IBM AIX multipath I/O (MPIO) resiliency and problem determination

Article Number: 86 | Rating: 5/5 from 1 votes | Last Updated: Wed, May 30, 2018 10:46 AM

IBM AIX multipath I/O (MPIO) resiliency and problem determination

Using the new features of the IBM AIX MPIO lsmpio command for storage area network (SAN) fabric problem determination

Gary Domrow
Published on February 19, 2018

0

Fibre Channel (FC) based storage area network (SAN) fabrics are widely used to attach computers to storage devices. Typically, users configure the SAN fabric with more than one path between the computer and the storage device. This is accomplished by using multiple FC ports on both, the computer and the storage device. Using multiple paths might provide greater bandwidth, but more importantly multiple paths provide redundancy to protect against failures of SAN fabric components.

Using multiple paths to access a storage device is referred to as multipath I/O (MPIO). The IBM® AIX® operating system has supported MPIO for many years. IBM frequently updates the MPIO features of AIX to assist clients using increasingly complex SAN configurations.

Configurations for multipath redundancy

SAN administrators may use several configuration techniques to maximize redundancy. They may configure an AIX logical partition (LPAR) to use ports from two different FC host bus adapters (HBAs). Thus if a single adapter fails, the other adapter provides access to the storage device. Likewise, using multiple FC ports on the storage device protects against the failure of a single port. Many storage devices are designed with two or more controllers. Connecting to one or more ports on each of the controllers protects against the failure of an entire storage device controller.

Finally, the SAN administrator might configure two distinct SAN fabrics to provide even more protection. For example, connect AIX LPAR FC port A to SAN fabric 1, which is connected to port 0 on each of two storage device controllers. Connect AIX LPAR FC port B to fabric 2, which is connected to port 1 on each of the two storage device controllers. In such a configuration, the failure of any single item or an entire FC switch in a fabric does not break all connectivity between the AIX LPAR and the storage device.

Figure 1 shows such a configuration with redundant ports and redundant SAN fabrics.

Figure 1. Redundant SAN Configuration

Types of SAN errors

If a component has a solid, permanent failure, the failure can be easily detected either by the MPIO software in AIX or by the firmware in the storage device. All the MPIO paths that use the failed component indicate a **Failed** state, and error log entries are created against that component, and the MPIO software ceases using the failed paths for user I/O.

For example, if the fiber connected to a particular AIX FC port is pulled out, all currently used paths originating at that port move to the **Failed** state. The AIX system error report shows errors for a particular FC port and shows disk path failures for paths using that particular port. The AIX lsmpio command shows many paths in the **Failed** state for one particular port. Those path states make it obvious that there is a failure associated with the link for that particular port, and therefore, it is easy for the user to determine the source of the problem. Likewise, for such failures it is also easy for the AIX MPIO software to detect and react to the failure by discontinuing the use of any path that uses the failed component.

A second class of failure presents greater difficulty, both for the SAN administrator and MPIO software. If a component is in some way degraded but still partially functional, it can result in recurring intermittent failures. For example, perhaps the optical source on an FC port is becoming weak, a connection between the fiber and the port of an FC switch is not secure, or some device on a SAN fabric is malfunctioning in a way that is flooding that SAN fabric with extraneous traffic. All these types of errors may allow some FC packets to be transmitted correctly on the path, while other packets on that same path experience errors, causing a connection to partially work.

Processing intermittent errors

Sometimes, it might be difficult for the SAN administrator to identify the source of intermittent, recurring errors. Likewise, the host MPIO software may have difficulty in identifying the source

component causing failures. So, it may continue to attempt to use the path that contains the failing component. To compound the issue, some error recovery actions may cause errors to occur on paths that are actually healthy. For example, the AIX LPAR may send a *LUN Reset* request to a device to clear an error state. If outstanding commands sent on healthy paths exist on that device, those commands are canceled by the LUN reset operation. Thus the host LPAR detects induced errors on healthy paths.

