Service Monitoring | Probe System | Distributed Database

Technology Reports

A Service Operations System Monitoring User Perception of Quality in Real Time

As can be seen with smartphones, use of mobile terminals is diversifying and it is becoming difficult to maintain stable service quality using only conventional NE monitoring methods. To deal with this issue, we have developed a service operations system equipped with functionality to monitor service quality directly. In this article, we describe the need for service monitoring and give an overview of this system.

Core Network Development Department **Shinya Aoyama**
Hideki Toube
Hiroki Muramatsu
Network Service Operation Department **Saki Iwami**

1. Introduction

In recent years, mobile terminals have become established as part of our life infrastructure and, as can be seen with smartphones, a variety of ways of using them are being offered.

Till now, stable quality in communication networks has been maintained by monitoring Network Equipment (NE). NE monitoring refers to detecting faults on the network, analyzing them, and carrying out recovery measures based on information in notifications sent periodically by NE, or from queries sent periodically to NE regarding their operational state. However, when using NE monitoring to understand the state of service provision, the monitoring staff must often use know-how and experience to decide whether services will be affected from the NE state and warning information, and there are limits to what can be done to make the state of service provision more visible. Further, to maintain stable service quality amid diversifying usage, as described above, requires monitoring each user's perception of quality in real time as well as sophisticated analysis.

As such, NTT DOCOMO has developed functionality to monitor the services themselves together with conventional NE monitoring. In this article, we describe the need for service monitoring and give an overview of this system.

2. The Need for Service Monitoring

2.1 NE Monitoring Issues

Conventional NE monitoring had the following issues.

1) Issue (1): Difficult to Determine Effects on Services Accurately

The effects of a fault detected in equipment on services were determined based on the know-how and experience of the monitoring staff. There are ways to associate faults detected in equipment with the effects they have on services and manage them, but there are many potential patterns of equipment failure, so it is difficult to do this reliably. Also, with the diversifying ways in which mobile terminals are being used recently, and the various services being provided, it is becoming much more difficult to associate equipment with services.

2) Issue (2): Difficult to Detect Silent Failures

With NE monitoring, warnings are only sent for faults that the NE itself detects. Because of this, silent faults that the NE cannot detect itself are difficult to detect. Also, in cases such as errors in configuration of NE, the NE is operating correctly according to its configuration, but the service may not be being provided correctly. Such faults are difficult to detect, and are often come to light only by reports from users. As a result, corrective measures can be delayed and at times can become large-scale faults with many users affected.

3) Issue (3): Difficult to Investigate the Cause of Difficult-to-reproduce Phenomena

To analyze faults discovered through user reports can be difficult utilizing NE warnings alone. In such cases, a test call is usually placed to try to reproduce it, or logging is configured on a specific phone number before waiting for the phenomenon to reoccur. However, when phenomena are not easily reproducible, they cannot be checked. This can prolong the investigation and delay finding a solution.

2.2 Service Monitoring Overview

As described in section 2.1, there are limits to detecting the effects on services accurately when using only NE monitoring. Thus, a means of monitoring whether the service itself is being provided properly (service monitoring), regardless of the state of NE, is needed. The differences between NE monitoring and service monitoring are shown in **Figure 1**.

Considering the issues with NE monitoring, elements required of service monitoring include the following.

- Making service-status visible

  Whether a service is being provided or not must be based on objective facts and is not something for conjecture. Further, it is essen-

