Featuring workflow management

Featuring Workflow Management

An overview of the distinctive features of workflow processes and their consequences for workflow management

Gido A.J.F. Brouns

Eindhoven University of Technology
Department of Mathematics and Computing Science
Eindhoven, The Netherlands

March 15, 2000

Abstract

Workflow Management is a rich and diverse technology that is now being applied across an ever-increasing number of industries. The competitive advantages of applying workflow technology are beginning to emerge as organizations focus on restructuring and optimizing their business processes.

Workflow management has developed into a field of research that is characterized by input from many disciplines, including formal methods, Information Technology (IT), Operations Research (OR), organizational theory, user-machine interaction and office logistics. Still, despite major efforts on the part of the software industry, many fundamental problems remain unsolved. These problems are not primarily focused on the software, but rather on the underlying concepts and techniques.

We explain the basics and characteristics of workflow management and *workflow management systems* and illustrate the importance of effective workflow management in present-day operational management. Furthermore, the relationship between workflow management and the concept of BPR is addressed.

We identify the crucial operational issues workflow management seeks solutions for, and point out that workflow processes exhibit specific features that distinguish them from business processes in general. Special attention is paid to the differences between workflow processes and manufacturing processes, which form the traditional field of study of business processes. In this context, the concept of *quality of service* is introduced.

It turns out that the differences are significant to such an extent that it is inconceivable that industrial companies can have their ERP system — the system that controls the manufacturing process — manage their internal administrative processes as well.

However, the current generation of workflow management systems does not provide suitable or sufficient functionality either to satisfactorily manage the special characteristics of workflow processes. There remains a clear need for intelligent mechanisms for workflow process control.

In this context, some generic rules are suggested as regards the design and control of workflow processes.

Key words: workflow processes, workflow management (WfM), workflow management systems (WfMSs), quality of service (QoS).

1 Introduction

Workflow Management (WfM), often just denoted by the term ‘workflow’, promises a new solution to an age-old problem: managing and supporting business processes. What is new about workflow is the way it employs the power of Information Technology (IT) to support structured work.

Although large paperwork offices and organizations have existed for many decades, workflow has only recently gained tremendously in interest. Two main reasons can be identified. First, in the view of many organizations, especially industrial ones, workflow had never been something one could ‘score on’. Traditionally, the production system took care of ‘bringing in the cash’. The administrative part of the manufacturing process was seen as cumbersome and unchangeable. Of course, it had to be taken care of, but that was about as far as the commitment should stretch.

Recently, mainly because of fierce competition, many large industrial companies have been forced to cut their expenditures. On the one hand, various decision systems for inventory management and production planning assist them in optimizing their manufacturing processes by minimizing inventory costs and maximizing output — taking into account the restrictions imposed on the production system. Also, a common action to defy fierce competition is cutting profit margins on finished products. Furthermore, costs can be reduced by laying off administrative staff. Although these last two possible courses of action could be advantageous in the near future, they cannot be defended on a long-term basis. In general, marginal profit margins and an eroded middle management constitute a serious impediment to company growth.

Yet, on the other hand, more and more companies are starting to realize that being in absolute control of the administrative part of the organization could not only be a key factor in the pursuit of cost reduction, but that this could also pave the way for a more effective company as a whole. Administrative processes are indeed makable and changeable, and at the heart of effective control lies the effective engagement and use of company personnel. In pure paperwork organizations, the various administrative processes form the ‘production system’ themselves, but it is important to note that even the vast majority of these organizations tended to take the actual way of controlling their internal processes for granted.

Second, the recent and ongoing unprecedented developments in the IT sector — the well-known IT boom — have made it possible to build sophisticated information systems that are capable of regulating and controlling the division and execution of work-in-process (WIP) and future work. The strength of IT is that it serves both as a technology push and as a technology pull. The information systems used in workflow are commonly known as *workflow management systems*. Some examples of software packages currently on the market are ActionWorkflow, COSA, CSE/Workflow, FlowMark, SAP Business Workflow and Staffware. The wide range of products shows that the software industry has also discovered the emerging field of workflow management.

To conclude, administrative as well as industrial organizations are increasingly starting to recognize the potential effects a full understanding and control of the internal administrative processes can have on the overall performance of the organization. It might seem a bold statement, but it is not unlikely that for companies operating in highly competitive markets, it could be the case that the company which does not manage its administration effectively in the not too distant future, will be ousted by competitors that did manage.

In this paper, we will focus on pure paperwork organizations or offices. Paperwork offices include the administrative departments of industrial organizations. Therefore, all the concepts to be discussed are also directly applicable to workflow management in industrial environments. Before we explore further the concept of ‘workflow’, we will briefly discuss the relationship between workflow management and the popular concept of BPR.

1.1 Reciprocity between workflow and BPR

Workflow is often associated with *Business Process Re-engineering* (BPR), the term that is widely used to indicate the field of improving and optimizing existing business processes. Most of the theoretical contributions to this field originate from outside the academic world and concern the discussion of specific re-engineering projects carried out for commercial companies.