The default AIX active-active path control module (PCM) in the newest technology levels of AIX (AIX 7.1 TL5 and AIX 7.2 TL2) was improved to better detect and respond to such intermittent errors. The new improvements, include the following items:

- Better distinguish between the real errors and the errors induced by error recovery
- Additional use of the **Degraded** path state
- Enhancements to the lsmpio command output

When there are intermittent errors in the SAN fabric, the AIX MPIO software may need to initiate recovery actions, such as sending a LUN reset, that could induce additional errors. The enhancements in the new TL attempt to distinguish these induced errors from errors caused by the SAN issue to better identify the good paths and the problematic paths.

The new AIX MPIO software also makes greater use of the **Degraded** path state to limit the use of paths that recently experienced errors. In the lsmpio list of path states, a degraded path is identified as **Deg** in the **path_status** column. The MPIO software avoids selecting a degraded path for I/O unless no other good paths exist to choose from. A path stays in the **Degraded** state from the time it experiences certain errors until there have been five successful health check commands processed on that path.

Another feature of the recent improvements to the AIX MPIO software is the expansion of the output of the lsmpio command to include error counts per adapter and per remote port, allowing the SAN administrator to better deduce the source of intermittent, recurring errors. The remaining sections of this article explain these improvements.

The AIX lsmpio command

The lsmpio command implements the -a and -r flags to show the local and remote FC ports used by the disks. The -a flag (for adapter) shows the local adapters; the -r flag (for remote ports) shows the target device ports that each adapter is connected to and the number of paths in each possible state, as shown in Example 1. The lsmpio command in older technology levels of AIX supports these two flags. Note that

the lsmpio command depends on support from the PCM. The AIX default active-active PCM implements all the features of the lsmpio command. Other PCMs (such as the AIX active-passive PCM or SDD PCM) implement fewer of the lsmpio features. The lsmpio examples in this article are all from the AIX active-active PCM.

Example 1. lsmpio command output for the -a and -r flags

```
# lsmpio -ar

Adapter Driver: fscsi0 -> AIX PCM

    Adapter WWPN:  10000000c97080b1

    Link State:    Up

                    Paths      Paths      Paths      Paths

    Remote Ports        Enabled   Disabled     Failed    Missing        ID

    500a098396a7d4ca         29         0         0         0    0x62000

    500a098386a7d4ca         20         0         0         0    0x62100

    500507630418d5a0         10         0         0         0    0x62300

    50050763041815a0         10         0         0         0    0x62800

    50050763041855a0         10         0         0         0    0x62900



Adapter Driver: fscsi1 -> AIX PCM

    Adapter WWPN:  10000000c99d9791

    Link State:    Up
```

```
                Paths      Paths      Paths      Paths

Remote Ports        Enabled   Disabled     Failed    Missing        ID

500a098396a7d4ca        29         0         0         0    0x62000

500a098386a7d4ca        20         0         0         0    0x62100

500507630418d5a0        10         0         0         0    0x62300

50050763041815a0        10         0         0         0    0x62800

50050763041855a0        10         0         0         0    0x62900
```

The new AIX technology levels added the -e flag (for errors) to the lsmpio command, to be used with the -a and -r flags. When the SAN administrator specifies the -e flag, the lsmpio command output includes error counts for the local adapters and for the remote ports. By examining these error counts, the SAN administrator can deduce the SAN fabric components that causes intermittent errors. Example 2 shows the output of the lsmpio command with the -e flag along with -a and -r flags.

