

It looks like AI (artificial intelligence) and its associated opportunities are real. Where is AI leading us to? We are very concerned that robots can create new robots emulating biological reproduction without being biological and yet physical. AI has already gone beyond virtual space into physical space such as autonomous driving on human roads. Technology is neutral. It is up to us to make good or bad uses of technology. The key message here is that business opportunities are huge and widely open to small and medium-sized companies as well as globally large companies. The current moment in 2024-07 is akin to the years preceding 2000 when the demand for business PC reached a pinnacle.

Nvidia CEO did mention in a presentation prior to Computex 2024 in Taipei that Nvidia was going to enable robots to learn the laws of Physics. Nvidia has a good track record of involving in physics for its GPU card design and development based on its powerful software ecosystem (CUDA). How do we capitalize on this trend by analyzing information disclosed by technology leaders such as Nvidia, AMD, Intel, ARM, Microsoft, Google and so on.

Observation 1 is that AI developments rely on PC technology and will continue to develop in parallel with PC technology. PC technology is largely dependent on semiconductors which are now hitting Moore's Wall which reflects the limit of the laws of Physics. Quantum computing will unlikely be available within the next 10 years. We shall look at how PC technology makes advances over the next 10 years.





Observation 2 is the emergence of new computing protocols in the near future (such as CXL) to allow plenty of computer memories to be accessible to computer processors. This may allow a processor with 16

cores to access 16 memories without sharing a pool via a bottleneck. This is not hard to achieve.

Observation 3 is the immediate emergence of AI-PC with computer processors that can execute AI models in 8 bits of integer instead of 32 bits of floating point. Computing resources go up 4 folds immediately. Time to compute in integer is much shorter than floating point. This is a big boost to AI model performance and power efficiency, and a realistic way of emulating the effects of the brain.

The three observations are reasonable and realistic implying that PC technology is going to support further development and a full-scale deployment of AI or machine learning applications.

The PC industry has visioned doubling of datacenter computing power consumption in 5 years and 10-fold increase at the edge (outside of data centers) over the same period up to 2028. Nvidia is a vision leader of AI-PC and has become the most valuable company in the world of all industries recently. The meteoric rise in value of Nvidia is not an illusion despite stock investors may have illusory perceptions. The edge is in a very early stage of AI at present. It is where all business opportunities for most small and medium sized companies are.

This article was largely based on observations of news coming out of the globally leading companies around the Taipei Computex 2024 period. The same theme underlined the Compucon CPD seminar program in June. For more similar articles, please visit https://www.cnz.co.nz.

