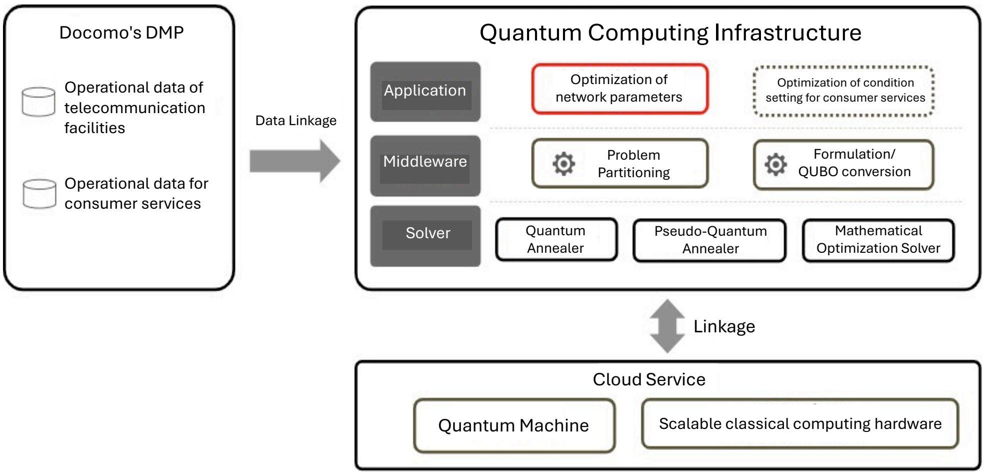
Topics

June 27, 2024
NTT DOCOMO, Inc.

Developing Quantum Computing Infrastructure for Service Optimization

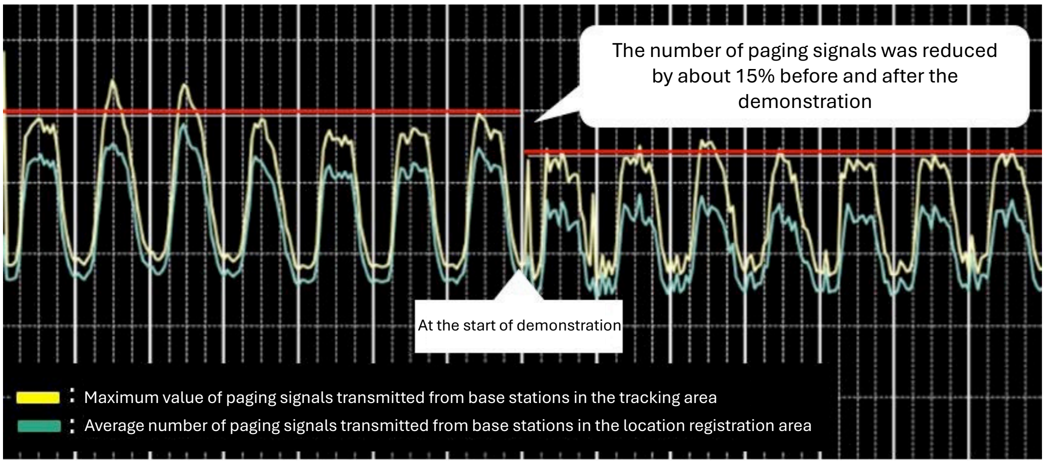
～Application to base stations to optimize communications services start in July～

NTT DOCOMO, INC. (hereinafter: DOCOMO) aims to optimize its services, and by focusing on quantum computers [1], which are expected to be a next-generation computing resource, we have developed a quantum computing infrastructure (hereinafter: the infrastructure) that utilizes quantum annealing technology. The infrastructure will be gradually applied to base stations nationwide starting in July 2024 with the goal of optimizing communication services.



【Figure 1: Overview of quantum computing infrastructure】

DOCOMO will utilize the infrastructure to reduce the load on base stations and ease congestion during times of high call volume. To be more exact, the infrastructure will reduce the number of paging signals [3] transmitted from base stations to customers' devices and mark the world's first [4] development of an algorithm that reduces paging signals using a quantum computer. To test the usefulness of the infrastructure, we conducted demonstration experiments at base stations in the Tokai, Chugoku, and Kyushu areas, demonstrating [5] and confirming that paging signals can be reduced by up to 15%. This equates to being able to connect approximately 1.2 times more devices than it's currently possible during periods of high call volume.



