A Survey of FPGA Benchmarks

Raphael Njuguna, rnnjugun@wustl.edu (A project report written under the guidance of Prof. Raj Jain)

Abstract

New markets are emerging for the fast growing field-programmable gate array (FPGA) industry. Standard and fair benchmarking practices are necessary to evaluate FPGA systems and determine their potential to support target applications. This paper provides an extensive survey of FPGA benchmarks in both academia and industry.

Keywords: FPGA, Benchmark, Performance, Evaluation, RAW, VPR, MCNC, IWLS, PREP, Toronto 20, LINPACK, DSP, BDTI, MATLAB, MediaBench, OpenFPGA, Smith-Waterman, BLAST, EEMBC, Dhrystone, MiBench, OpenCores.

Table of Contents

1. Introduction
2. Historical Background
3. Benchmarks for Traditional FPGA Systems
   3.1. RAW Benchmark Suite
   3.2. VPR Benchmark
   3.3. MCNC Benchmark suite
   3.4. IWLS 2005 Benchmarks
   3.5. PREP Benchmark Suite
   3.6. Toronto 20 Benchmark suite
   3.7. LINPACK Benchmark
4. Benchmarks for Hybrid-FPGA Systems
      4.1.1. BDTI Communications Benchmark
      4.1.2. MATLAB Benchmarks
      4.1.3. MediaBench Benchmark Set
   4.2. Benchmarks for FPGA-based Biological Systems
      4.2.1. OpenFPGA.org
      4.2.2. Smith-Waterman Algorithm
      4.2.3. BLAST
   4.3. Benchmarks for FPGA-based Embedded Systems
      4.3.1. EEMBC Benchmarks
      4.3.2. Dhrystone Benchmark
      4.3.3. MiBench Benchmark Suite
5. Sources of FPGA Benchmarks
   5.1. Conference Benchmarks
   5.2. Open Source Benchmarks
   5.3. Synthetic Benchmarks
1. Introduction

In the recent years, field-programmable gate array (FPGA) systems have gained popularity in many applications such as digital signal processing, high performance computing, biological applications, just to name a few. FPGA, a reconfigurable digital logic device, facilitates rapid prototyping and design verification that enable designers to develop robust hardware and software solutions. A typical FPGA design flow involves: creating an electronic circuit design, placing and routing connectivity of the design onto FPGA architecture, verification and validation of the design, and configuration of the design into an FPGA device [WikipediaFPGA].

The FPGA community relies heavily on benchmarks to evaluate performance of their hardware and software solutions. Therefore, standard and fair benchmarking practices are necessary to evaluate FPGA systems and determine their potential to support target applications. For instance, an end user may study benchmark results published by various FPGA vendors to select an FPGA device that is suitable for the intended application. This survey will explore utilization of benchmarks to evaluate systems that contain FPGA devices and their associated software design tool chains.

2. Historical Background

The art of benchmarking in FPGA industry is as old as the industry itself. Shortly after FPGA was born in 1984, a benchmark suite consisting of ten combinational benchmark circuits was reported at the International Symposium on Circuits and Systems (ISCAS'85) [Hansen99]. Four years later ISCAS'89 benchmark suite contributed sequential circuits into the FPGA community [Brglez89]. The need for challenging and updated benchmarks led to the introduction of MCNC'91 benchmarks, which were published at the MCNC International Workshop on Logic Synthesis, 1991 [Yang91]. A series of benchmark suites published for conferences and workshops soon followed; they included LGSynth91, HLSynth92, PDWorkshop93, Partitioning93, just to name a few. Microelectronics Center of North Carolina (MCNC), working under ACM/SIGDA grant, maintained free electronic distribution of the aforementioned benchmarks [Brglez93].

Over the years, conference benchmarks have flooded FPGA community because they are more readily available than real industrial benchmarks. Nonetheless, a few non-profit organizations have afforded FPGA industry with benchmarks that span over diverse applications. For instance, PREP'94 benchmark suite was published by a consortium of companies in the programmable logic industry to demonstrate performance and capacity of programmable logic devices [Kliman94]. On the other hand, EEMBC, a non-profit organization formed in 1997 to develop benchmarks for embedded systems [EEMBC08], is an invaluable resource for FPGA system designers implementing soft-core processors. Open-source organizations such as OpenCores allow FPGA community to share real designs, which can be used as benchmarks.

