White Paper
Y. Bot
bot@bqr.com

# Sparing for availability and sparing for confidence - use for cloud computing services

A. Segal

*BQR reliability engineering Ltd., Rishon LeZion, Israel*

Y. Bot

*BQR reliability engineering Ltd., Rishon LeZion, Israel*

ABSTRACT: Cloud computing services, Web hosting companies and Data Centers guarantee high service availability to the users (up to 99.999%). While industry standards exist for tier I – IV data centers, following the standard does not guarantee the required availability. Therefore, reliability and maintenance optimizations are crucial. Standard spare optimization strategies include sparing for confidence (where the goal is to minimize the probability of a stock-out), and cost optimization (where spare and down-time costs are minimized). However, in the case of cloud services a different approach is needed: sparing for availability. The availability of data centers depends on reliability and maintainability of many sub-systems (power supply, HVAC, safety systems, communications, and the servers themselves). As a result, the number and cost of spares required for the specified availability may be reduced by prior optimization processes of the above mentioned sub-systems. In this paper we discuss the relevant sparing calculation models, along with several examples from the field of cloud computing and data centers. Furthermore, we show that LORO (level of repair optimization: repair / discard, etc.,) increases system availability, thereby reducing the required number of spares.

# Sparing for availability and sparing for confidence - use for cloud computing services

White Paper
Y. Bot
bot@bqr.com

## 1 INTRODUCTION

Worldwide Cloud and data center business is estimated at 50 billion dollars annually (Centaur Partners 2015) and rising. The main purpose of these services is to provide the end users with robust computing and data storage capabilities which are highly reliable. This demand manifests itself in data center service level agreements (SLA) which guarantee up to 99.999% availability (See for example SLAs of Microsoft Azure for specific services). The financial loss for organizations that use cloud services reaches millions of dollars per hour of down-time (Ponemon Institute 2013, Patterson A. 2002). Similarly, data center SLAs include heavy penalties for failing to reach the promised availability. Therefore, great efforts are invested by the data center owners in order to fulfill the required availability.

Several industry standards exist e.g. TIA-942 and Telcordia GR-3160 which provide detailed guidelines for designing a range of data center systems including power supply, cooling systems, communications etc.

However, even if one follows the tier IV guidelines, achieving the required availability is not guaranteed. Therefore, availability and maintenance optimizations are a crucial part of data center design and operation.

One important parameter that affects data center availability is the number of spare parts in stock for each replaceable data center component. A shortage of spare parts can increase the system restoration time in case of a failure thus reducing availability. On the other hand, an excess of spare parts has a financial cost. Therefore, finding the right stock size is very important. The following section reviews different sparing strategies in the context of data centers and comparison of the different strategies is discussed in section 3.

## 2 SPARING STRATEGIES

### 2.1 *Cost optimization*

A well-known sparing strategy is cost optimization in which the minimal maintenance financial cost is searched for. The cost of maintenance actions can easily be defined by accounting for manpower, spares, support equipment etc. In most industries it is also possible to define the cost of system down-time. Then, the optimal balance between costs of maintenance actions and down-time is searched for.

In the case of data centers the cost of down time includes direct costs (as defined in the SLA) as well as substantial indirect costs due to loss of reputation (Wiboonrat M. 2008). It is very difficult to assess the indirect damage therefore the cost optimization approach is not advisable for data center optimization.

### 2.2 *Sparing for confidence*

Another standard way for spare optimization is sparing for confidence i.e. calculating the number of required spares such that the probability of a stock-out drops below a specified threshold (Louit D. et al 2011).

In the case where many identical components are stationed in the field, the failure events can be approximated by a Poisson process for which the probability of no stock-out is given by:

White Paper
Y. Bot
bot@bqr.com

$$P = e^{-m\lambda t} \sum_{i=0}^{k} \frac{(m\lambda t)^i}{i!} \qquad (1)$$

where $P$ is the probability of no stock-out, $k$ is the number of spares, $m$ is the number of components in the field, $t$ is the operation time and $\lambda$ is the component failure rate.

For example: given a data center with 160,000 hard drives that have a failure rate of 3.4 failures per million hours (Schroeder B. 2007) and a period of 7 days between stock replenishments (168 hours), the stock size required for $P \geq 99.999\%$ is 135.

However, there is a difference between probability of no stock-out and the system availability. The difference is especially evident in data centers due to their high redundancy design, therefore, a better approach is sparing for availability (S2A) i.e. optimizing the stock in order to ensure the required availability.

