Utilizing Edge Computing for Monitoring Plant Productivity in Print Industry

Vladimir Stanisavljević
University North, Croatia

Abstract

Automated monitoring of a whole production plant, equipped with a variety of different machines is a challenging task. Particular industries are introducing their own XML based schemas to ease the integration process. Print industry attempts to accomplish this with Job Description Format (JDF). However, a number of older print industry machines is rarely ready for such an integration. For integrating a real production plant, here is proposed a novel approach in utilizing a concept from Internet of Things (IoT) called *edge computing*, to enhance and integrate various printing and finishing equipment status in a unified manner. Edge computing assumes that a lot of processing is on a remote *node* and that the data is eventually aggregated to another location. For edge nodes small board computers (SBC) with wireless connectivity were used to collect data from machine sensors and store it locally. The data collected on the edge indicates status and operational speed over time of a machine and could be used for various analysis later. Edge node stores all data to a local database that could be accessed remotely or the node could be converted to a JDF compliant producer. The data from edges is then collected to establish a plant wide monitoring system that is a part of management information system. The concept presented here was successfully implemented in a real production environment.

Keywords: Internet of Things, edge computing, print industry, Industry 4.0, data aggregation, multi-source data
JEL classification: C81

Introduction

Extensive measuring of its own processes is a key move in improving productivity and becoming more competitive in any market, especially in a market with tight margins and declining revenues like traditional printing industry. Book and magazine circulations are falling for many years and are replaced by different print products. Surprisingly, some analysis claim that number of prints is changing at slower rate than anticipated a few years ago as the print on product packaging and for promotional purposes is replacing above mentioned traditional print-products. All printers are trying to adapt to this trend but market is facing race to the bottom for traditional products. Remaining printers are competing in a crowded market and are significantly trimming margins. Thus, more detailed knowledge of own processes is required to better predict production costs and timelines and improve productivity and offer competitive pricing. Strategic decisions about new equipment should be also based on a collected data about existing equipment.

Herein, a novel approach to data collection is presented. It utilizes concepts of *Internet of things* (IoT) (Vijayasanthi and Jayachandra, 2018) called *edge computing* for a number of various data collection nodes in a print plant. Existing print plant machinery is equipped with a *small board computers* (SBC) that act as data collection edge nodes. A node is attached to existing or added machine sensors and is used to measure typical activity of the machine-like operation speed or to count production passes. The data collected can origin from sensory input like optical sensors on machines for counting passing paper sheets or other items or can emanate from other networked sources within the machine or from an existing software service. The basic idea in the concept of edge computing (Hajibaba and Gorgin, 2014) is to distribute computing resources further from central computing units and closer to the points where the data is collected. The sensory data could then be processed on the edge node and later aggregated as more complex information to the main information system. The sensory data could also later be used in conjunction with other data sources as shown in (Zeid et al., 2019).

For more than a decade, many industries continue to introduce and standardize modern *eXtensible Markup Language* (XML)-schema based data formats for industry information exchange. An extensive catalog of such a schema could be found on http://www.oasis-open.org. These XML based systems have corresponding workflows with standardized data exchange mechanisms like XML-RPC (*remote procedure call*), SOAP or REST (Karanikolas, 2007), that ease a system integration somehow. In this manner, modern printing and finishing equipment has a number of advanced information integration capabilities, usually in a form of a so called *Job Description Format* (JDF) from CIP4 organization (http://cip4.org) that is used for, but unfortunately real-world print plants are rarely fully equipped with all machines implementing the standard, and even when this is the case the integration is rarely that easy as advertised. So that, the goal was to find a widely applicable and less expensive solution that will provide data needed for monitoring purposes.

Many of the machines in a print factory are in long-term operation and precede the era of the computer based inter-machine communication. Retrofitting an older machine with an OEM upgrade module, just for the basic monitoring, is prohibitory expensive. It can be said, that only the newly built plants can afford complete XML based integration. Return on investment, in buying some advanced print workflow system throughout a whole factory is doubtful. Thus, for many years to be seen, we can expect that many machines will stay in the production as they are. Some additional means for production MIS integration upgrade is thus extremely attractive.

