Peter Denning. 1994. The Fifteenth Level (Keynote Address). *Proceedings of ACM SIGMETRICS Conference on Measurement & Modeling of Computer Systems.* May, 1-4.

Appeared in *Proceedings of ACM SIGMETRICS Conference on Measurement & Modeling of Computer Systems.* May, 1994.

THE FIFTEEN LEVEL
(Keynote Address)

Peter J. Denning

Computer Science Department
George Mason University
Fairfax, VA 22030
pjd@cne.gmu.edu

There is a connection between operating systems and workflow and groupware systems that is not widely appreciated. It has profound implications for the design and performance evaluation of computer systems.

For over four decades the basic set of abstractions around which we structure operating systems has not changed. This is a testimony to the power and ruggedness of those abstractions, which have survived increases of processor speed and memory capacity exceeding six orders of magnitude, as well as an onslaught of parallel, distributed, and networked systems. This is about to change.

The burgeoning workflow software industry did about $200 million business in the U.S. in 1993 and is expected to reach $2.5 billion in 1996. Workflow systems are based on a different conception of work than that on which the entire architecture of operating systems is based. The new conception of work is concerned with the fulfillment of commitments made between people, whereas operating systems treat work as the execution of computational tasks. (1)

Since operating systems are the platform on which workflow systems rest, they will soon be forced to adapt to the new conception. This will compel the addition of a new abstraction, the business process, to the hierarchy of existing abstractions. It will also generate new meanings of performance evaluation and capacity planning.

Concerns and Breakdowns Experienced by Organizations

In the past decade, businesses that were not organized for consistent customer satisfaction have disappeared or met with wrenching hard times. The survivors have developed big concerns about how to reengineer themselves so that they can deliver products on time, offer services that consistently satisfy customers, maintain market credibility and reputation, and introduce new products and services faster than competitors. Many turned to information technology as a medium in which to do business and with which to manage business, but found little relief, prompting some doomsayers to argue that information technology is the cause of the crisis.

Those of us in scientific research and education institutions have been aggressive about using new technology, but we have fared no better. We have made computation a co-equal with theory and experiment as a paradigm of science. We are heavy users of electronic mail and the Internet, and we tout new discoveries arising in Internet collaborations. But our everyday concerns sound just like those of business people. In competing for research funds, we worry about track record, about reputation for keeping promises, about satisfying program managers, and about the credibility of our institutions, groups, and selves. In competing for scientific results, we worry about the accelerating change of the field, about being first to report our findings, and about our international standings. Our doomsayers openly wonder whether technology has accelerated the pace beyond our capacity to keep up.

The problem is not technology. By automating business and work processes that are not explicitly oriented toward customer satisfaction, organizations have accelerated the production of dissatisfied customers. It does not matter whether the organization is a business firm or a research institute. The problem is the design of processes for conducting business.

The information technology needed to support work tasks has traditionally come from the fields of operating and database systems. Because of the growing concern in organizations for management of work flows, operating-systems technology is being drawn into the management of organizations. The convergence of organizational work and computational work will produce major structural changes both in organizations and in the field of operating systems and will offer new challenges for performance analysts.

Work and Business Processes

A hundred years ago Frederick Taylor articulated the principle of "scientific management", a view that sees work as motions and activities of workers that can be planned and optimized by managers. With the introduction of the electronic computer 50 years ago, it became possible to draw a parallel between the motions and activities of workers and the flows and transformations of information; accordingly, the individual worker could be modeled as a function that processes input information into output information. These views were reinforced in the 1950s when Herbert Simon described management as decision-making by evaluating alternatives, and in the 1960s when Jay Forrester analyzed organizations as nonlinear feedback signal-processing systems. Today, these information interpretations of work have been embodied in workflow analysis systems such as Forester's Systems Dynamics and IDEF0, a dataflow analysis method originally designed for observing commodity flows in manufacturing plants. In all these systems, work is seen as the process of transforming the given inputs into the desired outputs.

Fernando Flores of Business Design Associates argues that the breakdowns experienced by organizations are a direct consequence of the input-output interpretation of work. He says that the effective management of organizations operating in a worldwide telecommunications network requires a shift in the conception of work (2).

The input-output conception of work pays scant attention to the human processes in which people make requests, agree on what will be done, who will do it, when it will be done by, and who must be satisfied with the result. As organizations and markets have burgeoned -- fertilized by information and communications technology -- failures to see and manage human processes have changed from an annoyance into a crisis.