### 2.3 *Sparing for availability*

As a first simple example we consider a data center with 200 clusters, each consisting of 20 racks and each rack holding 10 servers. Each server holds 4 hard drives (Total of 40,000 servers and 160,000 hard drives).

Assumptions: It is assumed that each server has 2 pairs of hard drives and each pair of drives is mirrored i.e. a failure occurs only when both hard drives of a pair fail. Furthermore, the servers, server racks and clusters were modeled in serial. This means that a single server failure constitutes a failure of the whole data center (more elaborate models will be discussed later). Hard drives are discarded upon failure and the stock is replenished on a weekly basis.

Calculation: The system described above was modeled using the apmOptimizer software. The S2A module of the apmOptimizer uses dynamic programming in order to build and compare sparing options in a bottom-up approach. For each sparing option the corresponding Markov chain is calculated in order to produce the resulting availability. This approach is preferred over simulations which are not practical for such a large numbers of components.

Results: Using the apmOptimizer S2A module it was found that 79 hard drives have to be placed in stock in order for availability of 99.999% to be reached. Note that 79 is lower than 135 (as recommended by the sparing for confidence approach), and it is also lower than the average number of hard drive failures per week (92). The reason for this is that redundancy allows the data center to stay up even when by the end of the week several hard drive pairs are running on a single hard drive.

A similar calculation was conducted for the diesel backup power generators. In this case 10 generators were modeled, each with a failure rate of 14.09 failures per million hours (IEEE std. 493 Gold Book) and average repair time of 1257 hours. Sparing for confidence demands 4 backup generators in order to have $P \geq 99.999\%$. However, tier IV

White Paper
Y. Bot
bot@bqr.com

data centers have redundant input power lines, each with an availability of 99.9536% (Eto J.H. et al 2008) therefore the use of generator backup power is extremely rare and the power system availability is larger than 99.999% without any backup generators. Therefore, the cost of 4 expensive diesel generators can be avoided.

A recent report (Ponemon Institute 2013) states the top root causes of data center failures: UPS system failure, Accidental/human error, cybercrime, weather related, water heat or CRAC failure, generator failure and IT equipment failure.

In light of the report mentioned above a more elaborate model was constructed which includes the main components of the power supply, HVAC and IT (see Figure 1).

| Reference Designator | Qty per parent | Reliability Model | K. | Inputted MTBF |
|---|---|---|---|---|
| DataCenter 40K servers | 1 | Serial | - | - |
| Power | 1 | Standby | - | - |
| PowerLine | 1 | Parallel | - | - |
| Grid | 2 | Serial | - | - |
| PowerUtility | 1 | Leaf | - | 1089 |
| 15KV Fuse | 1 | Leaf | - | 86330 |
| Diesel backup | 1 | Serial | - | - |
| Diesel Generator 3000KW emergency | 10 | Serial | - | - |
| Starter | 1 | Leaf | - | 212937 |
| Motor | 1 | Leaf | - | 106468 |
| HVAC | 1 | K out of N | 6 | - |
| CoolingTower | 16 | Serial | - | - |
| Pump | 3 | Leaf | - | 48733 |
| HeatExchanger | 1 | Leaf | - | 60606 |
| Clusters | 1 | K out of N | 5 | - |
| Cluster | 200 | Serial | - | - |
| Rack | 20 | Serial | - | - |
| Servers | 1 | Serial | - | - |
| Server | 10 | Serial | - | - |
| ServerPower | 1 | Parallel | - | - |
| PowerSupply | 2 | Leaf | - | 100000 |
| HardDrives | 1 | Serial | - | - |
| Hard Drive Duo | 2 | Parallel | - | - |
| HDD | 2 | Leaf | - | 292000 |
| Fans | 1 | Parallel | - | - |
| Fan Unit | 2 | Serial | - | - |
| Fan | 3 | Leaf | - | 200000 |
| RackUPS | 1 | K out of N | 2 | - |
| Battery | 11 | Leaf | - | 120000 |
| Network Switch | 1 | Leaf | - | 681009 |

Figure 1: project tree of the data center. Dark end items denote units that are replaced and discarded upon failure. Light end items are replaceable and repairable (the power utility is repairable but not replaceable).

Assumptions: Failure rates for the model were taken from IEEE std. 493 Gold Book as well as other sources (Eto J.H. 2008, Schroeder B. 2007). In order to account for the fact that server failures may affect some users but not others a critical failure was defined as a state in which 5 or more of the 200 clusters are down.

Y. Bot
bot@bqr.com

Results: 1 heat exchanger, 3 network switches, 114 UPS spare batteries and 159 fans were recommended by apmOptimizer in order to achieve system availability of 99.999%.

