

Profiling for the APinTA-PDEs project

David Acreman

July 19, 2021

Contents

1	Introduction	2
2	Target platforms	2
3	Profiling tools	3
3.1	Cray PAT	3
3.2	ARM Forge	4
3.3	Intel Parallel Studio	4
3.4	Profilers summary	4
4	Builds	4
5	Conclusions and recommendations	4

1 Introduction

This document will become the profiling report deliverable. Profiling is deliverable 5.1.1 and builds are deliverable 5.1.2. This report primarily focusses on Firedrake but the profiling tools evaluated could be used for profiling other codes used by the project.

2 Target platforms

The target HPC platforms for the project are Archer-2 (national tier-1), Isambard (national tier-2) and Isca (local tier-3). Isambard has two separate systems based on the ARM64 architecture: the XCI system which is an established production system, and the A64FX system which is a newer system. The A64FX system should be considered a stretch objective as the hardware and software are less well known to us. We have also purchased a local server which is designed to be similar to an Archer compute node, although it will run Ubuntu which will make it easier to build Firedrake. The server is in some sense a reference installation as it should use an unmodified Firedrake install script. Installations on other platforms will use non-standard components (e.g. MPI and maths libraries). An overview of the hardware in the target platforms is shown in Table 1.

System	Processor	Cores/node	Memory/node	Interconnect
Archer-2	AMD x86_64	128	256 GB DRAM	HPE Cray Slingshot
Isambard XCI	Thunder X2 ARM64	64	256 GB DRAM	Cray Aries
Isambard A64FX	Fujitsu A64FX ARM64	48	32 GB HBM	Mellanox Infiniband
Isca	Intel x86_64	16/20	128 GB DRAM	Mellanox Infiniband
Server	AMD x86_64	128	256 GB DRAM	None

Table 1: Target platform hardware specifications. The Isambard A64FX platform has high bandwidth memory (HBM) which should give better performance than DRAM but has a smaller capacity.

An overview of the software stacks on the target platforms is shown in Table 2. The common factor in the compilers is the GNU compiler which is available on all platforms. Each HPC platform has a compiler from the processor vendor (AMD, Intel, ARM and Fujitsu) and Cray systems also have the Cray compiler. On the Isambard A64FX system some modules are only available after

System	Compilers	MPI libraries	Maths libraries	Profilers
Archer-2	Cray, GNU, AOCC	Cray MPICH2	Cray libsci	Cray PAT
Isambard XCI	Cray, GNU, ARM	Cray MPICH2	Cray libsci	Cray PAT, ARM Forge
Isambard A64FX	Cray, GNU, ARM, Fujitsu	Cray MVAPICH2	Cray libsci	ARM Forge
Isca	GNU	OpenMPI	OpenBLAS	None
Isca	Intel	Intel	Intel MKL	Intel Parallel Studio
Server	GNU	MPICH?	???	None

Table 2: Software stacks on target platforms. AOCC is the AMD Optimizing Compiler Collection. Intel MPI and MVAPICH are MPICH derivatives which support an Infiniband interconnect. The A64FX system did not appear to have a perftools module.

running

```
module use /lustre/software/aarch64/modulefiles
```

On the Cray systems compilation is handled by wrapper scripts from the Cray Programming Environment. The wrapper script can run different compilers depending on which programming environment module is loaded at compile time (e.g. `PrgEnv-cray` calls the Cray compiler and `PrgEnv-gnu` calls the GNU compiler). The MPI library and maths library are then linked by the

wrapper script. On Cray systems the standard profiling tool is Cray Performance Analysis Tools (PAT) and on Isambard there is also the ARM Forge profiler¹.

The software environment on Isca is managed using Easybuild which has the concept of a toolchain. There are two toolchains on Isca: the GCC-foss toolchain and the Intel toolchain. Each toolchain has a different MPI library and maths library. Intel Parallel Studio has an MPI profiling tool called Intel Trace Analyzer and Collector (ITAC) which needs to use Intel MPI.

Recommendation: standardise on the GNU compilers for all platforms. This avoids having to manage too many different builds and takes the compiler out of the performance equation (different GCC versions notwithstanding).

3 Profiling tools

Within Firedrake there are two ways to obtain profiling information:

- Add a PyOP2 timed stage
- Request profiling information from PETSc

These options will be available on all platforms when Firedrake is the target application.

There are also external profilers which are already installed on the target platforms. Different platforms have different profilers available:

- ARM/Allinea Forge
- Cray PAT (Cray systems only)
- Intel Parallel Studio (Intel MPI only)

Plan for testing external profilers:

1. Test with a simple example e.g. Mandelbrot
2. Test with Python

See whether the profiler can spot load imbalance vs. overhead and identify the MPI call responsible.

We could get an evaluation licence for ARM Forge and/or Intel Parallel Studio on the server. We could also look at open source profiling software for the server.

3.1 Cray PAT

Although this is a Cray tool it does work with compilers other than Cray (the other PrgEnv modules load perftools-base). Experience building SWIFT showed that the `perftools-lite` module can confuse autogen and configure so only load the module prior to the make command.

- Cray PAT has two modes: standard and “lite”
- There is a graphical viewer (Apprentice 2) which reads the profiling output
- Can view a time line from a full trace

See also <https://docs.nersc.gov/tools/performance/craypat/>.

It’s not clear how to use Cray PAT with python/Firedrake. Loading the `perftools-lite` module and using the python from the `cray-python` module does not produce any profiling output. **Is there a limit on the number of MPI processes? If so I’ve not seen it.**

¹ARM Forge was previously known as Allinea Forge

3.2 ARM Forge

ARM MAP is the profiler which is part of the ARM Forge suite and it can profile C++, C, Fortran and Python². **Is there a limit to how many MPI processes can be used?**

3.3 Intel Parallel Studio

Previously working on Isca with the Intel build. Intel MPI is required for the Trace Analyzer and Collector profiler to work. **Is there a limit on the number of MPI processes? If so I've not seen it.**

3.4 Profilers summary

Table 3 summarises the capabilities of the profiling tools reviewed in this section.

Profiler	Max MPI procs	C/C++	Fortran	Python
Cray PAT	Unlimited	✓	✓	
ARM Forge	??	??	??	??
Intel Parallel Studio	??	??	??	??

Table 3: Summary of profiling tools

4 Builds

Plan:

1. Working builds on main target platforms
2. Optimised builds on main target platforms: use optimised maths libraries but beware threading

Platform	Build status	Profiling status
Archer2	??	
Isambard XCI	working build	
Server	??	
Isca	??	
Isambard A64FX	??	

Table 4: Status of Firedrake builds and profiling for the target platforms

- Isambard build of Firedrake: what's happening with BLAS/Lapack etc? Scalapack gets built as part of the PETSc build but is also available from `libsci`

5 Conclusions and recommendations

²<https://www.arm.com/products/development-tools/server-and-hpc/forge/map>