StorPool Pre-installation Checklist

2019-01-09

The following list specifies the configuration needed on each server or switch before we proceed with installing StorPool. Most of these settings are considered a good practice, even without StorPool, and a few of them are StorPool-specific.

1. Hardware compatibility

Check that the hardware in the server is compatible with StorPool. Check the "StorPool System Requirements" document. The latest HCL can be downloaded from here.

If the specific models of NICs, HBAs, SSDs have not been tested previously for compatibility, then we need to first make sure that they are compatible, by performing some additional tests on them.

2. Network Switches

StorPool may use a single tagged VLAN on the two separate networks for the storage traffic and the management access (API). As an addition to the explanation below, see the following diagram:
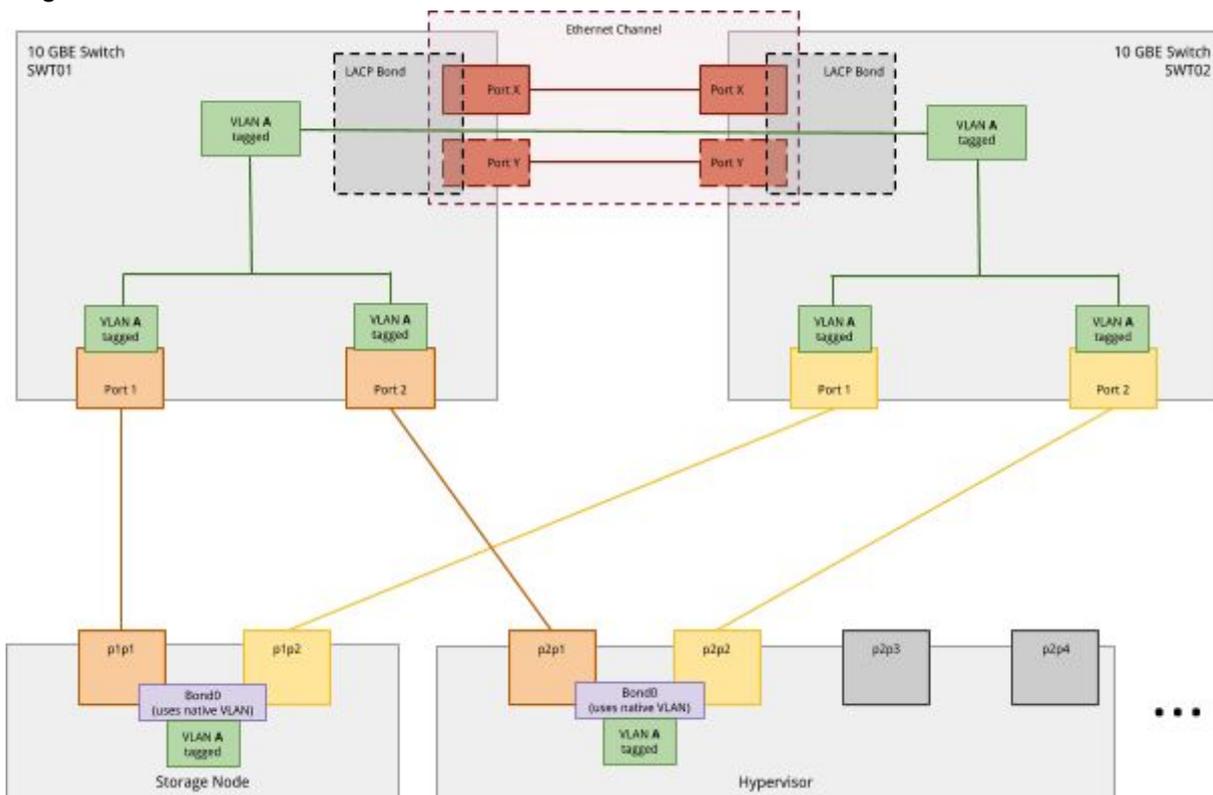
- **Storage network (tagged VLAN A)**
  StorPool uses one VLAN, exposed to two ports on each storage node and each host for redundancy and load balancing.
  The physical network is implemented as two 10Gbe (or faster) switches. Each host and each server have a port connected to each switch.
  The storage network is implemented as a VLAN over the two corresponding switches.

The VLAN is configured as tagged on the relevant ports and is also tagged on the inter-switch connection.
For all ports used by storage servers or initiators (clients) configure the following settings on the storage switches:
- ◦ Jumbo Frames (9000 bytes)
- ◦ Flow control RX disabled / off
- ◦ Flow control TX enabled / on
- ◦ Spanning-tree disabled / off

- **API network (tagged VLAN A)**
  The API network is used for provisioning and monitoring commands to StorPool. For example it is used by the cloud management system to create and attach StorPool volumes. StorPool maintains a floating IP address on this network for the StorPool API server (storpool_mgmt service).

  The API network interfaces and corresponding ports on the hosts should be configured in bonding for high availability. The recommended setup is with an active-backup bond interface using arp-monitoring with addresses from the same network.

