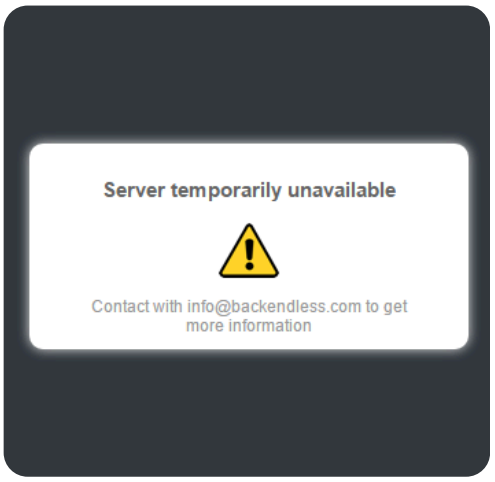
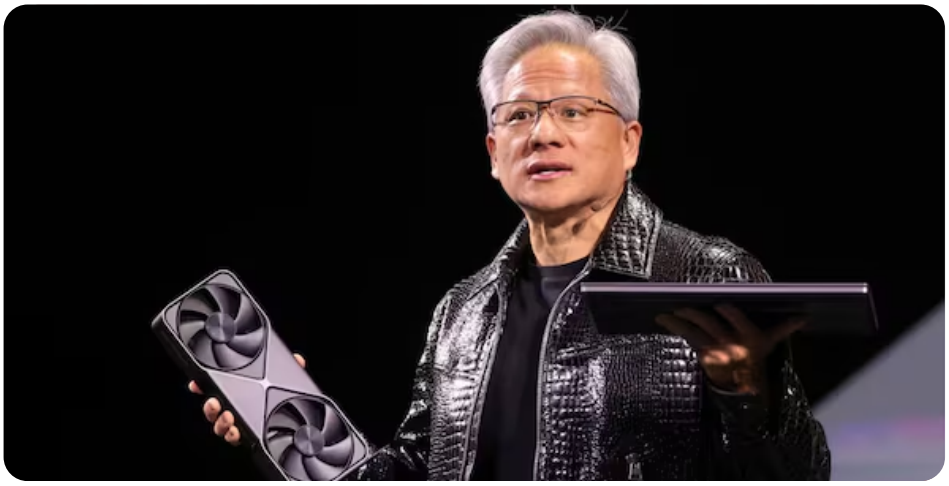
Department Of CSE

Byte Quest
Issue 164

28 Jan 2025





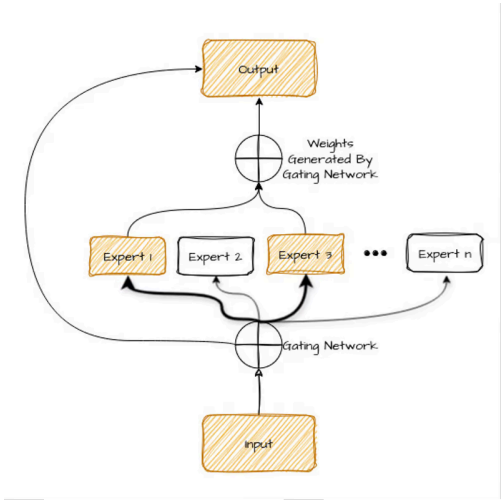
Department Vision

To be a center for academic excellence in the field of Computer Science and Engineering education to enable graduates to be ethical and competent professionals.

Department Mission

To enable students to develop logic and problem solving approach that will help build their careers in the innovative field of computing and provide creative solutions for the benefit of society.

