VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

DATABASE REMODELING AND IMPLEMENTING TABLE PARTITIONING AS A SOLUTION TO SMART CABLE GUARD SYSTEM

Humasak Tommy Argo Simanjuntak[1], Paul Wagenaars[2] and Guido Bakema[3]
[1]Department of Information System, Faculty of Electronics and Informatic Engineering, Institut Teknologi Del, Sumatra, Indonesia
[2]DNV GL, Arnhem, the Netherlands
[3]HAN University of Applied Sciences, Arnhem and Nijmegen, the Netherlands
E-Mail: humasak@gmail.com

ABSTRACT

One of the services that DNV GL Netherlands provides to its customers is to continuously monitor distribution cable circuits for upcoming defects. This system monitors over 100 circuits with an average of 5 discharges per circuit per minute. Regardless of the amount of circuits and pulse discharges per minute, the system should be able to complete all processes within one hour or even less. Meanwhile, the current database design using SQL Server cannot adapt to the new requirement of the user and low query performance for inserting and retrieving Partial Discharges (PD) data. Thus, to meet the new requirements, database remodeling should be done. To increase query performance for inserting and retrieving PD data, a new design to store PD data were introduced with remodel database structure and implementing partitioning.

Keywords: partial discharges, low query performance, remodeling database, partitioning.

INTRODUCTION

Partial discharges (PD) is a localised dielectric breakdown of a small portion of a solid or fluid electrical insulation system under high voltage stress, which does not bridge the space between two conductors. Nowadays, many companies focus on electrical system (example: electricity or power company) build an online system to monitor partial discharges. PD online system is an effective predictive maintenance for generators at 4160 volt and above, as well as other electrical distribution equipment. The systems do equipment analysis and dignostics during normal production, so corrective action can be planned to reduce unscheduled downtime (Paoletti and Golubev, 1999).

A new, unique and innovative measuring system for the on-line acquisition and analysis of partial discharges (PDs) in medium-voltage power cables is developed by DNV GL (formerly KEMA), together with Eindhoven University of Technology, with support from the Technology Foundations of the Dutch government STW, and the Dutch Utilities Alliander, Stedin and Enexis.

The system which is called Smart Cable Guard (SCG) (formerly Partial Discharge Online Sysytem), checks the 'heartbeat' of the medium voltage (MV) cables continuously by monitoring and locating partial discharge activity. At both sides of the cable, a Measurement Unit (MU), which consists of a control unit and one or more sensors are installed to measure the complete MV cable connection on partial discharge activity during daily use.

A MU measures and stores one hour-PD pulses locally. After collecting one hour-PD pulses, each MU in one cable circuit will send the measurements recorded in two pulse files to the server in DNV GL Control Centre by GPRS or LAN. In the Control Centre, one computer collects data from all SCG measuring systems. The data will be analyzed and interpreted for final presentation and interpretation, as depicted in Figure-1. Each pulse file will be processed by the Data Combiner and stored into the SCG database.

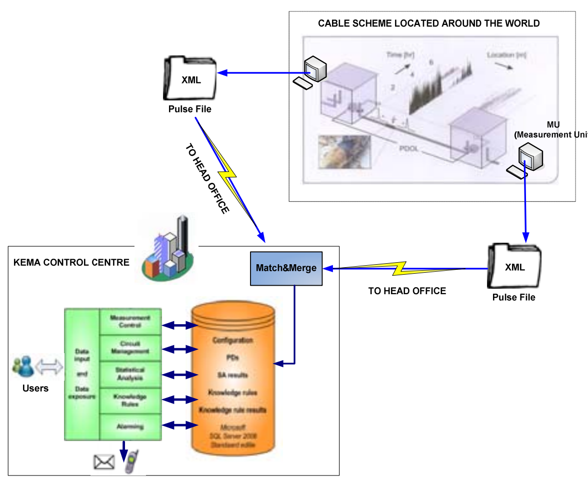

Figure-1. SCG system.

The descriptions in detail for each subsystem in Figure-1 are as follows:

1. Match and Merge is a stand-alone application that combines pulses into PDs. The Match and Merge application monitors (polls) a configurable location at defined intervals and detects new pulse
2. Data input and Data exposure, which is responsible for interacting between the users' web browsers and the core components of the SCG system.
3. Database, runs on a dedicated MS. SQL Server 2008 environment. The database contains all SCG persistent data. Logging and tracing data is not stored in the database but in files.

SCG system monitors the condition of the cable circuits periodically by locating partial discharge activity.