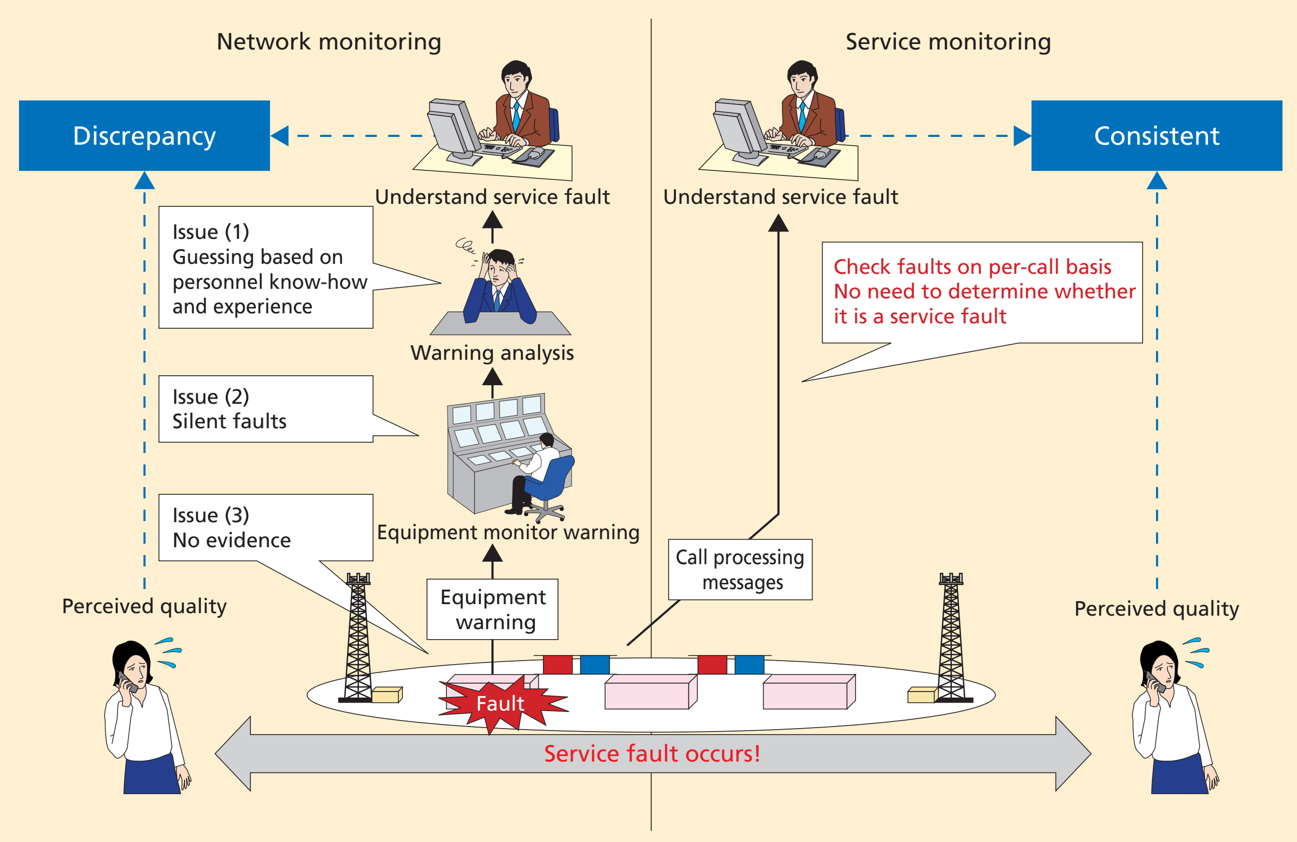

Figure 1 Differences between network and service monitoring

tial to have an accurate understanding of the extent to which services are being affected.

- Detecting service faults in real time

  There are potential phenomena that will have an effect on service quality, even if they are not detected as faults in NE, so service quality itself must be detected in real time.

- Checking the cause of phenomena

  To prepare for problem phenomena that are difficult to reproduce, some evidence that can be checked must be left behind.

To summarize, rather than the actual state of the NE, the behavior of the NE is what is important. That is, understanding how call processing is occurring, and treating that as the most basic issue.

2.3 Functions Required for Service Monitoring

The functions required for service monitoring are shown in **Figure 2**.

- Service fault monitoring function

  This function detects service faults (service interruption or decrease in service quality) and displays them in the form of warnings. To show clearly the effects of the service fault, warnings must show the name of the service affected, the scope of effect, and the number of occurrences in real time.

- Service quality management function

  This function generates service quality indices, and enables them to be stored and searched. It clarifies levels of satisfaction for defined service quality indices, and enables flexible retrieval of quality index data so that the causes can be analyzed when problems occur.

- Communication history display function

  This function stores, searches and displays Call Detail Recordings (CDR) that enable phenomena to be checked and problems analyzed in cases such as when users submit complaints.

