Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

<table>
<thead>
<tr>
<th>Product</th>
<th>IBM*</th>
<th>PowerVM</th>
<th>System z10</th>
<th>z/OS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIX*</td>
<td>IBM eServer</td>
<td>PR/SM</td>
<td>WebSphere*</td>
<td>zSeries*</td>
</tr>
<tr>
<td>BladeCenter*</td>
<td>IBM (logo)*</td>
<td>Smarter Planet</td>
<td>z9*</td>
<td>zVM*</td>
</tr>
<tr>
<td>DataPower*</td>
<td>InfiniBand*</td>
<td>System x*</td>
<td>z10 BC</td>
<td>z/VSE</td>
</tr>
<tr>
<td>DB2*</td>
<td>Parallel Sysplex*</td>
<td>System z*</td>
<td>z10 EC</td>
<td></td>
</tr>
<tr>
<td>FICON*</td>
<td>POWER*</td>
<td>System z9*</td>
<td>zEnterprise</td>
<td></td>
</tr>
<tr>
<td>GDPS*</td>
<td>POWER7*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license there from.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Windows Server and the Windows logo are trademarks of the Microsoft group of companies.

InfiniBand is a trademark and service mark of the InfiniBand Trade Association.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply. All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.
Agenda

- Purpose
- Workflow Overview
- Collecting Data
  - Data Sources
  - Architecture
  - ziomon Usage
- Analyzing Data
  - ziorep_config
  - ziorep_utilization
  - ziorep_traffic
- Hints & Tips
- Data Analysis on Other Platforms
How files end up on Storage Devices

**Linux kernel**

- file system
- Block Devices / Device Mapper
- SCSI stack
- zfcp module
- qdio module
- z/VM
- FCP device

**Hypervisor (optional)**

**System z hardware**

[Diagram showing the flow from file system through Block Devices / Device Mapper, SCSI stack, zfcp module, qdio module, z/VM, and FCP device to SAN.]
Motivation

- Performance issues can potentially stem from multiple sources
- Focusing on a single component will not tell the whole story
- Holistic approach, considering all components involved (Linux, hypervisors, HBAs, SAN with storage devices) is required
- Requires advanced data analysis features, e.g. for filtering and aggregation
- **ziomon** tools aim at providing all of the above
- Availability:
  - SLES: Part of the *s390-tools* package starting with SLES10 SP3 and SLES11 SP1
  - RHEL: Part of the *s390utils* package starting with RHEL 5.4 and provided as *s390utils-ziomon* beginning with RHEL 6
- Dependencies: Requires the *blktrace* package as well as CONFIG_BLK_DEV_I0_TRACE enabled in the kernel config
Workflow Overview

- Collect data using \textit{ziomon}
- Yield resulting files (.log, .agg & .cfg - all required)
- Use reporting tools to analyze data

ziom\textsc{on} Tools

Data

- .log
- .agg
- .cfg

Data

- .config

\textit{ziorep\_config}
\textit{ziorep\_traffic}
\textit{ziorep\_utilization}

"Target system"
Collecting Data: Components

- Components used by ziomon to collect data:
  - `ziomon_fcpconf`: zfcp configuration
  - `ziomon_zfcpdd`: Channel & fabric latencies, QDIO outbound queue utilization
  - `ziomon_util`: zfcp adapter utilization
  - `blkiomon`: blktrace component for SCSI subsystem latencies

- Note: Data collection only starts after the configuration has been captured
Collecting Data: Invocation

- Minimum invocation:

  ```
  user@larsson:~> zimon -d 5 -o data /dev/sda
  ```

- Mandatory parameters to use:
  - `-d` : total data collection duration in minutes
  - `-o` : output file – should be on a device not sampled by `ziomon`!
  - `<device>` : SCSI device to monitor

- `ziomon` will aggregate data during collection, defaults to 60 seconds. I.e. it will not collect data on individual I/O operations.

