INTERNATIONAL JOURNAL OF ADVANCE RESEARCH AND DEVELOPMENT

(Volume 5, Issue 4)

Available online at: www.ijarnd.com

Calculation on the false assumption makes incorrect result

Zin Mar Win[1], Swe Nwe Ahnigar Aye[2]

[1]*Lecturer, University of Computer Studies, Myitkyina, Myanmar*

[2]*Associated Professor, University of Computer Studies, Myitkyina, Myanmar*

ABSTRACT

This paper is only for reviewing a formula using in a particular subject. We have known a false assumption is an assumption as a kind of logic fallacy. And we have also known the false assumptions would produce incorrect results. Incorrect results cause to be incorrect decisions. In this paper, we propose a little but important problem that the calculating on a false assumption of total communication time for distributed query processing within the distributed database environment. To understand about our paper, we must first explain about the basic knowledge of a distributed database system. A distributed database system consists of a collection of sites, connected together via some kind of communication network. In the distributed query processing, there is two or more sites may be included. The primary objective in distributed systems is to minimize the number and volume of message between the participant sites. There will be many possible ways of moving data around the system to satisfy the request. To minimize the network utilization, it is important that an efficient strategy be found. There may be many possible strategies for processing the request. To evaluate that which strategy is to be efficient, we will calculate the total communication time for each strategy by using the formula. The choice of the best strategy is based on the result of we calculated. So, if we must have the correct answer, then we can make the correct choice. In this paper, we will apply a formula that is teaching now in our university to calculate the total communication time. We would like to present our paper with two aspects, one is 'the false assumption on the formula that we used' and second is 'suggested for that formula with other view'.

Keywords— *False assumption, Total communication time, Distributed database environment, Distributed query processing*

1. FOREKNOWLEDGE

An example of distributed database

Customer {cno,cname,city}	10,000 stored tuples at site X
Item {ino, itype}	100,000 stored tuples at site Y
Order {cno,ino,qty}	1,000,000 stored tuples at site X

Assume that every stored tuple is 25 bytes (200 bits) long.

An example of distributed query

Query: Get customer names for Delhi customers who ordering of item 'Pen'.

((Customer Join Order Join Item) where city = 'Delhi' and itype = 'Pen') {cname}

Estimated cardinalities of certain intermediate results;

no of Pen item = 10

no of orders by Delhi customer = 100,000

Communication assumption:

Data rate = 50000 bit per second

Access delay = 0.1 second

The formula that we used to calculate the total communication time

Total communication time = (total access delay) + (total data volume / data rate)

which becomes (in seconds)

Total communication time = (number of message /10) + (number of bit / 50000)

Message required

For data checking site to site - each data checking that required two messages

For data moving site to site - each table moving that required one message

1.1 The false assumption on the formula

From the above formula, we want to discuss the part, the calculation of the total access delay. According to the formula we can drive the total access delay (in seconds) with number of messages is dividing by 10. It means that,

If access delay = 0.1 second

Total access delay = number of message / 10

At that point, some people were assuming with like that:
In the first case, some people inferred that whenever the given access delay was changed there was no need to change the formula. They thought that the given formula is default. In the second case some people determined that if access delay = 0.2 second, then

Total access delay = (number of message /20)

Because of (if access delay = 0.1, Total access delay = number of message/10)

Indeed the assumptions on the formula had been wrong. One consequence of these assumptions is that the calculation will lead to produce incorrect result.
The true assumption of the first case is that the calculation of the total communication time, access delay is an important factor to get the result. If the access delay is changed the total communication will also change. And the formula we should be used is also change on the base of the given access delay.

For the second case, we want to explain the solution like this with some example:

Table 1: Examples of solution for the second case

Because of	Access delay	(so)Total access delay
1/10 =	0.1	Number of message/10
1/5 =	0.2	Number of message/5
3/10 =	0.3	Number of message×3/10
1/100 =	0.01	Number of message/100

If the calculation of total access delay is not correct, the total communication time will also be incorrect. The decision based on the incorrect result is never correct.

We will be proved that with the following example.
Query: ((Customer Join Order Join Item) where city = 'Delhi' and itype = 'Pen') {cno}
Strategy: For each Delhi ordering, check if item is pen
Access delay = 0.2 second
Notice: With this strategy, there is no data moving around the sites, so (total data volume / data rate) is not affected.
Formula: Total communication time = (number of message /5) + (number of bit / 50000)

By the false assumption:

$$\text{Total communication time} = \frac{100000 \times 2}{20} = 10000 \text{ seconds approx.}$$

By the correct assumption:

$$\text{Total communication time} = \frac{100000 \times 2}{5} = 40000 \text{ seconds approx.}$$

The two results are very different so we can see that the false assumption on the formula makes incorrect result and it will lead to decide incorrect decision.

1.2 Suggested for that formula with other view

Whenever change the access delay we cannot consider how to change the formula, we suggested with the other view. We can apply the relation between the division and multiplication. With the multiplication point of view, the above formula become

Total communication time = (number of message × access delay) + (number of bit / 50000)

So, we can calculate the above example with this formula that we get is:

Total communication time = 100000 ×2× 0.2 = 40000 seconds approx.

We can see the results with division or multiplication are the same. So you can choose one of the two formulas as your wish.

2. CONCLUSION

Indeed, the formula we applied in this paper was used for the demonstration of distributed query optimization process in distributed database management system. But on the other hand we would be used various formulas to produce the results for

various areas (mathematics, science, engineering, etc…).So we should be noticed that the formula we used with calculating on the false assumption will lead to incorrect result. An assumption is something that we assume on the formula, as if it was true, we've constructed a reality result that does not reflect what's happening in the moment.

3. ACKNOWLEDGEMENT

I would like to thank my colleague Swe New Ahnigar Aye for her major advice of language aspect and I really thanks to all thankworthy person.

4. REFERENCES

[1] An introduction to Database System, seven edition, Chapter 20, Distributed Database System,C.J.Date

