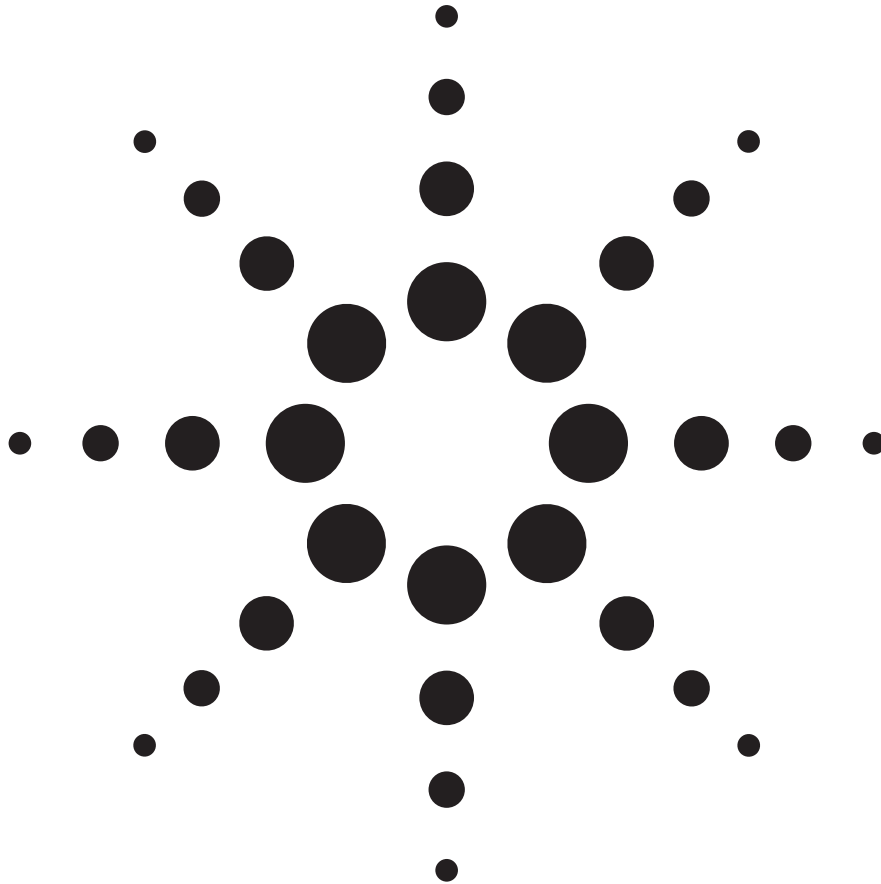


Storage Extension Test Challenges

Testing today's network devices for
geographically distributed storage

White Paper



Agilent Technologies

Introduction

Geographically distributed storage networks have introduced a number of new network devices and new technologies into an already complex system, that is the storage network. These new network devices are more complex and have more intrusive features, compared to a traditional Fibre Channel switch. These new storage network extension devices are challenging the traditional method of testing for storage networks; usually based on real servers, storage and backup devices. This article will discuss the limitations of the traditional test methodology, and why a new way of testing these devices is required.

The challenges of the traditional test methodology

With the emergence of this new generation of storage network devices, the limitations of the traditional test method, using real servers and storage, become more visible, especially in the areas of performance and robustness testing. These limitations are due to the fact that real servers and storage are not designed for test.

There are usually multiple software layers in the device that (application, operating systems and drivers) generate levels of indeterminism on the traffic injected in the storage networks, therefore reducing the repeatability of measurements.

Also, real server and storage equipments are designed to conform to protocol specifications, testing the robustness of the network by simulating disasters, faulty devices, injecting errors and protocol violations on real devices has always been a challenge. When errors do happen, the real devices typically do not provide the ability to drill down into the details to understand the root cause of the issues.

Lastly, as storage networks move to higher speeds, such as 4 Gbs and 10 Gbs, individual real devices may not load the network at 100% of its capacity, making maximum performance measurement even more difficult.

In the next paragraphs, we use several test cases for extension network devices to highlight the challenges associated with the traditional test methodology, and how they can be overcome.

