

A shift in the meaning of energy storage

December 25, 2024 Zhan Jia zhang m10707102@gapps.ntust.edu.tw

Conclusion:

Originally, energy storage systems were mainly used to cooperate with renewable energy, but due to the advent of the AI industrial revolution, a large amount of energy is needed, and reducing carbon emissions is no longer the most important, and the role of energy storage will be transformed into assisting various power plants to maximize power generation

Tesla unveiled their Master Plan 3 at the 2023 Investor Day, which mentioned that in order for the world to move towards a decarbonized economy, the world needs about 240TWh of battery energy storage, mainly used in transportation, energy, heating and other fields. But during Tesla's earnings conference in the second quarter of 2024, Musk said "*With the addition of a battery pack, you can now run your power plant in steady state. Steady-state means that basically any given power grid, anywhere in the world, can produce cumulative energy over the course of a year, at least twice as much as it currently does, and in some cases, potentially triple*"^[1] At the same time, Google and Microsoft have also revised their carbon emission standards, or announced that they will abandon carbon neutrality^[2] Carbon reduction goals are no longer the focus of energy storage development, the most important thing is to maximize the production of energy to develop AI, to meet the new round of industrial revolution, auxiliary renewable energy allows the grid to incorporate more renewable energy, auxiliary traditional generators, and increase the time for generators to be fully loaded.

cause

AGI

Huawei Founder Ren Zhengfei: We are about to enter the fourth industrial revolution, the foundation is large computing power, the fourth industrial revolution is magnificent, and its scale is unimaginable^[3]

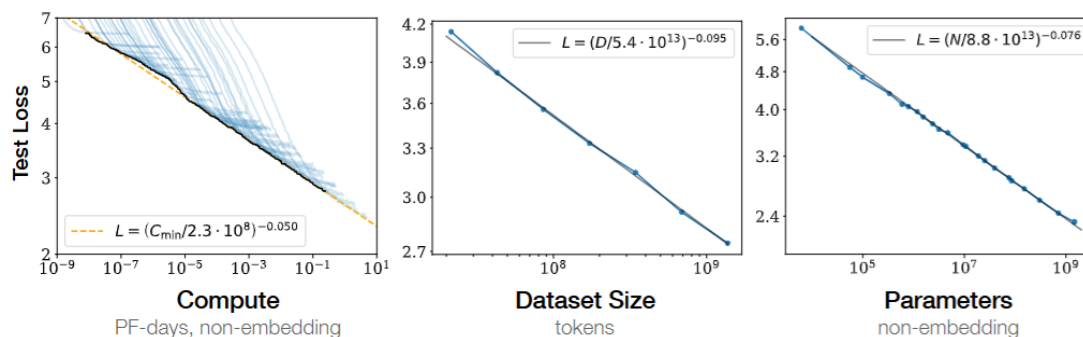
Definition of AGI (Artificial General Intelligence) There is no clear definition of AGI, but everyone's understanding is different. For Sam Altman, AGI is defined as: "An artificial intelligence equivalent to the average human being, like a remote colleague who can be hired." It can do anything you want a remote worker to do from a computer, including learning how to be a doctor, learning how to be a really good programmer."^[4] Demis Hassabis, CEO of Google DeepMind, defines AGI as "a general-purpose system capable of accomplishing any cognitive task that a human can do."https://www.youtube.com/watch?v=vd9GxG5Qn-k&ab_channel=Greylock

[5]

Once humans make AGI, AGI is expected to replace a large part of white-collar jobs, and many things can be handed over to AI to do, and the economic output value of society will soar rapidly. Nobel laureate Geoffrey Hinton said that almost all AI experts agree that humans can make AGI [6] Demis Hassabis agrees [7]. However, it is difficult to predict when this will happen. The tech giants' forecasts range roughly from 2025 to 2030 [8]. Next, we will describe why humans need a lot of computing power and power to build AGI.

The Law of Scaling

"""



[9]

Ilya Sutskever, former Chief Scientist at OpenAI:

"If you have a very large neural network and use a lot of datasets and computing resources to train it, then you can achieve amazing results reliably and predictably. And this simple idea is the core foundation of almost all AI advances since the deep learning revolution began. Big neural networks, big data, and a large amount of computing power are the basis for almost all artificial intelligence advances in the past decade, which is basically the so-called law of scaling, what is the law of scaling?"

