Measuring Information Latency

Rani Hoitash. Suffolk University. USA.
rhoitash@suffolk.edu

Alexander Kogan. Rutgers University. USA.
kogan@rutgers.edu

Rajendra P. Srivastava. The University of Kansas. USA.
rajendra@falcon.cc.ukans.edu

Miklos A. Vasarhelyi. Rutgers University. USA.
miklosv@rutgers.edu

Abstract. Many companies have recently digitized a substantial number of their business processes. It is often the case that some companies store their data on a digital medium and that this data is can be captured and stored in real-time or in a batch process. Measuring the time it takes a system to traverse and process information is of the utmost importance to internal and external stakeholders of the organization. However, it is often difficult for companies to identify information flow and process bottlenecks. In this paper we develop an approach for evaluating the level of companies' digitization. We formalize and provide an illustrative example for a model that can facilitate the measurement of information flow latency within organizations and discuss the benefits associated with the model.

The proposed latency measure is unique in that it facilitates the measurement of information flow and process latency. The measure is a very important in that it can assist management in improving the most critical business processes in their company. Moreover, an assurance service can be developed in order to increase the trust in the digitization level of business partners. Using the proposed latency measurement model a future assurance service can provide supply chain partners the confidence that their business partners can meet delivery and production schedules accurately and on a timely basis.

Keywords: Digitization, Latency, Data Flow Diagram, Assurance.

1. INTRODUCTION

The advent of Internet technologies, hardware solutions, telecommunication networks, and computer applications has grown rapidly over the last two decades. This fast growth resulted in a fast changing business environment and increased productivity. However, it has also created new challenges for companies forcing them to adjust their business strategies, reengineer their business processes, increase their service offerings, and reduce their costs. Technology has opened up many new channels of communication and commerce between companies and customers, companies and shareholders, and companies' employees. Trading partners can interact in a completely digitized environment and conduct many of their operations without any human intervention. Many forward-looking companies are leveraging new technologies to increase their productivity, reduce their cost and create more value for their shareholders.

With the fast ever-changing technology environment and with the high cost of acquiring, integrating, and implementing new technologies, businesses need to invest their limited resources in technology while maintaining the focus on their core competency. Therefore, in order to effectively invest in new technologies maintain a high level of service and maximize financial performance in the digital environment it is critical for companies to evaluate their existing level of digitization and find ways to improve their most critical business processes. Since computers are limited to their ability to process information rather then optimize business processes, it is up to managers to design business processes in which the usefulness of computers is maximized. Conceivably, most transaction processing today can be achieved flawlessly, nevertheless, the burden on processing is not constraint by computer power, but rather by the effectiveness in which business processes are designed and integrated.

Digitized seamless information flow and processes are highly desirable. A seamless way to capture data, propagate data and traverse it can potentially be associated with an increase in the level of productivity, service, and customers' satisfaction. However, the use of un-digitized data and processes can contribute to a number of deficiencies that could potentially result in added direct or indirect cost to organizations. Some of the benefits that are associated with digitized activities are as follows:

- Capturing digital data usually results in a broader scope of the captured data as opposed to a narrow scope of data that is captured by un-digitized activities. Specifically, capturing digitized data for the purpose of monitoring inventory quantities could be attained using digital and manual techniques. RFID (Radio Frequency Identification) is one technology that allows companies to record accurately the location and availability of inventory within its premises. This approach is currently the best technology for supporting perpetual inventory management systems. Information such as quantities, physical location within the warehouse, and timestamps are only a few of the benefits associated with this technology. Such richness of information could never be captured using un-digitized technologies.

- The probability of introducing corrupt information is associated with data that is not captured, communicated, processed, and stored digitally. Specifically, with the appropriate controls, digitized information could virtually be error free. Though, it is beyond the scope of the current paper, potentially the probability for errors in each process could be estimated by asking users to estimate the initial probabilities for errors in an un-digitized activity. This probability could than be applied at the lowest DFD level in which each data transformer performs a simple operation that does not need to be further decomposed and subsequently, probabilities can be propagated through the system (following the uni-directional graph).

- Lastly, un-digitized information could potentially lead to unnecessary monitory cost, which is the loss due to delayed propagation, process, and extraction of information. The loss of the timelines of information may result in production delays, reduction in customer dissatisfaction, idle machines and other potential costs.

So, it can be argued that digitized data communication and transformation could lead to a minimal cost through the use of continuous communication, on-line processing, and digital data stores. The above limitations of undigitized activities are significant enough to warrant the development of a model that can assist companies in identifying inefficiencies in the process and flow of information within their systems.

The main purpose of this paper is to propose a set of measures that can capture the level of digitization of companies. In particular, we develop a specific methodology for measuring information flow latency within companies. The approach is layered on top of data flow diagram (DFD) which we have determined to be specifically suitable for this task. The rest of the paper is structured as follows, in section 2 we address the need for flawless information flow and processing and discuss the environment that e-businesses face. In section 3 we present various issues related to information technology productivity and risks. In section 4 we introduce the concept of measuring information latency using DFDs and provide the corresponding logic for doing so. In section 5 we illustrate the process of measuring information latency using the proposed approach and section 6 follows with summary and direction for future research.

2.   E-BUSINESSES VALUE CHAIN

Corporations generate, store, share, and use large quantity of business data across the organization and beyond organizational boundaries. More and more data is being used for decision-making, operation planning, channel integration, customer service, and administrative tasks. Advances in network computing technologies in recent years created a new environment for business, namely, E-Business. E-Business activities often face a different value chain structure. For example, Slywotzky and Morrison (2000) suggested that customers could initiate the value creation process. Specifically, the following sequence of events may apply for digital businesses:

Customers>Channels>Offering>Inputs>Assets

According to this sequence of events customers make their decisions and communicate them digitally to the business, thus initiating the value creation activities. This type of new information flow can be modeled for highly digitized companies, such as Dell Computers (Slywotzky and Morrison 2000) where production and planning are triggered by customers' orders. Therefore, similar new business models require businesses to respond promptly and accurately to customers' demand.

As an illustration, let's think of a company similar to Dell which has an integrated data repository. A customer places an order on the corporate website

(therefore, initiates the value creation), the system (possibly an ERP system) instantaneously triggers an update to the production schedule, to the shipping schedule, initiates the ordering of parts that are needed for the production (assembly), and updates the sales report. This type of seamless information flow is highly desirable and could be completed with minimum human intervention and process latency. Such desirable level of digitization requires management to thoroughly analyze existing business processes and identify potential bottlenecks in the flow and process of information.