3. Benchmarks for Traditional FPGA Systems

FPGAs have traditionally been used in reconfigurable and parallel computing systems. Consequently, FPGA community has developed numerous benchmarks to evaluate hardware and software solutions that implement...
these systems.

3.1. RAW Benchmark Suite

RAW benchmark suite was published by MIT's reconfigurable architecture workstation project for performance evaluation of reconfigurable computing systems such as FPGA [Babb97]. It implements diverse algorithms in general purpose computing that include CPU and parallel processing benchmarks. Performance of FPGA is based on throughput and resources required to solve a particular benchmark problem. Benchmark results are reported using the following metrics: solution speed (kHz), speedup relative to reference software, and speedup per FPGA [Babb97].

3.2. VPR Benchmark

Versatile place and route (VPR) is a component-level benchmark program contained in SPEC CPU2000 package. It was published by Standard Performance Evaluation Corporation (SPEC) to evaluate compute-intensive integer performance of FPGA during place-and-route design process [SPEC]. VPR demonstrates speed and throughput of performing place-and-route design task. SPEC adopted VPR program from a research project that created it as a tool for packing, placement, and routing designs in FPGA [Betz97]. Although VPR program is not included in the latest SPEC CPU2006 package, it is still popular in the FPGA community.

3.3. MCNC Benchmark suite

Microelectronics Center of North Carolina (MCNC) benchmark suite was published for MCNC International Workshop on Logic Synthesis, 1991. It included logic synthesis and optimization benchmark sets from ISCAS’85 and ISCAS’89 in addition to some other benchmarks collected from industry and academia. The benchmark suite has standardized libraries with representative circuit designs ranging from simple circuits to advanced circuits obtained from industry. MCNC also maintained free electronic distribution of benchmarks originating from past workshops and conferences [Brglez93]. MCNC benchmarks are very popular in academic research. For instance, [Mishchenko06] evaluates runtime performance of their optimization approach using MCNC benchmarks.

3.4. IWLS 2005 Benchmarks

The IWLS 2005 benchmark suite was published by International Workshop on Logic and Synthesis (IWLS). It contains diverse circuit designs derived from past conference benchmarks, open source community of hardware designers, and industry to represent a variety of applications [Albrecht05]. The benchmarks were synthesized and organized into a standardized library with a common timing infrastructure, standard APIs and reporting formats to promote easy exchange of benchmarks and experimental results in the community [Albrecht05]. [Mishchenko06] demonstrates performance of their approach, technique for combinational logic synthesis, by comparing it with runtime of logic synthesis scripts using IWLS benchmarks.

3.5. PREP Benchmark Suite
PREP benchmark suite was developed and published by Programmable Electronics Performance Corporation (PREP) to demonstrate performance and capacity of programmable logic devices [Kliman94]. PREP benchmarks enable designers to estimate target devices that best suit a particular application early on in the design process. The benchmarks implement a variety of applications ranging from simple data path circuits to complex state machines that stress on full utilization of routing resources [Kliman94]. PREP benchmarks indicate performance and capacity of an FPGA device using average benchmark capacity (ABC) that represents the maximum number of instances of a benchmark circuit that can fit into a device and average benchmark speed (ABS) that represents mean speed of internal and external logics of the device [Kliman94].

### 3.6. Toronto 20 Benchmark suite

Toronto 20 benchmark suite originated from an FPGA place-and-route challenge that was set up to encourage FPGA researchers to benchmark their software design tool chains on large circuits [Betz]. Some academic researchers have adopted these benchmark designs to evaluate their work. For instance [Strukov06] uses the Toronto 20 benchmark set to compare area ratios of CMOL technology with CMOS and nanoPLA circuit architecture technologies. Similarly, [Marquardt99] uses Toronto 20 benchmark circuits to evaluate performance of T-VPack, a timing-driven packing algorithm on various FPGA architectures based on area-delay product evaluation metric.