Partial discharge data is recorded in pulse files, which are sent by each SCG system to the control centre at DNV GL. Later, all pulse files will be processed and stored in the database, which is located in the head office.

At this moment, the SCG system monitors over 100 circuits with an average of 5 discharges per circuit per minute. These discharges are stored in the relational database MS. SQL Server 2008. Because the SCG system is relatively new, the number of customers is expected to increase in the future. In the future, the SCG system should be able to handle 5000 circuits with the maximum of 1000 pulse discharges per minute. In fact, the time consumption to store partial discharges for 100 circuits with 1000 pulse discharges per minute is 67 minutes. However, regardless of the amount of circuit and pulse discharge per minute, the system should be able to complete all processes within one hour or even less. Meanwhile, the current database design cannot adapt to the new requirement of the user that is needed to be fulfilled as soon as possible to support their decision making process.

RESEARCH METHOD

This research aims is mainly to find solution to help the company to improve SCG System. Research adapt a software development model methodology which is called Scrum. This methodology emphasizes in communication and collaboration, functioning software, and the flexibility to adapt to emerging business realities.

Scrum is an agile approach to software development. Rather than a full process or methodology, it is a framework. So instead of providing complete, detailed descriptions of how everything is to be done on the project, much is left up to the team. This is done because the team will know best how to solve its problem (Mountain Great Software, 2010). Scrum is unique because it introduced the idea of “empirical process control.” That is, Scrum uses the real-world progress of a project - not a best guess or uninformed forecast - to plan and schedule releases (Danube, 2010).

Currently, Scrum methodology is successfully implemented for completing the project by adapting all rules and conditions in Scrum. As stated before, Scrum projects make progress in a series of sprints, which are time boxed iterations no more than a month long.

At the start of a sprint, team members commit to delivering some number of features that were listed on the project’s product backlog. At the end of the sprint, these features are done--they are coded, tested, and integrated into the evolving product or system. At the end of the sprint a sprint review is conducted during which the team demonstrates the new functionality to the product owner and other interested stakeholders who provide feedback that could influence the next sprint.

Overall, research was conducted by the following steps:

1. Conduct Literature study in order to learn and understand current SCG system.
2. Analyze the new requirement of the SCG system according to the number of measuring system and component of the circuit
3. Remodeling SCG database based on the analysis in point 2
4. Implementing Partitioning on the table.
5. In the end of the research, researchers did benchmarking test for insertion and selection time to prove the experiment result and give research conclusion.

ANALYSIS

In this section, the result of analysis and result finding are explained.

Analysis on new Smart Cable Guard System

Based on the analysis there are several business requirements that need to be fulfilled in the new database design. The requirements are related to:

a) Number of measuring system which installed on the circuit which accommodating possible circuit-core configuration. According to the SCG business process, there are four scenarios that may be occurred related to the number of measuring system on the one circuit. The four scenarios are:

a. One measuring system on three single-core cables or three cores cable. The component configurations for each core are guaranteed to be the same. One circuit can be consists of three single-core cables or three cores cable. The core of circuit is composed by many component that can be a type of cable, joint, rmu, termination, etc. In this scenario, there is one measuring system on three single-core cables or three cores cable. The component configuration for each core is the same means every core composed by some type of component on the different locations. The configuration for other core is copy from the other core. Figure-2 show the example of one circuit with three single-core cables or and Figure-3 show the example of three cores cable which installed one measuring system.

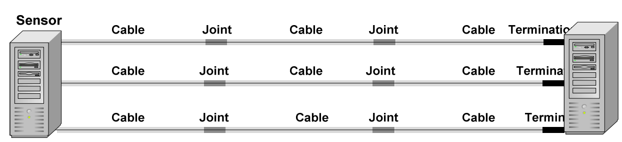

Figure-2. One measuring system on one circuit three single-core cables.

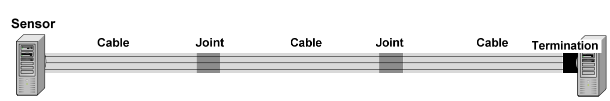

Figure-3. One measuring system on one circuit three cores cable.

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

b. Three measuring system on three single-core cables. The component configurations for each core are guaranteed to be the same. One circuit with three single-core cables is installed three measuring system (each core installed one measuring system). The configurations component for each core is the same means every core composed by some type of component on the different locations. The configuration for other component is copy from the other. Figure-4 show the example of one circuit with three single-core cables which installed three measuring system.

