Student Learning Using MongoDB Sharding on a Cluster of Ubuntu Raspberry Pi 4B Servers

Robert T Mason
William K Masters
7/6/2022

Welcome

Robert T Mason

William K Masters

Today's Overview

1. Cloud computing does not provide hands-on access to hardware (directly)
2. Building a cluster of server machines is expensive ($30,000)
3. Raspberry Pi 4B(s) are inexpensive alternative ($1,628) to build a cluster

Overview of Topic (for non-expert)

- This paper is a tutorial that demonstrates that Raspberry Pi 4b servers (with 8 gig of RAM) can be leveraged to build a cluster of low cost servers to run both Linux Ubuntu 20 and MongoDB Sharding (NoSQL database distributed processing).

The ability to purchase commodity servers (e.g., $3,000 per Dell server) to create a cluster of multiple machines is cost prohibitive for most faculty and students because it can cost upwards of $30,000 for 10 machines.

This cost does not include the other hardware components that are required for the cluster, such as cooling equipment, cables, rack, etc.

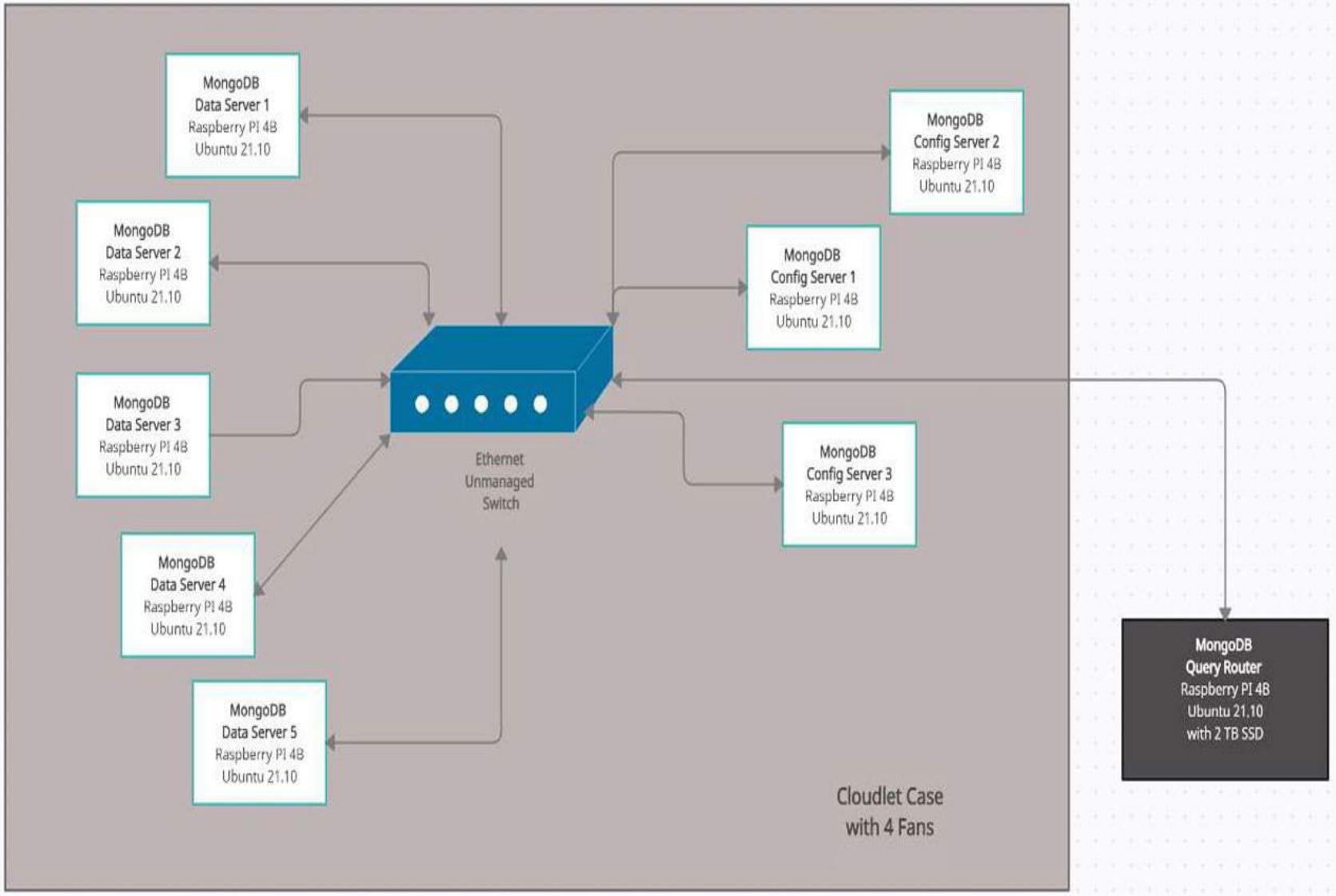

Figure 1: Rpi Model 4B Cluster Architecture

Table 1: Hardware Components

Quantity	Hardware Components
9	Rpi 4B with 8 gb RAM - 64-bit with quad-core Broadcom BCM2711 SoC - (9 x $99 = $891)
9	Rpi Power Adapters - (9 x $8 = $72)
1	Cloudlet Case with 8 slots and 4 cooling fans (to hold the Rpi 4Bs) – ($59.99)
1	Solid State Drive (SSD) - 2 Terabyte in size - ($194.99)
1	SSD Kit for Rpi, includes one cooling fan and case – ($77.99)
1	Unmanaged Switch with 8 ports (note: one of the 4B servers used WiFi) – ($21.99)
8	Ethernet Cables - 1 ft length (to plug the Rpi 4Bs into the switch) – ($17.01)
1	Ethernet Cable 3 ft. length (to plug the switch into the Comcast router) – ($1.04)
9	Rpi 32 gb microSD cards formatted with Ubuntu 20 – ($8.99 x 9 = $80.91)
8	Flash Drives with 128 gb of storage – (8 x $18.99 = 151.92)
2	Power Plugs – (surge protector cords to plug in the power adapters) – (2 x 29.99)

```
mongos> db.cryptoHistory.getShardDistribution()

Shard shard0000 at 10.0.0.???:270??
 data : 16KiB docs : 271 chunks : 2
 estimated data per chunk : 8KiB
 estimated docs per chunk : 135
...

Totals
 data : 94KiB  docs : 1500  chunks : 10
 Shard shard0000 contains 17.74% data, 18.06% docs in cluster, avg obj size on shard : 63B
 Shard shard0001 contains 21.26% data, 21.33% docs in cluster, avg obj size on shard : 64B
 Shard shard0004 contains 19.95% data, 20.86% docs in cluster, avg obj size on shard : 61B
 Shard shard0003 contains 20.23% data, 19.13% docs in cluster, avg obj size on shard : 68B
 Shard shard0002 contains 20.79% data, 20.6% docs in cluster, avg obj size on shard : 64B
```

Summary

- Raspberry Pi 4B(s) are inexpensive alternative ($1,628) to building a cluster of commodity servers
- Students are often more enthusiastic when they are able to conduct hands-on experimentation with hardware and software (Hills et al., 2019).
- **Future work in this area can inspire students to experiment with NoSQL databases by working on performance scaling, deployment automation, Edge processing and observability.**

QUESTIONS?