Product Based Production Workflow with the use of Database Technologies

Jessica Murphey*

Keywords: Architecture, Automation, Book, Database, Workflow

Abstract: Efforts to integrate Computer Integrated Manufacturing (CIM's) technology into the printing industry have experienced varying degrees of success over the past few years. To the extent that it is possible to have full automation, the process and products for which this makes economic sense must be identified and new architecture models developed to reflect this change. There are currently several different protocols, specifications, and standards available that attempt to facilitate automation; however, it is (as yet) not possible to reach consensus throughout the industry on the use of any one specific format or standard interface. Work in this area has focused on adjusting pre-existing file formats and architectures so that changes to established systems could be kept to a minimum. Research currently underway at the Institute for Print and Media Technology analysis the problem using reverse engineering principles; the supposition being that by beginning with the end product, all possible process channels are defined and the functionality of independent devices can be characterized. The goal is to establish whether or not there is indeed a better way for the printing industry to communicate data as the work environment surrounding printing production continues to expand into new areas of business. The research is still in the early stages of development, and the current results are inconclusive; however, indications suggest that formats that have traditionally been outside the field of printing could have a new significance for the industry as the market and the services it demands continue to change.

---

* Institute for Print and Media Technology, Technical University of Chemnitz, Germany

Introduction

Examination of the workflow solutions available on the market today are instructive in their ability to show that there is no *one* way to link independent devices together, nor is there one pervasive data file format which is used throughout the various print production steps. However, potential productivity gains and the money savings from the use of digital technologies will continue to push the printing industry towards standardization. An increasing number of manufacturers and publishers are now seeking to provide system solutions in an environment where automation was previously unknown and are finding the diversity of device formats (i.e., printing presses, finishing machines, packaging, postal software, bar coding, etc.) an obstacle. Without a file format specification for job and device characterization, metadata transfer is achieved through the exchange of various forms of ASCII text data, which supplies the necessary information about a job or a specific device. Whether that data is processed through a parser, a compiler, or a database, the outcome is essentially the same; raw data must be converted in order to be used for broader applications. True device connectivity, in which various finishing machines in a process chain communicate with one another, is not achieved; rather the job requirements are passed along as the job is transported to the next step in the process chain. The end result being that better manipulation of device and job data is needed.

A product-oriented workflow, by definition, is incapable of encompassing or anticipating all of the possible variations within today's print production workflows. What this approach offers is the opportunity to examine workflow architectures from the bottom-up (Has, 1997). Automation can only be achieved by isolating the recurring patterns in a process and defining the variable output channels; in order to begin this examination, some fundamental questions need to be addressed:

Can the identified process channel be automated?
Is there enough financial incentive to automate this process?
If so, what is the best way to automate this process?

In order to answer these questions, a research project has been undertaken to isolate and identify where the problems are within current print production processes and to theorize over possible solutions that could be used to develop a new model for device interaction.

The Product

After careful consideration, hard cover books, printed in editions of two thousand or less, were selected as the product for this trial workflow research. Low volume production of hard cover books presents some interesting problems for workflow automation and standardization. The first problem being that the covers for the books must be assembled off-line and then integrated with the book block construction. The second reason hard cover books were chosen, is because the physical product of a book is the combination of several different technologies, most of which do not have the capability to communicate directly with each other.

Paradoxically, a workflow that is tailored to the needs of single output creates greater flexibility for all participants (i.e. end customer, publisher, printer, and manufacturer). For the end customer, he is able to order only the amount of books that he needs. For the publisher, there is an increase in the range of publication material he can offer; for the printer there is a direct way to profit from low volume print runs; and for the manufacturer of the equipment, there is the capability to link into any given production configuration, thus saving development costs on machine connectivity.

Specifically this project has focused on the issues that the printer must address because most printers, using current configurations can not afford to print small editions. Prepress costs become disproportionate to the print run as well as the costs for finishing equipment, which must be reset for each new job. In order to recover their production costs, a workflow architecture that specifies the product requirements and is capable of maintaining itself would be ideal. This would enable printers to profitably fill customer needs while freeing personnel to work on jobs that require special handling.

