BE Project Review-1

Group ID : 21

Guide: Mr. Vishal Jaiswal

Group Members:

1.	Bharat Kothari	43110
2.	Animesh Landge	43132
3.	Shrijan Vats	43258
4.	Amod Dhopavkar	43304

BE Project

Sponsored by Veritas LLC

Title - Distributed Replicated Block Device (DRBD) Network Packet Tracing Module.

Problem Statement - Build a packet tracing module on DRBD to track the data packets transferred over a cluster in a network during data replication between multiple block devices and analyse the transfer of packets.

Purpose / Motivation of Project

- Opportunity to work on open-source project.
- First-hand industry exposure.
- Opportunity to work on an industry standard replication project.
- Learn about cluster networks
- Learn about system programming

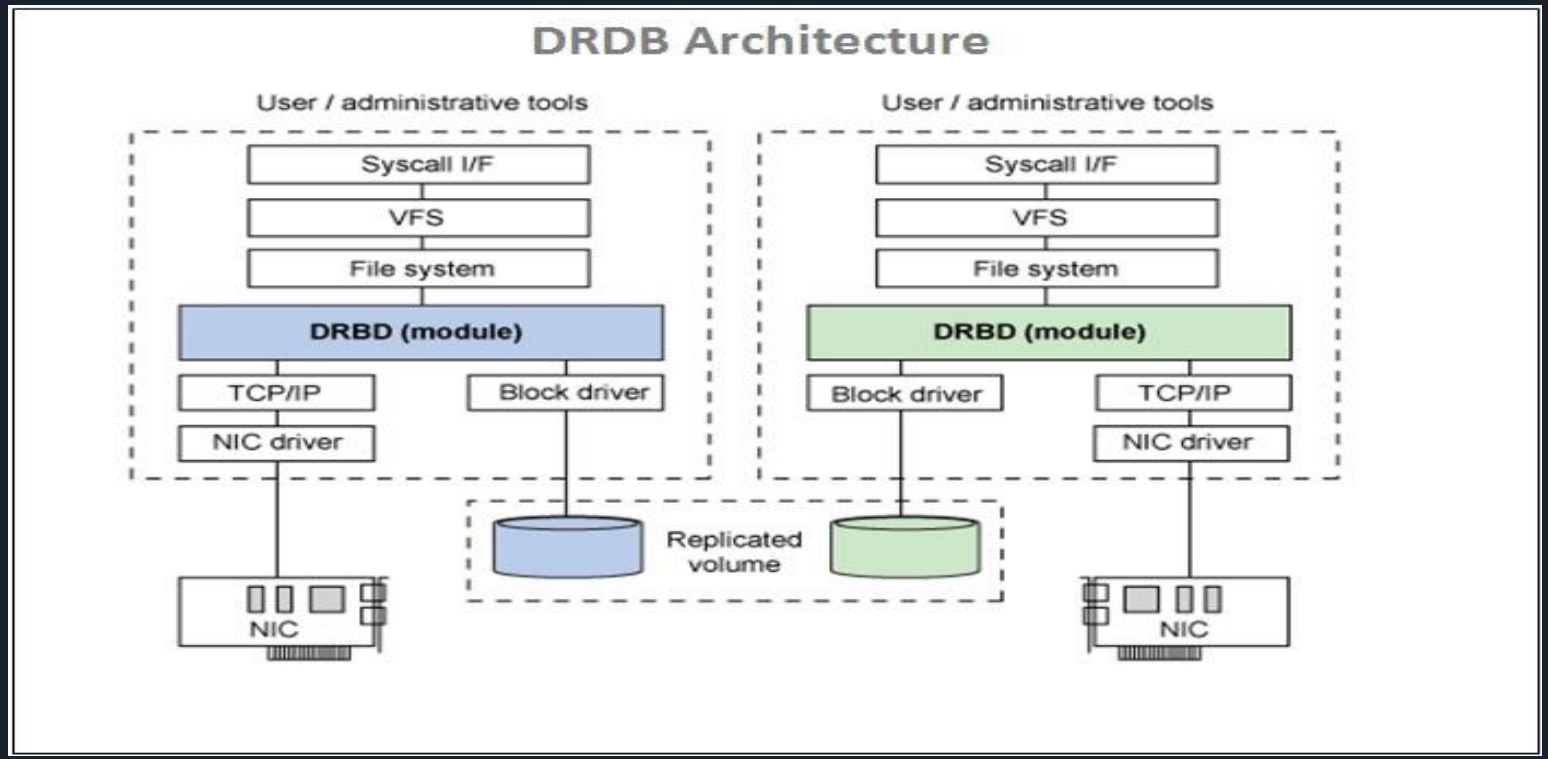
Objectives of Project

- Understand how replication in DRBD works - code flow how/where we put the packet on air in code.
- Writing kernel module to trace/keep track of logs in memory and on-disk.
- Enable tracing on the modules dynamically.
- Least or no performance impact is desirable.
- Writing utility/daemon to pull tracing information from kernel - parse the data for reporting/debugging problems
- Use container to run the analyser tool.
- Eg: plot a graph, create sequential logs combining primary and secondary logs.

