Scalable Root Cause Analysis Assisted by Classified Alarm Information Model Based Algorithm

Masanori Miyazawa and Kosuke Nishimura
KDDI R&D Laboratories, Inc.
2-1-15 Ohara Fujimino City, Saitama Prefecture, JAPAN
ma-miyazawa@kddilabs.jp

Abstract— **One of the important issues for telecom carrier is that the time required to identify the root cause of a failure has increased since the number and types of alarms caused by network or service failures has increased in a fixed-mobile convergence network environment,. To address this issue, this paper proposes a root cause analysis (RCA) mechanism which classifies alarms based on their types of failures, such as resource, performance and service failures, and then promptly identifies the root cause by using a hierarchical alarm information model. Our proposed mechanism which is the implemented into a prototype system was successfully demonstrated in a testbed. Its effectiveness was validated that our RCA mechanism handled 65,000 alarms within 550 seconds in a practical network consisting of 100,000 equipments. The results also show that the algorithm minimizes the overhead of RCA itself to apply large-scale environment, and thus the total RCA performance is limited only by the DB access.**

Keywords; Root cause analysis; fault management; resource management; classified alarm information model

I. INTRODUCTION

As the emergence of new services on top of telecom network is being accelerated by the spread of smartphones and tablet PCs, the environment surrounding telecom carriers such as Fixed Mobile Convergence (FMC) have been drastically changing with the increase of the complexities of such services and technologies. Under such conditions, telecom careers are expected to continue ensuring Service Level Agreements (SLA) including service availability ratio and so forth, and must compensate a customer when the SLA is violated. Thus, the guarantee of the SLA is important and a service provider keeps track of service qualities and identifies the sign of a problem before the occurrence of SLA-violation. However, due to the diversification of services and network technologies, the number and types of alarms generated by network elements (NEs) and/or their management systems have increased and it takes time to identify the root cause of a failure resulting in degradation of service qualities. Furthermore, it comes to a major concern that operators are required to have high levels of knowledge and skills across the fixed network and the mobile network to use the RCA system, while the possibility of human errors would increase. Therefore, it is strongly desired that an RCA is able to work without high knowledge in such complicated network. On the other hand, in order to facilitate efficient operations, Operation Support System (OSS) is already deployed to support all the management functions, such as resource management (RM), and fault management (FM) and so on. However, these functions may have to be newly developed or modified every time when a new equipment or service is introduced and OPErational eXpenditures (OPEX) in terms of the development period is a major problem. Especially, the FM functions with an RCA have to support alarms for a new equipment or service specification. However, most of existing OSS is lack of flexibility for a prompt development for modification.

In recent years, there are two major approaches reported to implement a RCA [1-4]. One is a rule-based approach [1,2] which identifies the root cause based on event correlation rules which are defined in advance. The other is a model-based RCA approach [3,4] which identifies the root cause using the relationship based on network topology information. The rule-based approach has a difficulty to promptly modify itself because the rules contain hard-coded network configuration using high-level policy language. In contrast, the model-based approach is easy to deploy and modify and is appropriate for a large-scale network if the network resource information is available. However, since the proposal basically focused on the network failure [4], no mechanism capable of an analysis based on the network performance and service failure has yet been considered. Under a complicated environment, it requires a consistent RCA mechanism to identify the root cause from a network failure to a service failure from easy operation and deployment point of view.

To address these issues, we describe a RCA mechanism to support several types of failures which is achieved by our proposed two functions; one is a hierarchical alarm information model to classify alarms based on their types of failures, the other is RCA mechanism to identify the root cause using the alarm information model. In addition, our proposed mechanism which is implemented into a prototype OSS was successfully demonstrated in a network testbed. Its effectiveness was validated that our RCA mechanism handled 65,000 alarms in a large-scale network consisting of 100,000 equipments.

