A Low Latency Library in FPGA Hardware for High Frequency Trading (HFT)



John W. Lockwood, Adwait Gupte, Nishit Mehta (Algo-Logic Systems) Michaela Blott, Tom English, Kees Vissers (Xilinx)



August 22, 2012, Santa Clara, CA

© Algo-Logic Systems Inc., All rights reserved



Outline

Introduction

– High Frequency Trading (HFT)

- Survey of HFT Platforms
 - Software, Hardware, and Hybrid Approaches
- Field Programmable Gate Arrays (FPGAs)

 Advantages and Disadvantages
- Algo-Logic's Low Latency Library
 - Implementation on NetFPGA 10G Platform
 - Exposure and Position Tracking Application
 - Protocols Supported
- Results



High Frequency Trading (HFT)

HFT is

- Trading of equities, options, futures at high speed in large volumes
- Earning money by exploiting the fleeting variation in stock price or demand

HFT accounts

- 70% of all trades in US Markets in 2010
- And it continues to grow

HFT involves

 Using computers to place orders based on pre-defined algorithms



Challenges in Financial Markets

Main Challenges

- Latency
 - Execute orders faster than other investors to capture fleeting variations in price and demand in the markets

Jitter

- Provide consistent and fair executions

Secondary challenges

- Throughput
 - Handle large volume of orders
- Flexibility
 - Adapt to changing risks and trading strategies



Recent Problems in HFT

Knightmare (Knight Capital)

- Test script executed live trades
- \$450M loss in 45 minutes

Nasdaq - Facebook IPO

- Order confirmations delayed
- \$62M loss in direct damages

BATS Failure

- Software bug in order auctions
- Forced to cancel IPO

Have not only hurt the banks/institutions financially, but also the credibility of the market



Latency in Current Approaches

Software

Linux 10GE NICs

- 15-20 µs for Half-Round Trip Time (½ RTT) through un-optimized kernel
- TCP Offload: 2.9 µs Transmit +
 6 µs Receive for ½ RTT

Datagram Bypass Layer (DBL)

3.5 µs for UDP and
 4.0 µs for TCP

Infiniband MPI

1 µs
 (excluding application layer)

Hardware

- Graphics Processor (GPU)
 - Optimized for throughput, but not optimized for low latency
 - Incurs additional overhead of passing data through PCIe bus

• ASIC

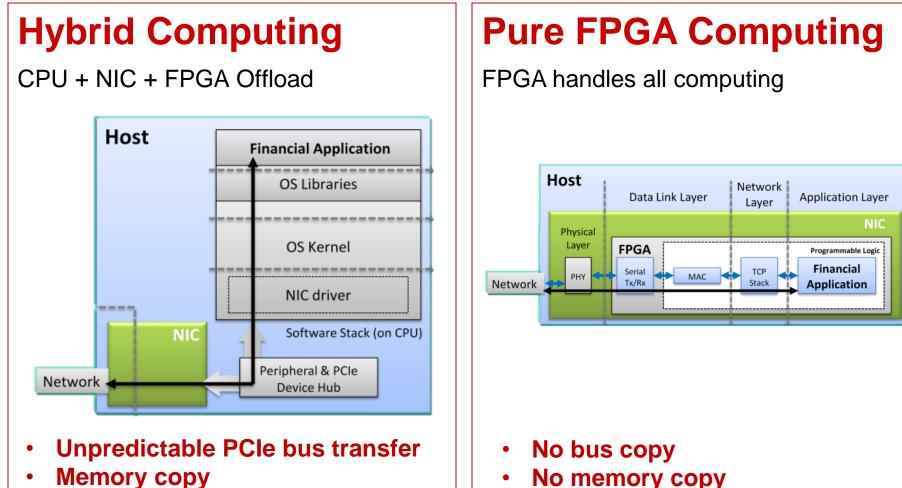
- Achieves sub-µs latency
- But lacks flexibility to handle new protocols and features

• FPGA

- Provides 0.2 µs latency w/TCP
- Has the flexibility to support new protocols and features



