Software Test-Suite for Computer Energy Efficiency Measurement

Webinar

Andoni Arregui, Thomas Wucher, and Joan Roig

2022-09-20

GTD GmbH

Table of Contents

1. Testing Tool Architecture
2. Test Setup and Procedure
3. Test Worklet Analysis

GTD GmbH

Who we are and what we do:

- We are a software company in the south of Germany
- Part of the European group GTD with headquarters in Barcelona, founded in 1987

We mainly work for:

- European Space Agency (ESA)
- German Aerospace Center (DLR)
- German federal institutions (UBA, BAM)

Since 2019 we have our software in space

Testing Tool Architecture

Test Suite Overview

The main characteristics of the Test Suite developed by GTD are:

- It is based on the **open-source** Phoronix Test Suite
- It runs a **series** of ***worklets***[1]
- It executes **native binaries** in Microsoft Windows, MacOSX, and Linux (incl. ChromeOS)
- Measuring power **does not alter** the **performance** or the **power demand**
- All the results are fused together in a **single meta-efficiency metric** value

[1] A *worklet* is a set of tasks, placed in a *workflow*, that is, in a sequence of activities that can be easily reproduced on a regular basis.

Phoronix Test Suite Overview

Phoronix Test Suite[2] is an open-source framework for automated benchmarking:

- Runs on Microsoft Windows, Linux, MacOSX (written in PHP)
- Automates the whole testing process from dependency installation to test results aggregation
- Has a huge collection of predefined test-profiles (i.e., worklets)
- Easily extensible for our use-case (e.g., new worklets)
- Results can be visualized by Phoronix Result-Viewer or exported to CSV
- Support for power measurement devices
- No external connection is required to execute the worklets

[2] https://www.phoronix-test-suite.com/

Phoronix Test Suite Changes

Several changes were necessary to make Phoronix Test Suite work for our purpose:

- Support merging power data after running all tests
- Define and implement support for calculating efficiency for lower is better result units
- Support power measurements on Windows
- Support the result viewer on Windows
- Make the installation of PHP more robust on Windows
- Add better CSV exports
- Other small bug fixes and optimizations

Worklet Structure

Each worklet consists of the following parts:

- Test definition
- Results definition
- Downloads definition
- Installation script
- Run script

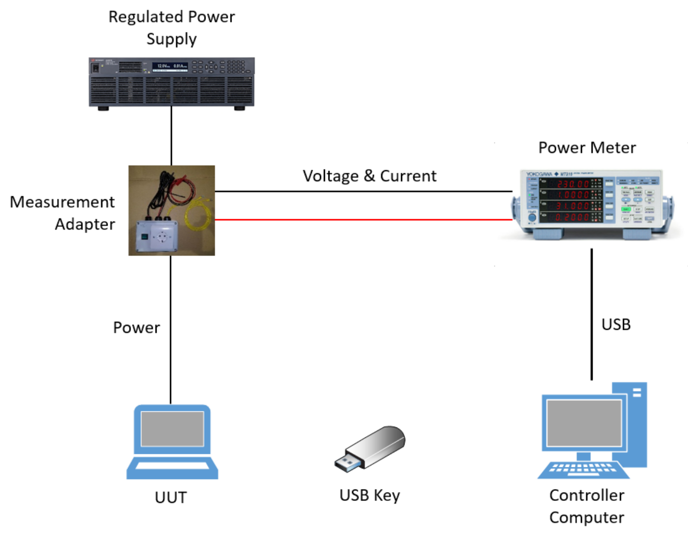
Definition files are written in XML, scripts are shell scripts.

Test-Suite Structure

A test suite defines:

- Series of worklets to be run
- Configuration options for each worklet

Test Setup


Regulated Power Supply
Measurement Adapter
Voltage & Current
Power Meter
Power
USB
UUT
USB Key
Controller Computer

Power Measurements

Reading data from power meters is done by using *sigrok*, an open-source signal analysis software suite.

