Chapter 1

Introduction

In this chapter, we discuss the basics of workflow technology and show the applicability of this technology. It provides a basic understanding of the various aspects of workflow and provides the basis we will build on in the rest of the book.

We have divided the chapter into four parts: in the first part, we define what a business process is, its relationship to workflow, and the different aspects and types of workflows; in the second part, we show the characteristics of applications built on workflows; in the third part, we show some more advanced applications of workflow; and in the fourth part, we state the requirements for a production workflow management system. We finish off by sketching the relationships between workflow and other information processing technologies.

1.1 Business Processes

Most people associate business processes with the typical operations in a bank or an insurance company: the loan process in a bank or claims processing in an insurance company. Such operations have led us to understand business processes as a sequence of activities performed by various persons, the visible result being various pieces of paper.

These activities are typically repeated over and over, following the same well defined pattern, the *process model*. This is the type of processing well established in manufacturing: assembly-line production of cars in the automobile industry, or customer-specific assembly of a personal computer in the factories of PC makers are well known examples. This similarity in processes has led to the metaphor of insurance companies and banks being nothing but paper factories, with paper as the final product.

What about the area of software development? It also is guided by a set of rules that determine the sequence of steps to be followed. In the waterfall model, the steps are Architecture, High-Level Design, Low-Level Design, Coding, Unit

Test, and System Test.

Or, let's look a the work of a database administrator. A database administrator performs a set of precisely prescribed processes consisting of a number of steps that may be executed sequentially or in parallel. User interaction in this type of process is minimal, and this user interaction is limited to one person, the database administrator.

This type of a business process allows us to conclude that all batch jobs being run can also be considered some kind of business process. The only difference from the database administrator example is that there is no user interaction. The steps in these batch jobs are described by some kind of scripting language that provides the flexibility to take appropriate actions if one of the steps in the batch job fails to complete correctly. The job control language (JCL) of the OS/390 operating system, one of the most successful operating systems for mainframes, supports such batch jobs, which in turn implement a business process. Code Example 1.1 shows a sample specification of a business process, using OS/390 JCL.

Code Example 1.1 Using the Operating System Job Control Language

```
//OVERDUE JOB
//*-------------------------------------------------------
//* STEP 1 : DETERMINE CUSTOMERS WITH OPEN BILL
//*-------------------------------------------------------
//EXTRACT  EXEC PGM=OVERDUE
//IUSERS   DD   DSN=CUSTOMER.DATA,DISP=OLD
//OUSERS   DD   DSN=OVERDUE.USERS,DISP=(NEW,PASS)
//*-------------------------------------------------------
//* STEP 2 : WRITE LETTER TO CUSTOMER
//*-------------------------------------------------------
//WRITE    EXEC PGM=PRINT,COND=(4,LT,EXTRACT)
//OUSERS   DD   DSN=OVERDUE.USERS,DISP=OLD
//*-------------------------------------------------------
//* STEP 3 : DELETE FILE
//*-------------------------------------------------------
//DELETE   EXEC PGM=DELETE,COND=EVEN
//OUSERS   DD   DSN=(OLD,DELETE)
//
```

The process is used for a credit card company to write a letter to all of the customers that have not paid their open bill. In the first step, all customers that have not paid their bill are selected. This information is saved in a file, identified by OUSER. The second step uses the data obtained in the first step and writes a letter to each customer. If all customers have paid their bill on time, the second step is skipped (COND=(4,LT,EXTRACT)). In the third step, the file is deleted, regardless of whether the second step has been carried out (COND=EVEN). This process shows all of the characteristics of a process. It consists of a set of activities

(we called them steps) that are carried out in a particular sequence. Some of the activities may be skipped as the result of a previous activity indicating so. Data is passed from activity to activity. We later call these characteristics control flow and data flow.

The examples above illustrate that the notion of business processes covers a wide spectrum; there is no such thing as a typical business process. The business of a user determines what a business process is: for the database administrator, it is the reorganization of a database; for the controller, it is the creation of the monthly balance sheet; for the CEO of a bank, it is the granting of a loan.

It is this variety of business processes that needs to be addressed by a production workflow management system and that we cover in the book.

1.2 Business Processes as Enterprise Resource

Enterprises are considering information more and more as an important resource [LA94]. They consider it now to be a major asset that is as important as the more traditional ones such as land, labor, and capital. The information includes all data about all resources that are needed to achieve the goals of the enterprise. It is generally agreed that this information should be structured and presented as formally as possible. The set of actions needed to come up with this formal specification is called "enterprise modeling."

The conceptual base for enterprise modeling is sometimes called a *hyper-semantic data model* [PTE89]. The result of modeling a particular enterprise is an *enterprise model*. An enterprise model consists of two parts: the *data model* and the *knowledge model*.

The data model describes the structure of all resources of the enterprise. It is thus somewhat like the syntactical component of the enterprise model, that is, it specifies the static aspects of an enterprise. In this sense, the data model describes *what* is available in the enterprise to reach the specified goals. Most enterprises build their data model via *semantic data models*. Some of the more prominent semantic data models are the entity/relationship model [Che76], EXPRESS [ISO], and the unified modeling language (UML) [Rat97].

The knowledge model describes the use of the resources and their connections. It is thus the semantic component of the enterprise model; it specifies the dynamic aspects of an enterprise. In this sense, the knowledge model describes *how* the enterprise uses its resource to achieve its goals. The knowledge model contains constraints, heuristics, and procedures. Constraints define the local and global consistency of resources; for data resources stored in databases, constraints define the valid states in the database. Heuristics describe how to derive data. Procedures define events and correlated actions, set sequences of actions, and describe business processes.

1.3 Virtual Enterprises

Traditionally, business processes are carried out completely within the boundaries of a company. The processes start within the company and finish within the company. External requests come in as phone calls, faxes, postal mail, or, to a limited extend as electronic messages, for example, in electronic data interchange (EDI) or extended markup language (XML) format. These requests are handled by appropriate personnel who start the appropriate business process for each request. As these business processes move to completion, actions, such as shipping goods to the originator, take place. Once the process is complete, the appropriate results are returned to the originator again in the form of phone calls, faxes, postal mail, or electronic messages.

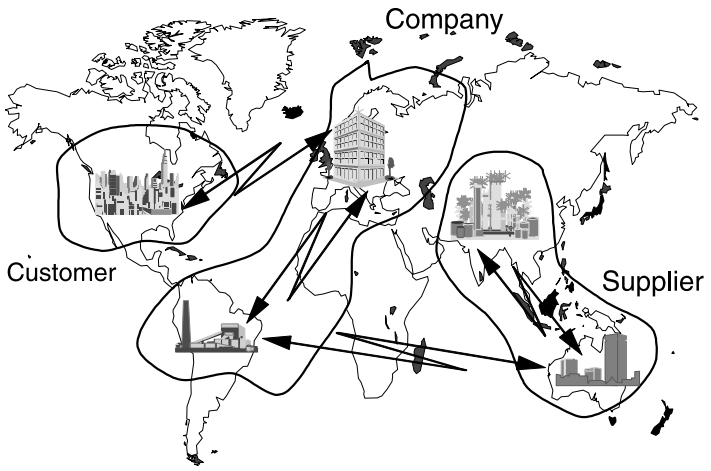
Inquiries about the status of a business process were usually received as phone calls, with the customer waiting for the result. The call center operator would immediately perform the appropriate query against the proper application system to determine the status and give the corresponding answers to the customer. This dialog would go on for some time until the customer was happy with the answer. As a result of all the involved manual interactions, the service to customers and suppliers was limited to certain operation hours.