Example 2. lsmpio -are output

```
# lsmpio -are

Adapter Driver: fscsi1 -> AIX PCM

    Adapter WWPN:  10000090fa020d7d

    Link State:    Up

    Connection Errors
```

```
Last 10 Minutes:                0

Last 60 Minutes:                0

Last 24 Hours:                  0

Total Errors:                   0

                  Connection Errors

                Last 10     Last 60     Last 24

                Minutes     Minutes     Hours

500507680b318662        0           0           0

500507680b318663        0           0           0

5005076802364af3        0           0           0

5005076802364af4        0           0           0
```

In this AIX LPAR, just one FC adapter is used, and it is attached to four remote ports. This output shows the current link state, and then a count of errors that occurred on the adapter. Following that is a table showing the error count for the remote port that was in use when the error occurred. Note that if the error count for the remote ports is zero or very small, it may indicate that the port is working fine or it may indicate that the port is not currently being used. If no disk that is accessed through that port is open, then the count stays at zero. If the port is on the non-preferred controller for all open disks, the port may only be used rarely. So nonzero counts, especially large nonzero counts, are more interesting than a small or zero error count.

Also notice that the output categorizes the error count into different time ranges. This gives an indication if the errors are still occurring or if perhaps there was a temporary issue that has not happened recently.

So how might the SAN administrator use the error count? The following sections discuss the output of various samples of the lsmpio command. To generate errors, a FC analyzer or jammer is placed at various locations in the SAN fabric between the AIX LPAR and the storage device, and set to occasionally drop frames. This simulates the failing components in different parts of the SAN. The following paragraphs examine and compare the output for different scenarios, providing an example of deducing where in the SAN fabric intermittent errors are being introduced.

In my configuration, the AIX LPAR had two FC adapters attached to a single SAN fabric. The AIX LPAR had visibility to ports on an IBM Storwize® V7000 storage device and an IBM XIV® Storage System device.

Now assume that you start to see a slow down in the I/O throughput. Looking at the system error report, there may be TEMP disk errors, indicating that certain Small Computer System Interface (SCSI) commands failed on their first attempt, but were successful when retried, perhaps down a different path. When the lsmpio command is used, you might see output as shown in Example 3.

Example 3. lsmpio output showing local adapter errors

```
# lsmpio -are

Adapter Driver: fscsi0 -> AIX PCM

    Adapter WWPN:  10000000c957167e

    Link State:    Up

    Connection Errors

    Last 10 Minutes:                    1

    Last 60 Minutes:                    1

    Last 24 Hours:                      1

    Total Errors:                       1

                        Connection Errors
```

	Last 10 Minutes	Last 60 Minutes	Last 24 Hours
5005076802364af3	1	1	1
5005076802364af4	0	0	0
5005076802264af4	0	0	0
5005076802264af3	0	0	0
5001738000330171	0	0	0
5001738000330173	0	0	0

Adapter Driver: fscsi1 -> AIX PCM

Adapter WWPN: 10000000c957167d

Link State: Up

Connection Errors

Last 10 Minutes: 83

Last 60 Minutes: 83

Last 24 Hours: 83

Total Errors: 83

Connection Errors

	Last 10	Last 60	Last 24

```
                Minutes       Minutes       Hours

5005076802364af3        16          16          16

5005076802364af4        14          14          14

5005076802264af4        12          12          12

5005076802264af3        10          10          10

5001738000330171        15          15          15

5001738000330173        16          16          16
```

Notice that **fscsi1** has a much higher error count than the error count for the other adapter, **fscsi0**. Also notice that the errors occurred across all of the remote ports accessed by fscsi1. These traits indicate that the SAN issue is associated with the AIX FC adapter **fscsi1.** It may be a problem with that adapter, the fiber connected to the adapter, or perhaps the port on the FC switch that the fiber is connected to. But something associated with that adapter is causing errors. In this case, I had the FC jammer connected between **fscsi1** and the FC switch.

Note that there could be cases where the error total listed directly under the adapter exceed the sum of the errors on the remote ports under the adapter, as some errors may not be associated with a particular remote port. This also indicates a possible problem with the local FC adapter.

Now consider the lsmpio -are output shown in Example 4.