Within the field, there remains a clear challenge to identify a well-formulated methodology that delivers measurable results. Almost every organization may benefit from such a methodology to reduce costs, increase flexibility and manage its business processes. An appropriate level of abstraction is the *structure* of the business process.

Here we identify the smaller steps within the process, as well as the relations between these steps. The content of a step is only of importance insofar as it influences the ‘improvement driver’ in a specific case. Within this methodology, it is not important *how* a step is carried out, but *that* it is carried out. Similarly, *how* the decision that the step should be carried out was taken is not important, but *that* the decision was taken. The configuration of these steps is the subject of the re-engineering effort.

One of the advantages of the abstraction of the work content is that many mathematical theories become available. More specifically, graph theory, queueing theory and process algebra are typical examples of suitable frameworks. The steps or activities can be treated as black boxes, represented by nodes or servers within a graph or network. As we will see, similar concepts arise in the modelling and analysis of workflows. We will also see that workflows possess typical characteristics that distinguish them from general business processes, and especially from traditionally studied manufacturing processes.

Furthermore, it is said that for many administrative processes, the development of workflow management systems will be an essential enabler for BPR efforts, since such systems create the opportunity to easily adjust existing processes. The introduction of a workflow management system usually also gives rise to a completely different operating procedure. Conversely, it can be said that specific BPR efforts will, either directly or indirectly, result in the acquisition of a workflow management system.

2 Exploring the concept of workflow

The term *workflow* refers to the automation of business processes that involve:

- the processing of *cases* and the execution of *tasks*,
- in a particular order,
- by particular *resources*,
- so that some *objective* is met.

2.1 Cases and tasks

Whenever work arrives from outside the organization, in terms of or as a result of an order, a request or a compulsory return, a *case* is opened. In workflow terminology, we speak of the 'arrival of process instances', where a *process instance* is the representation of a single enactment of the process, the epitomization of an individual case. A process instance is created, managed and eventually terminated by the workflow management system that supports the control of the workflow. We refer to the section on workflow management systems for an elaboration on the functionality and the characteristics of these systems.

Every process instance corresponds to a certain case, where every case consists of a list of one or more predefined *tasks*. Tasks are logical, indivisible units of work, which are either executed completely or not at all. Tasks do not correspond to individual cases. Instead, we use the term *activity* to indicate the execution of a task for a specific case.

An activity concerns the manual, automated or otherwise semi-automated execution of a specific task, where the last mentioned way of executing tasks is a combination of the first two ways. Process instances that only concern fully automated tasks are also termed *workflow instances* in literature.

Usually, the amount of work-in-process varies heavily over time, and the moments at which process instances arrive are usually so random that the arrival process of work can be described as a Poisson process. Furthermore, the nature and extent of arriving process instances differ immensely. Hence, three sources of uncertainty that can be identified are (1) the number of process instances that have to be processed in a given period, (2) the nature of these process instances, and (3) the amount of work incorporated into each of these process instances.

We note that one could also think of processes in which work arrives in *batches*. Then, upon arrival of such a batch, the number of process instances that have to be processed in the immediate future is known. However, if the interarrival times of batches are not fixed, the first source of uncertainty may still apply.

2.2 Resources and resource classification

A task is executed by a *resource* — and by one resource only. A resource is any entity (any means of production) that is capable of performing certain tasks. To every resource corresponds a finite set of tasks the resource can execute. Based on this principle, resources are classified into *resource classes*. Concerning a specific set of tasks, the resource class linked to this set is, by definition, an exhaustive set of resources that are capable of performing each of the tasks belonging to the considered set of tasks. Note that although a particular resource always fits into some resource class, it can indeed also belong to two or more resource classes. Also, at any moment in time, it is quite possible that a resource is busy executing more than one task simultaneously — if it is capable of doing so. Clearly, while discharging its duties, the capacity of the resource may not be exceeded.

Various types of resources exist. We can, for example, think of human resources, technical devices or machines, and software packages. However, unless otherwise stated, any two resources will be treated as if they displayed no distinctions whatsoever, if they exhibit the same capabilities and functionality — i.e. if they belong to exactly the same resource classes.

Therefore, resources are only identified by a set of required capabilities and a set of functional operations to be provided. In principle, a resource is characterized by two essential facts. First, it is capable of performing certain tasks, as discussed. Second, it might not be available at all times.

2.3 Resource management

Resources need to be properly acquired, deployed, and managed while operating. These activities are known collectively as *resource management*. To be more specific, in paperwork organizations, resource management is successively concerned with:

1. resource selection,
2. resource layout design,
3. resource assignment,
4. resource scheduling,
5. resource control.

2.3.1 Resource selection and layout design

Resource selection and *resource layout design* take place long before the actual business operations get under way. In these phases, decisions are taken in matters like who to hire, what machinery to purchase, where to put desks, and how to arrange and organize working cells. The acquired number and type of resources and the way the physical layout is designed will depend heavily on economic considerations, although other aspects — e.g. ecological, ergonomic and legal considerations — will also have their impact on the final decision taken by management. Allowing for various quantifiable aspects, mathematical analysis can assist in determining suitable configurations with regard to the selection and the layout design of resources.

