EUNIS 2009: AVAILABILITY AND CONTINUITY MANAGEMENT AT TECHNISCHE UNIVERSITÄT MÜNCHEN AND THE LEIBNIZ SUPERCOMPUTING CENTRE

Wolfgang Hommel[1], Silvia Knittl[2], and Daniel Pluta[3]

[1]Leibniz Supercomputing Centre, Boltzmannstr. 1, Garching n. Munich, D-85748 Germany, hommel@lrz.de
[2]Munich Network Management Team, Oettingenstr. 67, Munich, D-80538 Germany, knittl@mnm-team.org
[3]TU München, Boltzmannstr. 3, Garching n. Munich, D-85748 Germany, pluta@tum.de

Keywords
Availability and Continuity Management, ITIL, IT Service Management, ISO 20000, Identity Management

1. EXECUTIVE SUMMARY

Campus Management (CM) systems depend heavily on reliable IT services. The outage of a single component is critical for university business processes e.g. in the examination office. In turn, this may result in severe legal problems, for example when the students cannot sign up for a test.

At the Technische Universität München (TUM) and the Leibniz Supercomputing Centre (LRZ) in Germany, these issues have motivated a comprehensive programme to professionalize the IT service delivery and support. In this article, we focus on the Availability and Continuity Management processes. These two processes are essential for high-class IT Service Management (ITSM) and reflect our goal of guaranteeing the necessary quality of service to the university staff, students, guests, and affiliates.

Our efforts are based on ISO/IEC 20000, which is an international standard that serves as a guideline for the alignment of IT services with the business processes. It precisely defines core requirements concerning the processes of service continuity and availability management, which we extended by several special requirements elicited in our university environment. Currently, the LRZ is heading towards an ISO/IEC 20000 certification for several of the services it offers to the higher education institutions in the Munich area.

Throughout the presentation, we use the TUM's Identity and Access Management infrastructure, which is an essential part of the technical backbone of TUM's campus management infrastructure, to illustrate our availability and continuity management strategy. We first give a short overview of the activities that constitute professional IT service management according to ISO/IEC 20000. We then show how service level agreements (SLAs) help us to clearly define priorities and improve incident management as well as availability planning in a cooperative manner.

In the second part of our presentation, we discuss several concrete measures and components of our technical infrastructure that support our availability and continuity management strategy. Examples of the technical workflows and procedures we established include service load balancing mechanisms that, for example, are maintenance-aware and enable system and service updates without service interruption. The paper concludes with an outlook to our further ITSM activities and the current status of our ITSM tool suite deployment.

2. INTRODUCTION

A Campus Management (CM) system provides an integrated instrument for higher education institutions' administrative processes. TUM has strategically decided in 2007 to deploy TUMonline, which is based on CampusOnline (Haselbacher, Franz (2003)), as its new CM system within the project CM@TUM (TUM website (2009)). With the help of TUMonline the whole student lifecycle, from enrolment to graduation and alum status, will be electronically supported. TUMonline is also connected to TUM's comprehensive Identity and Access Management (I&AM) infrastructure, which was built within the project IntegraTUM and has been presented at EUNIS 2006 (Boursas, L., Hommel, W. (2006)). This I&AM solution also made the integration of further IT platforms possible, including groupware, file, and mail services. All of them are provided by the LRZ, which is the central IT service provider for Munich's universities as well as the Bavarian Academy of Sciences and Humanities.

Figure 1 outlines the association between TUMonline and the I&AM infrastructure, which is also operated by the LRZ. TUM's students, staff, and guests are able to access TUMonline via a web interface. During the login phase, the user's name and password is validated using an authentication service, which is also provided by LRZ. This authentication service retrieves the users' data from an LRZ-operated TUM directory service (meta-directory in Figure 1), which itself retrieves its data from TUMonline. TUM's members are also able to access various supplementary applications, like lab, wireless LAN, and file services etc. which are provided by TUM or LRZ respectively.