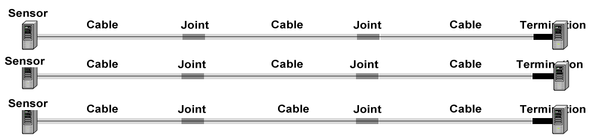

Figure-4. Three measuring systems on one circuit three single-core cables (core configuration same).

c. Three measuring system on three single-core cables. The component configurations for each core are different. One circuit with three single-core cables is installed three measuring system (each core installed measuring system). The configurations component for each core is different. Every core composed by different type of component in the different location. So, the configuration for other component is not copy from the other. Figure-5 show the example of one circuit with three single-core cables which is installed three measuring system.

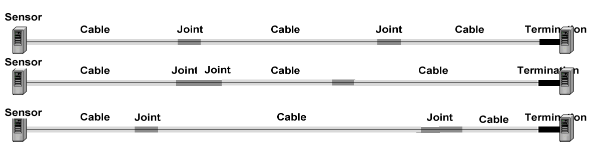

Figure-5. Three measuring systems on one circuit three single-core cables (core configuration different).

d. One measuring system on three single-core cables. The component configurations for each core are different. One circuit with 3 single-core cables is installed one measuring system. The configurations component for each core is different (same with scenario 3). Every core composed by different type of component in the different location. Therefore, the configuration for other component is not copy from the other. Figure-6 show the example of one circuit with 3 single-core cables which installed three measuring system.

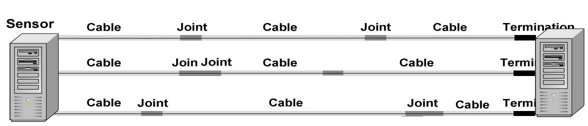

Figure-6. One measuring systems on one circuit three single-core cables (core configuration different).

If a circuit with three single-core cables has the same configuration for each core, then there is a requirement that the component configuration is stored only once. That means, there is no redundant data for this condition.

Instead of number of measuring system in one circuit, there is a possibility the measuring system is not active. The status of measuring system is not active means the measuring system is not working or still repaired from crash. The status of measuring system on the circuit must be stored in the database.

b) Type of the core component in the circuit

Every core in the circuit is composed by many components. The Component can be as component type or component group. One component type is a member of one or more component group and one component group consist of more than one component type.

Component type has information that support DNV GL analyst to know the location of discharges and generate knowledge rule according to the specific component. Example: PILC cable has length. Normally, not all components have added information. Therefore, it is possible for other component to store null value in the table if the components do not have added information.

Some business requirement on SCG system related to the core component which is composed by component type and component group are:

a. All components are exactly known include the length of component. This business requirement is normally occurred in one circuit. One circuit core composed of many components, each component is known, and the information about component is known. Figure-7 show the example of core circuit with all components are known.

Figure-7. All core components are known exactly.

b. Some or all components are only known as a component group. Every specific knowledge rule in this group must be applied. It is possible, a component in one core is only known by the group. There is no information about specific type of component. DNV GL analyst should apply specific knowledge rule in this group. Figure-8 shows the example of circuit core with some components are known by group.

Figure-8. Some or all core component are only known as group.

c. The component group and length is known. One component can be a member of many component groups. Example: in SCG system, there is a group cable that consists of XLPE cable and PILC cable. Another group is PILC circuit that consists of PILC

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

joint and PILC cable. Figure-9 show the example of circuit core with component group and length is known.

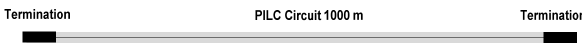

Figure-9. Only known is group and length of core component.

d. Unknown components but with joints at known locations. One circuit core can consist of unknown component. The only known component is joint and its location. Figure-10 show the example of circuit core with unknown component but with join at known locations.

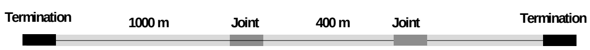

Figure-10. Unknown core components but with joint at known location.

c) One Utility can be in more than one region

Utility is the company which has circuit. A utility can be located in more than one region, but not all utility has data about region. Every utility has contact person who notified by PD-OL system if there is something happen with the circuit. One or more contact person is available in one utility. It is also possible utility which is located in more than one region may have one contact person. A contact person is allowed to access PD-OL system to monitor what the current status of the circuit.

