Performance Evaluation of ESX Cloud with Cloud Native Apps

Arun Kalyanasundaram

Mentors: Tariq Magdon-Ismail, Ravi Soundararajan. Manager: Priya Sethuraman

MOTIVATION

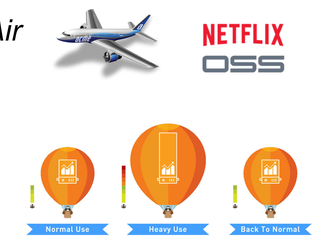
- ESX Cloud is designed for Cloud Native Apps*.
  *Think sheep not pets!
- However, does it deliver on its promise?

GOALS

1. Evaluate the performance of a real world Cloud Native App on **ESX Cloud**.
2. Compare its performance with **vSphere**.

OUR APPROACH

- **Cloud Native App Benchmark**: *Acme Air*
- **Test Bed**: *Apache Jmeter*
- **Auto Scaler Service**: Scale up based on app performance.



Block diagram of our benchmark deployment

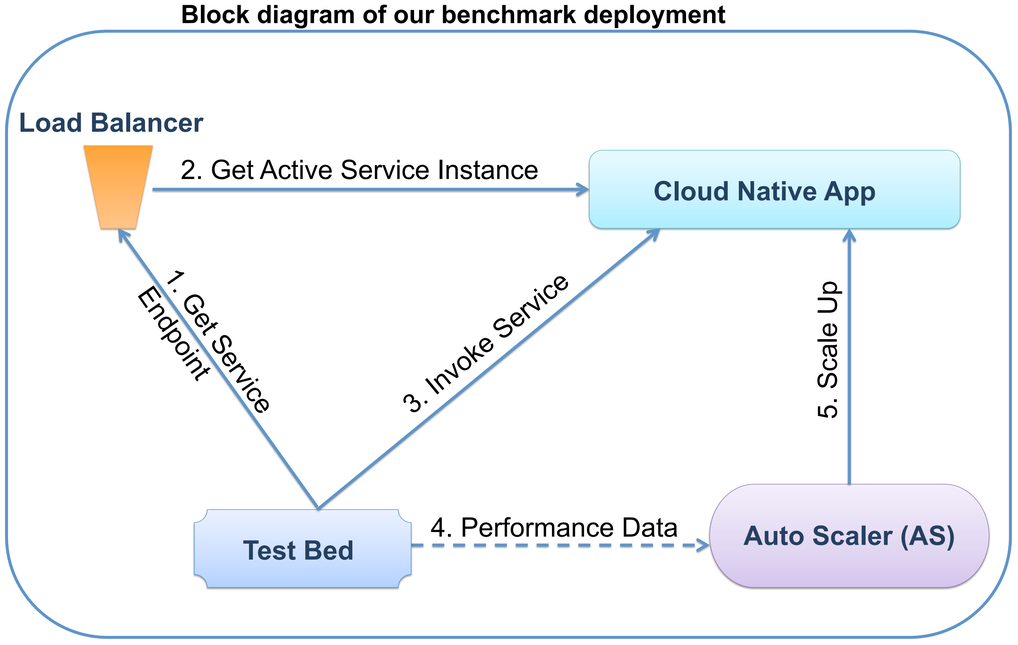


Figure 1: The *test bed* invokes a service on the *cloud native app* by acquiring the address of an active service instance from the *load balancer*. New service instances are spawned by the *auto-scaler* based on the performance data from the *test bed*.