Test Case 1 – Credit Characterization

As a part of the initialization process, a transmit buffer-to-buffer-credits (BBC) is determined between end devices. The sending device can only send if there are transmit BBC available. After sending a frame, the source device decrements its transmit BBC. When the destination device receives a frame it sends back a credit (R_RDY) to indicate that it has processed the received frame and is ready to receive more frames. Upon receiving the R_RDY, the source device will increment the transmit BBC. The distance extension network devices in this scenario pass the signals through transparently.

Due to the speed of light delays, across long distances there could be idle periods between the transmissions of frames as end devices wait for the transfer of the frame and credit. This is known as “data droop”, where the maximum throughput can not be maintained on the extension link.

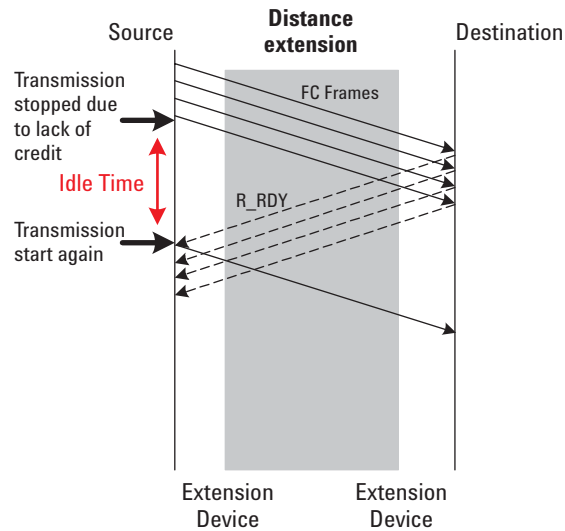


Figure 1. Flow control between end devices

One way to compensate for drooping is for the end devices to support larger BBCs. Assuming the transmission delay is the only delay in the system, then, with a BBC of 256, it is possible to extend a 1 Gbps link 500 km without suffering from credit starvation. Theoretical extension limits for other BBC values are shown in Figure 2.

However, as extension devices introduce their own delays and have different performance characteristics, the theoretical limits generally cannot be reached. Thus, before deployment, it is important to validate that the expected throughput values can actually be achieved.

To do this validation using real devices is difficult because real devices typically have a fixed BBC value, which is dependent on its design. So to test the actual extension link throughput, with different BBC values, at different transmission rates, requires many real devices to be used in turn.

On the other hand, test platforms work on the principle of emulating real devices. Emulation means the test platform supports enough of the protocol to convince the extension device under test that it a real server or storage device. For this test case, the test platform will emulate different devices by changing the BBC. Having this flexibility, means the optimum credit value can be found easily, and more importantly, can be tested using just one piece of equipment.

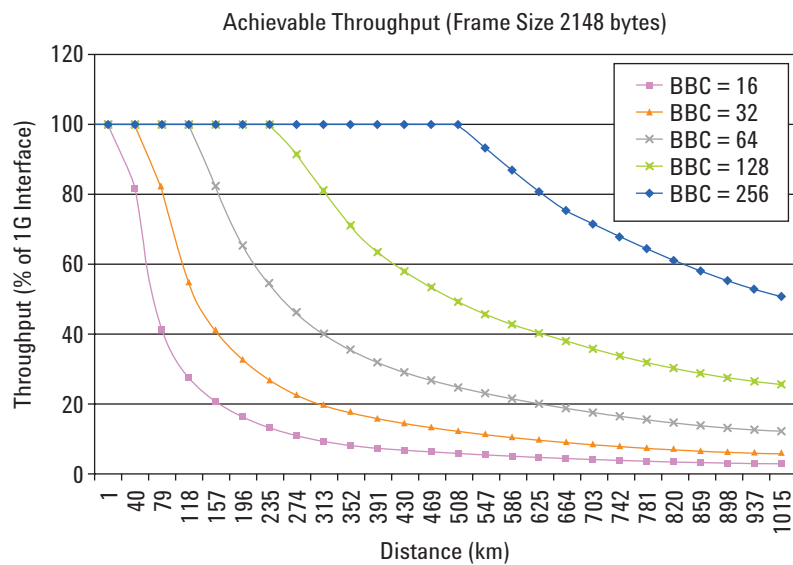


Figure 2. Theoretical extension limits

Test Case 2 – BBC spoofing

Another problem with testing using real devices is the difficulty in generating specific scenarios or simulating negative test cases. This can be demonstrated by the Buffer to Buffer Credit (BBC) spoofing test.

