Specification of a Service Management Architecture to Run Distributed and Networked Systems

C. Mayerl, Z. Nochta, M. Müller, M. Schauer, A. Uremovic, and S. Abeck

Cooperation and Management IT Research
Institute for Telematics
University of Karlsruhe (TH)
Zirkel 2, D-76128 Karlsruhe Germany
Phone: +49-721-608-6390, Fax: +49-721-608-4046
email:{mayerl|nochta|abeck}@cooperation-management.de

Abstract. Due to the complexity of running distributed and networked systems, we can recognise a trend of virtual re-centralisation of distributed systems. Since outsourcing IT resources, IT service providers are born. The aim of service providers is to run IT components offering qualified IT services to customers, who specify quality demands as service level agreements (SLAs). Service providers use management tools to assure the quality of IT services and to run IT effectively and efficiently. The management tools available today do not meet all demands of IT service providers fulfilling service level agreements customers. This paper specifies an integrated service management architecture supporting IT service provider to assure the quality of IT services.

Introduction

The migration of IT operators towards IT service providers shifts the focus from IT components to the needs of customers. Customers are interested in IT services and their quality properties, but they are no longer interested in fact whether systems are running on a mainframe machine or a toaster. Buzz words like *IT outsourcing* or *application hosting* describe this trend [Fis99, Pau99]. Service providers have to integrate network, system and application components delivered by developers into a networked system. Examples for components are business and office software running on workstations or mainframes connected by network components like hubs, routers and switches [HAN99]. There are also developers delivering management tools to support an operation of distributed systems [Slo94]. Customers focus on daily needs and choose services offered by providers. They need communication and application services to process business critical information. Customers also need support services in order to use IT services correctly. The quality demands of customers with regard to IT services are specified in service level agreements (SLAs). The aim of a provider is to assure agreed service levels instead of paying penalties. Figure 1 shows a scenario of IT service provision.
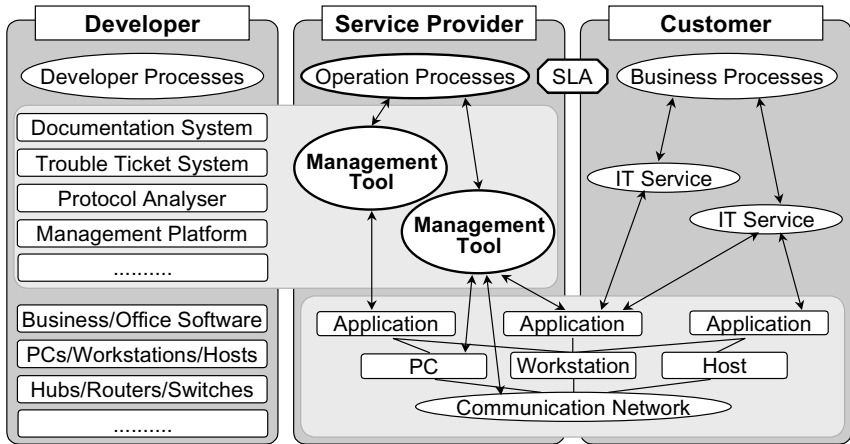
Figure 1: Service level agreement (SLA) between provider and customer

To be effective and efficient, IT service providers use computer based management tools to monitor and control IT components. Although IT management tools like Hewlett Packard Open View, Cabletron Spectrum, Tivoli NetView provide a necessary technical view on distributed components, they do not satisfy the demands of an IT service provider which has to fulfil SLAs [Ghe97]. For example, available management tools do not support co-operation and co-ordination of distributed operation personnel, which is necessary to fulfil quality demands.