Faculty Coordinators: Dr. Bhargavi Peddireddy(Assoc. Prof.) S. Komal Kaur(Asst. Prof.)
Student Coordinators: Sarath Kumar NVS 1602-22-733-061 Ryan Gabriel 1602-22-733-039

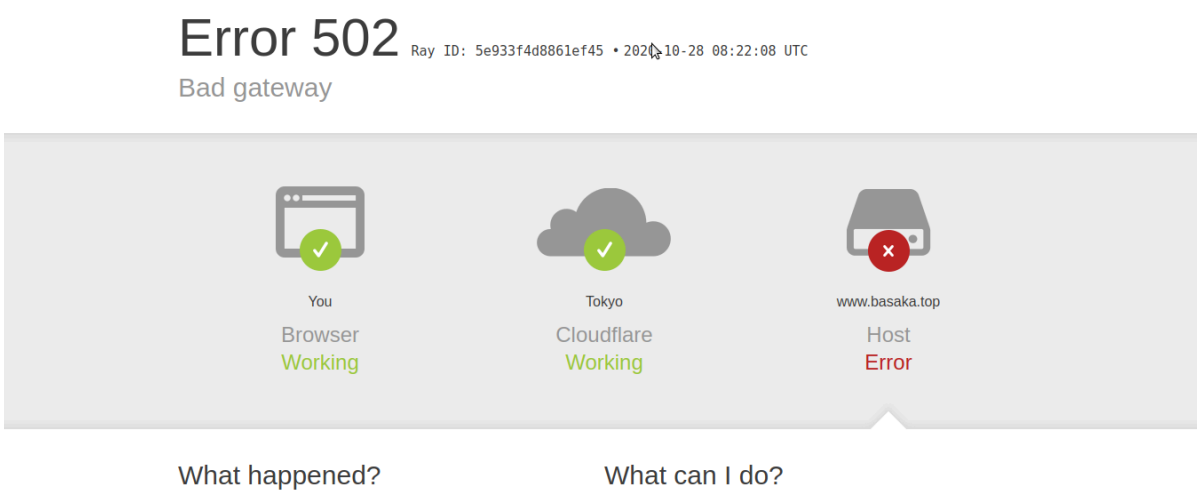


On **Jan 20th 2025**, a small startup, **Deepseek** released a model called **Deepseek R1**. To everyone's surprise it was capable enough to compete and in some areas **defeat** the latest version of **OpenAI's model O1**. The icing on the cake was that it was able to defeat O1 while also being **cheaper to build and run**. It is soo cheap to run that you could **run it on your PC** provided it has a **GPU** that is not too old. DeepSeek is able to give such performance at such low costs because it uses and approach known as **Mixture of Experts (MoE).**

The mixture of experts (MoE) technique addresses this challenge of **demanding high computational resources** by **breaking down large models into smaller**, specialized networks. It can be split into 4 components. The **input** is the problem or the data that you want the AI to handle, these inputs are passed on to **specific experts**, which are **smaller AI models**, each trained to be really good at a specific part of the problem. The **Gating Network** decides which expert is suited for which task based on the input and finally we get the **output** once the experts are done with their work. This approach is **efficient** and **flexible** as **only** experts that are good at a specific scenario are considered, which saves **time and computational power** and we can also add more experts or change their specialities based on our use cases . Since each expert focuses on what they're good at, the overall solution is usually more **accurate and reliable**. Despite challenges in training complexity and model size, MoE offers a promising method for advancing AI capabilities like with the case of DeepSeek.

It was once said that an entrepreneur can never fill a stadium, but then during **(Consumer Electronics Show) CES 2025**, **Jensen Huang**, the **CEO of nVidia** was welcomed by a stadium full of fans cheering at the mere sight of the entrepreneur. What was about to be unveiled that made the audience loose their mind? It was the launch of **5000 RTX series GPUS** , the new regular consumer market lineup of GPUs .The Nvidia RTX 5000 series GPUs are the newest high-performance graphics cards to supercharge your creative projects. Rendering 3D models, editing high-resolution videos, or creating detailed digital art? These cards are designed to make your workflow faster, smoother, and more enjoyable. The RTX 5000 series is packed with advanced features that set it apart like the **Blackwell architecture,** a successor to Hopper and Ada Lovelace microarchitecture .Blackwell is fabricated on the **custom 4NP node from TSMC**. 4NP is an enhancement of the **4N node** used for the Hopper and Ada Lovelace architectures. The Nvidia-specific 4NP process likely adds **metal layers** to the standard TSMC N4P technology. The **GB100** die contains **104 billion transistors**, a **30%** increase over the 80 billion transistors in the previous generation Hopper GH100 die. As Blackwell cannot reap the benefits that come with a major process node advancement, it can achieve power efficiency and performance gains through underlying architectural changes. The 5000 series also improved upon Ray tracing which creates realistic lighting, reflections, and shadows, making your projects look stunningly lifelike and with the improved ray tracing capabilities, rendering and previews in software like Blender and Unreal Engine are faster and more visually accurate and with other AI innovations like **DLSS 4 (Deep Learning Super Sampling)** use AI to boost viewport framerates in creative apps such as Blender and Unreal. While the GPUs themselves are pretty powerful in the markets they target, many have pointed hours after the event that the gains that nVidia has promised are not huge like the previous generations of the GPUs. Is **hardware tech** about to touch the peak or are there any twists like the introduction of **M series** chips from **Apple** that revolutionized **power efficiency and raw power** in large handhelds and laptops

