Accelerating Scientific Applications with the Quantum Edge: A Drug Design Use Case

Vincenzo De Maio(✉) and Ivona Brandic

Vienna University of Technology, Vienna, Austria
vincenzo@ec.tuwien.ac.at, ivona@ec.tuwien.ac.at

Abstract. To address increasing demands of computational resources, scientific applications started to support different type of hardware accelerators, e.g., GPUs, TPUs, ASICs. However, due to the limitation to scalability of hardware resources posed by the Moore Law, a conspicuous amount of research is focusing in integration of Quantum Computers in the overall computing continuum. The high amount of data and the high heterogeneity of quantum architectures necessitate development and integration of additional hardware/software layers to facilitate integration of quantum hardware into the overall computing continuum. In this work, we discuss the possibility of applying edge computing to address this issue, laying the fundamentals to the concept of the Quantum Edge. Further, we present and analyse a drug design use case, identifying challenges and future research directions for Quantum Edge.

Keywords: Scientific computing · Drug design · Quantum computing · High performance computing

1 Introduction

Scientific applications are used by scientists from different domains to simulate complex scientific phenomena and speed up research in different fields. Typical examples of scientific applications are molecular dynamics [14], material science [8], and drug design [4]. Due to their high computational demands, scientific applications rely on HPC resources [15] for their execution.

However, we are currently entering the Post-Moore era, which poses serious limitations to scalability of classic HPC systems. Considering the growing storage and analytics demands, scientific computing recently focused in exploiting different specific-purpose hardware accelerators. Accelerators include devices such as TPU, GPUs, FPGAs and ASICs, each one designed for a very specific problem and with its own specific limitations. To further improve capabilities of future HPC systems, a considerable amount of research is considering the exploitation of so-called Non-Von Neumann architectures, such as Neuromorphic and Quantum Computing [12].

A. Bienz et al. (Eds.): ISC High Performance 2023 Workshops, LNCS 13999, pp. 134–143, 2023.
https://doi.org/10.1007/978-3-031-40843-4_11

Among Non-Von Neumann architectures, quantum computing clearly stands out, due to characteristics such as (1) theoretically proven speedup for many computationally-intensive problems, (2) natural 3D modelling of different scientific problems, and (3) wide availability of different quantum systems and frameworks [24]. At one hand quantum computing can provide significant speedup to many scientific computations, such as computational chemistry, combinatorial optimization and drug design. On the other hand, current state-of-the-art NISQ (Noisy Intermediate Scale Quantum) architectures suffer from drawbacks, such as limited number of qubits and high amount of error, limiting adoption of quantum computing on a larger scale.

Hybrid classic/quantum systems target integration of classic von Neumann hardware with quantum hardware. The main idea of hybrid classic/quantum systems is to use the classic HPC facilities for the preparation of the input data, adaptation of the classic code for the execution on the quantum machine and for the data post processing. The quantum machine is utilized for the execution of very specific program parts that can benefit from the quantum architecture. Once a program is executed on the quantum hardware, classic hardware will perform error correction on its output. However, how to deal with issues such as management of streaming data, privacy and integration between classic and quantum hardware still remains an open challenge.

Edge computing has proven effective in addressing challenges of processing streaming data [11,12], as well as ensuring privacy of geographically distributed systems [20]. Therefore, in this work we investigate the use of Edge computing to address challenges of integrating quantum in the computing continuum.

First, we describe our target use case and identify which components would benefit from quantum execution. Afterwards, we identify challenges of executing target use case on hybrid classic/quantum systems. Finally, we define our concept of Quantum Edge and describe which Edge computing methodologies could be applied to our specific use case. To demonstrate the applicability of Edge Computing in hybrid systems we focus on a computer-assisted drug design use case, which is one of the widely used applications in scientific computing [2].

The paper is organised as follows: First, we introduce background notions in Sect. 2 and our target use case in Sect. 3. In Sect. 4, we identify challenges towards implementation of current use case in hybrid classic/quantum systems. The concept of Quantum Edge is described in Sect. 5. Finally, we describe related work in Sect. 6 and conclude our paper in Sect. 7.