The reason that no hard drives are required in the stock is that given the extra redundancy in the model (5 cluster failures out of 200 clusters), the hard drives contribute a negligible amount of failures to the system. Further inspection of the model and results reveals that many UPS batteries are required in stock due to the fact that battery supply is on a periodic monthly basis (compared to weekly supply of hard drives).

This demonstrates the sensitivity of system availability and the corresponding sparing policy to specific details in the data center design and logistics.

## 2.4 *Level of repair optimization*

Another important factor is the choice of repair policy (repair / discard and location of repair shops, stocks and vendors).

For example: using the apmOptimizer level of repair optimization (LORO) module it was found that reduced costs can be achieved by replacing the UPS battery vendor with a different one that supplies cheaper batteries once every 60 days. The long replenishment period demands a larger initial battery stock in order to sustain the required availability. However, the increased initial cost is balanced by the cheaper cost of batteries for the data center lifetime (32,089 battery replacements occur during ten years).

Therefore, S2A and LORO can be used together in order to find the cheapest combination that fulfills the availability requirement.

## 2.5 *Performance optimization*

Inspection of data center SLAs reveals that their definition of availability ($A_{SLA}$) is different from the system availability:

$$A_{SLA} = \frac{Average\ number\ of\ available\ servers}{Number\ of\ servers} \quad (2)$$

Standard reliability models are not equipped for calculation of $A_{SLA}$ however it was found that the apmOptimizer performance module, originally developed for the oil and gas industry allows calculation of $A_{SLA}$. Instead of calculating the flow of oil through different pipe configurations and bottlenecks, the module can calculate the “flow” of operational computing units. For example: The servers are analogous to oil sources and the power utility is analogous to the main oil line (power supply failure stops all computing units).

Calculation: The reliability model for the data center was updated in order to reflect the actual case where a single server failure does not cause a rack failure and a rack failure does not cause a cluster failure. In this case the standard system availability is extremely high. Next the performance calculation was carried out.

White Paper
Y. Bot
bot@bqr.com

Results: Using the apmOptimizer performance module it was found that $A_{SLA}$=99.994%. In order to achieve $A_{SLA}$=99.999% another network switch and 67 UPS batteries had to be added to the stock.

## 3 DISCUSSION

Several methods for data center optimization were presented: cost optimization, sparing for confidence, S2A and performance.

Sparing for confidence is a straightforward method for spare optimization but it disregards system redundancy and therefore recommends a higher amount of required stock components than actually required.

S2A has the advantage of accounting for the system reliability model as well as the logistics of the data center. Furthermore, it can be used together with LORO in order to further optimize the data center cost and availability.

While S2A can help data center owners to optimize the power supply, HVAC and safety systems, it is not clear how to define the reliability model for the servers. S2A can also be useful for service providers that use cloud services, know where their specific servers and racks are located and how failure of specific servers and network switches percolates through the system until a service failure occurs.

Finally, data center optimization for performance was explored. This procedure is very useful for data center owners that have to meet availability requirements as defined in the SLA.

## 4 SUMMARY

Several methods for spare optimization of data centers were discussed: cost optimization, sparing for confidence, sparing for availability and sparing for performance. It was shown that substantial financial savings can be obtained by choosing the correct optimization method. Sparing for availability and performance were found to be better compared to cost optimization and sparing for confidence. Furthermore, improved results can be obtained by combining the optimizations with a level of repair optimization.

apmOptimizer™

Sparing for availability and sparing for confidence - use for cloud computing services

BQR Reliable Engineering Optimized Maintenance

White Paper

Y. Bot
bot@bqr.com

REFERENCES

Centaur Partners 2015. Introduction to Centaur Partners SaaS Market Overview

Eto J.H. & LaCommare K.H. 2008, Tracking the Reliability of the U.S. Electric Power System: An Assessment of Publicly Available Information Reported to State Public Utility Commissions. ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

Louit D., Pascual R., Banjevic D. & Jardine A.K.S. 2011. Optimization models for critical spare parts inventories—a reliability approach. *Journal of the Operational Research Society* 62, 992–1004

Patterson A. 2002. A Simple Way to Estimate the Cost of Downtime. *Proc. LISA '02: Sixteenth Systems Administration Conference*, 185-188

Ponemon Institute, sponsored by Emerson Network Power 2013. Research report: 2013 Cost of Data Center Outages

Schroeder B. & Gibson G.A. 2007. Understanding disk failure rates: What does an MTTF of 1,000,000 hours mean to you? *ACM Transactions on Storage* 3, 3

Wiboonrat M. 2008. An Optimal Data Center Availability and
Investment Trade-Offs. *Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*