Edge computing assumes that a lot of processing is on a remote *node* and that the data is eventually aggregated to another system on another location. For an edge node a *small board computer* (SBC) with wireless connectivity was used to collect data from machine sensors and store it locally. The data collected on the edge indicates status and operational speed over time of a machine and could be used for various analysis later. Edge node stores all data to a local database that could be accessed remotely or the node could be converted to a JDF compliant producer. The data from edges is then collected to establish a plant wide monitoring system that is a part of *management information system* (MIS).

IoT like approach to data collection

Recently, a number of affordable but quite powerful computing devices has emerged, thanks to mass produced mobile technology. They mostly utilize ARM or Intel Atom processors and are sometimes referred as *small board computers* (SBCs). One well known example of such a computing device is Raspberry Pi (Severance, 2013).

This work describes a relatively novel concept of data integration that utilizes concepts IoT - *edge computing* for a number of various data collection nodes in a print plant. The concept could be applied to a number of other industries as well. In converting a production machine to a "Thing" the following approach was used: All relevant printing and finishing machines are equipped with a small board computer of Raspberry Pi type. Different models were investigated and put on different machines. The speed of all models equipped with about 1 GHz multi-core ARM processor was found to be quite adequate.

Except for older machines, additional sensors were rarely need to be add to a machine for measuring machine activity. Instead, the most of the machines we had examined had a number of sensors on different positions and all that was required was to find and hi-jack/piggy-back an existing sensor with the appropriate data to the SBC. The computer module will be further referred as *data collection (DC) node*. Most of the sensors on any examined machines were of optical pass counter type. They measure exitance or pass of an object trough a beam of LED light. Small opto-coupling circuit had to be used to galvanically isolate a DC node from a machine and to transform standard industrial 24 V to Raspberry Pi 3,3V logical levels. This also assures minimal intrusion to existing sensor operation (just about 10mA of electric current is taken by the circuitry and a digital input on a Raspberry Pi). An example of installation of the sensor into a real printing machine is shown on the Figure 1.
The top shown data collection module on the figure is wireless only *Raspberry Pi Zero W* that is put into a relatively slow lamination machine. The wires from DC node are connected to existing sensor switch board of an appropriate paper pass measuring sensor. The speed of the machine is about thousand units per hour. The lower data collection node on the figure is in a bigger print machine that is capable of printing about ten thousand sheets per hour. Faster wireless only Raspberry Pi 3 A+ was used in this machine.