II. MANAGEMENT ARCHITECTURE

OSS architecture provides overall management functions including service, resource, fault and service quality management, and this have been formulated by the TeleManagement (TMF) as the Frameworx. The service management (SM) system manages service information using a service information database (DB) defined by the TMF Shared Information and Data modeling (SID) model [5]. The RM system organizes various types of managed objects including

NEs, servers, physical/logical network topology and applications using a resource information DB also defined by the TMF SID model. The service quality management (SQM) system has functions to monitor and maintain service qualities, and to detect the service problems. This system also manages performance information, such as packet loss, throughput and delay as a network performance, memory and CPU utilization as an application performance and calculates a Key Performance Indicators (KPI). The FM system is to handle of all alarms of network/application resource, performance and service quality sent from NEs and other management systems, and then identifies the root cause of a failure by the analysis of all alarms. Our proposed RCA mechanism is designed based on a model-based approach. Although it is necessary to understand network topology and resource information for analysis of a root cause, a FM system is generally lack of resource and service information. Therefore, we enable our architecture to share resource and service information between RM systems using a web service (WS) interface [6].

III. FAULT MANAGEMENT WITH ROOT CAUSE ANALYSIS

A. Proposed alarm information model

As the number of alarm types generated by network resource, performance and service failures is increasing in a future network and service environment, the key function of our proposed RCA is to identify the cause of a failure based on different alarm types, and eventually execute the causal correlation among these types of alarms. To achieve this, we introduce a hierarchical alarm information model which classification and management of the alarms as depicted in Fig. 1. The alarm class is a base class to describe different types of alarms. The main purpose of this class is to manage common identifiers and attributes such as an alarm ID in all alarms. A resource alarm class is a subclass of the alarm class, which stores physical, logical and application alarms by their resource failures, and has two main attributes; one is a resource ID which is stored to identity each resource in the RM system and the other is a root cause flag to identify whether the alarm is a root cause or not. A performance alarm class, which is also a subclass of the alarm class, stores physical, logical and application alarms of the performance failures. This subclass has the same attributes as the physical resource subclass. The service alarm class manages alarms regarding a services failure and has a service ID, which is a unique identifier to distinguish service information in the SM system, and a root cause flag as same as the others. Since this model is designed to be able to separately manage each root cause flag alarms in the types of alarm, it is expected to solve the root cause between resource, performance and service failures, and to identify the original root cause using the proposed RCA algorithm.

B. Proposed RCA algorithm

Figure 2 and 3 shows a flow chart of our proposed RCA algorithm based on a hierarchical alarm information model. This algorithm is mainly divided into three processes; one is an information collection process (Fig. 2) which collects resource or service information related to received alarms, the second one is an alarm classification process which classifies alarms based on the type of an alarm (Fig.3), and the last one is a RCA process which has two functions; an identification of the root cause based on the alarm types defined in a hierarchical alarm information model inside each class and a localization of the original root cause by a causal correlation among these different types of alarms (Fig.3). To clarify these processes, we show an example focusing a physical link down, and an example of instances of alarm DB as shown in Fig. 2.

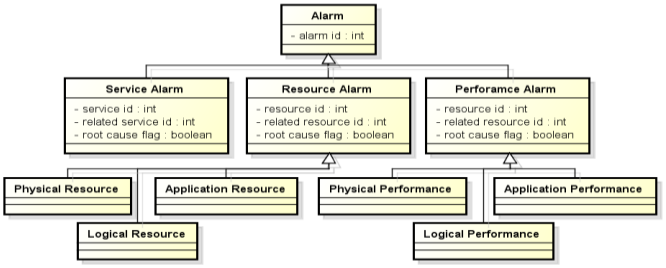

Figure 1. Hierarchical alarm information model.

1) Information collection process

When the FM system receives an alarm message (Fig.2 (1)) from a NE, a server or management system, the information collection process starts querying the RM system for resource information regarding the alarm and the resource name including the alarm is utilized to query via the WS interface [6] (Fig.2 (2)). The RM system then looks the resource information associated to the resource name up in the resource information DB. Resource information contains a resource ID as a unique identifier, and the resource type which represents hierarchical resource information such as hardware, physical layer, logical layer and so forth (Fig.2 (3)). When the resource ID is identified, the FM system restore it and queries the hierarchical resource information to identify the affected resource information, and then also keeps them as related resource IDs which are utilized to analyze the relationship among alarms. For example, if a physical link (PL) down occurs, this failure affects other logical links (LL) on top of it, such as LL_1 down and LL _2 down. On receiving the physical link down alarm, the FM system gets the resource ID of 1 (Fig.2 (3)) and then queries the related resource IDs so as to locate logical links related to the link, and obtains the related resource IDs of 2 and 3 which represent LL_1 and LL_2 (Fig.2 (4)). The system eventually understands the relationship among these resources in this process. Additionally, service information affected by the failure is queried to the SM system.

