Research on Low-Power Main Control Chip Architecture Based on Edge Computing Technology

Guoyu Cui1, *, Wei Ye2, Zhanbin Hou1, Tong Li3, Ruolin Liu2
1Beijing SmartChip MicroElectronics Technology Company Limited, Beijing 100192, China
2State Grid Zhejiang Electric Power Corporation Information & Telecommunication Branch, Hangzhou 310020, China
3State Grid Jilin Electric Power Company, Changchun 130021, China
*Corresponding author: cui_guoyu2020@foxmail.com

Abstract. With the continuous development of the times, chip technology has also been innovated. In order to meet the needs of edge computing networks and data communication networks, it is necessary to make chips with shorter delay time and lower power consumption, which in turn brings challenges to chip architecture design. In the field of chip design, the problem of chip performance has been overcome. In many edge design networks, the core of the design is to allow the chip to meet low power consumption requirements, thereby extending the use time of edge devices. This article conducts research on a low-power master chip architecture to enable it to meet the needs of edge computing networks and enhance product competitiveness.

1. Introduction

In recent years, the speed of the development of information technology has changed with each passing day. The Internet of Things and artificial intelligence have entered people's field of vision. The development of these technologies not only brings innovation, but also makes people's lives undergo earth-shaking changes. However, with the continuous development of the Internet of Things technology, the number of network devices that need to be invested is also increasing, thereby greatly increasing the amount of data transmission. In this context, the cloud computing model can no longer meet the requirements of data security and real-time. In order to improve this situation, researchers put forward an edge computing model. By using edge computing, some cloud computing can be carried out on terminal edge devices, thereby shortening processing delay time, reducing power consumption, and satisfying data transmission bandwidth requirements. Therefore, it is necessary to design the product chip so that it can meet the above requirements. In the chip process, the chip area is gradually reduced and the performance is gradually improved, which in turn leads to an increase in chip power consumption. When designing the chip architecture, low power consumption needs to be considered. This article designs a low-power main control chip. Compared with the traditional chip architecture, it has lower power consumption, which can extend the use of equipment and reduce cost.
2. Edge computing overview

2.1. Edge definition
Edge computing is a new type of computing paradigm, which is a type of computing performed at the edge of the network, and computing at the data source is the core of this technology. China's relevant alliance defines it, that is, edge computing is a development platform, which is based on the edge of the data source network, and through the integration of multiple core capabilities, including storage, network, etc., so that edge intelligent services can be provided and digital The industry's delay safety and application intelligence are satisfied. The edge computing is shown in Figure 1.

![Figure 1. Conceptual diagram of edge computing](image)

2.2. Overview of Edge Computing and Chip Development
American semiconductor company AMD launched two processors in 2018, namely Ryzen V1000 series and EPYC 3000 series. The processors are all oriented to edge computing and the CPU architecture is ZEN. ARM announced a project in 2018 to develop a smart chip to enable artificial intelligence and machine learning for edge-side devices. Google designed an Edge TPU, which can run on edge devices, making high-performance computing possible on the basis of low power consumption and small physical footprint. In the same year, Intel Corporation launched an artificial intelligence processor, the model is Xeon D (Xeon D), the Skylake architecture is the processor architecture, which can be applied in the edge or restricted environment, thereby making it intelligent support and power consumption Wait for the problem to be resolved. At the Intel Artificial Intelligence Conference, the second generation of a neural computing stick was launched, which can be used in edge networks to enable computer vision equipment and artificial intelligence algorithms to be built. Intel believes that in the field of vision, edge-side artificial intelligence can play more roles, such as the field of medical imaging and machine vision. Based on the visual processing unit, the second generation of Intel Neural Compute Stick optimizes product performance, thereby reducing the cost of developing deep neural network inference applications. Not only that, Intel is also developing a new network system chip to enable it to be used in edge computing and 5G networks. The product is expected to be launched in 2020.

2.3. Advantages of edge computing
Compared with cloud computing, edge computing has many advantages. For example, all data does not need to be uploaded to cloud computing, and only data needs to be exchanged in the edge design. Therefore, its advantages include:
(a). Bandwidth cost. In order to meet the needs of the Internet of Things, it is often necessary to access more Internet of Things devices, which in turn puts greater pressure on network transmission. In edge computing, because there is no need to exchange a large amount of data, it will not put great pressure on network transmission. Because edge computing handles small data, no matter whether it is storage or data calculation, it does not require much investment. Therefore, it is clear that the use of edge computing can not only reduce the amount of data transmitted on the network and reduce bandwidth pressure, but also increase the computing efficiency of cloud computing and enable it to be applied in more fields.