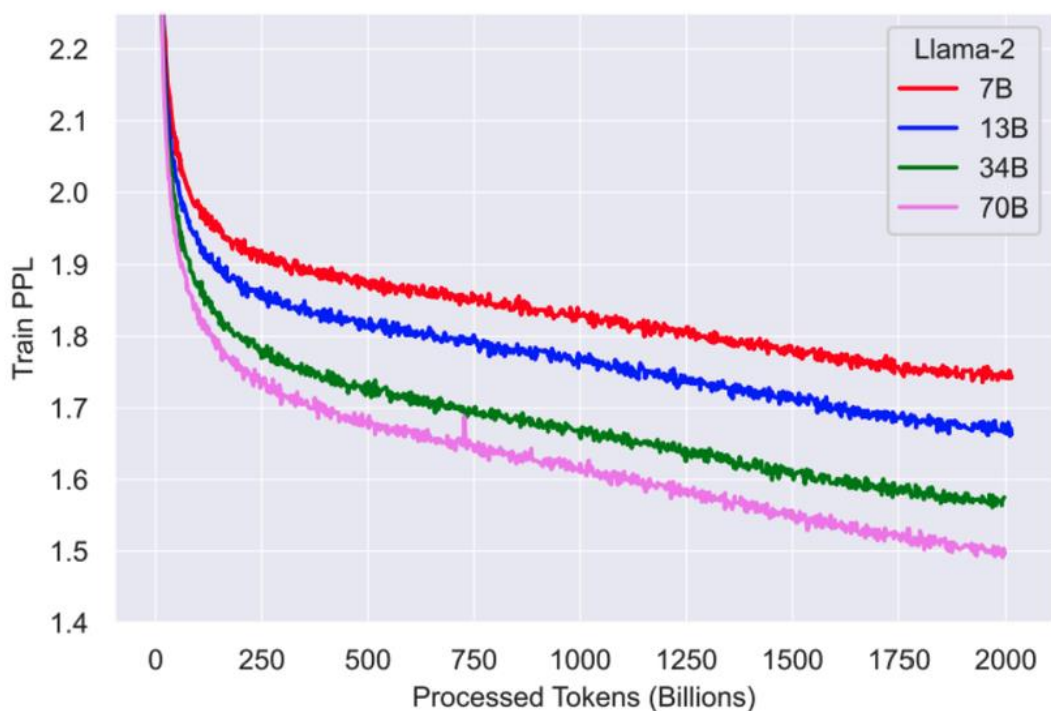
[10]

The above figure is mainly about the core results of the scaling law of the AI model, the y-axis is the loss of the model, and the smaller it is, the more accurate and intelligent the model predicts. The graph on the far right shows the expansion of the number of parameters, and the x-axis is the logarithm of the computing power. The researchers are using a very large data set, so they don't have to worry about running out of data, and they train all the models for a long time to make sure that they basically reach the upper limit of learning performance (convergence state), and you can observe that there is a limit for each number of parameters, and then as the number of parameters increases, the limit value gets lower and lower, that is, as the number of parameters increases, The smart upper limit of the model will also be improved

in the middle of the graph to do a similar operation, but the amount of data and the number of parameters are interchanged, that is, a very large model is trained, so the size of the model is not a limiting factor for

performance, but with a dataset of different sizes to train, and stop in advance, measure the test loss at the point with the smallest test loss, and once again get a very clear power-law relationship between the test loss and the size of the data set, This means that each amount of data will have the limit of the best model that can be obtained for that amount of data, and then as the amount of data increases, a better model can be obtained. The

learning curve of different size models is drawn, and each model is given a very large amount of data and trained for a long time for all different model sizes, so each model can learn to the limit, on the picture, the end of the light blue thin line is horizontal, that is, with the increase of the amount of calculation, the model has reached the limit, and there is no way to continue to get better, it can be observed from the yellow dotted line, no matter how many parameters are given or how much data is given. There will be a limit of computing power that cannot be broken, and then as the computing power increases, this limit will become lower, which means that if you have more powerful computing power, you can get a smarter model



