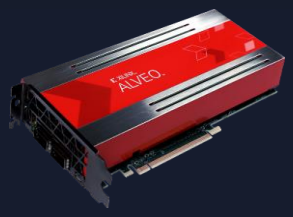
Hive Acceleration

Xilinx Alveo powers bzQAccel

SOLUTION BRIEF

XILINX ALVEO™



- 4x speed-up
- 1-click install
- No setup change
- No query re-write
- Extendable to other engines

INTRODUCTION

There is insatiable need for speed from analytical queries. At the same time, with Moore's law coming to an end, CPUs have hit a performance wall. Many studies have clearly established that RAM/disk contention and network I/O severely limit benefits of scaling up/out beyond a point.

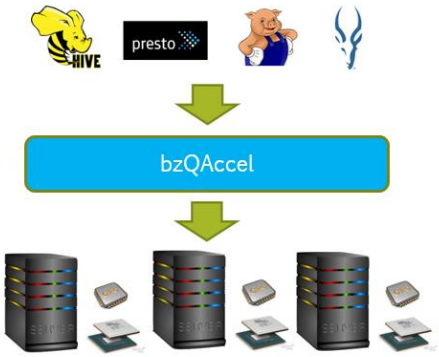
FPGAs have emerged as a very popular choice for accelerating variety of algorithms and are well suited to take on Big Data work-loads. Realizing the need for speeding up analytical queries, BigZetta has augmented Hive with FPGA based acceleration. With this BigZetta solution (bzQAccel) has demonstrated speed-ups of 4x (and more) over several TPCH benchmark queries.

KEY BENEFITS

- **Speed-up:** Up to 4x speed-up of analytical queries
- **1-click install:** In a single step, bzQAccel can be installed on top of any Hive distribution
- **No setup change:** User does not need to change query or existing setup

SOLUTION OVERVIEW

BigZetta's solution, bzQAccel, acts as middleware between query processing and its execution on hardware (CPU+FPGA).



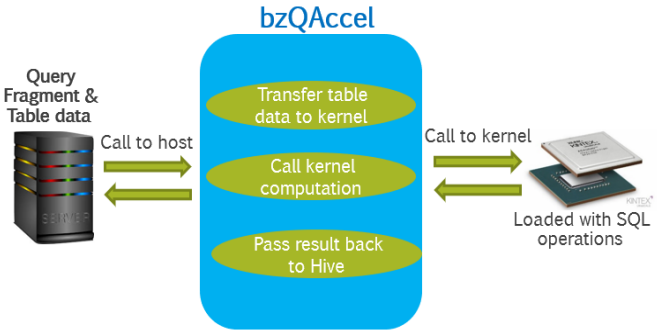
bzQAccel has been architected to provide seamless integration with Hive and is agnostic of Hive version being used. User does not need to make any changes to either query code or cluster setup. In fact, the technology can be deployed on top of a running Hadoop cluster. Only change a user would see would be faster execution of queries.

Hive Acceleration

Xilinx Alveo powers bzQAccel

SOLUTION DETAILS

The key aspect of FPGA based Hive acceleration is the interaction between the CPU (Host) and the FPGA (Device). BigZetta's Query Accelerator (bzQAccel), has been designed to manage the interaction between Hive code running on CPU and offloading of runtime critical operations onto FPGA. Figure below shows the interaction between CPU and FPGA through bzQAccel layer:

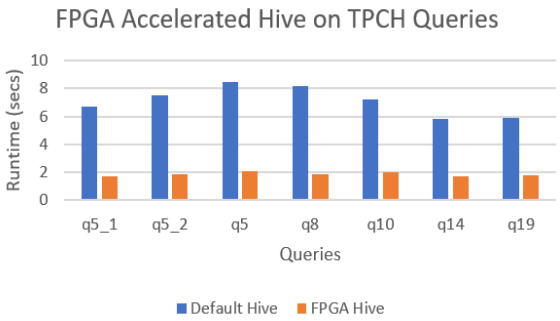


Key features of bzQAccel:

- Acts as **middleware** between Hive and underlying hardware
- **Optimizes** query execution plan suited for FPGAs
- Provides **fastest execution** of the plan on FPGAs
- For different queries, **no need to recompile** either the host code or the FPGA kernel
- **Minimal penalty** of data movement (from CPU to FPGA and back)

RESULTS

On a cluster of machines (CPU + Xilinx® Alveo™ Data Center accelerator cards), several TPCH benchmark queries were run with default Hive and one augmented with bzQAccel. FPGA accelerated Hive consistently outperformed default Hive in runtime by **4x across all the queries**.



TAKE THE NEXT STEP

Learn more about Xilinx Alveo accelerator cards
Learn more about BigZetta Systems, visit www.bigzetta.com
Reach out to sales – sales@bigzetta.com