



## GPU's and its spread across Industries



**Banking** - 2008 financial crisis triggered enormous amount of procedural regulations and local/cross-border compliances in Banking domain. Resulting into tremendous demand on IT and HPC (High Performance Computing) resources in order to run some of the voluminous and complex operations on day to day basis - such as Monte Carlo Simulations (generative models), run by large investment banks & hedge funds to keep risk and uncertainty of the financial instruments under control; portfolio restructuring for assets management; re-design & enforcement of new trading strategies with real-time market based prediction; fraud analysis & detection etc. to name a few.

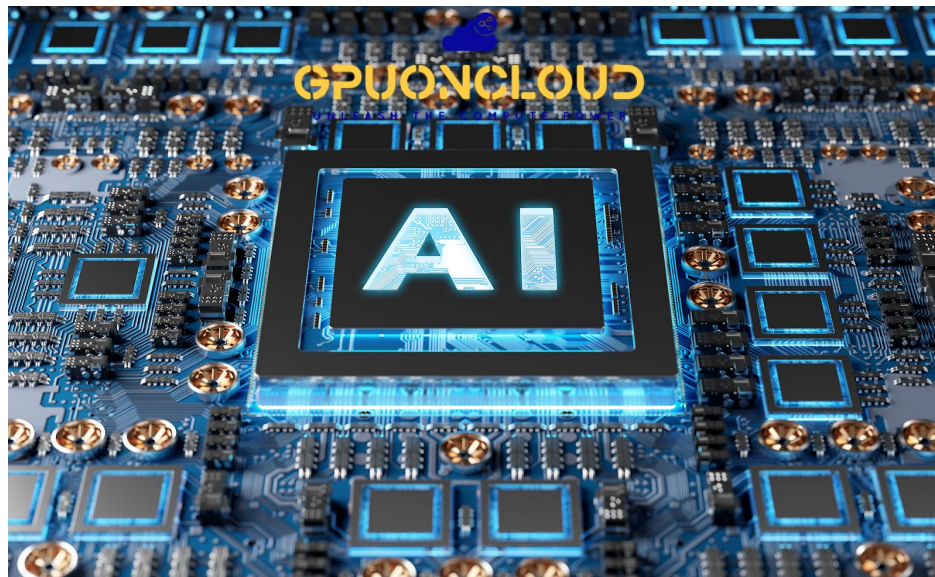
**Retail** - Industry has evolved from providing traditional retail experiences e.g. going to the brick & mortar store, looking for a product of your choice and then making a purchase, to offering a buyer personalized choice of products, basis consumer preferences, clubbed with associated product offers of the top rated products at an affordable price and discounts basis purchase history. E.g. customers getting personalized skin care treatment without having to see the dermatologist - possible with HPC enabled AI applications which takes a selfie of their plain face and then providing recommendations dealing with specific skin problems and a personalized skin care regimen recommendations.

**Healthcare & Life Sciences** - Advancements in healthcare domain driving the need for HPC powered with AI (Artificial Intelligence) based solutions. Identifying and decrypting simple to complex diseases from a wide variety of medical diagnostic samples collected in the forms of ECG, MRI, X-RAY, real-time scan & images, blood & other samples etc. calls for detailed and voluminous dataset to be collected, curated for processing, labelled, analyzed with set of various related & unrelated permutations and combinations in order to identify the diseases, proactively take due course of medicines, actions and

eventually prevent it from happening. E.g cancer classification & detection with right set of treatments across various stages, along-with corrective and preventive therapies. Needless to state the clinical trials & research programs are increasingly demanding the GPUs (Graphics Processing Units) to develop the drugs in fast pace e.g. finding medications that reduces Ebola infectivity in a day compared to many months time in past.

**Manufacturing & Supply Chain** - Having a rightly priced product, meeting desired quality & service is the key in manufacturing. Manufacturing Industry has started leveraging HPC & AI models to improve supply chain efficiencies, reduce inventory costs, predict assembly lines performance and machines breakdown cycles etc. e.g. analyzing variety of data points including material characteristics, various possible configurations, designs, and predictive test results leads to faster go to market, improved quality and reduced cost for a manufacturer.

**Travel & Hospitality** - The most obvious way of deploying the HPC & AI models is for a customer service desk. In particular technology is doing wonders when it comes to sending direct messages. Providing online chat services and responding to simple requests & questions. e.g. AI chatbots made available at social media platforms being available 24x7 to provide instantaneous responses to customer queries - greatly appreciated by hospitality industry.



Prior to the GPU evolution, various industries leveraged traditional multi-core CPUs, thereby increasing the datacenter footprint, investments in enterprise class HPC systems with INTEL (Xeon), IBM (POWERx), AMD (EPYC™ - zen microarchitecture) CPU's, along with associated expertise. However with GPU's evolution its most common to see GPU's such as kepler's K40's, K80's, followed by Pascal, Tesla & Radeon Instinct™ GPU's marking its presence to the tune of 100+ numbers in banks datacenter, leading a way to quantum computing - if/when it commoditizes. And it's not a surprise to see 1000+ GPUs for giant investment banking entity.

It's predominant to notice, AI programs running at slow pace, acute shortfall of right skilled talent to build and integrate enterprise class AI systems - making ROI justification questionable in few of the large giants & conglomerates. While developing separate clusters for AI Training and Inferencing is preferred to justify ROI, the need to re-train the AI models on day to day basis is equally essential. Other approach to justify the ROI, is to opt for cloud based HPC platforms offered by AWS, GCP, Azure and players including [GPUONCLOUD](#) - India based startup with innovative and differentiated offerings in Scalable Instant Jumpstart Cloud platform for AI/ML/DL powered by both AMD and NVIDIA GPU's, while taking care of enterprise class data Confidentiality, Integrity & Availability pertaining to specific industry.