

Harnessing Generative AI: Key Innovations and the Power of Polymer Capacitors



The rapid rise of generative AI is revolutionizing how content is created and consumed, driving demand for increasingly complex hardware solutions. This article delves into the advancements in hardware technology necessary to support generative AI services while emphasizing the pivotal role of high-performance capacitors in ensuring reliable operation. With a focus on Panasonic conductive polymer capacitors, it explores how innovative designs enhance computational efficiency and stability, ultimately powering the future of AI-driven applications.

Utilizing Generative AI and the Strength of SP-CAP™ Capacitors

As generative AI services rapidly evolve and gain popularity, the complexity of the hardware required to support them is also increasing significantly. This shift calls for an examination of the factors contributing to the growing computational demands and how hardware is adapting in response. To fully understand these changes, it is important to explore the background of generative AI development, review examples of compatible hardware, and analyze the role of capacitors in enhancing computational performance.

Development of Generative AI

Generative Artificial Intelligence has gained significant attention recently, transforming how AI creates new content rather than just analyzing it. Unlike traditional AI, which focuses on identifying images, generative AI like “ChatGPT” produces content based on user input. Since its launch in November 2022, ChatGPT became the fastest web service to reach 100 million registered users by January 2023, generating immense global interest. This technology supports diverse applications, from answering questions and drafting documents to composing music and generating code. While it holds great promise, it also raises concerns about copyright, job displacement, and its potential expansion into robotics and heavy industry.

Exploring Generative AI Mechanisms and Concerns Over Inaccurate Responses

Generative AI is revolutionizing content creation through two key processes: learning and inference. In the learning phase, a model is built by analyzing large datasets to capture their features, which can take months for high accuracy. In contrast, the inference phase quickly generates responses to user inputs using these pre-trained models, typically within seconds, which is critical since users tend to shy away from slow services. Despite its potential, generative AI faces challenges, particularly in generating inaccurate content—a phenomenon known as “hallucination,” where the AI produces plausible yet false information. One effective strategy to combat this issue is to enhance the data volume and computational parameters during the learning phase. This approach has reportedly improved the accuracy of generative AI models, such as OpenAI’s ChatGPT, highlighting the ongoing evolution of this promising technology.

Hardware Supporting Generative AI



Currently, services like ChatGPT and other generative AI applications operate as web services, leveraging numerous servers in cloud data centers for processing. Traditionally, each server relied on one or two CPUs for computational tasks, but by the 2010s, the benefits of circuit miniaturization for CPU performance had diminished, making further enhancements difficult. This led to the adoption of accelerators (coprocessors), which involve adding high-efficiency processors to servers to boost performance. This shift has enabled the practical use of AI that requires extensive data processing for web services, with GPUs or ASICs typically serving as the chosen accelerators for generative AI.

How Processors Are Used In Generative AI

1 CPU (Central Processing Unit)

CPU is essential in a server as it is responsible for the fundamentals of web services including versatile software processing. Meanwhile in specific processes, processors with higher parallel processing performance such as GPU and ASIC can provide higher-efficiency processing than CPU. For this reason, in AI, a combination of these processors in addition to CPU has become common place.

2 GPU (Graphic Processing Unit)

As its name suggests, its conventional main application was image processing for games. However, as similar parallel processing is used in learning and inference, use of GPU has spread widely in AI data processing since the middle of 2010s. The software flexibility of GPU is also considered attractive for responding to constantly evolving AI algorithms, thus it has gotten favorable reception from many AI researchers and AI service engineers. Therefore, GPU is functioning as a foundation for the evolution and development of generative AI.

3 ASIC (Application Specific Integrated Circuit)

As these processors are designed for specific applications, they can provide high performance while saving electricity, reducing operation cost, which are considered to be attractive. However, as they are not as flexible as GPU and the processes are fixed to a certain degree, they are developed upon sufficient examination of target processes and development scale. As this development is considered to require a lot of capital and labor, ASIC is utilized mainly by major cloud vendors that have high financial and development abilities. However, there are processor manufacturers that encourage user companies to utilize ASIC by providing sufficient technical support.

Market and Equipment Insights

The demand for enhanced data processing in hardware is surging, fueled by the need for higher accuracy and faster service delivery in generative AI. Consequently, high-end specifications in generative AI processors are now common, featuring configurations like two CPUs and eight accelerators. Newer models are raising the bar further with eight CPUs and closer integration of CPUs and accelerators to enhance communication speed and power efficiency. Strategies such as deploying mega ASICs for faster learning and embedding AI functions in mobile device processors are also gaining traction. As the landscape evolves rapidly, a key focus remains on leveraging cutting-edge technology for peak performance and optimal configurations.

