Automated Supervision of Data Production – Managing the Creation of Statistical Reports on Periodic Data

Anja Schanzenberger and D. R. Lawrence*

Contents

Abstract

Data production systems are generally very large, distributed and complex systems used for creating advanced (mainly statistical) reports. Typically, data is gathered periodically and then subsequently aggregated and separated during numerous production steps. These production steps are arranged in a specific sequence (workflow or production chain), and can be located worldwide. Today, a need for improving and automating methods of supervision for data production systems has been recognized. Supervision in this context entails planning, monitoring and controlling data production. Two significant approaches are introduced here for improving this supervision. The first is a 'closely-coupledd' approach (meaning direct communication between production jobs and supervisory tool, informing the supervisory tool immediately about delays in production) – based upon traditional production planning methods typically used for manufacturing (goods) and adopted for working with data production.

The second is a 'loosely-coupled' approach (meaning no direct communication between supervisory tool and production jobs is used) – having its origins in proven traditional project management. The supervisory tool just enquires continuously the progress of production.

In both cases, dates, costs, resources, and system health information is made available to management, production operators and administrators to support a timely and smooth production of periodic data. Both approaches are theoretically described and compared. The main finding is that, both are useful, but in different cases. The main advantages of the closely coupled approach are the large production optimisation potential and a production overview in form of a job execution plan, whereas the loosely coupled method mainly supports unhindered job execution and offers a sophisticated production overview in form of a milestone schedule. Ideas for further research include investigation of other potential approaches and theoretical and practical comparison.

* Middlesex University, School of Computing Science
Trend Park, London, N11 2NQ, United Kingdom
E-mail: anja.schanzenberger@gmx.de, dave7@mdx.ac.uk

I. Introduction

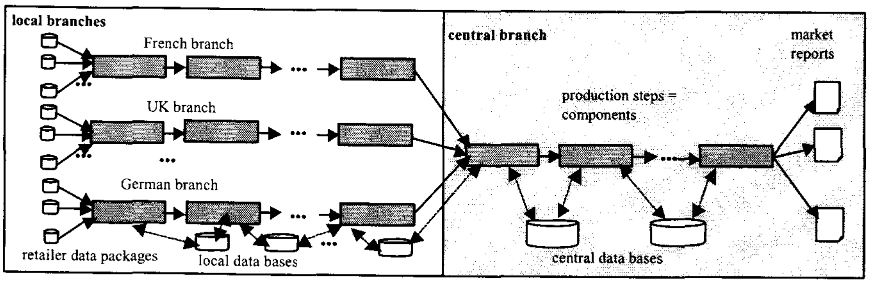
This research focuses on effective improvements to automated supervision of distributed data production systems. Two of these enhancements are introduced and 2 Anja Schanzenberger and D.R. Lawrence compared in this paper. Data production systems are specialised data processing systems. They typically comprise multiple production steps. Their goal is to periodically process and analyse large quantities of data and report pertinent statistics.

They are be deployed in such areas as government, administration, ma rket research, and other businesses with an interest in statistical analysis based on periodic data observation. It is important to note that data production is periodic, i.e. data is observed over more than one period. Automated planning takes advantage of this periodic behaviour. The approach to modelling 'goods' production is a useful reference when considering the supervision of data production systems, but there are also some important differences. For example, no parts-lists can be used. Many aggregations and separations of the produced data-packages are normal. Data packages are mixed together or are divided into other data packages and thus are not constant (changing primary keys). Moreover, there is a need to control the many deviations which are normal at run time in data production. Deviations can arise from the dynamic time scheduling of data (e.g. delayed deliveries), or they can arise out of dynamic changes of input data (statistical principle: different samples can lead to the same statistical result). This report focuses on the supervision and management of data production rather than on the data production techniques themselves. Thus, how the data production process is carried out will only be briefly described. Automated supervision means planning, controlling and monitoring of data production processes to enable all participating contributors (manager, operators and administrators) to control production and provide IT-aided tools for decision support in every production situation.