Problem in the current database

Based on the analysis on the current PD-OL database design (designed by Humiq), there are some problems which occurred regarding to the business requirement. The problems are:

1) Number of measuring system on the circuit

a. Data about three measuring system which installed for each core in one circuit stored in the current PD-OL database design with modifies the name of the circuit. Every circuit is only having one core. Therefore, if there are more than one measuring system on the circuit, it seems like every core is one circuit, and this is not true according to the business process. Figure-11 show how the circuit data stored in current PD-OL database

GUID	CIRCUIT_NAME	CORE_ID
8C6FBFB0-39D3-DA9B-88FD-E9E0A7628C63	ZZ_OFF Circuit 1	1
8DA8686A-17C5-4456-BFAA-2CA62F163EC6	ZZ_OFF Circuit 2	1
90B43CE8-87C7-7B78-EE2B-2A8A03AA5604	Circuit 3 Phase 2	1
947659E7-0899-401A-8E95-5C6FD0E36F4F	Circuit 4 Phase 2	1
94AD567D-EB46-4C98-B475-4E991BECCA61	Circuit 5 Red Phase	1
94DD5A58-7B5D-4F98-82AD-6BAB150D00F1	Circuit 6 Blue Phase	1
95BD4157-FD16-AFE0-0BE0-6CA8BA1932AE	Circuit 7 Yellow Phase	1

Figure-11. Each core has GUID in circuit table.

b. One measuring system on three single-core cables or three cores cable. The component configurations for each core are guaranteed to be the same. For each core in one circuit, data about configuration component stored three times (three rows) in current SCG database. It means, there will be data redundancy on core component table, because configuration for each core is same. According to the requirement from DNV GL Analyst, if component configuration in three single-core cable guaranteed to be the same, then data is stored only once in table.

c. Three measuring system on three single-core cables. The component configurations for each core are guaranteed to be the same. For each core in one circuit, component configuration stored three times (three rows) in current SCG database. There is data redundancy on core component table, because configuration for each core is same.

d. The status of measuring system which installed on the circuit is stored by adding the string "ZZ_OFF" to the Circuit Name. If there is a circuit with prefix "ZZ_OFF", that means the measuring system is not active anymore.

2) Type of the core component in the circuit

Based on the requirement related to the type of the Core Component in the circuit, current PD-OL database can handle the all components are known exactly include the length of component? If the other requirement occurred, then current PD-OL database cannot handle this business requirement. Figure-12 shows the current relation of Core_Component, Components, Component_Type and Component_Name.

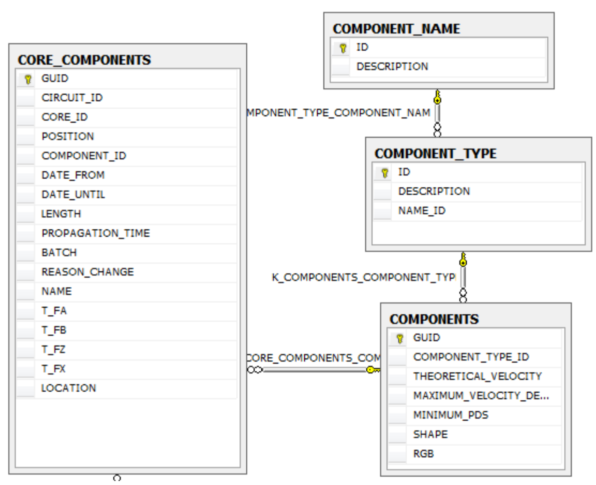

Figure-12. Database diagram for core component.

One problem that cannot be handled by the current SCG database design is if some or all core components are only known by the group, such as cable or joint. The group of component (cable or joint) is stored in component_name table, and there is no relation between core_component and component_name. Therefore, this condition is not possible in current SCG database.

3) Session table is a history table which stores all configuration information about sensor and measurement mode which installed in one circuit.

According to the business logic and requirement from DNV GL analyst, this condition is not good logic because on the specific time, one or more partial discharges comes from one or more measuring system which is installed in the core of the circuit. DNV GL analysts also want the database design gives flexibility to the DNV GL analyst to access measuring system which installed in one circuit. Figure-13 shows the relation between Session, SI_Result, Partial_Discharge_Records, and Partial_Discharges table in current SCG database.

