

Best Practice Guide - Modern Accelerators

Hardware accelerators offer certain advantages over general-purpose Central Processing Units (CPUs) as they provide a greater computational throughput when applications exhibit high degrees of data parallelism, due to their highly parallel architectural design and high-bandwidth memory systems. These specialized chips are usually more energy-efficient, i.e. are capable of delivering much higher compute performance as compared to CPUs at the same power cost. These characteristics make accelerator technologies appealing both for system vendors and users. This PRACE Best Practice Guide (BPG) on "**Modern Accelerators**" follows the style of previously published guide on "Modern Processors" (<u>https://prace-ri.eu/training-support/best-practice-guides/modern-processors/</u>), and provides an update on certain contemporary accelerator technologies, again, in a hybrid fashion of a field guide and a textbook.

More specifically, this guide starts with introducing the architectures of considered accelerator technologies, namely:

- Graphics Processing Units (GPUs)
- Field-Programmable Gate Arrays (FPGAs)
- Vector processors

This is then followed by a discussion on their suitability for different HPC applications, and a description of their accompanying emerging programming models (e.g. CUDA, SYCL, HIP, etc.) and development environments. This BPG then outlines certain hints and best practices for application porting and tuning built upon the available literature and the expertise of authors involved.

Finally, the guide provides a brief information on the available hardware infrastructure at PRACE HPC sites featuring the discussed accelerator technologies, and concludes with a description of the case applications used.

The **PRACE Best Practice Guide** on **"Modern Accelerators**" can be accessed via the following link: <u>https://prace-ri.eu/training-support/best-practice-guides/modern-accelerators/</u>

<u>KEYWORDS</u>: accelerators; GPUs, FPGAs, vector processors, programming models, application porting and tuning, accelerator suitability, performance analysis