

# Research Support for Key Technologies for Distributed Computing Grid Investment Analysis

Geli Zhang<sup>\*1</sup>, Tong Li, Xiaohui Wang, Wei Cheng, Heng Zhang, Yujie Wang  
Dept. Technical and Economic Center, State power economic research institute, Beijing, China

## ABSTRACT

With the high proportion of new source loads such as distributed power sources, electric vehicles, and energy storage, as well as the rapid development of new models and formats such as microgrids, integrated energy, and virtual power plants, the distribution network is gradually transforming from a power network that simply receives and distributes electricity to users to a power network that integrates source grid load storage interaction and flexibly couples with the higher-level power grid. In recent years, State Grid Economic and Technological Research Institute has focused on original technology research and development in the field of distribution networks, fully supporting the construction of original technology source areas. It has developed a distribution network simulation calculation and analysis platform with completely independent intellectual property rights, supporting the construction of new distribution systems in multiple regions. The development of new source loads may lead to local concentration and overall imbalance, resulting in significant differences in load rates between different regions and lines of the distribution network. This often leads to issues of insufficient or unusable equipment capacity, which is not conducive to improving the overall quality and efficiency of the distribution network. The State Grid Economic and Technological Research Institute, in collaboration with the United Nations Grid Tianjin Electric Power Company, has established a flexible interconnection technology system for distribution networks, developed the world's first set of corresponding flexible interconnection power electronic equipment, and implemented it in the Beichen National Industry City Integration Demonstration Zone in Tianjin. After the project is put into operation, the load balancing degree of the demonstration zone's power grid has reached over 99%, the power supply capacity of the power grid has increased by nearly 40%, and important industrial enterprises have achieved "zero power outages".

**Keywords:** distributed computing, Power grid, high-quality development

## 1. INTRODUCTION

Energy security and safety are closely related to the national economy and people's livelihood, and cannot be ignored as the 'great importance of the country'. State Grid Corporation of China firmly shouldered the responsibility and mission of safeguarding national energy security, strengthened planning and overall guidance, fully served the high-quality development of new energy, accelerated the construction of ultra-high voltage and main grid, strengthened the construction of distribution network, improved the disaster prevention and disaster resistance capacity of power grid and the digital and intelligent level of power grid, and worked hard to write a new chapter of high-quality development of power grid to better support and serve Chinese path to modernization.

## 2. DISTRIBUTED COMPUTING FOR GRID HIGH-QUALITY DEVELOPMENT

### 2.1 Distributed computing technology

Distributed training of data parallel large models is a widely used technique in the fields of machine learning and deep learning. With the continuous growth of data volume and the increase of model complexity, traditional single machine model training can no longer meet the demand for efficient and fast training. Therefore, adopting distributed training to accelerate the model training process has become a trend. Data parallelism is a strategy that divides a large dataset into smaller sub datasets and trains them simultaneously on multiple computing nodes. Each computing node uses the same model parameters, but operates on different subsets of data. By calculating gradients separately at each node and

---

<sup>1</sup> \*litong@chinasperi.sgcc.com.cn

aggregating them, the global model parameters can be quickly updated. This data parallel approach can significantly improve training speed and solve the bottleneck problem during single machine training.

Distributed training of large-scale models refers to the method of training large-scale models on the basis of data parallelism. Large scale models have a huge network structure and parameter count, therefore requiring larger computing resources and storage space for training. In order to achieve distributed training of large models, it is usually necessary to use distributed storage systems and computing frameworks to support parallel processing of data and the training process of models.

The principle of data parallelism is based on the concept of parallel computing, where multiple computing nodes perform computing tasks simultaneously. Each computing node has its own model copy and locally calculates the gradient of the data batch it is responsible for. After the calculation of each small batch is completed, gradients are aggregated between nodes, and the model parameters are updated by taking the average or weighted average. This iterative process alternates between nodes until the convergence condition for training is reached[1-3].