Flores says that work is the fulfillment of commitments by one person to the satisfaction of another person. The fulfillment of a commitment takes place in the context of a closed loop that connects a customer and a performer. The loop consists of four segments connecting four events. First comes the formulation of a customer's request, culminating in its delivery to the performer. The second segment is negotiation on the conditions of satisfaction of the customer, culminating in an agreement between customer and performer. The next segment is performance, culminating in the performer's declaration that the work is done. The final segment is acceptance, culminating in the customer's declaration of satisfaction. (3)

In carrying out the work agreed to, a performer makes new requests. The performer thus becomes the customer of others, who agree in turn to take on pieces of the work. In this way a network of performers and customers comes into play for the fulfillment of the original request. In fact, either the customer or the performer can make further requests during any segment of the workflow, and thus the network can grow from any of the four segments. The loop connecting an external customer to the organization itself easily decomposes, fractal-like, into webs of constituent loops, until the roles of everyone in the organization are accounted for. The loop notation, structural elements, and networks are called ActionWorkflows and are copyright by Action Technologies, Inc. and Business Design Associates.

Incomplete work flows invariably cause breakdowns, and persistent incompletions give rise to complaints and bad moods that interfere with the ultimate purpose of work -- satisfaction of a customer. Loops can be incomplete in many ways. If no one has taken responsibility to see that a particular workflow completes satisfactorily (a "missing customer"), or if no one has taken responsibility for fulfilling a request (a "missing performer"), a satisfactory completion cannot be guaranteed. Many purely information-flow offices are like this: A form in an in-basket obscures the customer who used it to make a request, and the anonymity of an office hides the performers. A missing request, misunderstood conditions of satisfaction, or delayed delivery can leave the customer expecting actions that the performer will not supply or the performer supplying actions that the customer does not want.

In their consulting work with organizations, Business Design Associates have shown that organizational mood, market position, and customer satisfaction are directly and positively correlated with the degree to which everyone in the organization is contributing to the completion of the loops in which they participate. They have shown that persistent incompletions of loops leave trails of dissatisfied customers. Persistent dissatisfaction can produce bad moods -- most commonly distrust, resentment, and resignation -- that interfere with productive work and foster additional work through new work flows designed to deal with complaints. They have shown that workflow maps can be used to help people see and overcome such breakdowns.

It should now be apparent that every organization depends on three kinds of processes: materiel, information, and workflow. The first two are of the traditional input-output kind. Materiel processes deal with the movement of materials to particular places, where the materials are transformed, manipulated, consumed, or combined into new items. Information processes deal with the movement of information to particular places and people, who transform, manipulate, or consume it. Workflow processes are equally important; they deal with the requests for work to be done, agreements about what will be done, who is responsible for doing it, when it will be done, and whether the customer is satisfied with what has been done. The processing of materials and information are the consequences of people making and fulfilling commitments.

Technologies for modeling material and information processes are mature; queueing networks are an example familiar to this audience. But technologies for drawing maps of workflows and tracking the events that constitute them has not been widely available or appreciated.

It is crucially important to note that redesigning work flows for closure and simplicity is not a simple technological problem. The work flows are "the way things are done around here" -- part of the culture of the organization. Redesigning them means an alteration of that culture. Technology by itself is not sufficient to bring this about.

Operating Systems

The field of operating systems has historically provided much of the technology on which organizations depend for conducting their business. The first operating systems, built in the 1950s, were viewed as control programs that allocated scarce and expensive resources (processor time, main memory space, disk storage space, and access to input/output devices) among several contending users. These systems were seen as managers of work tasks submitted to the system by the users. The designer's horizon coincided with the boundaries of the system.

During the 1960s, the concern of operating-system designers was identifying abstractions that could be used to construct virtual machines for users; the abstractions that emerged included stack-structured execution of programs, interrupts, processes, virtual memory, files, directories, input/output streams, and "shell" programs for interpreting user commands. These abstractions constitute a hierarchy that has changed little since the 1960s (4,5).