We don't know why AI models have a scaling law, it's a rule of thumb maybe it will continue, maybe it will stop, so let's look at the current situation. The most advanced open source model in June 2024 Meta's Llama 2 training data shows that even the model with the smallest parameter amount of 7B still loses plummeting throughout the training process, and there is no sign of slowing down [11], Mark Zuckerberg said in April 2024 [12], their latest model (Llama 3) Even with the smallest parameters, it is still learning and not yet saturated, that is to say, so far, if the model is given more data or more training, the model is still continuing to get smarter, so major technology companies are still continuing to expand the training scale and build larger computing centers

2GW+ computing power center

Because of the scaling law of artificial intelligence, we only need to invest more computing power, that is, energy, to get a smarter model. As a result, almost every big tech company has started to build several high-capacity AI computing centers. Oracle founder Larry Ellison in September 2024 Report to the meeting [12] Investors were told that they were already building an 800MW hash centre and were designing a 1GW hash centre. Ellison said there are currently several big tech companies, and possibly just one country, vying for "technological supremacy" for AI models in the coming years. To remain competitive in the race for cutting-edge AI models, the cost won't be cheap, he added.

"The entry price is about \$100 billion. Let me repeat, about \$100 billion. This is the price that anyone who wants to participate in it will have to pay in the next five years," he said [13]

September 2024, Elon Musk's artificial intelligence company xAI has deployed about 100,000 NVIDIA H100 GPUs, about 150MW, at its hash center in Memphis and it took only 122 days [14] , Mark Zuckerberg said in the October earnings conference that they are now training models with more than 100,000 H100s and plan to deploy 350,000 Nvidia H100 GPUs by the end of the year [15] , according to the proportion, it is estimated that the computing center of Facebook at the end of the year will be 525MW, Microsoft and open AI, the companies that currently buy the most GPUs, will not have smaller computing centers than other companies, and they are expected to restart the Three Mile Island nuclear power plant in Pennsylvania, where there was a nuclear accident, CEO of Alphabet in September Sundar Pichai said, Google is expanding its computing infrastructure and working on 1GW+ of the hash center [17] Personally, I think there will be several 1GW computing centers built within two years.

In other words, the law of scaling continues to evolve and has not yet reached its limits, so companies are still expanding their computing centers. At present, almost every company has announced that it is designing and building a 1GW computing power center, which is reasonably estimated to be in 2025 There will be the first 1GW Hash Center. according to **Zuckerberg's attitude, we will not reach the limit of scaling in the short term, Personally, I think it is likely that we will see 10GW next of the hash center** . Zuckerberg, December 2024 Announced the construction of a 2GW computing center and the construction of a new 1.5GW natural gas power plant to power the computing center [17]

And that's just the power needed to train the model, and in the real world, the power needed for inference. The training process is to allow the model to learn and acquire wisdom, while the inference process is the power consumption of the model to provide services, such as when we use ChatGPT on a daily basis. There is still a lack of publicly available information to make it difficult to estimate these needs concretely , But Amazon Claude ai, founder of Dario Amodei Say most of the cost in inference, not just the cost of training the model [18], it is possible that the power consumption of inference will be more than that of training.

Lack of electricity

Since both model training and inference generate a lot of power demand, Zuckerberg mentioned in an interview in April 2024 [19] He believes that the biggest bottleneck in the development of artificial intelligence is not chips, but energy. Musk also said that he thinks there will be power shortages in the future [20] °

Based on the above information, it can be predicted that one of the bottlenecks for the development of all the top technology companies in the United States in the future will be energy production. Therefore, it can be speculated that in the future, the world's most abundant capital and top engineers will invest a lot in electricity production.

The current trend that can be observed is that many companies have already started to buy and build nuclear power generation themselves. In addition, Japan's Mitsubishi Electric, the world's top manufacturer of gas generators, also reported a 50% increase in orders in 2024 compared to the previous year due to the development of artificial intelligence. [21] Siemens Energy, which also makes gas generators and transmission transformers, has risen more than Nvidia [22]

Power System Challenges Raised by Training Models:

Musk said that for a hash center with a capacity of 50MW, power demand fluctuations can be reduced from 50MW to 20MW and back to 50MW in 100 milliseconds [23] As far as I know, there is no other way than an energy storage system to compensate for such rapid power fluctuations. According to reports [24] Tesla's energy storage system is also ready to be put into the power supply equipment of the computing center, which is already under construction.