2 Background

2.1 State of the Art in Scientific Computing

Scientific computing is a branch of computer science, spanning different disciplines (i.e., finance, biology, chemistry, engineering) with the goal to develop *standardized*, *robust* and *accurate* simulations of different phenomena. Simulations can be decomposed in *tasks* (i.e., aggregate data from different sources, average a set of samples). Tasks can be combined into *workflows*, represented as DAGs (Directed Acyclic Graphs) [1] where edges model interdependencies between tasks. Data-intensive workflows are defined as *extreme-data workflows* [11].

Current research trends in scientific computing go towards the convergence of artificial intelligence, data science and physical simulations [5], which rely on HPC systems for their execution. However, given the limitation to hardware scalability posed by the Post-Moore era, together with the increasing data that have to be processed and growing computing demands of scientific applications, scientific community is faced with the challenge of scaling HPC capabilities beyond limitations of Von Neumann's architectures [10].

2.2 Hybrid Classic/Quantum Systems

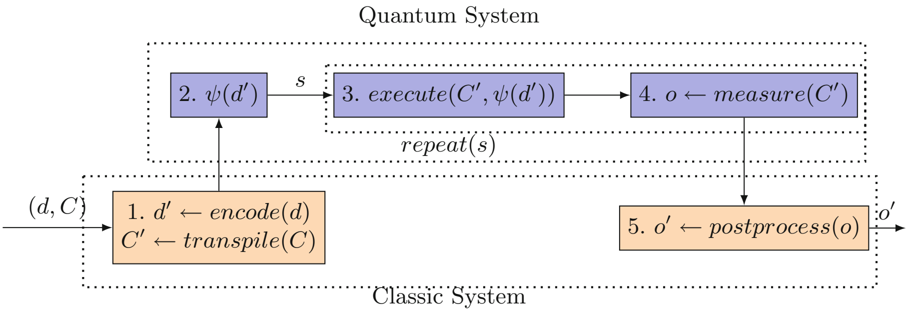

Fig. 1. Hybrid Classic/Quantum Systems.

Hybrid Classic/Quantum Systems define a class of systems that combine classic and quantum computers to solve a problem. The main advantage of this approach is the exploitation of classic computers for specific tasks (e.g., error correction, data encoding) and capabilities provided by quantum machines [24].

Hybrid quantum systems are depicted in Fig. 1: in step 1, data d are encoded on the classic system for execution on the quantum system. At this stage, different data encoding techniques [26] can be applied to transform input in a quantum state. Also, high level circuit description C is transpiled [27]; in step 2, quantum state ψ is prepared based on encoded data d'. In step 3, circuit C' is executed with the given input. Execution and measurement (step 4) are performed s time, due to the intrinsic nondeterminism of quantum computation, in order to create a probabilistic distribution of the output. Finally, in step 5, postprocessing of the output o is applied. Postprocessing can range from error correction [23] to noise mitigation [22], to address limitation of current NISQ (Noisy Intermediate Scale Quantum) machines.

3 Use Case: Computer-Assisted Drug Design (CADD)

Traditional drug design is a long, complex and costly process, with huge impact on pharmaceutical companies profit, as shown by the recent pandemic. Computer-Aided Drug Design (CADD) provides a variety of tools and methods that assist in the various stages of drug design, i.e., (1) discovery of a candidate drug, (2) evaluation of efficacy and safety and (3) drug-target interactions simulations. Recent development of quantum computing affected many branches of scientific computing, including CADD [7].

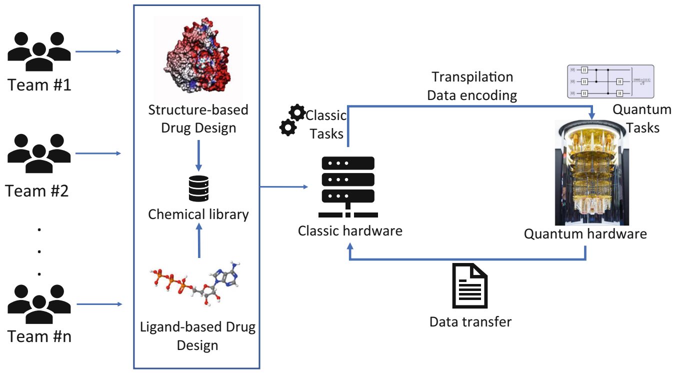

Fig. 2. A Hybrid CADD Workflow.

