PREGEL-large scale graph processing

Name: Anudeep vattikuti,

Date: 05/09/2014.

Bibilographic Data:

http://www.labouseur.com/courses/db/papers/malewicz.sigmod10.pregel.pdf

http://googleresearch.blogspot.com/2009/06/large-scale-graph-computing-at-google.html

http://horicky.blogspot.com/2010/07/google-pregel-graph-processing.html

Basic Idea:

To overcome problems with the algorithms and faults, Google invented the PREGEL .By using this PREGEL can solve the problems and give the desired outputs.

The main idea of the PREGEL is for fault tolerant of graphs, people can use this method by using c++ API. When compared to the remaining outputs it give good results.

The high level language is used by the PREGEL. The large graphs include the social networks like Facebook , twitter. And used in many web graphs.

The PREGEL overcome the many problems over the networks and graphs and it is one type of best graph processing invented by Google.

And the PREGEL when compared to the other parallel graph system, it give the less fault tolerant. And the main idea is to modify the graphs and used for the large classification of graphs.

The results after processing the large scale graph using the PREGEL is easy to understand and impressive. It includes all types of activities where the graph has nodes, edges, vertices.

The Implementation of Idea:

The implementation of the PREGEL is similar to the previous graphs, and it uses the multiple processing by dividing the network graph into the sub parallel graph .And each divided sub graph have large number of nodes. The execution of PREGEL is similar to the bulk synchronous processing.

PREGEL is used to reduce the usage of bandwidth, it is non scalable have lot of vertices. Each graph contains number of incoming links and outputs using the nodes and vertices. At first the graph considered from network or web is divided into parallel groups using nodes and it has multiple partitions.

After this all the multiple partitions are processed in order of supersteps. Inside the supersteps include the information received from the previous steps and send the information to processing units and the cycle repeats until the output occurs.

The PREGEL algorithms not include the deadlocks, the faults in the network graph are minimized using the mechanism of storage memory in the graph periodically for each iteration if it shows any bad content the process stops working and again restart it until the output obtained without any errors.

The PREGEL is the very effective graph processing when compared to the previous graphs. The graphs are divided into multiple network including nodes, vertices. We can say in one line as the PREGEL is one type of graph parallel processing used for networks or web graphs.

Analysis:

My analysis for the PREGEL idea and implementation is very effective and impressive, when compared to the previous graphs. The output from the processing is used for the large scale processing networks or web designs.

In past days the graphs used for networking having faults and outputs are not effective, but at present storing entire graph at standard place in not possible, we must divide the graph into parallel parts.

Here in the entire execution of PREGEL graph analysis is very simple for large graph outputs. In the multiple processing is best results are obtained because it using the multiple processing. I suggest that one can use the PREGEL processing if we have large graph operations.

The operations performed in the PREGEL are based on bulk processing model. It is operated in the parallel sequence of supersteps. Overall the analysis of the entire processing model it is the best method.

Finally in my analysis the PREGEL is one of the best processing methods for large graph processing methods including networks and web models.

For accurate results, fault less tolerance, impressive outputs we can use PREGEL processing.

Comparison to the Idea and Implementation:

In the comparison of the database and PREGEL, they are entirely different. Here the database management system is a relational database, where PREGEL is a graph database.

In graph database are faster for data sets and these are structure oriented, they not requires the joint operations where these are performed in the relational database.

In the relational database is designed by using the joint operations, where in the graph database use nodes, vertices. In DBMS having no graphs, no parallel operations required.

The PREGEL is graph processing used for high level networks, web graphs. In the model graph is divided in sub graphs that are performed in parallel operations.

The DBMS is relational database uses joint operations and commands. It is designed for the small database systems. For operations large processing it performs faults, not impressive.

When comparing both databases, PREGEL graph database obtaining the outputs are more reliable and accuracy for high level processing. For the small database design relational database is best option moreover graph database.

Advantages and Disadvantages:

The graph database having more advantages when compared to the relational database system, graph database the outputs obtained are more accurate and reliable. Relational database is designed using joint operations.

In the relational database system is uses sub queries, joint operations, while using the joint operations is the main disadvantage for the relational database. While in the graph database nodes, vertices are used for parallel operations.

In the graph database operation the graph is divided into parallel operations and stored in different places to increase the speed of operations. And for relational database is not used for large network or web operations.

Both the databases having the respective advantages and disadvantages depending upon the processing operations, graph database having more advantages than the relational database.

The major disadvantage of relational database is when performs large queries, output is time taken and not performed without fault tolerance. In the graph database output obtained are accurate and performs for large processing graph like networks, web models.