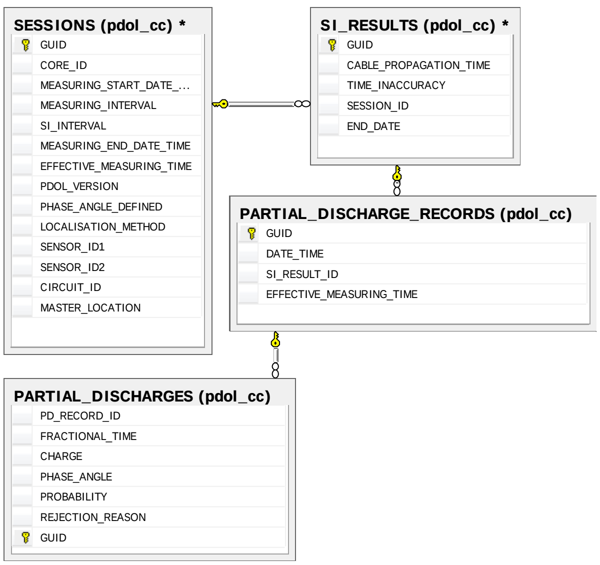

Figure-13. Database diagram for partial discharges table.

4) Every specific knowledge rule must be applied to the type of the core component. The Current SCG database is not stored the data about knowledge rule and the SCurve that related to the specific knowledge rule.

5) The current SCG database is not stored information about utility that is located in more than one region, but it has stored information about contact person in one utility. One utility has one or many contact person. Furthermore, a contact person need user login to access the application. Information about user login is not stored in the contact table.

Partitioning on PD table to improve query performance

When database table grows in size to the tens of gigabytes or more, it more difficult to load data and to maintain the data. Increasing of table size causes SELECT, INSERT, UPDATE, and DELETE operations take much longer time. Even the data that must be loaded or removed can be very sizable, making INSERT and DELETE operations on the table impractical.

The expected number of records in the PD table in the future is:

5000 circuit * 1000 pulse * 60 minutes * 24 hour = 7.200.000.000 records per day

Furthermore, assuming a size of 16 bytes for a single discharge, the database volume is approximately 38 Tbyte per year for storing discharges. This large amount of data gives impact significantly to the performance of SCG system.

Therefore, there is a need to find solution, how to improve the performance of database, especially processing data on PD table. The solution is not only improving the performance of database, but includes how to manage and organize large amount of data in the table.

Based on analysis on SCG database, current transaction table that mostly accessed by SCG system is Partial Discharges (PD) table. This table contained large amount of rows and there are many transaction for every hour. It has size to tens of gigabytes or more. This table accessed frequently by DNV GL Analyst to plot partial discharges data to the chart on SCG application and to calculate Statistical Analysis data. Some conditions that must be considered in PD table are:

a) Loading new data to PD table is per hour. Now, number transaction on PD table is approximately:
100 circuit * 5 pulse * 60 minute = 30.000 rows per hour.

Time that is needed for inserting data 100 circuit is still possible in one hour. In the future, number transaction on PD table is approximately 300 million rows per hour. The previous research shows time needed for inserting PD data is more than 1 hour. This is not feasible, because SCG system must insert PD data from all circuit in one hour to avoid many data queues in the system. Process which takes a long time when inserting PD data to the table is rebuilding data and index on PD table. Every inserting process must rebuild index in the table. Therefore, data partitioning is needed to implement on PD table to improve query performance (Microsoft, 2014). Therefore, rebuilding data and index on PD table only applied to specific partition, not to all data in PD table.

b) All PD data from previous year are stored in one PD table. There is no history table and data archiving in PD table or SCG system. Therefore, every inserting data will rebuild index for all partial data in PD table. This process will take a long time to rebuild index from millions rows data.

c) Most of query that used to access data on Partial Discharges table based on specific circuit and specific range of time (normally per week).

Example

SELECT * FROM [PARTIAL_DISCHARGES]
WHERE CIRCUIT_ID=1 AND StartDateTime
BETWEEN '2010-06-25 00:00:00' AND '2010-06-30 23:59:00'

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Based on this query, the database engine evaluate index for all data. If PD table contains many rows data from many years, it will take a long time to evaluate the data, even index implemented on PD table. To reduce time of database engine for evaluating all index for PD table, partitioning can be implemented based on the query. Amount of data that accessed by DNV GL analyst is based on range of specific time. So, partitioning can be implemented to eliminate other rows that is not needed in the query (at least reduce to scan data or index)

Table partitioning is a data organization scheme in which table data is divided across multiple storage objects called data partitions or ranges according to values in one or more table columns (Tripp, 2005). It is a feature designed to improve the performance of queries made against a very large table. It works by having more than one subset of data for the same table. All the rows are not directly stored in the table, but they are distributed in different partitions of this table. When user queries the table looking for data in a single partition, then due to the presence of these different subsets, user should receive a quicker response from the server

According to the (OSIsoft, 1995), there are two types of partitioning, Horizontal partitioning and Vertical Partitioning. Based on analysis, the method that will be chosen to improve query performance on PD table is horizontal partitioning because Each transaction (insert operation) per hour increase number of rows significantly, Data in PD table is static data, and Accessing data on PD table need most of the column that appear in PD table.