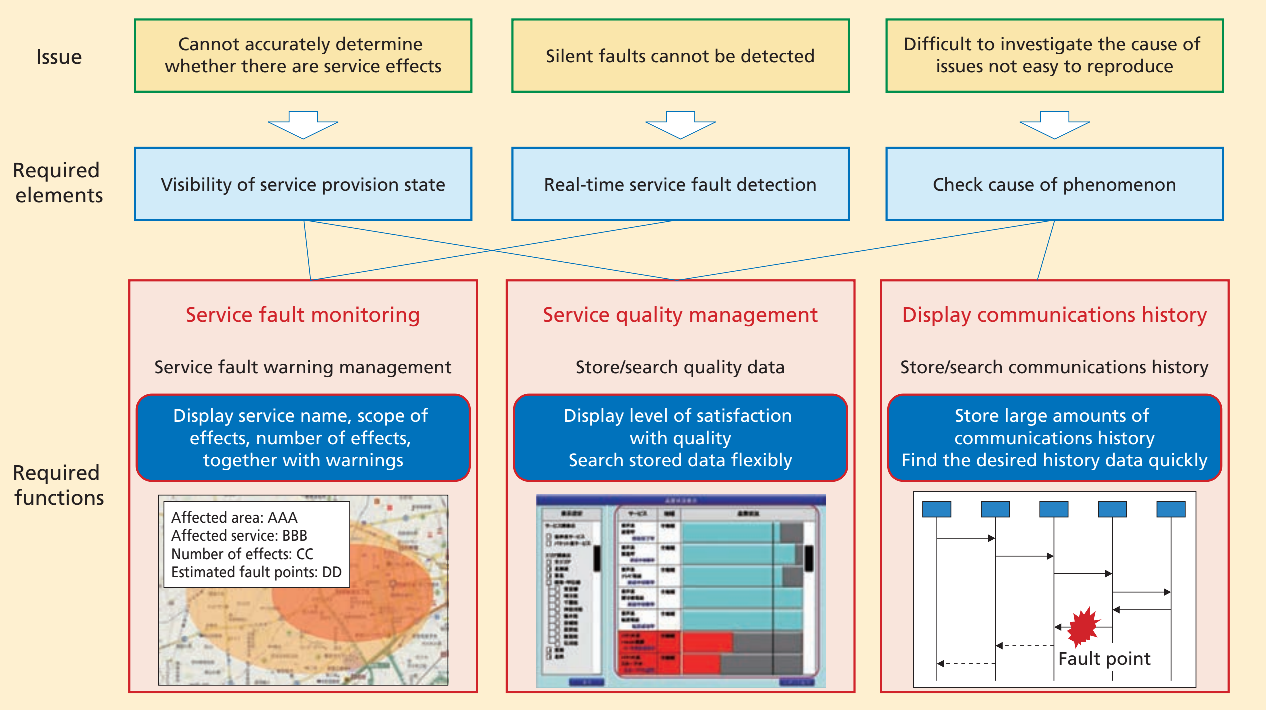

Figure 2 Functionality required for service monitoring

It must be possible to specify time, service type, subscriber number and other details and immediately display a search result from among the large amount of call history data. Also, when necessary, it must be possible to display the communications messages (raw data) corresponding to a CDR containing a fault.

3. Service Monitoring Operation System Implementation

3.1 Implementation

1) Service-fault Monitoring and Service Quality Management Functions

A common means of managing service quality involves computing and managing Key Quality Indicators (KQI), which are service quality indices, based on Key Performance Indicators (KPI), which are network quality indices[1]. KQIs are computed by combining multiple KPIs, and by combining them in units of the item being monitored, they can compute quality indices by service or by building. The scope of effect of a service fault can be monitored clearly by statistically analyzing KQIs and detecting abnormalities in real time.

There are two main methods for obtaining a KPI:

- Obtaining it as NE traffic

An advantage of obtaining KPI as NE traffic is that no investment in special equipment is necessary. Also, as long as simple signals can be counted per piece of equipment, KPI can be generated easily. However, call processing is ultimately the most important task for NE, and generating traffic data must be done under the constraints of excess resources that will not affect call processing. Thus, performing more-detailed analysis, such as obtaining information by terminal type or protocol, can raise concerns that it will place pressure on processing load or storage capacity. Also, since faults must be detected immediately, if traffic information is used for service monitoring, KPI must be generated in real time from the traffic data, and this can also result in increased load for NE.

- By obtaining message data on the network using probes[*1]

Obtaining KPI using a probe involves using an external server to perform the counting process. The counting process load does not affect call processing, so all of those resources can be devoted to generating the KPI. Thus, counting in more detailed units is possible, together with analysis of more-detailed quality estimates, such as sound quality, peak throughput, and response time per protocol.

