>--SET VSWITCH--switchname--.-GRAnt--userid--+-----------------+-.-------><
                            !                  !      <-------< ! !
                            !                  '-VLAN----vlanid---' !
                            !-REVoke--userid---------------------!
                            !         <------------<            !
                            !              (1)                  !
                            !-PORTname----portname---------------!
                            !       <---------<                 !
                            !            (2)                    !
                            !-RDEV--.---rdev------.--------------!
                            !       '-NONE--------'             !
                            !-CONnect----------------------------!
                            !-DISCONnect-------------------------!
                            !-QUEuestorage--numberM--------------!
                            !-CONTRoller--.-*-------.------------!
                            !             '-userid1-'           !
                            !-IPTimeout--nnn---------------------!
                            !-NONrouter--------------------------!
                            '-PRIrouter--------------------------'
```

ibm.com/redbooks

Using RACF Authorization

```
RAC RDEFINE VMLAN SYSTEM.SW1 UACC(NONE)
Ready; T=0.01/0.01 15:18:43

RAC PERMIT SYSTEM.SW1 CLASS(VMLAN) ACCESS(UPDATE) ID(LNXSU4)
Ready; T=0.01/0.01 13:34:47

RAC PERMIT SYSTEM.SW1 CLASS(VMLAN) ACCESS(UPDATE) ID(LNXSU4) DELETE
Ready; T=0.01/0.01 13:34:57

RAC SETROPTS CLASSACT(VMLAN)
Ready; T=0.01/0.01 13:43:23
```

ibm.com/redbooks

HyperSwap Function

Continuous availabilty and disaster recovery for Linux guests

- Virtual devices associated to real disk can be swapped
- Relies on Peer-to-Peer Remote Copy (PPRC)
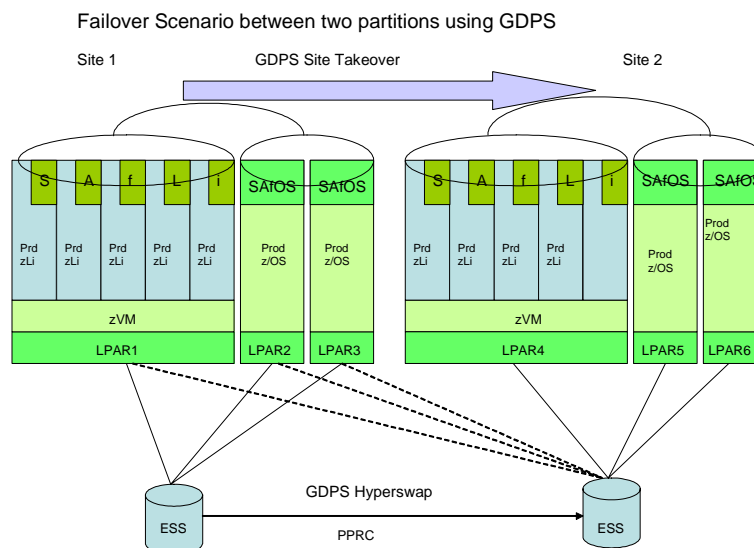- GDPS intends to exploit Hyperswap

Requires:

- GDPS
- IBM Tivoli System Automation for Linux
- PPRC

Redbooks

Failover Scenario Using GDPS

Failover Scenario between two partitions using GDPS

Redbooks