  Alternatively the API could be running in a separate VLAN or to be configured on a different set of ports of the storage nodes. It would typically use private addresses.

3. BIOS and Firmware updates

Check if there are BIOS and firmware updates available from the hardware vendor. These might include firmware for the HBAs/RAID controllers, NICs, BIOS, IPMI module, CPU microcode updates, etc.

Please consult the [StorPool System Requirements](#) document for recommended firmware revisions, note that sometimes the newest available revision might not be the recommended one.

Specifically running HBAs, RAID controllers and SSDs with older firmware has caused problems previously. Specific models of hardware also have a recommended maximum version due to bugs we've seen with newer versions.

4. BIOS settings

- General operating profile, CPU frequency select,Memory speed and Performance Bias settings should be set to "Maximum performance".
- All C-states are usually enabled (with at least C1 and C1E enabled).
- Turbo frequency should be enabled
- Intel NICs require SR-IOV to be enabled for hardware acceleration in 18.02+ releases
- Node Interleaving should be either disabled or set to NUMA.
- Channel interleaving should be set to enabled.
- FakeRAID (Intel RST) must be disabled and drives that are connected to the motherboard must be in AHCI mode

5. Memtest

The best way to test ECC memory is to boot from a *memtest86+* ISO (http://www.memtest.org/#downiso). Note that by default the ECC tests are disabled and it is highly recommended to turn them on.

Another option for testing the memory is to use *memtester* and test only the memory not used by the running kernel on the node. Please note that the tested amount of memory must be 300 - 500 MB less than the actual free memory, otherwise the OOM killer may be triggered. If swap is configured, stop it during the test with *swapoff -a.* For dual socket servers *sysctl kernel.numa_balancing=0* will be needed during the test to not interfere with memtester.

*** Important note:** The *mcelogd* service must be installed and configured before using the user-space tool "memtester".

6. Burn-in test

The server must undergo a burn-in test.
- SSDs must be fully overwritten at least once.
- Hard drives must be randomly written for a couple of hours.
- All drives must pass a full (known also as long) smart test and should not report any errors from S.M.A.R.T. after the burn-in tests.

7. OS installation

The servers must be pre-installed with the Linux distribution which will be used in production. The supported distributions are:
- CentOS 6.x, CentOS 7.x (prefered)
- Debian 8 Jessie, Debian 9 Stretch
- Ubuntu 16.04 LTS, Ubuntu 18.04 LTS

StorPool recommends the use of the latest stable version of *CentOS 7*. If you are using another Linux distribution (e.g. RHEL, OEL, SuSE) just let us know. We can support StorPool on all Linux distributions with good build and packaging systems. The supported architectures are:
- x86_64 (amd64)
- ARM64
- POWER 8/9

During the OS installation:
- Disable swap
- Don't use fakeraid such as Intel Matrix RAID. If you have to use RAID, use regular MDRAID (software RAID) instead.
- On storage nodes leave the drives for the StorPool storage system unpartitioned.

- If NVMe drives are to be used for StorPool, leave them unpartitioned

8. Drives

Decide which drives and partitions will be used for StorPool. Send us a list with their `/dev/disk/by-id/scsi-*` names. Please note that any data stored on these drives/partitions will be erased.

Make sure all drives intended to be used as StorPool drives are in JBOD mode. In case the HBA or RAID controller does not support JBOD the best workaround is to export every disk as a RAID0 volume.

If used, Avago(LSI) Megaraid controllers (e.g. 3108, 9261, Dell H730) should be configured as follows:
- Each SSD must be presented as JBOD
- Each HDD must be in Unconfigured Good state. StorPool will configure the RAID controller during installation.

9. Partitions & Alignment

To be able to collect its reports, the StorPool needs at least 16G free space in */var/spool/storpool*.

In case some of the StorPool drives will be on a partition, check that the partition is 4K aligned. Run blockdev --report and check that *StartSec* is divisible by 8 and that *Size* is divisible by 4096.

SATA and SAS SSDs and HDDs can be used for the operating system and StorPool on separate partitions. NVMe SSDs must be fully dedicated to StorPool.

10. OnApp Hypervisors & Backup Servers

All OnApp hypervisors and the backup servers (if present), must have access to the storage networks and to the API network. The StorPool initiator (storpool_beacon, storpool_block, StorPool CLI and integration with OnApp) must be installed and configured on them by the StorPool team.

11. Hand over for installation

After performing these steps please hand the server over to StorPool, by providing:
- SSH remote access details to the nodes - IP addresses, ports
- IPMI access to the nodes
- Access to the switch/switches (optional, but highly recommended, as it will be of great help for debugging network issues as well as much faster setup)
- IP addresses for the API network - one IP address per storage node, one address per API user (hypervisor, etc), one IP address for floating IP of storpool_mgmt service
- List of drives from /dev/disk/by-id/ to install StorPool