- Use `-i` to specify an arbitrary interval length for aggregation. E.g. to invoke with highest possible resolution (2 seconds) use

  ```
  user@larsson:~> zimon -d 5 -i 2 -o data /dev/sda
  ```

- Use `-l` to specify size limit for `.log` file.

  ```
  user@larsson:~> zimon -d 5 -l 137M -o data /dev/sda
  ```

  Will wrap around in case specified limit is too low (like in a ring buffer). The overwritten data is aggregated into a single (bigger) interval in a `.agg` file.
Collecting Data: Specifying Devices

- Example:

```bash
user@larsson:~> multipath -ll
36005076303ffcc562000000000000010cc dm-0 IBM,2107900
[size=5.0G][features=0][hwhandler=0]
\_ round-robin 0 [prio=2][active]
\_ 1:0:0:1087127568 sdb 8:16 [active][ready]
\_ 0:0:0:1087127568 sda 8:0 [active][ready]
```

- Multiple ways to specify devices for collecting data:
  - Provide a list of single path devices:
    ```bash
    user@larsson:~> ziomon [...] /dev/sda /dev/sdb
    ```
  - Specifying multipath devices will identify all underlying devices (recommended):
    ```bash
    user@larsson:~> ziomon [...] /dev/mapper/36005076303ff...
    ```
  - Mixing single path and multipath devices is possible
    ```bash
    user@larsson:~> ziomon [...] /dev/mapper/36005076303ff... /dev/sdc /dev/sdd
    ```
Collecting Data: Specifying Devices (continued)

- **Notes:**
  - SCSI tape devices are supported, too
  - `blktrace` (as used by `ziomon`) requires 2MB `vmalloc` memory per device and processor
  - Check current configuration:

```bash
user@larsson:~> grep Vmalloc /proc/meminfo
```

- Use kernel parameter `vmalloc` to increase, e.g. `vmalloc=512M`
Data Analysis

- Use the `ziorep_*` commands to generate reports and investigate data:
  - `ziorep_config`: Insight into the zfcp configuration
  - `ziorep_utilization`: Adapter utilization statistics
  - `ziorep_traffic`: Traffic statistics

- Note: Always pass on all files (.log and .cfg, and possibly .agg) when forwarding data to others for analysis
Data Analysis: Configuration Report Overview

- `ziorep_config` can take a `.cfg` file as input

```
user@larsson:~> ziorep_config -i sample.cfg
```

- Displays information on zfcp-related hardware (HBAs, single and multipath devices) data was collected on.

- Skip the `-i` parameter for a report on the current system.

- 3 different report types available:
  - Adapter report (default, option `-A`): HBA details
  - Multipath report (option `-M`): Multipath/device mapper details
  - Devices report (option `-D`): SCSI device details

- Selection options to limit output to a subset only (can be combined):
  - `-p <port>` limits output to specified WWPN, e.g. `-p 0x5005123456789000`
  - `-l <lun>` limits output to specified LUN, e.g. `-l 0x401040a600000000`
  - `-m <mdev>` limits output to specified multipath device, e.g. `-m 36005076303ffc562000000000000010a6`
  - `-a <adapter>` limits output to specified adapter, e.g. `-a 0.0.3c07`
  - `-s <scsi_host>` limits output to specified SCSI host, e.g. `-s host0`
  - `-d <device>` limits output to specified device, e.g. `-d sde`
Data Analysis: Adapter Configuration Report

`ziorep_config -A`

System z

Linux

mpA | mpB | mpC

z/VSE

Storage Area Network (SAN)

Enterprise Storage

LUN A | LUN B | LUN C

Host 0 | Host 1

HBA 1 | HBA 2 | HBA 3

LUN A | LUN B | LUN C

WPN 1 | WPN 2 | WPN 3 | WPN 4
Data Analysis: Adapter Configuration Report (continued)

```
user@larsson:~> ziorep_config -A -i sample.cfg \ 
   -a 0.0.1800 -a 0.0.1820
Host:    host0
CHPID:   64
Adapter: 0.0.1800
Sub-Ch.: 0.0.0842
Name:    0xc05076ffe4805611
P-Name:  0xc05076ffe4805611
Version: 0x00005
LIC:     0x0000014c
Type:    NPort (fabric via point-to-point)
Speed:   8 Gbit
State:   Online

Host:    host2
CHPID:   66
Adapter: 0.0.1820
Sub-Ch.: 0.0.08a2
Name:    0xc05076ffe4805e11
P-Name:  0xc05076ffe4805e11
Version: 0x00005
LIC:     0x0000014c
Type:    NPort (fabric via point-to-point)
Speed:   8 Gbit
State:   Online
```
Data Analysis: Multipath Report

```
ziorep_config -M
```

**System z**

**Linux**
- mpA
- mpB
- mpC

**z/VSE**

**Storage Area Network (SAN)**

**Enterprise Storage**
- LUN A
- LUN B
- LUN C
```bash
user@larsson:~> ziorep_config -i sample.cfg -M \n-m 36005076303ffc52a00000000000005502
```