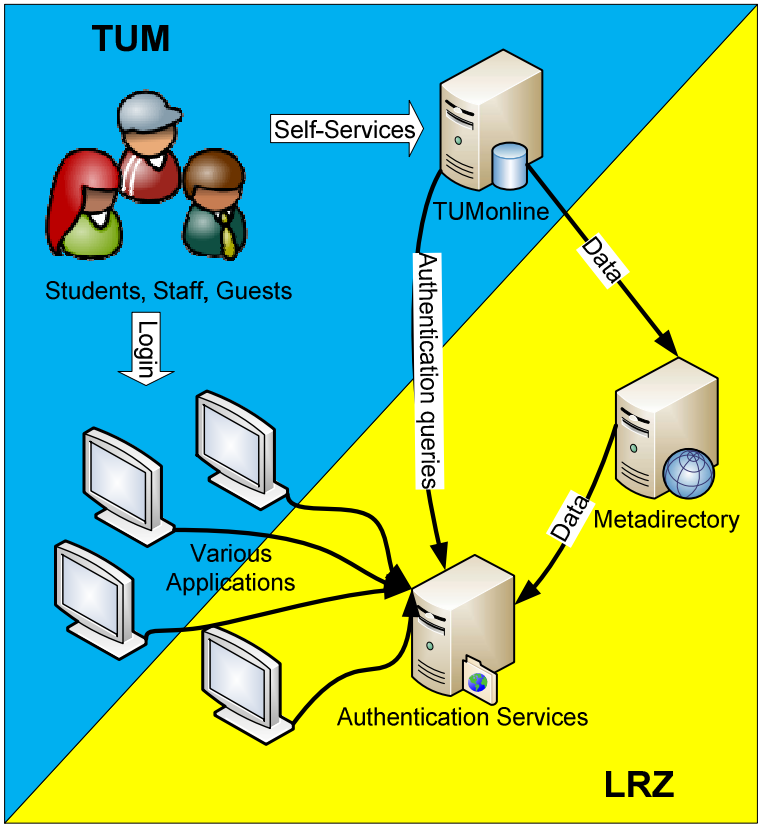

Figure 1: Small outline of IT services at TUM and LRZ

As can be seen from our scenario in Figure 1, the CM obviously depends on reliable IT. An outage of the authentication services would prevent TUM's members from using the CM, and therefore no online application or course management would be possible. Closing the gap between IT and business needs is the primary goal of IT Service Management (ITSM). ISO/IEC 20000 (ISO20k) is a common standard for IT Service Management. Part 1 of ISO20k, named "Specification for Service Management", defines minimum standards to effectively deliver services in order to meet customer requirements (ISO20000-1 (2005)). The guidelines defined in part 1 are must-haves for organizations

that are planning to obtain an ISO20k certification. Part 2, called "Code of Practice for Service Management", describes best practices (ISO20000-2 (2005)). ISO20k allows both – persons and organizations - to be certified. In our case, the LRZ is heading towards an organizational ISO20k certification.