In a fast-paced global market, companies can no longer operate in this traditional way of doing business. Consumers select those companies where they can do business any time they want. To support the cycle time needed by customers and suppliers, the business processes of customers and suppliers must interact with a company's business processes without any human intervention. Figure 1.1 depicts the interactions between a company and its customers and suppliers, as well as the interactions between the different parts of a company, such as the sales head office and the manufacturing plant. These interactions range from simple inquiries to complex interactions between companies to the situation where organizational units of different companies are tied together by sharing data and business processes and operate like an independent enterprise (*virtual enterprise*).

Customers no longer make phone calls or mail order forms when they want to do business with a particular company. They no longer have to deal with an operator at a call center when they want to know the status of the order they had placed a few days ago. They enter the company's URL on their Web browser and are then connected to the company's home page, from which they can select the appropriate application system. The Web browser runs on such diverse devices as a personal computer, a mobile phone with a built-in monitor, or even a microwave oven.

Business processes of different companies honor each other's requests without any manual intervention. No interaction is required to have a company's business processes talk to the appropriate business processes of customers, suppliers, distributors, or even government agencies, such as the tax department.

Figure 1.1 Connected Business Processes



The levels of interaction between the different business processes covers a wide spectrum:

- An activity in a business process invokes another business process. That new business process executes totally independently of the original business process.
- An activity in a business process invokes another business process and waits until that new business process has completed.
- An activity in a business process invokes another business process, and an activity later in the business process waits until that new business process has completed.

Very complex scenarios can be built with these basic structures. We discuss these structures in detail in section 3.4.3 on page 87. Typical business processes that are made up of these basic structures are the business processes that define the interactions between the manufacturing company and its suppliers. A manufacturer of personal computers may have a business process for the assembly of the personal computers. One of the activities is the ordering of motherboards to be delivered by a supplier when the amount of available motherboards falls below a certain level. In this case, a request is sent to the supplier; the request starts the appropriate order entry process. When production of the motherboards has completed and they have been shipped, the final activity in the order entry process sends the appropriate billing request back to the manufacturer. The manufacturer receives the billing request and starts the appropriate accounts payable process.

To make sure that business processes are carried out correctly, each business process has a process administrator assigned to it. If an out-of-line situation occurs, the process administrator is notified so that corrective actions can be taken. A typical out-of-line situation occurs when a time limit is associated with a process and the process takes longer than the time specified. Tools are available to the process administrator to query the state of the business process, assign a higher priority to the business, or shift work from one person to another.

In the case of inter-enterprise business processes, it is also possible that the problem is associated with a process that is running at another company. In this case, it is desirable that a process administrator can also query the state of that other process so that the process administrators of both companies can jointly work on the appropriate corrective actions.

The level of interaction we have discussed so far reflects interaction between independent companies. Each company runs its business in its own way with its own distinct business processes. Each business process is associated with data that is part of the enterprise data. When data is needed by processes of other companies, the data is shipped together with the invocation of the other process.

Within an enterprise, the difference between processes and subprocesses is not visible at all. All programs associated with the different business processes access, in general, the same set of enterprise data. Communication between the applications that make up the business process is not visible to the outside.

Virtual enterprises are enterprises in the sense that they operate as any other enterprise. However, they are not organizationally the same as a typical enterprise. They are made up of parts of different companies, where each company contributes parts to the overall enterprise. In the extreme case, one company contributes by manufacturing the product; another company is responsible for distribution; a third one, for marketing; and a fourth one, for billing.

In this case, the different pieces of the virtual enterprise operate as one enterprise. This collaboration is made possible by the use of common business processes and common data. Each activity within a business process can then be performed by one of the participating companies. The data is accessible to all of the companies. The Internet provides the communication backbone for running the processes and accessing the data.

It is interesting to observe that the same methods for carrying out business processes can be observed if parts of a company are outsourced. For example, a new, independent company is formed to perform a certain task. In this case, existing business processes need to continue to be performed even though there are two independent companies.

Each level of interaction we have discussed requires that the different systems work together without any friction independently of the type of implementation that has been chosen by the involved parties. This requirement mandates that the protocols used to communicate are standardized. Without standardization, bilateral

agreements would have to be negotiated and then implemented on both ends for every new set of interactions.

1.4 Processes and Workflows

The *process model* describes the structure of a business process in the real world. It defines all possible paths through the business process, including the rules that define which paths should be taken and all actions that need to be performed. This model is a template from which each *process* is instantiated; that is an instance of the process model is created. Some people use the term *process instance* for the instances that are created from process models. An individual process (process instance) is carried out according to a set of values that determines the actual path through the process. In an insurance company the process model might be called ClaimsProcessing, and from this model a large number of processes are created, one for Joe Smith, one for Laura Miller, and so on.

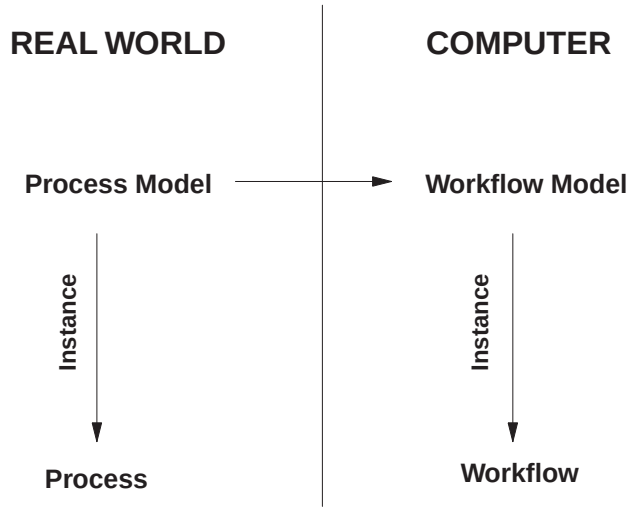
Processes need not necessarily be run on a computer. A large number of business processes are performed without a single step being performed by a computer. A typical example is that of a manager circulating a document for each employee to read. Whenever an employee receives the document in her mail basket, she reads it, signs it, determines who has not seen it, and then puts it into the mail basket of somebody who has not yet signed it. The employee who signs the document last puts it back into the mail basket of the manager.

Business processes thus may consist of parts that are carried out by a computer and parts that are not supported through computers. As shown in Figure 1.2, the parts that are run on a computer are called a *workflow model*.

A workflow model may be just a small part of a larger process model, or it may encompass the whole process model. The computer programs that carry out the workflow can be a general-purpose workflow management system or specialized applications that implement the process model. The workflow model is a template for creating workflows in the same sense that the process model is a template for creating processes.

In this book, we follow the convention of using the term *process* when talking about processes and process models, and *workflow* when talking about workflows and workflow models. In fact, we even follow the conventions of even using the terms process and workflow interchangeably. The context will make it obvious which one we are talking about. In cases, where the meaning could be ambiguous, we use the full term.

