Evolving Finance

with Groq and Real-time AI

Introducing Groq

The fastest path to compute

Headquarters

Mountain View, CA

Applications

AI, Machine Learning, and converged HPC/AI

Employees

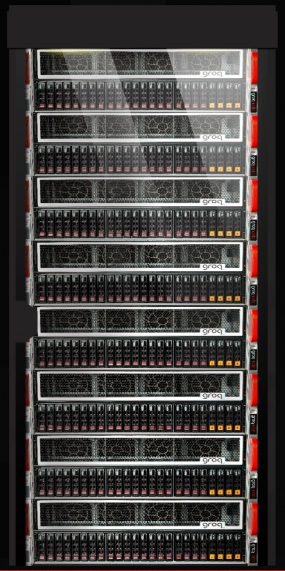
250

Flagship Products and Partners

GroqCard™

GroqNode™

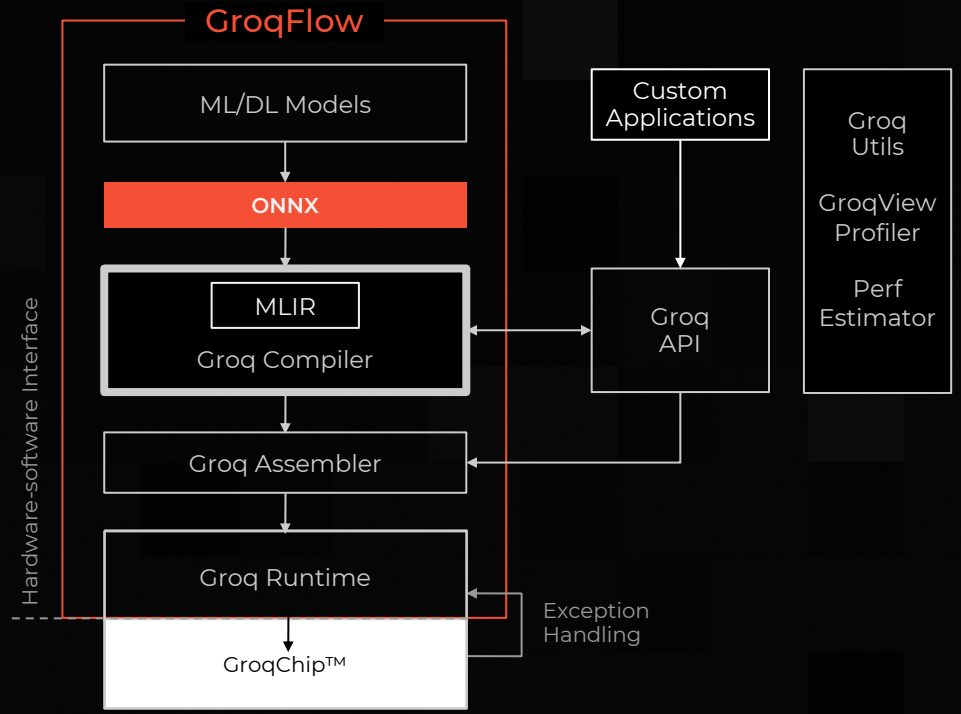
GroqRack™



GroqChip™

Dell R750XA

THE COMPONENTS OF GroqWare™ Suite



A Diverse Suite of Development Tools

Out-of-the-Box

Groq Flow provides out-of-the-box support for standard Deep Learning models

Fine Grained Control

Groq™ API provides finer grained control of GroqChip in order to support custom applications

Productivity Tools

GroqView™ Profiler provides visualization of the chip's compute and memory usage at compile time

Performance Estimator provides predictions for model compiler performance even before compilation

GROQ AND STAC-ML'S LSTM

Benchmarking Models

STAC

STAC provides comparative benchmarking for hardware to the finance industry

STAC	Newest benchmark is AI based, STAC-ML. ▪ Uses historical training data to make real-time trade decisions Results available in STAC report
Why LSTM?	Real-time, risk, automatic trading and stock advisers Used by hedge fund managers, portfolio managers, traders
Additional Usages	LSTM has a wider, universal appeal for any time-based analysis Fraud, Cybersecurity, Fault detection and simpler NLP models They are used in many verticals beyond just finance Other non-LSTM areas to explore: ▪ Anomaly detection (CC Fraud, Insurance, Cyber Security, etc.) ▪ Realtime preventive maintenance i.e. Industrial

Supporting the Financial Services Market

A growing list of applications requiring the very lowest latency and highest throughput

Not STAC Benchmarks

Fraud

Fraud Detection & Anti-money Laundering

Fewer false positives

Improved accuracy without performance compromise

Trading

Real-time Trading

Exceptionally low latency

Time-based predictions from tick and news feeds

Analysis

Quant Analysis Backtesting

High throughput and performance at scale

Automation

Customer Support

Recommendation engines with low latency language processing

Application	Performance Over GPU
Value At Risk (VaR)	10x
LSTM Real-time Trading / Advisor	20x
Options Calculation	20x
Fraud and Anomaly Detection	70x
Quantitative Analysis Backtesting	10x
Topic Modeling of SEC K-10 Filings	8x

YOU + groq™ + MAXELER™ a groq company

Simple development environment powered by GroqWare™ Suite

Detailed tuning with our low level APIs gives **total control**

We work closely together to **tackle the hardest problems**

DEPLOY GROQ PRODUCTS TO

Cloud **OR** On-prem

© 2022 Groq, Inc. |

groq™

Thank You.

MORE INFORMATION
www.groq.com

CONTACT US

Cameron McCaskill
cmccaskill@groq.com

Amr El-Ashmawi
amr@groq.com

Come say Hi.

STAC SUMMIT — Where strategy meets technology™ — London / November 10, 2022

SC22 Dallas, TX | hpc accelerates.

THE AI SUMMIT NEW YORK