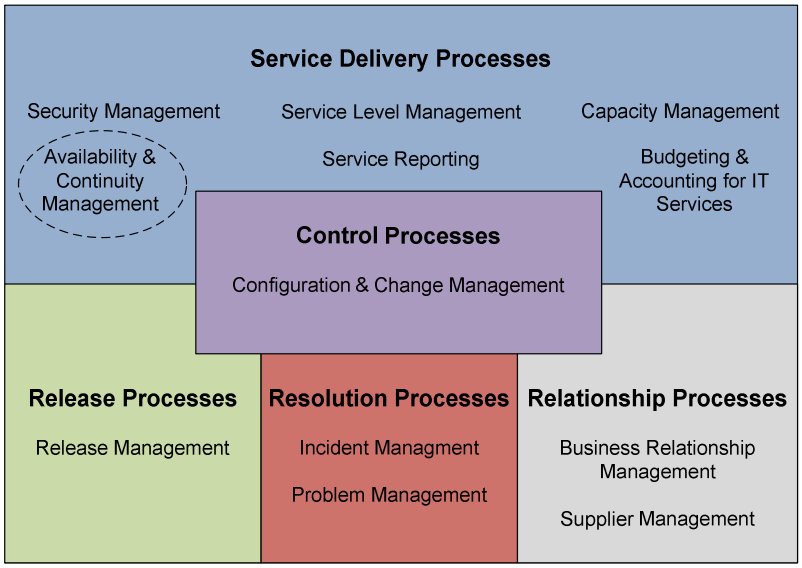

Figure 2: The ITSM processes in ISO/IEC 20000

Figure 2 shows the main ITSM processes of ISO20k. We already have demonstrated TUM's approaches for the efficient implementation of the Resolution Processes part – the introduction of a Service Desk – at EUNIS 2007 (Hommel, W., Knittl, S. (2007)). The need for an appropriate tool support to facilitate the Control Processes with the help of an efficient Configuration Management Database that spans the involved organizations TUM and LRZ was presented earlier (Hommel, W., Knittl, S. (2008)). In the following we concentrate on our Availability and Continuity management approaches.

Availability and Continuity management in general aims to assure the agreed continuity and availability to customers under all circumstances. Therefore, the following activities should be undertaken according to ISO20k: First of all, it is necessary to define the customers' requirements on availability and continuity of IT services. The customers' business priorities, service level agreements, or assessed risks are forming the basis. One method to assist this definition phase is the Component Failure Impact Analysis (CFIA). It provides vital information for designing the IT services to ensure the availability demands. Table 1 shows a simplified CFIA of our scenario at the point in time before we had implemented corresponding actions within our Availability and Continuity management. The more X indicators (which means failure of component causes failure of service) a service has the more susceptible and vulnerable it is to outages. Therefore, we have to define a set of actions to decrease this failure probability in the subsequent step.

Configuration Item	**Service** CFIA state e.g. enrolment via TUMonline
TUMonline	**X**
Data synchronization TUMonline → Meta-directory	**X**
Meta-directory Service	**X**
Authentication Service	**X**
Service CFIA state '**X**': Failure of CI causes Service outage Service CFIA state '**A**': Alternative CI available to provide Service Service CFIA state '**B**': Alternative CI available, but recover actions needed first	

Table 1 CFIA from TUMonline's perspective

After having defined the corresponding requirements, the next step is to plan both the availability and the service continuity activities. Therefore, it has to be defined first whether the availability, serviceability or maintainability has to be analyzed. Where availability means the ability to perform the required function over a stated period of time, serviceability contains arrangements with external suppliers to assure availability of components under their care, and maintainability refers to the possibility of a component to be restored to an operational state. To improve the availability, the IT management has to decide, for example, whether the hardware for critical systems should be provided redundantly. By steering the maintainability the planned downtime can be influenced.

For planning service continuity issues, a service continuity plan has to take into account the dependencies between services and configuration items. Vital parts of a continuity plan are, for example, the definition of preventive actions being taken like backup procedures. It has also to be assured that such backups of data are available for restoration in case of a failure or a disaster.

The last activity after having defined and planned Availability and Continuity procedures is to monitor the agreed availability levels and to test the established plans and implemented actions on a regular basis. In the following we describe the implementation and results of our Availability and Continuity management.

3. OUR IMPLEMENTATION

The application of the three discussed activities - defining requirement, planning activities, monitoring, and testing - in our scenario (as shown in Figure 1) will now be described. We therefore depict the Availability and Continuity management from TUMonline‘s point of view.