- sigrok is cross-platform, supporting Windows, MacOSX, Linux
- sigrok is vendor-independant, supporting devices from different manufacturers
- We added support for
  - Yokogawa WT310(E), WT333, WT500
  - Hioki PW3335, 3332

Power meters connected via GPIB are read using a specifically created Python application using librevisa.

Software Components on controller computer

On the controller the following pieces of software are installed:

- PHP (and its dependency, the Visual C++ redistributable package)
- Driver for USB-connected power meters
- NTP server (for time synchronization)
- Phoronix Test Suite configuration (e.g. to configure the result viewer)

Software Components on EUT

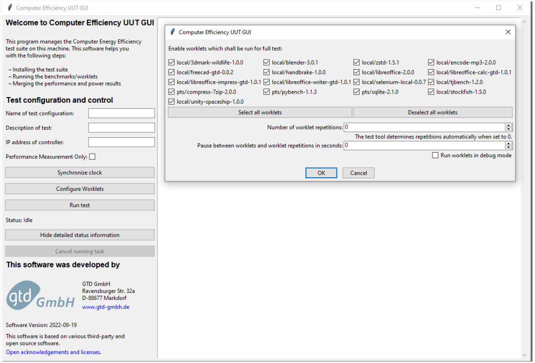
On the EUT the following pieces of software are installed:

- PHP (and its dependency, the Visual C++ redistributable package)
- NTP client (for time synchronization)
- Third party applications used for testing (Libreoffice, FreeCAD, Blender, ...)
- Cygwin (to run the worklet scripts, only on MS Windows)
- Phoronix Test Suite configuration
- Phoronix Test Suite
- All relevant worklets according to their definition files

Graphical User Interface

The GUI of the test software is

- written in Python and packaged with pyinstaller
- developed using the tkinter widget toolkit
- only a frontend to various shell and batch scripts



Welcome to Computer Efficiency UUT GUI

This program manages the Computer Energy Efficiency
test suite on this machine. This software helps you
with the following steps:

– Installing the test suite
– Running the benchmarks/worklets
– Merging the performance and power results

Test configuration and control

Name of test configuration:

Description of test:

IP address of controller:

Performance Measurement Only: □

Status: Idle

This software was developed by

GTD GmbH
Ravensburger Str. 32a
D-88677 Markdorf
www.gtd-gmbh.de

Software Version: 2022-09-19

This software is based on various third-party and
open source software.
Open acknowledgements and licenses.

Enable worklets which shall be run for full test:

☑ local/3dmark-wildlife-1.0.0
☑ local/freecad-gtd-0.0.2
☑ local/libreoffice-impress-gtd-1.0.1
☑ pts/compress-7zip-2.0.0
☑ local/unity-spaceship-1.0.0
☑ local/blender-3.0.1
☑ local/handbrake-1.0.0
☑ local/libreoffice-writer-gtd-1.0.1
☑ pts/pybench-1.1.3
☑ local/zstd-1.5.1
☑ local/libreoffice-2.0.0
☑ local/selenium-local-0.0.7
☑ pts/sqlite-2.1.0
☑ local/encode-mp3-2.0.0
☑ local/libreoffice-calc-gtd-1.0.1
☑ local/tjbench-1.2.0
☑ local/stockfish-1.5.0

Number of worklet repetitions: 0

The test tool determines repetitions automatically when set to 0.