Properties of IT Services

For each IT service property customers can specify an individual level determining the fee they have to pay for using that service. There are two kinds of services offered by a service provider, as shown in Figure 2: **Main services** are application and communication products, such as SAP R/3, standard office tools or me-too-products. They comprise functionality provided by network, system and application components implemented as hardware and/or software. When providing services, an integration of used network, system and application components is important. **Support services** are services within, for example, customer management, problem management [CCT95] and change management [CCT98]. They support the use of main services and increase their quality. On the one hand support services can be ordered implicitly, for example by specifying reaction time for a breakdown of SAP R/3 services. On the other hand they can be offered explicitly, for example, help desk service supporting the use of main services [Wal99]. Customers specify properties, such as availability of support teams, reaction and repair times and frozen zones for changes of the system. Figure 2 shows the coherence between the service properties from customers' viewpoint and the underlying technical and organisational resources.
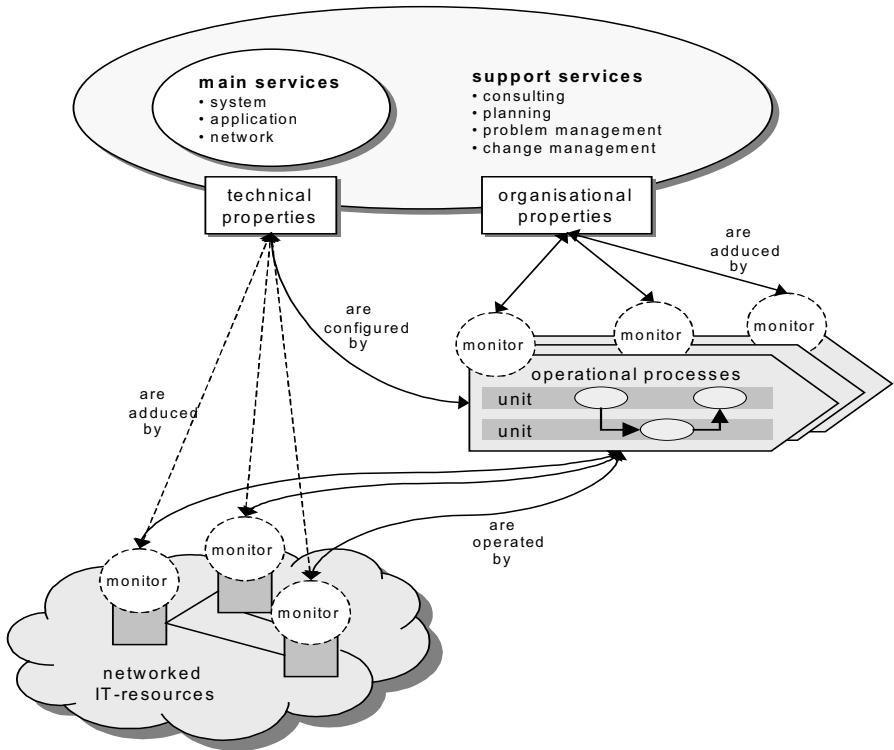
Figure 2: Services and their technical and organisational properties

As mentioned above main services provide technical SLAs adduced by monitored IT resources. Beyond this, technical properties are configured in the context of IT operations management processes. The organisational SLAs of support services are adduced by organisational units executing operational processes. The service model simplifies the quality management of services by structuring and categorising resources and their properties. Within one category a service includes characteristic quality parameters (SLA-parameter) based on IT or organisational resources. Properties of resources define the quality of service as seen from the customers' viewpoint and are the basis of SLAs. A prerequisite for guaranteeing a property of a service is to assure the properties of technical and organisational components. For an evidence of fulfilling SLAs, properties of resources have to be measured and correlated to a service [YKMY96]. Resource properties that cannot be measured and assigned to a service and to an individual customer cannot be assured as a quality of service attribute specified in SLAs.