The Economic Basis

Market research (Hübler, Reuter 1998) shows that low run hard cover editions are a niche market that is still difficult for most printers to capitalize on. The problem being, that as total production and turn around times have been reduced, the amount of processing time and material handling within finishing departments remains at a constant. Post-press, due to the mechanical nature of the document handling involved, presents physical problems in terms of automation. This makes low volume, finished, hard cover books too expensive for production. Illustration 1 provides a visual comparison of the time required for post-press in both long and short print runs today in contrast to the estimated time required for a profitable post-press time.

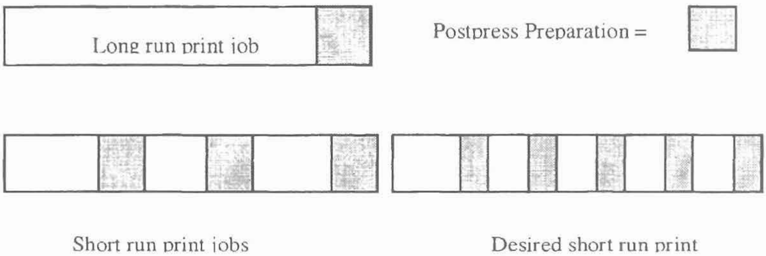


Illustration 1. Example of post-press setup time in relation to print runs.

A German cost study of production costs for single book production (using conventional printing technologies) shows that the cost-per-book with conventional printing technologies for editions under two thousand to be over 19 DM per book.

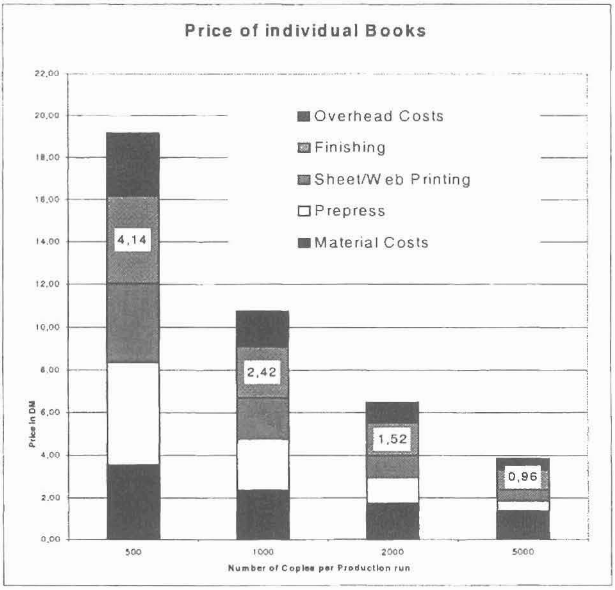


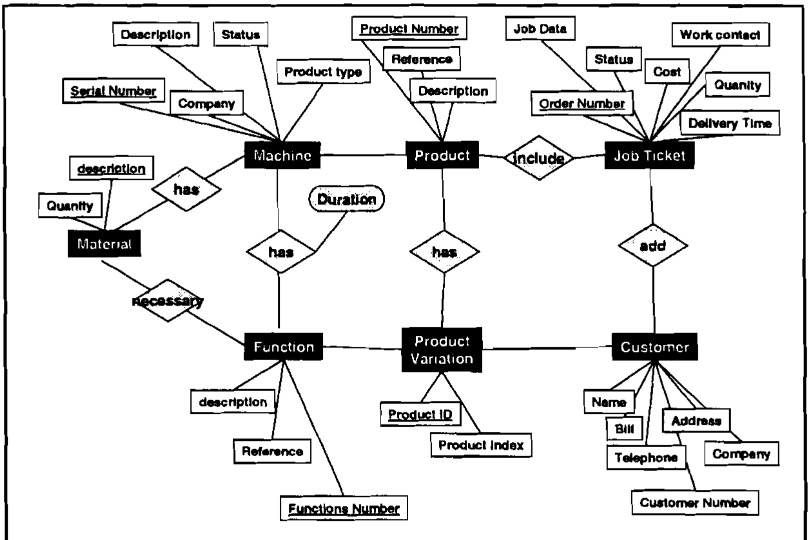
Illustration 2. A cost study of single book production.

