Rebellions Inc (리벨리온)

Towards the next era of AI

by delivering rebellious compute power at lower energy than ever

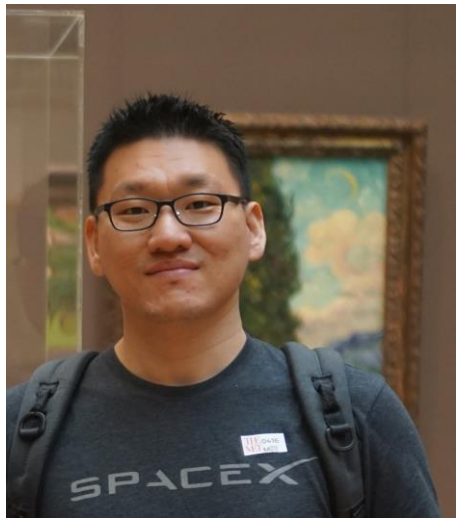
Founders

Jinwook Oh

AI Chip Lead Architect
IBM TJ Watson, USA

PhD in KAIST EE, 2013
BS in SNU EE, 2007

Sunghyun Park



Vice President
Morgan Stanley, USA

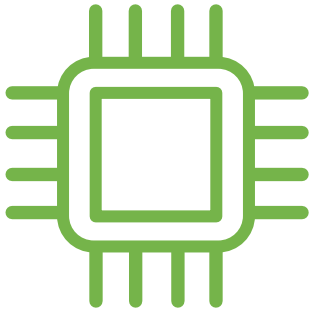
PhD in MIT EECS, 2014
BS in KAIST, 2009

Hyo-Eun Kim

Chief Product Officer
Lunit, South Korea

PhD in KAIST EE, 2013
BS in KAIST EE, 2005

Vision



Rebellion
AI Core

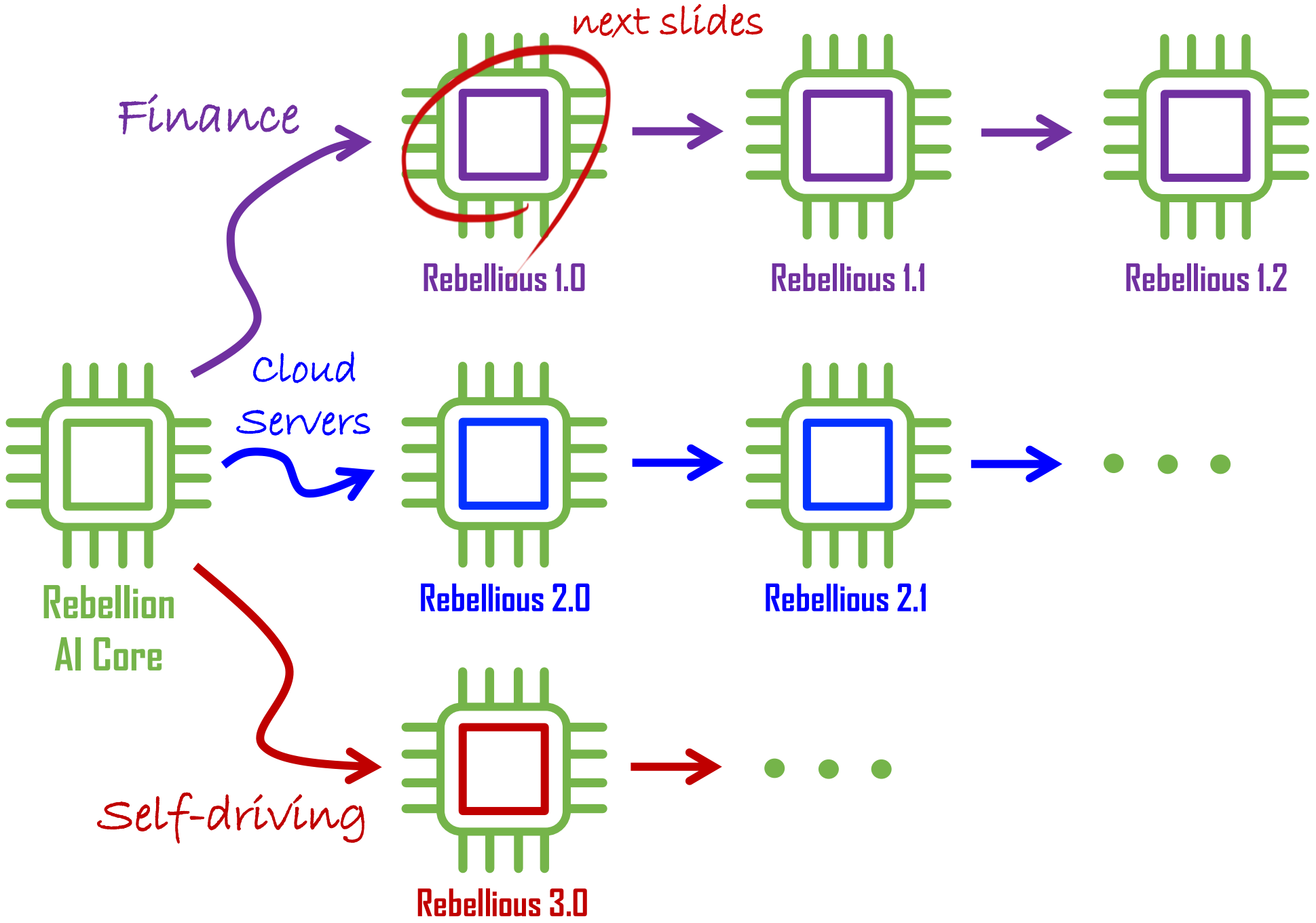
Domain-Specific, AI-Powered