Businesses that operate in the E-Business environment not only have to digitize their own operations and information infrastructure but also need an assurance that their trading partners, who are now part of their integrated value chain, are also sufficiently digitized. Papazoglou et al. (2000) extended Porter's (1985) view of the traditional value chain to an intra-organizational view. He presented a model in which each organization focuses on its core competency, and multiple organizations contribute directly to each other's value chain. The author identified four key driving forces that would enable successful development and deployment of integrated value system applications: (1) new business models (2) cross enterprise interoperability (3) adoptable and flexible business processes (4) organization infrastructure. Another example for the joint need of companies to become digitized is the Merge-in-Transit (MiT) strategy. In this model shipments are collected from multiple origin points and are consolidated in transit to a single shipment to the customer. O'Leary (2000) showed that the reengineered process of MiT requires timely digitized data interchange and results in the following benefits: Reduced transportation costs, improved customer service, lower obsolescence, and lower capital requirements.   MiT and the integrated value chain are only two examples in which a digitized environment is essential. It is therefore expected that companies will seek some time of assurance from their business partners regarding their level of digitization.

Digitization of companies entails substantial investments in IT, which in turn may or may not result in the appropriate return on investment (ROI). Although it may seem intuitive that investments in IT contribute to an increase in productivity, discerning the effect of IT on productivity is a difficult task. The value of IT investment has been the subject of many research papers and will be discussed bellow.

3. INFORMATION TECHNOLOGY AND PRODUCTIVITY

Information Technology introduces a great dilemma to management. On the one hand IT investments can lead to an increase in productivity, enable the company to better compete with its peers, and resolve unfulfilled user needs and deficiencies in older technologies. Conversely, IT investments can diverge critical resources from the organizations' core competency, they may not necessarily be optimal, and often introduce organizational risks. The issue of the value of IT has generated substantial interest among both academics and practitioners. This interest was primarily triggered by the large investments in IT during the 80s and 90s. Since many of the benefits of using IT are intangible in nature, the productivity of IT is very hard to measure. IT is often viewed as a black box in which it is hard to separate the portion of productivity in a particular process due to IT and the productivity due to other factors. Brynjolfsson (1993) presented the productivity paradox, which indicated that computers were not delivering the value promised. He raised some concerns regarding the measures used in the estimation of IT productivity. He therefore suggested four possible reasons for the failure to capture improved productivity, namely: measurement error, time lag, redistribution of activities among firms, and misallocation and over-consumption of information by managers. Kumar (1997) used a model of real option valuations to show the value of improved responsiveness resulting from IT investments. He was able to quantify the intangible benefit of IT through the improved responsiveness resulting from IT investments.

Although the benefits of IT are hard to measure, companies continue to heavily invest in IT. Investing in the right technology is not the topic of this paper, however, we do propose an approach that could potentially assist management in allocating their investments more effectively. Through the use of the digitization metrics management can direct IT investments to the most critical business processes in the organization.

IT Risk

In additions to its benefits, IT introduces new risks to the organizations. The greater the reliance of companies on IT the more important it is to create the

appropriate controls to maintain the reliability and security of the system. System reliability is defined by the AICPA/CICA (2000) as follows: "A reliable system is one that is capable of operating without material error, fault, or failure during a specified period in a specified environment". As part of developing a new system assurance service (SYSTRUST). The AICPA identified four principles that can be used evaluate the reliability of a system: (1) Availability (2) Security (3) Integrity (4) Maintainability.

In the current paper we do not directly address the issue of system reliability. We do note that potentially it is possible to measure the probability of errors given digitized vs. un-digitized activities. This type of study could complement the current study through enabling managers to allocate scarce resources towards IT investments based on both latency of the information flow and the probability that errors could occur.

In the following sections we present our approach for measuring information latency using data flow diagrams.

4. DIGITIZATION METRICS

In this section we introduce a new approach to measure information latency within a system. Effective processing and information flow leads to reliable data transformation while efficient processing and transformation leads to digitized frictionless processing and flow of data. We believe that processing and delivering information digitally should result in reduced risk through increased reliability and reduced cost through minimizing human intervention. Non-digitized business processes often result in information degradation and latency, which in turn can be reflected in the bottom line financial performance of organizations.

Identifying the digitization level of companies could lead to many intangible benefits. Being able to analyze business processes analytically, and assign processes a numerical ranking should be extremely valuable. The assigned values could assist management in channeling investments to improve critical business processes, and consequently generate higher cost savings. Therefore, resolving process bottlenecks through improving non-digitized business processes and attesting/assuring on the current level of digitization may result in lower cost of capital to the company, and

higher value to its shareholders. Additionally, if process and flow latency values are available, auditors, as part of an assurance engagement could examine these values and attest on management's assertions regarding their accuracy.

The flow of information within and between organizations can be defined as "digitized" as long as there are no discontinuities in the real time flow of data, including the initial capture of information, information processing, information storage, and information communication. The digitization of companies can be evaluated based on the level of degradation and disruptions in the initial capture of information, communication and information flow, and the processing of that information. Disruptive activities can occur by either capturing a narrowed scope of the data (or data is that is already aggregated), a manual intervention, or the loss of timeliness (not real time communication). Disruptive flows of information can result in higher processing cost, time delay, and substantial increase in the potential risk to the organization. In this study the flow of information within organizations is modeled using a revised data flow diagram approach.

Latency can be defined as a delay that information flow experience from the source to its destination. Latency is quite often introduced because information needs to traverse through the system and/or because of the need to process information. Frequently, latency depends on the level of digitization, the computer configuration and infrastructure of the system. Specifically, manual flow of data (physically moving paper from one department to the next) and manual data processing (performing data manipulation manually) introduces the greatest information latency within a system. However delays can exist in digital processes and flows due to the lack of suitable infrastructure (networking capabilities), and/or system configuration issues (such as batch processing). Our approach for measuring information latency is independent of the level of digitization of companies and can be applied to information processes that are partially digitized and partially manual. In the following section we will discuss our motivation for using data flow diagrams for that purpose.

A data flow diagram (DFD) graphically describes the flow of data within a system. It is used to document existing systems and to plan and design new ones. A DFD is composed of four basic elements: data sources and destinations (external processes), data flows, transformation processes (data transformer), and data stores.