In contrast to the decrease in cost as volume increases digital printing remains at a constant of 11 DM, this would be cost savings for the printer of over 8 DM per book. Another consideration is the savings that would be realized by manufacturers who no longer need to pay development costs each time a new third party product is incorporated into their product line. The goal, then, is to realize these economic savings through the automation of post-press departments.

The Database

The architecture for this product-based workflow is centered around database technology. As stated earlier, the diversity of file formats and interface protocols throughout the industry, prevents any one direct solution. Therefore, the focus has been shifted to isolating a universal middle format--in this case an SQL conformant database definition. The use of database technology would not require conformity, just well defined inheritance properties. The idea is to create a model for practical implementation within an acceptable cost range. A hybrid object-oriented, relational database was selected where the functionality of the various devices can be defined and then saved as indexes within the database, this would allow for the concurrent creation of production hierarchies which can then be combined into any number of configurations to address real time requirements.

In this product oriented database (see illustration 3), each component device, or application in the process, feeds information to a central database. Each networked device is registered with it's own IP (Internet Protocol) address in order for the database to receive the transferred SNMP data. Once defined within the databank, it is possible to assign metadata tags, or identifiers, to each component. The stress being focused on the communication to and from all of the various devices (from their native formats) and the applications within the process chain that relies on the machine data via the database. In this way, metadata can assign hierarchies of events to the predefined output channels. Using currently available database technology, it is possible to define a list of



Illistration 3 A flowchart detailing the necessary Object classes within the PODB-Database

output priorities generated from the compiled component information. This list would be dynamic and have the ability to adjust production in response to identified errors within any given channel. A prototype of this kind would allow for functionality that could be combined to handle the scheduling and optimization of an entire system by focusing on known output channels. This functionality should be directly accessible from a web browser interface, thereby allowing the customers to choose and monitor their output preferences.

Functional definitions within the database have been constructed using the mandates of the IPP consortium™ (Bergman, Lewis. 1998). In particular by using the SNMP MIB specification, which deals with printer finishing device sub-units; however, additions have been made in order to create complete product descriptions. Definitions have been created with the desire to adhere, whenever possible, with already existing specifications. Utilizing the SNMP and HTTP standards allows for an easy transfer of device functionality and once uploaded to the Product Oriented Database (PODB) is able to be incorporated into greater applications.

The prototype database for this workflow is Oracle 8.0.5. ™ running off of Linux 6.0™ which was selected in order to meet the high demands of data transformations and querying that a centralized workflow would require; while encouraging the specifications acceptance on an open platform for single book production. Interface methods have yet to be defined. The use of Linux also allows for multi-processing to be incorporated into the workflow and thereby creating a fast, more accurate, feedback channel.

The key to the database is to structure common data definitions and well-defined transformation rules through the mixture and adaptation of finishing code that has already been defined by IPP and ANSI SQL for device definitions, which will isolate where process connectivity is poor and show how feedback could be improved. Production data is being requested regularly from several different departmental databases that lie outside of the physical printing environment, for example:

- Order Entry
- Inventory
- Shipping
- Service
- Production

Each of these departments require access to a Common Metadata Repository (CMR)*, where common semantics, metrics, and rules, are applied to the metadata being sent from the central database to the departmental databases. In order to minimize the amount of data that must be sent, metadata should be subdivided into three tiers of complexity-- core, basic, and advanced:

- Core functionality defines the Components
- Basic functionality defines Service Requirements

Advanced functionality defines the querying capabilities for the Management Requirements