Example 4. lsmpio output showing SAN fabric errors

```
# lsmpio -are

Adapter Driver: fscsi0 -> AIX PCM
```

Adapter WWPN: 10000000c957167e

Link State: Up

Connection Errors

Last 10 Minutes: 4

Last 60 Minutes: 33

Last 24 Hours: 33

Total Errors: 33

Connection Errors

	Last 10 Minutes	Last 60 Minutes	Last 24 Hours
5005076802364af3	2	18	18
5005076802264af3	2	15	15
5005076802264af4	0	0	0
5005076802364af4	0	0	0

Adapter Driver: fscsi1 -> AIX PCM

Adapter WWPN: 10000000c957167d

Link State: Up

Connection Errors

Last 10 Minutes: 3

Last 60 Minutes: 33

Last 24 Hours: 33

Total Errors: 33

Connection Errors

	Last 10 Minutes	Last 60 Minutes	Last 24 Hours
5005076802364af3	3	19	19
5005076802264af3	0	14	14
5005076802264af4	0	0	0
5005076802364af4	0	0	0

Notice that both fscsi0 and fscsi1 experienced errors and therefore, the SAN administrator knows that the issue is not associated with a single AIX HBA. Also, note that two different storage device ports have errors indicating that the issue is not associated with a single port on the storage device. Thus the issue must be within the SAN fabric. Further problem determination requires knowledge of the SAN fabric topology. In this case, I had (somewhat artificially) plugged two of the storage device ports into a second switch, and inserted the jammer in the link between the switches [the inter-switch link (ISL)]. Causing errors in that ISL affected both AIX HBAs and both ports on the storage device attached to the second switch. But this shows how errors spread across multiple ports on both the AIX end of the connection and the storage end of the connection indicate an issue exists within the SAN fabric.

Finally, consider the output of the lsmpio -are command as shown in the Example 5.

Example 5. lsmpio output showing storage port errors

```
# lsmpio -are

Adapter Driver: fscsi0 -> AIX PCM

    Adapter WWPN:  10000000c957167e

    Link State:    Up

    Connection Errors

    Last 10 Minutes:                3

    Last 60 Minutes:                23

    Last 24 Hours:                  23

    Total Errors:                   23

                        Connection Errors

                    Last 10        Last 60        Last 24

                    Minutes        Minutes        Hours

    5005076802264af4            3            23            23

    5001738000330171            0            0             0

    5001738000330173            0            0             0

    5005076802364af3            0            0             0

    5005076802364af4            0            0             0

    5005076802264af3            0            0             0
```

Adapter Driver: fscsi1 -> AIX PCM

Adapter WWPN: 10000000c957167d

Link State: Up

Connection Errors

Last 10 Minutes: 2

Last 60 Minutes: 22

Last 24 Hours: 22

Total Errors: 22

Connection Errors

	Last 10 Minutes	Last 60 Minutes	Last 24 Hours
5005076802364af3	0	0	0
5005076802364af4	0	0	0
5005076802264af4	2	22	22
5005076802264af3	0	0	0
5001738000330171	0	0	0
5001738000330173	0	0	0

Note that the error count here occurs only on one particular remote port. Both **fscsi0** and **fscsi1** are able to use some remote ports without errors (or with very few errors) while attempts to access one particular storage device port result in many errors. Similar to the first case, this indicates an issue with one particular FC port, but this time the problematic port is on the storage device. The FC jammer was connected to the fiber from the storage device port to the FC switch.

Also, note that the lsmpio command accepts a -z flag to zero out the error counters. Running the lsmpio -az command resets all of the above counters to zero. This may be useful for determining if a problem is still occurring or if a problem has been resolved. The SAN administrator could make a hardware change, run lsmpio -az to zero out the error counters, then wait a few minutes before running lsmpio -are again. If the error count is zero, then the problem is likely resolved.

Conclusion

The recent technology levels of the AIX operating system contain improvements in the AIX MPIO software, including improvements to the lsmpio command. This article discussed those improvements and demonstrated how a SAN administrator can use the new features of the lsmpio command during problem determination.

Related topics

- IBM AIX MPIO: Best practices and considerations
- IBM AIX Knowledge Center: lsmpio command

Posted - Wed, May 30, 2018 10:46 AM. This article has been viewed 13841 times.

Online URL: http://kb.ictbanking.net/article.php?id=86