Demartek Best Practices for SPECSFS Testing

Suzanne Stone

Demartek

SNIA Emerald™ Training

SNIA Emerald™ Power Efficiency Measurement Specification

Version 3.0

February-March 2018

Demartek Background

- Test lab specializing in real-world, hands-on research and analysis of data center technologies
- ISO 17025 accredited test lab
- EPA-recognized test lab for ENERGY STAR Data Center Storage testing
- SNIA Emerald Recognized Tester
- Website: www.demartek.com/TestLab

Real-world, Hands-on Research & Analysis

Demartek

SPECSFS Deployment

- SPECSFS should be deployed on 3 distinct networks:

  1. Control Network: Completely isolated network.
     - Your DNS/AD, each load client, and each prime client should have a connection to this network.
     - Provision a separate (non-production) DNS/AD and if possible make sure it is highly available.
     - Make sure the prime FQDN is in the DNS/AD for forward and reverse lookups.

SPECSFS Deployment

- SPECSFS should be deployed on 3 distinct networks:

  2. Storage Network: Completely isolated network. The storage and each load client should have a connection to this network.

  3. Production Network: Designate 1 server that will live on both your production and isolated control network to give you access to your environment, accessible via RDP or SSH. This should be the only connection your environment has to non-test networks.

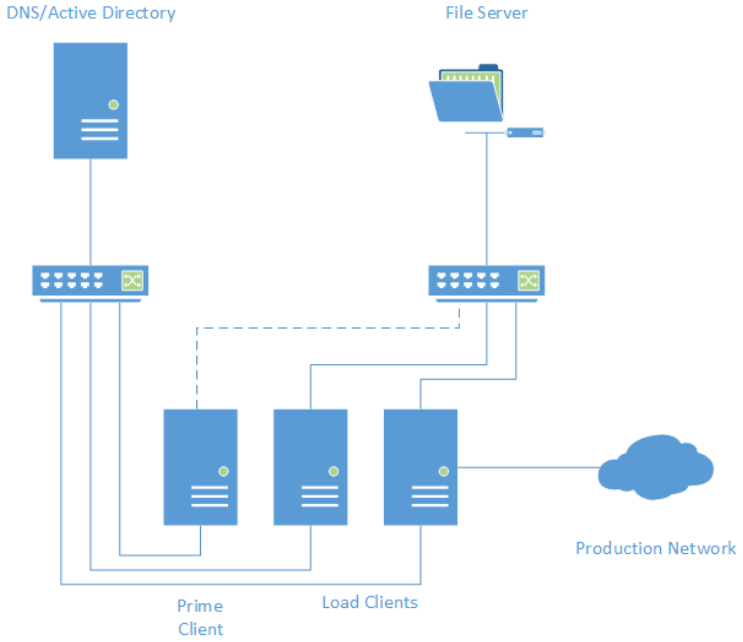
SPECSFS Deployment

- The storage should only be on the storage network, and each load-client should have 2 connections – one on the storage network and one on the control network.

SPECSFS Deployment

- If prime is not a load client (applicable to Windows environments), then only give it access to the control network.
  - Make sure subnetting/routing is set up correctly.
  - One machine should be on the production network and control network giving access to your environment.

SPECSFS Deployment



SPECSFS Deployment

- For AD/DNS, make it a completely separate domain, no delegations or access to the production network.
  - Due to not starting with 3 separate networks, our Active Directory/DNS got messed up. After much troubleshooting, we had to deploy a new AD/DNS and connect all servers to it.
  - Takeaway: Deploy on 2 isolated networks for starters to avoid messing something up, and back up your AD/DNS after initial deployment.

SPECSFS Deployment

- We eventually switched to a Linux environment, and chose CentOS. However, CentOS 7.4 DNS configuration is completely different from RHEL.
- For your Linux DNS, consider using a more standard flavor of Linux where you are familiar with the DNS setup.

Virtualization

- Virtualization is not recommended. We would recommend a bare-metal deployment of Windows for testing of SMB fileshares.
  - Initially, we used Hyper-V VMs and experienced constant networking issues. Isolating the control network to an internal (no actual NIC) network between VMs did not solve the problem.
  - An alternative is VMware VMs instead of Hyper-V. Presumably there are fewer roadblocks here than with other hypervisors.

Virtualization

- A lot of strange errors disappeared when we ran one big bare-metal load client.
  - Before we had file-lock errors that killed a test, though SPECSFS does not actually use file-locks. Again, switching to bare-metal with no virtual switches fixed this problem, but the performance suffered.
  - It may still take multiple servers or additional tweaks to make sure you are getting the most out of your storage with bare-metal deployments.