“””

“””

Because Taiwan will need a lot of electricity, and then batteries can be used to help

With the development of artificial intelligence (AI) technology, future large-scale applications will significantly increase the global demand for electricity. As this cycle continues to expand, the demand for various forms of power generation resources will also increase, including gas-fired power generation, nuclear power, renewable energy, and eventually even coal-fired power generation.

In this context, battery energy storage has become an important solution to cope with peak power demand. Generators are often not in continuous operation, but are started or shut down based on load demand, limited by capacity factors. Therefore, the use of batteries to support the demand during peak power hours is an effective solution. Relevant literature points out that batteries can be stored during non-

peak hours and discharged during peak load peaks to achieve the effect of load regulation, effectively reduce the load of generators and improve grid stability.

Taiwan's power grid has been using pumped storage to reduce night spikes, with about 10GWh of pumped storage in Taiwan, according to the power trading platform^[25] As of December 2024, at least E- batteries have been procured^{reddReg} There are 3GW, while E-dRegis at least 2.5 times the power, and it is reasonable to assume that within two years, Taiwan will add more than 7.5GWh of battery storage, which will become increasingly important.

“””

How the U.S. assesses peak power capacity

What is ELCC? How is it assessed?

In the United States, when assessing whether a power system can operate stably during peak loads, it is common to use a method called payload carrying capacity (Effective Load Carrying Capability, ELCC)

^[26] standard method. This approach focuses on the reliability of the power system and analyzes the extent to which new resources can meet the increased demand by gradually adding power generation equipment.

Specifically, ELCC's calculations start with a basic power system model. First, the generation capacity in the system is gradually increased, while the load demand is increased simultaneously, until the new generation resources are added to restore the overall reliability indicators of the system (e.g., Loss of Load Expectation, or LOLE) to the same level as the original system. The purpose of this process is to quantify the actual contribution of the new generation resources to ensuring the reliability of the system.

When performing an ELCC analysis, it is critical to clarify the composition of the power system. This approach requires a detailed assessment of the capacity of each generator and its probability of failure, as well as the addition or retirement of power equipment. For example, the introduction of new equipment may increase the overall capacity of the system, while the retirement of old equipment may negatively impact system reliability, and the core strength of the ELCC approach is that it supports power system planning. By calculating the actual contribution of each generation resource, decision-makers can plan the allocation and use of resources more scientifically, ensuring that the system is resilient for future development while meeting current load demand. Compared to traditional capacity-based approaches, ELCC more accurately reflects the true value of resources during peak load demand periods. As a result, it is widely used in resource assessment for the U.S. power industry

Evaluate the ability of energy storage to be added to the power system

Since photovoltaic systems and wind power generation have developed much earlier than energy storage systems, there has been a lot of literature discussing the introduction of renewable energy ELCC's approach [27][28][29] For energy storage to join ELCC, I would like to quote Elon Musk in the 2024 Tesla Q2 financial report [30]

.....

I think people don't understand just how much demand there will be for storage..... I think, are underestimating this demand by product order of magnitude.

grid is, if the power plants can operate at steady state is at least two to three times the amount of energy it currently produces. Because there are huge gaps -- because there's a huge difference in the -- from peak to trough in terms of energy or power generation. So in order for a grid to not have blackouts, it must be able to support the load at the worst minute of the worst day of the year. The coldest or hottest day, which means that for the rest of the time -- the rest of the year, it's got massive excess power generation capability, but it has no way to store that energy.

Once you add battery packs, you can now run the power plants at steady state. Steady state means that basically any given grid anywhere in the world can produce in terms of cumulative energy in the course of the year, at least twice what it is currently producing, in some cases, maybe three times.

.....