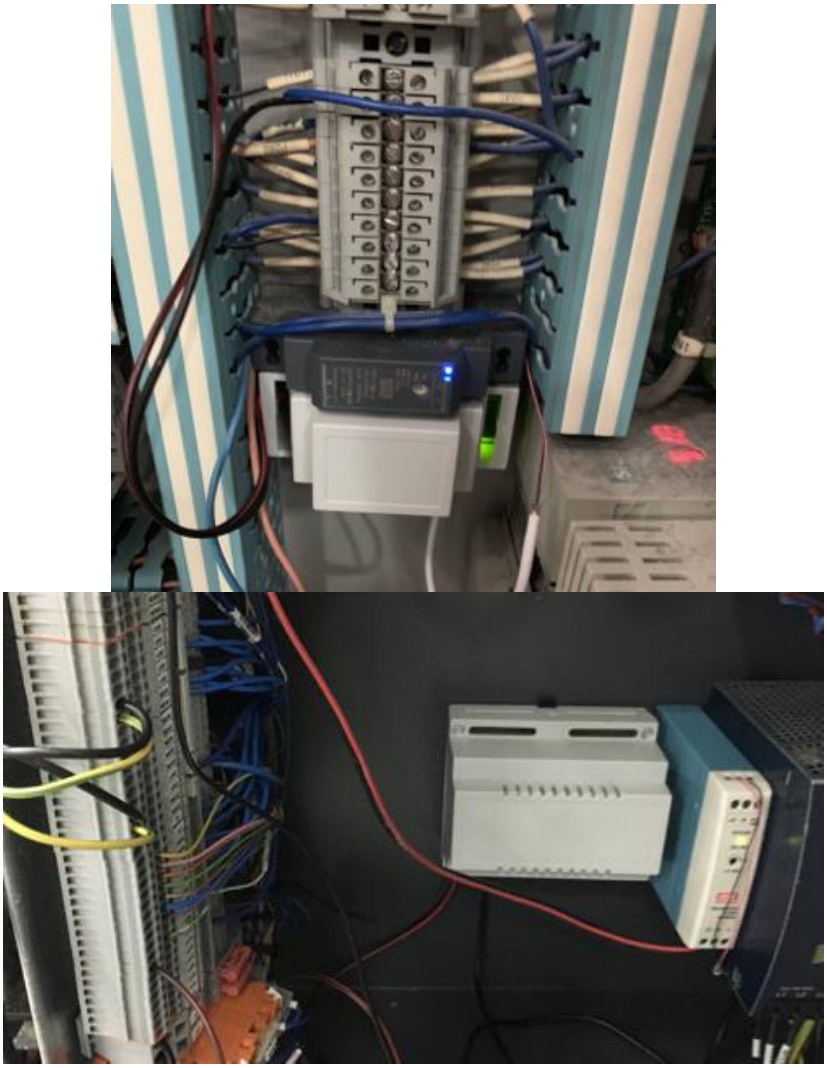
At any larger printer company, that offers a complete service for many different kinds of printed products, the print plant consists of a number of machines for printing (like sheet-fed presses and web presses) and of a much larger number of different machines for finishing the print product. Finishing machines usually provide actions like folding, sewing, binding, laminating, embossing and then packing products like books, magazines or leaflets. At our partner institution we easily identified more than a dozen of different machine types for printing and finishing. The approach explained before was successfully integrated to a number of print and finishing machine types:

- Smaller conventional (sheet-fed) print machine without a networking
- large web print machine

and on the following types of finishing machines:

- folding machine (different generation and complexity)
- sewing machine
- lamination (plasticizing) machine
- foil packaging machine
- binding machines of all major kind (sewed hard cover, perfect/glued soft cover and wire stitched soft cover)

Figure 1
Examples of Data Collection Edge Nodes within a Finishing and a Print Machine



Source: Authors' work

Almost all machines in all production phases before the modification were not networked and there was no wired network among all of them. Some printing machines were networked by wire to a prepress department mostly for color management purpose. As there already was a wireless network for connecting some administration computers, we decided to start experimenting with the integration of all DC nodes over the standard 802.11b/g/n 2,4 GZH wireless network. Only 3 access points were installed in the plant consisting of four main halls of a few hundred square meters of area. Every production machine had at least one (metal) power cabinet where we put our data collection edge computers. In many months of operation, we realized that the wi-fi connection was sufficient and that the nodes were accessible almost all the time. The most of the communication problems was solved by repositioning modules within the cabinet to a point better oriented toward access point antenna.

Developing software for the data acquisition

The software that is run on the DC node is written in C++ programming language and writes data to local PostgreSQL database. As the data collection rate on some machines is quite high it was decided to avoid any possible overhead of an interpreted language like Python. The database is accessed using C language based PostgreSQL *libpq* library (https://www.postgresql.org/docs/current/libpq.html) that comes preinstalled on Raspbian stock PostgreSQL distribution.

The DC node code jumps to an *interrupt service request* routine on interrupts generated on Raspberry Pi General Purpose Input-Output (GPIO) pins (Severance, 2013) on every sample of data acquisition with minimal overhead. Then the speed of production is calculated relative to previous data sample representing production pass. No data averaging was done in that calculation and no noticeable jitter was present even on the fastest machines. As the timing in the data acquisition was quite accurate and was mostly ascribed to uneven load of paper sheets or other parts on conveyer line and calculated speed was just slightly drifting around the data shown on the machine console (usually rounded to 100) the decision was made to sample all data samples and do an averaging later in database by a SQL query. A simple estimate was made that a 32 GB SD card will be sufficient for many years of continuous sampling of the data even on the fastest web printing machine. As this Is an ongoing research project the idea was to keep all the information as long as possible as it could be found relevant for later production analysis. In the real-world production all machines have quite large periods of constant rate operation. Later, the speed and count data could be transformed to a period average and thus be significantly reduced by different coding of data entries.