Service element	Expectation of service companies	Hardware task
① Accuracy	Want to get more users through higher-accuracy AI	Larger-scale data processing is required. →Higher-performance hardware is required.
② Time	Want to launch and update services quickly. Want to reduce response times.	Learning and inference time must be reduced. →Higher-performance hardware is required.

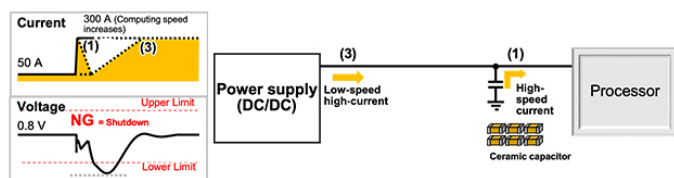
To ensure the stable operation of high-performance processors, it is essential that the accompanying electronic components meet stringent requirements. In particular, capacitors are critical for efficiently powering the processors required for generative AI applications. Innovations from Panasonic Industry offer effective solutions to these challenges, highlighting advancements in capacitor technology that significantly enhance both computational performance and reliability.

The Evolution of AI Services and the Essential Role of High-Performance Capacitors in Powering Generative AI Processors

Capacitors play a critical role in generative AI hardware, particularly in the power lines of processors that perform essential calculations. Just as people require food to think and move, processors need a steady current supply to execute their tasks. This supply is managed through the power supply lines, but processors often demand such high currents at rapid speeds that the power supply may struggle to keep up. When this happens, the voltage necessary for stable processor operation can drop significantly, leading to instances where the processor ceases functioning. Consequently, this cessation halts both learning and inference processes. To mitigate these issues, capacitors act as reservoirs, swiftly supplying current to the processor. However, if a capacitor's capacity is insufficient, the voltage can still drop, resulting in processor failure.

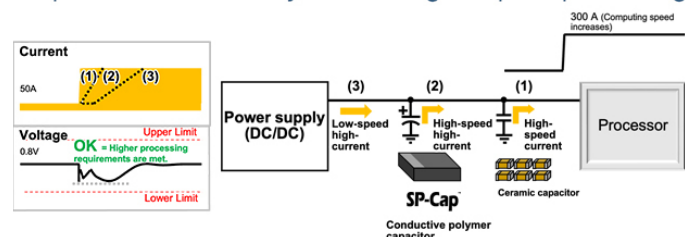
When Conductive Polymer Capacitors Are Not Used

- Current supply is insufficient and the voltage drops
- The voltage cannot be stabilized and it deviates from the acceptable range
- The system shuts down and data processing is aborted



When A Conductive Polymer Capacitor Is Used

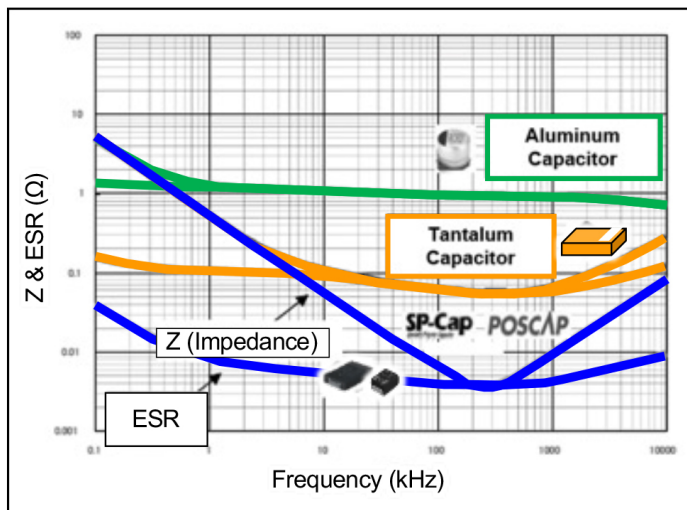
- Sufficient current is supplied, preventing voltage drop
- The voltage remains at an acceptable range.
= The voltage stabilizes
- Operation can smoothly transit to higher-speed processing



Features of Conductive Polymer Capacitors and Their Advantages for Generative AI Processors Compared to Other Capacitor Types

In power supply design, the best approach is to utilize a combination of capacitor types and ratings tailored to the design requirements. Conductive polymer capacitors offer distinct advantages for AI processors over other types. They are well-balanced capacitors known for their low equivalent series resistance (ESR) and enhanced safety compared to standard electrolytic and tantalum capacitors. Additionally, they provide stable capacity with fewer components when compared to ceramic capacitors. These benefits make conductive polymer capacitors a popular choice in power supplies for AI processors, which demand stringent electrical performance and reliability.