2.3.2 Resource assignment

The next phase, *resource assignment*, takes place just before the start of the business operations the organization is involved in. Resource assignment is concerned with the question of how to allocate tasks to resources, or vice versa. Usually this assignment is a prerequisite to resource scheduling.

2.3.3 Resource scheduling and control

The last two phases, *resource scheduling* and *resource control*, are active during business operations. However, changes in the business process may require temporary re-activation of one or more of the other phases.

Resource scheduling can be seen as a part of resource control. It is concerned with the assignment of activities to resources in time and space. Resource control itself is concerned with managing resources just before, during and just after the execution of activities, by making sure they do what they are supposed to do by instructing them what activities to perform, where to perform them and possibly how to perform them.

2.4 Resource flexibility

Recalling that there are three sources of uncertainty concerning the delivery of work to the workflow, it is clear that the organization needs some degree of flexibility with respect to the resource capacity to keep the average *throughput time* of process instances (the average time process instances reside in the workflow) within limits. Some possible courses of action include the introduction of overtime and overcapacity, temporary staff employment, and multi-employment and multi-skilling of staff. Despite such solutions, it might still be impossible to have most of the work-in-process finished on time, let alone *all* work.

Actually, this is a day-to-day reality in most organizations. For example, specific crucial resources might be (temporarily) unavailable or not obtainable at all. It is also a common phenomenon that certain information that is needed to start up some activity is unavailable at the time the activity should be carried out. Requested information might be stored in some other department within the organization, or not yet obtained from an external individual or organization.

As a matter of fact, it is estimated (see PLATIER[9]) that as regards process instances passing through real-life workflows, on average over 95% of the throughput time of such a process instance consists of waiting time — i.e. time in which no activities related to the process instance are being carried out — and that, consequently, on average less than 5% of the throughput time of such a process instance consists of processing time.

As a result, other courses of action are sought to reduce the average throughput time of process instances. Reducing this throughput time is of the utmost importance, since work-in-process is directly related to cost, possibly in terms of a loss of revenue. For example, an insurance company that is able to process insurance claims within, say, 24 to 48 hours instead of 48 hours to 1 week, could increase its market share considerably, recovering the capital expenditures amply.

It should be noted that ‘revenue’ does not necessarily refer to financial issues. Passing some deadline in a criminal trial could, for example, mean that the case is to be dismissed, without any conviction.

Also, not meeting a deadline may require a predefined ‘escalation procedure’ to be invoked. In terms of capacity, such a procedure is usually highly demanding.

The amount of work-in-process is also directly related to the number and size of storage facilities needed to store all the files related to the cases in hand. However, the costs of purchasing and possessing these facilities are, in general, just a fraction of the costs of exceeding deadlines.

If overrunning a deadline is heavily penalized, whereas staying below it does not deliver any observable benefits, then the organization will most likely be inclined to express the desire to keep control of the *maximum* throughput time of process instances, instead of the average throughput time. So, wherever we mention the average throughput time, one can also think of the maximum throughput time, depending on the circumstances.

2.4.1 Quality of Service consideration

In general, the available capacity will not be sufficient to execute all tasks down to the last detail. Depending on the progress of the process and the current pressure of work, decisions must be taken on how much capacity to be assigned to the work-in-process.

However, the throughput time of a process instance is usually highly correlated with its *Quality of Service* (QoS), a performance measure that indicates the quality of the actual outcome of a case compared to the most favourable outcome which was achievable in that case. For instance, a taxation office cannot expect to detect large-scale tax fraud without thoroughly examining tax returns. On the other hand, sifting through a tax return does not necessarily lead to the detection of fraud. In other words, an increase in capacity by no means guarantees an increase in QoS, but it *could* very well result in a QoS increase. And QoS is, in its turn, related to revenue.

Therefore, it may be clear that in many purely administrative processes, including tax collection and the administration of justice, not just the throughput time and processing costs of a process instance are of interest, but also (and often especially) the QoS. The *QoS consideration* concerns taking decisions on the progress of a process instance based on the current amount of work-in-process and related (processing) costs on the one hand, and the quality of service and related revenues on the other. Such decisions are taken on-the-fly, so each time the state of the process changes — e.g. because of an arrival of a process instance or the completion of an activity — new decisions need to be taken. In the decision-making process, throughput time, quality of service and costs have to be weighed against each other. Ideally, the workflow management system takes care of this on-line decision making, using advanced techniques for the determination of an appropriate quality of service level. Unfortunately, the current generation of workflow management systems does not offer such functionality at all.

The consideration of QoS is a characteristic aspect of workflow processes. However, many more such aspects that distinguish these processes from general business processes can be identified.

2.5 Characteristics of workflow processes

Workflow processes have some characteristics that distinguish them from general business processes, and from processes in the domain of manufacturing in particular. As stated before, manufacturing processes had traditionally been the spearhead of most industrial companies, so it goes without saying that almost all — sometimes even all — engineering efforts of industrial companies to optimize their internal business processes — whether these efforts have a scientific character or not — had always been aimed only at the manufacturing processes.

