

The logo for GroqWare, featuring the word "groqware" in a lowercase, white, sans-serif font with a trademark symbol (™) to the upper right.The logo for Groq, featuring the word "groq" in a lowercase, white, sans-serif font with a trademark symbol (™) to the upper right.

## GroqWare™ Suite.

A comprehensive software stack designed to accelerate a variety of HPC and ML workloads.

The foundation of our software-defined hardware approach is the GroqWare suite. Groq™ Compiler, Groq API, and Utilities comprises the extremely versatile software stack made to efficiently run a wide array of HPC and ML workloads. Groq Compiler, co-developed with the Tensor Streaming Processor (TSP) architecture, is an efficient and flexible tool to deploy state of the art deep learning models trained in frameworks such as PyTorch, TensorFlow, or ONNX. Groq API provides customers granular control of GroqChip™ processor. Finally, Groq provides utility tools such as GroqView™ profiler and visualizer to not only enhance the developer workflow, but simplify it all together.

### Key Features

**Open source, simple stack** eases deployment implementations with an open source driver/runtime and support for industry standard AI/ML frameworks.

**Groq Compiler** provides out-of-box support from regression to deep learning models.

**Groq API** enables finer grained control of GroqChip™ in order to support custom applications.

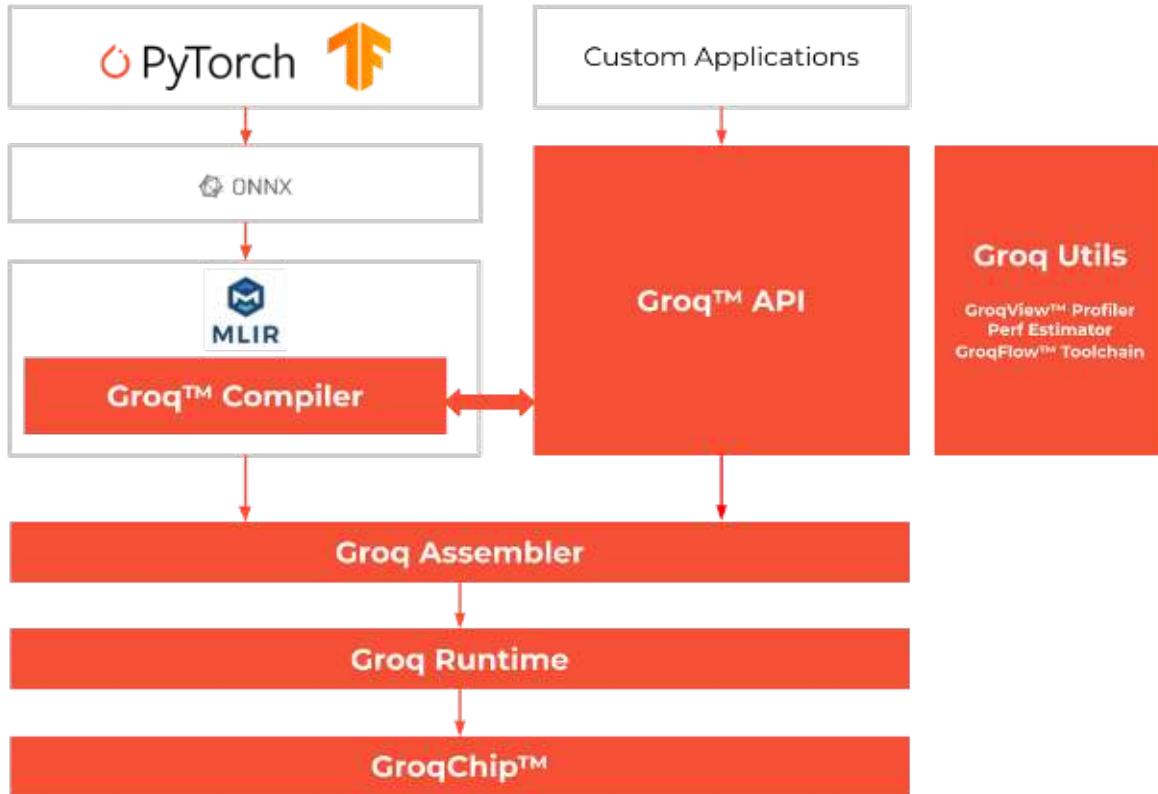
**GroqView Profiler** delivers a visualization of the chip's compute and memory usage for the program at compile time for debuggability and optimization.

**GroqFlow™ Tool Chain** enables a single line of Pytorch or TensorFlow code to import and transform models through a fully automated tool chain to run on Groq hardware.

**Performance estimator** enables accurate performance predictions for deep learning models even before compilation.

© 2022 Groq, Inc. All rights reserved. This document is approved for public release. Distribution is unlimited. For informational use only. Groq, the Groq logo, TruePoint, RealScale, GroqView, GroqChip, GroqCard, GroqNode, GroqRack, GroqWare, GroqFlow and other Groq marks are either registered trademarks or trademarks of Groq, Inc. in the United States and/or other countries. Other names and brands may be claimed as the property of others. Reference to specific trade names, trademarks or otherwise, does not necessarily constitute or imply its endorsement or recommendation by Groq. Other names and brands may be claimed as the property of others.

## Software Stack



## Specifications

Item	Description
Availability	Broadly available to Groq customers
OS	Ubuntu, RedHat and Rocky Contact Groq at <a href="mailto:info@groq.com">info@groq.com</a> for the latest versions
Framework Support	Current frameworks supported: PyTorch, TensorFlow, and ONNX Contact Groq at <a href="mailto:info@groq.com">info@groq.com</a> for additional frameworks.
Model Support	Computer vision, natural language processing, linear algebra, and more See current model examples available at <a href="https://github.com/groq">github.com/groq</a> .
Numerics	INT8, INT16, INT32 & TruePoint™ technology MXM: FP32 VXM: FP16, FP32

Groq Inc. HQ  
400 Castro St. Suite 600  
Mountain View, CA 94041

Mailing Address  
PO Box 1778  
Mountain View, CA 94041

[www.groq.com](http://www.groq.com)  
[support@groq.com](mailto:support@groq.com)

For more information visit [groq.com](http://groq.com) or contact us at [info@groq.com](mailto:info@groq.com).