Availability measurement of grid services from the perspective of a scientific computing centre

H. MARTEN, T. KOENIG

Karlsruhe Institute of Technology
Steinbuch Centre for Computing
Hermann-von-Helmholtz-Platz 1, D-76344 Eggenstein-Leopoldshafen

E-mail: holger.marten@kit.edu, tobias.koenig@kit.edu

Abstract. The Karlsruhe Institute of Technology (KIT) is the merger of Forschungszentrum Karlsruhe and the Technical University Karlsruhe. The Steinbuch Centre for Computing (SCC) was one of the first new organizational units of KIT, combining the former Institute for Scientific Computing of Forschungszentrum Karlsruhe and the Computing Centre of the University. IT service management according to the worldwide de-facto-standard “IT Infrastructure Library (ITIL)” [1] was chosen by SCC as a strategic element to support the merging of the two existing computing centres located at a distance of about 10 km.

The availability and reliability of IT services directly influence the customer satisfaction as well as the reputation of the service provider, and unscheduled loss of availability due to hardware or software failures may even result in severe consequences like data loss. Fault tolerant and error correcting design features are reducing the risk of IT component failures and help to improve the delivered availability. The ITIL process controlling the respective design is called Availability Management [1]. This paper discusses Availability Management regarding grid services delivered to WLCG and provides a few elementary guidelines for availability calculations of services consisting of arbitrary numbers of components.

1 Introduction

This document provides the theory for availability calculations of grid services provided by the Grid Computing Centre Karlsruhe, GridKa [2], a Tier-1 centre in the framework of the World Wide LHC [3] Computing Grid Project, WLCG [4]. Service availability definitions given in the WLCG Memorandum of Understanding [5] as well as experience with the currently used grid middleware stack are taken into account to design the grid services at the GridKa Tier-1 with required availability.

The following chapter outlines some terminologies like availability, reliability and maintainability used in the scope of ITIL Availability Management and gives a brief introduction into the WLCG Memorandum of Understanding. Chapter 3 introduces the concept of internal and external views of IT services and demonstrates the logical service breakdown into Service Components and Configuration Items. For both these component classes elementary equations are derived that allow overall availability calculations. Finally, the benefit of these estimations is demonstrated and substantiated by means of a few simple example services.

1 To whom any correspondence should be addressed.
2. Objectives and scopes of Availability Management

Availability Management is an ITIL process in the early design phase of an IT Service Lifecycle. Its goal is to ensure that the level of availability delivered in all services is matched to or exceeds the current and future needs in a cost-effective manner. This is usually done by creating and maintaining an up-to-date availability plan for all services negotiated with all involved parties. In the sense of a request oriented design, customers usually provide their requests in terms of Service Level Requirements (SLR). In order to define a common language and avoid future misunderstandings, the SLR together with respective service availability measures are formally documented in a Service Level Agreement (SLA) to be signed by the customer and the provider. The IT provider himself will / must also formally agree on the different service levels with each of his infrastructure suppliers as well as with his own personnel. This is done by setting up respective Operational Level Agreements (OLA). SLRs, SLAs and OLAs require a set of terminologies for common understanding and verification by all involved parties.

2.1. Terminologies

**Availability**
The availability is the ability of an IT service or component to perform its required function at a stated instant or over an agreed period of time. It depends on the
- Availability of its components and possibly other services,
- Changes in the future,
- Quality of maintenance and support,
- Quality, pattern and extent of operational process and procedures,
- Security, integrity, and availability of data.

**Reliability**
The reliability of an IT service can be quantitatively described as “absence of operational failure”. The reliability of an IT service is determined by the reliability of each component within an IT infrastructure delivering the service and the level of resilience designed and built into the infrastructure.

**Maintainability**
Maintainability relates to the ability of an IT infrastructure component to be retained in, or restored to, an operational state. It can be divided into 7 stages:
- The anticipation of failures,
- The detection of failures,
- The diagnosis of failures,
- The resolution of failures,
- The recovery from failures,
- The restoration of the data and the service,
- The levels of preventive maintenance to prevent potential future failures.