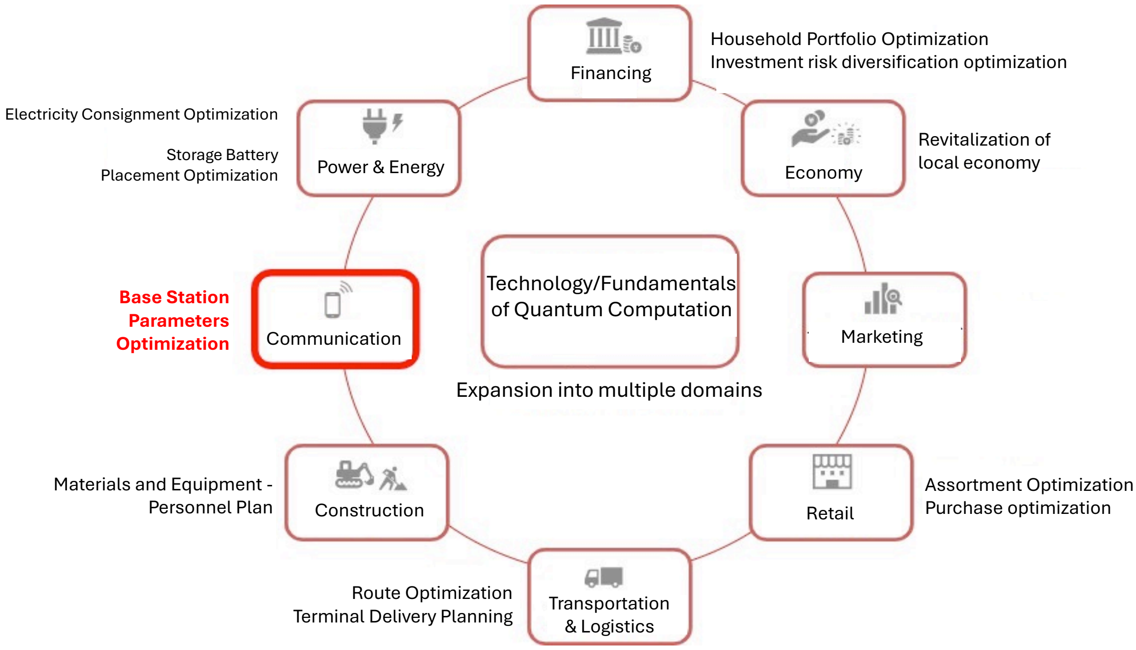
【Figure 2: Number of paging signals reduced in the demonstration experiment】

(Vertical axis: number of paging signals; Horizontal axis: date (one week before and after the demonstration experiment))

In mobile communications, a paging signal is sent from the base station to the customer's smartphone or other terminals so they can receive calls and their location can be determined by receiving a response from the terminal. The data volume transmitted through paging signals has been increasing, and if there is no response from the terminal, data must be resent over a wider area. In addition, the recent proliferation of smartwatches and IoT devices has resulted in an increased volume of paging signals, which has put a burden on base station facilities, generating an issue.

To solve this issue, we will utilize the quantum-computing-based infrastructure to enable data processing on a scale that surpasses conventional computers, thereby reducing the number of paging signals. Specifically, based on statistical data consisting of a large volume of log data that records at what base stations have terminals received/ended calls, the system optimizes groups of base stations that collectively send paging signals (hereinafter: tracking area [6]). Utilizing the characteristics of quantum computers allows us to predict the number of paging signals for all base stations' combinations simultaneously, making it possible to determine the optimal tracking area in an abbreviated time to find a terminal with fewer paging signals from a high number of combinations. This minimizes the number of signals, provides a safety margin for base station resources, and contributes to the stability of communication quality so that communication is unaffected even when the call volume is high.

In the future, DOCOMO aims to provide convenient services by deploying the infrastructure not only in the telecommunications domain but also in various other domains of services offered by DOCOMO.



【Figure 3: Image of quantum computing infrastructure utilization】

[1] Quantum computer: A next-generation computer that is expected to be able to process information at a speed and scale faster than a conventional computer by utilizing "superposition" and other characteristics of "quantum", which makes up atoms and other matter to enable parallel computing.

[2] Quantum annealing: Algorithm for solving optimization problems using quantum fluctuations

[3] Paging signal: A signal sent by a base station to a terminal to determine the location of the terminal when a call is received in mobile communications. The location of the terminal is determined by the response from the terminal to the base station

[4] DOCOMO survey as of June 27, 2024

[5] Appendix 2 provides details of the demonstration experiment.