2) Alarm classification process

In the alarm classification process, the received alarm is categorized based on the hierarchical alarm information model and the classified alarm is registered to an instance of the appropriate subclass of the alarm class. The classification rule depends on the alarm type and is defined in advance.

3) Root cause Analysis process

As for the RCA process, a classified alarm is processed through three different routes based on their types. For example, the physical link down alarm is classified in the previous process (Fig.3 (5)) and stored into an instance of resource alarm class. After that, the FM system extracts alarms with the related resource IDs of 1 which corresponds to the resource ID

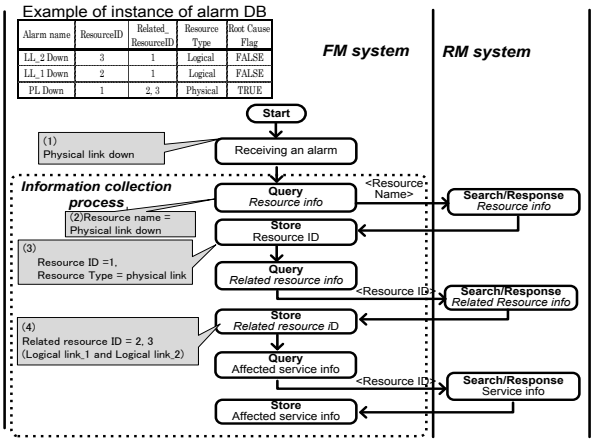

Figure 2. Information collection process

of the physical link down alarm (Fig.3 (6)). In this case, the FM system identifies two failures of LL_1 and LL_2 with the related resource ID of 1, and the failure of physical link is related to these failures. After that, the FM system looks up the resource types of these alarms and resolves the hierarchical relationship among three alarms based on the hierarchy of resource type. Since the resource types of a physical link down and two logical link downs are identified as in a physical layer and a logical layer respectively (Fig.3 (7)), the system identifies the physical link down alarm as the root cause. Then, the root cause flag of the physical link alarm is changed to "TRUE" and the flags of other alarms are kept with "FALSE" (Fig.5 (8)). An RCA process is executed within a pre-determined period so as not to associate alarms with no relationship. This period depends on the propagation time of alarms in a network. The root causes of performance and service failures are also identified by the same process as a resource failure. Finally, to identify the original root cause across different alarm types, the system find out alarms which satisfy both conditions of (1) and (2) from the root cause alarms as identified above.

$$R\_ID_A(\Delta t) = R\_ID_B(\Delta t) \quad (1)$$

$$A\_ID_A(\Delta t) \mathrel{!}= A\_ID_B(\Delta t) \quad (2)$$

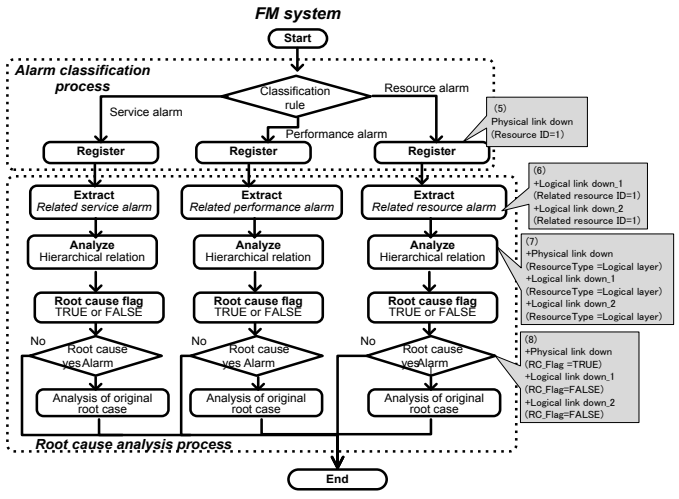

Figure 3. Alarm classification process and RCA process

where R_ID_A and R_ID_B, denote a resource ID with the root cause in a subclass of A and that of B respectively and A_ID_A and A_ID_B denote an alarm ID with the root cause in a subclass of A and that of B respectively. Δt denotes pre-determined period as described above. If processing alarm meets the conditions in other alarms, the system identifies these alarms with the same cause, and then determines the original root which satisfies the condition as follows.

Resource Alarm > Performance Alarm > Service Alarm.

This condition indicates the fundamental relationship among different subclasses and the FM system identifies an alarm as the original root cause of a failure at the end.