<table>
<thead>
<tr>
<th>Path</th>
<th>Drive</th>
<th>Device Name</th>
<th>Mapper Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0.1800</td>
<td>0x5005076303c52a</td>
<td>/dev/sda</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1810</td>
<td>0x5005076303c52a</td>
<td>/dev/sde</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1820</td>
<td>0x5005076303c52a</td>
<td>/dev/sdi</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1830</td>
<td>0x5005076303c52a</td>
<td>/dev/sdj</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1800</td>
<td>0x5005076308c52a</td>
<td>/dev/sdd</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1810</td>
<td>0x5005076308c52a</td>
<td>/dev/sdh</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1820</td>
<td>0x5005076308c52a</td>
<td>/dev/sdl</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1830</td>
<td>0x5005076308c52a</td>
<td>/dev/sdp</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1800</td>
<td>0x5005076310c52a</td>
<td>/dev/sdc</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1810</td>
<td>0x5005076310c52a</td>
<td>/dev/sdg</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1820</td>
<td>0x5005076310c52a</td>
<td>/dev/sdj</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1830</td>
<td>0x5005076310c52a</td>
<td>/dev/sdo</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1800</td>
<td>0x5005076318c52a</td>
<td>/dev/sdb</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1810</td>
<td>0x5005076318c52a</td>
<td>/dev/sdf</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1820</td>
<td>0x5005076318c52a</td>
<td>/dev/sdk</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
<tr>
<td>0.0.1830</td>
<td>0x5005076318c52a</td>
<td>/dev/sdn</td>
<td>/dev/mapper/36005076303ffc52a00000000000005502</td>
</tr>
</tbody>
</table>
Data Analysis: Device Configuration Report

Linux

mpA  HBA 1  Host 1

mpB  HBA 2

mpC  HBA 3

z/VSE

System z

Storage Area Network (SAN)