The DFD is a directed graph consisting of the four primitives mentioned above. In our model we look at a sequential system where we assume that data flows between two data transformers (processes) are direct in the sense that data transformers that use the data flow, as input cannot be called before the data transformer that uses the data flow as output. The basic elements of the DFD are presented in figure 1.
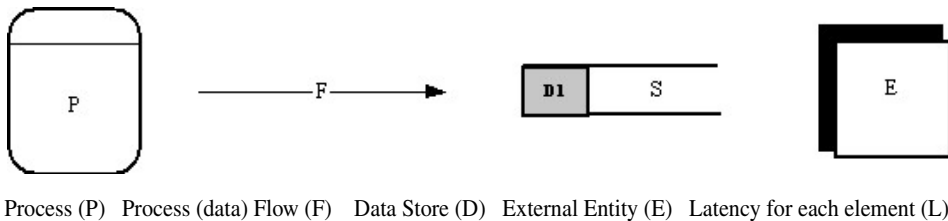

Process (P)   Process (data) Flow (F)   Data Store (D)   External Entity (E)   Latency for each element (L)

Figure 1. Elements of DFD

- Data source- represents external entities (another department, business, person…) that are not part of the system (or process) but belong to the outside world (external to the system that is being modeled). They are used to show where the input to the system is coming from and where the output of the system is going to.

- Data flows- data flows are represented as arrows connecting one data transformer to another. They represent a flow of data between the data transformers they connect. The flow of data in this case is unidirectional in the direction of the arrow.

- Data transformers (processes) - data transformers denote a transformation from an arbitrary number of input values to an arbitrary number of output values.

- Data stores- data stores provide a representation for storage of data.

DFDs are specifically useful for modeling information flow within a system. In the process of considering the most appropriate diagramming technique to model information latency we have evaluated multiple techniques. These techniques included ERD (Entity Relationship Diagrams) which are designed for logical data modeling rather then logical process modeling, EPCs Event Driven Process Chain (Keller, N¨uttgens and Scheer 1992 that have become a widespread technique for process modeling since the success of ERP systems such as SAP, and other process

modeling techniques geared towards BPR (Business Process Reengineering). We have concluded that the most suitable technique to diagram information latency in multiple level of detail is DFD.

The traditional DFDs have some limitations, including: (1) No indication for which event triggers a process, meaning, it does not indicate why a process begins. (2) A DFD represents only information flow and processes and may exclude some business processes that do not require information flow. (3) In a DFD there is no way to represent the logic of the flow such as conditional events. However, the main advantage is that the flow of information can be diagramed in various levels of detail. Following a top down approach, the diagram can be viewed from a general view to a specific view. Using the extended DFD model, we intend to measure information flow and process latency by drilling down only as much as needed to unambiguously categorize the data flow representation, processes and data storage.

DFD research was primarily used for the purposes of software design and maintenance. Ward (1986) introduced an extension to the DFD entitled a transformation schema. He introduced the ability to represent control and timing aspects of the system. We are currently not a ware of any studies that adopted the use of DFD for the purpose of process measurement. Though, traditionally, DFD have been used solely for documentation purposes their unique design could potentially facilitate numerous applications such as representation and aggregation of internal controls.

As discussed earlier, information latency resulting from manual information processing, and delayed propagation may be costly to organizations. We devised a method to measure the information latency within a system in the following way: every flow, process, and storage of information is given a latency value, which represents the time it takes to accomplish the task. These values are additive, and therefore are added to one another to determine the total information latency for each outflow of a system represented at the level 0 diagram. The latency values are assigned to each element in the DFD by the system analyst and can be added along the directed graph.

In the process of measuring information latency the system analyst need to assign latency values. There are various techniques that can be used by the

analyst to determine the latency values. Latency values for manual processes and information flows need to be extracted from the users of the system through, questioners, interviews, manual log records or observation. Any of the above or a combination of them can facilitate the extraction of these values. As for digital flow and processes, the analyst will need to examine log files, and time stamps to assess the duration of each task. If logs and timestamps are not available, the analyst will use some of the extraction techniques for the manual processes as described above. Once values are collected, the analyst will allocate them to particular elements on the DFD and will calculate the latency.

Similarly to Larsen et al. (1994) we view DFDs as directed acyclic graph at each level of the DFD hierarchy. Doing so we can uniquely identify an acyclic path for each output at each level of the DFD. Though each path can be unambiguously identified, a particular output can have multiple paths that lead to it. Therefore, we sum the latency values by traversing backwards through the DFD. The multi-step process of calculating the aggregated latency values is described below followed by a formal notation.

- Latency values are given to information flows at the level 0 DFD (this is the level in which a system represented in the context diagram is represented by its top level processes). By construction, in DFDs the latency values of information inflows and outflows are not expected to differ in different levels of representation (e.g. level 0, level 1…).Therefore in our model we absorb the latency of the information inflows into their target process.

- Each process at the level 0 DFD diagram is drilled down to the extent to which a unified level of latency can be identified with the information inflows and outflows. In other words, if ambiguity regarding the information latency of data inflows or data outflows remains after drilling down into a process we should keep drilling until this ambiguity is resolved. Unique information latency is identified when no process in the system has a different latency value contributing to the outflow of the data. Even if one unique latency value is identified for the process we may still need to drill down to a child level diagram in order to find out the value of that latency. Specifically, if that process comprises of a number of sub-processes, the value of the process will be found by further drilling down.

- Once unique information latency is identified for each process, latency values are aggregated to the level 0 diagram. Since we are interested in identifying the information latency for each outflow at the level 0 diagram, different outflow latency values will be propagated to the aggregated (level 0) diagram. The minimum value of each process's latency will be assigned to the parent process. The differences between minimum values of outflow latencies will be added to the outflow latencies that were assigned at stage 1.   The procedure facilitates an analysis processes with multiple number of inflows and latency values.

- The values at the level 0 diagram will be added along the information flow and a latency value is calculated for each outflow.

During the system development stage (or the documentation of existing systems) an analyst would document the process flow in the system using DFDs. Latency values are assigned to the elements in the DFDs based on the description provided the previous section. In our design, we assume the following regarding elements and their latencies.

- Each flow traverses between two points A and B. For each flow shown on the diagram we can assign a measure of latency.

- If a latency value is assigned to a process it is required that this value is the unambiguous value that it takes a particular process to produce an output. Specifically, we assign latency to a process only if all the outputs/outflows are generated with that latency.