(b). Interaction delay. Now that the Internet of Things is becoming more and more popular, the amount of data that needs to be uploaded has increased significantly. If the processing data needs to be uploaded to the cloud computing center, it will inevitably cause the network to face greater bandwidth pressure, and it will also cost more when searching for massive amounts of data. More time. In edge computing, because it is closer to the data source, data processing can be faster, because the data transmission process is reduced, so it can have a faster data processing speed. If the application has strict requirements for response time, the immediacy and timeliness of edge computing can meet the needs of this type of application, so it can be used in fields such as video surveillance and autonomous driving.

c). Data Security. Analyzing the cloud computing model can make it clear that when using this model, all user-side data needs to be uploaded. In this process, data security issues are the core. Uploaded data includes search history, financial account passwords, etc. In the process of uploading data to the data center, commercial secrets or personal privacy data are at risk of leakage. In the edge computing model, this problem can be effectively improved. The calculation and collection of data can be completed locally without uploading data to the cloud, so that data will not be leaked due to transmission. Therefore, it has high security.

3. Research on Low Power Chip Architecture Based on Edge Computing

3.1. Low power design technology
The chip mainly includes analog modules, multimedia modules, DSP sub-modules, on-chip memory modules (SRAM), CPU modules (ARM), etc.

According to relevant research materials, it is clear that the power consumption of a chip consists of two parts, one part is static power consumption and the other part is dynamic power consumption. The former is related to the transistor. In the off state, the transistor may have leakage current, thus generating static power consumption. If the chip process is low, such as before 180nm, the main power consumption of the chip can be ignored. However, if the chip process is relatively High, then you need to consider the static power consumption of the chip. The latter can be referred to as switching power consumption. In the circuit, if a logic flip occurs, there will be energy consumption, which is dynamic power consumption. When the logic is inverted, the load capacitor will be charged and discharged. Therefore, in a single circuit, the energy loss generated by the capacitor during the charging and discharging process can be approximated as power consumption.

3.2. Low power mode design
When designing the chip architecture for low power consumption, design the chip application scenarios. This article designs three chip application scenarios so that the main control chip can meet the power consumption requirements of different scenarios. Active, Sleep and Stop are three scenarios designed for this article, and different scenarios correspond to different working modes. When the edge device is working normally, the corresponding scene is Active, when the edge device is in battery-powered mode, the corresponding scene is Sleep, and when the edge device is in the very low power consumption mode, the corresponding scene is Stop. When switching between different modes, the module used is the PMU module. After completing the programming of the relevant control registers, you can use the PMU to switch the different working states of the edge device, such as turning off the power module, turning off the clock module, and so on. In different working states, the chip working states are:
(a). Stop mode
In this mode, the clock of the main control chip will be switched to a low-frequency state, and the PMU will turn off the two power domains. All modules working in this power domain will be in a power-down state, which in turn makes it have lower power consumption. The chip will retain two wake-up functions, one wake-up function is key-press wake-up, one wake-up function is timed wake-up, and the other wake-up functions are in a restricted state. On this basis, only two parts will generate power consumption, one is the button wake-up module, the other is the timing wake-up module, and the remaining modules are in a low-power state.

(b). Sleep mode
In this mode, most functional modules will be turned off. At this time, the interrupt processing module will work in low-frequency mode. In this mode, multiple modes can be used to wake up the device, including comparator wakeup, key wakeup, etc.

(c). Active mode
In this mode, the main control chip is in the normal working mode. Analysis of this mode shows that the chip can work in frequency division mode or high frequency mode. When the CPU is turned on, the CPU will complete the working mode setting of the surrounding modules by running software.

Analyzing the three modes, it is clear that the most used mode of functional modules is Active, so it has larger power consumption, and the mode that uses the least functional modules is Stop, so it has smaller power consumption.

3.3. Low power consumption design
When designing the low-power architecture of the main control chip, based on the three modes, through the use of multi-power domain design, the low-power requirements can be met. When designing the chip, divide its internal power supply into three parts, namely VCC33, VDD18 and VDD18LC. When the chip turns on the VCC33 power supply, it can make a variety of modules work, including comparator (COMP), display driver (DISP DERIVER), temperature sensor (PTAT), etc. When the chip turns on the VDD18 power supply, it can provide working voltage to the digital circuit, so that the digital function module can obtain the working voltage, such as RAM controller, M0 CPU, etc. When the chip working mode is M0 CPU, there are two power domains for shutdown mode, namely VCC33VDD18 and VCC33, so that the quiescent current of the chip can be reduced. When the chip is in normal operation, VDD18LC needs to be in normally-on mode, so it is clear that although it belongs to the working voltage domain of digital circuits, it cannot be locally turned off. There are multiple circuits working in this voltage domain, including Button detection circuit, compensation circuit, oscillation circuit, etc.

3.4. Low power circuit design
Analyzing the circuit can make it clear that the circuit scale and dynamic power consumption are proportional, and the circuit flip frequency and dynamic power consumption are proportional. In the main control chip, 30% of the chip's dynamic power consumption comes from the clock dynamic power consumption. Therefore, if you want to reduce the power consumption of the chip, you need to reduce the signal flip frequency and clock frequency. When designing the architecture of the main control chip, this article chooses to use the following techniques to reduce dynamic power consumption, including:

(a). Design asynchronous circuits
In the synchronization circuit, the global clock controls the entire system. When the clock pulse occurs, the flip-flop is run, which in turn generates dynamic power consumption, although in this process, there is no new data that needs to be accessed. In an asynchronous circuit, there is no invalid power consumption. When the system needs to work, the circuit power supply will start, so that the system module can work normally. After it finishes its work, the system module will stand still, and there will be consumption leakage current. But no dynamic power consumption is generated. When designing the main control chip, asynchronous circuit design can be used, so that some circuits, such as button detection circuits, can reduce dynamic power consumption.