Enterprise Storage

ziorep_config -D
Data Analysis: Device Configuration Report (continued)

```
user@larsson:~> ziorep_config -i sample.cfg -D \n-l 0x4055400500000000 -l 0x4055400000000000
```

```
0.0.1800 0x5005076303c52a 0x4055400000000000 host0 /dev/sg49 /dev/sdax 67:16 Disk 2107900 IBM 0:0:0:1073758293
0.0.1800 0x5005076303c52a 0x4055400500000000 host0 /dev/sg108 /dev/sdde 70:192 Disk 2107900 IBM 0:0:0:1074085973
0.0.1800 0x5005076303c52a 0x4055400000000000 host0 /dev/sg74 /dev/sdbw 68:160 Disk 2107900 IBM 0:0:1:1073758293
0.0.1800 0x5005076303c52a 0x4055400500000000 host0 /dev/sg118 /dev/sddo 71:96 Disk 2107900 IBM 0:0:1:1074085973
0.0.1800 0x5005076303c52a 0x4055400000000000 host0 /dev/sg56 /dev/sdbe 67:128 Disk 2107900 IBM 0:0:2:1073758293
0.0.1800 0x5005076303c52a 0x4055400500000000 host0 /dev/sg122 /dev/sdds 71:160 Disk 2107900 IBM 0:0:2:1074085973
0.0.1800 0x5005076303c52a 0x4055400000000000 host0 /dev/sg87 /dev/sdcj 69:112 Disk 2107900 IBM 0:0:3:1073758293
0.0.1800 0x5005076303c52a 0x4055400500000000 host0 /dev/sg89 /dev/sdcl 69:144 Disk 2107900 IBM 0:0:3:1074085973
0.0.1810 0x5005076303c52a 0x4055400000000000 host1 /dev/sg80 /dev/sdcc 69:0 Disk 2107900 IBM 1:0:0:1073758293
0.0.1810 0x5005076303c52a 0x4055400500000000 host1 /dev/sg113 /dev/sddj 71:16 Disk 2107900 IBM 1:0:0:1074085973
0.0.1810 0x5005076303c52a 0x4055400000000000 host1 /dev/sg103 /dev/sdcz 70:112 Disk 2107900 IBM 1:0:1:1073758293
0.0.1810 0x5005076303c52a 0x4055400500000000 host1 /dev/sg124 /dev/sddu 71:192 Disk 2107900 IBM 1:0:1:1074085973
0.0.1810 0x5005076303c52a 0x4055400000000000 host1 /dev/sg46 /dev/sdau 66:224 Disk 2107900 IBM 1:0:2:1073758293
0.0.1810 0x5005076303c52a 0x4055400500000000 host1 /dev/sg122 /dev/sdds 71:160 Disk 2107900 IBM 1:0:2:1074085973
0.0.1810 0x5005076303c52a 0x4055400000000000 host1 /dev/sg70 /dev/sdcj 69:112 Disk 2107900 IBM 1:0:3:1073758293
0.0.1810 0x5005076303c52a 0x4055400500000000 host1 /dev/sg80 /dev/sdcl 69:144 Disk 2107900 IBM 1:0:3:1074085973
0.0.1820 0x5005076303c52a 0x4055400000000000 host2 /dev/sg110 /dev/sddg 70:224 Disk 2107900 IBM 2:0:0:1073758293
0.0.1820 0x5005076303c52a 0x4055400500000000 host2 /dev/sg120 /dev/sdq 71:128 Disk 2107900 IBM 2:0:0:1074085973
0.0.1820 0x5005076303c52a 0x4055400000000000 host2 /dev/sg69 /dev/sdbr 68:80 Disk 2107900 IBM 2:0:1:1073758293
0.0.1820 0x5005076303c52a 0x4055400500000000 host2 /dev/sg71 /dev/sdbt 68:112 Disk 2107900 IBM 2:0:1:1074085973
0.0.1820 0x5005076303c52a 0x4055400000000000 host2 /dev/sg119 /dev/sddf 70:208 Disk 2107900 IBM 2:0:2:1073758293
0.0.1820 0x5005076303c52a 0x4055400500000000 host2 /dev/sg116 /dev/sdcm 71:64 Disk 2107900 IBM 2:0:2:1074085973
0.0.1820 0x5005076303c52a 0x4055400000000000 host2 /dev/sg92 /dev/sdco 69:192 Disk 2107900 IBM 2:0:3:1073758293
0.0.1820 0x5005076303c52a 0x4055400500000000 host2 /dev/sg117 /dev/sddn 71:90 Disk 2107900 IBM 2:0:3:1074085973
0.0.1830 0x5005076303c52a 0x4055400000000000 host3 /dev/sg125 /dev/sdvc 71:208 Disk 2107900 IBM 3:0:0:1073758293
0.0.1830 0x5005076303c52a 0x4055400500000000 host3 /dev/sd97 /dev/sdct 70:16 Disk 2107900 IBM 3:0:1:1073758293
0.0.1830 0x5005076303c52a 0x4055400000000000 host3 /dev/sg99 /dev/sdcv 70:48 Disk 2107900 IBM 3:0:1:1074085973
0.0.1830 0x5005076303c52a 0x4055400500000000 host3 /dev/sd14 /dev/sddk 71:32 Disk 2107900 IBM 3:0:2:1073758293
0.0.1830 0x5005076303c52a 0x4055400000000000 host3 /dev/sd12 /dev/sddy 128:0 Disk 2107900 IBM 3:0:2:1074085973
0.0.1830 0x5005076303c52a 0x4055400500000000 host3 /dev/sd12 /dev/sdki 71:0 Disk 2107900 IBM 3:0:3:1073758293
0.0.1830 0x5005076303c52a 0x4055400000000000 host3 /dev/sd12 /dev/sdxi 71:240 Disk 2107900 IBM 3:0:3:1074085973
```
Data Analysis: Adapter Utilization Report

- Ficon Express card with FCP CHPIDs has a CPU, bus and adapter chip
- Utilization report provides two sections:
  - **FCP channel** report provides details on
    - CPU utilization
    - Bus utilization
    - Adapter utilization
  - **Virtual HBA** report provides details on
    - QDIO utilization, queue full and failure conditions
    - Throughput
    - I/O requests numbers
- Purpose: Identify potential bottlenecks within the adapter
- Note: Percentages reported are for entire FCP channel – no fraction according to utilization given when FCP channel is shared by multiple OS images! I.e. high utilization during phases of inactivity point to activity on other OS images sharing the same channel.
Data Analysis: Adapter Utilization Report

- Use '-s' to get a quick overview of the available data

```
user@larsson:~> ziorep_utilization -s sample.log
```

