Di-Plast Data Infrastructure Wiki - **beta**

29 November 2022

Preface

The following guideline is meant to give a minimum complicated roadmap for all decisions to be made when setting up a data acquisition infrastructure for a production plant in the plastics industry. To achieve this goal several strategies can be followed with different advantages and disadvantages. While there is no perfect solution this wiki will try to illustrate the necessary decision and their consequences from the point of view of a typical plastics processing company based on the experience from the pilots and other real-world examples. Since many of the described topics are interlinked, the given order is just one possibility.

Information Security

One of the first questions that has to be taken into account is the general topic of information security. Since this wiki is focussing on the production environment or shopfloor we will not deal with every possible aspect. What is always important is to make sure that your data can not be stolen and that your network is safe from intruders. To ensure the latter, most companies are separating the production network from the office and/or server network via a firewall to make it harder for potential attacks to reach the critical company data from the production network since it is often easier to break into the production network via a network port in that environment. To ensure the safety of the gathered data several topics have to be taken into account. First of all, the company data is protected by the aforementioned network clustering. For the process data gathered on the shopfloor level it is far more complicated to protect this data. An intruder within the shopfloor network will most likely be able to steal data from the machines or the directly connected systems (i.e., data bases). The value of the data produced in real time is often quite low and since there are no correlations with other data describing material or business processes a loss of the real time produced data is not critical if it only happens for limited time. To protect the data within the data bases those can be moved to the server infrastructure in a different network cluster or onto a cloud storage making the network architecture more complicated. Since information security is a specifically complicated and individual topic it is highly recommended to design the system with an expert. In the following the focus lies on the technical aspects of the data acquisition leaving out the context of information security.

Interfaces

Digital interfaces for data acquisition are key for a working digital infrastructure. During the last years many different interfaces were used to establish communication between machines, sensors and machines or from machines to different software infrastructures. Whereas machines typically give different data points from all sensors and actors within the machine (up to several 100). Within the last years there was a focus on implementing a standardised interface within the production environment. For this application the OPC-UA standard is already widely used and developed to cover a wide range of different machine types and industries.

Details on OPC UA standards can be found here: Home Page - OPC Foundation

To implement and use the OPC UA standards in the plastics industry the European Plastics and Rubber Machinery Association has defined the so called EUROMAP Standards based on OPC UA for several kinds of plastic conversion machines. Many more so-called companion specifications for additional equipment like pelletizers, drying and conveying equipment or dosing units are available or under development.

More information on the EUROMAP standards can be found here: EUROMAP

Besides those quite new and well-defined standards there are plenty of other interfaces in the field. The older and quite unsafe OPC-DA, all kinds of MODBUS types, and approximately 20 different fieldbus systems.
So, one of the most challenging tasks in setting up a good data acquisition is dealing with all different interfaces or converting them to a few ones. The recommendation here is to define a target configuration and then make a roadmap for every single machine to reach that target. These target configurations may vary between different machine types and the importance of the data they deliver. For example it can male sense to

upgrade all injection moulding machines and extruders to OPC UA while all energy measurement equipment stays on MODBUS. The overall target should be to limit the variety of interfaces to a maximum of 5 with exceptions when necessary. To change or adapt the interfaces several technical alternatives are available that will be described now.

1.1 Retrofitting

Within a typical production environment machines vary extremely in age and feature set. Since older machines often don't have suitable interfaces for data acquisition the retrofitting of old machines is often without alternatives. In many production environments machines as old as 40 years can be found but even quite new machines below 5 years of age are often lacking digital interfaces since it was not specified when the machine was purchased. In this context there are two alternatives for retrofitting. The more recent machines often only need an upgrade to the software, or the interfaces have to be unlocked. This is usually a rather fast, simple and affordable process with costs usually below 10.000 EUR and with a minimum of downtime. The used hardware in older machines often is not capable of running with up-to-date software because in contrary to more recent hardware it is not built upon an integrated operation system but individually programmed for the specific hardware. In this case most of the hardware need to be overhauled or completely replaced making this the far more costly alternative but yielding the highest possible upgrade level. Depending on the individual machine typical costs can range from 25.000 EUR up to 50.000 EUR and even above. As a result of the retrofit the machine is equipped with an up-to-date software interface that can be quite easily used to acquire the process data.

Retrofitting is usually carried out be the machine manufacturers themselves. If the manufacturer is not willing or able to retrofit the machine automation specialists that are capable of exchanging the hardware as well as setting up new software are also an option even though the costs can be much higher since those companies are not familiar with the individual machine.