Main services can cover a range from central data centre operations to desktop services. These services can be differentiated by separate modules.
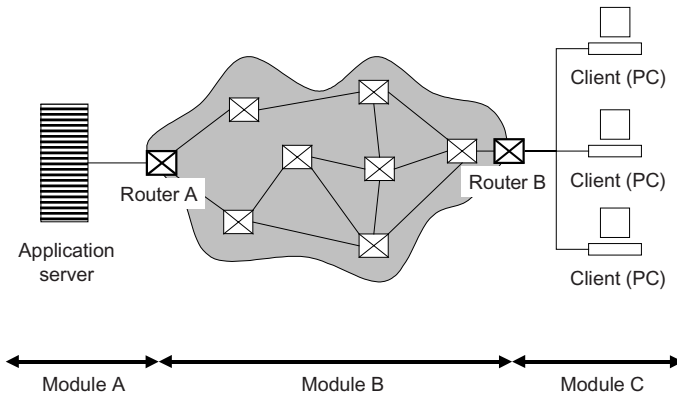
Figure 3: Modules of a main service

Figure 3 explains technical modules of a main service: Module A covers central application components including hardware up to a central router entry point (Router A). Module B contains wide area network (WAN) resources and comprises network resources from the central router (Router A) to an entry point at the local customer's router (Router B). Module C covers resources from the local router up to the desktop equipment on an end-user's desk. Each module has characteristic technical properties within its area of expertise. Technical properties are represented by SLA-parameters, for example availability or response time of an application server, throughput of WAN components, and throughput of local area network (LAN) components. Definition of these SLA-parameters and their measurement checkpoints depends on allocated modules. Important checkpoints are service access points between different modules such as Router A and Router B.

A specification of SLAs for support services can be based on organisational or functional modules [CCT95]. Modules differ from each other by e.g. employee qualification. End-user support provides technical support for registered users, and it is performed by three functional organisation units: first-level support to manage customers' requests, second-level support to solve problems of module A and B professionally, and location support to maintain components in module C. These units handle incoming and outgoing service requests. The first level support acts as a single point of contact for end-users. Important properties of this unit are: Availability of support staff within service hours and resolution time of problems. The second-level support is responsible for the final resolution of service requests which are received from the first-level support. Main tasks of this unit are: solving problems, documenting actions, informing first-level support and documenting the result of actions. SLAs like second-level average problem, request and inquiry resolution time determine the quality of this service. In location support, organisational units deliver onsite support and services to a customer in their local computing environment. There are SLA-parameters, like response times and average resolution time for problems detected in module C.

Specification of Provider Accepted Management Tools

SLAs specify technical and organisational quality demands on services offered by an IT provider. To assure the quality of services, it is necessary to take a closer look at the operational processes [AM97, AM99]. Assuming that a provider guarantees a service level, for example, availability of 98% of SAP R/3 system, it is not enough to run highly available IT resources. It is necessary to use accurate service operation processes, too. An important process to enable a high availability is the problem management process. Providers must react fast to calls by customers and should be able to solve problems in a service level oriented manner, so that guaranteed service levels will not be violated. Therefore, operating processes are often more critical to SLAs than IT resources. Processes have to be modelled in order to understand and implement them. We assume that agreed services are mapped on operation processes and processes realised by process oriented management tools.
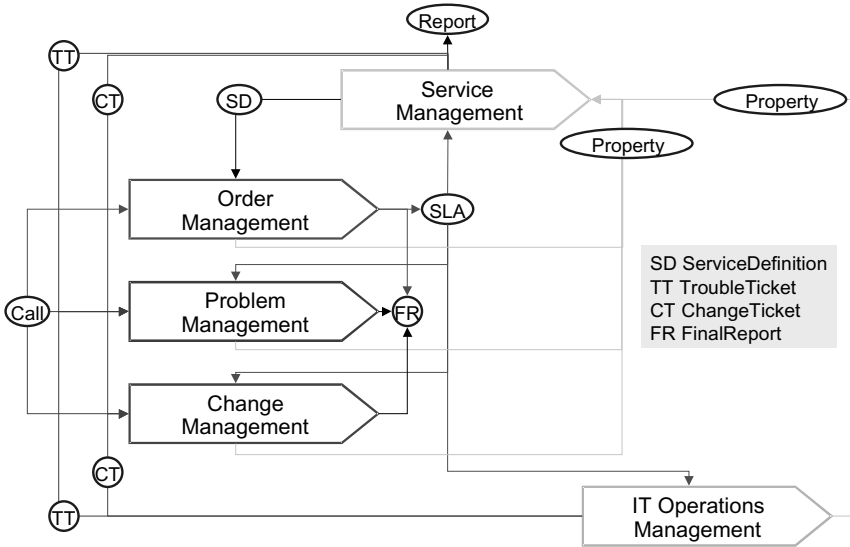