2) Communication History Display Function

Communication history includes CDRs and communication message data. CDRs provide a summary of call processing. They can be generated easily by the NE processing the calls and do not add much to processing load. To generate communication message data, a large amount of call processing messages are stored on an internal disk and associated with CDRs for search purposes. This process cannot be done practically without effecting call processing. A probe does not perform call processing, so it is able to generate communication message data and also associate it with CDRs.

3.2 Implementation Policy

Use of mobile terminals has been diversifying recently, and it is becoming necessary to obtain KPIs for more-detailed service units. For this development, we studied methods of retrieving KPI and communications history directly from the NE. **Table 1** shows the results of comparing different methods. We decided to use probes based on the results of this comparison, and because probes also have the following advantages.

- When there is a service faults it is analyzed based on communications message data, so a function to reference this data is essential.
- Probe products are widely available on the market, and prices are tending to decrease. Also, a probe product suited to the application can be selected and purchased.
- In most cases, probe products on the market provide interfaces to output KPI to external systems.

*1 **Probe**: A device that retrieves data flowing on the network, analyzes the protocols, and computes a KPI.

3.3 System Organization

The service monitoring system structure is shown in **Figure 3**, including a Data Capture (DC) component and a MoniToring (MT) component.

The DC component has a structure that allows suitable probe products to be selected from among the many on the market, able to generate KPI and to support the various protocols in use on mobile networks, according to requirements. The MT component provides central management of the KPI obtained in the DC component, and forms the base for service monitoring.

The MT component must also be customizable flexibly, in terms of KQI generated and fault detection functions, according to monitoring policies and the services being monitored. Because of this, and in order to be able to match operability with other operations equipment and integrate with operator workflow, we decided to develop our own solution.

1) MT Component Implementation Issues

Four issues arose in developing the

Table 1 Comparison of retrieval methods

Type	Item	Network retrieval	Probe retrieval
KPI	Real time	△	○
	Count signals	○	○
	Estimate quality	×	○
	Count units	By equipment, etc.	By device By terminal type By protocol, etc.
Communications history	CDR	○	○
	Raw data	×	○