- Processes are drilled down to the level of detail in which every process have a latency value associated with it.

- After drilling down, we introduce a buffer for each input flow. The buffer absorbs that latency of the inflow into the process so that the new latency value becomes part of the process. If we can uniquely identify the latency for each process at the level 1 DFD we will stop drilling. Otherwise we will continue to drill until on unique latency value can be assigned to each process.

For each process P we need to identify a unique latency value L. Additionally, for each data flow F we need to identify a unique starting point (when does the process output that flow) and a unique flow latency. Unique inflow latency can be assigned independently of the process latency. However, process output flows may differ and are therefore ambiguous when multiple outputs (F1…Fn) exist for a particular process P. Additionally, a process P with more then one outflows and more then one inflow can have an ambiguous latency value because it is not clear which inflows are needed to generate a certain outflow. Specifically, in figure 2 bellow it is not possible to identify which inflows are needed (F1, F2) to produce outflows F3 and F4 and whether it takes the same time for process P to produce/ calculate the necessary F3, F4 outflows. For this reasons we conclude that a more detailed representation is needed. Thus, we need to solve the ambiguity by drilling down into a more detailed representation of process P.
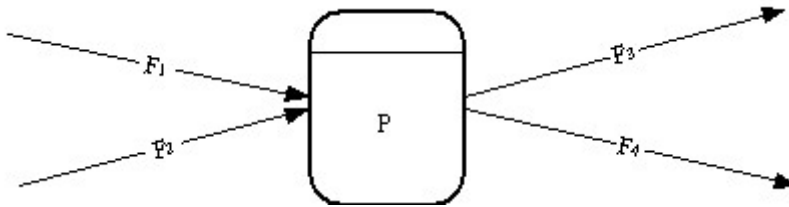

Figure 2. Process outflow ambiguity

As mentioned above, we drill down because we cannot identify a unique latency for each outflow from a process. Once we are able to identify a unique latency value for each outflow, we can proceed to the calculation of the latency. The procedure for calculating the latency values is formalized bellow.
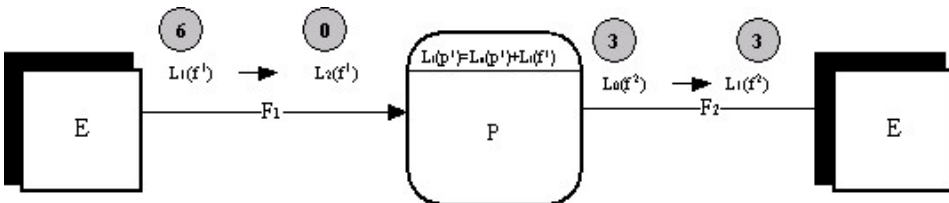

Figure 3. Unambiguous DFD

Figure 3 presents the procedure for calculating latency values using an unambiguous DFD. Each flow from A to B is assigned a latency value L0(f), this value represents the initial latency value of the flow. If a process has only 1 outflow, then there is not ambiguity associated with the latency of that outflow.

Specifically, as long as there is only 1 output, if there are more then 1 inputs to that process, the maximum value of these inputs is used. If $L_0(f)$ is the latency flowing from an external entity, then $L_0(f)= L_1(f)$, otherwise $L_0(f) \leq L_1(f)$. The latency of $L_1(f)$ (in this example $L_0(f)= L_1(f)$) is absorbed into process P, so that $L_1(f)$ becomes $L_2(f)$ and equals 0. Specifically, after absorbing the latency of the inflow into process P, the new latency of the inflow becomes 0 while the process latency is increasing by $L_1(f)$. The new process latency equals the initial value plus the latency of the inflow $L_1(p)= L_0(p)+ L_1(f)$. Absorbing the latency into the process can remove the ambiguity (if exists) arising from multiple inflows. In the illustration above the latency of the process from the input to the output would comprise of $L_1(p^1)+L_1(f^2)$.

The less trivial scenario involves an example similar to the one on figure 2. In figure 4 we display an DFD in which it is not possible uniquely identify the output latency for F3 and F4 at that level. Therefore we drill down to provide a more thorough representation. In this example the ambiguity results from process P1. In order to calculate the process latency for flows F3 and F4 we need to know which inputs are needed to generate that output (F_1 and/or F_2) and how long it takes to process P^1 to process the needed data for F_3 and F_4. The final latency values are displayed on figure 4. In order to derive these values there is a need to drill down into process P^1 to the level 1 DFD. The level 1 DFD diagram displays all the sub-processes of process P^1 and the inflows and outflows from that process. Figure 5 displays the level 1 DFD that was derived from figure 4. Process 1 was broken into four sub-processes with one output from each one of these processes. If any of these sub-processes had more then 1 output, a further drill down into level 2 of that process would have been necessary in order to resolve the latency ambiguity. As mentioned, at the level 1 DFD diagram, we absorb the inflow latency into the process. Specifically, F_1 which have a latency of 6 seconds is absorbed into sub-process $P^{1.1}$ so that the new latency for that process equals to the sum of the initially assigned latency (5) and the inflow latency (6). So that $L_1(p^{1.1})= L_0(p^{1.1})+ L_1(f^1)=11$.

Similarly, the value for process $P_{1.2}$ is calculated as 21. Since process 1 has two outputs, F_3 and F_4 there is a need to assign a different latency values for each one of these outputs.
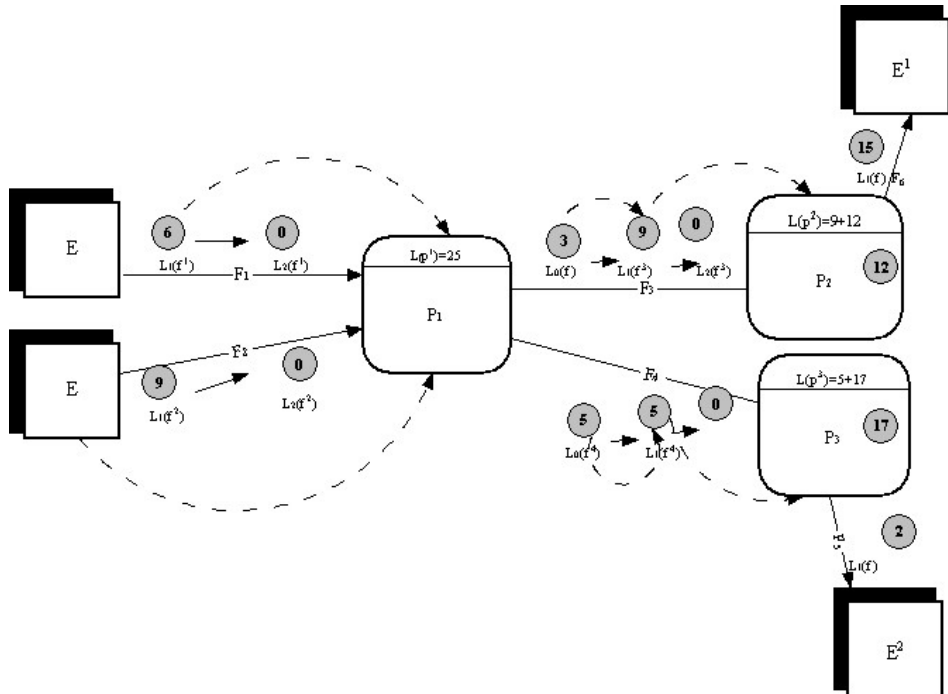
Figure 4. An Ambiguous level 0 DFD