BBC Spoofing is a new feature offered by some extension devices. With BBC spoofing the extension device takes over the flow control between the end devices. When end devices send frames to each other, it is intercepted by the extension device and buffered. The extension device at the ingress will respond to the source device with a R_RDY as soon as a frame is received, and will transmit the frame across the link to the replicated extension device on the remote side. At the egress, the extension device will send frames to the destination device only at the rate which the destination device can accept frames.

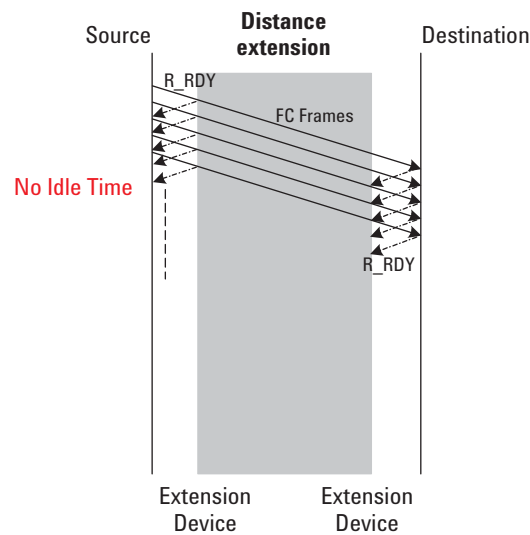


Figure 3. Flow control with BBC spoofing

With BBC spoofing, the source device is not waiting for the destination device to respond with R_RDYs. Thus a higher level of utilization is maintained on the extension link.

With the BBC spoofing feature, extension devices maintain the flow control. If congestion occurs due to a slow destination device, it is important to understand the implications. Are frames lost because buffers in the extension device overflow? Or, does the extension device throttle the sending device?

To create congestion in the extension link can be very difficult using real servers and storage systems. Application level performances and SCSI protocol handshaking can mean that the actual link utilization is lower than expected. Thus, a number of servers and storage systems may be required to generate concurrent load to congest the link, this is especially true as the link speeds increase. Each of the servers and storages systems must be controlled separately. This method can be very expensive in both management time required and equipment costs. Another problem with this approach is that the servers and storage devices will not be able to help drill down into the issue. For example, if frames are lost due to the extension device, using a real server and a real storage device, this will appear on the server or the storage device as an aborted operation, but is the aborted operation due to a frame loss, or due to some other application specific issue? This can be very difficult to identify.

In contrast, a test platform is a hardware test system that is tailored for flexibility, and thus able to recreate test scenarios. The test platform does not have issues with application level performance limitations – the application on the test platform is also designed for testing. So the dedicated test platform has the ability to easily create both stress conditions for the fabric, as well as the ability to inject negative test scenarios. This capability is very important in the testing of storage network devices; a network device can only be considered tested if its robustness has been validated under stress and invalid network scenarios.

Another advantage that a dedicated test platform brings in this case is the ability to drill down into the cause of the issue. Using the same scenario, if the extension device is causing frame loss due to congestion, the test platform will be able to highlight this issue. The test platform will track the number of frames sent and the number of frame received, and thus able to determine that frames have been lost. Even more importantly, the test platform can be configured to recreate the congestion scenario on demand.

Test Case 3 – Failover recovery

Geographically distributed storage networks use a number of different technologies. The technologies include SONET/SDH, WDM and FCIP. Regardless of the type of network used for extension, failures can always occur. If FCIP is used, this could be in the form of one of the IP routers being overloaded and causing IP encapsulated FC frames to be dropped. In the case of SONET/SDH or DWDM, one of the links on the ring could be broken. These failures in the extension link can be the cause of longer latency or in the worst-case application timeouts. Thus, understanding implications for the storage application if failure occurs is a test that should be performed.