A shift in the meaning of energy storage

I think this is a very important moment, symbolizing a shift in the meaning of the existence of energy storage systems The original power system has very little growth in electricity consumption, and the meaning of energy storage is to be able to buffer the unstable energy generated by renewable energy to match the real electricity consumption and reduce carbon emissions, but with the advent of ChatGPT, everything has changed, the demand for electricity has increased wildly, and energy storage has become used to maximize power generation to provide enough energy to AI

On the other hand, Tesla is thinking about evaluating this matter, which means that this thing is needed and valuable In the short term, it may be to increase peak power supply capacity and increase reserve capacity, and in the long term, grid energy storage is to maximize power generation to AI supply of energy, therefore, currently, the New York Independent System Operator (NYISO) 、 PJM Interconnection Texas Public Utilities Commission (PUCT) Many regions are actively studying the capacity and continuous discharge time requirements of energy storage systems [31, discussed the U.S. Federal Energy Regulatory Commission (FERC) Published in the first place 841 The impact of the order on the participation of energy storage systems in the wholesale market. The order requires regional transmission organizations (RTO) and Independent System Operators (ISOs) (ISO) to establish rules that allow energy storage resources to

participate in wholesale energy, capacity, and ancillary services in a non-discriminatory manner.

[32] It is pointed out that the PJM power market in the United States uses the ELCC model to scientifically evaluate the capacity value of new energy storage. [33] In this paper, a method considering ELCC and levelized LCOE (LCOE) The multi-objective capacity allocation optimization model of the combined wind power energy storage system aims to improve the comprehensive benefits of wind power and energy storage systems. [34] A simplified deterministic method is proposed to quantify the contribution of energy storage to capacity value in the case of increasing the penetration of a high proportion of renewable energy. The study highlights that accurately measuring the contribution of energy storage systems to grid reliability is critical to maintaining grid stability amid increasing renewable energy penetration.

Resources

[1] https://www.tesla.com/ns_videos/Tesla-Master-Plan-Part-3.pdf ,

Once you add battery packs, you can now run the power plants at steady state. Steady state means that basically any given grid anywhere in the world can produce in terms of cumulative energy in the course of the year, at least twice what it is currently producing, in some cases, maybe three times., 2024 Tesla Q2 earnings transcript, <https://www.fool.com/earnings/call-transcripts/2024/07/24/tesla-tsla-q2-2024-earnings-call-transcript>

[2] AI consumes high energy, Google, Microsoft's carbon emissions are rising instead of decreasing... Google announced that it would give up carbon neutrality, and Microsoft bought carbon capture carbon rights for self-insurance.

<https://esg.business today.com.tw/article/category/180687/post/202407110035>

[3] [Ren Zhengfei: We are about to enter the fourth industrial revolution, <https://www.ithome.com/0/720/163.htm>].

[4] [How Sam Altman Defines AGI ,

https://www.youtube.com/watch?v=vd9GxG5Qn-k&ab_channel=Greylock

[5] [Hern, Alex. "AI Has the Potential to Cure All Diseases, Says DeepMind Chief." *The Times,* December 20, 2024. Accessed December 20, 2024. https://www.thetimes.com/business-money/technology/article/ai-has-the-potential-to-cure-all-diseases-says-deepmind-chief-9hbdp5kpm?utm_source=chatgpt.com®ion=global.]

[6] [Geoffrey Hinton: Almost everybody I know who's an expert on AI believes that they will exceed human intelligence , <https://x.com/BBCNewsnight/status/1791587541721780400>]

[7] [Demis Hassabis - Scaling, Superhuman AIs, AlphaZero atop LLMs, AlphaFold , <https://youtu.be/qTogNUV3CAI?t=306>]

[8] Bigwigs from all walks of life have set a "deadline" for AGI, and the 27-year-old founder has prepared the AI for the "final human test"! <https://mp.weixin.qq.com/s/V9oNGADteTE4WzUju-zQ2Q>

[9] Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). *Scaling Laws for Neural Language Models*. arXiv preprint arXiv:2001.08361

[10] Ilya Sutskever's Law of Scaling <https://youtu.be/BjyZcSiVg5A?t=206>

[11] Mark Zuckerberg says that the model is still learning,
<https://youtu.be/9TU0XjJqpOg?t=17>

