Home Hadoop Cluster

Hardware and Networking

Brian Baillod

Introduction

I wanted to set up a home Hadoop cluster for educational and demonstration purposes. Alex Breshears and Matt Harris provided some excellent guidance as they had already done this. I took the time to document this as I occasionally get requests from people who want to do this themselves.

Choosing Server Hardware

Capable used traditional 1U servers can be procured on EBay for less than $200 each. It takes some time to search and find the best deal. You will need 4 or more servers – minimally one management node and 3 workers.

I was looking for inexpensive servers between $100 and $200. These seem to fall into the following specification ranges:

Height: 1U or 2U
CPU Cores: 2-8 cores, I don't care too much about the speed or model. My servers have two Intel L5420 Quad Core Xeon 2.5 GHz chips
Memory: 16-48GB of RAM. These days more RAM is preferable.
Disk: I wanted some I/O parallelism so I was looking for 4 or more drives per server. 73GB or 146GB are economical. SAS or SATA is fine. I would rather have 4x146GB drives than 1x1TB drive.

Many servers are sold without hard drives or without enough RAM so purchasing and installing components yourself is a potential solution.

For reference, the servers I bought are Dell CS24-SC which was a private model only sold to Facebook. I bought them from Belmont Technology Remarketing. Here's an article about them: https://hurtigtechnologies.com/2014/06/the-definitive-guide-to-the-dell-cs24-sc-server/

Chicago Vendor:
Belmont Technology Remarketing
1300 Thorndale Ave
Elk Grove Village, IL 60007
They sell on EBay and via their website. If you live in Illinois you can pick servers up at their warehouse.
http://www.belmont-techsales.com/

Example servers available on EBay – May 2015

I think these would work well for traditional servers. The IBM servers are nice – better than mine. More drives, more memory and a DVD drive. The Sunfire servers are less powerful but have a nice price.

IBM x3650 M1, 2x X5450 3.0ghz Quad Cores, 32gb RAM, 6x146gb 15k SAS HDD, DVD-ROM $199 each

- IBM x3650 M1 2U Server
- 2x Intel Xeon X5450 3.0ghz Quad Core Processors
- 32gb RAM (8x4gb)
- ServeRAID 8k 256mb Raid Controller
- 6x146gb 15k RPM 3.5" SAS Hard Drives
- DVD-ROM Optical Drive

Sunfire X4100 M2 Server 2x AMD Opteron 2218 Dual core 2.6Ghz 32GB 4x 73GB SAS $105 each

- Form Factor : Rack-mountable 1U
- Model : X4100
- Processor Manufacturer : AMD Opteron 2218 Processor
- Type : Dual core
- Processor Speed : 2.60Ghz
- Number of Processors : 2
- Memory Type : DDR-2 SDRAM
- Memory installed : 32GB
- Hard drive : 4x 73GB SAS 10K Drive 2.5”
- Optical drive : None

New Small Form Factor Servers

A variety of new small form factor machines are available which can be used as servers.

Hadoop Cluster built using Intel NUC devices. Completely silent, draws < 100 Watts of power. 4 Cores, 4 GB RAM. Total cost $4300.

Hadoop cluster built using 5 Firefly devices that have 4GB RAM, 4 Core ARM chips, and 16GB Flash storage. $129-$199 per server. Details at:
http://en.t-firefly.com/en/firenow/firefly_rk3288/