System-on-Chip

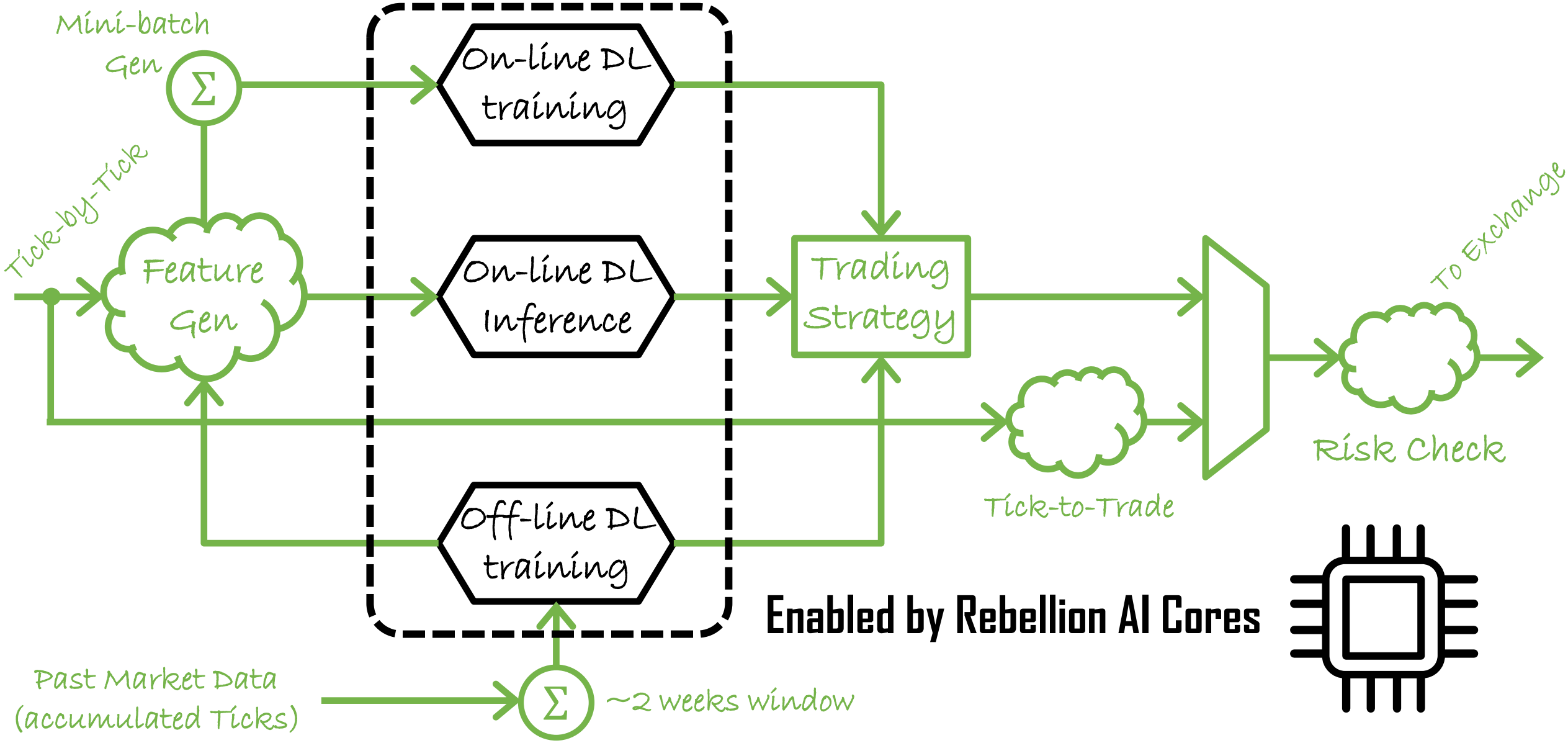
+

End-to-End Optimized Software

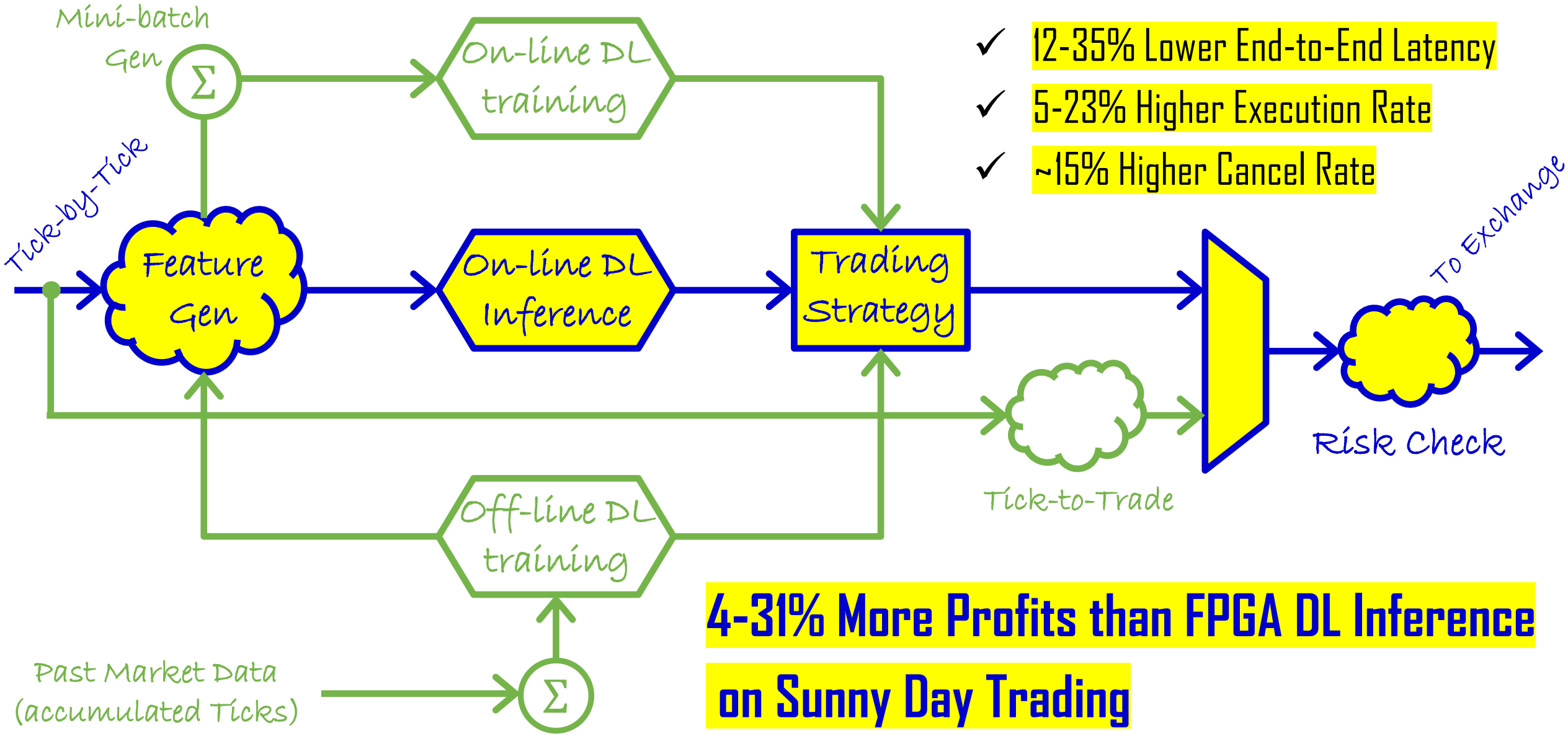
Vision


next slides
Finance
Rebellious 1.0
Rebellious 1.1
Rebellious 1.2
Cloud
Servers
Rebellion
AI Core
Rebellious 2.0
Rebellious 2.1
Self-driving
Rebellious 3.0