As can be seen in the CFIA, TUMonline is directly dependent on the authentication services, which itself are dependent on the I&AM meta-directory. These two services are therefore critical components in the entire system. Since TUMonline is a vital business application for TUM, it is necessary to have the highest possible availability guaranteed. Having noticed this, TUM will set up a SLA with LRZ addressing these availability guarantees. Comparing the availability demands of the meta-directory and the authentication services, it can be derived that on the meta-directory side, short outages might be acceptable for TUM from the TUMonline perspective, since there is no direct access to it. But with an outage at the authentication services no user is able to log in to TUMonline and to the other attached services anymore. That means that even short breakdowns are not acceptable. Therefore, the authentication services need stronger availability actions to be implemented.

In the planning phase, it was decided to have several technical measures in place, to ensure that the requirements are met. First of all, it was decided to integrate both services in a daily backup procedure. The backup data are furthermore mirrored to an external place to ensure the continuity of the services in case of disaster, such as fire or flooding. This action changes the state of this two services in the CFIA of Table 1 to “B = Alternative CI available, but recover actions needed first”. To further improve the availability, both services, meta-directory and authentication services, are provided on multiple redundant machines which are attached to different rack power supplies. This redundancy ensures that one system is still available to provide the service in case of a breakdown of the other system. This changes the state to “A=Alternative CI available to provide Service” in Table 1. Concerning the authentication services it is furthermore important to keep in mind that the access of the various connected clients must provide high performance rates. To improve the performance in accessing this service, a comprehensive load balancing mechanism is used. This load balancing mechanism allows a dynamic switching of the established connections to the other system in case one system has had a breakdown.

Next to the performance advantages of the described load balancing infrastructure above, the overall services' quality also profits from the redundancy which offers high-availability (HA) load balancing. The technical implementation of the HA load balancing mechanism is shown in Figure 3. The load balancer now introduces a new component, which is critical itself in case of failure. Hence, the load balancer is also deployed redundantly. In our implementation the access to the authentication services is separated from the clients' access via the outer network and a management access within the inner network (Server LAN 10.a.b.c in Figure 3). Besides paying close attention to security issues, this enables us to have management access to the authentication services even if the outer network is down. In order to have access to the running authentication server in case of maintenance, for example to upgrade the system, without reconfiguring the load balancer each time, we have introduced a unique binary semaphore mechanism. This mechanism allows us to disable the customers' connections via the load balancer on the fly and thus being able to keep a minimum of one server in the running state. For the load balancer, this takes the service itself down and therefore clients will automatically be switched to the other system, but the system itself is up and running for our maintenance purposes.