Data Summary
------------

Aggregated range: none
Interval length: 2 seconds
HBA/CHPID: 0.0.591d/42 0.0.5b1d/43
WWPN/LUN (dev): 0x5005076303000104/0x4021402100000000 (/dev/sdi)
  0x50050763030b0104/0x4021402100000000 (/dev/sdk)
  0x5005076303100104/0x4021402100000000 (/dev/sdm)
  0x5005076303100104/0x4021402100000000 (/dev/sdn)

- Use '-b <begin>' and '-e <end>' to limit output to specified timeframe
- Use '-i <length>' to aggregate data to bigger intervals
- Use '-c <chpid>' to limit output to specified adapter
- Notes:
  - Check the interval length to figure out available aggregation values
  - Aggregated range only listed in case of wrap-around in .log file (.agg exists)
# Data Analysis: FCP Channel Utilization Report

```plaintext
user@larsson:~> ziorep_utilization sample.log -b "2012-12-11 10:58:55"

CHP|adapter in %|--bus in %|--cpu in %|
ID min max avg min max avg min max avg
2012-12-11 10:58:55
42 3 3 3.0 20 20 20.0 2 2 2.0
43 3 3 3.0 12 12 12.0 1 1 1.0
10:58:57
42 3 3 3.0 18 18 18.0 2 2 2.0
43 4 4 4.0 15 15 15.0 1 1 1.0
10:58:59
42 2 2 2.0 16 16 16.0 2 2 2.0
43 3 3 3.0 17 17 17.0 1 1 1.0
10:59:01
42 1 1 1.0 16 16 16.0 2 2 2.0
43 4 4 4.0 15 15 15.0 1 1 1.0
[...]

CHP Bus-ID |qdio util.|queue|fail|-thp in MB/s|-I/O reqs-|
ID max avg full erc rd wrt rd wrt
2008-12-11 10:58:55
42/0.591d 99.2 14.0 12 0 1.0 2.7 5 692
43/0.5b1d 91.4 9.3 0 0 0.5 3.3 4 517
[...]
```
### Data Analysis: Virtual HBA Utilization Report

```bash
user@larsson:~> ziorep_utilization sample.log -i 60 -c 43

<table>
<thead>
<tr>
<th>CHP Bus-ID</th>
<th>qdio util.%</th>
<th>queu</th>
<th>fail</th>
<th>thp in MB/s-</th>
<th>I/O reqs-</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>max  avg  full erc rd  wrt rd wrt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012-12-11 10:52:55 43/0.0.5b1d</td>
<td>99.2 13.4 302 0 0.2 2.6 880 15K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:53:55 43/0.0.5b1d</td>
<td>98.4 10.0 203 0 0.3 3.1 163 1.3K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:54:55 43/0.0.5b1d</td>
<td>99.2 12.6 178 0 0.2 3.1 1.3K 10K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:55:55 43/0.0.5b1d</td>
<td>99.2 4.2 79 0 0.3 3.0 729 4.6K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:56:55 43/0.0.5b1d</td>
<td>0.0 0.0 0 0 0.0 0.0 168 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:57:55 43/0.0.5b1d</td>
<td>99.2 8.9 136 0 3.7 1.9 2.9K 11K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:58:55 43/0.0.5b1d</td>
<td>100.0 12.6 177 0 0.6 2.8 335 15K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:59:55 43/0.0.5b1d</td>
<td>99.2 14.0 177 0 0.3 3.1 367 19K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:00:55 43/0.0.5b1d</td>
<td>99.2 13.9 188 0 0.3 3.8 1.5K 14K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:01:55 43/0.0.5b1d</td>
<td>99.2 14.1 162 0 0.4 3.4 1.8K 14K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11:02:55 43/0.0.5b1d</td>
<td>99.2 8.5 150 0 0.3 3.1 1.1K 8.2K</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

[...]

Complete your sessions evaluation online at SHARE.org/SFEval
Traffic Analysis

Linux kernel

Hypervisor (optional)

System z hardware

<table>
<thead>
<tr>
<th>file system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block Devices / Device Mapper</td>
</tr>
<tr>
<td>SCSI stack</td>
</tr>
<tr>
<td>zfcp module</td>
</tr>
<tr>
<td>qdio module</td>
</tr>
<tr>
<td>z/VM</td>
</tr>
<tr>
<td>FCP device</td>
</tr>
</tbody>
</table>

SAN

Channel latency

I/O Subsystem latency

Fabric latency
Data Analysis: Traffic Report

- `ziorep_traffic` aims at identifying latencies of all components involved
- Provides two views on data
  - Regular view listing min/max/avg/std (default)
  - Detailed view providing histograms
- Supports option `-s` for a data summary (just like `ziorep_utilization`)
- Supports device selection options `-c/-u/-p/-l/-d/-m`
  (similar to `ziorep_config`)
- Provides data aggregation capabilities
- **Notes:**
  - Numbers reported consider *serialized* traffic! I.e. parallelism of requests is not accounted for. Hence throughputs are likely to divert from what other tools will report!
  - `ziomon` has a performance impact of approx. 5 percent on CPU utilization
user@larsson:~> ziorep_traffic sample.log -i 60 -l

0x4021402f00000000 -l 0x4021402d00000000

2012-12-11 10:52:55
10:54:55

© 2012 IBM Corporation

Data Analysis: Traffic Report - Default View

Linux on System z: zFCP Performance Analysis
Data Analysis: Traffic Report – Data Filtering Options

- To focus on individual devices, use '-d'