Back To Table Of Contents

### 3.7. LINPAC Benchmark

LINPAC Benchmark is a product of LINPAC software project that contains a collection of FORTRAN subroutines for solving various systems of linear equations [Dongarra03]. It measures floating-point rate of execution of a computer (Mflops/s) and may be used to compute theoretical peak performance for the machine [Dongarra03]. In an effort to explore viability of FPGA implementation in floating point scientific computing, [Turkington06] compares sustained floating point performance of FPGAs to standard commodity microprocessor based on sustained performance LINPAC benchmark set.

FPGA benchmarks for traditional reconfigurable and parallel computing systems were discussed. Evidently, some of the benchmarks are very old relative to modern FPGA technology. However, subsets of these benchmarks are still actively cited in today’s literature. Modern FPGA technology has progressed to implement other applications that were formerly realized by specialized architectures from other fields. This has led to a new concept of hybrid FPGA systems.

Back To Table Of Contents

### 4. Benchmarks for Hybrid-FPGA Systems

FPGA industry has grown and expanded to support diverse applications such as digital signal processing, biological systems, and embedded systems. Accordingly, research efforts and resources have been applied to develop benchmarks that are capable of evaluating these hybrid-FPGA systems.

Digital signal processing (DSP) industry is turning to low power FPGAs to implement their DSP applications. FPGA designers have adopted representative benchmarks from media and telecommunications industry that allow them to evaluate their FPGA-based DSP solutions.

### 4.1.1. BDTI Communications Benchmark

The benchmark was published by Berkeley Design Technology Incorporated, which develops signal processing benchmarks. It is an application-oriented benchmark based on orthogonal frequency division multiplexing (OFDM) receiver and designed to measure performance of signal processing engines [Bier07]. The benchmark enables designers to evaluate performance of FPGA platforms that implement high performance digital signal processing applications such as image processing. Results are reported in high-capacity and low-cost metrics, which represent maximum number of channels per chip and lowest cost per channel respectively [Bier07].

### 4.1.2. MATLAB Benchmarks

Nowadays, MATLAB is a common language for implementing DSP applications. Researchers have come up with tools that map DSP-MATLAB applications onto FPGAs e.g. AccelFPGA. MATLAB benchmarks are used to evaluate performance of these conversion tools. For instance [Banerjee03] uses MATLAB benchmark designs to test an approach of automatically converting floating point to fixed point based on resource consumption and frequency (MHz).

### 4.1.3. Mediabench Benchmark Set

MediaBench benchmark is a representative of multimedia and communications applications [Bishop99]. The benchmark set was introduced for performance evaluation of solutions that implement microprocessor architectures and ILP compilers for multimedia and communication systems [Lee97]. [Jones05] developed an architecture that combines VLIW processor and hardware functions that can support signal and image processing applications algorithms. The architecture was implemented on an Altera’s Stratix II FPGA and evaluated using signal processing benchmarks from MediaBench benchmark suite.

FPGA benchmarks for DSP systems were discussed. They represent diverse applications from media and telecommunications industry. Likewise, popular algorithms and applications from biological discipline have been designed into benchmarks to evaluate FPGA-based biological systems.

### 4.2. Benchmarks for FPGA-based Biological Systems

FPGA-based solutions enable biologists to explore huge bioinformatics’ databases with accelerated run times and high computing power. Therefore, they have adopted benchmarks to compare performance of various implementations.

#### 4.2.1. OpenFPGA.org
OpenFPGA is a non-profit organization that was formed in 2005 to promote progress in reconfigurable computing technology. Worldwide members from commercial, government, and academia share information about FPGA hardware systems and applications. One of its goals is to develop OpenFPGA benchmark suite to evaluate FPGA systems \cite{OpenFPGA}. It is mostly cited in performance evaluation of FPGA-based biological applications. For instance, \cite{Storaasli07} evaluate their FPGA implementations using human genome sequencing benchmark from OpenFPGA and publish their results at openfpga.org.

### 4.2.2. Smith-Waterman Algorithm

Smith-Waterman algorithm \cite{Smith81} is a computational intensive sequence alignment algorithm that identifies common molecular subsequences in bioinformatics. \cite{Zissulescu03} built a tool chain for mapping MATLAB-based applications onto FPGA-based platforms; they evaluated their methodology using Smith-Waterman algorithm, a computational intensive sequence alignment algorithm from bioinformatics. \cite{May07} evaluated performance of an FPGA implementation for detecting Ribonucleic acid (RNA) structures using Smith-Waterman Accelerator FPGA Design benchmark.