Figure 4: IT operation process model

Figure 4 models the most important operation processes. Service management process defines services and their components, and reports SLAs. Order management process uses the service structure to co-ordinate service levels with customers. SLAs are inputs of all other operation processes. Problem management handles unplanned incidences endangering SLAs [CCT95]. Change management plans and co-ordinates changes on IT resources [CCT98]. IT operations management processes contain tasks to configure, monitor and control components of a networked system.

In order to describe provider demands on management tools, we have introduced an abstraction layer in our modelling scheme which describes demands on tools in an abstract way. We speak of Process Orientated Management tools (POMs) [AM99]. The POM-layer meets the demands of providers as well as those of software developers. A detailed process description and deficiencies identified are used as catalysts to investigate what kind of (tool) support a provider needs to improve the

way service is delivered. Hence, consultants and operative staff elaborate in workshops the appropriate data and functionality required to operate efficiently. Depending on the type of support, different POM types can be deployed [MSS99]. In order to improve communication capabilities among process participants, horizontal communication POMs are applied. Horizontal communication can be either directed in terms of pre-defined rules or spontaneously used (non-directed) to get additional information needed to proceed. Besides horizontal communication we propose the use of vertical communication POMs to monitor, control and configure the underlying provider's IT infrastructure. Operative staff can deploy execution POMs if there is a need to compute a result based on input parameters. In principle, POMs encapsulate data and functionality with respect to the context of a process it is used in. It can be described in a pictorial way, such as a form, so it could be used to hold discussions with providers. Described in a object-orientated way, for example, using the Unified Modeling Language (UML), POMs could be used as specification for software developers [Bur97].

Integrated Service Management Architecture

Following methods like PRODEM [MSS99] IT service providers can specify tools to manage a networked system. An investigation of provider demands results in an integrated service management architecture (SMA). This section specifies the most important building blocks of SMA and the relationships between them. Figure 5 shows an overview of SMA.
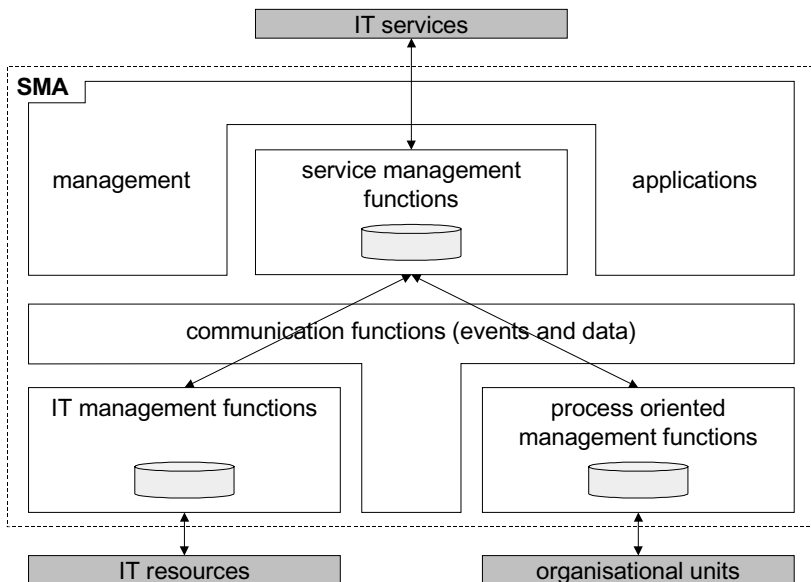
Figure 5: Service Management Architecture (SMA)

SMA integrates IT management functions, process oriented management functions and service management functions. Management relevant information is communicated between building blocks and processed by different management applications. The following sections describe different models as part of SMA extended by process management and service management aspects. These models detail building blocks shown in figure 5.

