# FHE That Ships: Our Journey From Research to Real Deployment

Ajay Joshi CipherSonic Labs

FHE Hardware Acceleration Summit Stanford University











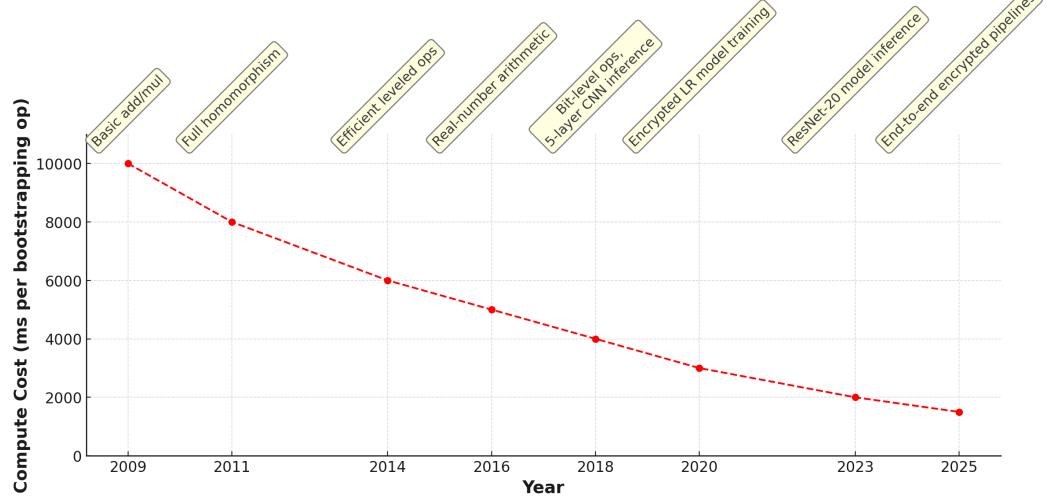








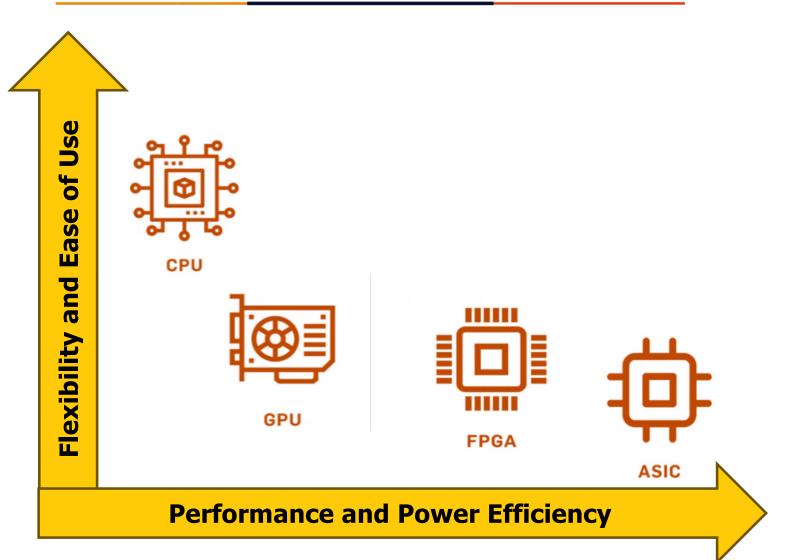
### FHE Landscape



Algorithmic, hardware, and software optimizations have contributed



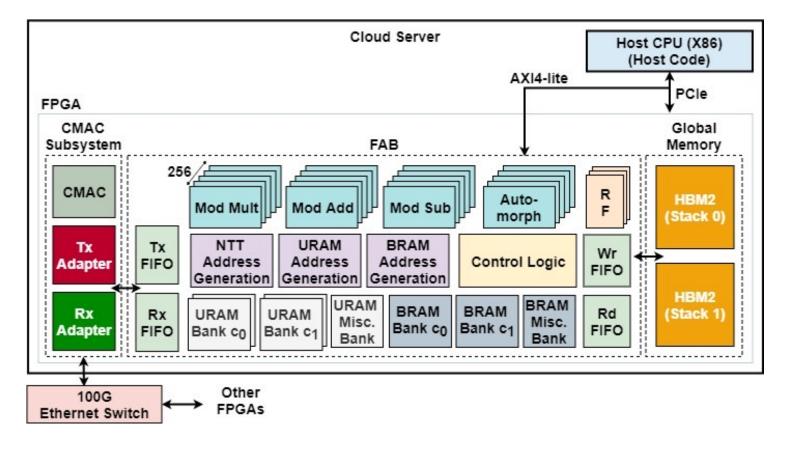
### Why FPGAs are a natural fit?