Miscellaneous Hardware

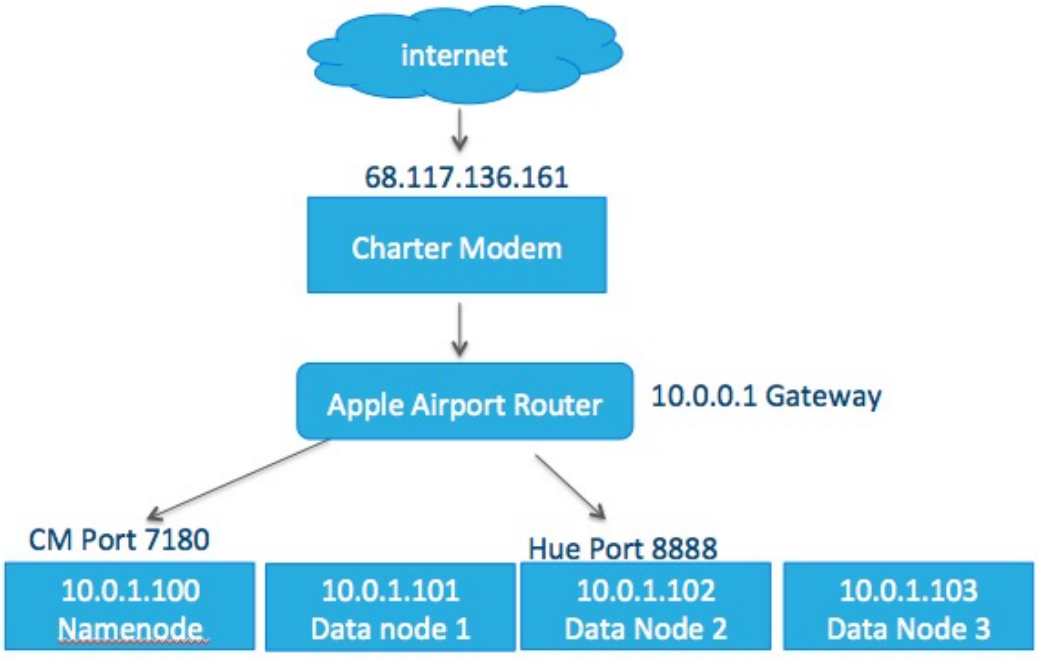
You'll need a gigabit Ethernet switch, a half dozen Ethernet cables, and a power strip/surge suppressor. You will also need occasional use of an external keyboard, mouse, and VGA monitor to plug into the individual servers.

I used a 5 port Gigabit Switch
TP-LINK TL-SG1005D 10/100/1000Mbps 5-Port Gigabit Desktop Switch, 10Gbps Capacity $14.99
http://www.amazon.com/gp/product/B000N99BBC/ref=oh_aui_search_detailpage?ie=UTF8&psc=1

Networking

I have a Charter Cable Modem providing external connectivity and an Apple Airport router which assigns IP addresses and routes the Ethernet traffic.