Figure 2 shows an example CADD hybrid workflow exploiting quantum machines. CADD is a multidisciplinary process, which involves experts from biochemistry, molecular biology, cell biology, bioinformatics and HPC scientists. To this end, typical CADD systems allow collaborations of different teams [18]. Also, CADD involves different families of techniques, namely structure-based drug discovery [3] and ligand-based drug discovery [13]. ML-based techniques are widely applied in both cases [18,19].

All identified CADD techniques require processing of large datasets, such as ChEMBL (https://www.ebi.ac.uk/chembl/) and DrugBank (https://go.drugbank.com/). Also, CADD relies on accurate simulation of target interactions, as well as computation of ground state for weakly-interacting molecules. Such simulation rely on molecular dynamics (MD) simulations and require a huge amount of computational resources, which is provided by remote distributed clusters [15]. However, as proven by [4,10], MD simulations are a very common use case of quantum computing, due to the fact that molecular interactions can be easily modelled by quantum mechanics. Moreover, different computation that are employed by CADD, such as eigenstates calculation, approximate optimization and quantum machine learning, can be easily mapped to Variational Quantum Algorithms, the main candidates to achieve the quantum advantage [8].

3.1 Current Limitations of Hybrid Classic/Quantum CADD

From our analysis, we conclude that executing CADD on hybrid classic/quantum systems can not only allow to overcome the scalability limits of HPC systems, but also to simulate processes that cannot be easily simulated by classic HPC [7]. However, several additional layers between HPC and quantum systems are necessary to enable a seamless integration between classic and quantum nodes, to address challenges arising from their integration.

First of all, current quantum hardware requires very specific facilities, which are not accessible to common data centers. This causes a higher degree of geographical distribution, which hinders latency of communication between classic and quantum systems. Moreover, due to the need of continuous data exchange between quantum and classic hardware, a hardware/software layer focused on encoding data coming from datasets/sensors for their use on quantum hardware, and for streaming results from quantum hardware to HPC systems for further processing, would significantly speed up data processing in hybrid classic/quantum systems. This additional layer could also be used for filtering and preprocessing data to further improve efficiency of streaming data. In the next sections, we will analyze some challenges of hybrid classic/quantum systems, coming from our analysis of CADD.

4 Challenges

4.1 Efficient Data Encoding

CADD requires continuous interaction with different pharmaceutical datasets, in order to evaluate target interactions. Information coming from different databases has to be encoded into a quantum state to enable processing by different quantum algorithms (Sect. 2.2).

Efficiency of data encoding is critical for CADD, since it might constitute a bottleneck in the quantum processing and potentially cancel the effect of a quantum speedup [26]. This is specifically true if we consider the geographical distribution of quantum hardware, which requires very specific facilities to be executed, usually not available in typical HPC infrastructures.

Moreover, considering the limited availability of quantum hardware, data needs to be filtered, aggregated and pre-processed at different system layers, to avoid wasting precious computational resources.

4.2 Security and Privacy

In CADD, privacy is a huge concern, since pharmaceutical companies are usually not willing to share their intellectual property and lose a competitive advantage on their competitors. As a consequence, different privacy-preserving and secure schemes for drug design have been developed to allow pharmaceutical companies to collaborate on the same datasets and or computational clusters without sharing private information [17].

Moreover, execution on quantum nodes requires to (1) encode data coming from different databases, (2) translate high level description of a quantum circuit for the target quantum machine, and (3) send both data and circuit over the network to be executed on the quantum machine. Communication between classic and quantum hardware could be exposed to privacy and security leaks, which might have disastrous effects for the companies involved in CADD.

Considering the huge development of quantum drug design [7], and its importance for pharma industry [2], it is important to guarantee secure and privacy-preserving multi-party interactions between classic and quantum hardware to ensure large-scale adoption of quantum technology.

