

# Galapagos: A Full Stack Approach to FPGA Integration in the Cloud

Naif Tarafdar, Nariman Eskandari, Paul Chow

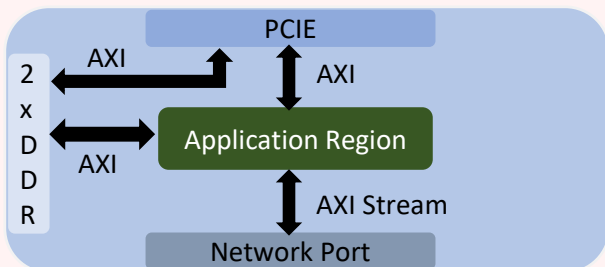
Electrical and Computer Engineering, University of Toronto, naif.tarafdar@mail.utoronto.ca



## Abstract

- Integrating FPGAs at scale is challenging as this requires the orchestration, and communication of many FPGAs
- We developed a hardware stack that abstracts individual FPGAs and clusters of FPGAs
- We created a heterogeneous communication layer
- We developed a prototype of a multi-FPGA CNN platform

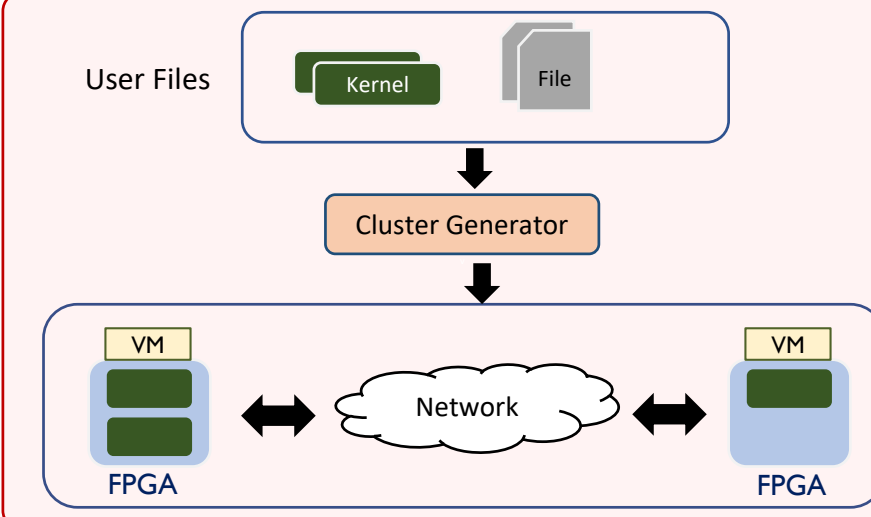
## Hypervisor (Single FPGA Abstraction)



## User Defined Cluster

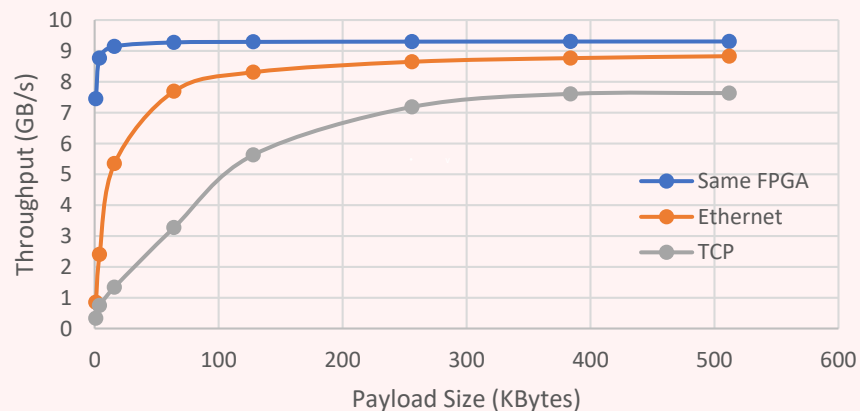
- User describes cluster as collection of streaming kernels
- User specifies network protocol (e.g TCP or Ethernet)
- User can choose to add additional communication layers such as MPI on top of user specified network protocol

## Cluster Generation Flow

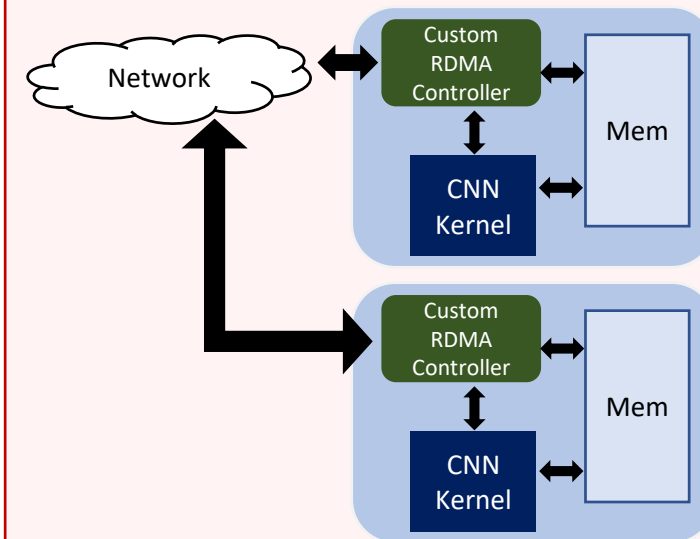


## Results

- Utilization on KU15 FPGA:
  - 20 % with TCP
  - 15 % with Ethernet



## Telepathy: Multi-FPGA CNN



## Conclusions/Future Work

- Our multi-FPGA framework makes the orchestration of data center scale applications easy
- Our heterogeneous communication layer allows easy communication between multiple-FPGAs, CPUs etc.
- We plan to explore other multi-FPGA applications

## Acknowledgements

We would like to thank Xilinx, CMC, Huawei, NSERC for funding our research