For every output (F_3 and F_4) we calculate the maximum latency it would take to get to that output by summing up all the elements (process, flow, and store) that lead to that output. The reason is that some outputs can have multiple paths along the diagram, and the longest path that lead to each output should be used.

$L^{max}(f^i) = MAX\{SUM\ L(e_i....e_n), SUM\ L(e_1....e_n), ..., \}$

In the example bellow there are two paths for outflow F_3 so that:

$L^{max}(f^3) = [MAX\{SUM\ L(e_i....e_n) = MAX[11+3+9=23, 21+1+9=31] = 31$

And there is only one path to outflow F_4 so that:

$L^{max}(f^4) = 21+1+3 = 25$

After establishing the latency for F3 and F4 (as 31 and 25 respectively) we need to assign a latency value to process P^1 at the level 0 DFD. This value should help in resolving the differences between the latencies of outflows F_3 and F_4 at that level. This goal is achieved by taking the minimum value of F_3 and F_4 and assigning that value to process P^1 as $L(P^1)$.

L(p)= MIN($L^{max}(f^1)$ .... $L^{max}(f^i)$) and in our example

L(p^1)= MIN($L^{max}(f^3)$, $L^{max}(f^4)$)=MIN(25,31)=25

In order to represent the differences between the time it takes to produce outflows F_3 and F_4 we add the differences between the value of that process and the process outflow to that outflow. So that:

$L_1(f^i)$ = $L_0(f^i)$+ ($L^{max}(f^i)$ - [MIN($L^{max}(f^1)$ .... $L^{max}(f^i)$)] =

$L_1(f^i)$ = $L_0(f^i)$+ ($L^{max}(f^i)$ - L(p)).

Specifically,

$L_1(f^3)$ =3+(31-25)=9

$L_1(f^4)$ =5+(25-25)=5

Now the latency for process P^1 is 25 and the new latency for $L_1(f^3)$ and $L_1(f^4)$ is 9 and 5 respectively. It is now possible to calculate the latency for each outflow from the system. The latency into E^1 is calculated as:
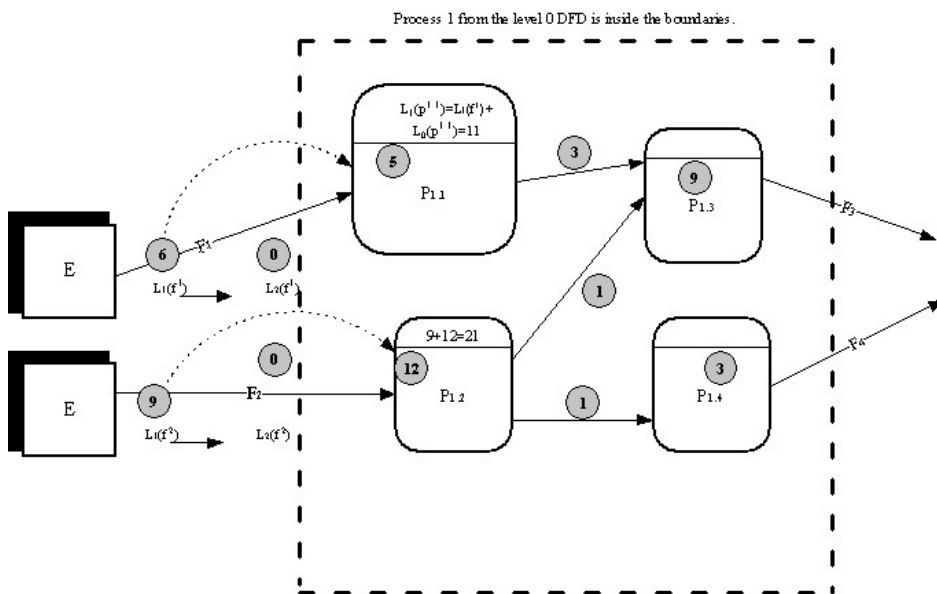
L(E^1)= 25+12+15=52

L(E^2)= 25+17+2=44.

Figure 5. Level 1 DFD

Though many companies are partially digitized, most companies are not completely digitized. Many processes require human intervention, are done manually and are entered and processed using batch input and processing techniques. Using the proposed structure for measuring information latency can be very beneficial for companies. Management could use the proposed measure of information latency through identifying critical business processes for their organization and assessing the cost benefit analysis from having the digitized. Specifically, management will have to balance the cost of not digitizing manual or batch processes with the cost of investing in the technology for having frictionless process.

5.   ILLUSTRATIVE EXAMPLE

The following illustration is designed to demonstrate how to measure information latency in a system using DFDs. In this illustration we arbitrarily assigned latency values to flows and processes. However, as mentioned earlier the analyst should extract these values in a systematic manner. Figure 6 displays an illustration of the context diagram of the FilmMagic Video Rental stores presented by Kendall and Kendall (1999). The context DFD represents the entire view of the video rental system including four external entities to the system: Video purchase system, customer, management, and accounting. Figure 7 presents the level 0 DFD in which the main processes are illustrated. As mentioned earlier information inflows and information outflows (from and two external entities to the process) are static across the different DFDs levels. In other words, the same input and output to the system are illustrated in the context diagram, the level 0 DFD and child diagrams. In this example will illustrate how latency values are aggregated from child processes in the drilled down diagrams to the level 0 diagram. At the level 0 DFD we assign latency values to the information inflows and to the information outflows these values are referred to as L0(f). Information inflows in this illustration are generated by the customer and by the video purchase system, whereas information outflows are directed to the customer, management and the accounting entities. For example, a customer provides information about the video rental item, their customer ID, and payment. Subsequently, the customer receives a receipt, promotion letters, and a video card.