The data is collected and stored locally on the DC node continuously, even if there is no network connection because of some failure or of some obstacle when a large vehicle or fully loaded fork-lift enters a production hall. This is where the edge computing concept shows its advantage. The data collection is not interrupted and just a communication glitch is reported. In a few months of operation only one DC node, out of a few dozen, required some reposition within cabinet closer to some holes to have more stable connection. All other modules were working continuously without losing connection. But, even in a case of the lost connection data continues to be collected and could be later transmitted to an aggregation point on demand when the connection is re-established. After a few months of continuous operation almost every DC node had collected at least hundreds of thousands if not millions of data points.

The machines (and thus the cabinets where the DC nodes are) turn-off from time to time because of various reasons and a proper shut down procedure on DC is rarely followed, but no data corruption has been noticed in any database. Generally, in everyday operation PostgreSQL had shown as performant and reliable DBMS on all kinds of DC nodes. However, some large averaging operations and report queries that were attempted on a DC node were too slow and were replaced with more efficient solution on a server and in some cases SQL queries were reformulated in a more efficient way. new techniques for data aggregation were utilized to aggregate the collected data to a central system.

The conclusion from many months of continuous operation is that standard Wi-Fi wireless communication is sufficient and its use has eased montage of the modules tremendously. In the case of existing wired network in the machine the Raspberry Pi 3B o 3B+ with ethernet port was used but as the ethernet port on Raspberry Pi is used over USB interface that somehow slows system operation even these modules were connected over Wi-Fi.

Mixed data collection

Any business metrics almost always includes all kinds of data collected from other environmental and machine sensors, but some detailed description from human operators and workers is also required. The data collected from various sources could then be filtered and crosschecked. Preferably, the data is collected from automated data capture directly from machines as described before, but some human explanation of the things happening in the production is inevitable, as the automated data hardly can explain problems and failures in production and this helps building more complete information then just from sensors. The input from operators is especially valuable in a case of failure or other problem that stops the production.

The acquired information could be further used to improve production process and general productivity. The other goal of our project was also to identify data collection points for human operators and enhance them including sensors (existing and new) that can provide advanced data in a typical industrial environment (in our case a print plant) and to complement the sensory data to other data sources and use combined data for further analysis.

The number of different production machines and operating workers and procedures in any plant was case substantial. Beside a number of different core machines for main production, there is a number of other helper machines in any industrial plant. Some of them are just for preparing production by transporting input material, and some are just for inter-processing. Many of them are for preparing production and some for finishing or packing a product.

Human data collection

In many cases an automated tracking of the all processes in a whole print plant is unassailable. Thus, human operators that will enter their observations throughout the production in an ERP are to do some kind of manual or semi-automatic evidence. The manual tracking could then be verified or supplemented by data collected from sensors or software diagnostic within machines. Data collected to an ERP will always be hybrid, coming from many different sources of different level of automation or intelligence. In a typical mid-size print plant, we had easily identified hundreds if not the thousands of nodes for data collection that include existing sensors, neglected manufacturing positions and various working points for humans. Putting a data collection node everywhere is always quite expensive. We combined nodes for human data entry and edge computing nodes as stated before.

Further automation of machine data collection

The first step was in redefining and enhancing a production machine by interfacing some selected sensors to a SBC. This computer will be called *data collection (DC) node*. DC nodes are wirelessly networked together and routed to a data aggregation system. Integrating all the data from distributed databases from many scattered nodes is the main topic addressed in this paper. For this purpose, built in mechanisms for distributed database organization were used. Beside of the data from sensors, local network traffic between a control console and actual press actuators was investigated to read some data but also JDF remote procedure calls or HTTP based protocols implemented on newer presses. We are currently investigating a conversion of our sensor network to JDF producers for consumption by other software systems. We have to note that machines covered in the system are in different departments and measure different data.