[6] Tracking area: Group of base stations that collectively send out paging signals to terminals

For inquiries regarding this matter, please contact:
NTT Docomo, Inc., Crosstech Development Dept. Head of FinTech Technology Development Email: xt2_quantum-pj@ml.nttdocomo.com

Appendix 1

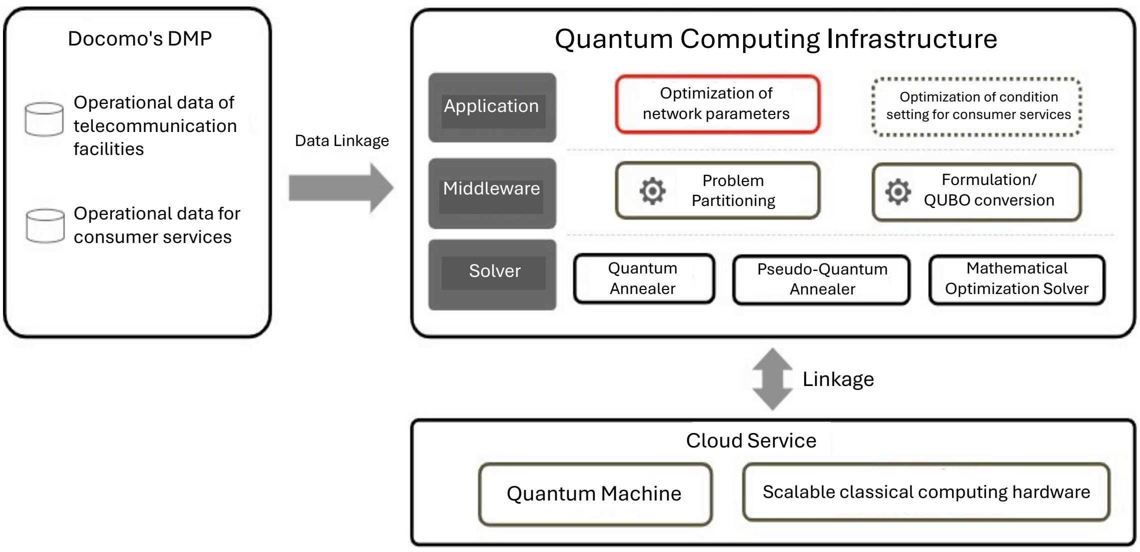
Overview of Quantum Computing Infrastructure

1. Functional Composition of the Quantum Computing Infrastructure

An overview of the quantum computing infrastructure is shown in Figure 1 below. The functionality of the infrastructure to perform optimization calculations consists of the following three layers:

- Solver [7] layer: Equipped with software to solve optimization problems.
- Middleware layer: Equipped with software for the common processing of all applications.
- Application layer: Equipped with software for problem-specific processing. In addition, the following two functions are provided to link the infrastructure with external systems.
  - Data acquisition function linked to DMP [8] owned by DOCOMO.
  - Linked to cloud services.

The aforementioned functions allow the optimization of the entire data-obtaining process, securing computing resources, performing optimization calculations, and obtaining the results to be completed within the infrastructure, which in turn allows rapid optimization of services when necessary.



【Figure 1 Revisited - Overview of Quantum Computing Infrastructure】

2. Features of the Infrastructure

The infrastructure was built with both stability and flexibility to utilize combinatorial optimization technology and, as a result, improve the service quality during operation. Generally, quantum computers are special hardware cooled to cryogenic temperatures and, as a result, require frequent maintenance. Quantum computers cannot be used during maintenance. This presents an issue as the operation of services that require optimization would be halted. Therefore, we were able to avoid operational downtime due to maintenance by introducing multiple solvers and increasing redundancy.

In addition, we have improved the analysis environment flexibility by introducing a combination of quantum computers manufactured by multiple companies. Combinatorial optimization problems (mathematical problems that search for the best combination of possible values for variables) are characterized by two factors: the scale of the problem (the number of variables to be considered) and the degree of coupling [9]. On the other hand, different types of solvers have different characteristics depending on the computer. Some solvers are good at large-scale problems, and others are good at problems with high-degree coupling.

In order to handle optimization problems of various problem scales and degrees of couplings, DOCOMO introduced an annealing-type quantum computer specialized for combinatorial optimization problems and designed the infrastructure so that multiple computers with different problem scales and coupling requirements can run on it. This makes it possible to select the optimal solver for the characteristics of the problem and solve the problem with high accuracy.

[7] Solver: annealing quantum computer and other software with algorithms for solving mathematical programming problems

[8] DMP (Data Management Platform): Platform for managing a variety of stored information

[9] Cohesiveness: a measure of the extent to which the variables to be optimized influence each other.