Proposed partitioning on PD table is applied on multiple FileGroup. Before creating partition on PD table, FileGroup added to the database. Number of FileGroup that is added to database is four. Each FileGroup has two physical files (.ndf) to store the data.

The proposed structure of the database for partitioned table after adding 4 FileGroup can be seen in Figure-14.

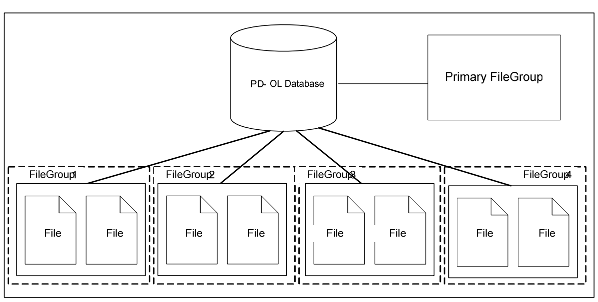

Figure-14. Proposed SCG database with adding four FileGroups.

The Partition boundary values that applied to the partial discharges table is based on month. Therefore, number of partition is 12 partitions. Figure-15 show the general structure of PD table after partition is applied.

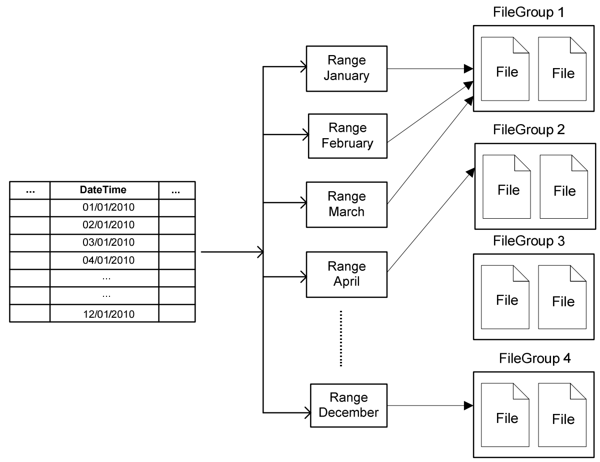

Figure-15. Partial discharges partitioned table.

Based on partition boundary value that applied to the PD table, every partition store data for every month. If there is insertion process in the specific time, then the data will be inserted to the one partition according to the month in PDDateTime column. Example: If there is data with PDDateTime='01/01/2010', then this data stored in Partition 1 (January) immediately. Rebuilding index only applied to the Partition 1. Therefore, insertion process will be faster, because rebuilding index only applied to the data in the partition that involve in insertion process.

Implementing partition on PD table improve performance of properly filtered queries using partition elimination (OSIsoft, 1995). Query processor can eliminate inapplicable partitions from the query plan, and just access the partitions implied by the filter on the queries.

Example: Number of data in PD table is 1200 rows. Every partition has 100 rows. If where clause query specify PDDateTime='01/01/2010', then database engine only access data in partition 1. Database engine eliminated data in other partition. In the other word, retrieving data from partition eliminate 1100 rows in other partition and only scan data from partition 1.

Proposed SCG database

Based on the current database and analysis on SCG system, a new database design is modelled and suggested to answer the problem in the current database

The new database design that answer the requirement and followed the best design is described as follows:

A) Introduce system under test table

Introduce one table that called SystemUnderTest. System under Test means a system that tested by SCG system with one measuring system is installed. The Primary Key of this table is SystemUnderTestID. One SystemUnderTestID means one measuring system installed. One circuit with three single core cables can be installed with three measuring system, that means each

ISSN 1819-6608

core has one measuring system (each core in the circuit has one SystemUnderTestId). It also possible one measuring system installed in the all three core of circuit. Figure-16 show the new relation SystemUnderTest table with other tables.

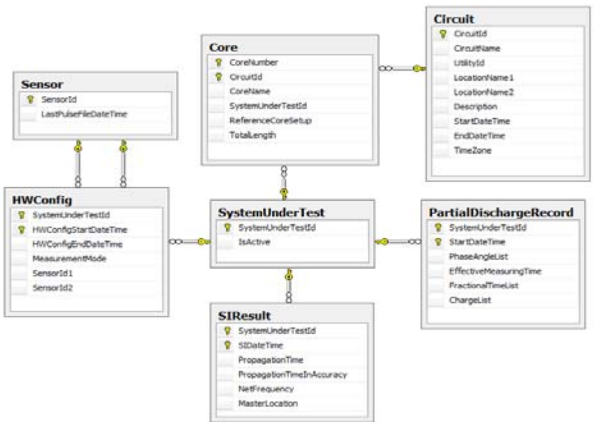

Figure-16. Proposed systemundertest table related with other tables.