General Electrolytic and Tantalum Capacitors

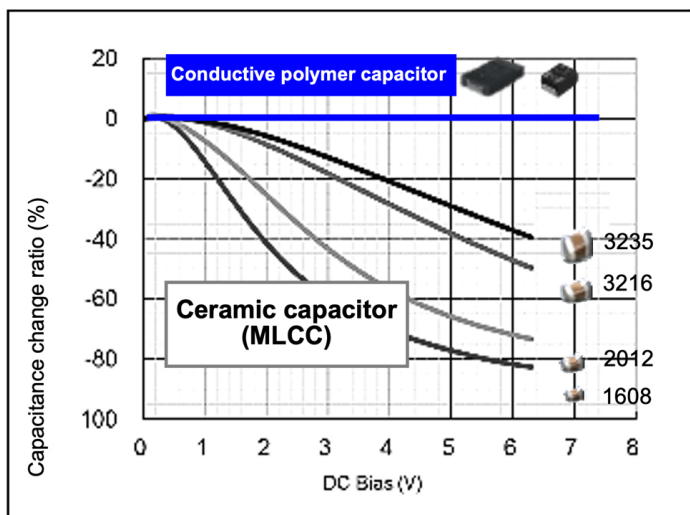


One-digit Lower ESR and Impedance
 - Suppress voltage fluctuation at high-speed current

Reliability and safety risks during use at high temperatures	
Aluminum electrolytic capacitor Liquid → dry-up 	Tantalum capacitor Solid (MnO ₂) + Weak self-healing → fire-up *O ₂ = Flammable
Conductive polymer capacitor Solid → no dry-up 	Conductive polymer capacitor Solid (Polymer) + Strong self-healing → no fire-up

Excellent Safety
 - Contribute to highly-reliable design

Ceramic Capacitor (MLCC)



Stable High Capacitance Not Influenced by Voltage or Temperature
 - Prevent voltage drop caused by insufficient power supply

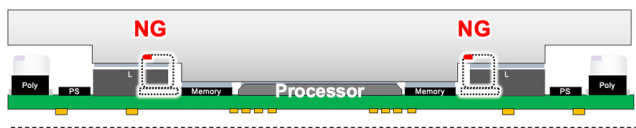
Ceramic capacitor (MLCC)	Conductive polymer capacitor
MLCC100uF x13 MLCC22uF x11 MLCC10uF x5 	SP-Cap 2V330uF x4 MLCC100uF x0 MLCC22uF x3 MLCC10uF x1
Transient Response 23A (15A/us, No-Delay) Sw Power Supply Vin=12V, Vout=1V, L=0.19uH@500kHz x1phase 	
Ceramic capacitor (MLCC) [Vpp -51 mV, +53 mV]	Conductive polymer capacitor [Vpp -53 mV, +31 mV]

A Fewer Number of Components
 - Contribute to simpler design

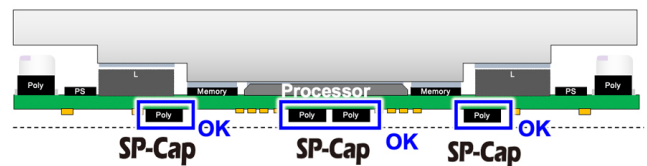
Advantages of SP-Cap™ Capacitors for Compact Size and Low Profile Benefits

AI processors require significant currents to achieve high computing power, steadily increasing processor sizes. Consequently, power supply dimensions are also growing to accommodate these high current demands, often resulting in limited capacitors' space on the board surface. Panasonic addresses this challenge with its SP-Cap Polymer Aluminum Capacitors, which come in various low-profile options, 2mm in height or less. These capacitors can be mounted on the back of the board, effectively solving the mounting space problem as it becomes more pressing in modern designs.

Conventional Can Type Capacitor

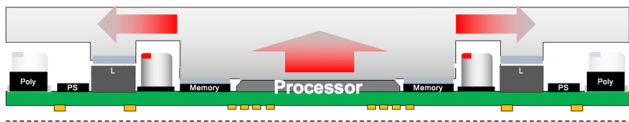


Conductive Polymer Aluminum Electrolytic Capacitor (SP-Cap)

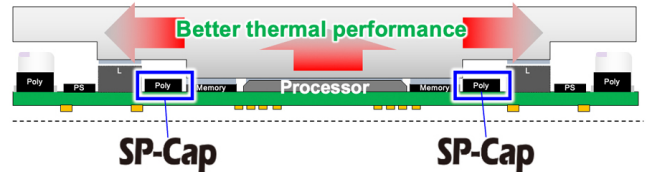


On the other hand, even if it is possible to place the capacitor on the surface, the heat dissipation mechanism can be designed to be simpler and with better heat dissipation properties than when a tall capacitor is used, thereby improving the reliability of the hardware

Conventional Can Type Capacitor



Conductive Polymer Aluminum Electrolytic Capacitor (SP-Cap)



Unmatched Reliability and High Voltage Options in Polymer Capacitors

Long-term reliability is crucial for AI hardware, particularly during the intensive learning phase when processors run at full capacity for extended periods. These processors are engineered for durability to minimize failure risk. As processing power increases, thermal management becomes more critical; thus, components in proximity must also endure high temperatures and maintain reliability over time. To address these demands, Panasonic has created conductive capacitor products designed for prolonged use in high-heat environments, ensuring dependable performance even in elevated temperatures around the processor.

