Improving Resource Usage in HPC Clouds

Tuesday 1 October 2019 16:45 (15 minutes)

HPC-as-a-service is a new cloud paradigm that represents easy to access and use cloud environments for High Performance Computing (HPC). This paradigm has been receiving a lot of attention from the research community lately since it represents a good tradeoff between computational power and usability.

One of the key drawbacks associated with HPC clouds is low CPU usage due to the network communication overhead [2, 3]. Instances of HPC applications may reside on different physical machines separated by significant network latencies. Network communications between such instances may consume significant time and thus result in CPU stalls.

In this paper we propose the scheduling algorithm that overcomes such drawbacks to increase the HPC task capacity in the Ethernet-based HPC cloud by sharing CPU cores between different VMs. The algorithm observes parallel tasks'behavior and packs tasks with low CPU usage on same CPU cores.

We fully implemented and evaluated our algorithm on 15 popular MPI benchmarks/libraries. The experiments have shown that we can significantly improve the CPU usage with negligible performance degradation.

Summary

During the past decade public clouds have attracted tremendous amount of interest from academic and industrial audiences as the effective and relatively cheap way to get powerful computational infrastructure without the burden of building and maintaining physical infrastructure.

Although clouds are less powerful than server clusters or supercomputers [1], they are becoming more popular as a platform for High Performance Computing (HPC) due to the low cost and easy to access. Cloud providers are starting to support this interest and come up with a new cloud paradigm - HPC-as-a-service. This paradigm represents a service that gives cloud resources for computationally heavy applications.

Several papers [2, 3] have shown that one of the main performance bottlenecks in HPC clouds issues from communication delays within the DC network. Such bottleneck is due to the insufficient network performance in HPC clouds. While supercomputers use fast interconnections like InfiniBand or GE, HPC clouds mostly use Ethernet.

However, this bottleneck brings important impact on the behavior of the applications in HPC clouds –communication heavy HPC applications tend to underutilize the CPU. This happens because most of computationally heavy applications use network to exchange messages between physical machines. And since cloud network is not fast enough for HPC, such applications spend a lot of time idling for messages to pass through the network [2, 3].

Such behavior of HPC applications also leads to a highly regular execution and network usage pattern, i.e. HPC applications show tendency to alternate computations with frequent network communications [5]. This communication pattern contribute to the idle CPU usage since the slowest message delivery dictates the overall performance degradation of an application.

These specifics of the behavior of HPC applications can be used in HPC clouds to improve the resource utilization by sharing the same CPU core between different applications, i.e. providing more virtual CPUs than there are physical ones. In this research we are proposing a scheduling algorithm that increases the resource utilization and the HPC task capacity of an Ethernet-based HPC cloud. The developed algorithm observes

network behavior of HPC tasks and uses a greedy heuristic to share CPU cores between such tasks, thus improving the overall CPU usage and increasing the number of tasks performed via HPC-as-a-service.

We have performed experiments with 15 popular MPI benchmarks/libraries and show that we can significantly improve CPU usage with negligible performance degradation.

[1] Netto, M. A., Calheiros, R. N., Rodrigues, E. R., Cunha, R. L., & Buyya, R. (2018). HPC cloud for scientific and business applications: Taxonomy, vision, and research challenges. ACM Computing Surveys (CSUR), 51(1), 8.

[2] Gupta, A., Faraboschi, P., Gioachin, F., Kale, L. V., Kaufmann, R., Lee, B. S., ... & Suen, C. H. (2016). Evaluating and improving the performance and scheduling of HPC applications in cloud. IEEE Transactions on Cloud Computing, 4(3), 307-321.

[3] Gupta, A., & Milojicic, D. (2011, October). Evaluation of hpc applications on cloud. In 2011 Sixth Open Cirrus Summit (pp. 22-26). IEEE.

Primary author: PETROV, Ivan (Lomonosov Moscow State University)

Co-authors: CHUPAKHIN, Andrey (Lomonosov Moscow State University); SMELIANSKY, Ruslan (Lomonosov Moscow State University); ANTONENKO, Vitaly (Lomonosov Moscow State University)

Presenter: CHUPAKHIN, Andrey (Lomonosov Moscow State University)

Session Classification: Distributed Computing. GRID & Cloud computing

Track Classification: Distributed Computing. GRID & Cloud Computing