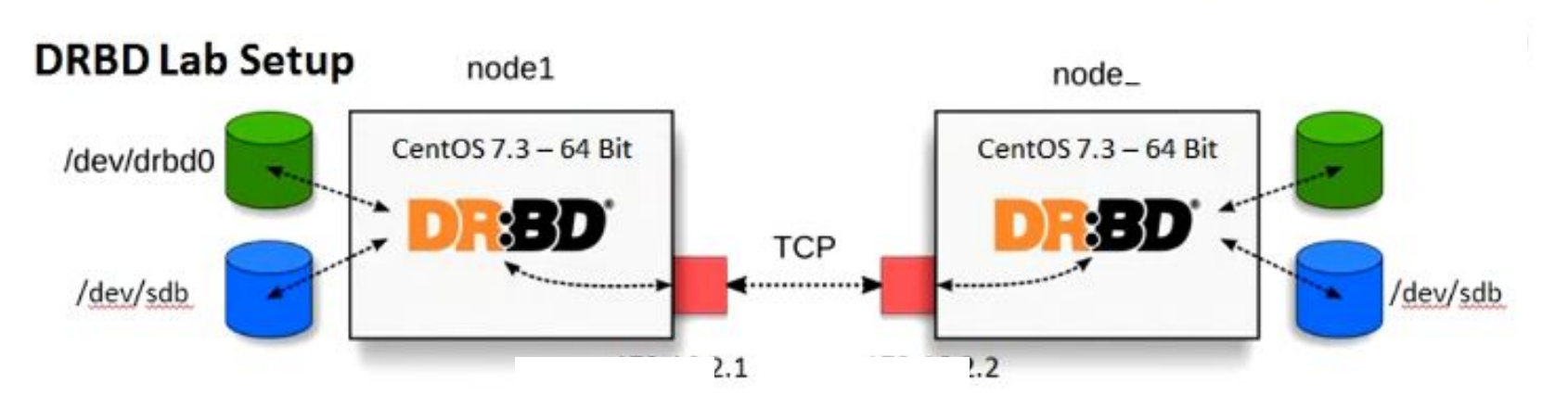
Scope of Project

- Provide an enterprise grade solution to analyze the packet transfer in replication phase at multiple data nodes.
- Calculate the accuracy of data replicated at multiple data nodes in a cluster.
- Determine the bandwidth required over a network based on the accuracy of the transfer of packets.
- Determine the number of replication nodes required in a data cluster.

DRBD Architecture diagram

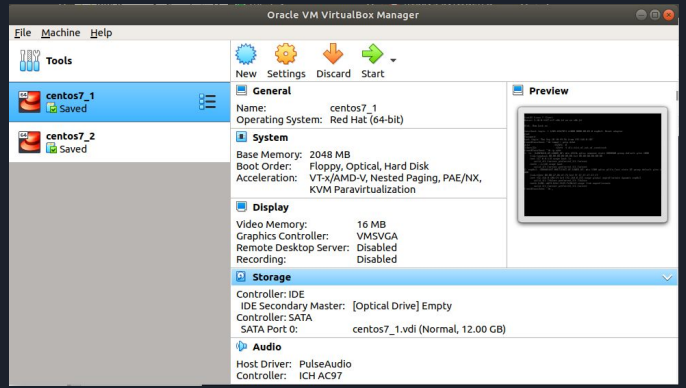


DRBD Lab setup



List Of Required Hardware / Software

- This project uses Oracle VM VirtualBox environment to simulate a cluster of DRBD computers.
- We simulate the network on Linux Machine having i5 8th gen 8 GB RAM
- The cluster contains of two VMs running centOS 7.



DRBD vs RAID

RAID is designed for two or more disks connected locally whereas
DRBD is designed to replicate a block-device over a network.

We use RAID (software or hardware) to increase the reliability of the local storage over 2 or more disks.
DRBD would sit on top of your RAID and replicate that data to another server for failover purposes.

DRBD can also allow two servers to access that data at the same time,
which you can't really do with RAID in a cluster-aware file system.

Initial Phase

1. Created a cluster of two centOS nodes on VirtualBox.
2. Installed necessary packages for running drbd modules on the machines.
3. Cloned the drbd repository on the two machines.
4. Made kernel modules from the cloned repo and added the modules to running kernel.
5. Understood the basics of makefiles and kernel modules.
6. Wrote a simple character device driver in C and added and tested the module by adding to kernel.

Human Efforts in Hours and Cost Hours

Project is still in it’s initial Learning phase.

Team members are giving 10 hr/week currently but can increase according to the requirement.

References

- LINBIT DRBD 8.4 Repo
- DRBD User Guide -
- https://developer.ibm.com/tutorials/l-drbd/
- https://en.wikipedia.org/wiki/Distributed_Replicated_Block_Device
- https://www.centos.org/
- https://mariadb.com/files/DRBD.pdf
- Installation help: https://youtu.be/zJ42PezwfSk\
- http://derekmolloy.ie/writing-a-linux-kernel-module-part-2-a-character-device/

THANK YOU !