APE – Tflops Computers for Theoretical Particle Physics

Karl Jansen[1], Norbert Paschedag[2], Dirk Pleiter[1], Hubert Simma[2] and Peter Wegner[2]
[1](John von Neumann Institute for Computing NIC/DESY, D-15735 Zeuthen, Germany)
[2](Deutsches Elektronen-Synchrotron DESY, D-15735 Zeuthen, Germany)

Abstract

The commissioning of several large installations of APEmille computers in Europe will have been finished in autumn 2001. All these machines together make another 2 Tflops of computing power available for numerical simulations in theoretical particle physics.

Keywords: parallel computing, lattice gauge theories, custom hardware

1 Introduction

For many problems in modern particle physics lattice gauge field theory offers the only known way to compute various quantities from first principles. Much progress has been made during recent years, e.g., in calculating the light hadron spectrum, the light quark masses, the running coupling constant $\alpha_{\rm s}$, or leptonic decay constants like $f_{\rm B}$. Further-on, lattice simulations allow the study of non-perturbative phenomena like chiral symmetry breaking, confinement, or studies of phase transitions in the early universe. (See [1] for an overview.) However, progress in this field is severely limited by the tremendous amount of required computer power. In order to make these resources available, various research groups engage in the development of supercomputers which are specifically optimized for their applications. One of the most successful projects is the array processor experiment (APE). The APE100, the second generation of APE supercomputers developed at the Instituto Nazionale di Fisica Nucleare (INFN), has been the leading workhorse of the European lattice community since the middle of the 1990s. Once the commissioning of several large European installations of APEmille computers will be have been completed, another 2 Tflops of computing power will be available. This third generation of APE machines has been developed by INFN in collaboration with DESY.

2 Features of the APEmille architecture

APEmille is a massively parallel computer which is primarily optimized for simulating QCD. The architecture is SIMD and all nodes run strictly synchronous at a clock frequency of 66 MHz. The communication network has a three-dimensional topology and offers a bandwidth of 66 MBytes/sec/node. The smallest APEmille unit is a processing board (see Fig. 1) with $2\times2\times2$ nodes. The largest stand-alone systems built until now consist of $4\times8\times8$ nodes. Three different ASICs have been custom developed for APEmille. Program execution is controlled by a control processor, which also performs the global integer arithmetics. Computations using local integers and all floating point operations are done by computing nodes. At each cycle they are able to complete the operation $a+b\times c$, where a, b, c are single precision complex operands. This gives a peak performance of 528 Mflops per node. Each of them has 32 MBytes of local memory and a very large number of 512 registers. Remote communication is implemented as direct memory access and controlled by the communication processors.

The processors are controlled by very long instruction words (VLIW). This allows efficient scheduling of the micro-code at compile time. Much effort has therefore been put into the development of software tools for generating efficient code. Application programs are written in the high-level language TAO. This language is very well suited for writing codes for QCD simulations. Loading of the executables and all other operating system services are handled via PCs running Linux. One host PC per four boards is directly attached to the APEmille

backplane. It uses a PCI bus to communicate with the processing boards. The system services are controlled by one master PC per machine (see Fig. 1).

3 Operating APEmille

DESY Zeuthen will install four APEmille machines with a total peak performance of 550 Gflops. Each machine consists of 256 nodes and 8 host PCs. Because of the very moderate power consumption of 3.7 kW per machine, i.e. less than 30 W/Gflops, air cooling is sufficient. During more than half a year since the first machine entered production status we did not encounter a single major problem. Normally up-times of 1-2 months or more can be achieved. Hardware maintenance is typically limited to simple replacement of aging modules and is therefore cheap, both in terms of hardware costs and manpower.

Scientists can apply for computer time at the John von Neumann Institute for Computing (NIC). All current projects focus on simulation of QCD with dynamical fermions. Because of the limits in available computing power most of them were until then forced to apply the so-called quenched approximation, i.e. the effects of vacuum fermion loops had been neglected. However, with the currently available computing resources it will not be possible to lower the masses of the dynamical quarks. It will be even more difficult to reduce the lattice spacing and to do simulations closer to the continuum limit.

4 Outlook

APEmille allowed a major upgrade of computer power available to the European community of lattice gauge field theory. For many groups this opens the door to simulations of QCD with dynamical fermions. From our experiences running APEmille we expect it to become as reliable as its precursor. The development of a future generation of APE computers, apeNEXT, has been started [2] and a prototype is planned to be running by end of 2002.

References

[1] T. Bhattacharya *et al.*, "Contents of Lattice 2000 Proceedings," Nucl. Phys. B (Proc. Suppl.) 94 (2001) [hep-lat/0104009].

[2] R. Alfieri *et al.* [APE Collaboration], "apeNEXT: A multi-Tflops LQCD computing project," [hep-lat/0102011].

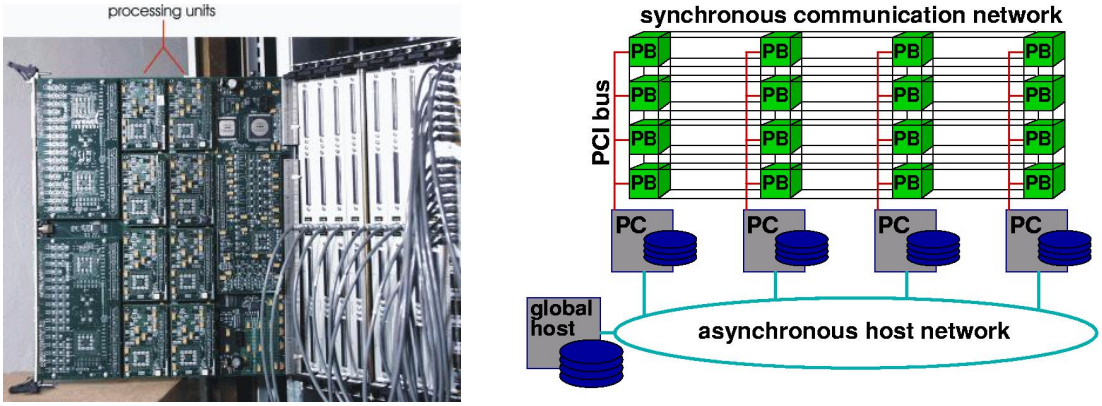

Figure 1: An APEmille processing board (left) and the architecture of a 128 node machine.