### 4.2.3. BLAST

Basic Local Alignment Search Tool (BLAST) is maintained and distributed by National Center for Biotechnology Information (NCBI), a resource for molecular biology. BLAST enables biologists to match nucleotide and protein sequences now available in huge databases; significant matches may suggest relationships between organisms \cite{BLAST}. BLAST utilizes heuristics to speed up Smith-Waterman algorithm, which produces more accurate sequence matches but is very slow for massive bioinformatics databases \cite{WikipediaBlast}. Computer vendors use Blast as a benchmark because it is very compute-intensive \cite{Sotiriades07}, \cite{Jacob07} demonstrated a speedup of 37X with their FPGA implementation of BLASTP compared to BLASTP software that is used for analysis of protein sequences.

Popular benchmarks for FPGA-based biological systems were examined. Similarly, benchmarks are needed to analyze performance of FPGA systems that contain embedded processors to enable them function as system-on-a-chip systems (SOCs).

### 4.3. Benchmarks for FPGA-based Embedded Systems

A move towards system-on-a-chip systems (SOCs) has introduced embedded processors into FPGA industry. For instance, Altera and Xilinx provide Nios II and Microblaze embedded processors respectively to facilitate design of systems on their FPGA platforms \cite{Orecchio07}.

### 4.3.1. EEMBC Benchmarks

Embedded Microprocessor Benchmark consortium (EEMBC) is a non-profit organization that maintains standard benchmarks for embedded systems dealing with automotive, consumer, digital entertainment, java, networking, office automation, microcontrollers, and telecommunication applications \cite{EEMBC}. For instance \cite{Sheldon07} explores Design of Experiments (DOE) paradigm approach to optimize a soft-core microprocessor for a particular application; they evaluate their approach using EEMBC benchmarks and monitors speedup for benchmark applications compared to a base core. Similarly, \cite{Lysecky05} uses EEMBC
benchmarks to compare performance and energy consumption of an FPGA soft processor core implementation with standard hard-core processors.

### 4.3.2. Dhrystone Benchmark

Nowadays, designers are adopting FPGAs to implement system on chip systems. Processor core is a key component for such a system. [Hempel07] has developed SpartanMC, an FPGA processor core that optimizes the resource usage for FPGA-based SOC's and evaluated its performance against other cores using Dhrystone benchmark whose results are reported in Dhrystone MIPS/MHz. Similarly, [Shannon04] evaluate performance of SnoopP, a profiling software tool that allows designers to measure system design on-chip, on Xilinx Virtex II FPGA with the MicroBlaze processor using Dhrystone benchmark. Monitor execution time (ms).

### 4.3.3. MiBench Benchmark Suite

MiBench is a set of freely available embedded application programs for embedded processor performance developed at University of Michigan at Ann Arbor. It represents diverse commercial applications categorized into automotive and industrial control, consumer devices, office automation, networking, security, and telecommunications [Guthaus01]. [Dimond05] evaluates performance of customizable multi-threaded FPGA soft processor and compiler generation system using media and cryptographic benchmarks from MiBench suite. [Xu04] propose and evaluate a memory saving code compression architecture using MiBench benchmark set. Monitor relative performance versus compression ratio, memory space saving for the benchmarks. [Dimond05] evaluates performance of customizable multi-threaded FPGA soft processor and compiler generation system using media and cryptographic benchmarks from MiBench suite.

Benchmarks for a rising trend of hybrid FPGA systems were considered in this section. In particular, the survey examined benchmarks in digital signal processing, biological systems, and embedded systems to illustrate efforts by FPGA industry to evaluate hybrid FPGA systems.

### 5. Sources of FPGA Benchmarks

An ideal benchmark for a system should represent workload executed by end users of the system. Obviously, this is a major challenge for FPGA community because real customer designs are rarely published to the public domain. This survey found that majority of FPGA benchmarks originate from past conferences, open source organizations, synthetic benchmark generators, and FPGA vendors.