- Reconfigurable architecture for FHEspecific optimization
- Custom memory hierarchy for primitivelevel operations
- Compute units optimized for lowerfrequency, highefficiency execution



### FAB highlights



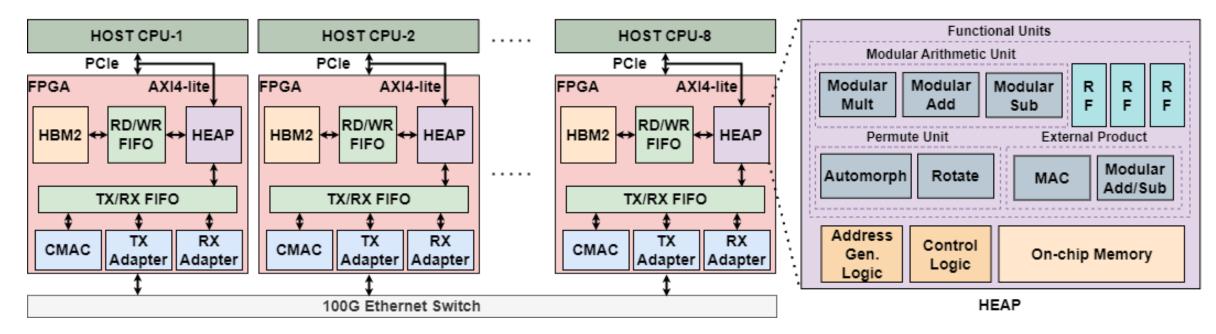
R. Agrawal et al., "FAB: An FPGA-based Accelerator for Bootstrappable Fully Homomorphic Encryption," in Proc. IEEE International Symposium on High-Performance Computer Architecture (HPCA) 2023.

- Single FPGA architecture mapped to Alveo U280 FPGA accelerator card
- First ever 128-bit secure CKKS bootstrapping implementation on FPGA
- Bootstrapping and logistic regression model training
  - ~100x faster than CPU
  - ~10x faster than GPU
- When scaled to 8 FPGAs, performance limited by bootstrapping performance



### **HEAP** highlights

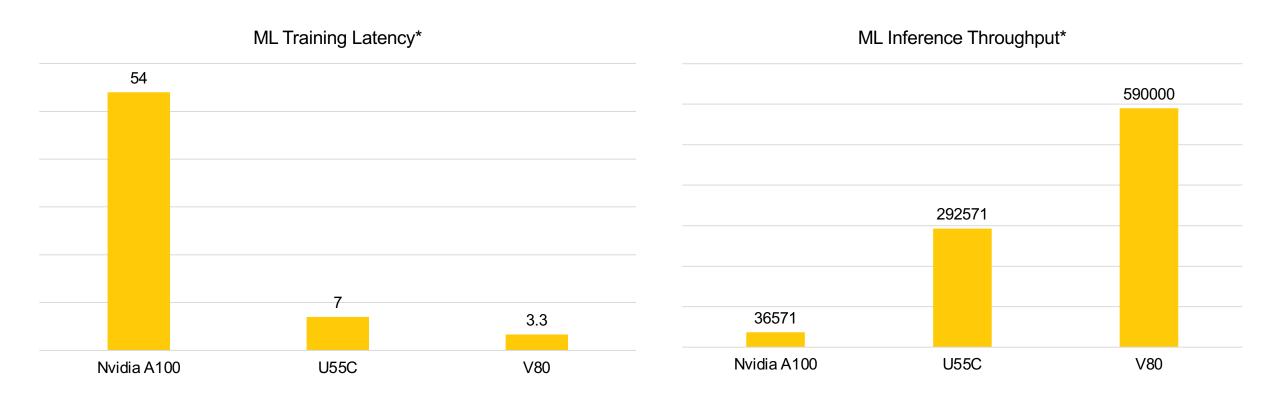
- Scalable bootstrapping accelerator → Multi-FPGA system
  - Parallelizable bootstrapping using scheme switching
  - ~15x better bootstrapping performance compared to FAB
  - ~11x better performance for logistic regression over 8-FPGA FAB



