

CLIC (Cloud In Cresco): towards HPC/HPDA-as-a-Service

Marco Faltelli, Alessandro Peloso, Francesco Iannone, Matteo Fois, Massimo Celino and Giovanni Ponti

ENEA, Lungotevere Thaon di Revel, 76 – 00196 Rome, Italy

Abstract. The demand for High-Performance Computing (HPC) services is rapidly growing, driven by the advent of new accelerators and increasingly complex tasks, with artificial intelligence (AI) being a major focus. Despite this, HPC cluster users continue to rely on bare-metal resources for their computations. In this position paper, we argue that the time has come for HPC to transition to virtualized services. Virtualization has demonstrated its scalability, reliability, and efficiency while also abstracting the underlying hardware's intricate details. Additionally, virtualized environments offer significant flexibility, allowing for easy adjustment to varying workloads and security through isolation, ensuring that different virtual machines can operate securely on the same physical hardware. We present a PoC architecture that shall be deployed in our virtualized HPC cluster based on OpenStack (as part of the IPCEI-CIS project), showing the future challenges and issues that may emerge

Keywords. HPC, cloud, GPU, IAAS

1. Introduction

The IPCEI-CIS (Next Generation Cloud Infrastructure and Services) project aims at creating a multi-provider, European-sovereign cloud-edge continuum between different partners, such as telco operators, industry operators, and R&D centers. ENEA is part of IPCEI-CIS through the DataCLEEN project: DataCLEEN's goal is to create a highly scalable and reliable cloud infrastructure based on HPC technologies.

ENEA is already a major HPC provider in Italy through the CRESCO infrastructure, with CRESCO8 being launched in April 2025; DataCLEEN will enable the creation of the first cloud HPC cluster called CLIC (CLOUD In Cresco), which is expected for the end of 2025.

We believe that creating a cloud HPC environment brings significant benefits to the community and poses significant challenges in engineering and maintaining such an infrastructure.

In this paper, we first describe the motivations behind DataCLEEN's virtualized HPC infrastructure. Then, we describe our proof-of-concept architecture for the future CLIC supercomputer, focusing on the challenges and some preliminary results, which show very close performances compared to bare-metal supercomputers.

2. Why cloud HPC

Adopting VMs in HPC clusters brings a series of benefits to both users and maintainers, which are described here.

2.1 The users' side

Unified access to resources: HPC suppliers usually provide users with isolated clusters with different login portals, requiring them to choose a cluster a-priori based on their needs. Virtualization offers a unified login portal, with hypervisor controllers selecting the best cluster for each workload.

Integrated data handling: A unique infrastructure for all the clusters brings the significant benefit of having a unique data storage platform that is geographically distributed and redundant among the clusters. In this way, new clusters and storage can be added modularly.

Specialized environments: HPC is traditionally related to topics like science materials, fusion power, bioscience, and weather forecasts. Lately, AI/ML has gained a lot of interest because of the increasing hardware requirements of LLMs. Each of these different fields requires specific software and packages, and it is hard to provide a catalog that can satisfy all possible applications. Through virtualization, users can create and replicate VMs equipped with specific software for their needs. For example, AI users can create VM images equipped with CUDA, PyTorch, Numba, and LLM-specific libraries.

Failure resilience: HPC jobs often involve many nodes and long durations, increasing the chances of hardware failure. Virtualization minimizes the consequences of failures through periodic snapshots and live migration of instances.

2.2 The administrators' side

Enhanced security: virtualization implies isolation both at a software level and at the network level, as well as advanced monitoring and logging features.

Easier maintenance: through multiple availability zones and migration capabilities, administrators can operate in isolated parts of the cluster without impacting the users.

Better resiliency: in case of any fault (hardware, network, software), VMs can be migrated to a different availability zone, thus giving the cluster a high degree of resiliency to failures.

3. A primer on CLIC

3.1 Infrastructure

Our PoC infrastructure is based on two ENEA HPC clusters: 20 nodes of the dismissed CRESCO4 cluster and 11 nodes from CRESCO5F. The CRESCO5F nodes feature 64 AMD EPYC 7313 CPU cores, 256GB RAM, and Infiniband Connect-X6 100Gbps NICs. Seven of these nodes are equipped with NVIDIA GPUs, including A100 and H100 models.

OpenStack is the main component of our infrastructure and it orchestrates our cluster's different computing, storage, and network resources through an Infrastructure-As-A-Service (IAAS) paradigm. OpenStack is composed of different services that can be added to our infrastructure modularly; each service is responsible for a certain feature in OpenStack. OpenStack relies on Canonical frameworks like juju and MAAS: the former permits us to manage OpenStack and its dependencies, while the latter discovers and provisions the available resources.