When **Apple Vision Pro AVP,** the AR VR (XR) headset that in theory could make interacting with tech more immersive than ever before was released to the public on **Feb 2nd 2024**, tech enthusiasts were shocked by the technical marvel that they had just witnessed. Weeks later people started to discredit the AVP, due to lack of features that made it useful in real world, which resulted it into being touted as a glorified Ski Mask by the general public. There were rumors floating around that **Samsung**, **Google** and **Qualcomm** would also team up to make a AR VR headset like the vision pro. On **January 29th 2025** these rumours turned into reality when a popular tech youtuber, **Marques Brownlee** made a video titled **" I Tried Samsung's Secret Android XR Headset!"**. Sites like Twitter and Reddit had lots of online discourse on the kinds of features that the XR Headset could do, while also speculating its price. In the tech demo, Marquess was able to use the XR headset to do regular tasks like **browsing the web, watch videos** and not much as it was sneak peak and then revealed the bombshell that **all PlayStore apps** could work on the headset which was a stark contrast to the Vision Pro which could only run **AVP made apps**. Internally it seems to work very similar to how **picture in picture mode in Android** or a **non-fullscreen mode on Windows**, i.e forcing a reactive UI for apps that don't have native XR support like Chess.com . The XR Headset also had **gemini** which seem to do wonders like **talk** with the users very **naturally** (thanks to the **multi-modal** nature of Gemini), **understand and guide** through the home **UI** and within XR supported Apps like google maps. You could ask gemini to navigate to a location that you are watching on the TV or use Android's latest feature **Circle To Search ,** which allow users to search anything just by drawing a circle around what you want to search with your hands in the air like a wizard. While the product is still under development, the video has excited a lot of tech enthusiasts including myself to try it try it out and tell whether its just hype or the next iPhone.



On **Feb 7th 2025** when gamers across the world went to homes to relax for the weekend and hop onto their game on their PlayStations ,the **PlayStation Network (PSN)** suffered a 20-hour outage , impacting online services and even digital game access for PS4 and PS5 consoles. **Sony** attributed the disruption to an "O**perational Issue**," explicitly denying a cyberattack. While the specific technical cause remains undisclosed, the incident highlights the complexities of maintaining large-scale online gaming infrastructures. Sony offered a 5-day PlayStation Plus extension as compensation. The outage, though not a hack, echoes past PSN security concerns, notably the **2011 breach**. This recent event, regardless of its cause, highlights the **critical need** for **robust network resilience** and incident response in online platforms. For computer science students, this event serves as a real-world example of the challenges in ensuring **high availability and reliability in distributed systems**, and the importance of clear communication during service disruptions. The PSN outage becomes a compelling narrative – a real-world mystery revealing the hidden vulnerabilities within even the most sophisticated distributed systems like the **CloudFare disaster during Sept 17th 2024**. From whispers of authentication failures to database anxieties and network congestion nightmares, the incident throws into sharp relief the critical need for robust defenses and masterful incident response in our hyper-connected world. While services were eventually restored and a five-day extension offered as recompense, the PSN story serves as a potent reminder: in the age of always-on services, reliability isn't just a feature, it's the very foundation of trust, a lesson etched in digital silence for the next generation of tech architects.