# Performance evaluation of InAccel ML scalable suite

Evaluation for the Accelerated ML suite of InAccel for Spark

The FPGA-Accelerated ML suite for distributed systems is a fully integrated AMI/AFI that in used to speedup machine learning applications for Apache Spark distributed systems. The current version of the suite allows the acceleration of logistic regression and k-means clustering machine learning applications for Apache Spark. The suite provides all the required APIs that allows the utilization of the Amazon F1 instances without any changes of the original code.

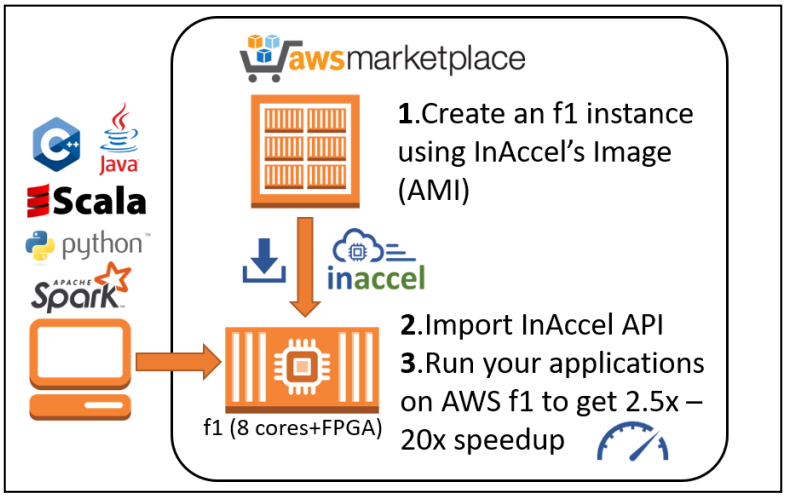
The Spark code for the utilization of the hardware accelerator through our accelerated machine learning library is shown in the figure. When the Spark user wants to utilize the hardware accelerator, the only change that needs to be made is the replacement of the Spark mllib library with the mllib_accel library. Therefore, the user can speedup the execution time of the Spark application with a simple replacement of the libraries that wish to accelerate.

## Accelerated ML suite for Spark

Features:

- Up to 12x speedup compared to 32-core contemporary processors
- APIs for C/C++, Python, Java, and Scala
- APIs and libraries for Spark integration
- Highly scalable system

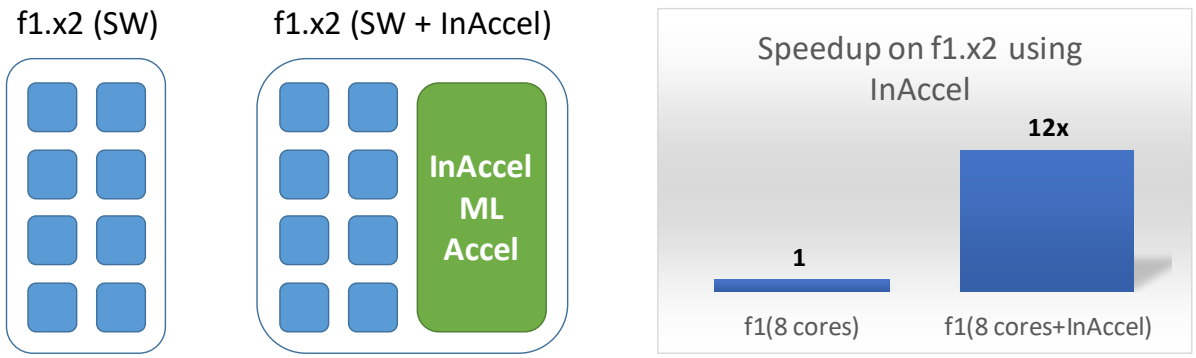
| | |
|---|---|
| **Version** | 1.0 |
| **Operating System** | Linux/Unix, Ubuntu 16.04 |
| **Fulfillment Methods** | Amazon Machine Image |
| **Price** | **f1.x2**<br>$1.4/hour or<br>$500/month ($0.7/hour)<br>**f1.x16**<br>$8.4/hour or<br>$3000/month ($4.2/hour) |



