Network Design for the Low Cost Cluster

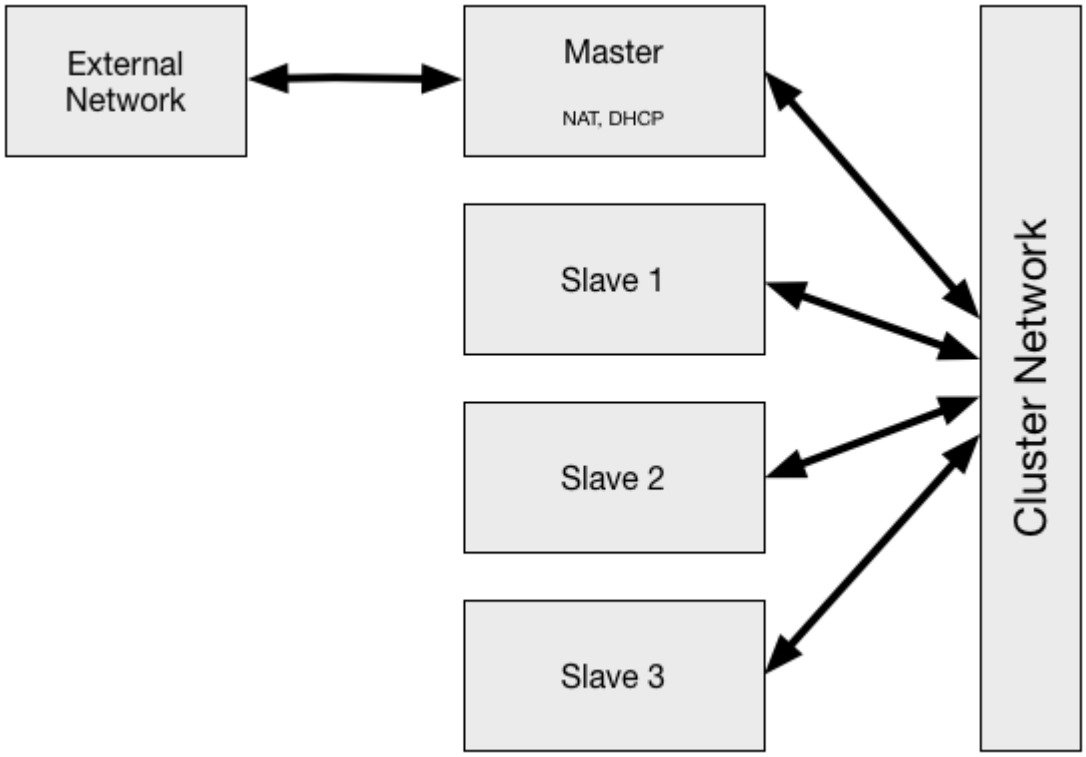
June 19, 2016 (https://diybigdata.net/2016/06/network-design-for-the-low-cost-cluster/)
michael (https://diybigdata.net/author/michael/)

Our first task in building any cluster is to first design how it will be set up, most notably how the nodes will interact with each other. The cluster we will be building will have 4 nodes, one master node and three slaves. Each node will be connected to each other by the ethernet switch. However, we want the node-to-node communication to be it's own network. This maximizes the throughput in the node-to-node communication, which is important for distributed computation, and also makes the cluster behave more like a single device to the external network. The benefit of doing this is that we can add or remove nodes to the cluster without a client ever knowing. However, this approach does present some challenges with how an external client (e.g., your laptop) will interact with the data analysis software, such as Hadoop, but we will deal with that later. My goal is really to create a "data analysis appliance", so requiring any client to understand the cluster's network topology is contrary to that goal.

To keep the cluster network isolated from the external network, a router (https://en.wikipedia.org/wiki/Router_(computing)) will need to be introduced. This could have easily been accomplished if we purchased a router rather than a switch, however, there is a significant cost difference between a router and a switch. The good news is that the master node can perform all the functions of a router if it has an additional ethernet port. For this is the reason we will purchase a USB ethernet dongle, which will be attached to the master node.

In order for the master node to perform the duties of a router, we will need to configure it to be both a DHCP server (https://en.wikipedia.org/wiki/Dynamic_Host_Configuration_Protocol) for the cluster network, and provide the NAT service (https://en.wikipedia.org/wiki/Network_address_translation)

between the external network and the cluster network. The DHCP server will assign IP addresses to each of the slave nodes, and the NAT service allows the slave nodes to reach out to the external network if needed. This overall cluster networking design is illustrated below.



Low Cost Cluster Network Diagram

It is a fairly simple design, but it is important to think these things through before building any cluster.

Low Cost Cluster (https://diybigdata.net/category/low-cost-cluster/), Set Up (https://diybigdata.net/category/low-cost-cluster/set-up/)

cluster (https://diybigdata.net/tag/cluster/), networking (https://diybigdata.net/tag/networking/)

« Building the ODROID XU4 Low Cost Cluster

Configuring the ODROID XU4 Operating System »

One thought on “Network Design for the Low Cost Cluster”

Pingback: Configuring DHCP and NAT in ODROID XU4 Cluster – DIY Big Data (http://diybigdata.net/2016/06/configuring-dhcp-and-nat-for-odroid-xu4-cluster/)

Leave a Reply

You must be logged in (https://diybigdata.net/wp-login.php?redirect_to=https%3A%2F%2Fdiybigdata.net%2F2016%2F06%2Fnetwork-design-for-the-low-cost-cluster%2F) to post a comment.

Recent Posts

Archives

January 2020 (https://diybigdata.net/2020/01/)

December 2019 (https://diybigdata.net/2019/12/)

October 2019 (https://diybigdata.net/2019/10/)

September 2019 (https://diybigdata.net/2019/09/)

November 2017 (https://diybigdata.net/2017/11/)

January 2017 (https://diybigdata.net/2017/01/)

November 2016 (https://diybigdata.net/2016/11/)

October 2016 (https://diybigdata.net/2016/10/)

September 2016 (https://diybigdata.net/2016/09/)

August 2016 (https://diybigdata.net/2016/08/)

July 2016 (https://diybigdata.net/2016/07/)

June 2016 (https://diybigdata.net/2016/06/)

Categories

Computer Science (https://diybigdata.net/category/general/computer-science/)

Data Analysis (https://diybigdata.net/category/low-cost-cluster/data-analysis/)

Data Analysis (https://diybigdata.net/category/general/data-analysis-general/)

General (https://diybigdata.net/category/general/)

Hardware (https://diybigdata.net/category/low-cost-cluster/hardware/)

Low Cost Cluster (https://diybigdata.net/category/low-cost-cluster/)

Personal Cluster (https://diybigdata.net/category/personal-cluster/)

Set Up (https://diybigdata.net/category/low-cost-cluster/set-up/)

Spark Performance (https://diybigdata.net/category/general/spark-performance/)

Uncategorized (https://diybigdata.net/category/uncategorized/)

Meta

Log in (https://diybigdata.net/wp-login.php)

Entries feed (https://diybigdata.net/feed/)

Comments feed (https://diybigdata.net/comments/feed/)

WordPress.org (https://wordpress.org/)