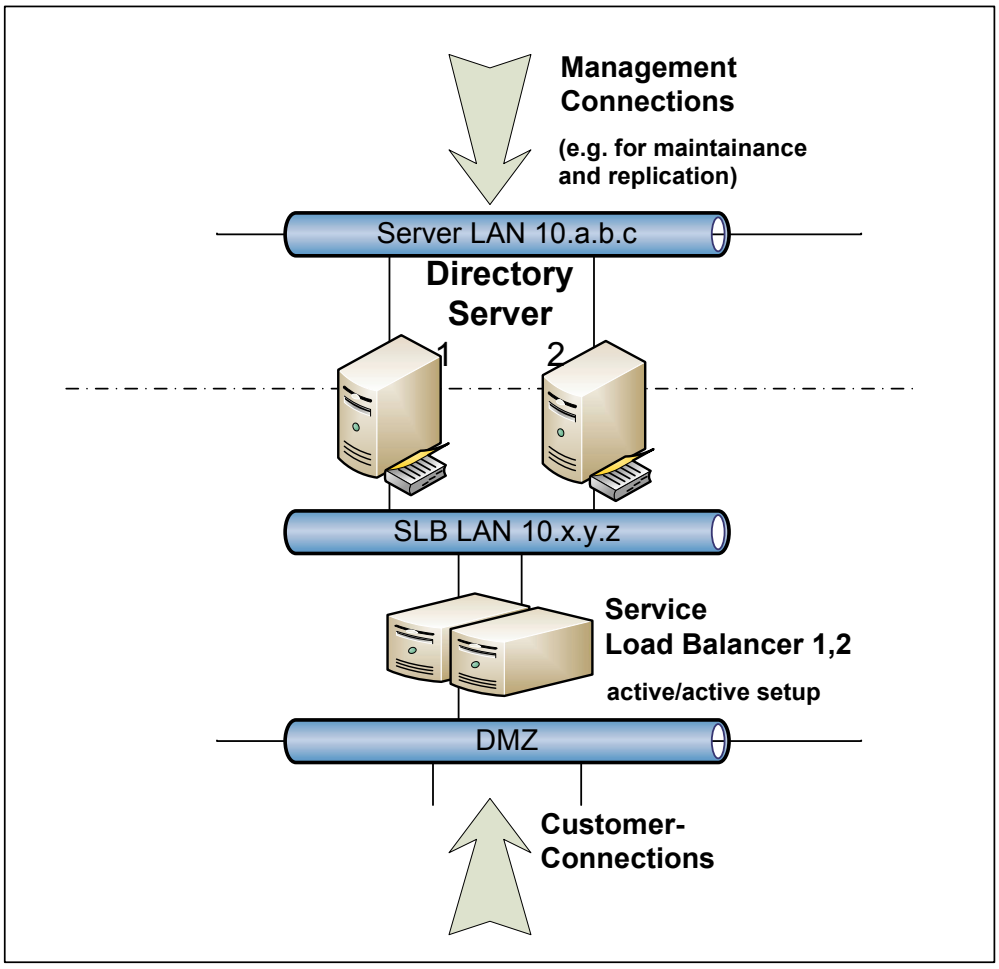

Figure 3: Availability management approach for authentication services

Having implemented these methods we are reaping the following business benefits:

- **Additional abstraction layer:** Both, the redundancy of the services and the access to the service, are hidden from the customers thanks to our load balancing infrastructure.
- **Independency:** Continuous performance and availability because no downtime is required any longer for all day standard and security maintenance, e.g. patch management.
- **Increased performance:** The response times of our service have been improved due to the load balancer and the redundancy of the services.

- **Increased Scalability:** Regarding the capacity management any further extensions (e.g. connecting additional server systems) are supported without the need to interrupt the service.
- **Security:** The load balancers additionally work as packet filtering engines that limit the network connections effectively.
- **Flexibility:** Relocating the service can be managed in the background, which means completely transparent for our customers.

Since both services, authentication services and meta-directory, need to be available as steady as possible even if there is a requirement to maintain it because of new released security updates or the upgrade to a new version, techniques are needed to meet this conflicting requirements. This is why we also have chosen to move our services to virtual machines. Herewith it is possible to maintain most of its system components during operation time. To assist the implementation of our Continuity management needs, we have implemented further activities, like various levels of emergency power supply. For the sake of brevity, we do not discuss this here any further. The results of our activities in Availability management, the resulting CFIA, can be seen in Table 2.

Configuration Item	Service enrolment via TUMonline	
	Previous Service CFIA state	**New Service** CFIA state
TUMonline	**X**	**X**
Data synchronization TUMonline → Meta-directory	**X**	**X**
Meta-directory Service	**X**	**A**
Authentication Service	**X**	**A**
Service CFIA state '**X**': Failure of CI causes Service outage Service CFIA state '**A**': Alternative CI available to provide Service Service CFIA state '**B**': Alternative CI available, but recover actions needed first		

Table 2 CFIA comparison before and after our implementation

Compared to the former infrastructure design and implementation all services operated by LRZ (see the yellow rows in Table 2) have been optimized regarding Availability and Continuity management. Additionally all these services have been integrated in a comprehensive monitoring infrastructure. This includes, for example, the control of disk space thresholds with automatic alarm mechanism and monitoring of network usage. To ensure the availability of our services even in the after regular office hours, we have implemented an automatic start mechanism, which starts our services automatically if a downtime has been detected by the monitoring system after e.g. a loss of power.