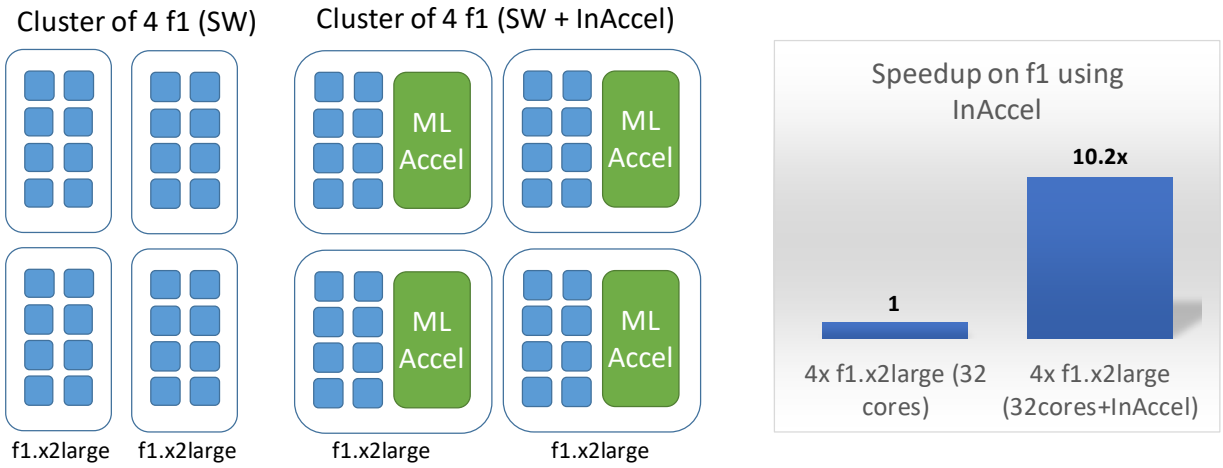
## Results

The following figures shows the speedup that can be achieved using the InAccel Accelerated ML suite for Spark. The figures below show the speedup achieved using Logistic Regression on a dataset of 6Million point (~20 GBytes). Different configuration are used to show the scalability of the InAccel scalable technology using our Accelerators.

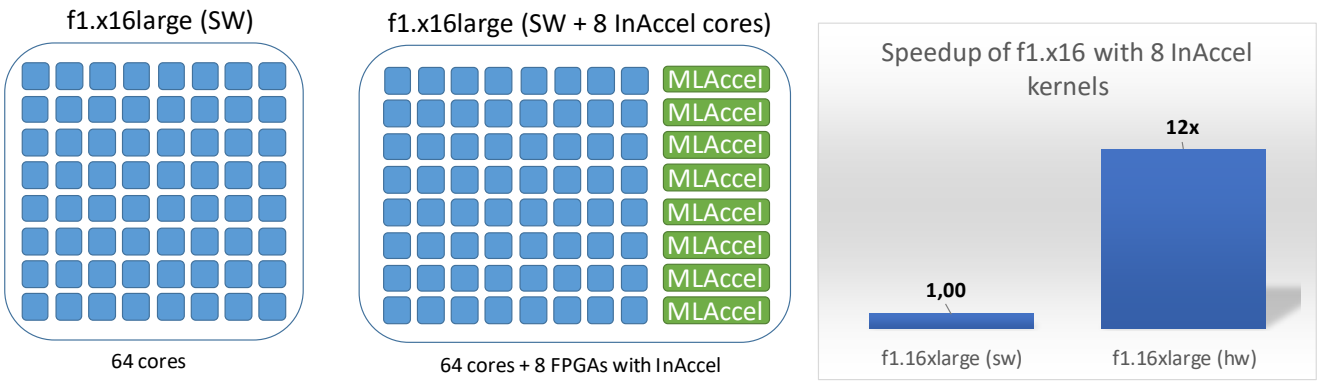
Comparison with one f1.x2large (without and with InAccel ML Accelerators)



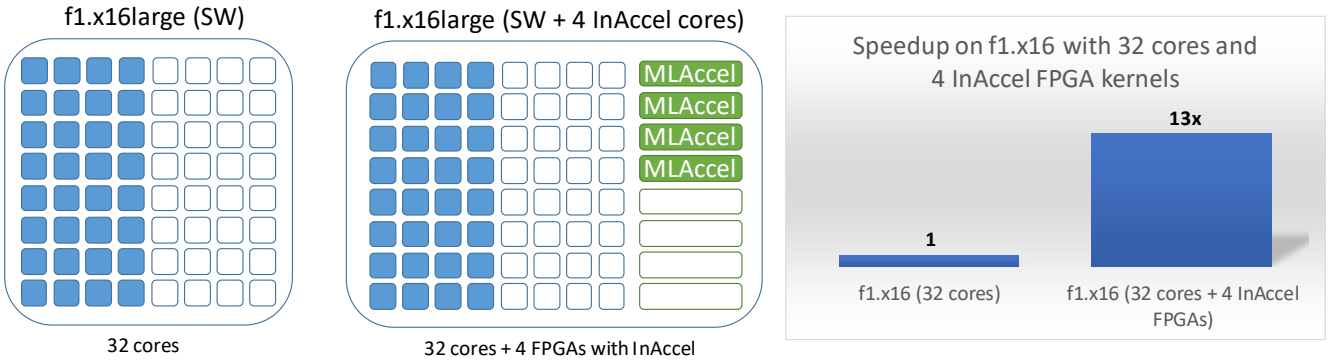

Comparison with a cluster of four (4) f1.x2large




**Comparison on large FPGAs (f1.x16) with 64 cores (without and with 8 InAccel FPGA kernels)**




## Comparison on large FPGAs (f1.x16) with 32 cores (without and with 4 InAccel FPGA kernels)




## Comparison on cluster of 4 nodes normalized per cost ($2/hour/node for r4, $1.6+$1.4 for f1.x2)