2.2. The WLCG Memorandum of Understanding

In the WLCG project a Memorandum of Understanding (MoU [5]) defines the capacities as well as the minimum levels of service availabilities at the Tier-0/1/2 centres. It proved to be the guideline for building individual availability service designs at the participating computing centres.
Table 1: Minimum service levels\(^2\) at Tier-1 centres according to the WLCG MoU [5].

Table 1 depicts the minimum service levels of Tier-1 centres according to the WLCG MoU. The response time in hours refers to the maximum delay before an action is taken to repair a service failure. The actual average time for a service to be available can be calculated from the percentage values in the fifth column of the table. It only accounts for the actual data taking period of the LHC accelerator, while column six is valid during the rest of the year. All these agreed parameters require a fault tolerant service design with an adequate level of staffing during prime and on-call service hours.

| Service                                               | Maximum delay in responding to operational problems | Average availability\(^2\) measured on an annual basis |
|-------------------------------------------------------|-----------------------------------------------------|-----------------------------------------------------|
| Service interruption                                  | Degradation of the capacity of the service by more than 50% | Degradation of the capacity of the service by more than 20% | During accelerator operation | At all other times |
| Acceptance of data from the Tier-1 Centre during accelerator operation | 12 hours | 12 hours | 24 hours | 99% | n/a |
| Networking service to the Tier-0 Centre during accelerator operation | 12 hours | 24 hours | 48 hours | 98% | n/a |
| Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres out with accelerator operation | 24 hours | 48 hours | 48 hours | n/a | 98% |
| All other services\(^3\) – prime service hours\(^4\) | 2 hour | 2 hour | 4 hours | 98% | 98% |
| All other services\(^3\) – out with prime service hours\(^4\) | 24 hours | 48 hours | 48 hours | 97% | 97% |

The “average availability” is defined in the WLCG MoU as (time running) / (scheduled up-time), i.e., it explicitly excludes scheduled downtimes for maintenance.

“All other services” are services essential to the running of GridKa and to those who are using it.

“Prime service hours” for GridKa are 8:00-18:00 during the working week (Monday till Friday), except for public holidays and other scheduled closures of the site.
Concerning the internal view, a service is composed of one or several Service Components (SC) which in turn may consist of one or several Configuration Items (CI). Although ITIL does not uniquely regulate the definition of Service Components and Configuration Items, the following definitions allow deducing some elementary calculation guidelines for the total availability $A_{\text{total}}$ of a service (Figure 1):

**Service Components (SC)**
A service consists of $m$ Service Components (SC) with individual availabilities $A_k$. These SCs describe the internal structure of a service in terms of building blocks, e.g. the external network, the frontend service, the internal network, backend storage etc. Service Components are logically (not necessarily technically) arranged sequentially, implying that the whole service fails if one of its components fails.

**Configuration Items (CI)**
A Service Component consists of $n$ Configuration Items (CI) with individual availabilities $A_{k,i}$ ($i=1,...,n$). CIs describe the internal structure of a Service Component and are always logically arranged in parallel, meaning that the respective Service Component fails if and only if all its CIs fail at the same time. This is the redundancy concept in the service design, and typical examples of Service Components are redundancy paths of a server through two dedicated internal network switches, or the individual nodes of a high-availability database cluster.

Given the definitions above, individual and total availabilities can be calculated in the following way:

**Availability of a Configuration Item (CI)**
The basic availability $A_{k,i}$ of a Configuration Item $i$ of a service component $k$ is

$$A_{k,i} = \frac{AST - DT_{k,i}}{AST},$$

where $AST$ denotes the Agreed Service Time of the whole Service, and $DT_{k,i}$ is the actual measured downtime of the configuration item during the $AST$. It should be noted that the $A_{k,i}$ defined above represent average measured availabilities but could also be replaced for the following equations e.g. by general information on the CI provided by the vendor.