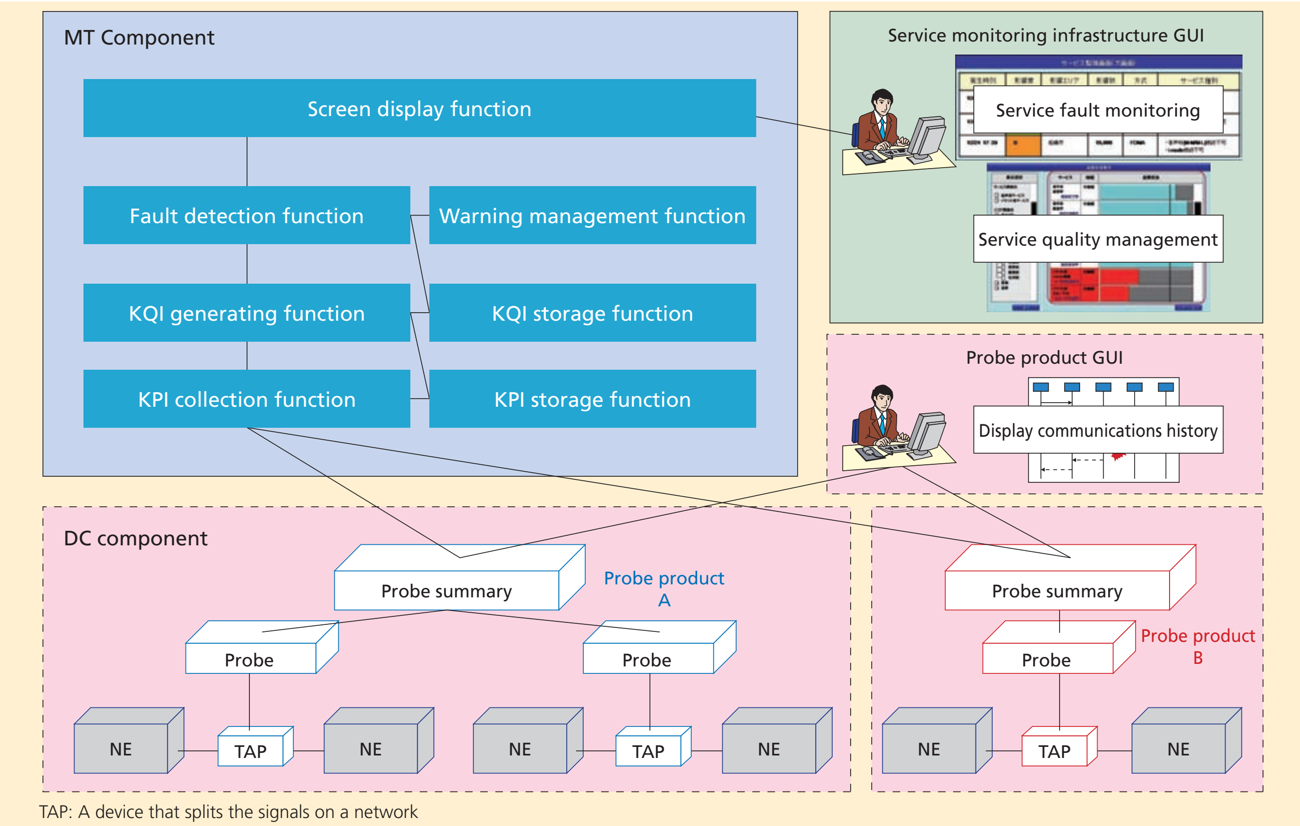

Figure 3 Service monitoring system structure

MT component from the perspectives of performance and operations.

- Real-time performance

  To implement service monitoring, fault decisions must be made in real time based on the accumulated KPI and KQI. To achieve this, a mechanism to reduce the number of searches is needed in order to perform database access and storage frequently and at real time.

- Productivity

  Service quality indices can change daily due to diversifying user needs, so monitoring staff must be able to change these indices (KQI generation, fault detection) in a timely manner. Thus, it is desirable for monitoring staff to be able to customize KQI and fault detection logic themselves, when they need to and without requiring development.

- Extensibility

  The system structure must be easily extensible in order to handle increased storage or processing as more KPI and/or KQI are needed, or the scope of monitoring increases.

- Availability

  A redundant system organization, able to recover quickly without interruptions in operation when a hardware or software fault occurs, is needed. It is also desirable that any effects are localized when a fault occurs.

2) Proposed Solutions to MT Component Issues

- Real-time performance

  Items for which faults are detected need to be grouped (e.g.: by service), and group information assigned to each KPI and registered in a DB. In this way, when a group fault is detected, the applicable KPI/KQI can be identified, reducing the number of extra searches (**Figure 4**).

- Productivity

  To provide an intuitive interface for monitoring staff, enable definition and editing of KQI from KPI formulas and of fault-decision logic, automatically generate SQL[*2] commands from these definitions, and store them in the DB (**Figure 5**). Thus, monitoring staff need only define formulas and decision logic, and can leave the DB access work to automatically-generated SQL commands.

- Extensibility, availability

  Use D3A[*3] infrastructure, which is also used for other operations equipment [2].

  When storage or processing need to be increased because more KPI/KQI are needed, or the scope of monitoring is increased, the system can be extended by adding hardware (IA Servers[*4]) to the affected element (KPI storage DB) using the D3A infrastructure.

  With D3A infrastructure, one of three structures can be selected according to functional characteris-

