

Evaluation of Mac mini to run Hadoop Cluster - [Abhishek Parolkar](#)

Why Mac mini vs 1U-rack server?

Hadoop's promise is to run on commodity hardware while a modern 1U x86 based server is over-specified for an average computing requirement. It also uses far more power and takes up far more space in a rack than it needs to, making it costly to host. The Mac mini's small size and low power consumption & less heat dissipation makes it suitable choice. It can run Linux :)

OS X vs Linux?

Linux because footprint can be smaller and widely maintainable because of community support. People from [mythic-beasts.com](#) says that we can use linux kernel optimized for intel mac hardware.

CentOS vs Ubuntu?

Cloudera Cluster Manager runs on CentOS only so Ubuntu is not a good choice.

ThunderBolt isn't supported yet on linux and might [not be supported until mid 2012](#).

ThunderBolt for Disk & Network Access?

We can have [Drobo Drives](#) connected via [FireWire 800](#) to mac minis while fastest commercial SATA drives can deliver [6Gbps](#). Faster Disk access with [FiberChannel Adapter](#) and [other technologies](#). Networking over ThunderBolt is not available because its not supported by network devices yet (hub/switch).

Can I just use wi-fi for network in Hadoop Infrastructure?

Current 802.11n networking on Mac mini can theoretically provide 300Mbps speed but in practice people have seen 130 Mbps ~ 150 Mbps.

New generation technology WiGig/802.11ad might delivers upto 7Gbps but not until end of 2012.

On other hand, Gigabit ethernet on Mac mini supports 1Gbps transfer, practically [it can deliver 500 Mbps data transfer](#). We might require to buy a [switch](#) if we choose to go only ethernet.

How much main memory & data storage can I get on Mac Mini ?

For main memory, 8GB DDR3 is supported for each Mini.

For storage, It supports 750GB (7200-rpm) hard drive per mac mini. Approx $750 \times 5 = 3.7$ TB with five minis

Power Consumption?

A mac mini requires about [85 W](#) (10 W if idle) with thermal 290 BTU/h , this is by far [lowest consumption](#) (A Dell 1U server would take 570+ W). The most efficient system in market would require Intel Atom Processor which can clock only up to 1.6 Ghz , it draws about [32 W](#).

Conclusion: Mac minis connected over gigabit ethernet looks good for basic Hadoop infrastructure

Resources

Linux Installation: <http://www.building43.com/blogs/2011/01/28/single-boot-linux-on-an-intel-mac-mini/>

From scratch linux on Mac Hardware: <http://www.madingley.org/macmini/walkthrough.html>

RackMac Mini Servers: <http://sonnettech.com/product/rackmacminiserver.html>

Hadoop Installation : <http://www.michael-noll.com/tutorials/running-hadoop-on-ubuntu-linux-multi-node-cluster/>

Hadoop@FB : <http://borthakur.com/ftp/RealtimeHadoopSigmod2011.pdf>,

<http://borthakur.com/ftp/SIGMODRealtimeHadoopPresentation.pdf>,

<http://hadoopblog.blogspot.com/2011/05/realtime-hadoop-usage-at-facebook-part.html>

Hadoop Performance Model : <http://pkghosh.wordpress.com/2011/11/06/simple-hadoop-performance-model/#more-474>

Mathematics of Hadoop Infrastructure:

<http://nathanmarz.com/blog/the-mathematics-behind-hadoop-based-systems.html>

Hadoop on Mac OS X: [http://wiki.apache.org/hadoop/Running_Hadoop_On_OS_X_10.5_64-bit_\(Single-Node_Cluster\)](http://wiki.apache.org/hadoop/Running_Hadoop_On_OS_X_10.5_64-bit_(Single-Node_Cluster))Tracking down the number of Firefox Addon users with

hadoop :<http://blog.mozilla.com/data/2009/08/10/tracking-down-the-number-of-firefox-addon-users-with-hadoop/>

Thunderbolt Technology Community: <https://thunderbolttechnology.net>

A Mac mini-like barebone cheaper box: http://www.logicsupply.com/products/nc9b_hm67 &

http://www.adlinktech.com/PD/marketing/Datasheet/MI-220/MI-220_Datasheet_1.pdf

BigData.sg – A community of BigData nerds from Singapore.