Rebellious 1.0: DL-based High Frequency Trading



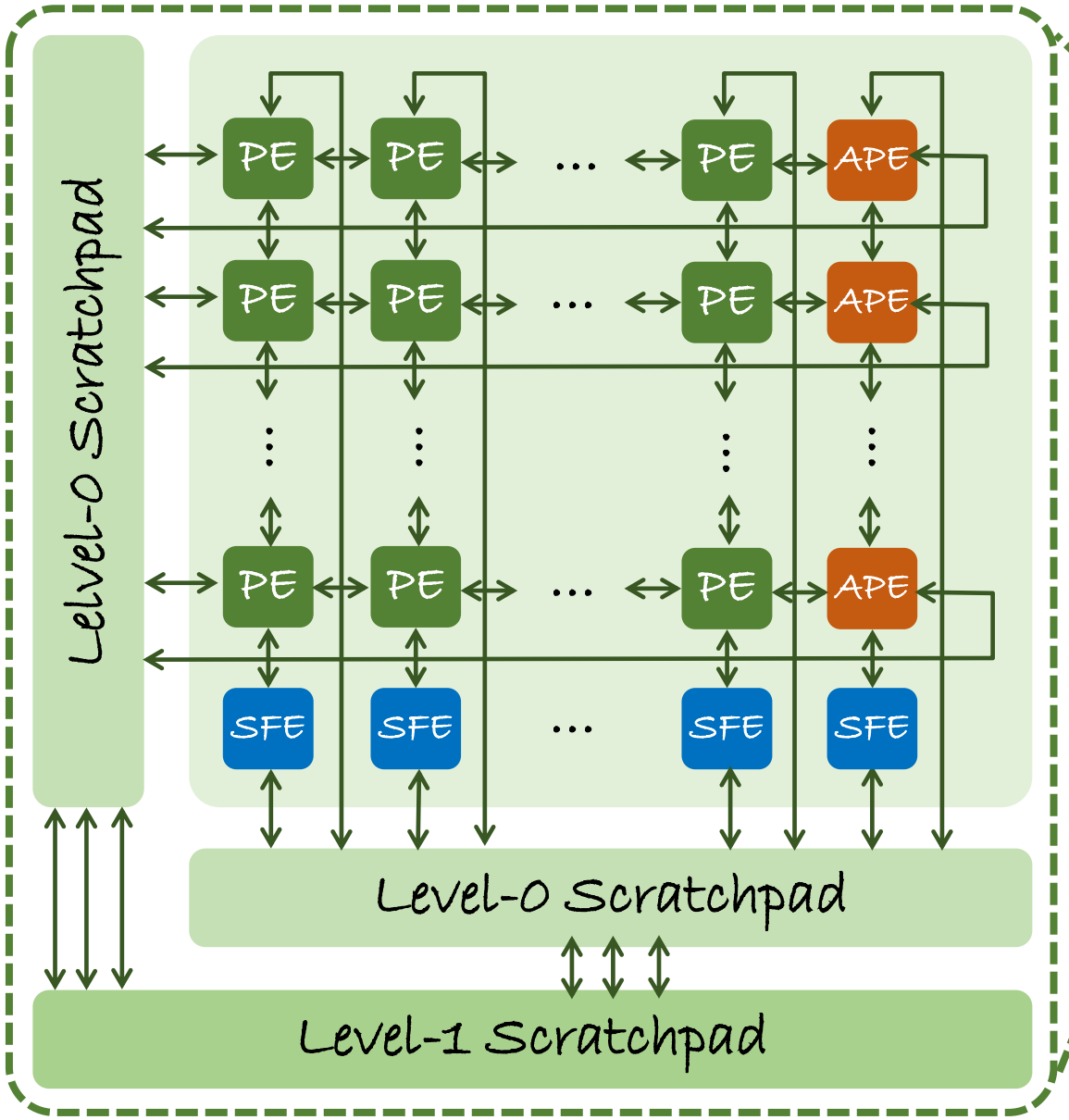
Latency-Sensitive Tick-to-DL Path