Pause between worklets and worklet repetitions in seconds: 0

□ Run worklets in debug mode

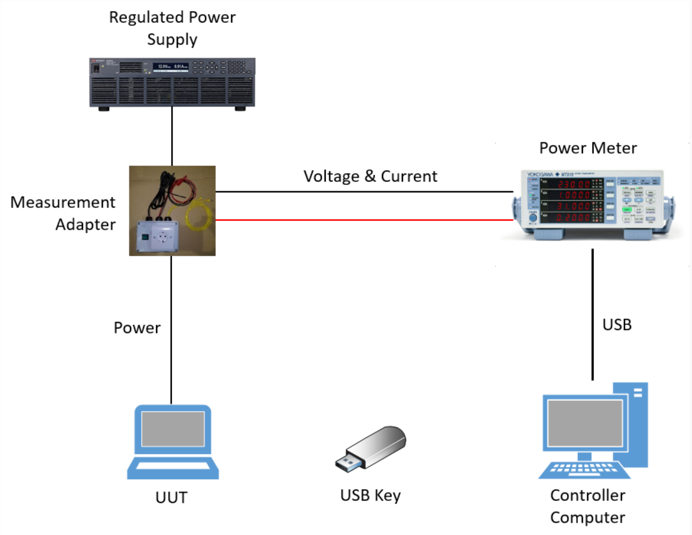
Command Line Interface

The CLI is implemented through scripts for each operating system to

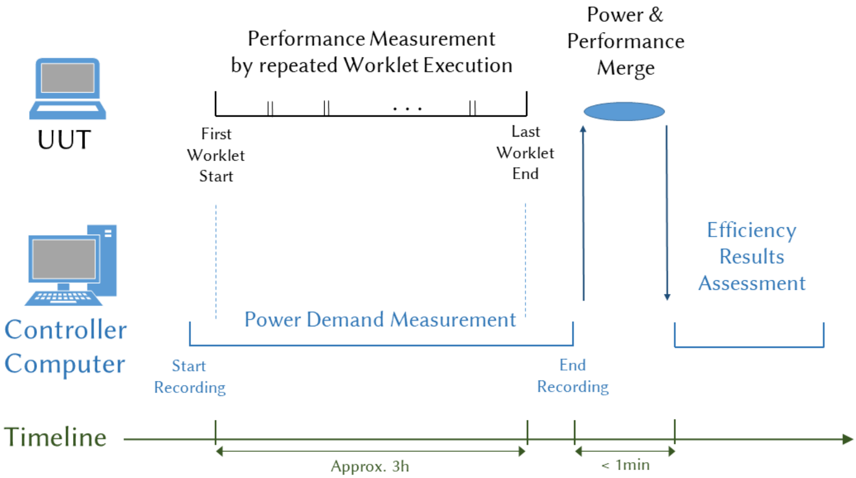
- Install the controller software
- Install the EUT software
- Check connection to the power meter
- Run the full benchmark
- Run the result viewer
- Start acquiring power measurements
- Start the NTP time server
- Synchronize the time to a given timeserver
- Uninstall the software
- Update an existing USB key to the latest version
- Run arbitrary PTS commands (development and debugging)

Test Setup and Procedure

Test Setup


Regulated Power Supply
Power Meter
Voltage & Current
Measurement Adapter
Power
USB
UUT
USB Key
Controller Computer