Therefore, much attention has been paid to the construction of sophisticated information systems — termed *Enterprise Resource Planning* (ERP) systems — that are capable of supporting and controlling these manufacturing processes. A rather naive way of thinking would be to use this type of system, although slightly modified, for the internal administrative processes as well. For, as will be demonstrated, workflow processes have special characteristics that require the service of an information system which is equipped with special functionality to correctly support and control the process, i.e. require the service of a workflow management system. However, the functionality offered by currently available workflow management systems is not yet satisfactory, as we have already illustrated by means of the QoS aspect.

The most important specific features of workflow processes are the following:

- the complete dependence on clients as regards the supply of work — i.e. an uncontrollable arrival process of process instances,
- the conflict of interest between the organization and its clients receiving services,
- the conflict of interest between the organization and its clients supplying basic information,
- multiple routings and outcomes — i.e. the absence of a fixed routing of process instances through the workflow and the phenomenon that apparently similar cases can have dissimilar outcomes,
- (sometimes) routing flexibility,
- the absence of strict causal connections between capacity and value added service,
- resource flexibility,
- (possibly) the absence of market forces,
- semi-automated execution of certain tasks.

2.5.1 Uncontrollable arrival process

No cases can be stored up, under penalty of extreme throughput times. Therefore, because of random arrivals, running out of work is theoretically possible. In such a case, resources are idle and temporarily needless. Flexibility of resources, i.e. having the option to either lay off resources temporarily or have them work on other type of work that *is* present, is highly desirable in this case. We refer to the subsection on resource flexibility for more on this.

In practice, however, a more common phenomenon is a superfluous supply of process instances. Process instances should not be rejected, on penalty of *zero* quality of service. Zero quality of service corresponds to extremely poor quality of service, and although it results in a temporary alleviation of the workload, it can have devastating consequences in the near future. One can think of negative publicity, legal actions against the organization, and an unintentional invitation to misconduct, which is also known as the 'moral hazard' effect.

For an example of such an invitation, consider a taxation office. For an individual or organization that deliberately returns invalid information with the sole purpose of avoiding taxes, and gets away with it year after year because the corresponding tax returns are not examined in detail by the taxation office, there exists no apparent reason to alter this 'strategy'.

2.5.2 Conflicting interests

Clients that receive service deviating from the predefined specifications will only report this if the difference does not work to their advantage. For example, if some person files a claim with his or her insurance company, involving a certain amount of money, and the insurance company transfers a substantially higher amount to his or her bank account, this person will be disinclined to mention it. If, on the other hand, the insurance company accepts the claim, but transfers a lower amount, the client will contact the company without hesitation to ask why the original amount has not been paid.

Furthermore, suppliers of basic information are more likely to provide specifications which are to their advantage than specifications which are to their disadvantage. Consider, for example, a client filing a claim with his or her insurance company, regarding the loss of some possession. When asked to specify the original value of the lost possession, the client will be more likely to state an amount that is higher than the actual value than an amount that is lower.

Because of these two types of conflict of interest, a painstaking verification of end-information supplied to clients and of basic information received from clients is highly recommended. For each case, it must be determined what the quality of the input data is, what quality risks are involved and to what extent these risks can or shall be reduced by means of checks.

2.5.3 Multiple routings and outcomes

When receiving an order, a request or a return, it is often unclear which standard procedure has to be followed. This becomes clear gradually during the execution of the process. Furthermore, besides the routing, also the eventual outcome of a case is often unknown beforehand. As a result of these uncertainties, it is also unclear what capacity engagement the corresponding process instance requires and which type of resources should be called into service.

Typical examples can be found in the world of criminal justice. Suppose a corpse has been found by the police, and that it is presumed to be a case of murder. Next, suppose the police have apprehended a suspect, who has been taken into custody. The continuation of this case is by no means certain, nor is the outcome of the case. Perhaps it turns out that the suspect has a watertight alibi, forcing the police to release this person and start looking for further clues. Or perhaps the pathologist's report states it was not murder after all, but merely a dreadful accident, or suicide. Then the criminal investigation can be dropped. Note that there does not exist a one-to-one correspondence between the routing and the outcome of the case.

Different outcomes of intervening (whether or not sudden) events and activities lead to different demands for capacity. If, in the example, it did turn out to be murder, but no progress has been made in the investigation for quite some time, the number of detectives on the case will probably be decreased, because of other cases also demanding attention. However, if after a while vital evidence is found after all, the number of detectives will probably be increased again, because now there is an improved chance of finding out the truth.

In manufacturing environments, the course and outcome of the process is fixed and known for every process instance, leaving aside exceptions like losses because of a machine breaking down. Process instances — usually products — have a fixed routing, and on completion of a production phase, it is clear what the next phase consists of. Therefore, ERP systems *do* assume full knowledge of the remaining process phases for every process instance in the production system.