15	Workflow process	10^{5}
14	User interface	10^{3}
13	User virtual machine	10^{1}
12	Directories	10^{-1}
11	I/O streams	10^{-2}
10	Devices	10^{-2}
9	Files	10^{-2}
8	Interprocess communication	10^{-2}
7	Virtual memory	10^{-2}
6	Local secondary storage	10^{-3}
5	Processes and semaphores	10^{-4}
4	Interrupts	10^{-5}
3	Procedures	10^{-6}
2	Instruction set	10^{-8}
1	Random access memory (RAM)	10^{-8}
0	Hardware electronics	10^{-12}

Operating systems have been viewed as "software that manages the flow of work inside a network of computers." The functions of operating systems can be interpreted as a hierarchy that spans operations at many time scales ranging from picoseconds at the hardware level to days at the user level (17 orders of magnitude). At each level, operations apply to objects visible at that level; the operations and objects are composed of operations and objects from lower levels. This principle of organization is much like the natural hierarchy of physical objects ranging from quarks to galaxy clusters. With the exception of graphical presentations, which replaced text-only shell programs in the 1980s, the levels of the operating-system hierarchy have been the same since 1970. In the 1980s, research on distributed operating systems showed how to extend the multimachine levels (8-14) so that their objects could be shared across the machines of a network. By the end of the 1990s, business processes may become the 15th abstraction in the hierarchy, and the interpretation of operating systems will shift to "software that assists in managing the flow of work in an organization."

During the 1970s, operating-systems designers were concerned with the efficient implementations of these abstractions on mainframe computers and with coupling them across data networks such as the ARPANET. In the 1980s, proliferating personal computers, desktop workstations, local-area networks, the worldwide Internet, and graphic displays focused the attention of operating-systems designers on extending the abstractions to systems of networked computers. Examples include abstractions that deal with the construction of large software systems from parts that run as processes or subsystems on their own computers. In spite of all the changes in hardware, software, and system scale that designers had to contend with during this period, the basic understanding of an operating system as "manager of work inside a system" did not change.

The gap between this interpretation of operating systems and the concerns of organizations has been brought into sharp focus by the widening realization that computers are for communications, not just for recording or processing. Users now complain vociferously about their inability to share files produced by different word processors, spreadsheets, drawing programs, and databases. They complain about their inability to deal with the growing volumes of information associated with the explosion in electronic mail, bulletin boards, and news groups. They complain about their inability to talk with distant collaborators and see their work. As a consequence, software vendors have begun to produce products such as network file systems, distributed databases, groupware, and work-sharing systems such as Lotus Notes. Although these subsystems have brought some order to the world of interacting computers, they are formulated in the old interpretation of work as information-processing tasks that flow inside a network of computers.

Add to this the new demands of mounting workflow management software on existing operating systems. New requests are coming to system administrators, system programmers, and network managers, for configurations, debugging, installations, upgrades, and maintenance; and to operating systems designers and performance analysts for more transparent and efficient support of the functions of workflow systems. At least one operating system manufacturer, MicroSoft, has declared that its next generation of operating systems will contain support for workflow. It will not be long before basic workflow concepts are taught in operating system courses.

These developments are forcing operating systems designers and performance analysts to understand the new conception of work. It is no longer sufficient to think of operating systems as tracking the flow of input-output specifying tasks; it is time to think of them as managers and facilitators of incomplete commitments in an organization. It is no longer sufficient to think of queueing networks as calculators of throughput and response time of input-output specifying tasks; it is time to extend them, or to find new models, to calculate the satisfactory production rates and cycle times of organizations.

The notion of a boundary between the computer system and the organization is disappearing and will be gone by the end of this decade. Operating systems will cease to be associated with the internal functioning of machines and will be associated instead with the functioning of organizations.

Raul Medina-Mora of Action Technologies, Inc., in Alameda, California, has been leading a design team to build technology that makes the connection (6). They call it the ActionWorkflow system. It has the following basic elements:

- A graphical work-flow mapping language that allows users to show, construct, and (with authorization) alter workflow maps.

- A database of workflow definitions that contains internal representations of business processes.

- A database of workflow transactions that contains records of each instance of a work flow created in response to a customer initiating a request.

- A workflow server that receives event notices from applications programs (including electronic mail) announcing events that represent progress in some workflow, and modifies the affected records in the transactions database. An interpreter running on this server notifies applications programs and other agents which actions are allowable next in a given workflow.

- A workflow reporting system that tracks and monitors workflows, answering questions from users about the status of particular requests and from managers about cycle times, congestion, throughput, and other performance metrics of business processes.

A New Kind of Performance Model

Queueing network models have become central to the performance evaluation of computer systems managed by traditional operating systems. These models suppose that a job (unit of work) is submitted by a user through an input port and that a response (work completion) is later received through an output port. Work is a specification of inputs to be acted on by a sequence of servers, producing desired outputs.