Test Suite Execution Workflow

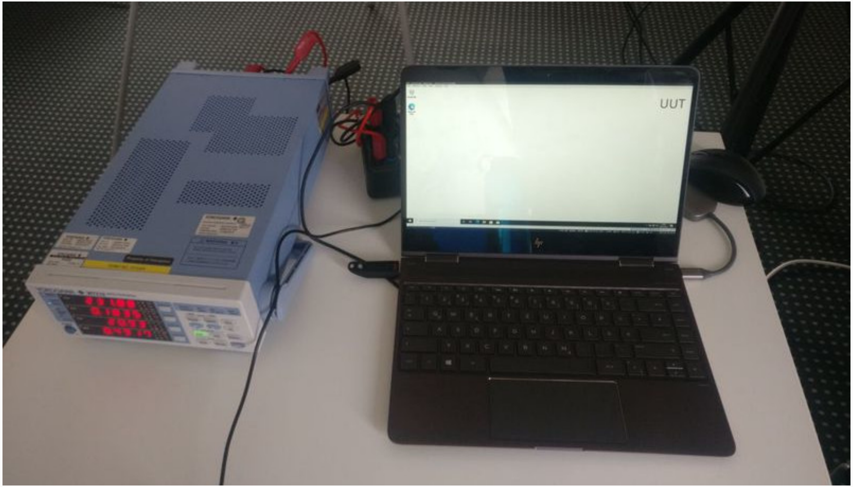


Test Setup and Procedure

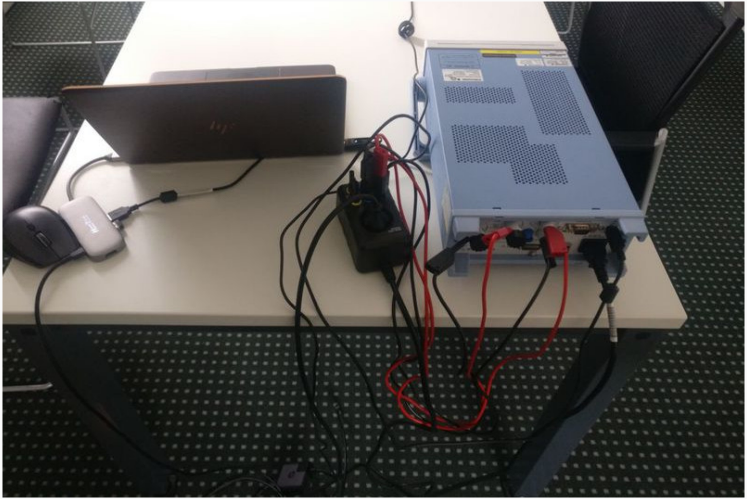
In a live demo we will show the following steps:

- Installation on the EUT
- Installation on the Controller
- Starting the software on Controller and EUT
- Synchronize the time
- Run a few selected worklets
- Merge the power and performance data
- Display the test results
- Uninstall the software

Demo Setup



Demo Setup



Test Worklet Analysis

Logic behind the worklets

- Re-using existing benchmark software
- Re-using existing real world applications
- Testing all relevant use cases for a computer
- Simple test setup (i.e. no complicated network setup needed)
- Open architecture for framework and worklets

Criteria for an existing worklet to be considered

- Supported on all relevant operating systems
- Free and Open-Source (exception possible)
- Actively maintained
- Simulates workloads of typical users
- Automation possible
- No additional testing setup needed
- Performance and power results are consistent and repeatable

Changes done to existing worklets

- Add support for missing operating systems
- Allow worklets to run offline
- Include as many dependencies as possible for offline install
- Automate the (un)installation as much as possible
- Check that all worklets run on all expected computer configurations

What worklets are available? (Rendering and CAD)

Blender

- Popular 3D modelling and video editing application
- Different 3D models available for testing
- Can render only on the CPU or via OpenCL, CUDA
- Result in seconds needed for the rendering

FreeCAD

- Newly developed worklet to simulate CAD usage profiles
- Uses an unmodified version of FreeCAD
- Creates a CAD model via FreeCAD’s API
- Result in seconds needed for creating the CAD model

What worklets are available? (Compression)

7-Zip Compression

- Compresses a file with the 7-Zip compression tool
- Result in million instructions per second (MIPS)

Zstd Compression and Decompression

- Compresses a file with the zstd compression algorithm
- Zstd is a more modern algorithm than used in other worklets
- Multi-threaded
- Result in MB per second of compression throughput

What worklets are available? (Office Tasks)

LibreOffice PDF conversion

- Converts several LibreOffice documents to PDF format
- Uses an unmodified version of LibreOffice
- Result in seconds needed for converting 20 documents

LibreOffice Spreadsheet/Presentation/Writer

- Newly developed worklet to measure spreadsheet performance
- Creates documents with various numbers, formulas, graphs, images, text
- Uses an unmodified verison of LibreOffice
- Result in seconds needed for creating a spreadsheet/presentation/text document

What worklets are available? (Web-Browsing)

Selenium

- Selenium is a browser automation tool, often used for testing websites
- Automates Google Chrome to measure Browser performance
- Runs the JavaScript benchmark Kraken
- Result in milliseconds for the average benchmark task

What worklets are available? (Artificial Intelligence)