Results

Some technical aspects of data integration from a number of distributed database tables is described in (Bagui and Nguyen, 2015). Basically, a local edge node database is aggregated to a central database by means of foreign data wrappers or sharding. The database mechanisms are quite efficient and leave bulk of data on the edge node and aggregate just some data on a demand. All DC nodes have the same database structure which is merged into a virtually single structure from

The data collected was first used for graphical representation of current machine state in a form of speed gauge and a graph in a web application. Other data from the edge nodes is use in other representations of larger data sets like speed over time and detailed analysis in production reports combined with human entered data. All this was included in a *Management Information System* (MIS) trough aforementioned database aggregation. For example, the latest data entry for current operational speed from all nodes is retrieved in about second and gauges are updated on a minute interval.

We are currently examining how to use the data collected to enhance monitoring of the production state and how to pinpoint some production problems to improve production. For example, a machine is not working at nominal capacity as the workers were not reporting repeated failures, but low operation speed was reported and later the real cause was pin-pointed by examining time-line of the production data from DC node. In the time-line manager has noted a number of short interruptions during a work at nominal speed and an effective production output average was not what the machine was reporting. Later discussion with the operating worker has revealed that the suction on the paper was not good because of the combination of environmental conditions and type of coating on paper sheets and thus a lot of paper sheets jammed to the entry to the folding machine.

Conclusion

The data collected from sensors from a machine is continuously stored on a data collection edge node and is on demand integrated to a management information system. Data is visualized by speed-meter gauges that show status of all machines. The idea was implemented by utilizing some principles of *Internet of Things* and edge computing. The edge node consists of a small board computer (in our case a number of contemporary Raspberry Pi-es). Nodes are wirelessly connected to central system on demand.

Built-in database mechanisms like sharding or foreign data wrappers were used wherever possible therefore avoiding using a complex middleware solution and lowering the implementation and integration costs.

Work presented here, is a part of an ongoing research in distributed software development for industrial applications with focus on print industry. The ultimate goal is networking of existing print (and other) industry machinery with affordable small computer modules as edge computing modules called data collection modules and trough them to integrate heterogenous industrial systems into a network of smart machines. For collecting data, a number of data nodes could be attached to almost any kind of a production machine in any industry.

The concept was successfully applied in a real print factory and data collected on various print and finishing machines have been integrated into a management information system and ERP for various analytics and for comparison with manually collected data. Other purposes of the collected data are still to be investigated.

References

1. Bagui, S., Nguyen, L. T. (2015), "Database sharding: to provide fault tolerance and scalability of big data on the cloud", International Journal of Cloud Applications and Computing (IJCAC), Vol. 5, No. 2, pp. 36-52.
2. Hajibaba, M., Gorgin, S. (2014), "A Review on Modern Distributed Computing Paradigms - Cloud Computing, Jungle Computing and Fog Computing", Journal of computing and information technology, Vol. 22, No. 2, pp. 69-84.
3. Karanikolas, N. N. (2007), "Low cost, cross-language and cross-platform Information Retrieval and Documentation tools", Journal of computing and information technology, Vol. 15, No. 1, pp. 71-84.
4. Severance, C. (2013), "Eben Upton: Raspberry Pi", Computer, Vol. 46, No. 10, pp. 14-16.
5. Vijayasanthi, M., Jayachandra, S. (2018), "IOT: An Overview on Internet of Restorative Things", in the Proceedings of the 2018 IADS International Conference on Computing, Communications & Data Engineering (CCODE).
6. Zeid, A., Sundaram, S., Moghaddam, M., Kamarthi, S., Marion, T. (2019), "Interoperability in Smart Manufacturing: Research Challenges", Machines, Vol. 7, No. 2.

About the author

Vladimir Stanisavljević is employed at University North, Croatia, as a Senior lecturer on IT and programming with an industrial experience. The author can be contacted at **vladost@unin.hr.**