As operating systems shift to become facilitators of work in organizations, the new notion of work (making and fulfilling commitments) will replace current notion (functional activation). Queueing network models are not be sufficient to enable estimation of performance metrics for the organization served by the operating system. (For example, the token corresponding to request that starts a secondary workflow must eventually be rejoined with the original requesting token. Analytic methods such as those proposed by Gelenbe may be useful here (7).)

The history of queueing network models can guide a speculation about the coming evolution of workflow models. Although queueing models had been studied by engineers since the early part of the 20th century, they did not receive serious attention as a means to estimate throughput and response time of computer systems until 1965, when Alan Scherr demonstrated that a machine-repairman model worked well for a time-sharing system. This initiated a period during which the networks-of-queues models long known to theorists were used as a notation to map out servers, queues, and flows within computer systems. These maps were very useful in identifying and correcting congestion points in saturated systems. After the mapping stage came a period of measurement, in which instruments were designed to gather data from real systems that could be compared to the outputs of models. In due course, instrumentation that routinely enabled measurements consistent with the models were widely available, and allowed capacity planners to annotate maps (queueing network diagrams of systems) with throughputs, queue lengths, and response times. After the measurement stage came the period in which the models were routinely used to anticipate performance metrics that would be observed if the system or its workload were changed. This stage was enabled by the invention of efficient computational algorithms that would calculate performance estimates rapidly from given system and workload parameters. The first such algorithm was Jeffrey Buzen's central server model algorithm in 1971, which triggered a torrent of algorithms work for queueing models that continues to this day.

I think these three stages -- mapping, measurement, and anticipation -- will characterize the progression of performance models for organizations adopting the new conception of work. We are just now entering the mapping stage, with notations such as ActionWorkflow loops proving to be very useful in locating the sources of breakdowns in organizations (analogous to locating queueing bottlenecks). As the new workflow models are deployed and used to observe and measure the completions of commitments in organizations, we will enter the measurement stage. Within the next few years, algorithms will be developed that enable workflow notations to be used as models for calculating the future effects of proposed organizational changes. Progress here may be gradual or there may be a breakthrough analogous to Buzen's algorithm.

The big opportunity -- and challenge -- facing the performance analysis field is to embrace these changes and take the lead with the measurement and modeling work that is needed to bring us to the third, mature stage.

A New View of Systems

The loop interpretation of work flows offers a new way of thinking about systems. In the traditional way of thinking, which views a system as a collection of interacting components, it is not clear whether human beings are outsiders that interact with the system or whether they are themselves components. (A whole field of inquiry, focused on computers and human interaction, has emerged to investigate this question.) The work-flow interpretation integrates the actions of people and computers. Human requests, agreements, and satisfaction appear as the driving force of work. Computers appear as tools for designing, tracking, and facilitating work.

Over the coming years I expect the widespread organizational concern for building market position by building credibility for delivery and a track record of customer satisfaction to fuel an explosion of software and systems to assist organizations with managing their work. The outcome will be the addition of a new category -- the workflow process -- to the hierarchy of abstractions managed with the help of operating systems, and analyzed by capacity planners.

References

1. Thomas White and Layna Fischer, eds. 1994. *New Tools for New Times: The Workflow Paradigm.* Book Division, Future Strategies, Inc., 909 Marina Village Parkway, Suite 178, Alameda, CA 94501.

2. Fernando Flores. 1992. "Offering New Principles for a Shifting Business World." Business Design Associates, 2200 Powell Street, Suite 400, Emeryville, CA 94608.

3. Peter Denning. 1992. "Work is a closed loop process." *American Scientist 80*, 4 (July-August). 314-317.

4. Peter Denning and Robert Brown. 1984. "Operating Systems." *Scientific American 251*(September). 94-106.

5. Peter Denning and Walter Tichy. 1992. "Operating Systems." In *Encyclopedia of Computer Science and Engineering.* 3rd edition (A. Ralston and E. Reilly, Eds.). Van Nostrand Reinhold.

6. Raul Medina-Mora, Terry Winograd, Rodrigo Flores, and Fernando Flores. 1992. "Workflow Management Technology: Examples, Implementations, and New Directions." ACM SIGCHI/SIGOIS *Proceedings of Conference on Computer Supported Cooperative Work* (CSCW). November.

7. Erol Gelenbe. 1991. "Product form queueing networks with negative and positive customers." *J. Appl. Probability 28.* 656-663.