Stockfish

- Open Source Chess Engine
- Plays several chess games using its neural network based AI
- Result in nodes per second

What worklets are available? (Image and Audio/Video processing)

tjbench

- JPEG compression/decompression benchmark part of libjpeg-turbo
- Measure performance of reading and writing images in JPEG format
- Result in megapixels per second

MP3 Encoding

- Encoding a file to MP3 format using lame encoder
- Result in seconds needed for encoding the test file

Handbrake

- Hardware accelerated video transcoding
- AMD, Apple, Intel and NVidia accelerators supported, with software encoding fallback
- Transcodes a H.264 encoded video to H.265
- Result in seconds needed for encoding the test file

What worklets are available? (Database)

SQLite

- Implements a simple file database
- Measures the performance of creating and querying a test database
- Can be configured to run with different amounts of threads
- Result in seconds needed for running a defined set of queries

What worklets are available? (Programming Languages)

PyBench

- Benchmark written in Python language
- Measures the performance of the Python interpreter on the system
- Average time needed for the test in milliseconds

What worklets are available? (3D Graphics)

3DMark Wildlife Extreme

- Graphics benchmark as known from the 3DMark benchmarks
- Renders several realistic scenes in 4K UHD resolution
- Result in average frames per second

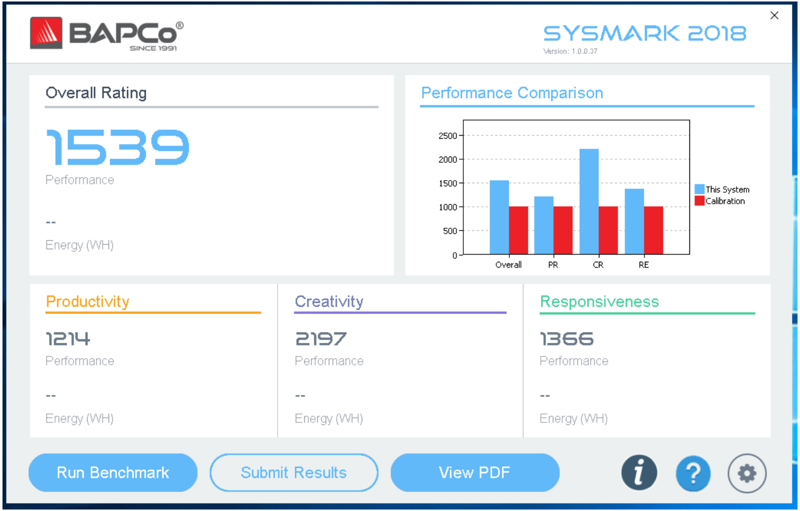
Unity Spaceship

- Graphics benchmark provided by Unity
- Renders several realistic scenes using the Unity engine
- Result in average frames per second

Comparison to SysMark

- SysMark benchmarks unmodified real-world applications
- Different scenarios for different usage profiles
- Focused on office and productivity workloads, e.g. no gaming
- Approximate run time is 1 hour
- Measures energy consumption
- Proprietary test framework and worklets
- Supports Windows only

SysMark Results



BAPCo® SINCE 1991

SYSMARK 2018

Version: 1.0.0.37

Overall Rating

1539

Performance

--

Energy (WH)

Performance Comparison

Productivity

1214

Performance

--

Energy (WH)

Creativity

2197

Performance

--

Energy (WH)