---

5 Without further notice we restrict ourselves to the analysis of a single service consisting of several components and configuration items in the following.
Availability of a Service Component (SC)
Given the basic availabilities $A_{k,i}$ of all its Configuration Items, the average availability $A_k$ of a Service Component becomes

$$A_k = 1 - \prod_{i=1}^{n} (1 - A_{k,i}).$$

The interpretation of this formula is straightforward. Since $A_{k,i}$ denote the (average) availabilities of all CIs within the agreed service time, $(1 - A_{k,i})$ represent the respective probabilities to fail. The product thus gives the probability that all CIs of the respective Service Component (and that were assumed to provide the same functionality in parallel) simultaneously fail at the same time.

Total availability of a Service
Finally, given the availabilities $A_k$ of the Service Components, its total availability $A_{total}$ is determined by the product of the $m$ availabilities of its components,

$$A_{total} = \prod_{k=1}^{m} A_k.$$

4. Example calculations – the benefit of redundancy
The achievement potential of the above simple set of equations is demonstrated by an example of the Grid File Transfer Service (FTS) in Table 2.

| k | IT Service Component      | Diagram | $A_{k,i} = \frac{AST - DT_{k,i}}{AST}$ | CI Availability | 1-CI Availability |
|---|---------------------------|---------|----------------------------------------|-----------------|-------------------|
|   |                           |         |                                        | $i=1$           | $i=2$             | $i=3$           | $i=1$           | $i=2$             | $i=3$             |
| 1 | Rack                      |         |                                        | 0,98            | 0,99              | 0,00            | 0,02            | 0,01              | 1,00              | 0,999600          |
| 2 | UPS                       |         |                                        | 0,95            | 0,97              | 0,99            | 0,05            | 0,03              | 0,01              | 0,999985          |
| 3 | WAN Router                |         |                                        | 0,97            | 0,97              | 0,99            | 0,03            | 0,03              | 0,01              | 0,999991          |
| 4 | LAN Switch                |         |                                        | 0,99            | 0,00              | 0,00            | 0,01            | 1,00              | 1,00              | 0,990000          |
| 5 | DNS external              |         |                                        | 0,99            | 0,00              | 0,00            | 0,01            | 1,00              | 1,00              | 0,990000          |
| 6 | DNS internal              |         |                                        | 0,98            | 0,00              | 0,00            | 0,02            | 0,01              | 1,00              | 0,999980          |
| 7 | FTS Nodes                 |         |                                        | 0,95            | 0,97              | 0,99            | 0,05            | 0,03              | 0,01              | 0,999985          |
| 8 | Glite FTS Webservice      |         |                                        | 0,99            | 0,99              | 0,99            | 0,01            | 0,01              | 0,01              | 0,999999          |
| 9 | Glite FTA Transfer Agents |         |                                        | 0,96            | 0,97              | 0,99            | 0,04            | 0,03              | 0,01              | 0,999988          |
| 10| SAN                       |         |                                        | 0,99            | 0,99              | 0,99            | 0,01            | 0,01              | 1,00              | 0,999900          |
| 11| DB Nodes                  |         |                                        | 0,96            | 0,98              | 0,99            | 0,04            | 0,02              | 0,01              | 0,999992          |

Service availability with redundancy:

$$A_{total} = \prod_{k=1}^{m} A_k = 0,9796$$

Service availability without any redundancy:

Product of elements of fourth column ($i=1$)

$$0,7443$$

Table 2: Comparison of the total availability $A_{total}$ of the Grid File Transfer Service (FTS) with and without redundant Configuration Items
The table shows the FTS setup consisting of eleven Service Components; some of them are redundant like the rack, the UPS, the WAN router, etc. But some of them left without any redundancy, like the LAN Switch and the external DNS. All CIs with estimated availabilities are ordered sequentially, giving a total availability of $A_{\text{total}} = 97.96\%$.

