

JUN 21, 2023 • NEWS & ANALYSIS

Renesas Picks Its Battle on Edge AI

Venturing into a broad and fragmented edge AI market is a safari fraught with challenges.



Edge AI

By [Junko Yoshida](#)

What's at stake?

In pursuit of the edge AI market, Renesas must bridge two disparate worlds. The more probabilistic AI realm deals with data and model creation. The embedded world – more deterministic in nature – does linear programming. Renesas must transit between these worlds without jeopardizing its status in either.

Embedded system designers are curious about AI, but they aren't necessarily interested in coding. Put bluntly, AI makes them uncomfortable.

Herein lies the dilemma for leading MCU/MPU suppliers, including Renesas, who covet the seemingly large edge-AI market.

Developing an ultra-low power MCU/MPU architecture equipped with reasonable flexibility and programmability is the basis. But to win the edge-AI race, MCU/MPU suppliers must have a targeted plan and a specific strategy.

First, they must know customers' wants and systems' limitations in each segment of the market.

Second, MCU/MPU suppliers cannot let themselves be pulled in too many directions. They must pick a spot that enables them to go up the stack and fully support their customers.

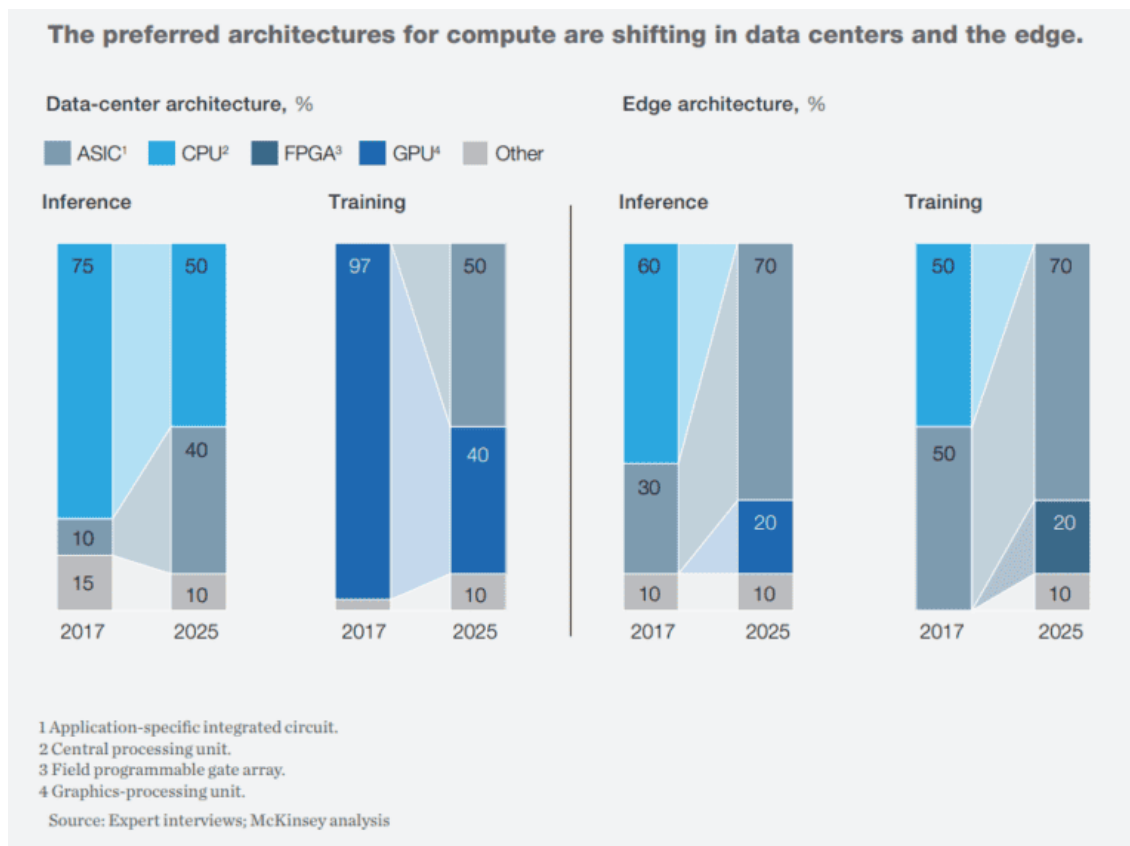
For now, Renesas has narrowed its edge-AI focus to the Heating, Ventilation, and Air Conditioning (HVAC) industry. "It's a very targeted approach," said Kaushal Vora, Renesas' senior director of business acceleration and ecosystems, in a recent interview with the *Ojo-Yoshida Report*. He explained that Renesas' edge-AI solutions for HVAC range from "a hardware, firmware reference design and sensor instrumentation to a production-grade data set." Renesas installs a professional grade lab where its customer OEMs' HVAC system is brought in, and "collects quality data for them."