Figure 1.2 Processes and Workflows



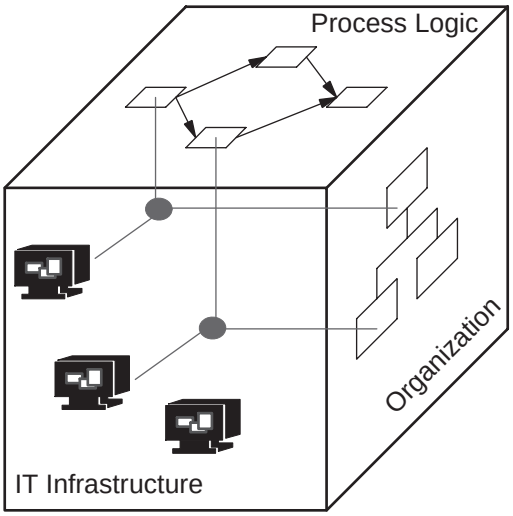
1.5 Dimensions of Workflow

Business processes, and therefore workflows, have three independent (orthogonal) dimensions. These three dimensions can be graphically depicted as a cube, as shown in Figure 1.3.

The cube's first dimension represents the *process logic*. It describes *what* in terms of which activities are to be performed and in which sequence they need to be performed. Each box represents either a program that is invoked or another process that is executed either as a local subprocess or as a remote subprocess within the same company or in another company. The arrows show the flow of control from one activity to the next. The flow can be sequential or parallel. If it is sequential, one activity is performed after the other; if it is parallel, two or more activities are executed at the same time, usually reducing the time required to perform a business process.

The cube's second dimension, orthogonal to the process logic, is the *organization* dimension. It describes the organizational structure of the company in terms of departments, roles, and people. This information is used to define *who* should perform each activity. A query can be specified for each activity to identify the set of people in the organization to whom the activity can be assigned. For example, if it has been defined, that an activity can only be performed by someone meeting a particular job function (role), then all persons assigned to the job function are selected and are assigned to perform the activity. If the activity does not interact with a user, the workflow management system carries out the activity on behalf of the selected user.

Figure 1.3 The Three Dimensions of Workflow



The cube's third dimension, orthogonal to the process logic and the organization, constitutes the dimension of information technology (IT) infrastructure. It describes *which* IT resources, such as programs that perform a particular activity, are required.

The execution history of a process is thus a sequence of triplets (activity, user, IT resource), or in other words, a series of points (trajectory) in the three-dimensional workflow space W^3 (*w*hat, *w*ho, *w*hich).

1.6 User Support

Several users are involved in running workflows: end users that perform the individual tasks, system administrators that have overall responsibility for the workflow management system and the actions it performs, operation administrators that are responsible that the workflow management system is up and running, and process administrators that monitor the expected processing of the individual business processes.

The individual tasks are typically performed though a graphical user interface (GUI). This user interface could be anything from a Web browser running on a network station, a window on a regular PC, or the tiny monitor of a hand-held device.

End users receive the requests for processing individual activities as *workitems*. These workitems are typically represented as icons. The representation of the icon shows, at least to a certain extent, the type of work that must be performed. These

workitems are made available to the user either upon explicit request or upon being pushed by the workflow management system onto the user's desktop. When a user wants to work on a particular activity, she clicks on the appropriate icon, which then starts the associated program. When the work associated with the activity has been completed, the workitem is automatically removed. The user can organize the individual workitems according to specified criteria into *worklists*. Criteria can be priority, age, or type of activity, such as confidential or public.

Operation administrators are provided with a graphical view of the status of each of the different workflow management system components. This interface can be provided either by the workflow management system or by a systems management system. The interface typically supports alerts for out-of-line situations as well as query and change capabilities for all or individual components.

The user interface for process administrators is geared toward managing individual process instances and acting as a focal point for all process-related events. The events are represented as workitems that can be started to perform appropriate process repair actions. The starting point is a graphical representation of the business process. By pointing to individual activities, the process administrator can dig deeper into the structure of the workflow. Corrective actions that could be performed are restarting the process, transferring workitems from one user to another, or repairing a broken process.

1.7 Categories of Workflows

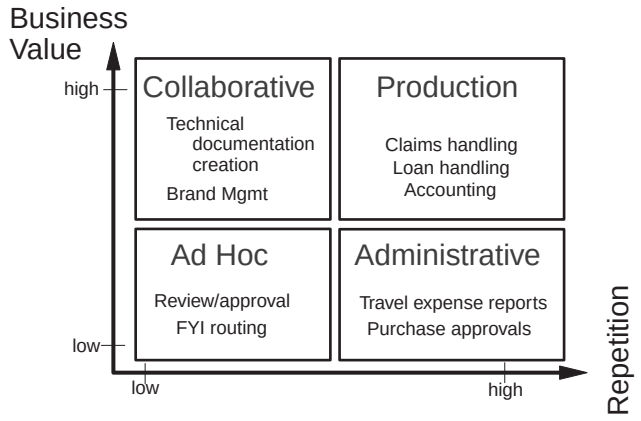
Many schemes have been developed to classify workflows. Figure 1.4 shows one of the more prominent ones. In this scheme, used by GIGA Information Group[Gro], workflows are categorized according to their value to the business and their repetition.

The *business value* defines the importance of a workflow to the company's business. A process of high business value is at the core of a company; it's a core competency of the company. The company has been founded to perform those processes; this means the company is defined as running those business processes. Typical examples of high-business-value processes are the granting of a loan for a bank or the manufacturing of a car by a car manufacturer.

The *repetition* measures how often a particular process is performed in the same manner. It is an indicator of whether the process is worth being modeled. This indicator is important since modeling of a process for the first time is nontrivial and is time consuming, which makes it typically an expensive task. The "for your information routing" business process, for example, is typically performed in large volumes; however the underlying process models are always different.

By using these two characteristics, we can distinguish the four different types of workflows as shown in Figure 1.4.

Figure 1.4 Classifying Workflow According to Business Value and Repetition



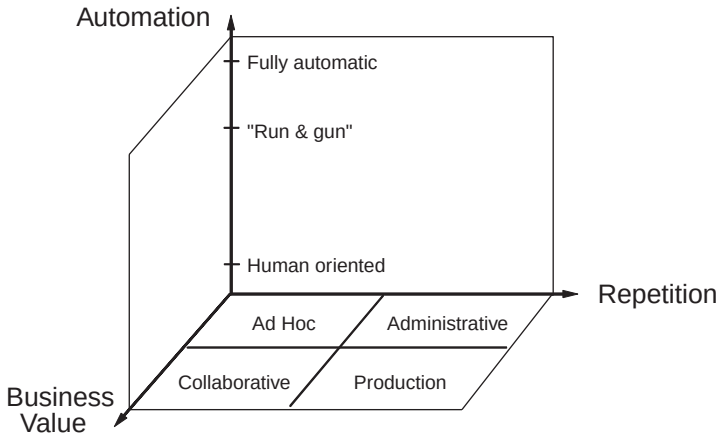
Collaborative workflows are characterized by a high business value but are executed only a few times. The processes associated with building a ship, creating technical documentation for a software product, or performing brand management for a consumer product fall into this category; all of them are extremely important for the success of the corresponding company. Their underlying process is generally rather complex and is created specifically for the particular task, often by customizing a given project plan. Changes to the underlying plan are fairly common.