FPGA Approaches



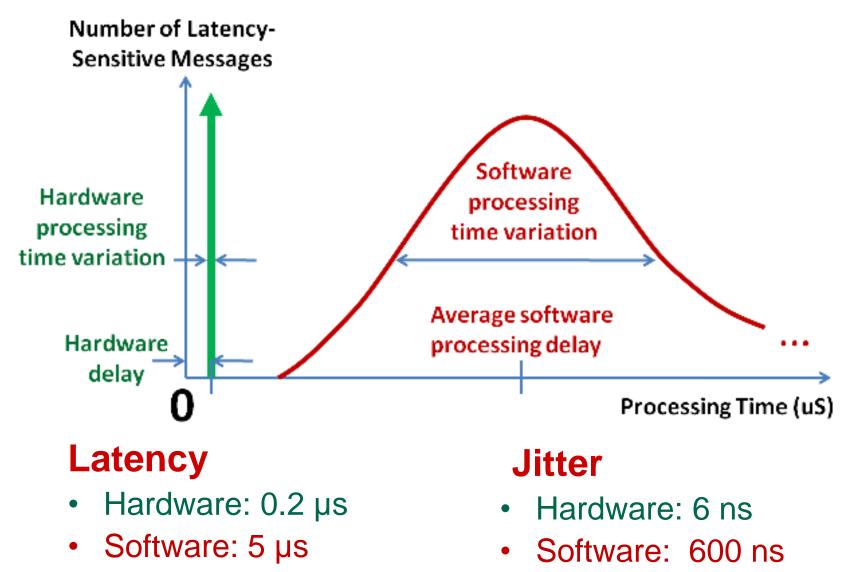
7

- Potential cache misses
- Amdahl's law

- No cache misses
- Parallel execution

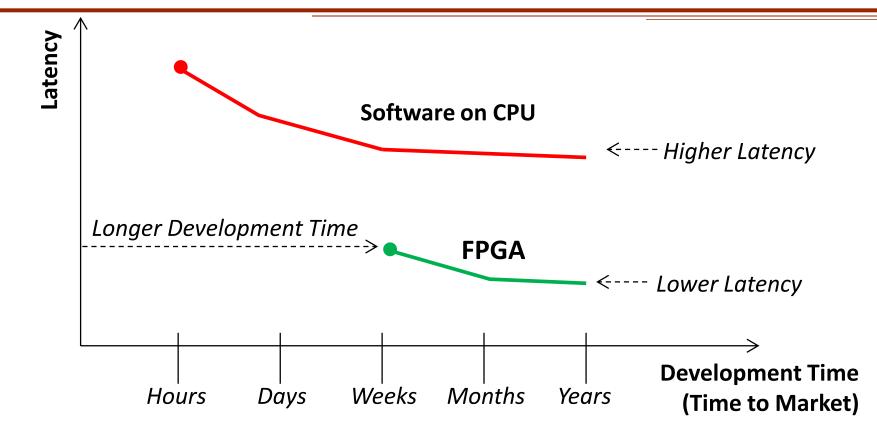


FPGA outperforms Software





Latency vs. Development Time



Software solutions

- Require less development time to get started

But FPGA hardware solutions achieve lower latency

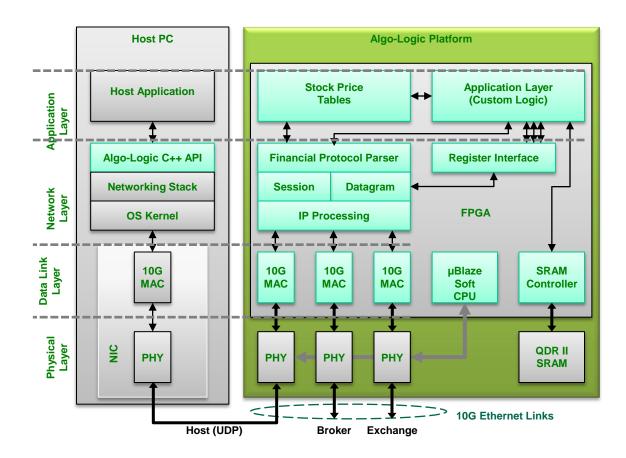
- That fundamentally cannot be achieved with software