Responsiveness

1366

Performance

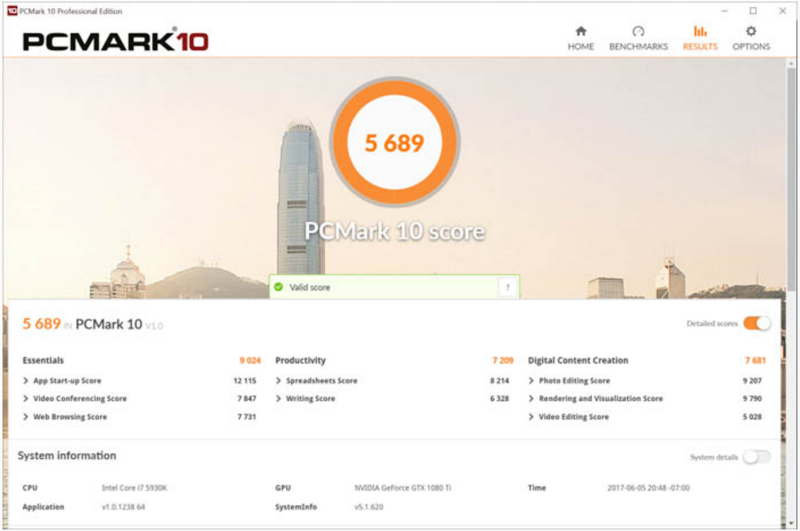
--

Energy (WH)

Comparison to PCMark

- PCMark does use some real-world applications but also synthetic workloads
- Different scenarios for different usage profiles
- Focused on office, productivity workloads
- Gaming workloads only in extended version
- Approximate run time is 30 minutes for the extended version
- Has a profile to measure battery life in hours for different scenarios
- Proprietary test framework and worklets
- Supports Windows only

PCMark10 Results



5 689 IN PCMark 10 V1.0

Essentials	9 024	Productivity	7 209	Digital Content Creation	7 681
App Start-up Score	12 115	Spreadsheets Score	8 214	Photo Editing Score	9 207
Video Conferencing Score	7 847	Writing Score	6 328	Rendering and Visualization Score	9 790
Web Browsing Score	7 731			Video Editing Score	5 028

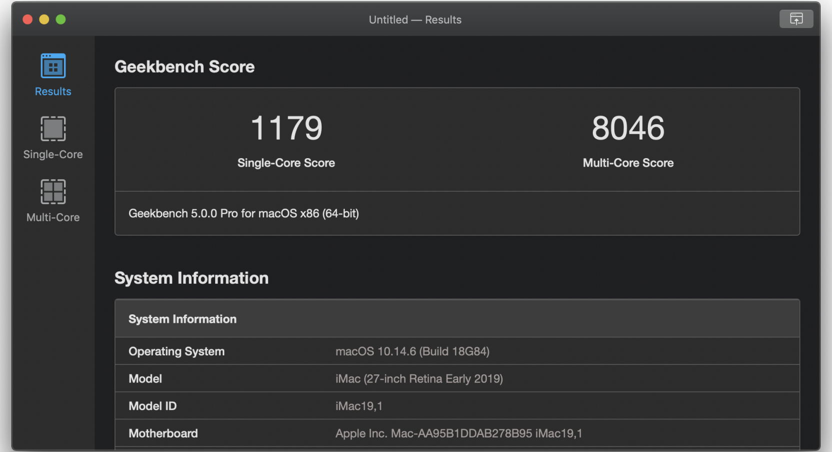
System information

CPU	Intel Core i7 5930K	GPU	NVIDIA GeForce GTX 1080 Ti	Time	2017-06-05 20:48 -07:00
Application	v1.0.1238 64	SystemInfo	v5.1.620		

Comparison to Geekbench

- Geekbench does not use real-world applications but claims to be simulating common tasks
- No different scenarios, only single core and multi core benchmarks
- Does not measure any energy consumption
- Proprietary test framework and worklets
- Supports Windows, MacOSX, Linux, Android, iOS

Geekbench Results



Geekbench Score

1179	8046
Single-Core Score	Multi-Core Score

Geekbench 5.0.0 Pro for macOS x86 (64-bit)

System Information

System Information	
Operating System	macOS 10.14.6 (Build 18G84)
Model	iMac (27-inch Retina Early 2019)
Model ID	iMac19,1
Motherboard	Apple Inc. Mac-AA95B1DDAB278B95 iMac19,1

Software Dissemination

The test software, test procedure, this presentation and other resources are available via the following website:

gtd-gmbh.de/computer-energy-efficiency-tool

clasp.ngo/tools/on-mode-computer-testing-tool