Each of the three networks (WDM, SONET, and IP) treats failures differently, and the expected network down time maybe very different as well.

To use real devices to calculate failure recovery time would be some variant of the “stopwatch” method. An extension network is setup, with I/O transfer started between a server and a storage device. Then the user creates a failure condition in the extension network and starts timing on the stopwatch. During the failure, the I/O will stop or slow between the server and the storage device. The test engineer monitors the system until I/O resumes to the original level and stops the stopwatch and calculates the failover recovery time.

The manual test method is very labor intensive. Also, human reaction times are highly variable and prone to errors. This type of analysis may work for systems with very long failover times; however, it is no longer sufficient in some of the new technologies introduced for distributed networks. For example, in a SONET/SDH network, which has Automatic Protection System (APS), the failover recovery time is guaranteed to be within 50 ms. In order to test this guarantee, an automated test method with accurate hardware driven timers is required.

A dedicated test platform will typically provide an automation interface that allows the user to control all aspects of its behavior; from changing behavior parameters, to I/O traffic profiles, to statistics collection. This means that repetitive tasks and time sensitive tasks can be automated. In the case of the failover testing, a test platform can be used to automatically start timers when the failover event starts, and when the I/O resumes automatically stop the hardware based timers; thus ensuring accurate and repeatable measurements.

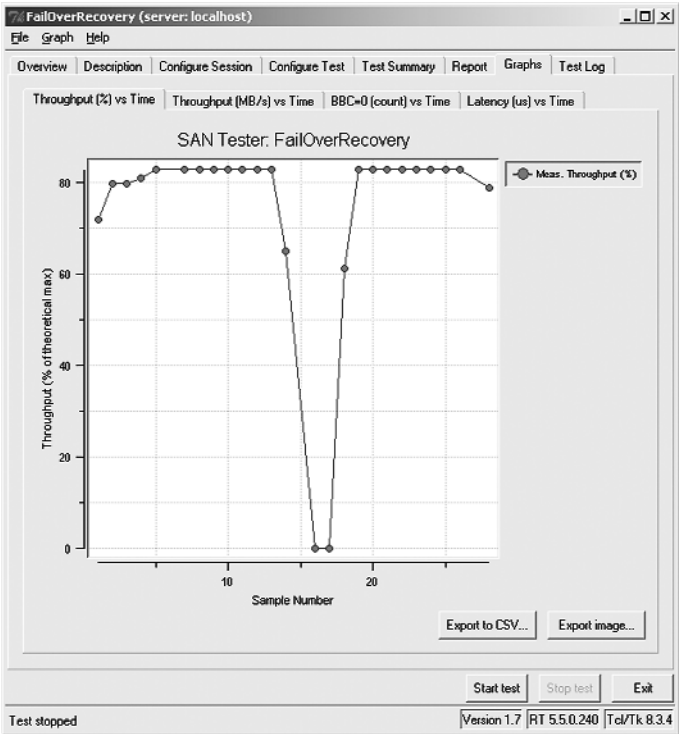


Figure 4. Hardware based failover recovery testing

Limitations of using test platforms

The challenges of testing distance extension devices with the traditional test methodology have been highlighted in this paper. A dedicated platform is required to fully test storage network devices as they become more complex. These test platforms provide flexibility in emulating multiple devices; ability to create stress, generate negative test scenarios, and simulate diverse network conditions. Also, with many automation features built into the test platform, testing can be simplified, and test time significantly reduced.

Given the number of advantages of using a test platform over real devices, and the ease of use of test platforms over managing multiple real devices, it is tempting to replace all real device test beds with test platforms. At the moment, this is not feasible, due to the need for interoperability testing with real devices. A distinction should be made between different types of test scenarios. Real devices should be used for interoperability test cases, but dedicated test platforms should be leveraged for device or network characterization and feature testing scenarios.