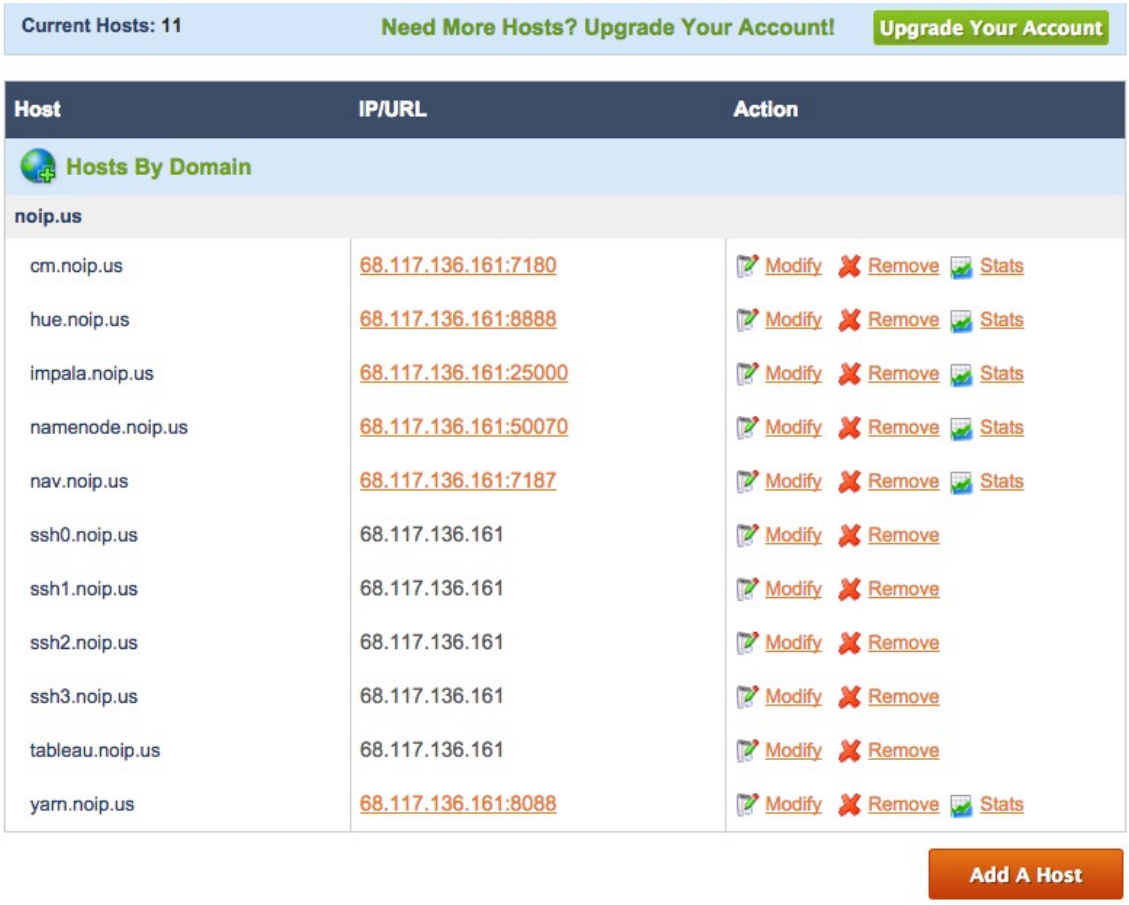
My servers have statically assigned IP addresses
Manager – 10.0.1.100
Workers – 10.0.1.101-103



I've created a companion document on how to install and configure Centos Linux on these servers that goes into more details about networking.

Dynamic DNS

I'm using a Dynamic DNS service from www.noip.com which allows you to use friendly DNS names to access your cable modem. Here are some of my DNS names that I have set up:



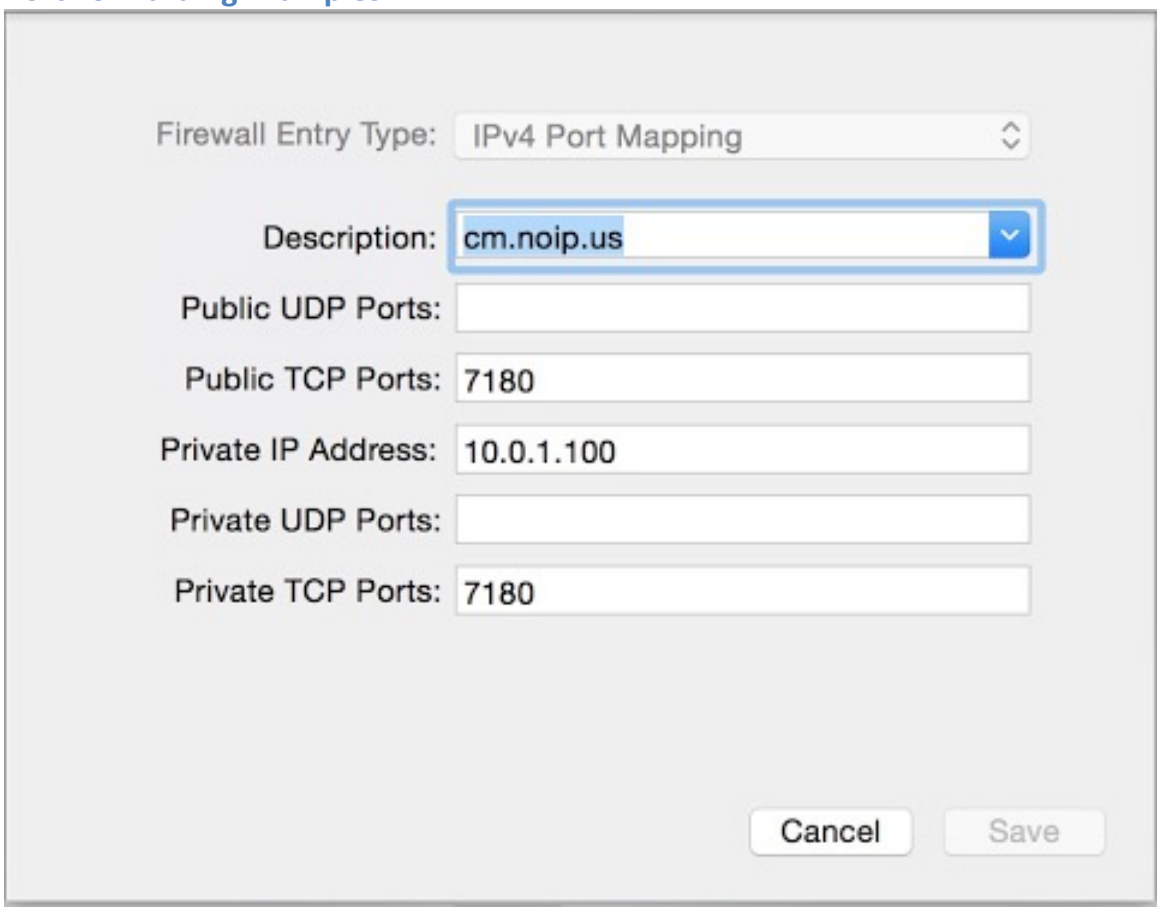
Current Hosts: 11 — Need More Hosts? Upgrade Your Account! — Upgrade Your Account

Host	IP/URL	Action
Hosts By Domain		
noip.us		
cm.noip.us	68.117.136.161:7180	Modify Remove Stats
hue.noip.us	68.117.136.161:8888	Modify Remove Stats
impala.noip.us	68.117.136.161:25000	Modify Remove Stats
namenode.noip.us	68.117.136.161:50070	Modify Remove Stats
nav.noip.us	68.117.136.161:7187	Modify Remove Stats
ssh0.noip.us	68.117.136.161	Modify Remove
ssh1.noip.us	68.117.136.161	Modify Remove
ssh2.noip.us	68.117.136.161	Modify Remove
ssh3.noip.us	68.117.136.161	Modify Remove
tableau.noip.us	68.117.136.161	Modify Remove
yarn.noip.us	68.117.136.161:8088	Modify Remove Stats