Information Model

Merely fulfilling SLAs, management information about status and properties of networked components is not enough. Therefore, an existing information model of IT management has to be extended by information about services and customers using services and also information about operation processes and the co-operation of organisational units.
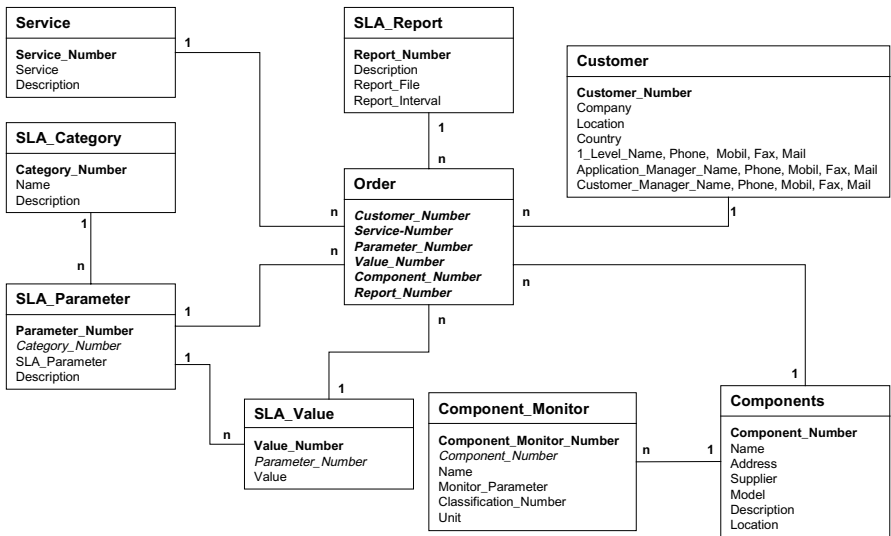

Figure 6: Service management information

Figure 6 shows information about service management, such as information about customers, services offered by an IT service provider, service levels (SLA_Value) for each property agreed by the provider, technical and organisational components as part of a service, measurement algorithms, and parameters to report SLAs (Component_Monitor). This model describes a networked system and its operation processes from the customers' viewpoint. To support an internal view of IT service providers, the information model is extended by **process-oriented management** information like availability of organisational units and their employees, qualification of organisational units, process definition driven by SLAs, process status, time stamps describing reaction, repair and work time etc.

Functional Model

Traditional IT management focuses on the fault, configuration, accounting performance, and security functions of IT resources. In order to meet customers' demands SMA integrates functions of service management to define a new service bundling of technical and organisational modules to a service. SMA also needs functions to manage customers' individual SLAs, to drill down SLAs to properties of components, to monitor technical and organisational components and to correlate measured information to a service. Finally, there must be functions to report and demonstrate the service quality.

Running a networked system effectively and efficiently with regard to SLAs, SMA integrates management functions to guide the execution of operational processes like problem management and change management and to manage involved organisational units. SMA triggers organisational units to act step by step without losing the flexibility to handle undefined exceptions, for example in the context of problem management. This process-oriented management concept is called the guided co-operation concept (GCC) [May98]. GCC describes processes, defines synchronisation points and functions to store, transport and process information. It guarantees also flexibility to handle exceptions. Figure 7 illustrates functions to report SLAs. IT management platforms, such as Tivoli NetView monitor networked systems. Process-oriented management tools e.g. Remedy's Action Request System, and also telecommunication switches including automatic call distribution (ACD) report information about properties of organisational units, such as support teams.
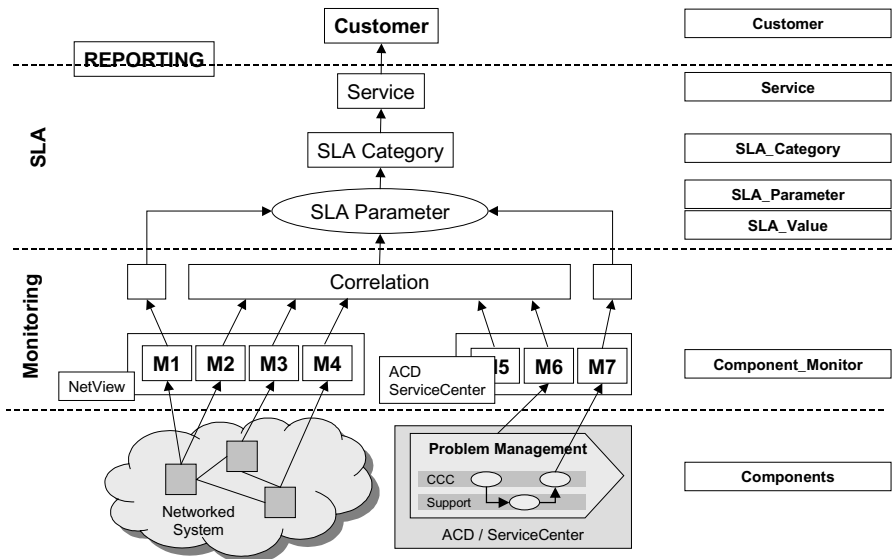
Figure 7: Functions to report SLA fulfilment