IV. PERFOMANCE EXPERIMENTS

A. Experimental configuration

In order to evaluate our proposed RCA mechanism, a network testbed simulating a telecom services was configured as shown in Fig. 4, which assumed three types of services including VoIP, VoD and VPN services. These services and network were managed by our prototype OSS which consisted of SM, RM, SQM and FM systems. In order to manage the network and applications, a network and an application management system were deployed, which collected network and application process information from NEs and servers. This resource information was finally stored to a RM system. The performance management system monitored the network resource performance and the application performance. Additionally, the SQM system collected the network performance data from the performance management system and calculated KPIs at one minute intervals, which indicated the quality of a service, and the calculated KPIs were compared with the pre-determined values to estimate the service availability. The FM system received all the alarms, and has the RCA mechanism. In order to improve an RCA processing, the FM system was divided into two servers; an alarm probe and an alarm DB were implemented into the server1 (CPU: Intel Pentium 1.80GHz, Mem: 2GB) and the RCA function was implemented into the server2(CPU: Intel Xeon 2.80GHz, MEM: 8GB). As the databases of RM and FM systems, XML DB [7] and in-memory DB [8] were used to easily change the resource model and tolerate heavy transactions.

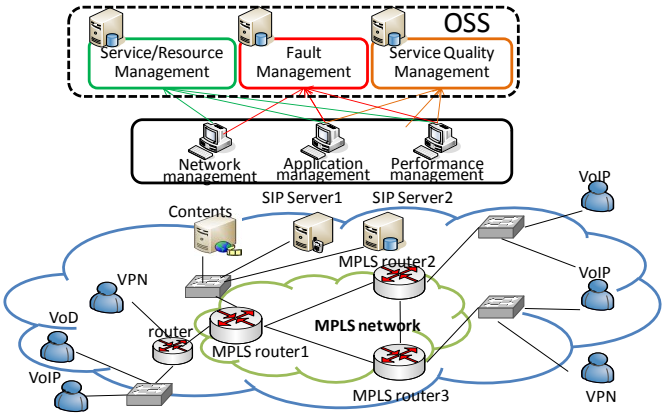

Figure 4. Network testbed

B. *A demonstration of RCA mechanism*

In order to verify the efficiency of our RCA mechanism, a link failure was intentionally generated by disconnecting the fiber between MPLS router1 and router2. This failure caused several logical resource failures and resource performance failures, and then the FM system received approximately 100 SNMP traps. Figure 5 shows a part of alarms received by the FM system and the results of RCA; (a) resource alarms, (b) performance alarms and (c) service alarms. When an alarm was received at an alarm probe, our RCA mechanism started as described in Section III. As a result, two link-down alarms from MPLS router1 and MPLS router2 were eventually identified as the root cause of the resource failure as indicated by highlights in Fig. 5 (a), and the root causes of the failure in performance and service types were identified as shown by highlighted rows in Fig. 5 (b) and (c). Then, the system identified the original root cause of these alarms as follows. When the root cause of performance failures with the alarm ID of 386621 was identified in the performance alarm type, the system looked alarms with the same resource IDs and different alarm IDs up in the resource alarm class and eventually identified the two link down alarms with the alarm IDs of 385680 and 385683 as the original root cause of these alarms.

Alarm name	Node	Root cause flag	Related_Resource_ID	Affected_service_ID	Resource_ID	Alarm_ID
MPLS Tunnel Down	MPLS router2	FALSE	61, 62, 200, 201	311	101	385685
MPLS Tunnel Down	MPLS router2	FALSE	61, 62, 200, 201	311	100	385684
Link Down	MPLS router1	TRUE	100, 101	311	61	385683
Link Down	MPLS router2	TRUE	100, 101	311	62	385680

(a) Resource alarm

Alarm name	Node	Root cause flag	Related_Resource_ID	Affected_service_ID	Resource_ID	Alarm_ID
MPLS performance	MPLS router2	FALSE	61, 62, 200, 201	311	101	386624
MPLS performance	MPLS router2	FALSE	61, 62, 200, 201	311	100	386623
Link performance	MPLS router1	TRUE	100, 101	311	61	386621
Link performance	MPLS router2	TRUE	100, 101	311	62	386619