4.3 Classic/Quantum Integration

As mentioned in [21], simulations of molecular dynamics for CADD processes can be executed on quantum computing, allowing either to speed up specific calculations or even to perform simulations that cannot be executed in classic systems. Also, development of Quantum Machine Learning [6] offers new possibility to ML-based drug discovery methods [9].

However, allowing such hybrid execution requires continuous communication between classic and quantum hardware. To enable such communication, there is a need of an intermediate layer capable to perform the translation between the two types of architectures. Also, considering the high heterogeneity of quantum hardware, each quantum task has to be allocated on the quantum node that guarantees the best performance.

5 The Quantum Edge

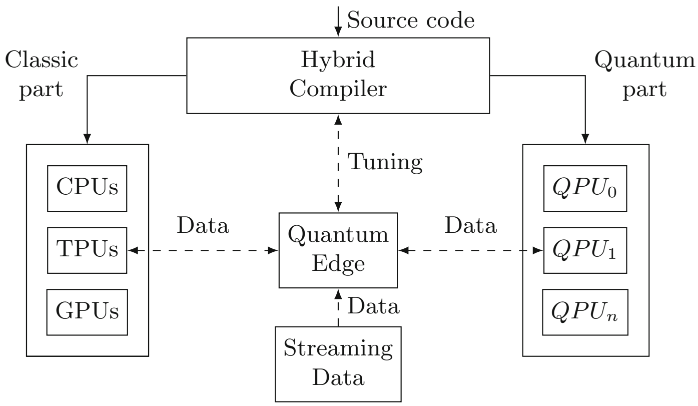

Fig. 3. The Quantum Edge.

We advocate the use of Edge computing to address challenges identified in Sect. 4. Edge computing is a computing model that works by placing processing logic in close proximity to data sources to reduce latency of their processing. The use of Edge nodes allows mitigation of effects of geographic distribution of quantum nodes. Also, it allows preprocessing and filtering of data at different system layers, improving data processing efficiency. In this section, we expand our concept of the Quantum Edge [12] and adapt it to CADD.

Our vision is summarized in Fig. 3. Since only very specific part of computation can be executed on quantum nodes, we employ a hybrid compiler, similar to Qiskit and PennyLane. Hybrid compiler will generate both classic and quantum part, which will be executed respectively in the classic and the quantum hardware. During execution, Edge infrastructure will be responsible for (1) secure and privacy-preserving encoding of streaming input data for the use on classic and quantum part, (2) offload data processing on different classic and quantum hardware, (3) collect data about execution on both classic and quantum hardware to fine-tune execution of both quantum and classic part.

5.1 Efficient Data Encoding

Since quantum machines' availability is limited in comparison to classical machines, allocating fixed resources to data encoding/decoding might result in underutilization of computational resources. Edge computing could be used not only to enable low-latency filtering, aggregation and pre-processing of data, but also to apply data encoding at lower layers of the network allowing their usage at different systems' layers. This model is particularly powerful for the processing of streaming data coming from IoT devices, since it allows developers to not worry about capacity planning, configuration and management of underlying infrastructure. Also, it increases the system elasticity, preventing under/over provisioning of classical infrastructure. Serverless paradigm could be applied to develop data encoding/decoding methods required by different quantum algorithms at the Edge. This choice will allow users to invoke required encoding/decoding methods on-demand, reducing infrastructure overhead and processing latency due to the use of Edge infrastructure.

5.2 Security and Privacy

The efficiency of Edge computing in executing data-intensive workflow computations has been proven by different works. For example, in [11], it is described how executing data-intensive tasks at the Edge allows to significantly speed up execution of scientific workflows. However, in a collaborative environment such as the scenario described in Fig. 2, it is also important to guarantee privacy and security, to protect intellectual property of different teams.

Edge is a natural solution in this context, due to the proximity of edge devices to data sources, which ensures an additional level of data protection by preventing user data to be moved to remote Cloud resources. This guarantees the required level of confidentiality, i.e., the the act of preventing unauthorized entities from reading or accessing sensitive materials [20]. In [25], it is shown how to apply privacy-preserving schemes to Quantum Federated Learning, which can be improved by means of Edge AI.