4. CONCLUSION AND OUTLOOK

The business alignment of IT is becoming vital also for higher education institutions. The deployment of a new campus management system at TUM and the coupling of this system with the Identity and Access Management solution provided by LRZ increased the dependency of strategic administrative processes at TUM, such as online student enrolment, on its IT components. This dependency resulted in the need for a professional Availability and Continuity management. In this article we have introduced our Availability and Continuity management which we have implemented aligned to the ISO/IEC 20000 standard. Our resulting concept includes planning activities, technical implementations, and monitoring measurements.

To better support the management in the future, we are proceeding with the introduction of Service Level Agreements between TUM and LRZ. By the time this article was written, we have already gained experiences how to integrate SLA details into TUM's IT service management tool. This will help us in future to be able to monitor the adherence of the SLA itself. SLAs are providing vital information for all stakeholders in IT Service Management. TUM's Service Desk will be able to better

prioritize incidents; it will be easier to manage the expectations of the organizations' members when SLAs are in place and IT service providers are having better means to plan their services regarding the provided availability. In order to better assist ITSM in future, methods related to the ones we discussed, such as Component Failure Impact Analysis or the automated impact analysis that have been proposed in (Schmitz, D. (2008)), obviously need better tool support. Therefore we are planning to implement a Configuration Management Database (CDMB) in each organization, which also is able to retrieve relevant information from the other side. We called the concept of a CMDB assisting inter-organizational IT Service Management inter-organizational CMDB (ioCMDB).

Acknowledgment

This work has been partly funded by the German Research Foundation (DFG) under contract WGI 554 975.

The authors thank the members of the IntegraTUM project team and the Munich Network Management Team for valuable comments on previous versions of this paper. IntegraTUM is headed by Prof. Dr. Arndt Bode. The MNM-Team, directed by Prof. Dr. Heinz-Gerd Hegering and Prof. Dr. Dieter Kranzlmüller, is a group of researchers from the University of Munich, the Technische Universität München, the German Federal Armed Forces University in Munich, and the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences. The web server of the MNM Team is located at http://www.mnm-team.org/.

5. REFERENCES

Boursas, L., Hommel, W. (2006), Efficient Technical and Organizational Measures for Privacy–aware Campus Identity Management and Service Integration, In *12th International Conference of European University Information Systems (EUNIS 2006)*, Tartu, Estonia, June, 2006.

Haselbacher, Franz (2003). The TUGonline Project. Retrieved May 19, 2009 from: http://www.eunis.org/activities/b_practices/award/Graz.pdf.

Hommel, W., Knittl, S. (2007). SERVUS@TUM: User–Centric IT Service Support and Privacy Management , In *13th International Conference of European University Information Systems (EUNIS 2007)*, Grenoble, France, June, 2007.

Hommel, W., Knittl, S. (2008). An Access Control Solution For The Inter–Organizational Use Of ITIL Federated Configuration Management Databases, In *2008 Workshop of HP Software University Association (HP–SUA)*, Infonomics–Consulting, Hewlett–Packard, Marrakech, Morocco, June, 2008.

Hommel, W., Knittl, S. and Pluta, D. (2008). *Stragegy and Tools for Identity Management and Process Integration in the Munich Scientific Network*. In Proceedings of the 14th International Conference on European University Information Systems (EUNIS 2008), Århus, Denmark, June 2008.

ISO20000-1 (2005). ISO20000 Part 1:2005: Information technology - Service management – Specification.

ISO20000-2 (2005). ISO20000 Part 2:2005: Information technology - Service management - Code of practice.

Schmitz, D. (2008). Automated Service–Oriented Impact Analysis and Recovery Alternative Selection, Ludwig–Maximilians–Universität München, Phd thesis, July.

TUM Website (2009). CM@TUM - Einführung des Campus Management Systems TUMonline. Retrieved May 20, 2009 from http://portal.mytum.de/iuk/cm/index_html.