2.5.4 Routing flexibility

Apart from the fact that the routing of a process instance is not fixed because of external factors, the routing can sometimes be determined by the organization itself, to a certain degree. In paperwork organizations, one sometimes has the choice between alternative processing procedures. Also, it is sometimes possible to join activities belonging to a certain case together, or to alter the order in which they are performed. Some activities might even ask for parallel execution, hereby reducing the throughput time.

The treated concept of QoS is also an example of routing flexibility. Certain activities are either performed — hence possibly reducing the quality risks, but implying a higher throughput time and more capacity — or left undone — hence keeping the throughput time and the demand for capacity low, but ignoring the QoS. The routing that is chosen depends on the outcome of a trade-off between QoS, throughput time, and capacity.

Concerning the operational control of manufacturing processes, such trade-offs are encountered on a much less frequent basis, really only when designing the manufacturing process. During operations, the QoS level, average throughput time and capacity demands remain fixed.

A conclusion extracted from PLATIER[9] is that it is useful to start off by designing the quality control system and then to design the logistic control system, i.e. the routing system, since quality control has its influence on the routing of process instances through the workflow.

2.5.5 Absence of causal connections between capacity and value added service

As stated before, extra capacity does not necessarily imply extra QoS. Regarding manufacturing processes, it is a common feature that the decision to put more effort (so as to mean more money, more capacity, more time, higher quality base materials) into a production phase will yield a product to a higher quality standard. Anyhow, it is inconceivable that the quality would drop.

In paperwork environments, however, such implications do not always hold. Consider, for instance, the example of the criminal case discussed in the subsection on multiple routings and outcomes. Suppose there is ample evidence to prove manslaughter, but that there are clues that point in the direction of murder, carrying a heavier sentence. However, these clues first have to be investigated in more detail, hence demanding capacity. If it is eventually decided to actually indict the suspect(s) for murder instead of manslaughter, but in the process some serious errors are committed by the prosecutors — e.g. in the form of technicalities or illegal interrogation methods — then the result might very well be a verdict of not guilty. Thus, instead of the QoS amounting to the penalty for manslaughter, it amounts to zero QoS (the accused shall be released), so here an increase in capacity has led to a serious decrease in the QoS.

2.5.6 Resource flexibility

Another issue is the great flexibility concerning the allocation of capacity to workstations. In manufacturing environments, resources are usually assigned to a fixed location and ordered to execute a fixed task. In paperwork environments, however, most resources are flexible in the sense that they can be put to work at various locations, performing various tasks after one another. On-line resource management, consisting of resource scheduling and resource control, is an important issue in workflow management, whereas in industrial environments resource management does not stretch beyond the first three phases of resource management — i.e. the successive selection, layout design and assignment of resources.

A shortcoming of all well-known WfM software packages is the lack of means for advanced facilities that provide the opportunity to model processes and capacities on a state-dependent basis, such that state-dependent decisions can be made. This functionality can indeed be supported by standard software, but it does require the availability of decision support routines.

2.5.7 Absence of market forces

Many paperwork offices are public organizations which are not focused on making profits. Instead, their existence was initiated and is financed by the government or a local council. The same holds for their goals, which are mostly focused on serving the public in some particular fashion, hereby maintaining a certain QoS level and average response time towards the public, and keeping operating costs within the budget.

Clearly, in general, commercial companies have a completely different set of objectives, e.g. maximizing profit under certain limiting conditions which can possibly be relaxed to some extent. This has its consequences for the information system that manages the business process, since the objectives partly determine the operating procedure of this information system.

2.5.8 Semi-automated control

In principle, the workflow management system controls the flow of work, but it should also allow (human) resources the latitude to take their own decisions, if appropriate. An important property of workflow management is that the user is given a high degree of process control capability. As it were, his own intervention determines the control flow. As a result, the process is guided by means of semi-automated control.

One reason why fully automated control is not desirable is that all human beings have their own special ways of working. Some find it pleasant to work on a *first-in-first-out* basis, while others prefer to order their work list according to the expected duration of the activities to be carried out, finishing the short ones first. And some just start working on whatever case catches their eye first. Forcing a specific, perhaps ‘optimal’, working procedure to be followed might prove to be counterproductive.

In addition to this, there always remain erratic events the potential occurrence of which was not taken into account when implementing the so-called *workflow definition* (see section 3.3) into the workflow management system. For example, imagine a routing no one has thought of, but which seems to be the appropriate operating procedure because of some ‘strange’ action by a client involved in the case. Surely, the workflow management system cannot detect such events, but human resources can. Whenever such an exceptional event is detected, appropriate changes can be made to the workflow definition and the workflow management system can be updated accordingly so that it will be able to manage similar events in the future.

2.5.9 Conclusion

We have seen that workflow processes exhibit some special features and that these characteristics have their consequences as regards the type of internal control to be applied. We have also demonstrated that, because of this, the systems used for the control of manufacturing processes — which were the traditional field of study of business processes — cannot be used to control workflow processes as well. Advanced (complementary) decision support routines for process control remain to be developed.

2.6 Objectives of administrative organizations

In the previous section we already noted that the overall objective of an administrative organization does not concern one single goal, but that it consists of a number of mutually highly dependent objective criteria. These are, in principle, set by the management of the organization or a higher authority. Putting all objective criteria together and quantifying them, we obtain the so-called *objective function*.