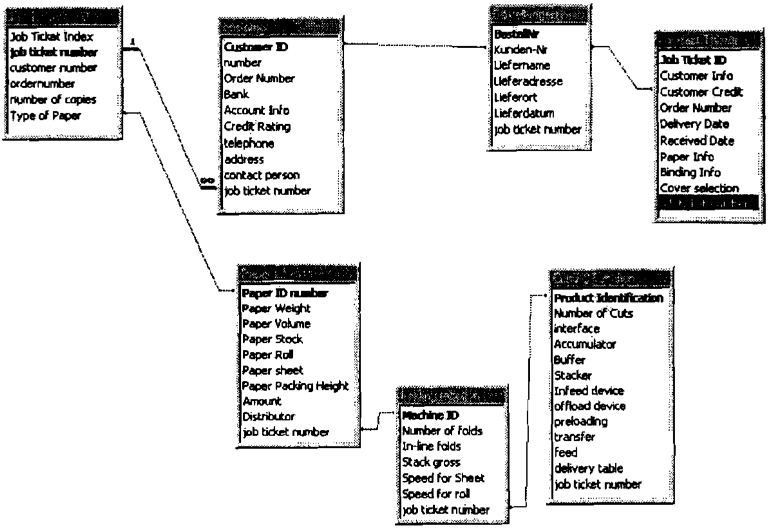


Illustration 4. Tabels within the PODB

In the above illustration (4), provides an example of how the different indexed tables within the database would be structured. In most cases, a C based query method will be used to communicate between the web interface and the PODB database, providing the greater flexibility. However specific instances will require physical access to the database, which will be restricted to the administrator. However, reading and writing to the database using SQL based definitions is more practical than maintaining the SNMP format.

An Example

An example of how all of these elements will come together in practical applications is presented here.

A printer of hard cover books has recently purchased digital printing equipment in response to increased requests from customers for low volume production, namely editions of fewer than five hundred. A WWW site is created to act as the interface between low volume customers and the printer. The customer is able to pick from an established selection of output formats that are displayed from the www site. These fields have been generated directly from the Product Oriented Database (PODB) in response to capacity querying. When the order is placed (a printed book consisting of 200 pages text, printed in black and white with approximately forty low resolution, 600 dpi, images in an A5 format), a series of different processes take place simultaneously:

- Customer Service receives job order and customer identification information;
- Accounting is able to substantiate contact and payment information;
- Production locates and collects the incoming job data and runs a check of:

  Page layout
  Typographic font
  Pictures
  Transfer curves
  Post Script or PDF ™ preflighting
- Inventory checks to assure that the needed materials are available.
- Hard Cover Production is put on stand by.

Metadata is then generated to perform a status check of all devices. Once cleared, the data goes into a job que while cover format specifications are created on the fly so that both halves of production are begun simultaneously. The desired book format triggers an automatic query and queuing for the particular process chain. This registers *busy* or *idle,* which then notifies scheduling as to whether the job can be run on the desired machine within the desired time. Each book block must be bar coded for later identification, so that it can then be attached to its cover. Off line cover integration is handled through the transporting of pallets. Processing information is collected through regular query and status information. This data is then used to refine quality control.

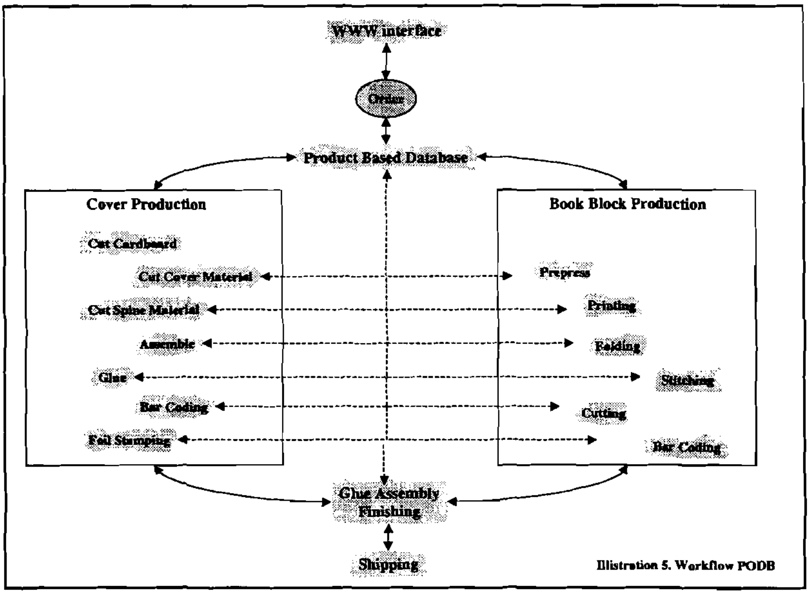


Illustration 5. Workflow PODB

The automated book system provides a maximum output of 195 books an hour with an average production speed of 0.8 meter of 2 up, duplex printed material coming off a web feed system. Paperweight is 80gsm. Normal printing area is 215 by 303 mm (before trim). The finishing equipment in this case, is an on-line glue binding machine with a maximum capacity of 350 books per hour, while calculating a minimum of 34 pages per book, giving a total production number of 175 bound books.