Process 1 in Figure 7 has 4 information outflows; consequently, we cannot identify a unique latency value that can apply to all outflows. Therefore we drill

down into process 1 and get to level 1 DFD as illustrated in Figure 8. Drilling into process 1 allows us to examine the detail sub-processes and flows that are unique to process 1. Process 1 still has 4 inputs and 4 outputs (as in level 0 DFD) and for illustration purposes is surrounded by a border. Process one result in the following outflows: cash transaction, rental receipt, rental information, and an update to the customer record. At this stage we use Figure 8 to try and find unique latency values for each outflow at the level 0 DFD. Again this is possible because there is a one-to-one mapping between parent level and child level diagram for all the inflows and the outflows to and from a process. We illustrate below how we calculate the latency value for each outflow from process 1. As illustrated in the previous section, we introduce a buffer at the level 1 DFD diagram to absorb the latency of the inflow (which is external to the process) into the process. Therefore, the first stage involves the assimilation of the inflows into the process. For example, the initial latency for process 1.5 was 30 seconds. After adding the customer ID latency into that process, then new latency value is 35. In a similar way we absorb the remaining three inflows into their corresponding processes.

- The "Cash Transaction" outflow latency can be uniquely identified by adding: "Find Customer Record" process latency (35 sec), "Customer Record" flow latency (2 sec), "Get Video Record" process latency (13 sec), "Rental Information" flow latency (5 seconds) to the "Get Consumer Payment" (40 seconds) process. (Total latency of 95 seconds).

- The "Rental Receipt" outflow latency is the aggregation of: the "Cash Transaction" outflow (95 sec) the, "Rental Information/Payment " flow (5 sec), "Update Customer Record" process (7 sec), the "Rental Information" flow (4 sec), and the "Produce Customer Receipt" process (20 sec), resulting in the total of 131 seconds. (There is an alternative path, but as mentioned earlier we take the MAX for each path).

- The "Rental Information" outflow latency is the aggregation of the "Cash Transaction" outflow (95 sec), the "Rental Information Payment" flow (5 sec) "Update Customer Record" process (7 sec), and the "Rental Information" flow (4 sec) latency, resulting in a total of 111 sec.

After identifying a unique latency value for each outflow at the level 1 DFD, these values need to be aggregated to the parent (level 0 DFD) level. To do that, the

latency value that will be applied to process 1 at the level 0 DFD is the MIN (95, 111,131). In other words, the minimum value[1] represents the shortest time that it would take a processes to provide an output. In our example 95 is the minimum value, and therefore the latency value that will be assigned to process 1 at the level 0 DFD is 95. This value is the true latency of the process to provide the "Cash Transaction" output. However, other outflows experience longer delays. The solution is to add the differences between the latency of the process (95) and the latency of other outflows (111 or 131) so that 131-95=36. The value of 36 states that it takes 36 seconds longer to create the rental receipt outflow. Therefore, we add 36 to the existing outflow value (15) to receive a total latency value of 51 (this value is $L_1(f)$). Similarly, for the "Rental Information" outflow the difference between 111 and 95 will be added to the existing flow latency (25) resulting in a total flow latency of 41 (111-95+25).

After aggregating these latency values into the level 0 DFD it is possible to calculate the latency for the "management report" and the "rental receipt" outflows (there are other outflows to which we did not calculate latencies). The Latency for the "Management Report" outflow is the sum of P1, P2 (both of which incorporate the latency of the inflow into the process) and the "management report" outflow (95+56+5=156). The latency for the "rental receipt" output is 95+51=146.

6. DIRECTIONS OF FUTURE RESEARCH

More than ever before, companies are driven by information. In this study we proposed a technique for measuring information latency within a system. Our approach is layered on top of a DFD allowing us to measure information latency at different levels of detail. The proposed technique could assist management in channeling their scarce resources to critical business processes. This in turn could lead to many tangible and intangible benefits to the organization.

New attestation services in the form of measuring and assuring the digitization levels of companies can emerge from this study. Attesting on the level of companies' digitization will require a more sophisticated evaluation scheme. Arnold et al.

---

[1] Using the minimum value is arbitrary, one could choose to assign the maximum value to the process and subtract the differences from the process outflow. The risk with this approach is that we may encounter negative flow values.

(2000) showed the value of attesting on a more elaborate scale than the traditional binary reports. Future attestation services will likely have attestation opinions on a continuous scale rather than issuing a binary opinion in the form of a qualified or unqualified audit opinion. This will enable auditors to attest on the level of automation of a particular sub-system, such as order fulfillment, or MRP. Through relying on such attestation services, trading partners, and customers, are more likely to establish business engagements with companies that are highly digitized. This can assure companies faster response time and smaller risk for their operations. Future research can also fine-tune our latency measure, as well as develop a probability network model for assessing information degradation levels. The digitization levels can in turn be studied using Activity Based Costing. This will result in quantify cost associated with un-digitized business processes.

7. REFERENCES

AICPA/CICA: SysTrust Principles and Criteria for Systems Reliability, Version 2.0, http://ftp.aicpa.org/public/download/temp/version2.0finalED.doc

ARNOLD V.; LAMPE, J.; MASSELLI, J.; SUTTON, S. (2000): "An analysis of the market for system reliability assurance services", *Journal of Information Systems*, 14, Supplement: 65-82.

BRYNJOLFSSON, E. (1993): "The productivity paradox of information technology", *Communications of the ACM*, vol. 36, n. 12: 67-77.

KELLER, G.; N¨UTTGENS, M.; SCHEER, A.W. (1992): Semantische Processmodellierung auf der Grundlage Ereignisgesteuerter Processketten (EPK). Ver¨offentlichungen des Instituts f¨ur Wirtschaftsinformatik, Heft 89 (in German), University of Saarland, Saarbr¨ucken.

KENDALL. K. E.; KENDALL, J. E. (1999): *System Analysis and Design,* Prentice Hall.

KUMAR, R.L. (1997): "Understanding the value of information technology enabled responsiveness", *The Electronic Journal of Information Systems Evaluation*, vol. 1, Iss. 1.