(b). Low-power bus technology
When designing the main control chip, the modules in the chip need to be interconnected. At this time, a standardized bus is required. By introducing the bus, the chip design can be modularized, making system integration more convenient, and it can be expanded when needed. System to make the system meet the demand. In the chip, the bus signal can be extended to the functional modules. The bus is mutually exclusive. When the master device is accessing, it will only access one device at the same time. On this basis, there will still be devices that are not accessed. The inversion signal is generated, which causes the bus to have higher switching power consumption. In the design of this article, in order to reduce the bus power consumption, choose to use the bus reverse technology. In the clock cycle, use the ideal module of the bus matrix to compare a value on the bus with the current value, choose whether to send the inverted code or the original code, and select the sending method according to the bus inversion signal condition. In the design, by adding the polarity signal, Let the bus data be correctly restored by the receiving module. By using this method, the dynamic power consumption of the bus can be reduced, thereby reducing the power consumption of the main control chip, so that it can meet the low power consumption requirements.

c). Gated clock technology

When designing the functional mode, let the clock gating be independent, that is, when the module is not working, the module clock can be turned off, thereby reducing the power consumption of the circuit flip; in addition, when synthesizing the chip code, design the clock gating Unit, and then make the chip not flip the signal, the clock will not flip. By using this method, the dynamic power consumption of the clock can be reduced.

4. Test Results

After completing the design, test the power consumption at room temperature. The test results are:

(a). Stop mode: In this mode, the I0 detection module, PMU, etc. are in working state, RAM is powered off, and 0.69 uA is the current intensity of the chip;

(b). Sleep mode: In this mode, the interrupt control module, IO detection module, etc. are in working state, 1.25uA is the current intensity of the chip;

(c). Active mode: In this mode, the CPU executes instructions from RAM, and 120uA/MHz is the current intensity of the chip.

Comparing the above data with other similar main control chips, it is clear that it has lower power consumption, so it can meet the design requirements of low-power main control chips.

5. Conclusion

With the continuous reform of chip technology, chips and edge computing have gradually embarked on a road of coordinated development. In the context of the continuous popularization of the Internet of Things and smart mobile terminals, the amount of data exchange and upload has also increased significantly, making cloud computing unable to meet low-latency requirements. By using edge computing, data analysis and processing can be completed at the edge, and then Meet the low latency requirements. If you want edge computing to develop, you must have chip support. When designing a main control chip based on edge computing, it is necessary to comprehensively consider its size, power consumption, and performance, so that it can meet the needs of edge computing. Analyzing the development trend can clarify that in order to meet the development of edge computing, the chip architecture needs to have a high degree of complexity, it needs to be able to adapt to a variety of scenarios, and it needs to be able to support artificial intelligence algorithms to improve the efficiency of resource utilization. This paper studies a master chip architecture based on edge computing, expounds on three different working modes, and designs low power consumption from the perspective of power supply and circuit for different modes. After completing the design work, its power Consumption is tested. The results show that the main control chip designed in this paper has lower power consumption than similar chips, which can effectively extend the use time of edge computing devices and meet the design requirements.
Acknowledgment
This work is supported by the Science and Technology Research Project of State Grid Corporation of China, project No.: 546816190007: Research on the key technology of microcontroller for intelligent acquisition unit based on edge computing and Near-Subthreshold.

Reference
[1] Dai Yueyue. Research on resource management issues in mobile edge computing [D]. University of Electronic Science and Technology of China 2019
[2] Li Yang. Research on several issues of energy-saving and efficient resource joint optimization in mobile edge computing [D]. Jilin University 2020
[3] Zhang Yaqiang. IoT event boundary detection and complex task scheduling optimization under edge computing [D]. China University of Geosciences (Beijing) 2020
[4] Chen Menggang. Research on Multi-user Dynamic Task Offloading Strategy of Mobile Edge Computing [D]. Southwest University 2020
[5] Li Yuqing. Research on collaborative optimization of task scheduling and resource allocation in mobile edge computing networks [D]. Shanghai Jiaotong University 2019
[6] Liu Lei. Research on on-board data distribution mechanism driven by edge computing [D]. Xidian University 2019
[7] Zhao Zhaoling. Application of Active Network Technology in 3TNet Edge Network [D]. Chinese People's Liberation Army Information Engineering University 2004
[8] Wen Wanli. Research on resource optimization technology of mobile edge storage, computing and wireless communication [D]. Southeast University 2019
[9] He Lingjuan. Study on unidirectional terahertz edge mode in semiconductor photonic crystals [D]. Nanchang University 2019
[10] Li Ning. Study on the measurement and characterization method of edge roughness of nano-scale semiconductor scribed lines [D]. Harbin Institute of Technology 2010