5.3 Classic/Quantum Integration

Edge nodes could act as intermediate layers in this scenario, either extracting ML-based models from performance data to identify the most suitable node or by performing different data pre-processing and filtering to enable data exchange between the different nodes.

In order to select the most suitable hardware for each quantum algorithm, we need to be able to predict performance of different computations on different quantum hardware. Considering the high variability of performance of quantum hardware, as well as the lack of datasets available, it is important to enable continuous collection of performance data of quantum and automatically tune the model while new performance data arrive. One might use machine learning (ML) on benchmarks suites such as [16] to train a performance model for quantum machines. Data collected can be used by on-line methods, which work by continuous improvement of a model using data coming from measurement, to address the limited amount of data available on quantum execution.

6 Related Work

Computer Assisted Drug Design has a wide application in pharma industry, as discussed in [2]. Most common methods in drug design are structure-based drug design [3] and ligand-based drug design [21]. AI-based methods are instead discussed in [18,19]. Applications of HPC to of drug design, molecular dynamics are described in [15].

Research efforts in the integration of Non-Von Neumann architectures in HPC systems are summarised in works such as [12,24], without considering drug design. Application of quantum computing to scientific applications are discussed in different works, such as [10] in the context of molecular dynamics simulations, [6] in the context of quantum machine learning , showing opportunities and limitation of applying quantum machine learning techniques to different problems, and [25] for federated learning, which is of particular interest for the Edge computing domain. Quantum Drug Design has been extensively discussed by [7], where different technological challenges of the NISQ hardware are described, but challenges related to hybrid classic/quantum systems and integration of classic and quantum hardware are not considered. Different applications of Variational Quantum Algorithms are described in [8], without considering HPC and drug design. Application of quantum computing to structure-based drug design are discussed by [4], while applications of quantum to ligand-based drug design are described in [13], without considering interactions with HPC systems.

7 Conclusion and Future Work

In this work, we lay the foundations for our concept of the Quantum Edge. First, we describe the challenges of executing scientific applications on hybrid classic/quantum systems, focusing on a CADD use case. Based on the specific challenges identified in the analysis, we describe how Edge computing could be applied to address identified challenges. In the future, we plan to investigate similar challenges in different scientific applications.

Acknowledgements. This work has been funded through the FFG Flagship project HPQC (High Performance Integrated Quantum Computing) # 897481.

References

1. Pegasus workflow management system. https://pegasus.isi.edu/
2. Research informatics and in-silico drug discovery. https://www.evotec.com/en/execute/drug-discovery-services/research-informatics-in-silico-drug-discovery. Accessed 04 Mar 2023
3. Anderson, A.C.: The process of structure-based drug design. Chem. Biol. **10**(9), 787–797 (2003)
4. Andersson, M.P., Jones, M.N., Mikkelsen, K.V., You, F., Mansouri, S.S.: Quantum computing for chemical and biomolecular product design. Curr. Opin. Chem. Eng. **36**, 100754 (2022)
5. Ang, J.A., Barker, K.J., Vrabie, D.L., Kestor, G.: Codesign for extreme heterogeneity: integrating custom hardware with commodity computing technology to support next-generation HPC converged workloads. IEEE Internet Comput. **27**(1), 7–14 (2023)
6. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S.: Quantum machine learning. Nature **549**(7671), 195–202 (2017)
7. Blunt, N.S., et al.: Perspective on the current state-of-the-art of quantum computing for drug discovery applications. J. Chem. Theory Comput. **18**(12), 7001–7023 (2022)
8. Cerezo, M., et al.: Variational quantum algorithms. Nat. Rev. Phys. **3**(9), 625–644 (2021)
9. Chan, H.S., Shan, H., Dahoun, T., Vogel, H., Yuan, S.: Advancing drug discovery via artificial intelligence. Trends Pharmacol. Sci. **40**(8), 592–604 (2019)
10. Cranganore, S.S., Maio, V.D., Brandic, I., Do, T.M.A., Deelman, E.: Molecular dynamics workflow decomposition for hybrid classic/quantum systems. In: 18th IEEE International Conference on e-Science, e-Science 2022, Salt Lake City, UT, USA, 11–14 October 2022, pp. 346–356. IEEE (2022)
11. De Maio, V., Kimovski, D.: Multi-objective scheduling of extreme data scientific workflows in fog. Futur. Gener. Comput. Syst. **106**, 171–184 (2020)
12. De Maio, V., Aral, A., Brandic, I.: A roadmap to post-moore era for distributed systems. In: Proceedings of the 2022 Workshop on Advanced Tools, Programming Languages, and PLatforms for Implementing and Evaluating Algorithms for Distributed Systems, pp. 30–34. ApPLIED 2022, Association for Computing Machinery (2022)

