Abstract - This paper presents the position that a formula can be used to assess the risks of a distributed datacenter. This formula may be compared against benchmarks to assist an administrator in making decisions concerning server upgrades, datacenter placement, etc. It can also be used to help the administrator decide what data to replicate and where to place the copies.

Keywords: Distributed Databases. Fault-tolerance, Fault-tolerance measurement.

1. Introduction

In the globalized world, companies, based on the proximity to the resources, markets and transportation hubs, consider opening branches and data centers in different locations over the world. These datacenters usually house large databases with data needed by their other branch locations. Hence, establishing the optimal distributed database system and the optimal configuration of redundancy and replication presents significant challenges.

Today, datacenters and computer systems are expected to be available all the time. The potential loss of revenue, customers and reputation is more and more dependent on the reliability of your datacenters. For some businesses, even small amounts of downtime may cost millions of dollars. As we now have global enterprises which support users, clients or customers in multiple time zones, it is no longer possible to bring systems down at night for extended periods of time to do maintenance and upgrades. Likewise we need to plan for ways to compensate for unplanned outages. In distributed database systems, when one site experiences an unplanned outage, there will be no way to access the data from that site, unless we have made provisions ahead of time for data replication. When one site or server is down, the replica of the data is used in another server or location, so that the system continues to function and respond in such a way that the users do not notice any problems.

In this paper we are proposing a formula that analyzes the risk to the servers. This analysis is calculated based on a variety of weighted factors such as an analysis of error logs for the servers, assessing reliability of the hardware components based on age and environmental factors, the size of the data, the frequently of data usage, and the data distribution model employed in the system. This formula may then be used to determine the risk of the system and to aid in the decision as to whether or not to replicate data, and if so, what to replicate and where.

2. Risks

The goal of any server or datacenter is to have high availability. The common measurement of “nines” has been debated over the years. Nonetheless, the closer to 100% availability that a server or datacenters obtains, the more satisfied the users become. The risks of server or datacenter outages come from a broad range of possibilities. The following graph shows the breakdown of the cause of outages [1]:

![Graph showing the breakdown of the cause of outages]

- **System Software** 27%
- **Hardware** 23%
- **Human Error** 18%
- **Network Failure** 17%
- **Natural Disaster** 8%
- **Unknown** 7%
3. Costs of Not Achieving High Availability

The cost and impact of not achieving an acceptably high amount of availability varies greatly depending on the type of business or entity involved. While unavailability of data may cost millions of dollars for some businesses, other outages may have zero cost to the entity—it all depends on the type of entity impacted. For example, if the entity impacted is a firm that trades stocks or a large retailer during heavy holiday shopping times, the measurable monetary loss would be great. However, if the entity is a repository for information and the availability was lost during a non-peak time, then the loss was negligible. There have been a variety of surveys and studies done that study this type of impact, such as those done by the Ponemon Institute [2]. However, even in the case where there was no monetary loss due to data not being available, there are unmeasureable consequences, such as reputation, that are still at stake.

4. Minimization of Risks and Costs through Database Replication

Distributed databases, where all or portions of the data are replicated on other servers or locations, have long been viewed essential in achieving an acceptable level of data availability. The costs associated with managing the entire system are often outweighed by the potential loss associated by unavailability. Synchronous replication and asynchronous replication are options that exist and are considered in the overall calculation of risk.

4.1 Synchronous Replication

This replication enables zero data loss disaster recovery by ensuring that the data stored at the secondary storage site is the exact mirror image of the data at the primary data site. In this method of replication, each update must be recognized and confirmed at both the primary site and secondary site before the application can continue production. Thus, the system ensures that the secondary site is always in sync with primary data center. This enables the secondary site to take over production immediately in case there is any disruption at the primary site. However this method of replication is very expensive in terms of time. Distance is also a major drawback of synchronous replication. The distance between primary centers and backup centers is restricted due to use of fiber channels which can only extend to 200km [4].

4.2 Asynchronous Replication

With asynchronous replication, updates are propagated to the copies periodically, not as they are made. One major advantage of this is that it is much more efficient in terms of transmission time—periodic updates mean less network overhead. In addition to this, asynchronous copies can extend to any distance without affecting the propagation. This means secondary sites can be located thousands of miles away from the primary site, ensuring that secondary data is away from any likely disaster region.

The greatest disadvantage of asynchronous replication is the time lag between data being stored at the primary site and the remote site. This may mean that transactions and data not replicated at the time of disaster will be lost. In the event of any unplanned outage, data on the secondary storage may not be current. [3]

The next subsections detail two approaches to permitting asynchronous replication. Which is best would be somewhat system dependent; the choice would be left to the individual systems administrator.

4.2.1 Primary Copy Asynchronous Replication

In this approach, one copy of the data is deemed the “primary copy” and always must be current. The other copies will not be updated immediately; the primary copy server will push out the updates to the remote copies periodically. This has the advantage of using a simple communication model. In a case when a server (either a server holding a primary copy or one holding a secondary copy) has a failure, the updates needed may be held in stable queues at the other sites until this server comes back up. This allows an element of fault tolerance to be built into the basic idea of replication [3].
4.2.2 Update everywhere

In the Update everywhere technique, any copy may be updated at any time. Periodically, the servers will exchange their updates with each other. This has the advantage of lower communication cost during updates (as compared with primary copy replication). However, it requires a much more complicated synchronizing algorithm, since different copies may have updated the same data.