Some considerations regarding to the SystemUnderTest table:

a. The status of system that tested by SCG system with one measuring system is stored in SystemUnderTest table. Therefore, it is not needed to modify the name of the circuit to change the status of the system.
b. According to the SCG business requirement, System under test table gives better logic to the SCG database design, because the PartialDischargeRecord table connected to the System Undet test, that means one or more partial discharges comes from one measuring system which is installed on core of the circuit, not from the history table or configuration table (Session table).
c. SystemUnderTest table connected to the HWConfig table. HWConfig table stored the history of hardware configuration (sensor) that installed on core of circuit. SystemUnderTest gives flexibility to the SCG system to configure the hardware that may connect to the system that is tested.

B) Core component

When a circuit installed one or three measuring system and the component configuration are guaranteed to be the same, CoreSetup table only store one component configuration. Adding SystemUndertest table gives possibility that each core in the circuit can be installed one measuring system. If the component configuration for each core is same with the other core in the same circuit, then attribute ReferenceCoreSetup stores value about reference core. Therefore, the same component configurations in one circuit are stored only once and solve the problem data redundancy in table CoreSetup.

The value of this column can be null if the configuration component for each core is different.

C) Type of the component in the circuit

Based on the business requirement, the core component can be composed by the component group or component type. Therefore, proposed SCG database design store all type and group of component in the component type table. Value in the component group table is a subset of the component type table. Every component type has drawing shape to describe specific component which is installed in one circuit. The drawing shape is used to visualize the specific component in the SCG application.

Specific knowledge rule is applied to the component type. One knowledge rule can be applied to the more than one component type, and it is possible different component type has same knowledge rule. Each knowledge rule which is applied to the component type is generated by some parameter that is stored in SCurve table. Figure-17 shows relation ComponenType table with the other table.

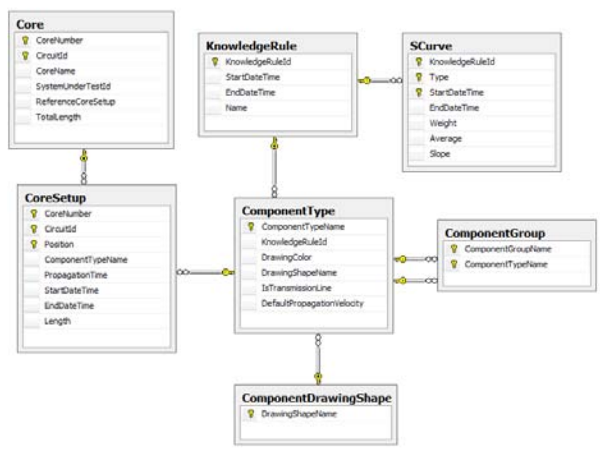

Figure-17. Proposed relation component type table with other tables.

D) Contact person in utility

Based on the business requirement, there is a possibility one utility can be located in more than one region. The proposed SCG database added region attribute to the utility table. One utility has one or many contact person. One person may as contact person for every utility which located in more than one region. If there are many contact people for one utility, then there is priority for each person.

If there is a problem occurred in the circuit, SCG system will send notification to the contact person of utility. Each contact person has one or many notification type and there is a priority for each notification. Figure-16 show relation between Contact table with Utility table.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

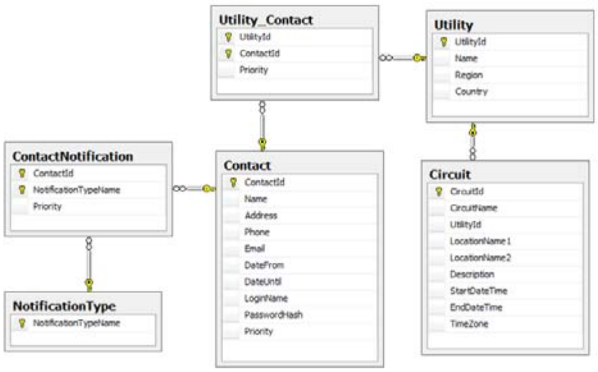

Figure-18. Proposed relation between contact table and utility table.