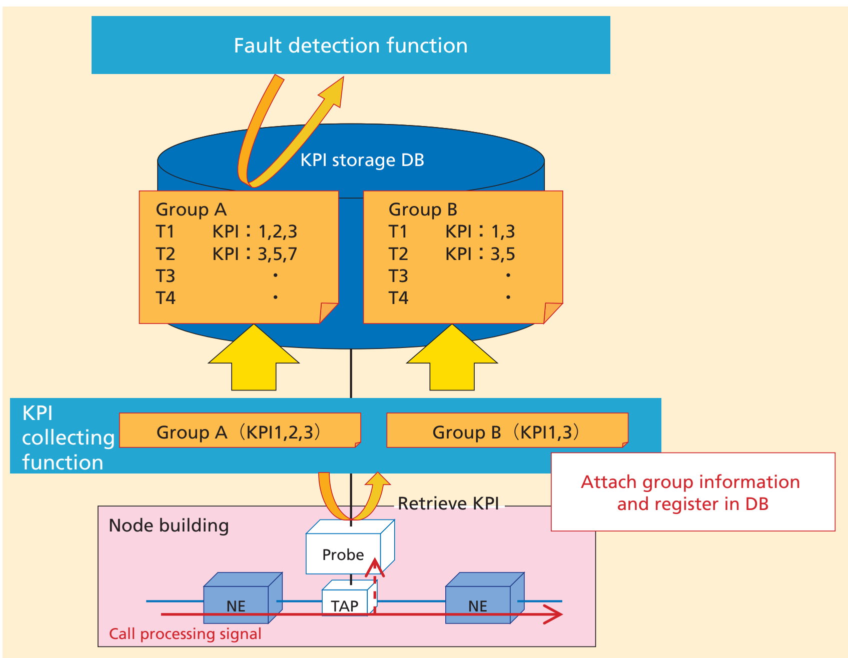

Figure 4 Decreasing load on the DB

*2 **SQL**: A programming language developed by International Business Machines Corp. in the United States for defining and manipulating databases.

*3 **D3A**: An architecture developed at NTT DOCOMO, which groups multiple IA servers to achieve high performance. IA servers use microprocessors from Intel Corp., which are less expensive than servers using microprocessors from other manufacturers. This allows low-cost, high-performance systems to be built easily.

*4 **IA Server**: A server equipped with an Intel microprocessor. Its internal structure is very similar to that of an ordinary PC, and it is less expensive than servers based on other types of microprocessor.

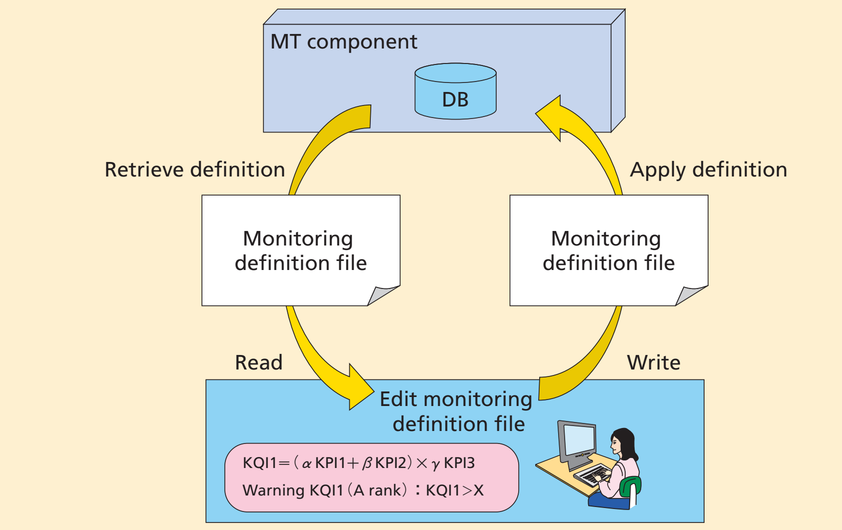

Figure 5 Customization with monitoring definition file

tics: n-ACT structure[*5] (selective execution) n-ACT structure (parallel execution) or n-ACT/m-SBY structure[*6]. Each function is arranged on each IA server, and by preparing them with an appropriate redundancy configuration, the effect of faults can be localized and high availability is achieved.