#### 5.1. Conference Benchmarks

FPGA community actively participates in conferences and workshops to discuss their work. Some of these conferences provide benchmark designs that consist of applications from diverse industries that can be potentially supported by FPGA systems. For instance MCNC’91 and IWLS’05 were published for workshops on Logic and Synthesis. On the other hand, Toronto 20 resulted from a competition that was held to encourage FPGA researchers to benchmark their tool chains. Conference benchmarks are widely circulated because they are freely available in public domains.
5.2. Open Source Benchmarks

Open source FPGA communities allow members to share benchmarks, methodologies, and results. OpenCores is an open source community that deals with semiconductor intellectual property cores [OpenCores]. Academia and industrial corporations use freely available designs to benchmark their products. Therefore, consumers and competitors can duplicate benchmark tests on their devices. For instance, Altera compared performance of their Stratix III FPGAs with Xilinx’s Virtex-5 FPGAs using OpenCores benchmark designs to enable customers to duplicate their benchmark results [Altera07].

5.3. Synthetic Benchmarks

In some cases, designers turn to automatic generation of synthetic circuits instead of existing benchmark suites that contain real designs. [Verplaetse00] presents an approach to generate synthetic benchmark circuits for evaluating new architectures and tools, which don’t have representative evaluation benchmark sets. For instance, [Chang04] implemented an algorithm for generating synthetic benchmarks and used them to study optimality and scalability of placer tools: dragon, Capo etc.

5.4. Industrial Benchmarks

FPGA manufacturers and vendors use real customer designs as benchmarks to demonstrate performance of their products over their competition. However, customer designs are usually confidential and provided under non-disclosure agreements. Therefore, FPGA industry seeks alternative benchmarks, which they can use to market their products, and enable customers and competition to duplicate their benchmarking tests. For instance, Altera compared their Stratix III FPGAs with Xilinx’s Virtex-5 FPGAs in terms of performance speed, resource utilization, and compilation time using largest of the most popular OpenCores benchmark designs, and latest vendor software with default CAD settings [Altera07]. In rare cases, some researchers have evaluated their FPGA solutions using real customer designs. [Metzgen05] implements synthesis algorithm on Altera’s Synthesis software within Quartus II and evaluates area reduction using 120 real customer designs using Altera’s benchmarking suite.

Sources of FPGA benchmarks were explored in this section. FPGA community faces a major challenge of obtaining real customer designs and publishing those benchmark results to the public domain. Nonetheless, FPGA community obtains majority of their benchmarks from conferences, open source organizations, synthetic benchmark generators, and FPGA vendors.

6. Summary

An extensive survey was conducted on FPGA benchmarks in both academia and industry. Evidently, FPGA community relies heavily on benchmarks to evaluate performance of its hardware and software solutions. Most of the benchmarks in academia originate from conferences and workshops. On the other hand, FPGA manufacturers and vendors emphasize on benchmarking with real customer designs. Unfortunately, real customer designs are confidential and bound by non-disclosure agreements. Therefore, FPGA industry seeks alternative benchmarks, which they can use to market their products and enable customers to duplicate their benchmarking tests. This has led to formation of non-profit organizations and open source communities that allow members to share benchmarks, methodologies, and results.
In some cases, FPGA designers automatically generate synthetic benchmarks to evaluate new architectures, which they cannot efficiently test using existing benchmarks. As FPGA industry expands to support new applications such as digital signal processing and embedded systems, benchmarks from these areas are adopted by FPGA community as well. The common metrics reported for FPGA benchmarks include: logic capacity, performance speed, resource utilization, power consumption, area required for placing designs etc.

7. References


A Survey of FPGA Benchmarks


8. List of Acronyms

ABC - Average Benchmark Capacity

ABS - Average Benchmark Speed

BLAST - Basic Local Alignment Search Tool

CAD - Computer Aided Design

CMOS - Complementary Metal Oxide Semiconductor

DSP - Digital signal processing

EEMBC - Embedded Microprocessor Benchmark consortium

FPGA - Field-programmable gate array

HDL - Hardware Description LanguageS

MCNC - Microelectronics Center of North Carolina

NCBI - National Center for Biotechnology Information

OFDM - Orthogonal Frequency Division Multiplexing