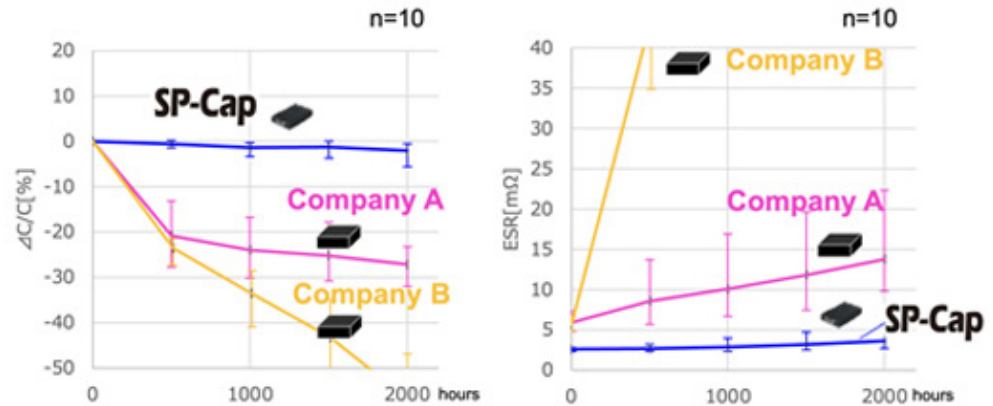
A comparison of ESR and capacitance changes in durability tests (105°C for 2000 hours) among similar capacitors reveals that, despite sharing similar specifications, Panasonic's SP-Cap™ capacitors exhibit minimal characteristic alteration, demonstrating exceptional reliability and voltage stabilization after long-term use.



Changes in ESR and capacitance during durability testing of SP-Cap™ Capacitors and similar products from other companies.

Durability Test 105°C 2000 H (rated voltage)

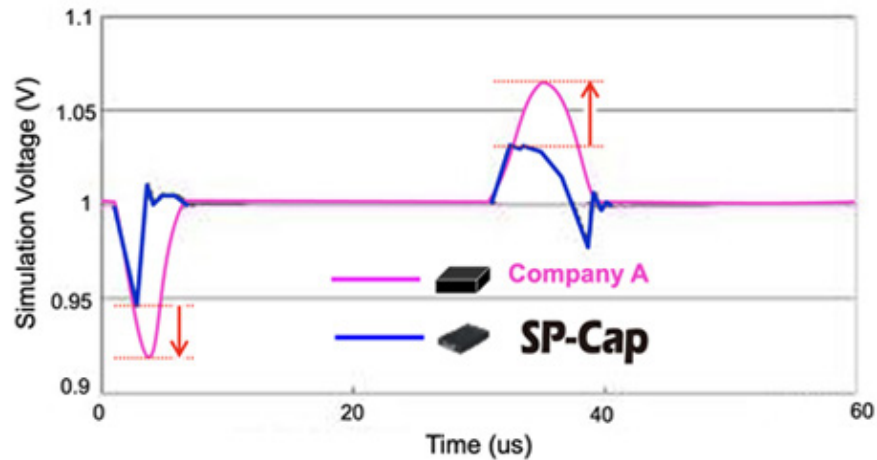
Comparison of characteristics of 2 V 470 μF 9 mΩ products



Even when the initial capacitance C and ESR are at equivalent levels, similar products deteriorate over long-term use.
 → SP-Cap demonstrates stable characteristics achieved by unique technology.

Differences in voltage suppression ability after testing.

Transient-load Simulation After the Durability Test



Performance lowers after long-term use due to capacitance and ESR degradation.
 → SP-Cap maintains good performance even after long-term use.

In conclusion, the evolution of hardware is essential for the rapid development and expansion of generative AI services. Among the key components, conductive polymer capacitors stand out for their ability to meet critical hardware requirements such as space efficiency and high reliability. These capacitors significantly enhance computational performance and play a crucial role in ensuring stable operation of evolving generative AI hardware, ultimately supporting the future of this transformative technology.

Contact us about SP-CAP™ Capacitors today!