4 Service Monitoring Operation System Implementation

The structure of the MT component system, using D3A, is shown in **Figure 6**,

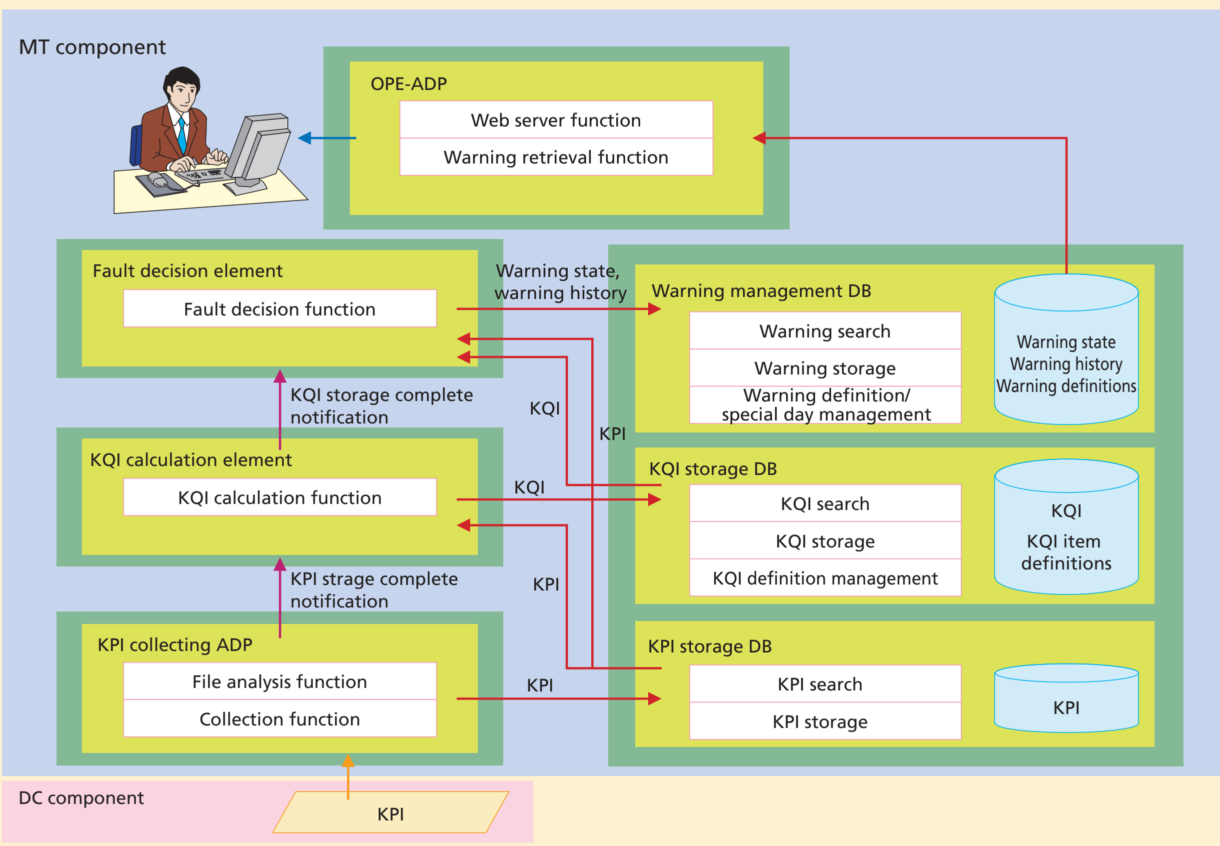

Figure 6 MT Department system structure

*5 **n-ACT structure**: N active servers operate in parallel, distributing the processing load. The other servers can carry on processing if one of the servers experiences a fault.

*6 **n-ACT/m-SBY structure**: N active servers and M standby servers are prepared.

and each part is described below.

- KPI collection ADaPter (KPI collection ADP)

    Collects KPI from the DC component.

    Also provides functionality to convert KPIs to a data structure used within the monitoring component, so differences in the interfaces for retrieving KPI from different probe products are handled in this adapter.

- KQI calculation ELement (KQI calculation EL)

    After receiving notification from the KPI collection ADP that KPI storage is complete, generates the KQI.

- KPI/KQI storage DB

    Functionality to store each of KPI and KQI in a DB.

- Fault detection element

    Functionality to analyze KPI and KQI, and detect faults.

    For fault detection logic, in addition to simple thresholds, various other types of logic can be configured to prevent issuing false warnings, such as comparing with data from the same time a week earlier or other historical data, or composite decisions that take into account issues such as low traffic volume.

- Warning management DB

    Maintains the results of detections by the fault management element (warning information, incident times and status).

- OPEration display ADaPter (OPE-ADP)

    Maintains the screen display function and provides the human interface for system monitoring staff.

5. Conclusion

In this article, we have given an overview of the need for service monitoring and a service operations system.

In order for the service operations system to monitor all users, many probes would be needed, and would incur high cost. Thus, we intend to monitor networks over a restricted area initially, and gradually expand the scope of our service operations system.

REFERENCES

[1] TMF GB917: "SLA Management Handbook, Release 2.5," Jul. 2005.

[2] K. Akiyama et al.: "Technology for Achieving an Economical Operations System—Distributed Data Driven Architecture—," NTT DOCOMO Technical Journal, Vol. 13, No. 2, pp.36-46, Jul. 2005 (in Japanese).