Information supplied by monitors depends on technical and organisational components. Therefore the next architecture layer correlates the different monitor information. The correlation can be done by logical rules specifying conditions.

Correlated data can be assigned to service categories described above, to services as well as to customers. Reporting functions support the configuration and illustration of SLA reports.

Organisational Model

IT management standards define roles, such as manager and agents of IT resources. In order to execute operational processes fulfilling SLAs it is necessary to define processes like problem management and change management and their tasks. Roles have to be specified by a service provider to execute tasks of operational processes. Defined roles should co-operate with each other. A role specifies a domain of tasks as part of a process and is occupied by an organisational unit at the execution time of a process. To be more efficient, organisational units need management tools to execute tasks and to co-operate with other units in a co-ordinated way. Service management specifies roles, like customer using services, suppliers of other services, like telecommunication providers, and also internal roles, for example process owners and operational roles. In the context of the problem management process an organisational model defines roles and their co-operation to execute operational processes running a networked system. These roles are: process owner (problem manager), first-level support (help desk or call center), competence centres for module A, B and C (second-level and local support). The co-operation between these roles has to be co-ordinated by defined communication protocols.

Communication Model

The vertical communication model between network, system and application components and roles executing operation processes is supported by IT management functions and protocols like SNMP or CMIP. These functions have to be extended by a horizontal communication between specified roles and organisational units. The horizontal communication model specifies how to transport management information between roles and when a role has to execute the next step. Therefore, the communication model defines communication mediums and functions like phone, mail, email etc. which should be used to co-operate. To trigger a role to do the next step, the SMA interacts with human beings using interaction functions of SMA. These functions can be supported by tools such as personal digital assistants or mobile phones using short message service (SMS). Information to be communicated can be differentiated as data or events. Data is transported without semantic changes to be processed by another role. Events describe a defined status, for example, of a process and trigger a next step processing received data.

Implementation Experiences and Outlook

This section describe a prototype implementing the SMA specified above [Sch00]. Due to heterogeneous networked systems and individual demands on IT service providers, there is no single general-purpose management tool which can support all

