HPC on the Cloud – A Norwegian Perspective

Nico Reissmann[1], Einar Næss Jensen[2], Jan Christian Meyer[3]

[1]Norwegian University of Science and Technology, nico.reissmann@ntnu.no
[2]Norwegian University of Science and Technology, einar.nass.jensen@ntnu.no
[3]Norwegian University of Science and Technology, jan.christian.meyer@ntnu.no

Keywords
Cloud Computing, High-Performance Computing

1. Summary

Cloud computing (CC) has become a viable alternative for business and web applications. As vendors start to offer products with bare metal servers and high-speed interconnection, it also becomes increasingly interesting to the high-performance computing (HPC) domain. This paper presents our experience of matching Fram, Norway's largest supercomputer, with equivalent Cloud alternatives. We compare the available features and necessary total cost of ownership (TCO) of Fram to four different products: Google Cloud, Amazon Web Services (AWS), Microsoft Azure, and IBM Softlayer.

2. Introduction

In recent years, CC has become increasingly popular in the commercial arena as a cost-effective alternative to run business and web applications. On one hand, it provides a fee-for-service model with 24/7 availability, and offers flexible utilization. On the other hand, it traditionally provided inferior network performance in terms of latency and bandwidth, as well as virtualization overhead and performance variation due to interference from virtualization. This rendered it an unsuitable replacement for conventional HPC systems. However, as vendors expand their services to include bare metal servers and high speed interconnections, CC becomes increasingly interesting to the HPC domain.

In this paper, we present the difference in total cost of ownership between Norway's largest supercomputer, called Fram, and Cloud systems of similar capacity. We match resources from four different providers with those of Fram, and estimate TCO using their public pricing information. Our comparison shows that an equivalent cloud-based system would be approximately four to seven times more expensive than Fram.

3. Fram Specification

An approximation of Fram's capabilities is necessary to establish a reasonable reference for comparison, and to accommodate the limited options of the vendors' pricing calculators. Table 1 shows Fram's different node classes. The compute nodes can be grouped by their respective amounts of memory and disk space. The significant majority of the nodes belong to compute A and only consist of 64 GB memory as well as 120 GB disk space. In addition, Fram contains 8 nodes (compute B) with 512 GB memory and 960GB disk space as well as 2 big-memory nodes (compute C) with 6 TB and 14.4 TB disk space. The nodes for accessing Fram (Frontend) are equipped with 128 GB of memory and 800GB of disk space. Moreover, all nodes have a 1 Gbps Ethernet interface and 1 EDR Infiniband high-speed interface. Fram's storage capacity is a 2.45 PB parallel Lustre file system with Remote Direct Memory Access (RDMA) capabilities over EDR Infiniband. It is tuned for parallel workloads, managing an average of 6.5 million file changes per day. Fram's nodes are connected to a RDMA capable 100 Gbps EDR Infiniband, and a 1 Gpbs Ethernet used for administrative purposes. The amount of outgoing data is ca. 150 TB per month on the compute nodes, and 100 TB on the frontend nodes.

***Table 1**: Node classes of the Fram supercomputer.*

Class	# Nodes	# Cores per Node	Memory [GB]	Disk Space [GB]
Compute A	940	32	64	120
Compute B	8	32	512	2 x 960 (Raid 0)
Compute C	2	28	6000	14400
Frontend	10	32	128	2 x 800 (Raid 1)

4. Feature and TCO Comparison

Table 2 shows the feature comparison between Fram and four different vendors. No vendor offers all features. In case of compute nodes, all products provide equivalent or partially sufficient alternatives to the nodes in Compute B and C, but no product provides Compute C nodes. IBM Softlayer is the only vendor that provides bare metal nodes. The only product offering a high-speed interconnect is Azure. All other products offer no such interconnect, or as in case of AWS, provide partially sufficient 100 Gbps Ethernet combined with RDMA over converged Ethernet (RoCE). In case of storage, the only product providing the possibility of a parallel file system is AWS.

***Table 2**: Feature comparison between products.*

	Compute A	Compute B	Compute C	Frontend	Storage	Network	RDMA
Fram	✓	✓	✓	✓	✓	✓	✓
Cloud	✓	✓	✗	✓	✗	✗	✗
AWS	✓	–	✗	–	–	–	✗
Azure	–	–	✗	✓	✗	–	✓
Softlayer	✓	–	✗	✓	✗	✗	✗

✓:equal or equivalent –: partially available or partially sufficient ✗: not available or insufficient

No product matches the two Compute C nodes. Even if these were excluded, no product would be able to meet all requirements. In particular, most products provide no adequate replacements for the required storage and network infrastructure, or RDMA capabilities.

***Table 3**: Monthly Cost comparison between products in MNOK.*

Fram	Cloud	AWS	Azure	Softlayer
1.687	6.8311	7.6003	12.3383	11.3959

Table 3 shows estimated costs per month for each product in million Norwegian Kroner (MNOK). We picked the closest option available from a vendor in cases where a product offered no sufficient alternative. The cost of Fram is computed using its total hardware cost without big-memory nodes, vendor and installation costs, and housing costs such as cooling, electricity, etc., which amounts to $58-1.1+23.87=80.77$ MNOK. Fram's lifetime is taken to be 4 years, which gives an estimated cost of $80.77/48 = 1.682708$ MNOK per month. Table 3 indicates that the cheapest alternative even without matching requirements is four times as expensive per month. The only alternative with high-speed interconnect is ca. seven times as expensive. The TCO comparison for IBM Softlayer was made with virtual nodes, as bare metal nodes are ca. ten times more expensive.

4. AUTHORS' BIOGRAPHIES

Nico Reissmann: https://www.linkedin.com/in/nico-reissmann-103446b9/

Einar Næss Jensen: https://www.linkedin.com/in/einjen/

Jan Christian Meyer: https://www.linkedin.com/in/jan-christian-meyer-16a0837/