Renesas offers full edge-AI service to customers who might not know how best to apply machine learning to HVAC problems. This requires some heavy lifting on Renesas' part, meaning, Vora acknowledged that Renesas can initially service "only a few strategic customers."

However, a few success stories with big HVAC customers could trigger interest in AI among more embedded-system customers. Renesas is designing similar "getting started packages" to hook more buyers on applying edge AI to their systems.

Scalability

To successfully add AI to a range of embedded systems, the key is scalability in edge-AI solutions. Executive vice president Sailesh Chittipeddi claimed that Renesas is ready to scale with an arsenal of MCUs and MPUs to address diverse customer requirements, "based on the memory, the performance and the power consumption that's needed."



For Renesas's simplest solutions, Chittipeddi said, "You can go with the 16-bit microcontrollers. As you get more advanced, we offer 32-bit, and 64-bit solutions. But then eventually [migrate] to the integrated neural processor." Renesas offers an AI accelerator called "DRP AI" device, added Chittipeddi.

When we pitch our AI solutions, we probe where they want to start and what they're trying to accomplish.

Sailesh Chittipeddi

Renesas also offers an FPGA solution, ForgeFPGA. "It's an FPGA consuming super-low power and costing at 50 cents or below," he noted. The company licensed an FPGA fabric from Flex Logix and optimized it with a very small lookup table – 5000 or less. "The beauty of using an FPGA is you can actually do parallel processing, which you cannot do with a CPU," explained Chittipeddi.

FPGAs are generally known to be expensive, but ForgeFPGA is comparatively cheap. Renesas is hitting the sweet spot, said Chittipeddi. While ForgeFPGA is initially available as a standalone product, it can be integrated into Renesas' other chip solutions as a System in a Package (SiP), added Chittipeddi.

But hardware doesn't drive AI chip sales. The key is how much vendors know about what their customers need. "When we pitch our AI solutions, we probe where they want to start and what they're trying to accomplish," said Chittipeddi.

Given the plethora of use cases, Vora added that an important criterion for chip suppliers is to grasp a system's restrictions. Renesas asks customers, said Vora, "Where is the data? Where does the data need to be operated on? Can you afford to transport the data to another layer of the network? What about security concerns when transporting the data? And what about the efficiency of operating the network?"

Knowing the variables helps scale Renesas' solutions.

Large Language Models

Rapidly advancing AI models are also a concern among customers of AI chips.

Some system vendors have suggested that the growing popularity of emerging large language models (LLMs), such as ChatGPT, has flipped the AI world. They wonder if the highly optimized hardware they recently adopted can do the math for LLMs.

Answer: It can't.

Although not exactly downplaying their importance, Chittipeddi claimed that LLMs didn't sneak up on Renesas. "We are working on it," he said, but declined further discussion.

For the time being, LLMs could be just a sideshow, diverting system designers from the real problems facing edge AI.

Vora said, "Generative AI has become so popular because OpenAI has demonstrated it is able to train the model to a point where the outcome is usable."

The lack of complex AI models isn't today's edge AI issue. The problem is the insufficient integrity of data sets.

He stressed, “People often equate AI to the complexity of the model.” But the lack of complex AI models isn’t today’s edge AI issue. The problem, in Vora’s view, is the insufficient integrity of data sets.