Ad hoc workflows show a low business value and low repetition rate. Generally, these workflows either have no predefined structure and the next step in the process is determined by each user involved in the process, or each business process is constructed individually whenever there is a need to perform a series of actions. For-your-information routing is a typical workflow without a defined structure. A manager, for example, sends a note to all his department members with information deemed to be important for them. The receivers of the note can do whatever they want with it, including routing the note to other people who should know about this particular information. Such a process terminates if nobody routes this note to anyone else.

Administrative workflows also show a low business value; however the repetition factor is high. These workflows are the typically administrative processes such as expense account or purchase order processing. In the case of expense account processing, the employee fills out a form, the employee's manager approves it, and the expense account department verifies the correctness and issues the appropriate bank order.

Production workflows are characterized by having a high business value and a high repetition factor. These are the workflows that implement the core business of the company, such as the loan process in a bank or the claims management process

Figure 1.5 Classifying Workflow According to Business Value, Repetition, and Automation



in an insurance company. It is their efficient execution that provides a company with a competitive advantage. Often, production workflows are still implemented without a workflow management system because the corresponding processes have been in place for decades. And because of their importance, the corresponding applications were typically implemented as transaction-processing monitor programs. The scalability, availability, and robustness of these environments are consequently a prerequisite for workflow engines that run production workflows; such a workflow management system is called a *production wfms*.

An additional distinctive dimension to the business value and the repetition factor is the degree of automation of a workflow as shown in Figure 1.5. This measurement shows the independence of a workflow from human intervention, that is, whether the activities of a workflow are mainly performed by humans or by the system. A highly automated process is computation intensive and typically integrates heterogeneous and autonomous application systems.

Fully automatic workflows have many similarities with batch procedures known for decades. They can be found in database administration, such as database reorganization, systems management (e.g. reacting to system events), and data warehousing (e.g. merging and cleansing of different data sources).

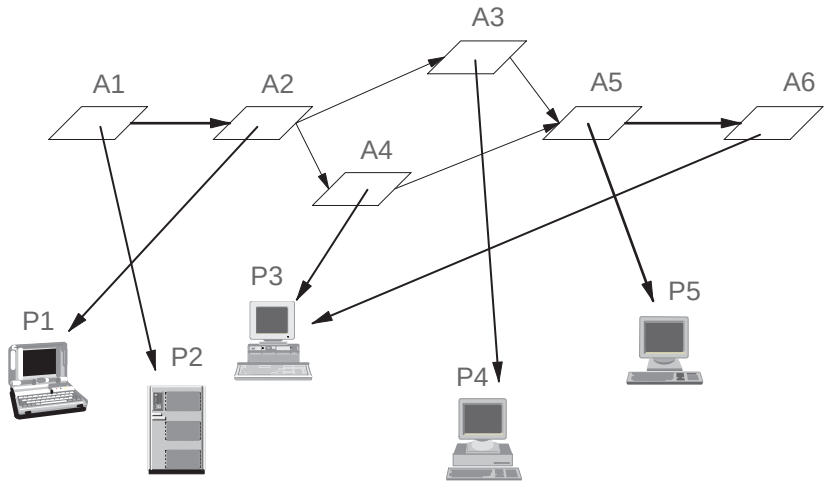
A process with a high degree of automation does not necessarily mean the process has no interactions with humans. But typically, a highly automated workflow is fragmented into sequences of activities, each of which is performed by a single user, intermixed with automatically performed sequences. The sequences of activities assigned to a particular user are often worked on in a *stream* environment, meaning the sequence is performed consecutively without any interrupt.

1.8 Application Structure

Applications that are built and executed according to the workflow paradigm are called *workflow-based applications*. They are no longer monolithic or executed on large central computers supporting dumb terminals, nor are they client/server applications where the still-monolithic application has been split into two pieces that are deployed to a large central computer and smaller intelligent workstations or net stations for better exploitation of resources.

These new applications execute on the network where the different pieces of the application are executed on different computers, each of them possibly running a different operating system, database system, or network protocol. Figure 1.6 illustrates this approach schematically.

Figure 1.6 Schema of Workflow-based Application



The first activity A1 is processed without any user interaction on workstation P2 running AIX; the second activity A2 is carried out on laptop P1 running Windows 98. Since these applications are not just carried out on a single processor but on different processors connected via a network, they are sometimes also called *networked applications*.

The heart of these applications is the business process that is executed by the workflow management system. Based on the context in which the business process executes, different paths are taken through the process model. Thus, different processes can take completely different routes through the network, are being worked on by different people, and are processed by different programs.

The business process can even be executed by different workflow management systems if a business process invokes a business process outside the domain of the

current workflow management system, for example, by sending a request via the Internet to some other company. Agreements between the different vendors and standards established by standards bodies are making this kind of execution happen.

1.9 Workflow and Objects

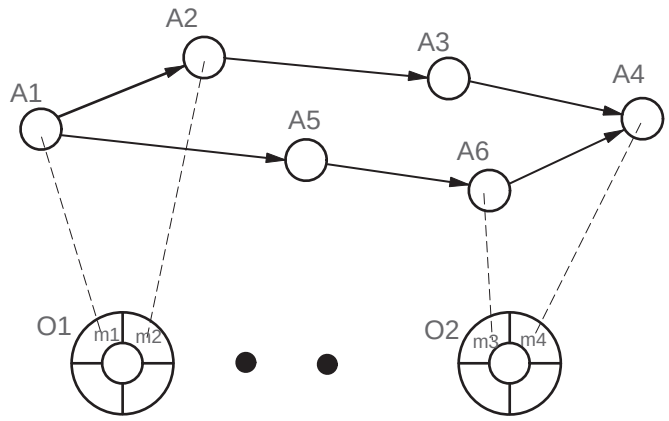
Enterprises are investing in object technology today to improve the productivity of their programmers and to enable even non-data-processing professionals to build applications by using visual builders. Vendors of standard software split their applications into business objects, allowing programmers to reuse the objects in a different context.

A cornerstone of object technology is the insight that the robustness of systems is normally achieved by encapsulating things that might become subject to change. If, for example, the order in which operations are to be performed can change or if operations can be added or removed, the guidelines of object technology consequently recommend the introduction of a dedicated control object. This control object encapsulates the scheduling of the various operations.

When the functions of the control object are examined, it is obvious that they are just a subset of the functions offered by workflow technology. The control object is represented by a process defined to the workflow management system. This representation enables vendors of standard software to glue together their business objects to form the original, monolithic business application.

Each activity of the process is implemented as a business object that is invoked by the workflow management system. As shown in Figure 1.7 an application thus consists of a business process and a set of business objects, where each activity is

Figure 1.7 Workflow and Objects



implemented as the invocation of a method of a business object. As illustrated, activity A1 invokes method m1 of business object O1, and activity A2 invokes method m2 on the same business object. This scheme should be compared to the more traditional approach in which a workflow-based application consists of a business process and a set of programs.