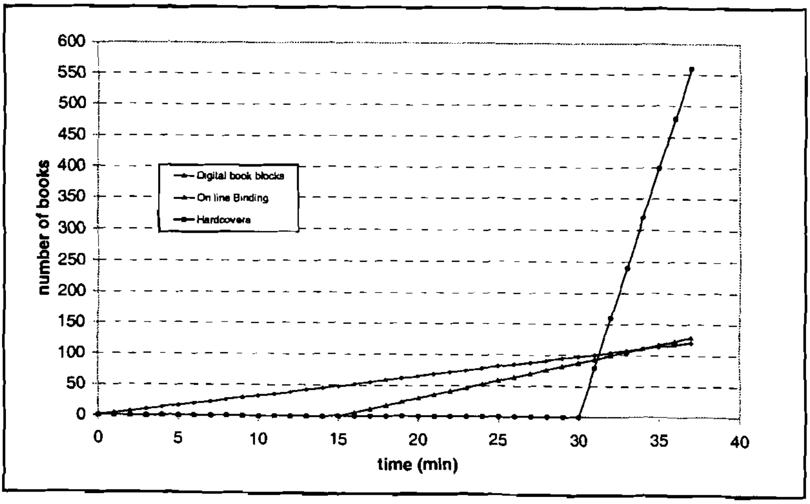


Illustration 6. Chart of Production Speeds for Process Segments

Tolerance between book covers and the book blocks must be within 2 to 3mm as cited by Liebau and Heinze (1997). In low volume production there is the possibility that each book could be of a varying thickness, thereby increasing the margin for error. These are all aspects that must be integrated into the automation process.

Conclusions

The scope of the product based approach to workflow is to develop an architecture that will facilitate the transfer of production and content data in an optimized format throughout any given database schema. The prototype database is currently in the early stages of development where the connectivity of the different devices in production can be tested in a virtual environment. Ideally each device in the process should have the capability to query the other devices in the process. Thereby, each device works with its established definitions through the database, which acts as a buffer between devices and facilitates the exchange of the SNMP data. Having observed the increasing roll of database technology in print production it is the opinion of the author that uniform database definitions will provide the key to complex interlinking of systems. Database functionality provides a realistic way to maintain process and data flexibility.

The next phase of development for this project is construction of the prototype.

Selected Bibliography

Bergman, Ron. Lewis, Harry.
1998 „Work in Progress –The Printer Finishing MIB for the management of printer finishing device subunits" (IPP, internet draft).

Bienz, Tim. Richard Cohn, and James R. Meehan
1996 „Portable Document Format Reference Manual Version 1.2" (Adobe Systems Incorporated, San Jose).

Duan, Stefan. Georg Lucas, Jürgen Schönhut.
1997 "Specification of the CIP3 Print Production Format" (Fraunhofer Institute, Darmstadt).

Has, Michael.
1998 "Automating Page Creation with the use of Contextual Information" (Taga, Michigan). 184 pp.

Hübler, Arved. Reuter, Susann.
1998 „Cost Study for Digital Printing " (Institute for Print and Media Technologies, Chemnitz), 12 pp.

Lewis, Harry.
1998 "The Large Mailroom Operations Specification" (IBM, Boulder).

Liebau, Dieter. Heinze, Ines.
1997 „Industrielle Buchbinderei" (Beruf & Schule, Itzehoe), 6 pp.

Nordqvist, Stig. Fredrick Fälström. Nils Enlund. 18 pp.
1998 "IFRAtrack 2.0 and ist Future Extensions-Glue for Integration in Business-wide Workflow Management Systems". (TAGA, Michigan).
Waters, Larry. Hiromichi Mishina.
1998 "Architecture Model and Control Parameter Coding for Graphic Arts (AMPAC)" (TAGA, Michigan).