Algo-Logic's Low Latency Library

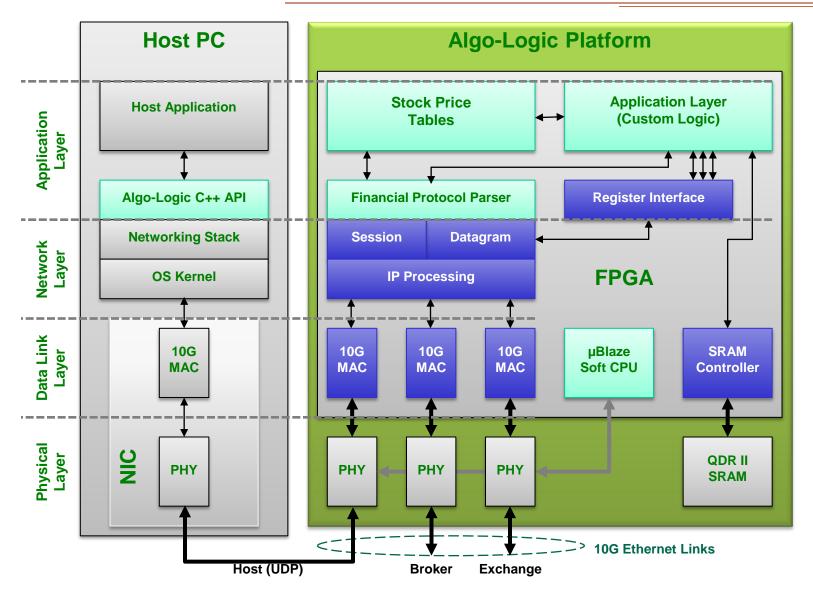
- Infrastructure
- Protocol Parsers

Market data in local memory



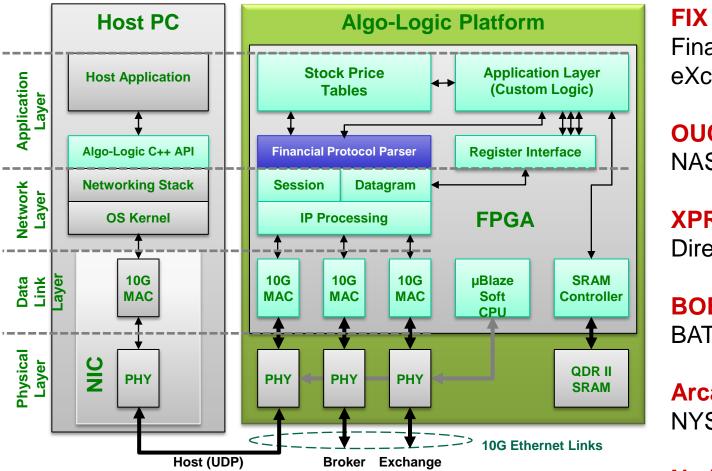


Infrastructure





Protocol Parsers



FIX Financial Information eXchange

OUCH NASDAQ

XPRS DirectEdge

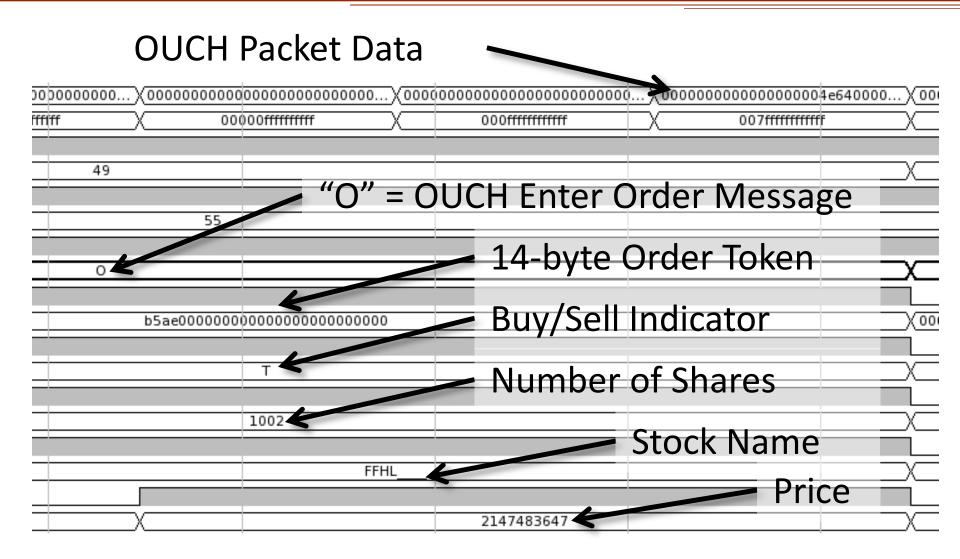
BOE BATS BZX

ArcaDirect NYSE Arca

Native Trading Gateway LSE

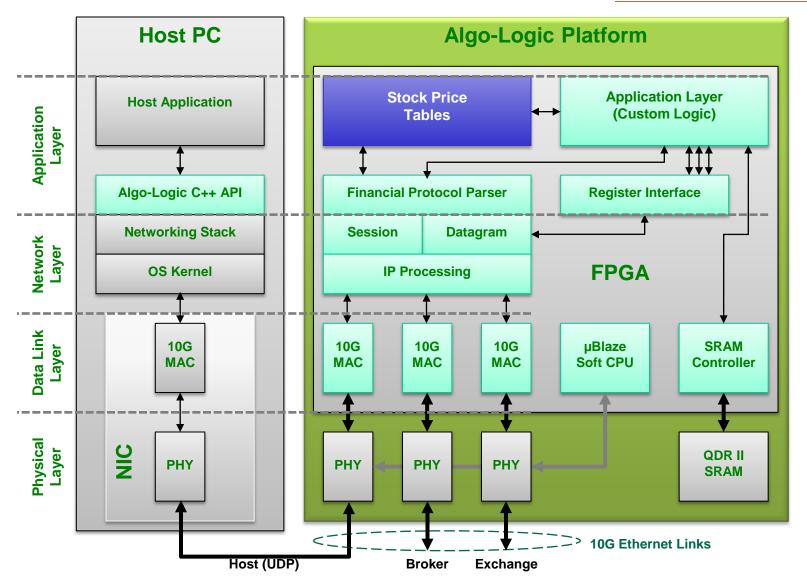


View of an Nasdaq order in the FPGA





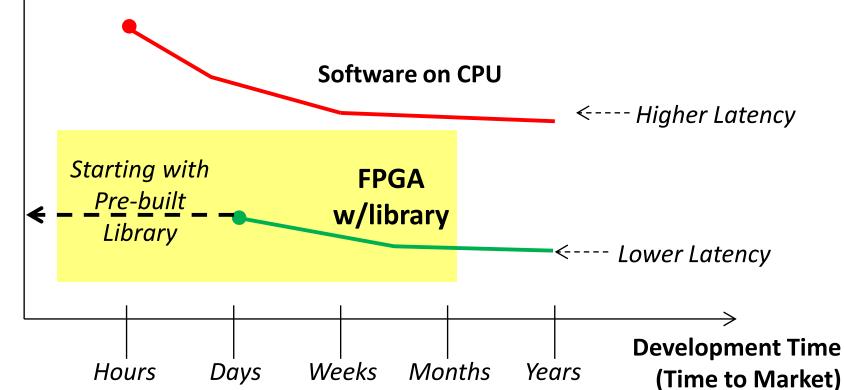
Market Data and On-chip Storage





Algo-Logic Reduces Time to Market





- Starting with a pre-built low-latency library
 - Reduces the initial development effort

But maintains fundamental benefits of FPGA solution

- Full data-path remains in low-latency hardware



Areas of Application

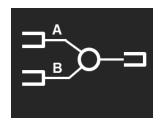
Algo-Logic's Low Latency Libraries provide lowest latency for Traders, Brokers, Market Makers and Exchanges in the areas of:



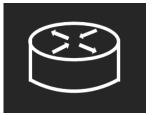
Trading Strategy



Internalization



Feed Processing



Smart Order Routing



Risk Management

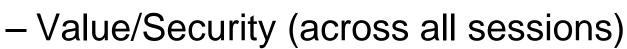


Matching

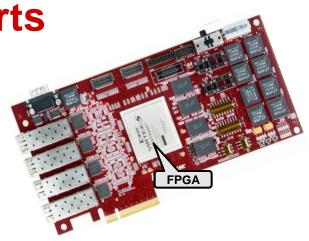


Operations performed in hardware

- Parsing FIX execution reports
- Update price per Security
 to calculate
 - Long Exposure
 - Short Exposure