Chapter 6 discusses the relationship between workflow technology and object technology in depth.

1.10 Application Operating System

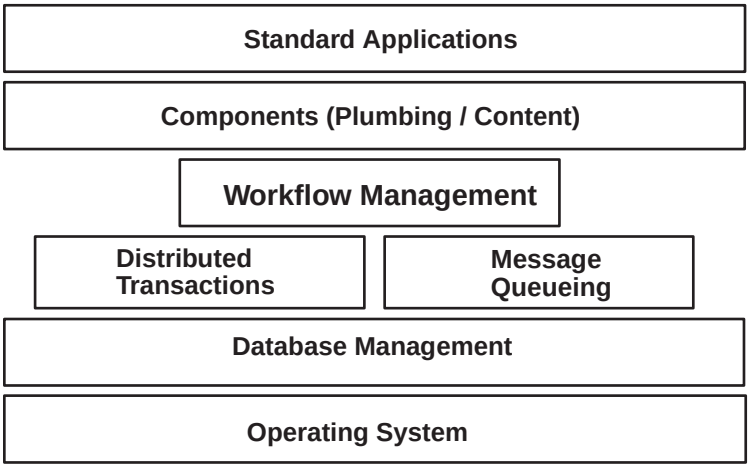
The purpose of an operating system is to provide an environment for the execution of programs. This environment includes the management of resources such as processors, memory, disk storage, and communication lines. The operating system effectively hides these resources by providing appropriate application programming interfaces (APIs) for the programs. A scripting language allows the user of such a system to specify the sequence in which the various executables are to be processed. The operating system will then make sure that the right resources are available to the program at the right time and at the right place.

When an application is made up of different components, the individual components must be called in the correct sequence (flow of control) and data must be passed correctly from one component to other components (flow of data). This knowledge is either built into each of the components or localized in some special components that manage for example the flow of control. If the application is distributed, the various components must exploit the underlying infrastructure to manage this flow of control and data. In addition, they must implement the various recovery mechanisms. The components must therefore not only be built to deal with the local operating system but must also consider the characteristics of the operating systems of other components they are communicating with.

With workflow, the component must only deal with the local operating system characteristics. The workflow management system takes care of the differences in operating systems, network protocols, and communication mechanisms. It just calls the component by using the invocation mechanism of the local operating system. If an operating-system-independent invocation mechanism, such as message queueing, is used, even this dependency can be removed. If the component is written in a portable language and does not perform operating-system-specific input/output operations, the individual component is portable.

Workflow management systems thus provide to applications the same functions that operating systems provide to the individual programs. It relieves the applications from dealing directly with operating systems and communication mechanisms by providing appropriate APIs. Workflow management systems are therefore sometimes referred to as *application operation systems*.

Figure 1.8 Software Stack



1.11 Software Stack

Figure 1.8 shows the position of workflow technology within a simplified software stack. This software stack shows only those components that are important for the points we want to make.

The base is the operating system, which provides an environment for the execution of programs.

On top of the operating system is a set of programs that are collectively called middleware. They provide the infrastructure necessary for application programs to manage data (database management system), communicate with each other via messages (message queuing system), and implement distributed transactions (transaction managers and TP monitors).

These are the components used by the workflow management system to provide the functions needed for an application operating system. The database management system is used to store the persistent information such as processes or organizational information. The message queuing system provides the infrastructure that is needed for communication between the different components of the workflow management system, either on the same computer or on different ones. The transaction manager helps the workflow management system process messages and make changes to the database under transaction control, providing the robustness required for an application operating system.

The workflow management system is then used by the next layer, which provides the object management infrastructure. This object support, together with workflow support, is then used by standard applications, such as payroll, general ledger, or groupware systems.

1.12 Document/Image Processing

Some business processes are centered around sets of documents that are routed from one person to the next. Document management systems have been developed to support this type of work. Because documents are quite often in the form of scanned images, these systems are also known as image-processing systems.

These systems focus on the management of documents/images, and offer a wealth of functions to manage and manipulate documents/images:

- Functions to scan images and annotate them with text that makes it possible to efficiently locate the images later
- A document store to support not only the simple storage of documents but also to provide functions to group related documents into folders and to search for documents using sophisticated searches
- Archiving functions to remove documents and images from the document library and put them into some slower, but cheaper, storage
- A document editor to create, modify, and annotate documents
- A forms editor to define forms and the rules that should apply to the data entered into the forms by appropriate users
- The capability to define how documents, forms, and folders are routed from person to person within the organization, and the type of action to be performed by each individual

In document/imaging systems, the workflow functions are in general an integral part of the system. A business process is typically represented as a folder. This folder is the entry into all activities associated with the process; it allows the user to perform queries into the document store and to perform actions such as adding a document, filling out a form, or determining the next user to work on the folder. These systems are sometimes also called *folder-management systems*, since the folder is the focal point. Since a folder typically represents a case, such as the loan in a bank or an insurance claim, the term *case processing* is used sometimes for this type of processing.

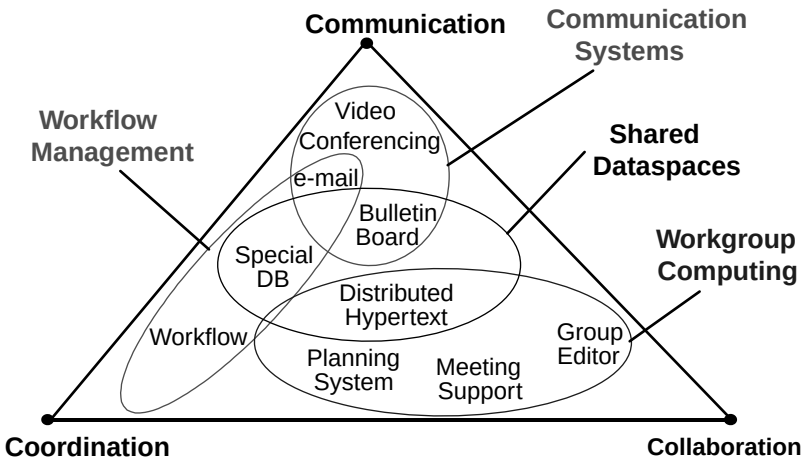
In contrast to these systems, workflow management systems focus on the underlying business process and the execution of the business as a set of activities. These activities can be anything from mailing a letter via e-mail or making a telephone call to filling out a form, for example, that automatically calculates taxes, or updates a personnel record in a corporate database.

It should be noted that document management systems have their own internal workflow management system. In fact, it is much better to implement them on top of a stand-alone workflow management system since that implementation integrates better with other systems and programs in the enterprise.

1.13 Groupware and Workflow

Workflow addresses only one aspect of groupware, systems that help people to work together on the same piece of work. These systems provide the implementations for computer-supported cooperative work (CSCW). Figure1.9 [TSMB95] classifies groupware according to the three aspects of CSCW: collaboration, communication, and coordination. Because each of the listed implementations reflects the three aspects more or less strongly, a charge diagram illustrates the strength of each aspect in a particular implementation. The closer an implementation is to an aspect, the more prominent the aspect is in the implementation. Video conferencing, for example, only has the aspect of communication; it does not have any coordination or collaboration aspects.