As the world collects more data, it must cope with “data gravity,” said Vora. “As data gets transported up the food chain, you start losing data gravity ... integrity.”

Vora observed, “If you can work with data at the source, you have the most accurate AI with integrity and explainability.”

Too often, AI truncates data as it feeds it to the cloud, changing outcomes and leading to faulty conclusions, cautioned Vora. “That’s a big part of problems with AI today.”

Efficiency in operating the network is another issue with edge AI, Vora noted. “A round-trip time — or latency — adds a few milliseconds,” also affecting outcomes.

‘Microverticals’

The semiconductor industry yearns to extract more value from the AI market.

Historically, the chip industry got 10-percent value from the smartphone market, and perhaps 15 percent in the PC era, estimated Vora. “If we learn from our missed opportunities,” he said, “we could extract 50 percent value from the AI market,” largely by going vertical, building complete stacks, and enabling the right ecosystem. This has been Nvidia’s approach.

Enabling Nvidia’s success in the AI market was CUDA – Compute Unified Device Architecture platform – a software framework developed by Nvidia. It enticed software developers to use CUDA-enabled GPUs for general-purpose processing.

Renesas has taken a cue from McKinsey’s report, “[Artificial-intelligence hardware: New opportunities for semiconductor companies](#),” In it, McKinsey advises semiconductor companies to focus on “microverticals” in the AI market.

By microverticals, Vora explained, it means picking a specific field out of diverse edge AI applications and building complete vertical AI stacks for customers.

Renesas has identified HVAC as an initial microvertical that presents a value creation opportunity.

Instrumental to Renesas’ microvertical approach was Reality AI, which Renesas acquired last summer.

Reality AI, based in Columbia, Maryland, was one of Renesas’ original ecosystem partners. It offers a range of embedded AI and TinyML solutions for advanced non-visual sensing in automotive, industrial, and commercial products. Renesas saw opportunity in combining Reality AI’s Edge AI software development environment with its own MCU and MPU portfolios. The acquisition enabled Renesas to offer “advanced signal processing, machine learning, and anomaly detection on every MCU/MPU Renesas core.”

Given the widespread discomfort for working with AI among embedded system designers, Vora explained, “The more vertical integration we can provide to customers, and the more guidance and tools we can provide them, the better ecosystem we enable for customers to define where AI is going to go.” Renesas’ plan is to gradually add microverticals and more ecosystem partners, identifying tools useful for other edge-AI applications.

From a starter package to subscription models

By working with customers, Renesas is also finding opportunities to monetize.

Offerings like the ‘getting started package’ create opportunities to bond customers more closely to Renesas’ support system.

Today, about 11 companies have gotten the “starter package” in microvertical fields, according to Renesas. They pay a one-time fee ranging from \$20,000 to \$70,000. In the next step, Renesas offers a tiered subscription model of recurring fees that range from \$36,000 to \$120,000 depending on the number of users’ compute requirements.

Renesas’ business is in making money by selling hardware. However, offerings like the “getting started package” create opportunities to bond customers more closely to Renesas’ support system.

This model, similar to Nvidia’s, is called “transfer learning,” explained Vora. “You take a model that’s been pre-trained for a particular task to about 80-90% accuracy. And then you can incrementally teach that model to do something the customer wants to do specifically. Some customers are comfortable with that, while others may not be.”

Either way, it creates “an additional opportunity for further monetization,” said Chittipeddi.

Bottom line:

Whether in the market for Internet of Things (IoT) or AI of Things (AIOT), there have always been too many use cases, variations, and fragmentation among AI applications. By attracting key ecosystem partners, Renesas appears to be navigating the edge-AI maze better than the average bear.

Junko Yoshida is the editor in chief of The Ojo-Yoshida Report. She can be reached at junko@ojoyoshidareport.com.

PREVIOUS

[Chiplets Are Still a Work in Progress](#)

Register Free Today: <https://ojoyoshidareport.com/subscriptions/>



The Ojo-Yoshida Report © 2021-2022

[Home](#) [About](#) [Partner Content](#) [This Month](#) [Contact](#) [Account](#) [Subscribe](#) [Podcasts](#)



 **Subscribe**