```bash
user@larsson:~> ziorep_traffic sample.log -i 60 -d sdi
```

<table>
<thead>
<tr>
<th>WWPN</th>
<th>LUN</th>
<th>I/O rt MB/s</th>
<th>thrp in MB/s</th>
<th>----I/O requests----</th>
<th>----I/O subs. lat. in us---</th>
<th>----channel lat. in us---</th>
<th>----fabric lat. in us---</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>max</td>
<td>avg</td>
<td>stdev</td>
<td>#reqs</td>
<td>rd</td>
<td>wrt</td>
<td>bidi</td>
</tr>
<tr>
<td>2012-12-11 10:52:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10:55:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10:56:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10:57:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10:58:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>6.9</td>
<td>2.7</td>
<td>1.159K</td>
<td>624</td>
<td>13</td>
</tr>
<tr>
<td>10:59:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>12.2</td>
<td>2.8</td>
<td>1.182K</td>
<td>320</td>
<td>6</td>
</tr>
<tr>
<td>11:00:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

- Specifying devices over a number of specifiers will use the union

```bash
user@larsson:~> ziorep_traffic sample.log -i 60 -d sdi -p
```

<table>
<thead>
<tr>
<th>WWPN</th>
<th>LUN</th>
<th>I/O rt MB/s</th>
<th>thrp in MB/s</th>
<th>----I/O requests----</th>
<th>----I/O subs. lat. in us---</th>
<th>----channel lat. in us---</th>
<th>----fabric lat. in us---</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>max</td>
<td>avg</td>
<td>stdev</td>
<td>#reqs</td>
<td>rd</td>
<td>wrt</td>
<td>bidi</td>
</tr>
<tr>
<td>2012-12-11 10:52:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10:55:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>4.4</td>
<td>1.5</td>
<td>857.7</td>
<td>114</td>
<td>2</td>
</tr>
<tr>
<td>10:56:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>14.5</td>
<td>2.7</td>
<td>1.151K</td>
<td>895</td>
<td>17</td>
</tr>
<tr>
<td>10:57:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>4.8</td>
<td>1.5</td>
<td>850.5</td>
<td>144</td>
<td>6</td>
</tr>
<tr>
<td>10:58:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>12.2</td>
<td>2.8</td>
<td>1.182K</td>
<td>320</td>
<td>6</td>
</tr>
<tr>
<td>11:00:55</td>
<td>0x5005076303000104:0x4021402100000000</td>
<td>0.0</td>
<td>2.9</td>
<td>1.2</td>
<td>769.3</td>
<td>170</td>
<td>7</td>
</tr>
<tr>
<td>0x5005076303000104:0x4021402200000000</td>
<td>0.0</td>
<td>18.1</td>
<td>2.4</td>
<td>1.082K</td>
<td>896</td>
<td>11</td>
<td>885</td>
</tr>
</tbody>
</table>

[...]
user@larsson:~> ziorep_traffic sample.log -i 60 -l \x04021402f00000000 -l 0x4021402d00000000 -D

|------------------------I/O request sizes in KBytes------------------------|
| 0    1    2    4    8   16   32   64  128  256  512   1K   2K   4K   8K  >8K |
|------------------------I/O subsystem latency in us------------------------|
| 0    8   16   32   64  128  256  512   1K   2K   4K   8K  16K  32K  64K 128K 256K 512K   1M   2M   4M  |
|------------------------channel latency in us--------------------------------|
| 0    1    2    4    8   16   32   64  128  256  512   1K   2K   4K   8K  16K  32K  64K 128K  256K |

|WWPN                LUN       |------------------------fabric latency in us------------------------|
| 0    8   16   32   64  128  256  512   1K   2K   4K   8K  16K  32K  64K 128K 256K 512K   1M   2M   4M   8M  |

2012-12-11 10:52:55
Data Analysis: Traffic Aggregation

- Use option '-C' to specify an aggregation level:
  - 'a'/u' aggregates by CHPID / bus-ID
  - 'p' aggregates by WWPN
  - 'm' aggregates by multipath device
  - 'A' aggregates all data

- Can be combined with device selection options '-c/-u/-p/-l/-d/-m'

- Note: Don't confuse device selection with aggregation options:
  - **Device Selection**: restricts use of data records according to the scope specified (e.g. data records for specific LUNs only)
  - **Aggregation**: generates combined results for given level of detail (e.g. combined result for all I/O handled by specified adapter)

⇒ **Filtering** applies to input data, aggregation shapes the resulting output
Data Analysis: Traffic Aggregation Examples

- Aggregate by bus-ID