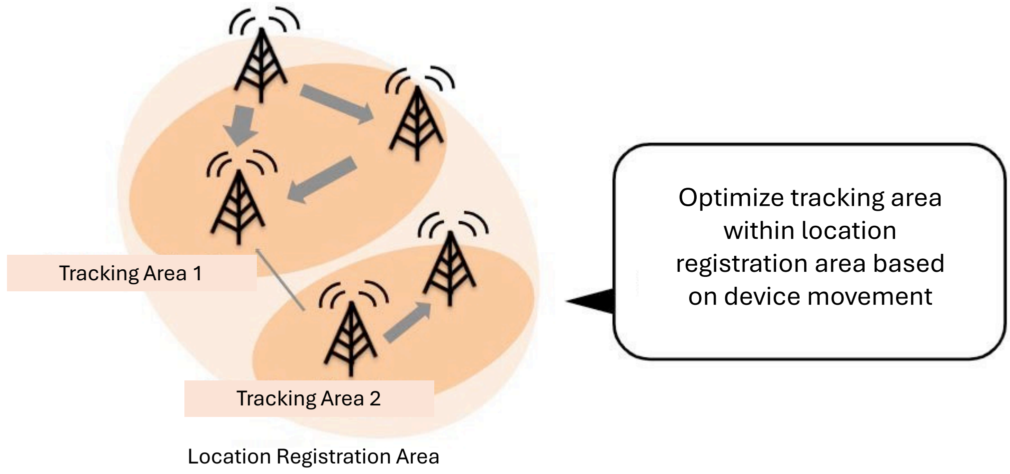
Demonstration Experiment Overview

1. The Aim of the Experiment

Currently, the number of paging signals emitted by the network to determine the device's location when a call is received on Docomo's 4G mobile network is increasing, so in this experiment, a quantum computer was used to optimize the grouping of base stations that simultaneously perform paging (setting of tracking areas) and to reduce the number of paging signals.

2. The Problem of Minimizing Paging Signals

When a terminal receives a call, the base station sends out a paging signal within the location registration area [10] to determine the terminal's location. By simultaneously sending a paging signal from a group of multiple base stations within the location registration area, the location of the terminal can be determined even if the terminal has moved from the base station with which it communicated immediately before. By grouping base stations that have a large number of terminal movements and paging them on a group-by-group basis, terminals can be found with fewer paging signals. In this demonstration, we verified the effectiveness of the annealing-type quantum computer for the combinatorial optimization problem (see Figure 4) of finding a group of base stations that minimizes the number of paging signals in a network.



【Figure 4: Base station grouping image】

3. Summary of the Demonstration

The following efforts were made in this demonstration:

- Gathering traffic data related to incoming base station calls and terminal movement over a wide area (in this case, Tokai, Chugoku, and Kyushu areas)
- Development of a mathematical model to predict the number of paging signals that will be generated under base station grouping conditions based on the conditions under which traffic data is obtained.
- Development of an algorithm (QUBO format [11]) for an annealing-type quantum computer to search for grouping conditions that minimize the number of paging signals predicted by mathematical models among all base station grouping conditions.
- Establishment of an operation to apply the optimal grouping conditions to actual networks.
- Establishment of a method to verify the effectiveness of applying the optimal grouping conditions to actual networks.

4. Utilization of the Infrastructure in this Demonstration

In this demonstration experiment, the quantum computing infrastructure was used to aggregate and analyze vast amounts of communication traffic data, perform optimization calculations using quantum annealing, and estimate the expected reduction effect all at once. With this, the time required for this entire process was significantly reduced.

In addition, the infrastructure enables the adoption of optimization methods appropriate for the mobility trends of terminals in each region. Specifically, it has been confirmed that in urban areas, terminal movement tends to occur in a complex manner within a small area, while in suburban areas, movement tends to be more even across a wide area. These differences in movement tendencies affect the scale of the problem and the degree of coupling of variables when applied to optimization problems.

In this way, the infrastructure does not only provide an environment suitable for the use of quantum computers but also contributes to the success of PoC as a tool that comprehensively supports everything from data acquisition and preprocessing to the practical application of optimization results.

5. Results of the Demonstration Experiment

As a result of demonstration tests in the Tokai, Chugoku, and Kyushu areas, the number of paging signals during peak hours has been successfully reduced by up to 15% and by 7% on average. This corresponds to approximately 1.2 times the number of terminals that can be connected during times of high volume of incoming calls. This allows more bandwidth for base station resources, reducing the risk of congestion during times of high volume of incoming calls and stabilizing the quality of communications.

The optimization problem for this demonstration requires an enormous number of calculations, and conventional computers could not keep up with the number of calculations in a realistic amount of time. However, by utilizing an annealing quantum computer, the optimal solution was calculated in only about a minute, making it possible to take this initiative to a nationwide scale.

The fact that a high-performance solution was obtained in a short time for a large-scale combinatorial optimization problem of base station grouping based on terminal movement indicates that quantum computing has the potential to make a significant contribution to improving the quality of communication networks. DOCOMO will continue to improve the functionality of the infrastructure to provide better communication services.

[10] Location Registration Area: Areas covered by multiple base stations are collectively referred to as a location registration area, and the location of terminals is managed in units of location registration areas.

[11] QUBO format: A method of finding a solution to a problem by expressing the problem in a mathematical expression using the values 0 and 1 and then optimizing the expression. By converting to QUBO format, the problem can be solved by an annealing quantum computer.