There are several types of objectives and we have come across the most prominent ones throughout section 2. We distinguish between completion time, efficiency and product quality objectives and objectives that concern the realization of certain external changes.

Completion time objectives concern controlling the average time it takes to complete a case and meeting one's commitments. Efficiency objectives are aimed at an efficient and effective deployment of company personnel and other available means of production. Product quality objectives are included into the objective function to represent assurance of a certain level of quality of service.

There is a wide range of objectives regarding the realization of external changes. For example, one can think of an insurance company seeking a larger market share, the department of internal affairs trying to reduce crime rates, or the tax department trying to encourage taxpayers to voluntarily meet their tax obligations. However, it may be clear that the achievement of such external effects — e.g. 'good behaviour' — makes its demands upon actualized quality of service, throughput times and resource management. Therefore, internal and external objectives are by no means independent of each other and therefore should not be treated as such.

2.6.1 Hard and soft constraints

Objectives are subject to a number of constraints. One can distinguish between *hard* constraints and *soft* constraints. Hard constraints are constraints that cannot be relaxed. Prominent examples of such constraints are the task to keep within the assigned budget and the obligation to operate in accordance with law. Some hard constraints are self-imposed by the organization, but most are not.

Soft constraints, on the other hand, can be relaxed. They represent most of the (usually self-imposed) constraints concerning ergonomic, ecological, sociological and similar aspects of the working environment. When using quantitative — i.e. mathematically oriented — techniques to optimize existing workflow processes, soft constraints are usually omitted in the analysis. Yet, if a soft constraint is to be incorporated into the mathematical model, then it should be re-defined as a hard constraint.

To every design of the workflow process corresponds a certain degree of realization of the overall objective, represented by the *value* of the objective function. This value determines the *performance* of the organization. Seeing the desired performance as an objective and the cost issue as a constraint, the appropriate mathematical approach will be to develop methods that try to reach the desired performance or optimize the performance, *given* that the costs stay below some predefined level.

Analysing workflows using mathematical techniques, it is assumed that particular workflow designs will yield particular (definite) objective function values. This means it is assumed that the *implementation* of a workflow design only depends on the design itself, and not on factors like time and space.

Not every workflow design can be implemented easily and some are more practicable than others. A certain design — virtually a collection of operational (decision) rules — may have been nominated ‘optimal’, but if it is evident that the implementation will be far from straightforward, then the design will not be a suitable one after all. Therefore, when applying a mathematical technique to construct suitable control rules, it is recommended to aim for ones that are likely to be actually implemented and followed during operations in the same way as was planned.

Therefore, it is our belief that workflows should be designed in such a way that simple (real-life) control rules are adequate to obtain good performance. Such rules will possibly not be optimal in the mathematical sense of the word — i.e. in terms of actually optimizing the objective function — but will be practicable and, hopefully, near-optimal.

It is important to note that most workflow processes one encounters in practice are of such complexity and size that it is not even possible to construct optimal control rules within a reasonable time, so even if we wanted an optimal, but surely impracticable, solution of the optimization problem, we would not be able to find it anyway.

3 Architecture of Workflow Management Systems

The organization of work within organizations continues to become more complex. This has given rise to the development of (automated) information systems that support the control of the workflow processes and their mutual tuning. *Workflow Management Systems* (WfMSs), as these systems are termed, manage workflows and offer a new model for the division of labour between people and computers.

WfMSs are used for flexibility, system integration, process optimization and control, organizational changes, maintenance improvement, and so on. The fact that there are many functionalities causes confusion with respect to the functionality that one might expect from WfMSs. This danger has been recognized at an early stage by the *Workflow Management Coalition* (WfMC). The WfMC is an independent organization that, among other things, standardizes workflow related terminology and constructs frameworks for the establishment of workflow standards. In addition, the WfMC defines standards for the exchange of data between workflow management systems and applications.

According to the Coalition, workflow standards are relevant because organizations making an investment in workflow software want to be sure that their investment is going to be protected. With standards, users can have confidence that essential criteria will be met, hence reducing the risks involved. This clearly becomes of paramount importance when workflow systems are required to interoperate with those of other organizations when business processes are conducted across company boundaries.

The Coalition has a membership consisting of a wide range of organizations from various branches of business, including Cap Gemini, Deutsche Telekom, Fujitsu, Hewlitt Packard, IBM, Lucent Technologies, Microsoft, Novell, Oracle, Royal Bank of Canada, SAS Institute and Siemens Nixdorf, to name only a few.

One of the basic principles of the WfMC is the *Workflow Reference Model*. It is a description of the architecture of a WfMS, identifying the most important system components and interfaces. This workflow reference model is depicted in figure 1.

We will briefly discuss figure 1. For an exhaustive elaboration on the characteristics and functionality of the individual components and their interrelations, we refer to WFMC[15] and WFMC[16], or LAWRENCE[8].