The development and evolution of new energy systems and new power systems exhibit prominent characteristics of long time span and strong uncertainty. High quality coordination of energy and power security supply and clean and low-carbon transformation is a complex systematic project, and multi-objective coordination during the transition period is crucial. Among them, the new power grid is the core platform for building new power systems and new energy systems. It focuses on the construction of backbone network structures, coordination of power grids at all levels, information technology driven development, and empowerment of digital intelligence. Continuously consolidating its four foundations is the key to building a strong Strengthen the construction of ultra-high voltage and ultra high voltage backbone network, guided by the concept of big energy, based on the energy demand and resource endowment characteristics of high-quality economic and social development, continuously improve the backbone network that adapts to the large-scale optimization and flexible scheduling requirements of multiple energy resources, and consolidate the physical foundation of digital power grid construction.gital power grid[4,5].

Continuously improving the backbone network structure and fully leveraging the energy allocation platform value of the ultra-high voltage power grid. Implement the requirements and deployment of the national plan for the development of large-scale wind and photovoltaic bases with a focus on desert, Gobi, and desert areas, adhere to the "trinity" of large-scale wind and photovoltaic bases, advanced coal-fired power, and ultra-high voltage channels, and make every effort to do a good job in the research of base transmission and power grid development pattern. Extend the main grid structure of North China and Northwest China to desert bases, improve the main grid structure of Northwest and Northeast China, accelerate the construction of the Sichuan Chongqing ultra-high voltage AC main grid, and support the safe and efficient operation of cross regional DC as depicted in Formula 1-2.

$$C_{\text{total}} = C_{\text{inv}} + \sum_{n=1}^N \frac{C_{\text{OM}} + C_{\text{R}} + C_{\text{C}} + C_{\text{Rec}}}{(1+r)^n} \quad (1)$$

$$B = A \circ R = [a_1, a_2, \dots, a_n] \circ \begin{bmatrix} r_{11}, r_{12}, \dots, r_{1m} \\ r_{21}, r_{22}, \dots, r_{2m} \\ \vdots \quad \dots \quad \vdots \\ r_{n1}, r_{n2}, \dots, r_{nm} \end{bmatrix} = [b_1, b_2, \dots, b_n] \quad (2)$$

## 2.2 Distributed computing technology for power grid

Before exploring the coupling, coordination, measurement, and evaluation of the digital economy and high-quality development, we first need to have a deep understanding of the theoretical foundations of both. The digital economy, as a new economic form in the new era, mainly relies on the development and utilization of data resources, as well as the deep integration and innovative application of information and communication technology. Its core lies in achieving digital, networked, and intelligent development of the economy through advanced technologies such as big data, cloud computing, and the Internet of Things. The digital economy has not only changed the production mode, organizational form, and business model The main elements that constitute a comprehensive evaluation are:

1. Evaluators. The evaluator can be an individual or a group. The setting of evaluation objectives, establishment of evaluation indicators, selection of evaluation models, and determination of weight coefficients are all related to the evaluator. Therefore, the role of evaluators in the evaluation process cannot be underestimated<sup>[6-9]</sup>.

2. The evaluated object. With the development and practical activities of comprehensive evaluation technology theory, the field of evaluation has also expanded from the initial comprehensive evaluation of economic statistics in various industries to later aspects such as technological level, quality of life, well-off level, social development, environmental quality, competitiveness, comprehensive national strength, performance evaluation, etc. These can all constitute the evaluated object.

3. Evaluation indicators. The evaluation index system reflects the quantity, scale, and level of specific evaluation objects from multiple perspectives and levels. It is a dialectical logical thinking process of "concrete abstract concrete", which gradually deepens, refines, perfects, and systematizes people's understanding of the overall quantitative characteristics of phenomena.