LARSEN, P.G.; PLAT, N.; TOETENEL, H. (1994): *A formal semantics of data flow diagram. Formal aspects of computing*, vol. 6, issue 6: 586-606.

O'LEARY, D.E. (2000): "Reengineering assembly, Warehouse and Billing Processes, for Electronic Commerce Using "Merge-in-Transit", *Information Systems Frontiers*, vol 1, num. 4: 379-387.

PAPAZOGLOU M.P.; RIBBERS, P.; TSALGATIDOU, A. (2000): "Integrated value chains and their implications from a business and technology standpoint", *Decision Support Systems*, vol 29, num. 4: 323-342.

PORTER, M. (1985): *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press, New York.

PARKER, M.; ROBERT, B.J. (1988): *Information Economics: Linking Business Performance to Information Technology*, Prentice Hall.

SLYWOTZKY A.J.; MORRISON, D.J. (2000): *How digital is your business?*, Crown Business Publishing.

WARD, P.T. (1986): "The transformation schema: An extension of the data flow diagram to present control and timing", *IEEE Transaction on Software Engineering*, vol. SE-12, n. 2, February: 198-210.
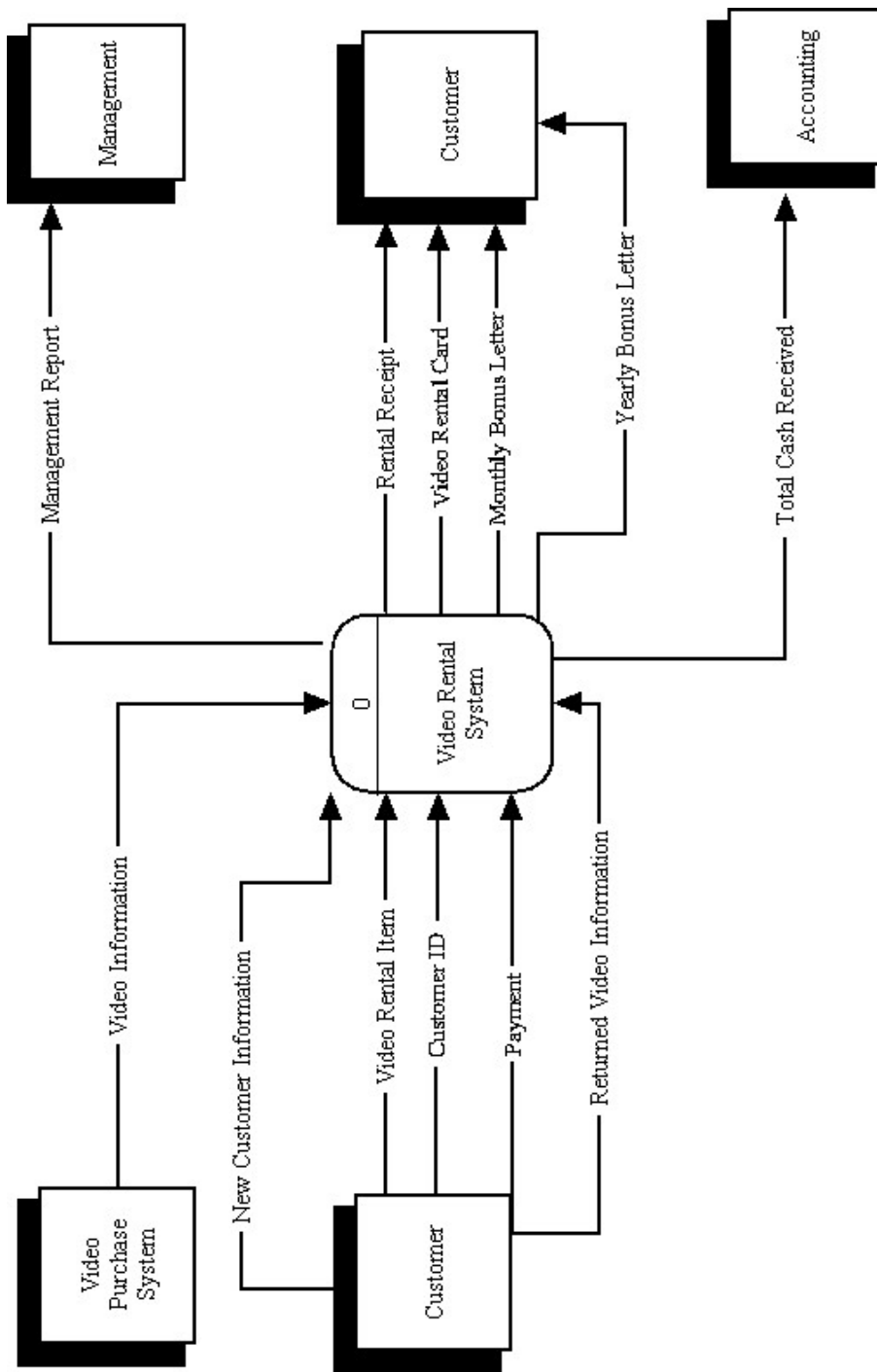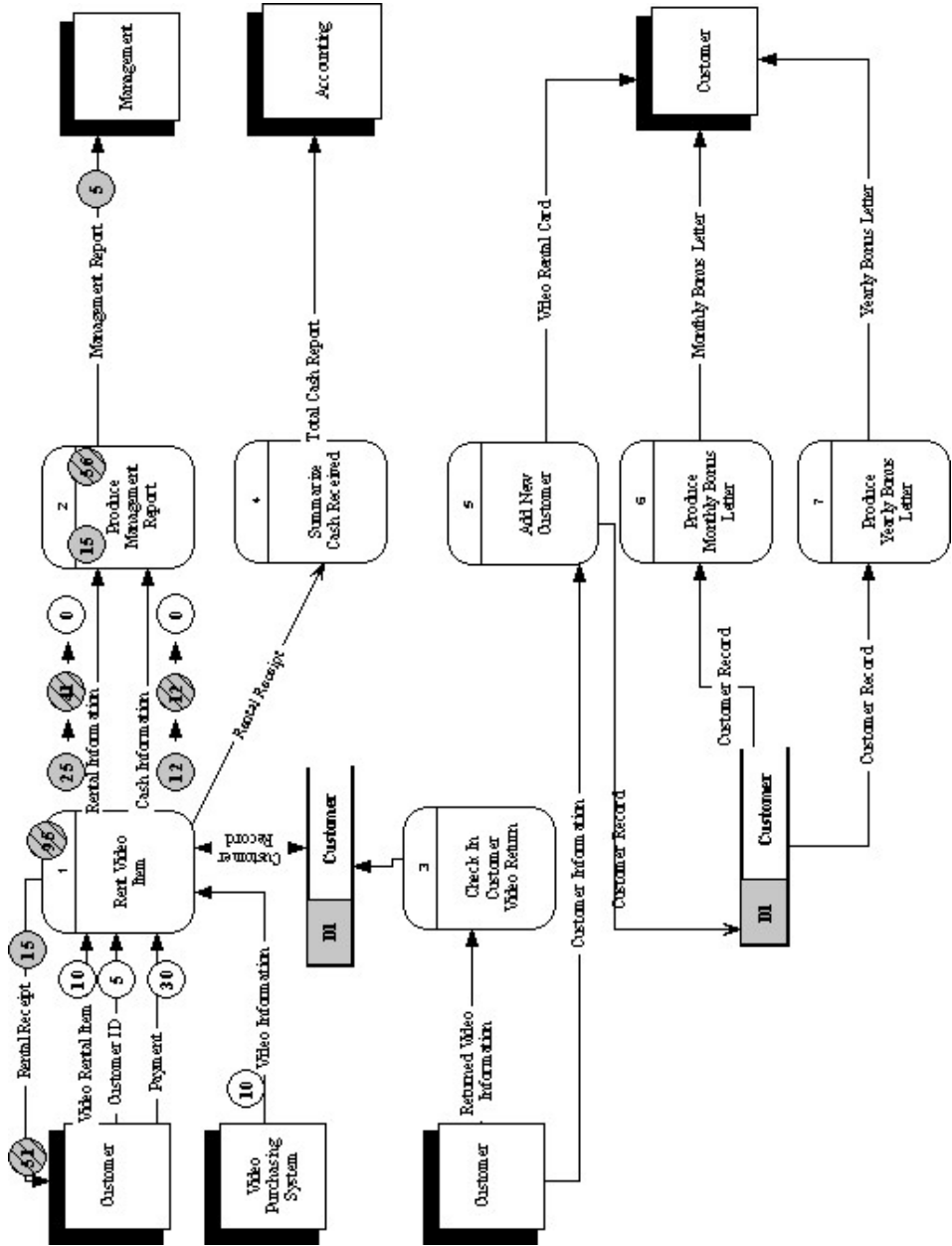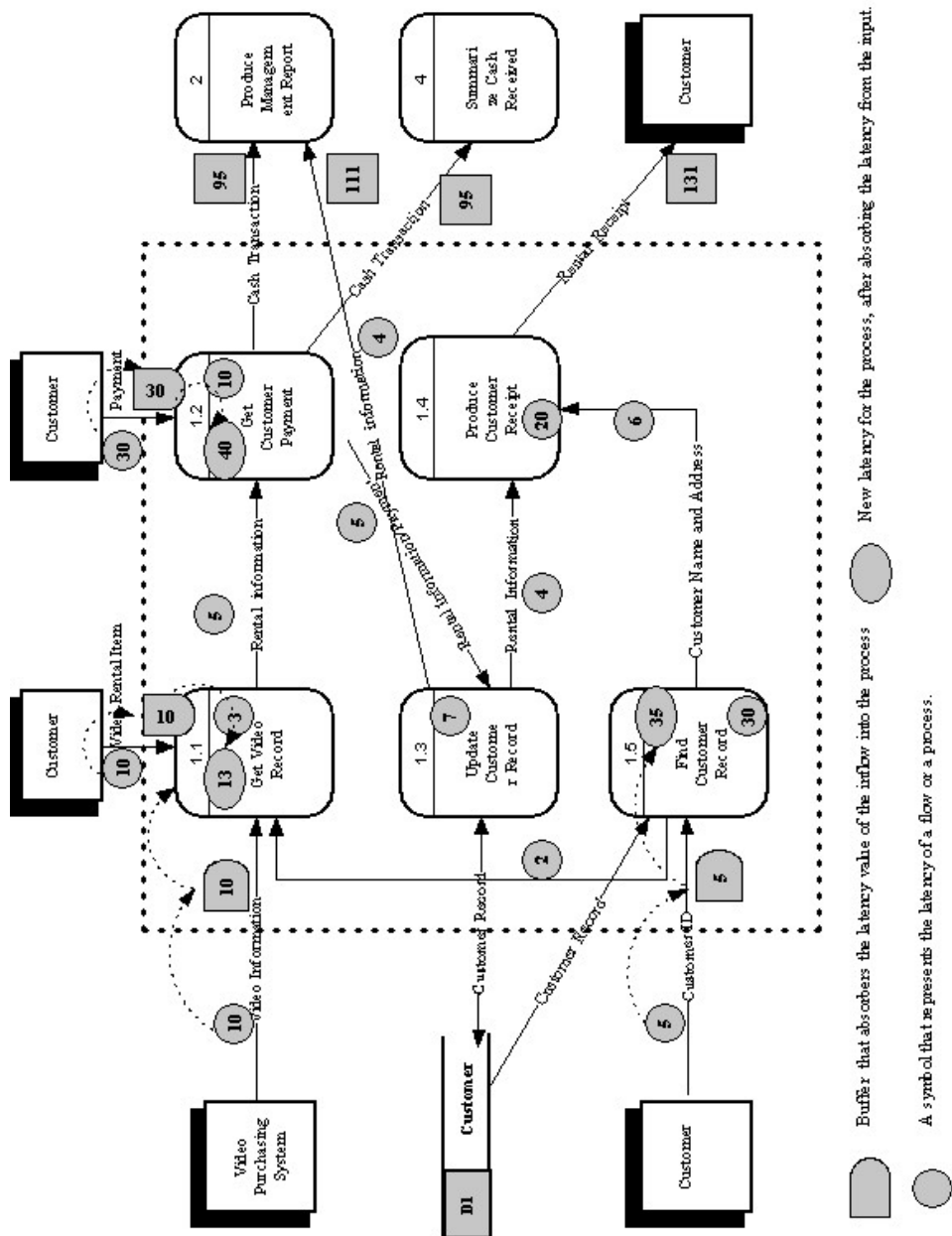
Figure 6. Context Diagram- Video Rental System

Figure 7. Level 0 DFD- Video Rental System

Figure 8. Level 0 DFD- Video Rental System