3.2 Challenges

3.2.1 Performance

The first requirement for CLIC is that performances should be equal (or very close) to bare-metal clusters. In other words, virtualization should not add a significant overhead. We underline that this is not a trivial question to be answered, as in HPC it is of paramount importance to effectively virtualize resources such as GPUs and Infiniband devices. As we show in Section 3.3, preliminary results highlight very similar performances between virtualized and bare-metal HPC hardware.

3.2.2 Licensing

Another challenge (at the moment of writing this paper) is GPU virtualization, which can be exclusively done on NVIDIA GPUs. Virtual GPUs need a license from NVIDIA, which must be bought separately. Otherwise, their performance is degraded twenty minutes after the VM startup, and after 24 hours, CUDA stops working. This is a substantial limitation in adopting vGPUs in HPC scenarios, especially since most of these clusters are used for non-profit, scientific research reasons. At the time of writing, AMD is planning to introduce virtual GPUs in the next year, which could be a viable alternative for GPU virtualization.

3.3 First results

Here, we showcase two tests focusing on fundamental HPC performances. The first one highlights the performances of a licensed and virtualized NVIDIA A100 GPU when running common HPC benchmarks, such as HPL and HPCG. We can see from Table 1 that the performance of the virtualized hardware is very close to that of the bare-metal one. Table 2 focuses on the communication overheads that the virtualization may cause. Surprisingly, the performances of the virtualized infrastructure perfectly match the bare-metal ones, with zero overhead. We believe these two tests make a first case for virtualized HPC clusters. We plan to do further performance tests in the near future.

| | vGPU | Bare-metal |
|---------------------|------|------------|
| HPL 21.4 (TFLOPS) | 9.73 | 9.85 |
| HPCG 24.09 (GFLOPS) | 228 | 230 |

Tab. 1: HPC tests on NVIDIA A100

| | VM | Bare-metal |
|-------------------|-----------|------------|
| OFED perf (read) | 96 Gb/sec | 96 Gb/sec |
| OFED perf (write) | 96 Gb/sec | 96 Gb/sec |
| MPI host-to-host | 88 Gb/sec | 88 Gb/sec |

Tab. 2: HPC communication tests with ConnectX-6 100Gbps Infiniband NICs

4. Future Plans

Our future plans include using the architecture described here for CLIC, the next ENEA HPC cluster, which will be distributed among four different ENEA research centers: the main cluster is going to be hosted in Frascati, while the other three will be hosted in

Casaccia, Portici, and Brindisi research centers. We foresee that such architecture will create innovative solutions for HPC and will foster collaboration with other partners of the IPCEI-CIS project.

Acknowledgement

This work was funded by the Italian Ministry MIMIT in the frame of the European initiative IPCEI CIS, under the PNRR/NextGenerationEU, as part of the project DataCLEEN (CUP: I38H23000710006).

Authors

Marco Faltelli marco.faltelli@enea.it

Marco Faltelli is a Researcher in the ENEA High Performance Computing Lab. In June 2023 he obtained his PhD in Computer Science under the supervision of Prof. Giuseppe Bianchi and Prof. Francesco Quaglia. He is a Microsoft Research PhD Fellowship recipient. His research activities combine computer architecture, HPC, and computer networks to design high-performing, scalable architectures for both network hardware and software solutions.

Alessandro Peloso alessandro.peloso@enea.it

Alessandro Peloso is a software architect in the ENEA TERIN department. His interests combine operative systems, cloud computing

Francesco Iannone francesco.iannone@enea.it

Francesco Iannone holds a degree in Physics. Currently he is researcher directory at ENEA specialising in High-Performance Computing (HPC). He works on the development and optimisation of scientific applications. His expertise includes parallel computing, numerical modelings, and support for HPC infrastructures within national and European projects.

Matteo Fois matteo.fois@enea.it

Matteo Fois holds a master degree in Applied Mathematics and a master course in HPC from the University of Rome Sapienza. He works as a researcher in ENEA's HPC laboratory of Portici, managing the compute infrastructure and collaborating with different research groups to help the development of parallel codes and simulations on ENEA's clusters.

Massimo Celino massimo.celino@enea.it

He studied Physics at the University of Rome Sapienza and holds a PhD from the University "L. Pasteur" of Strasbourg (France). He promotes and manages national and European projects in the field of HPC and cloud. He is currently the PI of the ENEA DataCLEEN project as part of the IPCEI-CIS cloud initiative. He is the author of more than 100 scientific papers in international journals.

Giovanni Ponti giovanni.ponti@enea.it

He is Director of the ENEA ICT Division and member of the GARR board of directors. PhD in Computer Engineering in 2010 (University of Calabria), during which he focused on data modeling and innovative data mining algorithms. Senior ENEA researcher for 15 years with research activities in HPC, Data Science, Cloud Computing, Big Data, Data Analytics and AI. In these contexts, he is the author of prestigious scientific publications (more than 100 with referee), member of the conference program committee and reviewer for important journals and conferences.