R. Agrawal et al., "HEAP: A Fully Homomorphic Encryption Accelerator with Parallelized Bootstrapping," in Proc. International Symposium on Computer Architecture (ISCA) 2024.



#### Performance – Credit card fraud detection



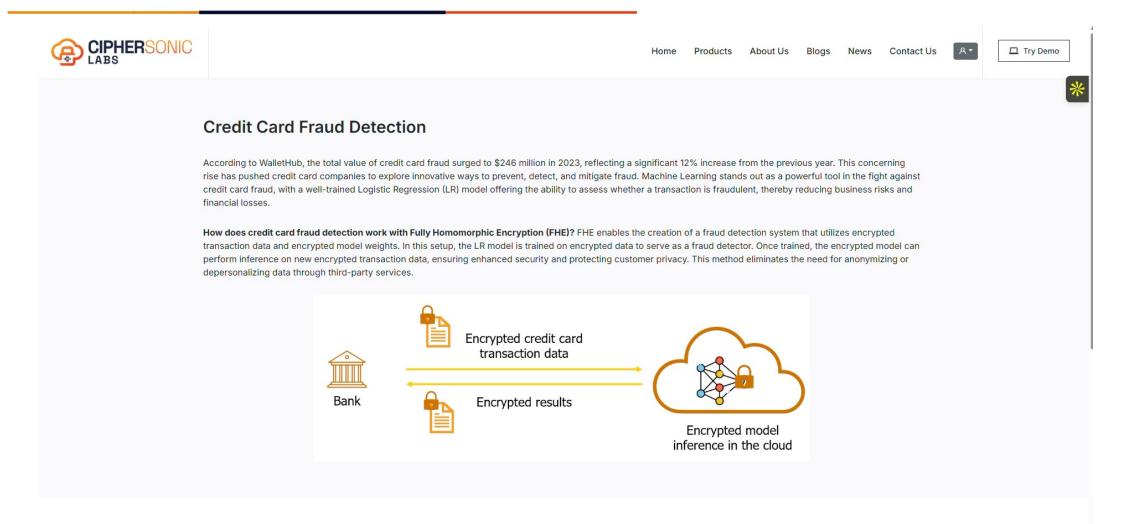
<sup>\*</sup>Latency measured in milliseconds per training iteration Solutions Brief with AMD:

https://www.amd.com/content/dam/amd/en/documents/products/accelerators/alveo/v80/cipherSonic-solution-brief.pdf



<sup>\*</sup>Throughput measured in credit card transactions per second

## MVP live on AWS (Looking for early adopters)



Choose the transaction you want to verify



### Call for action - Compiler stack for FHE on FPGAs

- What's missing today
  - No open-source, end-to-end compiler targeting FPGAs for FHE workloads
- What we need to build together
  - IR (Intermediate Representation) for encrypted ops targeting hardware
  - RTL & HLS codegen hooks for common FHE op-sets for CKKS
  - Integration with existing homomorphic libraries (OpenFHE, SEAL, Concrete)
  - Tooling for debugging and verifying encrypted dataflow
- Open sourcing is the way to go
  - This infrastructure should be open-source, accessible, and vendor-neutral
  - Collaborators can plug in passes, optimizations, backend targets



#### Summary

- FHE-based computing is becoming practical
- FPGAs provide a sweet spot for FHE acceleration
  - Existing clouds have FPGA COTS
  - Practical performance at a fraction of ASIC cost
- For widespread FHE adoption
  - Standardization of FHE algorithms
  - Proven use cases and success stories





<u>ajay@ciphersoniclabs.io</u> rashmi@ciphersoniclabs.io