```bash
user@larsson:~> ziorep_traffic sample.log -i 60 -Cu
```

```plaintext
<table>
<thead>
<tr>
<th>Bus-ID</th>
<th>I/O rt MB/s</th>
<th>thrp in MB/s</th>
<th>I/O requests</th>
<th>I/O subs. lat. in us</th>
<th>channel lat. in us</th>
<th>fabric lat. in us</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min</td>
<td>max</td>
<td>avg</td>
<td>stdev</td>
<td>#reqs</td>
<td>rd</td>
</tr>
<tr>
<td>0.0.591d</td>
<td>0.0</td>
<td>115.5</td>
<td>2.8</td>
<td>1.154K</td>
<td>338</td>
<td>12K</td>
</tr>
<tr>
<td>0.0.5b1d</td>
<td>0.0</td>
<td>76.1</td>
<td>2.6</td>
<td>1.112K</td>
<td>880</td>
<td>15K</td>
</tr>
<tr>
<td>10:53:55</td>
<td>0.0</td>
<td>104.1</td>
<td>9.2</td>
<td>2.113K</td>
<td>183</td>
<td>20</td>
</tr>
<tr>
<td>0.0.5b1d</td>
<td>0.0</td>
<td>12.4</td>
<td>3.1</td>
<td>1.224K</td>
<td>1496</td>
<td>163</td>
</tr>
<tr>
<td>2012-12-11 10:52:55</td>
<td>0.0</td>
<td>115.9</td>
<td>2.9</td>
<td>1.193K</td>
<td>14.1K</td>
<td>12K</td>
</tr>
<tr>
<td>0.0.5b1d</td>
<td>0.0</td>
<td>91.8</td>
<td>3.1</td>
<td>1.215K</td>
<td>11.7K</td>
<td>10K</td>
</tr>
</tbody>
</table>
```

- Aggregate all available devices in detailed traffic analysis mode

```bash
user@larsson:~> ziorep_traffic sample.log -i 60 -CA -D
```

```plaintext
<table>
<thead>
<tr>
<th>I/O request sizes in KBytes</th>
<th>I/O subsystem latency in us</th>
<th>channel latency in us</th>
<th>fabric latency in us</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 4 8 16 32 64 128 256 512 1K 2K 4K 8K &gt;8K</td>
<td>0 8 16 32 64 128 256 512 1K 2K 4K 8K 16K 32K 64K 128K 256K 512K 1M 2M 4M 8M 16M 32M &gt;32M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>0 8 16 32 64 128 256 512 1K 2K 4K 8K 16K 32K 64K 128K 256K 512K 1M 2M 4M 8M 16M 32M &gt;32M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>------------------------------------------</td>
<td>-----------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>0 8 16 32 64 128 256 512 1K 2K 4K 8K 16K 32K 64K 128K 256K 512K 1M 2M 4M 8M 16M 32M &gt;32M</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```bash
[...]
```
Data Analysis: Traffic Aggregation Examples (continued)

- Aggregate a selection of devices into a single interval

```
user@larsson:~> ziorep_traffic sample.log -i 0 -C A -d sdp -p 0x50050763030b0104 -c 42
```

