

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363539552>

# Celebrating 50 years of computer benchmarking and stress testing

Preprint · September 2022

CITATIONS

0

READS

215

1 author:



[Roy Longbottom](#)

UK Government

68 PUBLICATIONS 82 CITATIONS

SEE PROFILE

# **Celebrating Computer Benchmarking and Stress Testing**

## **1972 to 2022**

**Benchmarks**

**Stress Tests**



**From Supercomputers to Microcontrollers**



**Cray 1 1979**  
80 MHz, \$7M



**Raspberry Pi Pico W**  
125 MHz, \$6



From 1972 to 2022 I produced and ran computer benchmarking and stress testing programs. The Whetstone Benchmark, for which I became the design authority, also covered exactly the same time span.

### **Stress Tests 1972 to 1980**

I was a member of the engineering support group of the UK Government Central Computer and Telecommunications Agency (CCTA) from 1960 until my early retirement in 1993. For a long period I was responsible for designing and running contractually required acceptance trials for Government computers and those centrally funded for Universities. In the 1970s, major changes were required for stress testing under forthcoming Multiprogramming Operating Systems. For these, I wrote a series of 17 programs in Fortran (CPUs 5, disk drives 4, magnetic tape units 3 and one each for card readers and punches, paper tape readers and punches, and line printers). From 1972, these were used on many hundreds of acceptance trials up to 1990. Details of these trials, programs and some results are covered in my 1980 book "Computer System Reliability".

My hands on involvement included on site acceptance trials of latter day supercomputers IBM 360/85, IBM 360/195 and CDC 7600, in 1972. Then, in 1979/80, a Cray 1 and a CDC Cyber 205, for which I produced a new series of CPU benchmarks that compiled with automatic vectorisation. The latter two also had pre-delivery trials in the USA. Of these seven trials, three failed and required second trials, following appropriate delays. My stress testing programs were responsible for all of the failures, one due to an excessive number of CPU fault incidents, one due to reading the wrong files and the other due to the I/O interface not correctly transmitting my data patterns.

### **Benchmarks and Performance 1972 to 1993**

Besides for the Whetstone benchmark, I produced performance ratings from all CPU tests used during acceptance trials that I was involved in, covering 72 different processors.

From 1981 to 1987, I was mainly involved in dealing with performance of data processing systems, covering sizing, modelling, performance monitoring, general advice and attending user specified benchmarking sessions.

From 1987, I continued with data processing consultancy plus reinvolvement in University supercomputer activities, including acting as an independent adviser to a benchmarking NEC and Fujitsu systems, during 1992 in Japan.

## Whetstone Benchmark 1972 to 2022

The Whetstone benchmark was produced by my CCTA colleague Harold Curnow, with the initial official results obtained in 1972. The Fortran version became the first general purpose benchmarks that set industry standards of computer system performance. In its day, it was the equivalent of today's Geekbench. Then it was minicomputer manufacturers who said "Now who has the fastest computer", based on a single number score. I later took over design responsibility for this benchmark, but included some changes earlier. The main one was to produce performance measurements of the 8 test functions, to provide fairer comparisons and identify any underhand activities. I also produced versions for vector processors and multiprocessor systems. The benchmark was also run in all the acceptance tests and I continued collecting and reporting results until retirement. See <https://www.researchgate.net/project/Whetstone-Benchmark>.

## Later Benchmarks and Stress Tests

With my first access to a PC, in CCTA, I acquired IBM BASIC and Fortran compilers to produce new versions of the Whetstone benchmark. On retirement, with my own PC and C/C++ compiler and website, I started my PC Benchmark Collection in [roylongbottom.org.uk](http://roylongbottom.org.uk) (1996).

Initial benchmarks were mainly for measuring CPU performance, aimed at identifying best and worst characteristics, rather than a single number rating. These were followed by a number covering cache and RAM, then input output devices, networks and graphics. Stress testing programs, for all these areas, were produced as needed. All had parameters to specify running time, most with regular performance reports, to identify time related changes.

The website currently has 86 HTM reports and 75 compressed files containing benchmarks, source codes and descriptions, all FREE with no advertising. Also, 37 of these, or variations, are in PDF files at ResearchGate, along with benchmark codes.

Over the years I produced different and new versions of programs for 32 bit and 64 bit working, new compilers, new CPU architecture and multiple CPU cores. Each of these varieties covered a range of platforms including using Intel and ARM type processors, running via DOS, OS/2 (1997), and various flavours of Windows (e.g. XP 2005), Linux (2010), Android (2012) and Raspberry Pi (2013). Where appropriate, in each area there are around 20 benchmarking and stress testing programs, at both 32 bits and 64 bits.

Stress testing PCs started with those considering implications of overclocking then laptop performance issues. Next were those for evaluating newer technology heating effects, then advanced vector type architecture and unbalanced CPU arrangements (like big.LITTLE). My stress testing programs have parameters including for running time, alternative hardware, validation of correct results and regularly reporting performance (as opposed to a single result).

By 2019 I had produced a number of reports on Raspberry Pi 1, 2 and 3, including for comprehensive stress tests. Then, in 2019 (aged 84), I was recruited as a voluntary member of Raspberry Pi pre-release Alpha testing team. This has continued to the present time, leading to a further 9 PDF reports being produced, most based on those produced for the Alpha testing team. My latest, for Raspberry Pi Pico W, is described in my LinkedIn Posts. All are available from ResearchGate.