Some companies have specialised on retrofitting plastics processing machines. They can be found here (examples, no complete list):

[Main Page - SHS plus GmbH (shs-plus.de)]

[I4.0 retrofit for recording machine data | SICK]

[Kistler introduces new process monitoring system ComoScout | Kistler]

1.2 Edge Devices

As an alternative to retrofitting the use of so-called edge devices is a quite common strategy. An edge device is basically a computer that is connected to the machine and to the data acquisition. It acquires the data from the machine via the machine specific protocol or directly via individual analogue or digital signals and hands it over to the data acquisition system via a standardised interface like OPC UA or MQTT. Edge devices are usually a cheap alternative to retrofitting with prices starting below 1.000 EUR, but they need to be installed and the connection to the machine via a specific protocol can be challenging. Nevertheless, some companies that are offering process digitalisation rely completely on installing edge devices for each production line. This has the advantage that the machine does not need to be changed in any way and even if the edge device is not working the machine stays n production. The disadvantages are, that there is another device in the data acquisition pipeline that has additional energy consumption, needs to be serviced and I a potential source of errors.

Providers of edge devices that are familiar with the plastics industry can be found here (examples, no complete list):

Measurement Systems for Industry and Energy - iba (iba-ag.com)

Home - ENLYZE

Perinet

1.3 Middleware

Another alternative to set up a data infrastructure on the shopfloor are software middleware products. A middleware can be understood as a translator for data from one interface to the other. To make this possible most of the solutions are rather complicated to implement and require IT experts with programming knowledge to set up and maintain. Therefore, middleware yields a high flexibility to direct data between sources and drains. Middleware can be run on a server on premise or even on a cloud system. Therefore, it usually needs significant invest in a server system that is also redundant and fail-safe to guarantee a high availability. The software costs are depending on the license model as shown below. Good middleware solutions are compatible with many interfaces and thus make edge devices or retrofitting less necessary. In a typical shopfloor environment, a combination of all three is a realistic situation.

As for many other types of software, middleware products are available with different licensing models (examples, no complete list):

Perpetual license and service charge:
Kepware | Software for Industrial Automation and IoT

Software as a service:
Home | Cybus.io

Open Source:
Eclipse BaSyx

Sensoring

In addition to the already available sensors in production machines that are connected via the machine control additional sensors for in-line quality assurance or automation are widely used. To connect those sensors to the data acquisition three alternatives can be used. First, it is often possible to connect the sensors to the machine control to integrate them in the data acquisition. Second, many sensors can be operated stand alone. Those sensors, often called “smart sensors”, have an own interface to be integrated in the data acquisition via a standard protocol or middleware. The third alternative comes into place, when there is a higher level of automation where the process is extended with several sensors and actors to automatically process the product after the initial step, i. e. injection moulding followed by printing the part or mounting several parts together. In those cases, an additional control unit for the automation in used which can be connected to the data acquisition the same way as described under “interfaces” above. Either way the additional information can be integrated in the data acquisition, but one has to take care about the organization of the data base to have an understandable clustering of the data from the different sources aligned to the line setup.

To find suitable sensors for different processes the Di-Plast Sensoring Tool can be used:
Di-Plast-Wiki: Sensor Tool (uos.de)

Databases

After the data acquisition is working the next important step is data storage. It is very arguable how much and how long specific data should be stored but since the costs for data storage are extremely low it is an acceptable approach to store all numeric data in the first place. For imaging data, the decision is not that simple, and an often-used approach is to save the pictures for a limited time but the results that are derived from the pictures (usually numeric data) on the long run. In general, the storage of numeric data or short strings (words) is extremely efficient and needs only very low amount of storage space. Attached to these data different files or attachments have to be stored and linked to this data taking up a lot more space. To set up a useful data storage one should initially define what data types are available and what databases are the best fit for those data. In the following a database architecture and data acquisition scheme for the different interlinked databases can be set up and connected to the data acquisition system.

For the distinct tasks different types of databases can be used (examples, no complete list):

Time series data:
InfluxDB Times Series Data Platform | InfluxData | InfluxData

Relational data:
PostgreSQL: The world's most advanced open source database

Non-relational data:
MongoDB: The Developer Data Platform | MongoDB

Cloud Systems

As already mentioned under "information security" there is always the decision whether to run systems on premise (own servers) or in a cloud infrastructure. While cloud is becoming more and more popular there are still good reasons to host critical data that has to be available with very low latency on an own server system. A typical modern setup is a combination of on premise systems for low latency and good accessibility and cloud storage for high volume and good saleability.

Data Usage

After clarifying the above-mentioned topics and setting up a data acquisition infrastructure as well as a data storage infrastructure the use of the available data is the desired reward. There are many applications that can be tackled with production data:

- Reporting within the company or towards customers
- Predictive Maintenance
- Logistic optimisation
- Process optimisation regarding quality or resources
- Autonomous operation of production lines with AI
- …

For validating and analysing the data other Di-Plast tools are available:

Di-Plast-Wiki: Data Validation (uos.de)

Di-Plast-Wiki: Data Analytics (uos.de)

Other Guidelines

For the above-described topics there are several other guidelines and standards available. A collection of links can be found here (no complete list):

VDI/VDE 3714 Blatt 1 - 2022-09 - Beuth.de (link leads to part 1, parts 2 to 7 also available)

VDI/VDE-MT 3714 - 2022-11 - Beuth.de (necessary qualification to carry out the VDI/VDE 3714)

VDI/VDE 3715 on "Data as an asset - identifying, analysing, measuring and evaluating business data capital" is under development