service providers. SMA integrates different management tools to manage main services and operational processes. Managed objects are: application system (SAP R/3), network (WAN and LAN components), and server (Unix) as well as workstations (Windows NT). These resources comprise a main service with modules A, B and C as described above. We use Tivoli products providing network-, systems- and application- management in order to avoid many integration problems. These management tools collect raw technical data for SLA monitoring and reporting as well as daily system administration purposes. The management tools we used to implement SMA are: Tivoli Distributed Monitoring (DM), Tivoli NetView, and Tivoli Manager for R/3. DM manages end-user systems such as workstations and file servers. Management with DM requires demons running on the local systems which only monitor selected system properties like CPU usage, collect management data and implement a communication interface to management systems. NetView is a comprehensive management tool for heterogeneous devices on distributed networks.
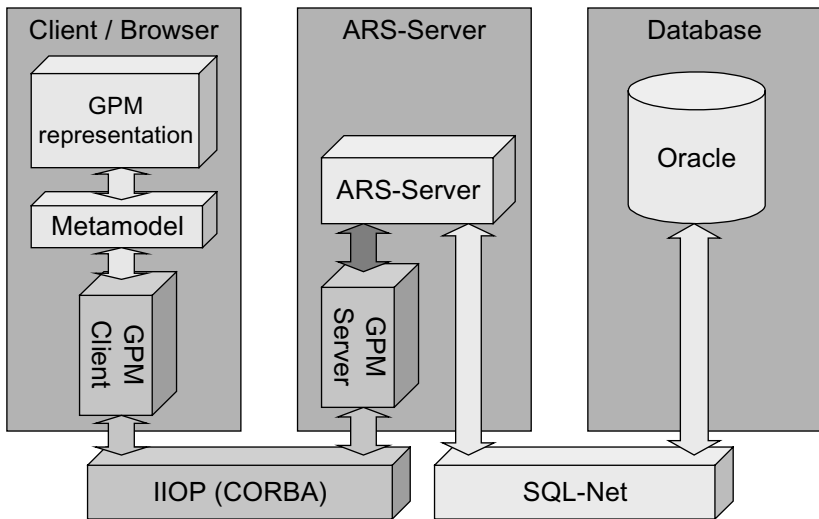

Figure 8: Architecture of the Generic Process Monitor

The Generic Process Monitor (GPM) we implemented monitors processes realised by process-oriented management tools like trouble ticket systems [Ben99, Aug00] such as Remedy ARS (Action Request System). Figure 8 shows the main blocks of the GPM architecture including ARS components. The main goal of GPM is to collect and visualise information about operational processes. GPM is based on CORBA. Central module of GPM is the GPM-server which is responsible for getting information from the monitored process oriented management tool dynamically. The GPM-Client consists of a representation layer to show management data on a web browser and a CORBA client interface. ARS can map IT problem-management processes into so-called forms. ARS is based on a client-server architecture that distributes the forms under predefined rules to according personnel. All information about rules, forms, sequences etc. are stored in a relational database, for example Oracle. In this scenario GPM server uses the open C-API of ARS to monitor the forms. At the moment the GPM prototype monitors ARS.

Transforming raw management data into service- and customer-oriented information is one of the major functions of SMA [MU99]. Customer-oriented information means well defined and formatted reports for high and midlevel management personnel controlling SLA parameters. High-level managers typically need long term reports (e.g. annual reports) to make strategic decisions unlike mid-level managers, who should get short term reports (daily, weekly) to control quality of daily business functions.

In our implementation we use InfoVista providing graphical reports in different formats. InfoVista has a Perl-API that makes it possible to generate reports. Since we use NetView to monitor and administer the network we renounce the network collector functionality of InfoVista. The prototype integrates the reporting functionality of InfoVista into SMA as shown in Figure 9.
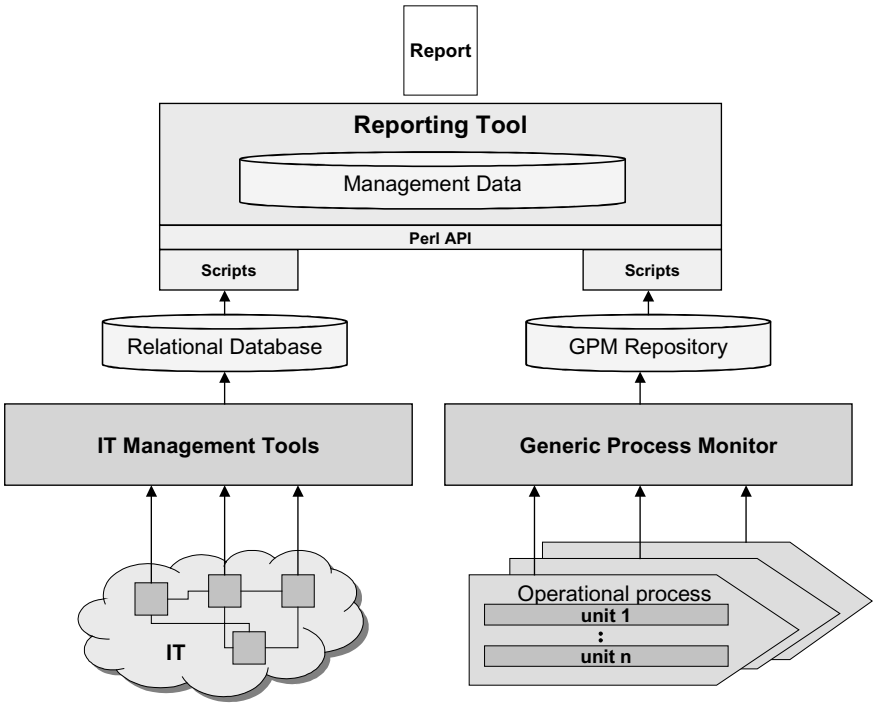
Figure 9: Service level reporting of IT management