The mathematical model for fuzzy comprehensive evaluation is established as follows:

1) Let the triplet (U, V, R) be the comprehensive evaluation space, where:

To evaluate the factor set, it represents the set of evaluation factors in comprehensive evaluation.

As a rating set for comments, it represents the combination of comprehensive evaluation and composition, and essentially divides the range of changes in the evaluated object.

For the fuzzy relationship matrix, R is the result of single factor evaluation, that is, the single factor evaluation matrix. It is also the object of fuzzy comprehensive evaluation. of traditional economy, but also promoted the leapfrog development of social productivity as depicted in Formula 3.

$$P = \frac{\sum_{n=0}^N (C_{0,n} - B_n)(1+r)^{-n}}{\sum_{n=0}^N (1+r)^{-n}} \quad (3)$$

### 3. DISTRIBUTED COMPUTING SYSTEM MODEL DIAGRAM FOR POWER GRID HIGH QUALITY DEVELOPMENT

It consists of a database, data warehouse, data warehouse management module, data mining tools, knowledge base, knowledge discovery module, and human-computer interaction module. The main inputs of the system are data from the database and knowledge and experience from the knowledge base. The data warehouse management module completes the creation of the data warehouse and various operations such as data synthesis and extraction in the data warehouse, and is responsible for the operation of the entire system Data mining tools<sup>[10-15]</sup> are used to complete various queries, multidimensional data analysis, and data mining in practical decision-making problems. The knowledge discovery module controls and manages the knowledge discovery process, using data input and information from the knowledge base to drive the data selection process, knowledge discovery engine process, and evaluation process of discovery The human-computer interaction module provides an integrated interface that connects users and systems through natural language processing and semantic queries.

Data warehouses provide a viable way of organizing data for decision support systems. The data involved in information systems mainly comes from specific daily business data within the department, which can be processed by general databases. This type of data is also known as basic data. The information required for decision-making is the overall trend reflected by basic data or the changing trend manifested over time. It is necessary to classify, extract, summarize, and process the basic data in order to obtain this information.

Logically, a complete data warehouse is defined by four parts: 1) Warehouse design. It is responsible for defining and setting up the data warehouse environment. 2) Data acquisition section It extracts and transforms data from external data sources, organizing and storing them in a data warehouse manner. 3) Data management section. It completes data updates, routine warehouse maintenance, and management of distributed data. 4) Data access section. It is aimed at end-

users and provides decision information and analysis reports to decision-makers in decision support systems as depicted in Fig.1.

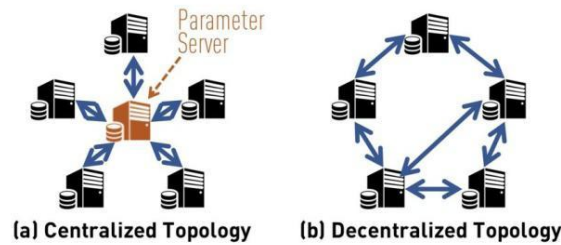


Figure 1. distributed topological diagram.

The process of establishing the decision support system can be described as follows: 1) Analyze decision requirements, determine decision themes, and describe and represent decision problems. 2) Determine the data source, reorganize the operable data records, databases, or file systems in heterogeneous environments, and establish a data warehouse. 3) Design or select effective data mining algorithms and implement them based on the category of the task to be discovered. 4) Call the data mining function to extract comprehensive data from ordinary historical data, and interact and collaborate with end-users to obtain macro level data and trend knowledge. 5) Test and evaluate the discovered knowledge, and apply consistency and utility processing to the knowledge. 6) Based on the requirements of end users, establish an integrated interface and application suitable for decision support, enabling users to apply the discovered knowledge in decision support[16. 17].