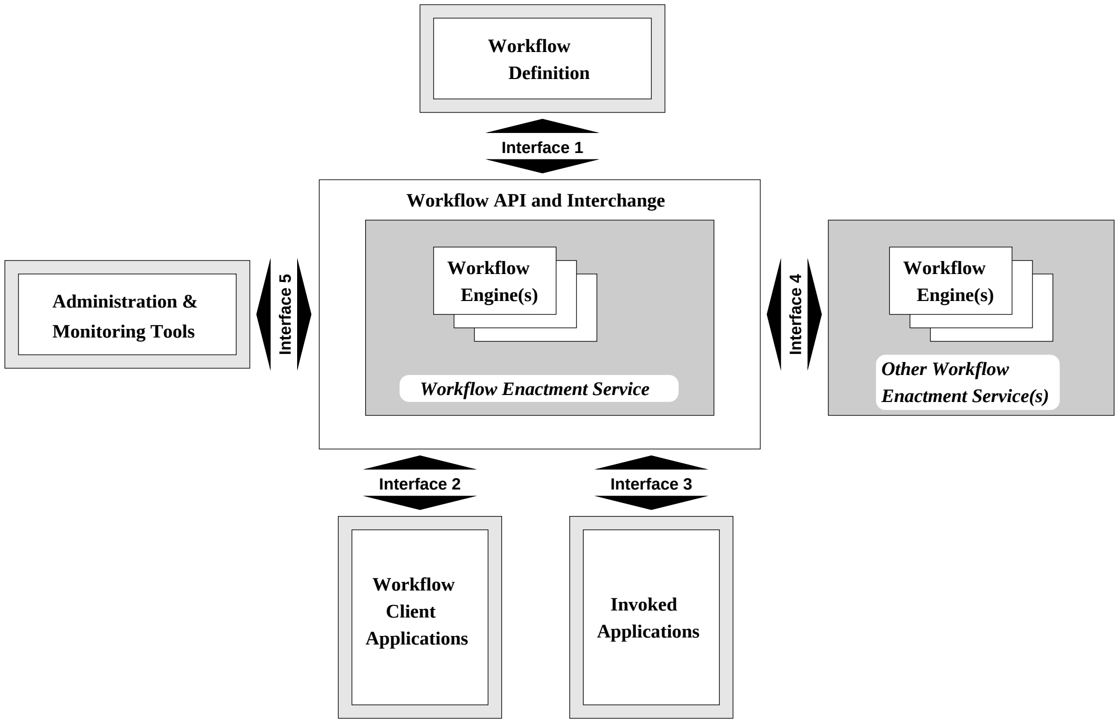

Figure 1: The Workflow Reference Model

3.1 Workflow enactment service

At the heart of any WfMS lies the *workflow enactment service*. The workflow enactment service makes sure new cases are created whenever process instances arrive, takes care that the correct activities are carried out in the correct order and by the correct — i.e. most suitable — resources, and makes sure cases are terminated when all specified tasks have been executed. If, for whatever reason, activities have not been performed as planned, e.g. by other resources than ordered, the workflow enactment service takes responsibility for ensuring that the process does not get locked, but that it is continued in a natural way.

The ability to change processes when necessary and the ability to respond to changing needs are important requisites of a workflow management system, as we have illustrated in section 2.5. This functionality is captured in the term *adaptive workflow*.

A workflow enactment service consists of one or more *workflow engines*. Enactment may either occur within a single (homogeneous) workflow domain or, using the facilities provided within the *Interoperability Interface* (Interface 4 in figure 1), across engines within several (heterogeneous) domains. Thanks to workflow interoperability, workflow engines can communicate and work together to co-ordinate work.

3.2 Workflow engines

A *workflow engine* is a software service that provides the run-time execution environment for process instances. It offers operational functions to support the execution of tasks, based on the *process definition* and using the previously discussed resource classification.

In particular, a workflow engine keeps track of the state of process instances as they progress through their defined task stages, pushing them along to subsequent tasks that need to be performed according to the *logic* that has been defined for the process. The process definition incorporates this process logic. Furthermore, a workflow engine links tasks that need to be performed with resources that can execute them and instructs these resources how — e.g. in what order and by what quality standards — the tasks they are assigned to are supposed to be executed.

3.3 Process definition

A *process definition* is a representation of a business process in a form which supports automation, such as enactment by a workflow management system. It contains all necessary information about the process to enable it to be executed by the workflow enactment service. This includes a — whether or not graphically depicted —network of process steps and their interrelations (the *process*), indicating all possible routings of process instances through the process. Moreover, criteria to indicate the start and termination of the process and of individual activities, references to applications which may be invoked during the executing of tasks, and the definition of any other workflow relevant data that may need to be referenced are included.

The process definition is usually specified in terms of organizational entities and role functions, rather than specific resources. It is the responsibility of the workflow enactment service to link these entities and role functions with specific resources within the workflow run-time environment. Here, a *role* is equivalent to a resource class. If, based on their capabilities, two resources belong to a same resource class, then they are said to be capable of fulfilling a same role.

By separating the charting of the workflow process from the charting of the resources within the organization, complexity is reduced and reuse is stimulated.