The IT management tools together with GPM monitor the networked system and the operational processes. Data collected by GPM are stored in its object-oriented repository. Technical data collected by the Tivoli management tools are stored in a relational database. Presently, the GPM repository includes metadata information collected from ARS. In the future, we will develop a generic repository served by a model engine which is under construction. This engine will enable to convert and store data from other process oriented management tools.

References

[AM97]   Abeck, S.; Mayerl, C., 1997, "Prozeßbeschreibungen als Basis für einen qualitätsgesicherten Betrieb von vernetzten Arbeitsplatzrechnern", APS'97, Koblenz, Germany.

[AM99]   Abeck, S.; Mayerl, C., 1999, "Modeling IT Operations to Derive Provider Accepted Management Tools", IM'99, Boston, USA.

[Aug00]   Augustin, G., 2000, "Generic Process Monitor (GPM)", University of Karlsruhe.

[Ben99]   Bender, P., 1999, "Implementation of a Process Monitor using Java and CORBA", University of Karlsruhe.

[Bur97]   Burkhardt, R., 1997, "UML – Unified Modeling Language", Addison-Wesley.

[CCT95]   Central Computer and Telecommunications Agency (CCTA), 1995, "Problem Management", London.

[CCT98]   Central Computer and Telecommunications Agency (CCTA), 1998, "Change Management", London.

[Fis99]   Fischer, J., 1999, "Outsourcing reaches desktop", Computer Zeitung, Volume 11, Germany.

[Ghe97]   Ghetie, I. G., 1997, "Networks and Systems Management – Platforms Analysis and Evaluation", Kluwer Academic Publishers, Boston, London, Dordrecht.

[HAN99]   Hegering, H.-G.; Abeck, S.; Neumair, B., 1999, "Integrated Management of Networked Systems", Morgan Kaufmann Publisher.

[May98]   Mayerl, C., 1998, "Process-oriented Approach to Develop Provider Accepted Management Tools", EUNICE'98, Munich, Germany.

[MSS99]   Mayerl, C.; Schäfer, A.; Schubert, L., 1999, "PRODEM: Method to Specify Service Provider Oriented Management Tools", MobIS'99, Bamberg, Germany.

[MU99]   Müller, M.; Uremovic, A, "Modeling and Implementation of a Service Management Environment to Report Service Quality", University of Karlsruhe.

[Pau99]   Paul, M., 1999, "Outsourcing of IT services at universities", Praxis der Informationsverarbeitung und Kommunikation, Volume 2, Germany.

[Sch00]   Schauer, Martin, 2000, "Specification of a Service Management Architecture to Provide Quality of IT Services", diploma thesis, University of Karlsruhe.

[Slo94]   Sloman, M., 1994, "Distributed Systems Management", Addison-Wesley.

[Wal99]   Wald, E., 1999, "Help Desk Management", MITP.

[YKMY96]   Yemini, S.; Kliger, S.; Mozes, E.; Yemini, Y., 1996, "High Speed and Robust Event Correlation", IEEE Communication Magazine.