Add A Host

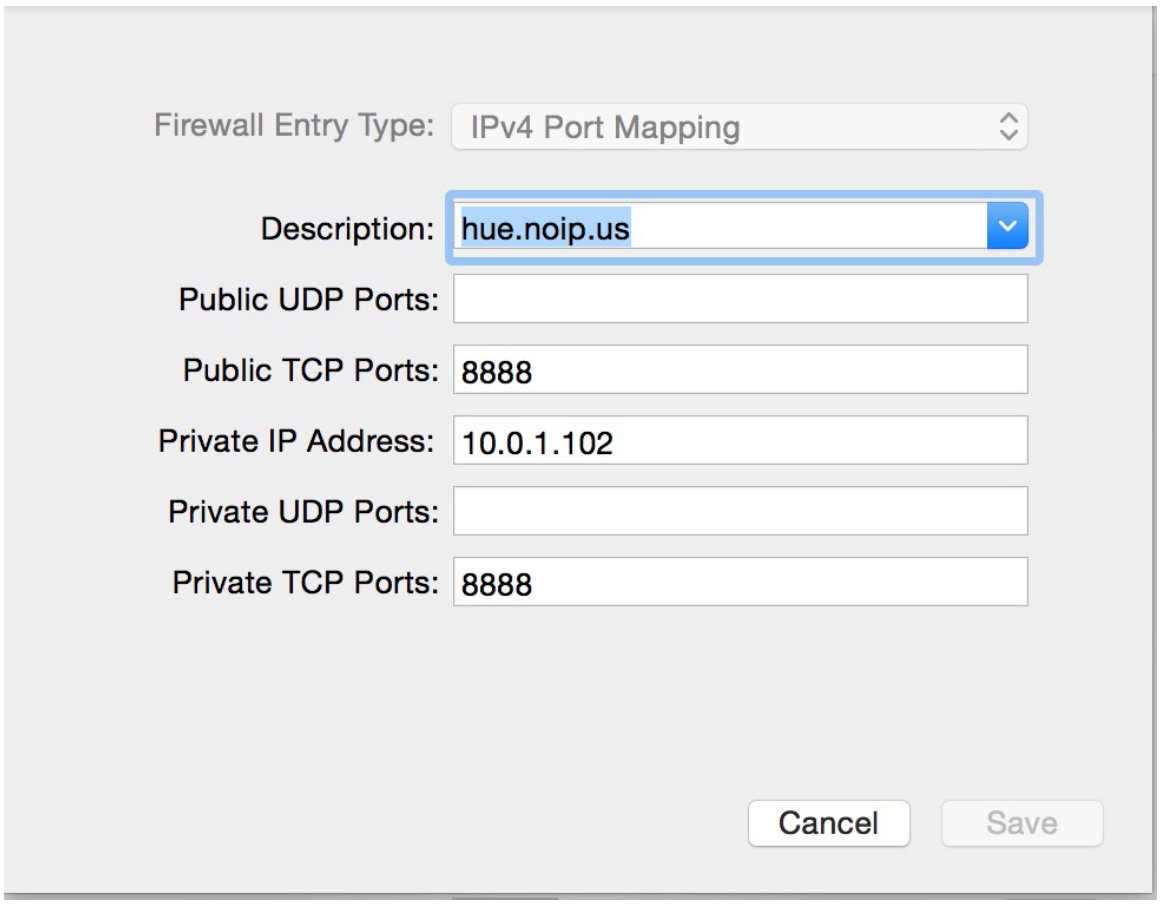
Dynamic DNS only gets the request to my router. I also need to set up port forwarding on my router to get the request forwarded to the correct server and port in my cluster.

Port Forwarding Examples:



cm.noip.us points to port 7180 on my router.

The router then forwards that request to server 10.0.1.100:7180
That is my manager node where Cloudera Manager is running.



I'm running Hue on one of the worker nodes so as to not overload the Mgt node. My router forwards hue.noip.us to port 8888 on server 10.0.1.102

Other port forwarding ideas:
nav.noip.us is navigator - forwards external port 7187 to 10.0.1.102 on port 7187

ssh0.noip.us is for ssh access

ssh root@ssh0.noip.us -p 9420 - I forward external port 9420 to port 22 on server 10.0.1.100
ssh root@ssh1.noip.us -p 9421 - I forward external port 9421 to port 22 on server 10.0.1.101

Building a Rack

I built a simple “Ikea Rack” using a table and towel bars from Ikea. You need one table and either 8 or 10 towel bars. I spaced the towel bars a little more than 3 inches apart on the inside of the legs to hold the 1U servers shown. The table is tall enough to easily fit five 1U servers. I used 1.25” wood screws and drilled pilot holes.

Table: $9.99 each
http://www.ikea.com/us/en/catalog/products/40104270/
Towel Bars: $2.99 each
http://www.ikea.com/us/en/catalog/products/50072645/