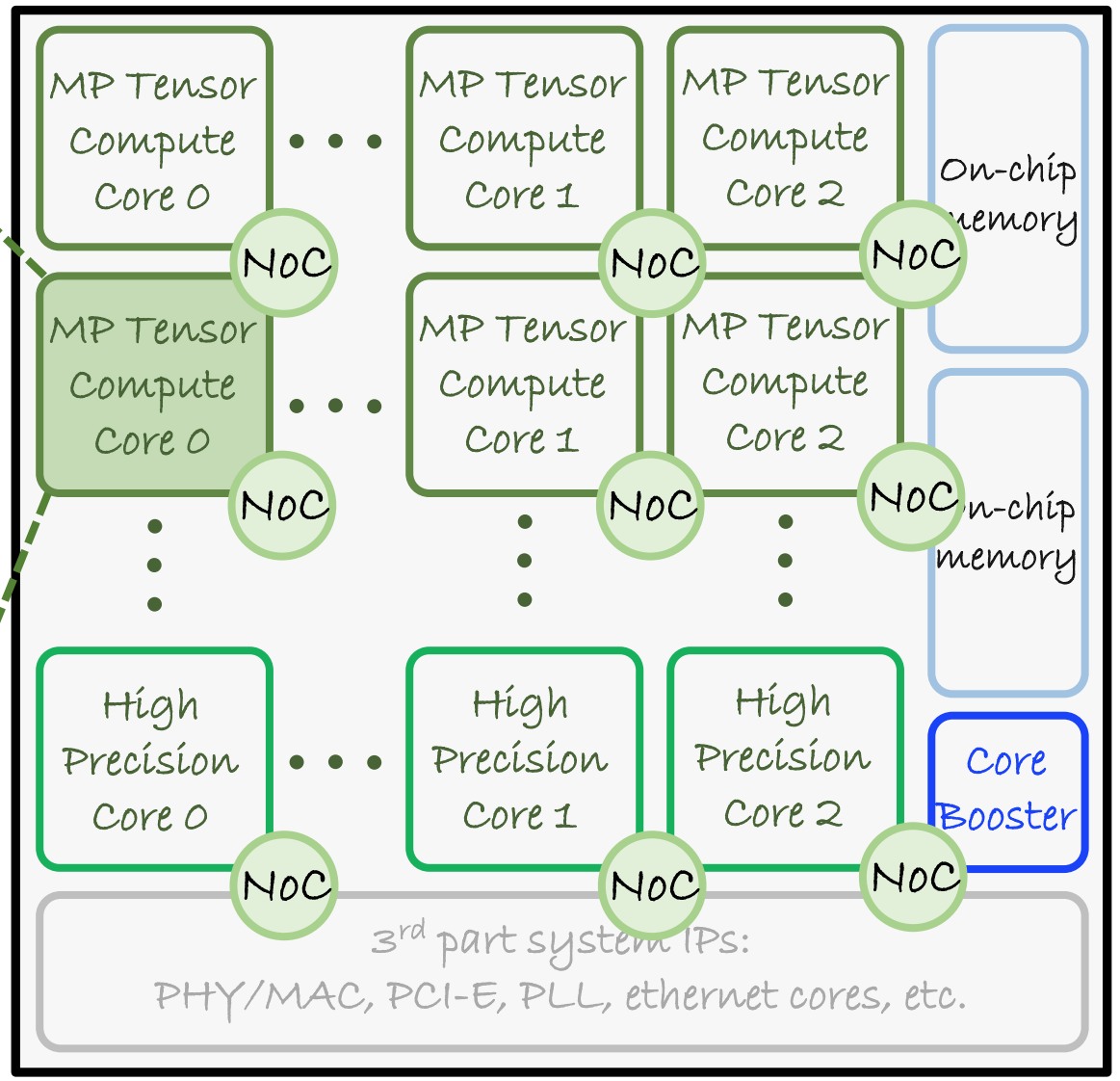


- ✓ 12-35% Lower End-to-End Latency
- ✓ 5-23% Higher Execution Rate
- ✓ ~15% Higher Cancel Rate