[12] [Oracle will build three small nuclear power plants, pushing the Blackwell Super Cluster TechNews](#)

[13] Larry Ellison, \$100 billion. This is the price that anyone who wants to get involved in the next five years will have to pay https://x.com/The_AI_Investor/status/1834773985990451509

[14] [Reduced from 4 years to 4 months! Musk's supercomputing construction speed frightens opponents Tencent News](#)

[15] Verbatim transcript of Meta's official earnings

conference https://s21.q4cdn.com/399680738/files/doc_financials/2024/q3/META-Q3-2024-Earnings-Call-Transcript.pdf

[16] Google is also building a 1GW+ hashrate center, <https://x.com/tsarnick/status/1837362656908562566>

[17] [Meta announces 4 million sq ft, 2GW Louisiana data center campus - DCD](#)

<https://www.datacenterdynamics.com/en/news/meta-announces-4-million-sq-ft-louisiana-data-center-campus/>

[18] Dario Amodei, founder of Claude AI, <https://youtu.be/7xij6SoCCII?t=725>

[19] Mark Zuckerberg's biggest bottleneck in the development of artificial intelligence in April 2024 is not chips, but energy. <https://youtu.be/i-o5YbNfmh0?t=36>

[20] Musk also said that he thinks there will be power shortages in the future,
<https://www.gvm.com.tw/article/112164>

[21] The demand for gas-fired generators has increased dramatically

<https://www.bloomberg.com/news/articles/2024-10-23/mitsubishi-power-sees-data-centers-boosting-gas-turbine-orders>

[22] [Siemens Energy beats Nvidia this year.](https://www.ft.com/content/46996cf4-472b-4729-bcb5-b66e2bbdded1) There's more

<https://www.ft.com/content/46996cf4-472b-4729-bcb5-b66e2bbdded1>

[23] Musk said that the super-large power changes in the computing center are

<https://www.bilibili.com/video/BV1jz421i7Kj?t=1321.9>

[24] Tesla Megapack is ready to be put into the hash center due to the moment of inertia demand <https://youtu.be/Jf8EPSBZU7Y?t=813>

[25] [Electricity Trading Platform](https://etp.taipower.com.tw/web/qse_info/qse_list), https://etp.taipower.com.tw/web/qse_info/qse_list.

[26] Public Utility Commission of Texas. (2024, September 26). 1219NPRR-17 PUCT Report. Retrieved from <https://www.ercot.com/files/docs/2024/10/01/1219NPRR-17%20PUCT%20Report%20092624.docx>

[27] IEEE. (2020). A Tutorial on PV Power Plant Capacity Credit Assessment. IEEE Transactions on Power Systems. DOI: 10.1109/TPWRS.2020.3008552

[28] IEEE. (2022). Evaluation of Resource Adequacy under High Renewable Energy Growth in the U.S. Midwest. IEEE Transactions on Sustainable Energy. DOI: 10.1109/TSTE.2022.3004921

[29] IEEE. (2022). Cumulative Capacity Credit Estimation for Renewable Energy Projects in Oman. IEEE Access. DOI: 10.1109/ACCESS.2022.3150845

[30] Tesla's 2024 Q2 earnings report meeting |

· <https://www.bilibili.com/video/BV1fw4m1r7ZC?t=3234.7>

[31] P. Maloney, As Grid Operators File FERC Order 841 Plans, Storage Floodgates Open Slowly, Utility Dive, 2018.

[32] Xia Qing. (2024). The practice of using the ELCC model to evaluate the value of energy storage capacity in the US PJM power market. The 9th Energy Storage Western Forum, September 2024. Retrieved from <https://baijiahao.baidu.com/s?id=1811311028329729738>

[33] Song, Q., Wang, B., Wang, Z., & Wen, L. (2024). Multi-objective capacity configuration optimization of the combined wind-storage system considering ELCC and LCOE. Energy. <https://doi.org/10.1016/j.energy.2024.131558>

[34] Jirutitijaroen, P., & Wangdee, W. (2022). Energy Storage Valuation toward 100% Renewable Electricity from Reliability Perspective. *IEEE Transactions on Sustainable Energy*.

<https://ieeexplore.ieee.org/document/10594815>