Katya Bespartochna
February 11th, 2020

Cloud-Cloud Computing Influence on Scalability

Cloud computing is becoming a network technology that is more integrated with the internet. The more it gets integrated, the more problems relating to load unbalancing like multi-variant, multi-constraint and degradation of performance and resources come up. Load balancing involves the methods that packets of information are being sent in from different locations of the internet. There have been a few models proposed to ensure the best load of that information.

Scalability is the Cloud Computing feature that tells how much a server is able to handle the growth and or decrease of resources in a business. These resources are monitored in virtual machines. As an example, one model shows a 'Two level load balancing Architecture' where a User connects through the request generator, Datacenter controller eventually getting to the load balancer. The second level shows the same load balancer distributing the same information in a bunch of virtual machines which had packets of information sent from the physical machines(Afzal, 2).

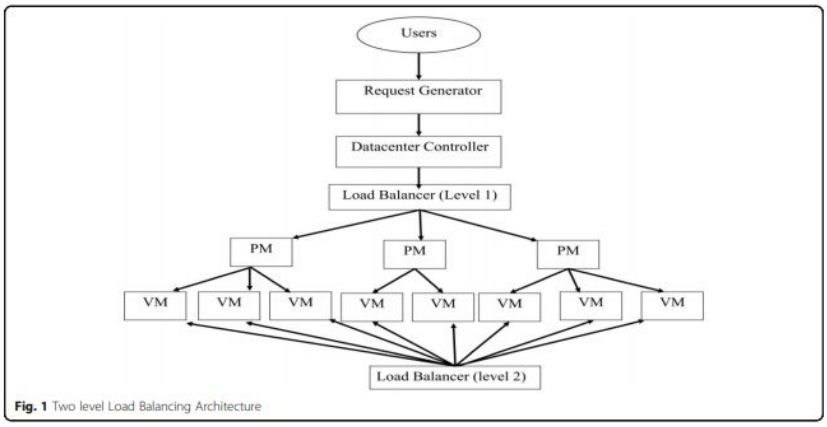

Fig. 1 Two level Load Balancing Architecture

Another study was done on a database center network in Italy. The information was gathered from a combination of sources such as services like HTTPs and DBs, Middleware/cloud platforms, and hardware layers(Elia, 2).

This architecture included 13,000 cores, 10 PB of disk storage 2.5 PB tape library which is complex and big enough architecture to get a substantial amount of traffic to monitor.

The cloud platform that was used was Openstack. Big data was tested on it because of its open source and big community. Other reasons openstack was used included Horizontal scalability, the high available amount of data, and Near-real-time and batch processing.

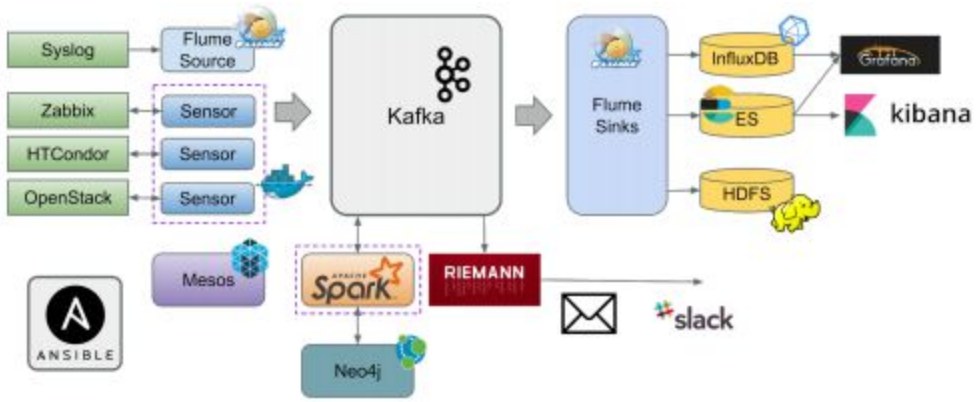

Figure 2. Architecture of proposed project.

Failures:

- Service malfunctions could be detected using the first source category but this information alone does allow to figure out the root cause.
- Disk controller fault

Solutions:

The information in various hosts most likely gets disrupted through anomalies. These are

outliers in data centers which are different from the rest of the data and can give a guide in how much and what resources are being lost or increased. The research did some tests on these anomalies through Root-causing analysis algorithms.

The article suggested implementing a 'newRCA uses machine learning and graph theory algorithms'.

This algorithm includes features like an event extractor, a graph builder and root-cause extractor and is meant to self-organizing (Elia,4).

Afzal, Shahbaz, and G. Kavitha. "Load balancing in cloud computing–A hierarchical taxonomical classification." Journal of Cloud Computing 8.1 (2019): 22.
Elia, Domenico, et al. "Developing a monitoring system for Cloud-based distributed data-centers." EPJ Web of Conferences. Vol. 214. EDP Sciences, 2019.