# Are ghiplets the next evolution stage in GPU designs?

edn.com/are-ghiplets-the-next-evolution-stage-in-gpu-designs/

When GPUs are implemented in a chiplet arrangement, can they be called ghiplets? Semiconductor industry professional Chetan Arvind Patil thinks so. In his blog "The Rise of Semiconductor Ghiplet," Patil outlines the benefits of GPUs in chiplet architectures compared to monolithic GPU designs currently struggling with lower yield rates for large and complex GPU devices.

In other words, all the rendering is handled by chiplets instead of a big-compute GPU. Take the case of AMD's Radeon RX 7000 series, one of the first chiplet-based GPUs, which incorporates a large graphics die and several memory dies. The Silicon Valley-based semiconductor firm is now considering using even more chiplets in a GPU design.

**Figure 1** The Navi 31 GPU, part of the Radeon RX 7000 series, is based on the RDNA 3.0 chiplet architecture and is built on TSMC's 5-nm manufacturing process. Source: AMD

The "ghiplet" approach has an inherent benefit for high-compute applications like artificial intelligence (AI): memory dies can be spread around the main GPU die. Thus, the ghiplets can help create a compute-memory balance. The modularity of chiplets helps optimize specific functions like memory handling.

Besides memory, chiplets offer faster and more efficient data handling with advanced interconnect implementation. This is specifically relevant for high-performance computing (HPC) applications in AI, data analytics, and scientific research.

GPUs have been slow to migrate to chiplets; GPU powerhouse Nvidia says that its recently launched graphics device Blackwell is not a chiplet. Despite much speculation, Nvidia has stuck to its monolithic guns for now. That's partly because GPUs are far more complex than CPUs.

However, as GPUs move beyond graphics and gaming, chiplets could bring a lot of flexibility and scale to GPU designs. So, AMD archrival Intel has entered the ghiplet space with its Max Series GPUs. It has over 100 billion transistors packaged in 47 different chiplets—Intel calls it tiles—with up to 128 GB of memory.

**Figure 2** The Max Series GPU, codenamed Ponte Vecchio, incorporates EMIB 2.5D and Foveros 3D packaging technologies and stacks chiplets or tiles on top of one another for greater processor density. Source: Intel

The suitability of the modularity in chiplet architectures for GPU designs seems a no-brainer. AMD and Intel have already jumped the fray, and Nvidia's move is much awaited. Especially when AI applications and their need for memory capacity paired with GPUs increasingly calls for a compute-memory balance.

**Related Content**