5. How to Determine Where to Replicate Data

Determining where to replicate all, or a portion, of a distributed database can be based on a variety of factors. Depending on the entity, the factors will include such items as: the proximity to primary/heavy users, the reliability/capacity of communication networks between sites, environmental factors such a typical seasonal weather related concerns, age/reliability of equipment at replicated site, comfort level of potential success of disaster recovery efforts, etc.

5.1 Hardware Risk Assessment

The follow are the factors that will help determine the hardware risk level of the server or data center in the distributed environment. These factors also help in determine the site for data replication.

a) Determine the access/query and update frequency for each table in the database. Likewise the number of applications that access each table, and the frequency with which those applications runs.

Analyzing this data will give an indication of which portions of the database have the highest need to be replicated.

b) Determine the amount of time, and the number of times, a server or datacenter had an unplanned outage within a given interval.

c) Determine the age and usage of the most vital hardware like CPUs, hard drives, fans, memory, etc.

5.2 Environmental Risk Assessment

Environmental risks should also be considered when placing datacenters and/or replicated parts of databases. Areas that are prone to hurricanes, earthquakes, floods, extreme hot or cold temperatures, etc. are less desirable locations for the placement of servers or data centers that need to have high availability.

While some of the above-mentioned environmental risks are hard to predict, and therefore assess, there are some environmental risks that we can assess. For example, we know that there is a high probability that at some point there will be an interruption to the power supply to a particular server or an entire data center. The cause of the interruption could be the result of a weather related event, or it could be accidental (for example a miscalculation of where a backhoe should dig during a construction project), or the interruption could be part of scheduled maintenance or upgrade to the electrical system. Hence, we also need to take into consideration other factors such as:

a) The presence and reliability of backup systems such as Uninterrupted Power Supply (UPS) devices. Likewise a determination on how frequently UPS devices are tested for functionality should be considered. For example, a datacenter has no protection against data loss nor the opportunity for graceful shutdown if the UPS devices also fail due to dead batteries or faulty failover to a fuel powered generator.

b) The presence of appropriate fire protection systems, water protection systems, cooling systems, alert notification systems, etc. The frequency of testing to make sure that these systems are fully functional also needs to be considered.

As an example, while we may not be able to assess the frequency with which ice sheets, created by an unusually bad winter storm, will fall from the roof of building and damage air conditioning units on the ground, we can minimize the likelihood that UPS unit will also fail due to a short in the failover system. Likewise, we can minimize the likelihood that the alert system will also fail due to an error in the script that indicates who should receive a text message or a pager alert.
5.3 Formula to Calculate Risk Level

A weighted formula is being developed to help determine the risk level for either an individual server or an entire datacenter. The formula takes to account the following:

1. Availability/Response time of the server/datacenter. This would include the percent of time the server/datacenter has been available in the last twelve months. Also, this would need to factor in size and frequency of data access as well as the replication scheme used.

2. Age of equipment. The value used in the calculation is the following on a sliding scale:
   - Less than 1 year old: 100 pts
   - Less than 2 years old: 80 pts
   - Less than 4 years old: 60 pts
   - Older than 4 years: 0 pts

3. Assessment of adequate cooling facilities, water protection systems and fire protection systems. The maximum points are 100 when everything is found to be adequate.

4. Assessment of reliability adequacy of UPS systems. The maximum points are 100 when it is found the UPS is completely adequate.

5. Assessment of other environmental risks as discussed in section 5.2. The maximum points are 100 when it is found that there are virtually no other environmental risks.

The weights of these four variables are:
1. Availability: 25%
2. Age of Equipment: 10%
3. Assessment of cooling, water and fire systems: 20%
4. Assessment of UPS: 40%
5. Assessment of other Environmental Risks: 5%

Example:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Weight</th>
<th>Calculations</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability/Response time in the last year</td>
<td>25%</td>
<td>Availability: 90 * 25% = 22.50</td>
<td></td>
</tr>
<tr>
<td>Age of Equipment: 15 month old server</td>
<td>10%</td>
<td>Age of Equipment: 80 pts * 10% = 8</td>
<td></td>
</tr>
<tr>
<td>Assessment of Cooling/Fire/Water protection</td>
<td>20%</td>
<td>Assessment Cooling/Water/Fire: 95 pts * 20% = 19</td>
<td></td>
</tr>
<tr>
<td>Assessment Electrical/UPS protection</td>
<td>40%</td>
<td>Assessment Electrical/UPS: 95 pts * 40% = 38</td>
<td></td>
</tr>
<tr>
<td>Assessment of other Environmental Risks</td>
<td>5%</td>
<td>Assessment Other Environmental: 90 pts * 5% = 4.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>Total = 92.00</td>
<td></td>
</tr>
</tbody>
</table>

As an example as to how this data may be used, suppose it has been determined that the minimum acceptable risk value is 93. The administrator may then recalculate with a replication schema deployed for the most vulnerable data. Suppose this increases the first metric to 95%: if so, the overall score would be increased to 93.25, which exceeds the minimum value. Thus, the administrator would be advised to replicate the most vulnerable data to achieve the desired score.

6. Conclusion

This paper has presented a position that the risk of which server/datacenter to replicate data to can be calculated by using a formula. Additional testing and simulation will allow the authors to continue to develop and fine tune the formula and the weights associated with each risk factor.

7. Bibliography