4-31% More Profits than FPGA DL Inference on Sunny Day Trading

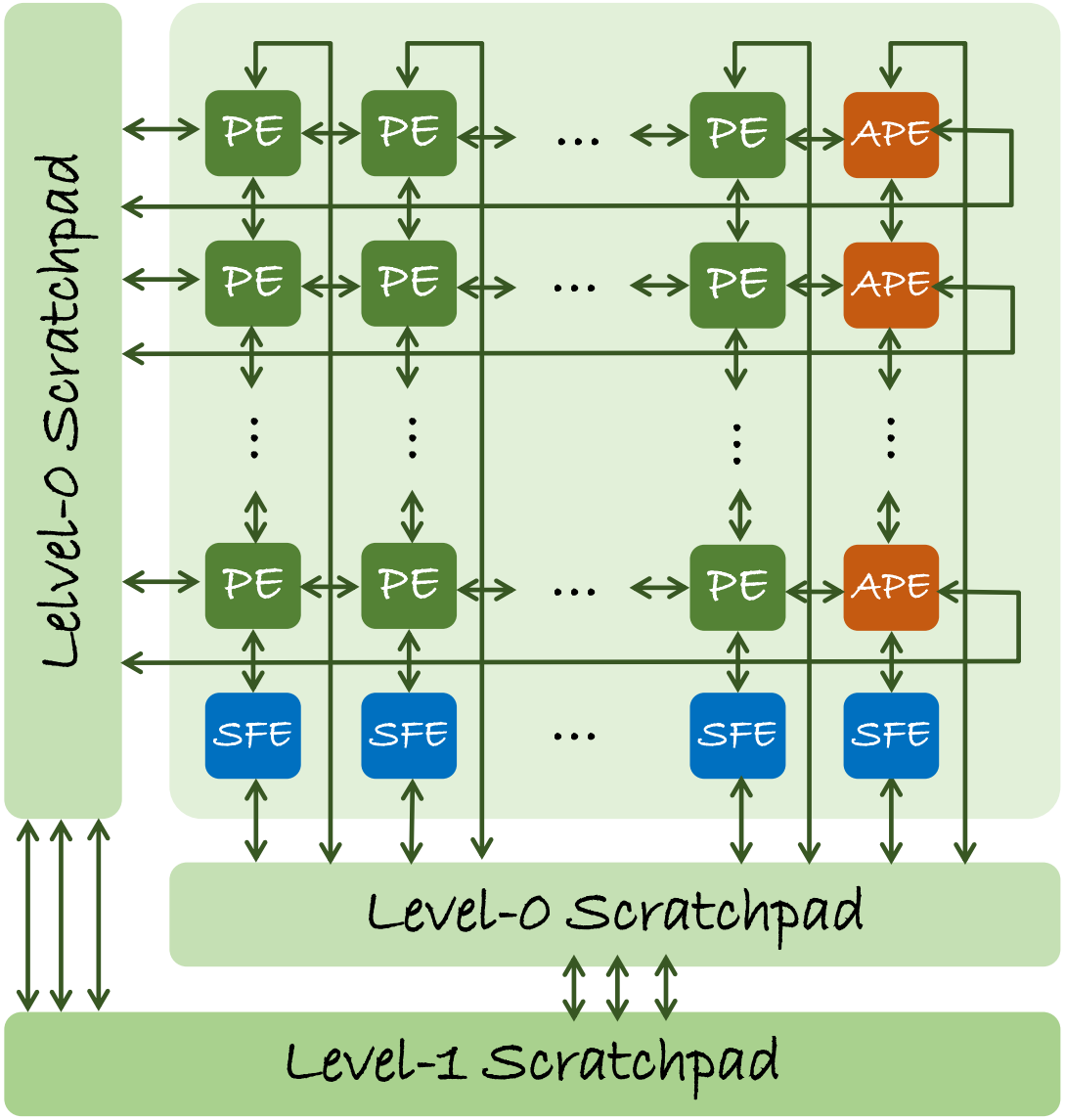
Rebellion SoC Overview





** Plug-and-Play Design platform **

Tensor Compute Core: Versatile-yet-Efficient AI Engine



Heterogeneous compute units on 2D Torus