BENCHMARKING TEST

Benchmarking test did to measure the inserting and retrieving time on Partial Discharges table that has been partitioned compared with
Benchmarking test was done on two operations:

- **a. Selection operation:** Retrieve data for one circuit on specific range of time. Data that retrieved for this operation is data in one week (9300 rows).
- **b. Insertion operation:** Inserting data for 500 circuits, 60 measurements in an hour (partial discharges come every minute within an hour). Number of partial discharges is 1000 partial discharges in one minute. Therefore, numbers of rows inserted are 30000 rows for one hour.

Benchmarking result for partition was compared with the previous structure of PD table. Table-1 show the result of benchmarking test.

Table-1. Benchmarking test result.

Operation	**Previous PD table**			**Proposed PD table with partition**		
	Test 1	Test 2	Avg	Test 1	Test 2	Avg
Insertion	51.172	50.472	50.822	23.407	24.033	23.720
Selection	41.395	40.395	40.895	36.882	36.101	36.491

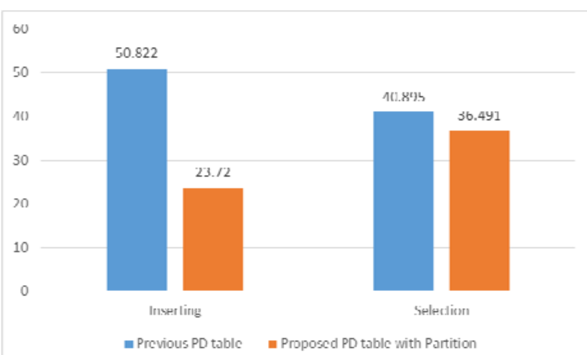

Figure-19. Benchmarking test result.

Table-1 and Figure-19 show the result of insertion and selection benchmarking test. Proposed partitioned PD table offered better performance for inserting and selecting data. Proposed partitioned PD over two times faster (27 second) than previous PD table for insertion and almost 4 second faster for selection.

CONCLUSIONS AND SUGGESTIONS

Poor system design can give huge affect to the system performance if there is a new requirement. Therefore, the system design should considerable predict the future requirement. So, it is need small changes if there is new requirement.

One main component in the system is a data store or database model. The database model must have a best design for business process and performance, especially in electronic system, such as Smart Cable Guard (SCG) System. The database design for SCG should considerable accommodate if there is a changes in cable configuration so data can be stored properly and performance of system still good.

Based on benchmarking test result, implementing partition on PD table improve query performance for inserting and retrieval operation. Partition elimination applied to the table when inserting and retrieval data, so it's reduce number of index and data that needs to rebuild (for inserting) and reduce number of rows that need to evaluate for selection.

Partitioning not only improve query performance, but it make easier for data archival and migration. If in the future, PD data need to maintain, implementing Partitioning is better to performed.
The list of suggestions for improvement of implementing partition on PD table is listed below:

1. Time for inserting and selection is not increase significantly. Therefore, it is better to implement partitioning with or without indexing in multiple file group on different hard disk to improve query performance significantly.
2. Proposed partitioning was implemented based on month. To Increase the performance of partitioning, need to implement partition on PD table based on week.

3. Every one hour, SA-1 Analysis is generated for one hour partial discharges. Because data of SA-1 related to the partial discharges on specific partition, then implementing partitioning on SA-1 table based on partitioned PD table can improve query performance for generating SA-1 analysis.
4. The number of data in partition is approximately 55 GB. Implementing partition benchmarking test on Tera Byte data need to be performed in order to get good benchmarking result.

REFERENCES

Sibarani E., Simanjuntak H., Lumbantoruan R. 2010. A Solution to Partial Discharges Online (PD-OL) System in KEMA. HAN University of Applied Sciences.

Paoletti G., Golubev A. 1999. Partial Discharge Theory and Applications to Electrical Systems. 1999 IAS Pulp and Paper Industry Conference in Seattle. IEEE.

Mountain Great Software. 2010. Introduction to Scrum - an Agile Process. http://www.mountaingoatsoftware.com/topics/scrum.

Danube. 2010. Scrum Methodology. http://scrummethodology.com/.

Microsoft. (2014). Partitioned Tables and Indexes. Microsoft MSDN Library

Tripp K. L. 2005. Partitioned Tables and Indexes in SQL Server 2005. Microsoft TechNet Library.

Nichter Z. 2005. How Using GUIDs in SQL Server Affect Index Performance. www.sql-server-performance.com.

OSIsoft. 1995. PI System. http://www.osisoft.com/software-support/what-is-pi/what_is_pi_.aspx.