Figure 1.9 Classifying Groupware According to Communication, Coordination, and Collaboration



This classification schema suggests four major groupings:

- Communication systems help people exchange information, quite often as a one-way communication. This exchange includes video conferencing, e-mail, and bulletin boards on which people post their opinions.

- Shared data spaces provide people with the facilities to share data. Bulletin boards are used as a means of sharing data as in the case of Question & Answer forums on the Internet. Distributed hypertext links documents, making it possible for a group to connect their individual data to a larger web of information. Special databases are typically used to allow people work together by sharing information.

- Workgroup computing helps small projects work together by focusing on the collaboration aspects. The most prominent example is a group editor that helps people jointly write one document, for example, the specification of a computer system. Meeting support helps people organize their collaboration; planning systems to coordinate this working together.

- Workflow management systems focus on the coordination of people performing different tasks to create the final work product. Special databases are typically used as part of the workflow-based application; this means the activity implementations work with these databases. A lightweight version of a workflow management system can also be built on top of an e-mail system by using scripting capabilities within the e-mail system.

1.14 Different Views of Applications

Different people within an organization view the company's applications differently. Figure 1.10 illustrates two major views.

Business administration and management typically view applications as some kind of business process that is performed by different parts of the organization and that results in the delivery of the required item. From their perspective the mechanics of achieving this result is immaterial; their major concern is the efficiency of the business process to achieve the goals specified in the business strategy. This concern causes them to mainly focus on business goals, such as how long it takes to process a customer order, how satisfied the customers are with the company's service, or how quickly a customer complaint can be processed. A second set of business goals deals with the efficiency of business processes, such as how efficient the business process is in terms of costs and people.

Figure 1.10 How Different People View Applications

In contrast to business administration and management, information technology (IT) personnel view applications as a combination of a vast amount of data stored in files and databases that are processed by a large amount of programs operating on the data and interacting with the users. These programs are so important for the enterprise that they are implemented as transactions. The IT people are mainly concerned with the availability and efficiency of the programs and the data those programs access. Consequently, they are concerned with the availability and efficiency of the underlying systems, such as database management systems, operating systems, transaction managers, or communication systems, on which those programs depend.

Workflow helps to bridge this gap because it connects these views into a set of workflow-based applications. These workflow-based applications consist of the process models that the business administration looks at and the activity implementations that are the transactions the information processing people see. Workflow thus helps to achieve a consistent and coherent view of applications.

1.15 Transactional Workflow

Applications that implement the core functions of the company's business are implemented as transactions. The ACID paradigm, the fundamental concept of transactions, when implemented makes these applications robust and safe. The ACID paradigm defines that a collection of operations within a transaction has the following properties:

- Atomicity—Either all of them are applied to the system or none of them are,
- Consistency—They lead to a new valid state of the system,
- Isolation—They do not affect (until explicitly made visible) operations outside the collection,
- Durability—They are not undone because of any later system failure.

The transfer of money from one bank account to another bank account is the prototype of a transaction. The transaction consists of two operations; the withdrawal of money from one account and the deposit of money into the other account. In this case, both operations must succeed or none of them may succeed (atomicity). As the amount of money remains the same, the new state is a valid state of the system (consistency). The new state of the two accounts is not available to any other operation, until the transaction has completed (isolation). The changes must not be lost under any circumstances (durability).

Workflows or parts thereof must also have transactional properties. Depending on the time required to execute this transaction, different techniques need to be applied.

In highly automated workflows, that is workflows with little or no user interaction, the individual activities are often carried out with programs that have transactional behavior. Because of the streamlike behavior (short-lived, no user interaction) of these workflows, collections of such activities can be grouped together to build a new transaction. Such a transaction is called a *global transaction*. Global transactions allow automated recovery in such workflows; whenever an error occurs, all effects of the different activity implementations are automatically undone. The definition of the transaction boundaries is part of modeling the workflow. We call such a set of activities that are part of the transaction an *atomic sphere*. A typical example of an atomic sphere is the funds transfer we discussed earlier. Enforcement of the global transaction, that is atomic sphere, can be achieved by the workflow management system using traditional techniques.

The typical execution time for an atomic sphere is in the range of tenths of seconds to seconds at the most. If the execution of a transaction takes longer, the standard techniques that are applied in atomic spheres can no longer be used efficiently. A number of advanced transaction models have been developed to cope with long execution times by weakening some of the ACID properties. Typically, those new transaction models require special programming by the application developer. This type of long-running transaction is called a *business transaction*.

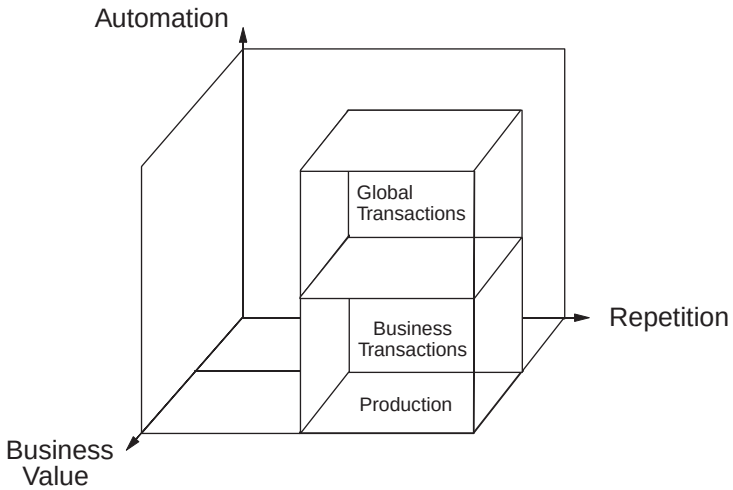
A typical example of a business transaction is our trip reservation example shown in Appendix A. After the client's itinerary is accepted, corresponding reservations for flights, hotel rooms, and rentals cars need to be made. To shorten the processing time, the three reservations are made in parallel. These three operations must be performed as a transaction. Either all reservations are successful or none of them should be successful.

The processing time of the transaction is measured in minutes or hours; however, none of the advanced transaction models can be directly applied here. The implementers of the different reservation systems had no knowledge of each other at the time they built the system, nor did they want to take this fact into consideration. Each of the systems performs its own transactions and exposes its results independently of each other. This behavior requires that the individual transactions need to be undone by compensation transactions that compensate the effects of the original transactions. The compensation of a flight reservation is the cancellation of the reservation. The definition of the business transaction boundaries and the compensation activities is, as for global transactions, part of modeling the workflow. We call the set of activities that are part of the transaction a *compensation sphere*.

Figure 1.11 shows the space that global transactions and business transactions occupy in the classification scheme shown in Figure 1.5 on page 12.

We extensively discuss the usage and properties of atomic and compensation spheres in Chapter 7.

Figure 1.11 Classifying Business Transactions and Global Transactions



1.16 Advanced Usage

Workflow technology not only applies to the traditional applications we have discussed so far, such as loan processing, it can also be used beneficially for other applications, such as for project management.

A project management system helps managers and team leaders plan and control the execution of projects. To support these tasks, it keeps a database in which all project-related data is stored.