Another area to be aware of is the application level behavior. Each application has very different behaviors and requirements. It is not possible for the test platform to completely replicate the behavior of a specific storage application. Thus it is very important to verify the application level behavior on the distributed storage network before it is actually deployed.

However, if used correctly, a dedicated test platform can be a very powerful tool in testing storage network devices, and qualifying network behavior.

Selecting a test platform

This paper talked about the test platform in general; however, there are a number of different test platforms available on the market. We will conclude our discussions with some thoughts on what to look for in selecting a test platform.

Obviously, the test platforms should be compared and contrasted to see how they perform in the areas of performance and robustness testing, as well as automation and usability. The ability of the test platform to cover a large percentage of the test cases in each of these areas is important.

As the storage networks move to higher speeds, such as 4 Gbps and 8 Gbps, does the test platform follow the speed curve? Are the interfaces flexible, so a single interface can cover not only the new interface speeds, but also be configurable to test the existing 1 Gbps and 2 Gbps interfaces?

How much flexibility does the test platform provide to allow the user to create negative network scenarios to fully test the response of the network device under adverse conditions? Does the test platform count or record the number of errors seen?

Also, the test platform should be assessed on its automation interface. How easy is the interface to use or to build internal automation scripts? Does the test platform already provided some pre-designed automation scripts that will allow you to get a jump start?

Lastly, the vendor's commitment to continuous improvement of the product is very important. What is the vendor's plan for growing the product in the next year, in terms of new features or new technology coverage? These are all important considerations in the final selection.

Agilent Technologies' Test and Measurement Support, Services, and Assistance

Agilent Technologies aims to maximize the value you receive, while minimizing your risk and problems. We strive to ensure that you get the test and measurement capabilities you paid for and obtain the support you need. Our extensive support resources and services can help you choose the right Agilent products for your applications and apply them successfully. Every instrument and system we sell has a global warranty. Two concepts underlie Agilent's overall support policy: "Our Promise" and "Your Advantage."

Our Promise

Our Promise means your Agilent test and measurement equipment will meet its advertised performance and functionality. When you are choosing new equipment, we will help you with product information, including realistic performance specifications and practical recommendations from experienced test engineers. When you receive your new Agilent equipment, we can help verify that it works properly and help with initial product operation.

Your Advantage

Your Advantage means that Agilent offers a wide range of additional expert test and measurement services, which you can purchase according to your unique technical and business needs. Solve problems efficiently and gain a competitive edge by contracting with us for calibration, extra-cost upgrades, out-of-warranty repairs, and on-site education and training, as well as design, system integration, project management, and other professional engineering services. Experienced Agilent engineers and technicians worldwide can help you maximize your productivity, optimize the return on investment of your Agilent instruments and systems, and obtain dependable measurement accuracy for the life of those products.



Agilent Email Updates

www.agilent.com/find/emailupdates

Get the latest information on the products and applications you select.



Agilent Direct

www.agilent.com/find/agilentdirect

Quickly choose and use your test equipment solutions with confidence.

For more information on Agilent Technologies' products, applications or services, please contact your local Agilent office. The complete list is available at:

www.agilent.com/find/contactus

Phone or Fax

United States:

(tel) 800 829 4444
(fax) 800 829 4433

Canada:

(tel) 877 894 4414
(fax) 800 746 4866

China:

(tel) 800 810 0189
(fax) 800 820 2816

Europe:

(tel) 31 20 547 2111

Japan:

(tel) (81) 426 56 7832
(fax) (81) 426 56 7840

Korea:

(tel) (080) 769 0800
(fax) (080) 769 0900

Latin America:

(tel) (305) 269 7500

Taiwan:

(tel) 0800 047 866
(fax) 0800 286 331

Other Asia Pacific Countries:

(tel) (65) 6375 8100
(fax) (65) 6755 0042
Email: tm_ap@agilent.com

Contacts revised: 05/27/05

Product specifications and descriptions in this document subject to change without notice.

© Agilent Technologies, Inc. 2006
Printed in USA, January 19, 2006
5989-4068EN