Block diagram of our cloud native app

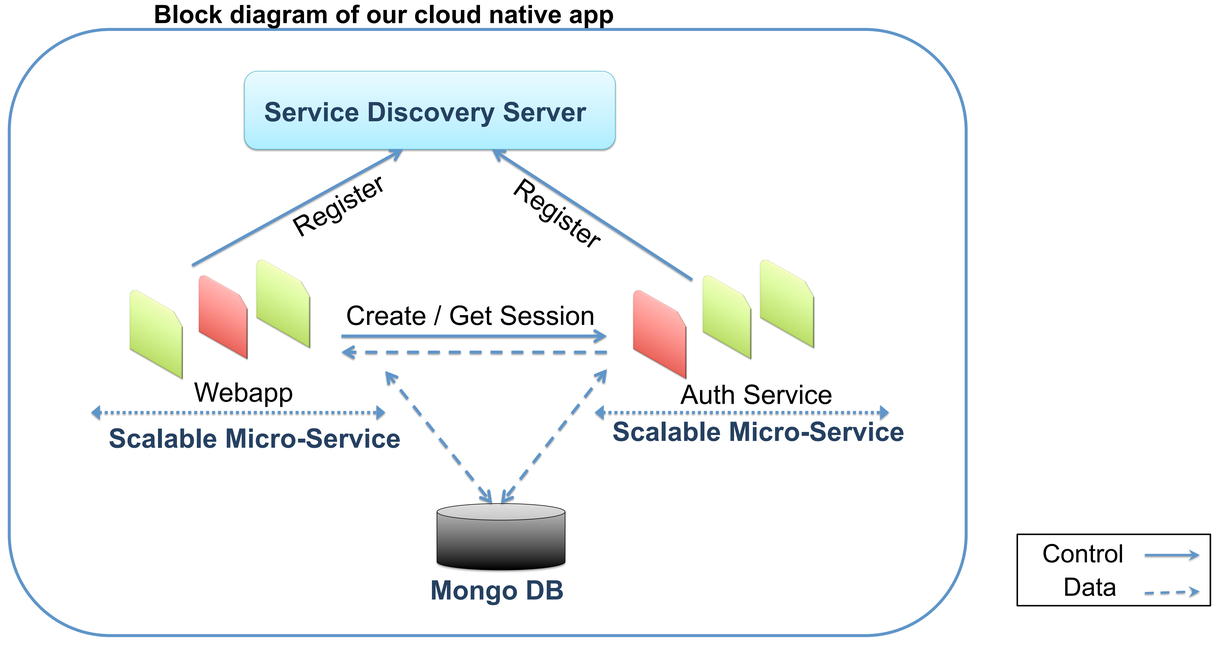


Figure 2: A *Service Discovery Server* keeps track of active service instances. *Webapp* micro-service is a user facing service which uses the *Auth service* to manage sessions. Both services use a shared Mongo DB database to persist data.

RESULTS: PERFORMANCE MEASURES

	ESX Cloud w/o AS	vSphere w/o AS	ESX Cloud with AS	vSphere with AS
Mean Response Time	116 ms	121 ms	306 ms	407 ms
Average Throughput	1169 / sec	1139 / sec	559 / sec	455 / sec
Dropped Connections	4 %	4 %	5.6 %	2.5 %

Table 1: Performance comparison between ESX Cloud and vSphere with and without the **Auto Scaler (AS)**. Workload: 300 users ramping up in 300 seconds for a total duration of 600 seconds.**

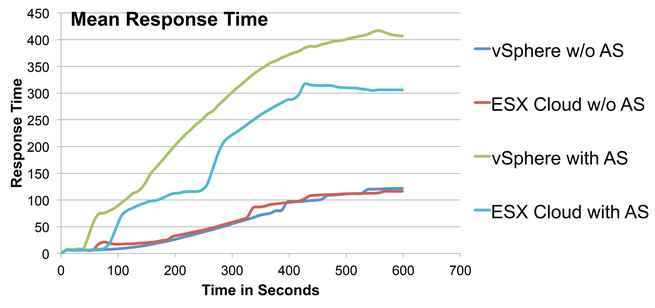


Figure 3: Plot shows the change in mean response time for the duration of the workload.**

** Results from a two node nested ESX host cluster

RESULTS: KEY TAKEAWAYS

- Both ESX Cloud and vSphere have a similar performance without auto-scaler.
- ESX Cloud outperforms better with auto-scaler.
- However, ESX Cloud has a higher fraction of dropped connections.

FUTURE WORK

- Evaluate the performance on larger environments and with larger workloads.
- Evaluate under High Availability (HA) scenarios.
- Develop new performance metrics.