Size of Environment

- We found that we were getting DNS failures, and a quick look at CPU showed that the loadclients (VMs at the time) were maxing out all hypervisor CPU.
  - Adding another server and migrating some VMs over to it reduced the load on the first hypervisor and eliminated the problem.
  - Takeaway: If the loadclients don't have enough resources, you may get hard-to-track down networking errors.

Size of Environment

- While considered a “workaround” for the networking problem, provisioning more resources for your loadclients also makes sure that storage is causing the bottleneck, not your loadclient resources.

- Follow tuning directions (registry hack for Windows max ports, etc.)

SFLOW

- Most GUI based SFLOW collectors are built to help a systems administrator get a general picture of what is going on in his network and is not geared towards taking precise measurements at precise intervals.
- We investigated the following sFlow collector software options for potential use, and found them to be lacking…

Inadequate SFLOW Collectors

- **SFLOWTREND PRO version 6.7:** Employs averaging, seems like smallest interval is 1 minute for consecutive data points. It also will list the total throughput received during an interval, instead of dividing by seconds, possibly causing confusion and errors.
- **Netflow/PRTG Network Monitor Freeware:** 100 sensors may be used for free after your 30-day trial is completed. Minimum polling interval for a sensor is 30 seconds. Only accepts flow SFLOW sample packets, not counter or extended SFLOW format packets. Must use SFLOW version 5, with “raw packet header” format.

Inadequate SFLOW Collectors

- **nTOP nProbe:** The free trial runs for 10 minutes in enterprise mode, after which you need to pay a fee. Enterprise mode is needed to run reports. A little difficult to learn: nprobe is used to set up a collector, which then forwards it to ntopng which will then send it to a browser-based interface. Daily, weekly, or monthly usage reports can be run, so it probably doesn't have the needed granularity, although I can't confirm because I never got the sensors to register any data.

Tested SFLOW Collector

- **Sflow Toolkit:** also available from Inmon; it receives the packet, parses according to each field and outputs to screen.
  - If you pipe it to a script to parse the data, or even just to an output file to do some simple greps on later, you have control over how your throughput is calculated without wondering what was averaged where.

Tested SFLOW Collector

SNIA GSI | GREEN STORAGE

Sflow Toolkit (cont'd)

- ifInOctets is the total bytes that have come into the port, ifOutOctets is total bytes that have come out of the port, unixSecondsUTC is your timestamp.
- To get throughput, take 2 packets and calculate:

$$\text{Throughput (units B/s)} = \frac{(\text{latest\_datagram\_octets} - \text{previous\_datagram\_octets})}{(\text{latest\_datagram\_timestamp} - \text{previous\_datagram\_timestamp})}$$

Tested SFLOW Collector

- **Sflow Toolkit (cont'd)**

  We chose Sflowtoolkit as it gave us the most control with the least effort. Writing a script or using grep or findstr to extract the data you need from its output is pretty easy.

Clock Pitfalls

- SFLOW packets have a Unix timestamp datafield in them. This time comes from the switch, so you need to synch time between your switch and the rest of your environment (and don't forget your time zone…)
  - To double check the Unix timestamp of our sFlow packets, we loaded them into column A of an Excel file and then used the following formula (This formula is for row 2, once input, copy and paste it for the rest):

    =(((A2/60)/60)/24)+DATE(1970,1,1)-(7/24)

    (The 7 in 7/24 was our hours from UTC before the DST timechange, you will want to adjust for your own location and time of year.)

Multiple Flow Pitfalls

- For our calculations with Sflow Toolkit, we had the switch generate counter SFLOW packets only. However, flow and extended flow SFLOW packets are also available.
  - If you are sending both types of SFLOW from your switch, you may complicate your parsing. (Each datagram has a field and sampleType, so make sure you are using all datagrams that list their sampleType as COUNTERSAMPLE if you employ our method of throughput calculation.)

Multiple Flow Pitfalls

- We simplified things by only collecting SFLOW from 1 port (the one directly connected to the storage. Our storage only used 1 port with a 40GbE NIC).
- We recommend only 1 switch port per SFLOW collector IP address, or you will have to figure out which switch port your SFLOW packet came from and calculating the throughput will be more complicated.
- Set up other SFLOW collectors at other IP addresses (or other ports), to collect any SFLOW from additional ports.

Q&A

SNIA GSI | GREEN STORAGE

Questions?

Contact

SNIA GSI | GREEN STORAGE

Demartek®

Demartek public projects and materials are announced on a variety of social media outlets. Follow us on any of the above.

www.demartek.com/TestLab