At the beginning of the project, the user specifies all project-related information, such as the individual tasks, the sequence in which the tasks need to be performed, how much of the tasks can be performed in parallel, and the resources, such as people, to perform these tasks. With this information, the project management system generates the appropriate charts, in particular, the GANTT charts (project plan), that show which resource must be available at which time. In addition, the user can specify which resources are actually available at which time. This specification enables the project management system to generate an adjusted GANTT chart that reflects the actual allocation of resources. This chart is refined, by different resource allocations, until the project plan fits the desired time frame.

During the lifetime of a project, actual data, such as the actual time or resource it took to perform a particular task, is collected and entered by the project manager. This information makes it possible to control the project progress and take appropriate actions, such as allocating more resources, if required. The quality of the project plan depends on the actual data being accurate, complete, collected, and entered into the project management system as early as possible.

Figure 1.12 Project Management

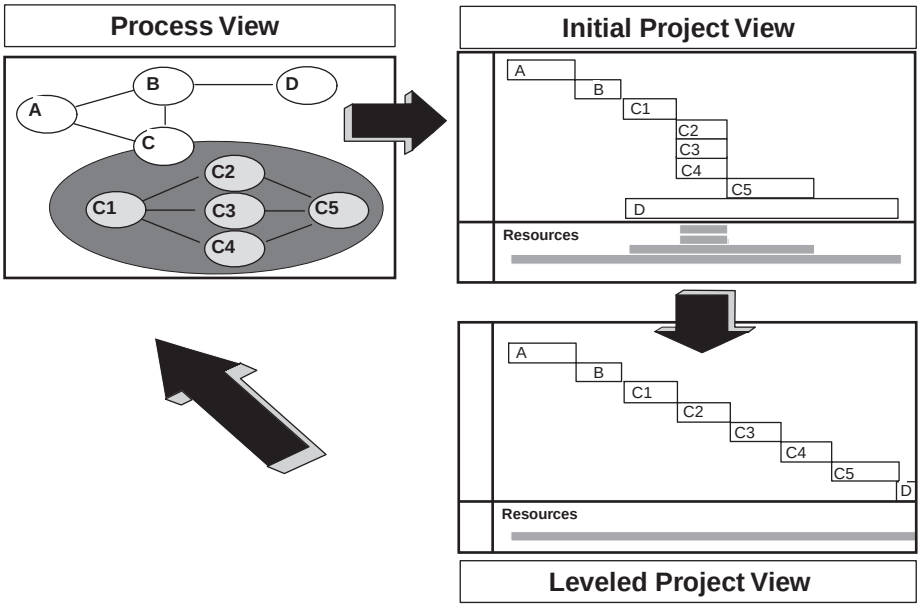


Figure 1.12 illustrates how workflow technology can significantly add to the capabilities of a project management system.

One models a project in the workflow management system as a business process either by modeling it directly in the workflow management system or by reverse-engineering it from an existing project plan in the project management system. We call this representation of a project the *process view*. When a new project is started, the process view is translated into the GANTT chart, which represents the *project view* of the project. This transformation is done under the assumption of unlimited resource and results in an initial project plan. The GANTT chart is then reworked using the available level of resources, until a satisfactory project plan has been found. We call the project to be leveled and the appropriate GANTT chart the *leveled project view*. When this stage has been reached, information in the GANTT chart that is process relevant, such as the amount and type of resources that are needed or the time scale of the project, is extracted and imported into the workflow management system. The workflow management system updates the process model with this information, for example, the assignment of roles to activities.

When the project is started, an appropriate process is created from the process model. This process now helps to control the execution of the process by proactively tracking the project progress. It collects the appropriate actual data from the project members and, if this data is not provided in a timely fashion, sends out reminders or even notifications to the project manager.

This is just one example of the advanced use of workflow technology. Chapter 9 provides an in-depth discussion of some additional examples.

1.17 System Requirements

In section 1.2 on page 3 we discussed process models as an enterprise resource. They therefore must be treated in the same way as data that the company owns. They must be protected against any accidental or intentional damage. Tight access control must be exercised to make sure that the structure of the business processes is not revealed to unauthorized users and that the structure of business processes cannot be altered with malicious intent. It is imperative that the business processes be executed exactly as defined.

Similar functionality has traditionally been delivered by transaction processing (TP) monitors. The structure of the business processes was an integral part of the application that was executed by the TP monitors. A production workflow management system therefore must provide the same level of operational and enterprise characteristics as that delivered by TP monitors. Actually, a production workflow management system must exceed these levels as workflow-based applications typically operate in a heterogeneous and distributed environment.

1.17.1 Operational Requirements

The following list highlights some of the more prominent operational requirements that a production workflow management system must fulfill.

- It must support *business transactions* and *global transactions*.

- It must be *reliable*. All internal operations must be performed as transactions. For example the invocation of an activity implementation, must be performed once and only once. The workflow management system must ensure that in the case of a failure, no completed action is undone and all actions that have not yet been completed are undone and started anew.

- It must show high *availability*. Actually, the workflow management system should support continuous operation; that means that the management of the underlying data and the communication software should not require the system to be quiesced even when changes, such as upgrading the different components of the workflow management system, are made to the workflow management system. High availability means that the workflow management system implements a system structure that allows the workflow management system to be up and running 24 hours a day 7 days a week.

- It must be designed for high *capacity*. Ideally, this would mean the support of an unlimited number of users working with the system and an unlimited number of processes maintained by the workflow management system.

- It must be *scalable*. This requires that the workflow management system be structured in such a way that additional resources either make the system perform faster or provide for higher throughput.

- It must provide the capability to record *process traces*. The trace information should include all relevant actions, such as the start of a process or the execution of an activity implementation. All trace information must be written to persistent storage. This trace information is usually called the *audit trail*. It allows for the monitoring of the system characteristics, such as the current number of actions being performed or the average duration of a process. It also helps check if the corresponding goals are met. And, it provides the data for statistical analysis, such as calculating the costs associated with carrying out a process.

1.17.2 Enterprise Requirements

Since workflow-based applications are typically executed in a heterogeneous and distributed environment, even within the same corporation, the workflow management system must also address issues that are important from an enterprise perspective.

- It must support *multiple platforms*. Platform refers to the operating system and network protocol the workflow management system operates on. In addition, it is required that the workflow management systems on the different platforms can communicate with each other. Platform also means that the workflow management system must be able to invoke applications that implement activities on different platforms using different invocation mechanisms.

  This flexibility is required to support the integration of a wide variety of existing programs, such as legacy applications running on a mainframe using IBM's Customer Information Control System (CICS) or IBM's Information Management System (IMS) as transaction monitor, or standard software such as SAP R/3.

- It must participate in *systems management*. This requirement has multiple facets. First, the workflow management system must be enabled to assist the systems management system in the distribution of workflow management system code to the various servers and clients. Second, it should assist the systems management system in the automatic distribution of the software each user needs to perform the assigned activities. Third, it must provide

the systems management system with information about its status. This information allows the systems management system to monitor the workflow management system and take proper actions if the workflow management system becomes unavailable for whatever reason.