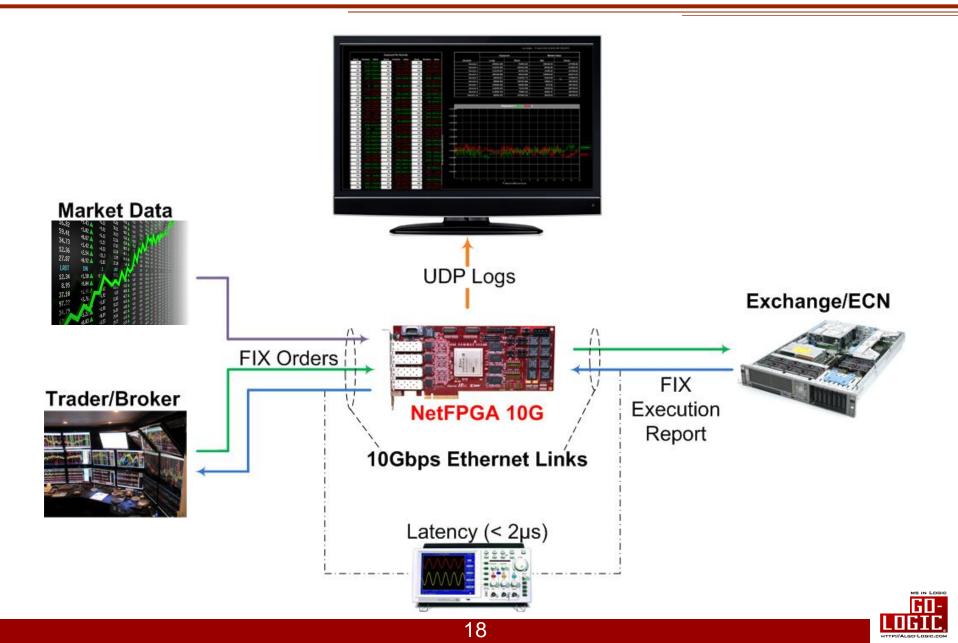


- Position/Security (across all sessions)
- Sending these values to the customer
 On a programmed, periodic basis





Example: Position & Exposure Monitor



Specifications & Performance

Parameter	Value (demo)
Hardware platform	NetFPGA-10G
Application	Position & Exposure calculation
Protocol	FIX 4.2
# sessions supported	10
# securities supported	100
Clock frequency	Line rate, 10 Gbps
Latency (Logic processing inside FPGA)	200ns
10GbE PHY delay	400ns (one-way)
Total Latency (pin-to-pin)	1µs



Results

Low latency library

- Implemented as FPGA gateware

Implements

- Order flow processing infrastructure
- Protocol parsing for all major exchanges
- Maintains market data in local memory

Demonstrated

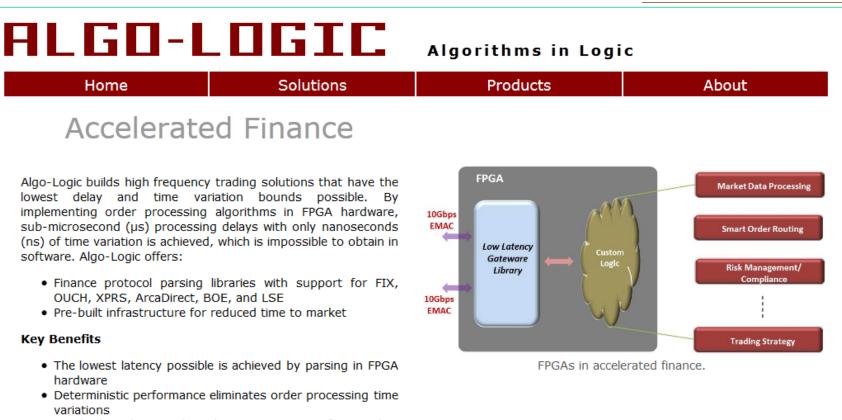
– Ten sessions of FIX 4.2 on NetFPGA 10G

Achieves

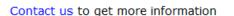
– Jitter-free processing with 200ns latency



More about Accelerated Finance



- UDP/IP-based control and monitoring interface makes operation of the hardware easy from graphical and command line interfaces
- Line rate processing at 10 Gigabits/second



Feedback

Careers

Copyright © 2009-2012 Algo-Logic Systems Inc., All Rights Reserved

21



Algo-Logic Systems, Inc.

ALGORITHMS IN LOGIC



Web

http://Algo-Logic.com

Email

Solutions@ Algo-Logic.com

Phone

(408) 707-3740

Office Address

2255-D Martin Ave. Santa Clara, CA 95050