Project success evaluation refers to the comprehensive evaluation of various indicators based on the experience of evaluation experts; Or use scoring methods to make qualitative conclusions about the success of the project. The specific approach is to conduct a comprehensive and systematic evaluation based on the evaluation conclusions of the degree of goal achievement and economic benefit analysis analyzed by the logical framework method, with project goals and benefits as the core. 1. Standards for project success

The success of project evaluation can be divided into five levels:

- (1) Completely successful. All objectives of the project have been fully achieved or exceeded; In terms of relative cost, the project has achieved significant benefits and impact.
- (2) Basically successful. Most of the project's objectives have been achieved; In terms of relative cost, the project has achieved the expected benefits and impact.
- (3) Partially successful. The project has achieved some of its original goals; In terms of relative cost, the project has only achieved certain benefits and impacts.
- (4) Not successful. The goals achieved by the project are very limited; In terms of relative cost, the project has generated almost no positive benefits or impact.
- (5) Failed. The goal of the project is unrealistic and cannot be achieved; In terms of relative cost, the project had to be terminated. Design indicator system and weight calculation, aggregation calculation method.

#### **4. SCENARIO FOR POWER SYSTEM DEVELOPMENT WITH DISTRIBUTED COMPUTATIONAL**

Distributed computing is a method of splitting a large computing task into multiple small computing tasks, distributing these tasks across several machines for computation, and then summarizing the computation results of each machine to obtain the final result. This computing method is opposite to centralized computing, which directly processes the entire computing task on one or more computers. The design methods and techniques of distributed computing are significantly different from centralized algorithms, mainly because there are essential differences in system models and structures between distributed and centralized systems as depicted in Fig.2.

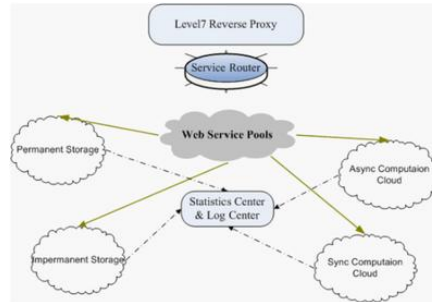


Figure 2. sever connection diagram.

Specifically, the process of distributed computing can be summarized as the following steps:

**Task splitting:** Firstly, it is necessary to split a large computing task into multiple small computing tasks, which can be processed in parallel to improve computing efficiency.

**Task allocation:** Next, these small tasks are assigned to multiple machines, each responsible for processing a portion of the tasks.

**Local computing:** Each machine performs local computing on the tasks assigned to it, which can be done in parallel to speed up the computation.

**Result summary:** After each machine completes local calculations, the results are sent back to a central node or the calculation results of each machine are summarized through the network to ultimately obtain the solution for the entire large computing task as depicted in Table.1.

Table 1. computational parameter.

Parameter	values
Initial investment cost (yuan/Wh)	1
Operation and maintenance cost (yuan/W)	0.04
System residual value rate (%)	5%
System capacity (MWh)	200
Discharge depth (%)	70%
Energy storage cycle efficiency (%)	75%
Annual average number of cycles (times)	600
Cycle life	3000-4200
Energy storage system lifespan (years)	3
Annual decay rate (%)	1.5
Discount rate (%)	6
Tax rate (%)	25%

## 5. CONCLUSION

Distributed computing is a method of splitting a large computing task into multiple small computing tasks, distributing these tasks across several machines for computation, and then summarizing the computation results of each machine to obtain the final result. This computing method is opposite to centralized computing, which directly processes the entire computing task on one or more computers. The design methods and techniques of distributed computing are significantly different from centralized algorithms, mainly because there are essential differences in system models and structures between distributed and centralized systems. This process is not only applicable to numerical calculations, but also widely used in fields such as data processing and machine learning. By utilizing the computing power of multiple computers or servers, it solves large-scale problems that are difficult for a single computer to handle.

## 6. ACKNOWLEDGMENT

This work was financially supported by the Science and Technology Project of State Grid Economic and Technological Research Institute Co. Ltd:Research support for key technologies for high-quality development of power grid (No. B3441324F00100ZQ0000003) fund.