3.3.1 Workflow definition

Put together, the process definition and the resource classification constitute the definition of the workflow process: the *workflow definition*. The interface between the workflow definition on the one hand and the workflow enactment service on the other (Interface 1), is termed the *Workflow Definition Interchange Interface*.

3.3.2 Process definition tools

Both process definition and resource classification are built using so-called *process definition tools*. Apart from these tools that support the charting of the workflow process itself, there also exist tools that offer analysis functionality. This concerns both the analysis of the workflow definition and the analysis of the performance that is attained when implementing this workflow definition. Unfortunately, the analysis functionality of present-day process definition tools is not very extensive.

3.4 Workflow applications

Applications may interface to the workflow enactment service via the *Workflow Application Programming Interfaces* (WAPIs). We identify the *Workflow Client Applications Interface* (Interface 2), the *Invoked Applications Interface* (Interface 3), and the *Administration & Monitoring Tools Interface* (Interface 5).

3.4.1 Workflow client applications

Workflow client applications are applications that interact with a workflow engine, requesting facilities and services from the engine. Common functions of a client application include worklist handling, and initiation, suspension, resumption and termination of activities, as well as manipulation of process definition data.

3.4.2 Invoked applications

Invoked applications are applications that are invoked by the workflow management system to automate an activity, or to support manual execution of an activity. Application invocation may be a function of the workflow engine c.q. the *worklist handler*, a software component that manages and formulates a request to the workflow enactment service in order to obtain a list of activities to be performed, whenever such a list is needed.

3.4.3 Administration & monitoring tools

The *administration & monitoring tools* are used by the workflow administrator, who has some special privileges which allow various workflow management system functions to be performed. Administrative functions may include setup and management of system users, re-assignment of activities to resources, and processing exception conditions. Monitoring tools offer facilities to track workflow events during workflow execution — e.g. the progress of work-in-process or the progress of a single process instance.

3.5 Conclusion

We have discussed the Workflow Reference Model, a comprehensive universal framework for the conceptual structure of WfMSs. Ideally, WfMSs offer modelling, execution and tracking functionality as well as decision support and analysis functionality. Because of lack of state-of-the-art decision support routines and analysis tools, the latter type of functionality is underdeveloped in present-day WfMSs. As regards the future development of such routines, we have stated that any (mathematical) approach should at least take into account the generic form of the objective function of a paperwork organization, as well as the fact that the difference between design and implementation is an essential one.

Literature consulted

[1] AALST, W.M.P. VAN DER, *Exploring the process dimension of workflow management*, Computing Science Report 97–13, Eindhoven University of Technology, 1997.

[2] AALST, W.M.P. VAN DER and K. VAN HEE, *Workflow management: modellen, methoden en systemen*, Academic Service, Schoonhoven, The Netherlands, 1997, *in Dutch*.

[3] AALST, W.M.P. VAN DER, A.G. DE KOK and W.H.M. ZIJM, *Overview Research BETA*, BETA Research Institute, Eindhoven University of Technology, 1999.

[4] BRAND, N.A. and J.R.P. VAN DER KOLK, *Werkstroomanalyse en -ontwerp*, Kluwer Bedrijfswetenschappen, Deventer, The Netherlands, 1995, *in Dutch*.

[5] BROUNS, G.A.J.F., *BETA-cursus Information Technology for Workflow and Operational Process Modelling*, internal report, Eindhoven University of Technology, 1999, *in Dutch*.

[6] HOPP, W.J. and M.L. SPEARMAN, *Factory physics: foundations of manufacturing management*, Irwin, London, 1996.

[7] KEYS, P., *Operational Research and Systems*, Plenum Press, New York, 1991.

[8] LAWRENCE, P., editor, *Workflow Handbook 1997*, John Wiley & Sons Ltd., Chichester, 1997.

[9] PLATIER, E.A.H., *Een logistieke kijk op bedrijfsprocessen*, Ph.D. thesis, Eindhoven University of Technology, 1996, *in Dutch*.

[10] REIJERS, H.A., *Business processes*, working paper, Eindhoven University of Technology, 1999.

[11] SCHEER, A.-W., *Business process engineering: reference models for industrial enterprises*, Springer-Verlag, Berlin, 1998.

[12] VERBEEK, H.M.W., T. BASTEN and W.M.P. VAN DER AALST, *Diagnosing workflow processes using Woflan*, Computing Science Report 99–02, Eindhoven University of Technology, 1999.

[13] VERNADAT, F., *Enterprise modeling and integration: principles and applications*, Chapman & Hall, London, 1996.

[14] WACKER, J.G., *A definition of theory: research guidelines for different theory-building research methods in operations management*, *Journal of Operations Management* **16**, 1998, pp. 361–385.

[15] WFMC, authored by D. HOLLINGSWORTH, *The Workflow Reference Model*, Document Number TC00-1003, Winchester, 1995.

[16] WFMC, *Terminology & Glossary*, Document Number TC00-1011, Winchester, 1999.

Acknowledgements

The author wishes to thank Prof Jan van der Wal and Prof Onno Boxma for their valuable comments on previous versions of this paper.