FP16
FP8
INT8/4

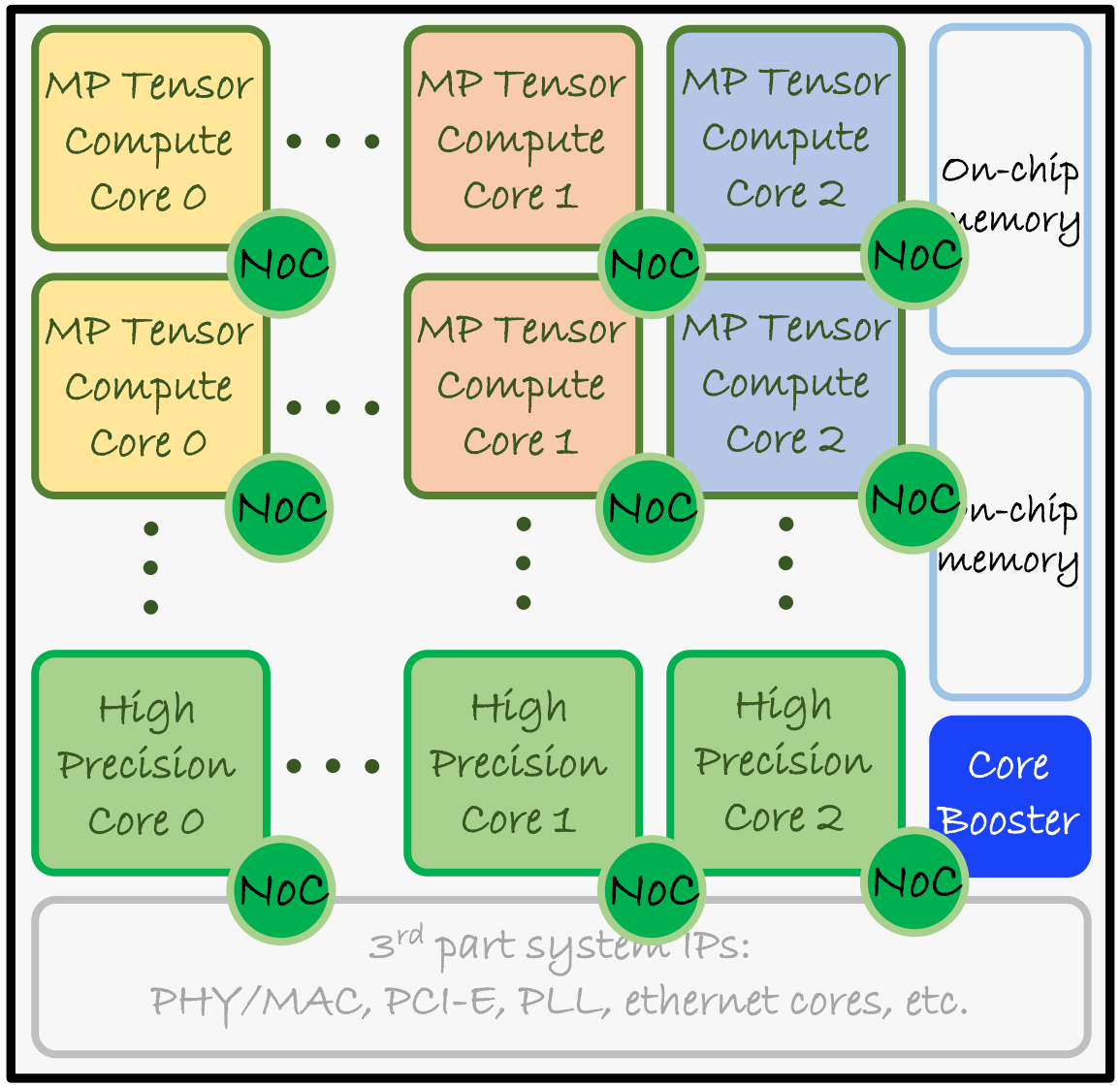
FP32
FP16

Analog Unit
INT1/2

- ✓ Mixed Precision
  - High Energy Efficiency > 1 W/core
  - High Throughput > 32 TOPS
- ✓ Modular Design within a Core
- ✓ Programming Model & Datapath Flexibility