## REFERENCES

- [1] Cambiucci W , Silveira R M , Ruggiero W V .Hypergraphic partitioning of quantum circuits for distributed quantum computing[C]//2023.DOI:10.48550/arXiv.2301.05759.
- [2] Yang K , Sun P , Yang A S L .A novel hierarchical distributed vehicular edge computing framework for supporting intelligent driving[J].Ad hoc networks, 2024, 153(Feb.):1.1-1.16.
- [3] Zhang S , Zhang W , Du J ,et al.Distributed reservoir computing based nonlinear equalizer for VCSEL based optical interconnects[J].Optics Communications, 2024, 563.DOI:10.1016/j.optcom.2024.130574.
- [4] Envelope F Y .A two-level network-on-chip architecture with multicast support[J].Journal of Parallel and Distributed Computing, 2023, 172:114-130.
- [5] Guoliang Z , Zexu D , Yi Z ,et al.Real-time performance evaluation and optimization of electrical substation equipment inspection algorithm based on distributed computing[J].International Journal of Low-Carbon Technologies, 2024.DOI:10.1093/ijlct/ctae136.
- [6] WuChangmao,XuZhengwei,HeXiaoming,et al.Proactive Caching With Distributed Deep Reinforcement Learning in 6G Cloud-Edge Collaboration Computing[J].IEEE Transactions on Parallel and Distributed Systems, 2024.DOI:10.1109/TPDS.2024.3406027.
- [7] Li J , Xie Z , Li Y ,et al.Heralded entanglement between error-protected logical qubits for fault-tolerant distributed quantum computing[J].Science China(Physics,Mechanics & Astronomy), 2024(2).DOI:10.1007/s11433-023-2245-9.
- [8] Shaheer S , Jayaraj P B .Distributed H $\infty$  Method Design and Operation using Dis- tributed Computing[J].ITM Web of Conferences, 2023.DOI:10.1051/itmconf/20235402003.
- [9] Adelman-Mccarthy J , Boccali T , Caspart R ,et al.Extending the distributed computing infrastructure of the CMS experiment with HPC resources[J].Journal of Physics: Conference Series, 2023, 2438.DOI:10.1088/1742-6596/2438/1/012039.
- [10]Li N .Research on Health Management Information Sharing Mechanism in Distributed Computing Environment[J].Applied Mathematics and Nonlinear Sciences, 2024, 9(1).DOI:10.2478/amns-2024-2276.
- [11]Yang J , Sun F , Wang H .Distributed collaborative optimal economic dispatch of integrated energy system based on edge computing[J].Energy, 2023(Dec.1):284.
- [12]He J , Zhang D , Liu S ,et al.Managing Information Updating with Edge Computing: A Distributed and Learning Approach[J].ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023.DOI:10.1109/icassp49357.2023.10095129.
- [13]Frank S , Hoch E N .DISTRIBUTED FILE SYSTEM FOR VIRTUALIZED COMPUTING CLUSTERS:US18049269;US202200018049269;US202218049269A;US202218049269[P].US20230085566A1; US2023000085566A1;US2023085566A1;US2023085566[2024-09-19].
- [14]Cuomo D , Caleffi M , Krsulich K ,et al.Optimized Compiler for Distributed Quantum Computing[J].ACM Transactions on Quantum Computing, 2023.
- [15]Jones G M , Jacobsen H A .Distributed Quantum Computing for Chemical Applications[J]. 2024.
- [16]Sazontev V V , Stupnikov S A .An Extensible Approach to Searching and Selecting Data Sources for Materialized Big Data Integration in Distributed Computing Environments[J].Pattern Recognition and Image Analysis, 2023, 33(2):147-156.DOI:10.1134/S1054661823020141.
- [17]lt,HTML&gt, Interface R M A ,et al.Demonstrating interactive, large-scale High Energy Physics data analysis workflows in distributed computing environments [J] 2024.