13. Ezz Eldin, R.R., et al.: Ligand-based design and synthesis of n'-benzylidene-3, 4-dimethoxybenzohydrazide derivatives as potential antimicrobial agents; evaluation by in vitro, in vivo, and in silico approaches with SAR studies. J. Enzyme Inhib. Med. Chem. **37**(1), 1098–1119 (2022)
14. Hansson, T., Oostenbrink, C., van Gunsteren, W.: Molecular dynamics simulations. Curr. Opin. Struct. Biol. **12**(2), 190–196 (2002)
15. Kutzner, C., et al.: GROMACS in the cloud: a global supercomputer to speed up alchemical drug design. J. Chem. Inf. Model. **62**(7), 1691–1711 (2022)
16. Li, A., Stein, S., Krishnamoorthy, S., Ang, J.: QASMBench: a low-level quantum benchmark suite for NISQ evaluation and simulation. ACM Trans. Quant. Comput. **4**(2), 1–26 (2023)
17. Liu, X., Deng, R.H., Choo, K.K.R., Yang, Y.: Privacy-preserving outsourced support vector machine design for secure drug discovery. IEEE Trans. Cloud Comput. **8**(2), 610–622 (2018)
18. Moingeon, P., Kuenemann, M., Guedj, M.: Artificial intelligence-enhanced drug design and development: toward a computational precision medicine. Drug Discov. Today **27**(1), 215–222 (2022)
19. Palazzesi, F., Pozzan, A.: Deep learning applied to ligand-based. In: Heifetz, A. (ed.) Artificial Intelligence in Drug Design. MMB, vol. 2390, pp. 273–299. Springer, New York (2022). https://doi.org/10.1007/978-1-0716-1787-8_12
20. Ranaweera, P., Jurcut, A.D., Liyanage, M.: Survey on multi-access edge computing security and privacy. IEEE Commun. Surv. Tutorials **23**(2), 1078–1124 (2021)
21. Ryde, U., Soderhjelm, P.: Ligand-binding affinity estimates supported by quantum-mechanical methods. Chem. Rev. **116**(9), 5520–5566 (2016)
22. Shaib, A., Naim, M.H., Fouda, M.E., Kanj, R., Kurdahi, F.: Efficient noise mitigation technique for quantum computing. Sci. Rep. **13**(1), 3912 (2023)
23. Sivak, V.: Real-time quantum error correction beyond break-even. Bull. Am. Phys. Soc. (2023)
24. Stein, S.A., et al.: A hybrid system for learning classical data in quantum states. In: 2021 IEEE International Performance, Computing, and Communications Conference (IPCCC), pp. 1–7. IEEE (2021)
25. Watkins, W.M., Chen, S.Y.C., Yoo, S.: Quantum machine learning with differential privacy. Sci. Rep. **13**(1), 2453 (2023)
26. Weigold, M., Barzen, J., Leymann, F., Salm, M.: Data encoding patterns for quantum computing. In: Proceedings of the 27th Conference on Pattern Languages of Programs, pp. 1–11 (2020)
27. Younis, E., Iancu, C.: Quantum circuit optimization and transpilation via parameterized circuit instantiation. In: 2022 IEEE International Conference on Quantum Computing and Engineering (QCE), pp. 465–475. IEEE (2022)