- It must provide *central administration*. Each one of the individual workflow management systems offers administration functions. These administrative functions fall into two categories: the administration of processes and the administration of the workflow management system. It is important that both types of functions are available centrally.

- *Standards compliance* is an important requirement for a number of reasons.

  1. In a large enterprise, it is likely that workflow management systems of different vendors are involved in the execution of workflow. This variety requires that the different workflow management systems implement some standard that defines the exchange format between different workflow management systems.
  2. Uncontrollable events, such as the disappearance of a workflow management system vendor, may require the installation of a new vendor's workflow management system. It is important that (ideally) the activity implementations and process models need not be adapted or even rewritten. This desired flexibility requires that the workflow management system honors the standard that regulates the invocation of activity implementations and complies with a common metamodel.
  3. Central administration requires that the workflow management system implements the standard that defines the information to be returned and the actions to be taken as the result of an administration request.
  4. Statistical analysis of process traces requires that the audit trail written by the different workflow management systems follow the appropriate standard. Otherwise, significant pieces of information are lost.
  5. Users typically interact with the workflow management system by using a graphical user interface (GUI). This GUI is very identical throughout the company. This standardization is particularly important in cases where different workflow management systems are installed within a company. User should not need to learn to use a new GUI every time they change jobs. This need for continuity requires that the workflow management system implements the standard that defines the functions that a client performs. Without this standard, the customer would need to invest in various differing GUIs.

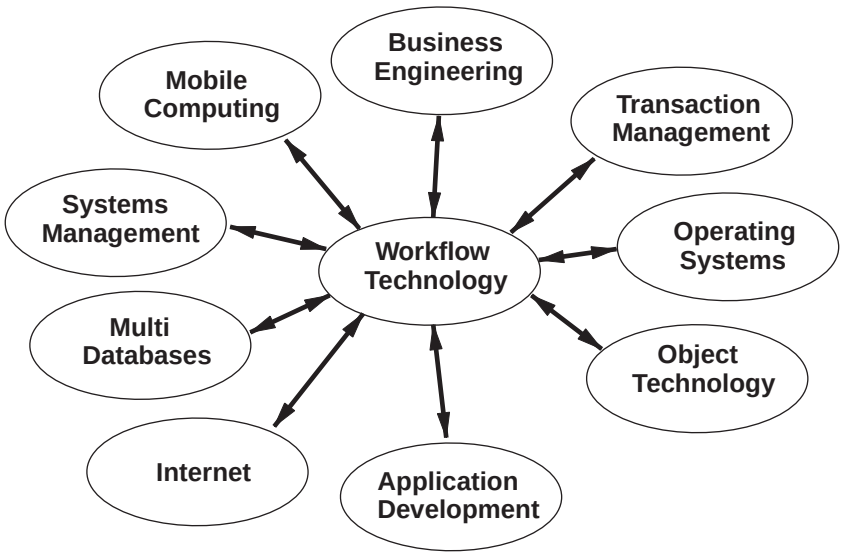
- A sophisticated *security* system is a definitive must for a workflow management system. We subsume under this requirement all aspects of security and authentication. These aspects include the protection of the workflow management system's data against any access coming outside the workflow management system, the encryption of messages that flow between different workflow management systems, the authentication of users when they access the workflow management system, and an elaborate access control scheme that allows the granular assignment of access rights to workflow management system entities such as process models, processes, and audit trail information.

1.18 Relation to Other Technologies

So far, we have sketched how workflow is related to object technology, transaction management, and project management. There are many additional software technologies that are more or less closely related to workflow. Figure 1.13 lists some of the more prominent ones. We start our discussion with business engineering and then continue clockwise around the circle.

Business Engineering is a collection of activities and techniques that helps define the optimal business processes for a company to make the company competitive. Those pieces of the collected information that can be managed by a computer are moved over to the workflow management system. We discuss business engineering and its relationship to workflow in Chapter 2.

Figure 1.13 Relations to Technologies



We described the support of global and business transactions in section 1.15 on page 20. Several later chapters address the relationship of workflow technology and transaction management. Chapter 9 elaborates on the use of transaction managers in applications, Chapter 7 provides an in-depth discussion of the transaction support needed in production workflow management systems, and Chapter 10 shows how the transactional support is realized in the workflow management system.

In section 1.10 on page 15, we described how a workflow management system is conceptually an application operating system and how the workflow management system is built on top of message queuing, database management, transaction management, and operating systems. In chapter 10, we show how such a workflow management system is implemented on top of those components and how it exploits each of these underlying systems.

The relationship between workflow and object technology has many facets. We discussed one of them, the use of objects as activity implementations, in section 1.9 on page 14. This application view is workflow-centric; we present an object-centric view of applications that use workflow in Chapter 6. We show how the structure of workflow-based applications can be derived from object-oriented analysis and design.

Development of applications is typically performed with the aid of an application development methodology particularly suited for the type of application that needs to be implemented. We devoted Chapter 11 to an in-depth discussion of a development methodology for workflow-based applications. We call this approach *process-based CASE* (computer-assisted software engineering). This approach starts with the global definition of business processes, using one of the many business process modeling tools, and derives from the collected information the workflow specifications and the structures of the databases that are used by the activity implementation. Another approach is to start with object-oriented analysis and design (OOA/D) and derive from this information the workflow specifications and database structures. We discuss this in detail in Chapters 6 and 11.

The Internet is the backbone of electronic commerce. It provides the mechanism that enables workflows to easily span companies all over the world. Users have access to workflow management system functions, such as starting a business process or querying the state of a particular business process, using their familiar browser, such as Netscape browser or Internet Explorer. For example, users can order a customized car and track its delivery. We elaborate on the relationship between workflow technology and the Internet in Chapter 2 in the context of business engineering and in Chapter 9 in the context of Web applications.

Applications often operate on multiple databases, where each database is managed by a different database management system. The involved database management systems must provide a particular behavior if the application needs to be executed as a transaction. The database management systems must operate as resource managers in a two-phase commit protocol. Not all database management systems

support this protocol, nor are all applications structured for the exploitation of these features: transaction integrity must be achieved differently. Workflow technology can provide the required transactional behavior by means of compensations spheres as discussed in section 1.15 on page 20.

Systems management systems help to manage the complete information technology infrastructure, including distributing code to the proper computers, observing the correct operation of applications, communication lines, processors, informing operations people about out-of-line situations, and taking predefined corrective actions. Applications must be instrumented to provide the proper information to the systems management system and to accept requests from the systems management system. Workflow-based applications are no exception. However, no particular instrumentation is required from the activity implementations; the workflow management system provides the appropriate hooks into the systems management, for example, to allow the systems management system to perform workload balancing. We discuss these aspects in detail in Chapter 10. In addition, the workflow management system's knowledge about IT structures and organizational information can be easily used to assist the systems management system in its tasks. The workflow management system, for example, knows which user needs which piece of code. This information is required by the systems management system to perform code distribution. We discuss those aspects in Chapter 8.

Not all participants in a business process use office workstations; sales or marketing people need to have access to information when they are on the road. Thus, workflow management systems must support the mobile user by providing access to the workflow management system at any time. In Chapter 6 we discuss the appropriate support in the context of standardization work of the OMG.