In the last row of the table, the availability of the FTS service is calculated assuming no redundancy at all. In this case the total availability is the product of the Configuration Item availabilities of the fourth column ($i=1$). This not only decreases the individual availabilities of the former redundant Service Components below 99\% but also decreases the total availability of the whole service to $A_{\text{total}} = 74.43\%$.

Several interesting conclusions can be drawn from these simple examples which are not surprising in general. However, the availability equations nicely substantiate the conclusions and help to quantify estimates:

- In case of sequentially arranged Service Components, the most easy and certainly most cost effective availability increase can be achieved by simplification. For example, if a sequence of seven Service Components with 99\% availability can be reduced to four, the total availability increases from 93.21\% to 96.06\%.

- Since the total availability of sequentially aligned Service Components is given by the product of the individual component availabilities $A_k$, the weakest component in the chain significantly limits the total availability of the whole setup, certainly giving the primary starting-point of overall optimisation. For example, only exchanging the host of the LAN switch in Table 2 by a less reliable one with 96\% availability would decrease the total availability from 97.96\% to 94.99\%. On the other hand, even an exchange by the best procurable machine with 100\% availability would still limit the maximum total service availability to 98.49\%.

- A comparison of the above examples with and without redundancy clearly demonstrates the benefit of redundancy concepts in numbers. Moreover, replacing single CIs by two or more identical ones allows performing scheduled maintenances of single CIs without total service outage. Although our availability calculations explicitly excluded scheduled maintenance, respective times can be easily taken into consideration by modifying the availability equation for the CIs (definition of $A_k$ above).

- Last but not least, cost optimisation is another important factor for the decision on how to assemble a service of single components. Multiplying the availabilities of SCs and/or CIs with corresponding cost factors for investment and operation allows estimating the overall cost of availability in different technical setups. It often turns out that redundancy with two low-availability but cheap components is more cost-effective than a single high-availability component.

5. Summary

This paper discusses a set of easy-to-use equations to estimate the total availability of services that are composed of several sequentially aligned Service Components and redundant Configuration Items that are operated in parallel. Together with respective factors for investment and operations cost they can provide the fundament for the assessment of alternative IT service configurations. Although not discussed in detail it should be noted that these equations are not restricted to the pure hardware setup but can and should include as well the software stacks required to operate the full service.

References

[1] Standard ITIL version 3 literature “Service Design” published June 2007, Office of Government Commerce (OGC) http://www.best-management-practice.com/Publications-Library/IT-Service-Management-ITIL/
[2] GridKa, the German Tier-1 centre http://grid.fzk.de/
[3] Large Hadron Collieder (LHC) [http://public.web.cern.ch/public/en/LHC/LHC-en.html](http://public.web.cern.ch/public/en/LHC/LHC-en.html)
[4] Worldwide LHC Computing Grid (WLCG) [http://lcg.web.cern.ch/LCG/](http://lcg.web.cern.ch/LCG/)
[5] Memorandom of Understanding (MoU) of the WLCG [http://lcg.web.cern.ch/LCG/mou.htm](http://lcg.web.cern.ch/LCG/mou.htm)
[6] “ITIL and Grid services at GridKa” H Marten and T Koenig 2010 *J. Phys.: Conf. Ser.* **219** 062018 [http://iopscience.iop.org/1742-6596/219/6/062018/](http://iopscience.iop.org/1742-6596/219/6/062018/)
[7] Service Catalogue of the Steinbuch Centre for Computing (SCC, in German) [http://www.scc.kit.edu/dienste/index.php](http://www.scc.kit.edu/dienste/index.php)
[8] GridMap – Visualizing the "State" of the Grid [http://gridmap.cern.ch/gm/#topo=regions&layout=tc&vo=OPS&serv=Site](http://gridmap.cern.ch/gm/#topo=regions&layout=tc&vo=OPS&serv=Site)
[9] GridKa monitoring web page [http://www.gridka.de/monitoring](http://www.gridka.de/monitoring)