Rebellion Core as a Design Platform, not an Ad-hoc

- ✓ Multiple Core Design Suites
  - Algorithm/HW Co-Optimization
  - Future-Proof Architecture
- ✓ Packet-Switched Mesh Network-on-Chip (NoC)
  - Scalable Network Backbone
  - Delivered by Tesla and SpaceX in 2019
- ✓ On-Chip Core Booster
  - Deep Learning Task/Resource Aware Controller
  - Much More Efficient than SW-level controls



**** Plug-and-Play Design platform ****

REBEL1.0 SW Programming Model

- ✓ Acceleration of all deep learning primitives across neural network topologies
  - Fine-grained, RISC-type ISAs
- ✓ Decentralized programming model for maximum flexibility
  - Carefully curated ISAs for programmable compute and memory agents
  - All efficient overlapping of data movement and compute
  - Hardware support for synchronization and dependency control
- ✓ Fine-grained programmability → high sustained utilization for target kernels

Hardware Development Status

Rebellions
Fabless

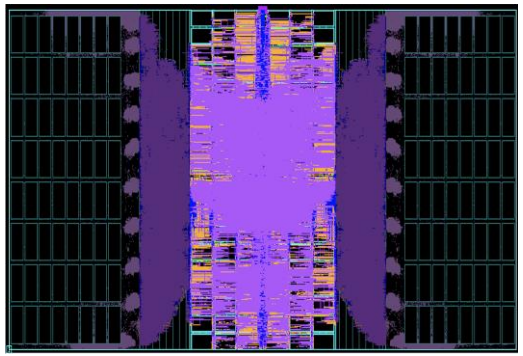
Successfully submitted final RTL on 2/15/2021

In Progress

ASICLAND
Design house

Tape-out on 5/21/2021

TSMC
Foundry



7nm ION (AI Core, DNC)

Initial post-layout simulation results are available now!

Software Development Status

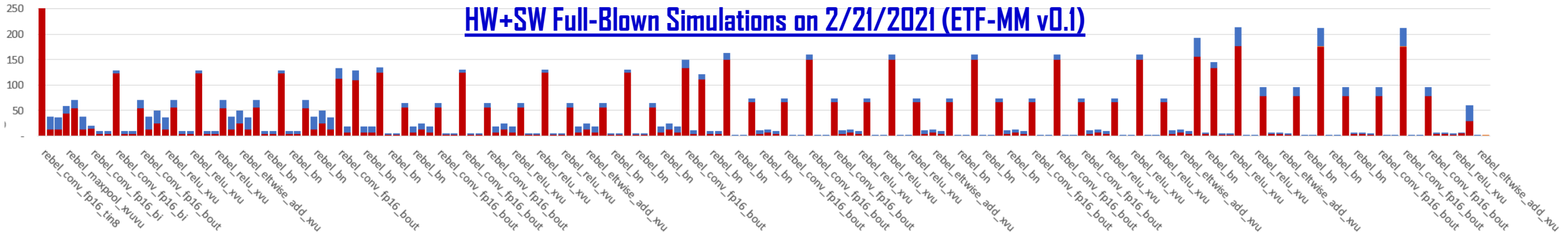
Front-end Compiler
- Design finished
- Opt finished

Back-end Compiler ↔ DL libraries (In Progress)

Runtime Driver
- Design finished
- Opt not yet

FPGA custom board for SW development

HW+SW Full-Blown Simulations on 2/21/2021 (ETF-MM v0.1)



리벨리온에서 함께 도전하실 팀원분들을 모십니다.

- Performance modeling and architecture simulator design (C++)
- Micro-architecture RTL design (Verilog, VHDL)
- UVM verification (SystemVerilog)
- Silicon physical design (Backend flow)
- FPGA engineering (Xilinx Vivado, etc.)

- Deep learning research (pruning, quantization, architecture search)
- Front-end and back-end Compiler design
- Runtime device driver design

contact@rebellions.ai