```
* |I/O rt MB/s|thrp in MB/s|----I/O requests----|----I/O subs. lat. in us---|--channel lat. in us---|---fabric lat. in us---|
min max avg stdev #reqs rd wrt bidi min max avg stdev min max avg stdev min max avg stdev
2012-12-11 11:51:55
* 0.0 115.9 2.9 1.181K 192K 42K 150K 0 154 315M 95.63K 1.048M 19 25K 3.725K 4.056K 68 969K 85.31K 109.5K
```

- Same selection, but detailed view

```
user@larsson:~> ziorep_traffic sample.log -i 0 -C A -d sdp -p 0x50050763030b0104 -c 42 -D
```

```
|------------------------I/O request sizes in KBytes----------------------------|
0    1    2    4    8    16    32    64    128    256    512    1K    2K    4K    8K   >8K
|------------------------I/O subsystem latency in us--------------------------|
0    8    16    32    64    128    256    512    1K    2K    4K    8K    16K    32K    64K    128K    256K    512K    1M    2M    4M    8M    16M    32M   >32M
|------------------------channel latency in us---------------------------------|
0    8    16    32    64    128    256    512    1K    2K    4K    8K    16K    32K    64K    128K    256K    512K    1M    2M    4M    8M    16M    32M   >32M
|------------------------fabric latency in us---------------------------------|
0    8    16    32    64    128    256    512    1K    2K    4K    8K    16K    32K    64K    128K    256K    512K    1M    2M    4M    8M    16M    32M   >32M
2012-12-11 11:51:55
* 13K 0 0 55K 614 663 1.6K 3.2K 5.8K 12K 100K 0 0 0 0 0 0
0 0 0 0 12K 8.0K 2.1K 3.3K 6.1K 7.6K 5.5K 18K 23K 36K 70K 826 157 21 7 4 1 0 3
0 0 0 0 0 0 38K 11K 1.0K 963 2.4K 13K 23K 36K 41K 25K 2.4K 0 0 0 0 0
0 0 0 0 0 0 11K 8.4K 2.6K 2.2K 3.6K 6.3K 7.7K 7.2K 21K 20K 39K 64K 693 126 0 0 0 0 0
```
**Data Analysis: Hints & Tips**

- Use `-i 0` to get a good starting point to identify anomalies quickly.
Data Analysis on Other Platforms

- Reporting tools do not depend on s390 architecture
- To compile on other platforms:
  - Extract source of `s390-tools` package, either from your distribution's source rpm, or download from http://www.ibm.com/developerworks/linux/linux390/s390-tools.html and extract via
    ```
    user@larsson:~> tar xvfj s390-tools-1.17.0.tar.bz2
    ```
  - Change to `ziomon` subdirectory, build and install
    ```
    user@larsson:~> cd s390-tools-1.17.0
    user@larsson:~> make
    user@larsson:~> sudo make install
    ```
- Run reporting tools on any given data
Data Analysis on Other Platforms (continued)

- Use option '-x' in `ziorep_utilization` and `ziorep_traffic` to export data to .csv format, e.g.

  ```bash
  user@larsson:~> ziorep_traffic data.log -x -i 300 -b "2012-11-27 08:45:21" -C a
  ```

- All data filtering options specified will apply to exported data
- Import in spreadsheet and perform further processing

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>timestamp</td>
<td>aggregated</td>
<td>CHPID</td>
<td>I/O rate in MB/s min</td>
<td>I/O rate in MB/s max</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>3a</td>
<td>1.49</td>
<td>88.2</td>
</tr>
<tr>
<td>2</td>
<td>07/16/09 02:02 PM</td>
<td>0</td>
<td>3a</td>
<td>1.44</td>
<td>67.25</td>
</tr>
<tr>
<td>3</td>
<td>07/16/09 02:07 PM</td>
<td>0</td>
<td>3a</td>
<td>0.53</td>
<td>103.14</td>
</tr>
<tr>
<td>4</td>
<td>07/16/09 02:07 PM</td>
<td>0</td>
<td>47</td>
<td>0.27</td>
<td>68.86</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Throughput Avg**

(higher is better)

**I/O Latency Avg**

(lower is better)
References

- Linux on System z on DeveloperWorks
  http://www.ibm.com/developerworks/linux/linux390

- Linux on System z Documentation
  Specifically:
  - Device Drivers, Features, and Commands
  - How to use FC-attached SCSI devices with Linux on System z (chapters 11&12)

- Linux on System z – Downloads
Questions?

Stefan Raspl
Linux on System z Development

Schönaicher Strasse 220
71032 Böblingen, Germany
Phone +49 (0)7031-16-2177
stefan.raspl@de.ibm.com