(b) Performance alarm

Alarm name	Node	Root cause flag	Alarm_ID	Service_ID
VoIP_Jitter	VoIP	FALSE	393476	311
VoIP_PacketLoss	VoIP	FALSE	393356	311
VoIP_SLA_Violation	VoIP_A	FALSE	393856	325
VoIP_SLA_Violation	VoIP_B	FALSE	393834	325

(c) Service alarm

Figure 5. Received alarms and the results of RCA

C. *Scalability evaluations of RCA mechanism*

Considering a management of all the alarms and their RCA in a telecom network, the scalability of an RCA mechanism must be evaluated and verified by measuring the RCA processing time for the number of alarms. In terms of the number of resources managed by the RM system, 100,000 equipments and 10,000 services were assumed and the total number of resources reached 500,000. Regarding the number of alarms, we used an alarm simulator to generate pseudo alarms of 500 alarms/sec, which are equal to the number of alarms after the alarm mask processing. Eventually, the scale of approximately 10,000 alarms/sec could be assumed as total number of alarms including the number before the alarm mask processing. The alarm mask processing was implemented in FM system, and the process does not affect the performance of our RCA. Figure 6(a) shows a relation between the number of alarms and the RCA processing time. In this evaluation, as the number of alarms increased from 100 to 65,000, the RCA processing time almost linearly increased. Our proposed RCA mechanism could analyze 65,000 within approximately 550 seconds and achieved to process approximately 120 alarms/sec.

Furthermore, to analyze the detailed time contribution of an RCA process, information collection processing time was measured. Figure 6(b) shows the relation between the number of alarms and the collection time which includes requests and responses by the web service interfaces and the time to look up in the RM system. A request also caused other requests between the RM and FM systems. When 65,000 alarms were generated, the total number of requests reached 200,000 since our information collection process executed three types of requests, such as resource, related resource and affected service information, per alarm. Since the collection processing time was approximately 550 seconds, the RCA mechanism took the most of time for the information collection process. Although the results also represent that our algorithm minimizes the overhead of RCA itself to apply large-scale environment, the performance of RCA would be able to be improved by a replacement of DB with higher performance.

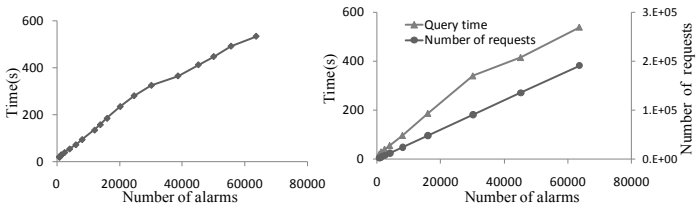

Figure 6. Results of scalability evaluation

V. CONCLUSION

In this paper, we proposed and demonstrated an RCA mechanism which classified types of alarms based on hierarchical information model information. The hierarchical alarm information model allowed operators not only to classify the alarms based on types, but also execute a root cause based on alarm types, and is expected to be extended without a large modification of a RCA algorithm. The results of demonstration and evaluation using a network testbed confirmed that the RCA mechanism identified the original cause from all the alarms and performed in ten minutes even against 65,000 alarms. The results also show that the proposed algorithm minimizes the overhead of RCA itself to apply large-scale environment, and thus the total RCA performance is limited only by the DB access. Therefore, in our future work, performance evaluation and tuning with the selection of appropriate DB is needed to maximize the total performance of the proposed RCA system.

REFERENCES

[1] G.Jakobson and M.D.Weissman. "Alarm Correlation". IEEE Networks, Vol.37, pp.52-59, 1993.

[2] L. Lewis. "A case-based reasoning approach to the resolution of faults incommunications networks." IM1993 , Page(s): 671–681.

[3] S. Kather, et al, “Fault Isolation and Event Correlation for Integrated Fault Management ”, IM 1997 Pages(s): 583-596

[4] K. Appleby, et al, “Yemanja-A Layered Event Correlation Engine for Multi-domain Server Farms” Journal of Network and Systems Management 10 (2) (2002) 171.194.

[5] Shared Information and Data modeling (SID), TM Forum, GB922, November, 2005

[6] Multi-Technology Operation Support Interface (MTOSI) Extensible Markup Language (XML) Solution Set, TM Forum, TMF854, December, 2005.

[7] http://www.modis.ispras.ru/sedna/

[8] http://www.sybase.com/products/databasemanagement/adaptiveserverenterprise