

Automatic Speech Recognition - ASR

A real-time ASR accelerator for data centers

Product Brief



Leverages the Achronix[®] Speedster[®]7t FPGA on the BittWare[®] VectorPath[®] PCIe accelerator card

Industry-leading WER, concurrent real-time streams, and lowest latency on a single card

Key Features

- Reduces costs up to 90%[†]
- Low-latency, real-time ASR for streaming applications
- Transcription of thousands of speech channels per PCIe card
- WebSocket API for drop-in replacement to cloud transcription services
- High-quality English language ASR as standard
- Optional customization for bespoke data sets
- Power dissipation suitable for 1U servers
- Standard dual-width PCIe card
- 7-10% word error rate (WER)

Equivalent Capacity (Simultaneous Speech Channels)



20×

=



1×

+



1×

One accelerator card attached to one server unit can replace 20 standard CPU servers



Reduces CapEx

- One accelerator card in one CPU server replaces 20 servers or 10 GPU accelerators, which means shorter cloud usage times or reduced equipment costs



Reduces energy

- Uses 90% less energy for the same number of speech channels, reducing energy costs and improving ESG ratings



Enables rapid scaling

- Runs on off-the-shelf industry-standard PCIe cards, enabling simple and rapid adoption
- Up to 20x more capacity can be achieved within existing rack space, avoiding the need for costly and time-consuming infrastructure expansion



Simple to integrate

- Delivered with complete acceleration stack including bitstream, drivers and efficient WebSocket API to facilitate simple integration into existing service provisions



Optional customization for bespoke data sets

- Requires at least 10k hours of audio data with accompanying transcriptions
- Each audio clip and transcript pair should refer to a maximum of 20 seconds of audio

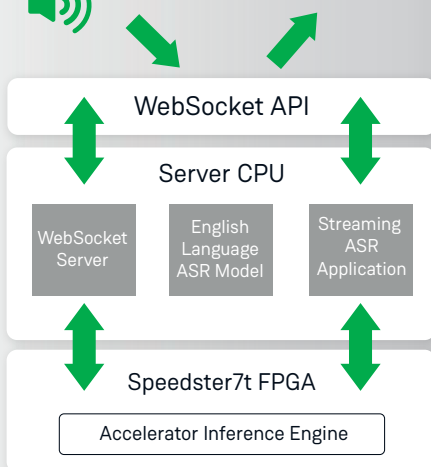
Arrange an Evaluation Today

Hardware (VectorPath S7t-VG6 Accelerator Card)	
Form Factor	• Standard dual-width PCIe card
Electrical	• Power (TDP) 230W
Environmental	• Operating temperature 5°C to 35°C
Cooling	• Standard, dual-width, passive heatsink

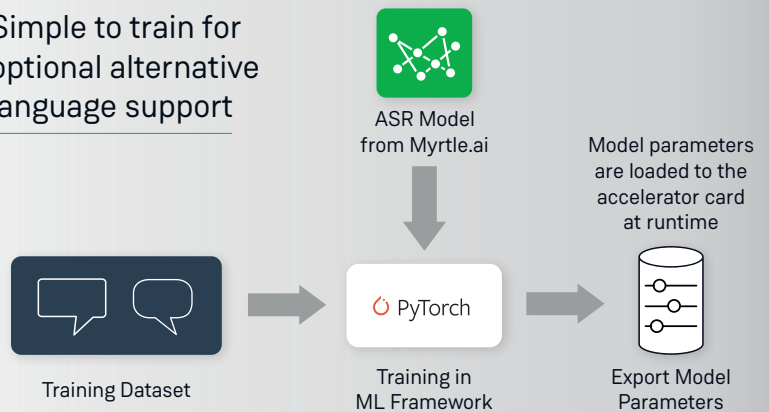
Software Stack	
Language Support	• English (other languages on request)
Vocabulary	• Standard (can be trained for specializations)
API	• Simple WebSocket API

Simple interface through WebSocket API

The quick brown fox jumps...



Simple to train for optional alternative language support



Optional customization for custom data sets

Key Features

- Enables support for alternative languages and additional vocabulary
- Compatible with the popular deep learning framework PyTorch
- Retrainable reference model to deploy with custom data sets and languages