

Heterogeneous Computing from Managed Runtime Programming Languages: Opportunities and Challenges

Parallel programming APIs, particularly those designed for GPU and FPGA computing, are typically tailored to specific architectures and programming languages. Examples include (i) CUDA, which allows developers to harness GPU computing power, and (ii) OpenCL, which provides access to a wide range of hardware accelerators, such as GPUs and FPGAs. These APIs are predominantly built for the C/C++ programming languages and are closely linked to GPU architectures.

However, many developers do not primarily use C/C++. According to the TIOBE index, in addition to C and C++, popular languages in 2024 include Python, Java, and C#, along with other languages that run on managed environments (such as the Java Virtual Machine, JVM). This raises the question: how can developers using managed runtime languages access heterogeneous hardware like GPUs and FPGAs? What solutions exist, and what challenges do they face in doing so?

In this talk, I will explore key features of parallel programming models and discuss the challenges that runtime environments encounter when aiming to harness the performance of parallel hardware. I will also present the current research conducted at The University of Manchester, which addresses some of these challenges. This includes proposals that combine the expressiveness of parallelism with functional programming, data-centric approaches, and object-oriented paradigms to maintain high-level abstractions while effectively utilising parallelism.

Short Bio

Juan Fumero is a Research Fellow working as part of the Advanced Processor Technologies (APT) Research Group at The University of Manchester. His research focuses on new approaches to bridge the gap between accessibility and exploitation of hardware accelerators such as GPUs and FPGAs, and programmability from managed runtime programming languages. He is also the main software architect of the [TornadoVM](#) project (a Java framework for automatically running JVM applications on heterogeneous hardware).

Furthermore, Juan Fumero participates in standardisation activities for the oneAPI Hardware Group, and the oneAPI Language Group to help shaping the evolution of oneAPI for managed runtime programming languages.

Juan frequently gives talks and seminars at various industry and academic events, such as DevOxx, JVMLS, JavaZone, among many others. As a result of his work, the University of Manchester awarded Juan the Best Research Outcome in 2022, and he has been recognized by Intel as an Intel Innovator.

Full Bio: <https://jifumero.github.io/>