1.1 A typical Scenario

The GfK Group is a leading market researcher. GfK Marketing Services [9], one of four main divisions of GfK Group, produces reports from periodical observation of =retailers world -wide (e.g. periodic reports concerning competition, demographic evaluation of subsidiaries or product 'hit' -lists). Local branches are available in more than fifty countries. Each country has a branch where the data is collected. GfK MS have established a very large distributed, component-based data production system as in figure 1.

Figure 1: Simplified workflow of a distributed Data Production System



Approximately ten sequenced components (=production steps) are located in each country's branch. This workflow is continued at the central branch and involves 3 Anja Schanzenberger and D.R. Lawrence roughly forty additional components to proceed production. Data sources are the roughly 30,000 different data packages per month delivered from an appropriate sample of retailers (ca. 10,000). Several local and central databases serve as storage for the periodically incoming data. Data is gathered, classified and formatted into a GfK internal uniformed format. After transmission to the central branch (ca. 750 GB p.a.) the extrapolation to statistical reports follows (ca. 1000 different reports per month). This extrapolation is accomplished via data warehouse technology and will be briefly described (briefly, since the report focuses on the supervision and management of data production and not data production itself). Although the process begins local, it ends centralised, so that reports can be provided at the international level. All the centralised data and processes are observed and processed by local staff through the use of web based tools. Although the majority of the local operators are just responsible for local reports, some international departments have been established to produce worldwide reports and use local data to do so.

There are approximately 10,000 jobs per day involved, whereas a job is defined as processing one single data package (or a data definition package) at one component. The duration time of a job can last from few seconds to several hours. An analysis of GfK's processes demonstrates justification for research into appropriate supervisory methods. Today, the majority of supervision there is conducted manually for each production step by permanent polling logs. Production planning is not automated and is prepared manually. Costs, dates and resources are only planned manually and can not be checked. Thus, management has no proven evidence with reference to the optimisation potential of data production. However, for GfK, the conclusion is that business success can not be reached sufficiently without automated supervision.

1.2 Academic literature discussion

If known technologies are applied, it is advisable to prove whether academic or commercial representatives can be used. The following methods have been discussed but have all been rejected for real-world implementation in data production. Nevertheless, this research has shown, all of them are very useful as role models. Data warehouse technology [10] can be used to extrapolate statistical reports (e.g. included in production components), but as the majority of this report focuses on supervision and management of data production, data warehouses will not be considered. In [5] the concepts of Production Planning (PPS) and Shop Floor Planning Systems (SFP) are introduced, but most of the available tools are not notably eligible, because they are exclusively made for goods production. Data production is, as explained before, not similar enough to justify the needed modifications in commercial representatives. Moreover, SFP are often made for small to medium sized organisations due to rapidly growing planning problems with high job volumes. The proces sing of jobs could be conducted by Job Scheduling Systems (e.g.[7]). However, the question is, if such a tool has a chance to evolve its efficiency. Assuming a re-synchronisation of the supervisory tool to its plan after each step to remark delays at once, means losing performance. Thus, in such a case it is not really effective. Workflow Management [1] and Business Process Management Systems [8] have also been discussed. For example, in [3] integrated workflow planning is introduced. However, these systems 4 Anja Schanzenberger and D.R. Lawrence usually lack automated supervision except of controlling the data and control flow [1]. A production overview is not sufficiently supported. Finally, traditional project management [6] includes a project overview, but lacks support for a sophisticated periodically repeated data production.

Ⅱ. Two Proposed Supervisory Approaches

Sophisticated approaches appropriate for automated supervision in data production were to be determined. Although some initial drafts have been considered (e.g. a system architecture based on web services can be found in [2]), this research shows that there are two key concepts worth serious consideration in real-world systems. These key concepts are introduced here (section 2.1 the 'closely coupled approach', section 2.2 the 'loosely coupled approach'). In both these cases, the same distributed data production system is assumed to be the observation base (i.e. the supervised object that has to be managed).

2.1 Closely Coupled Approach - Shop Floor Planning and Scheduling

This approach is an interesting concept for a close coupling of the jobs to a plan. The plan must be steadily compared with the actual process. The plan can be remodelled through reactive planning if variances are recognized. Strict planning of jobs allows to reach a high optimisation potential. For example, less important jobs can be rescheduled to less production critical times. Only planning enables the estimation of a difference between the current states and the planned states within any system. Without a plan, no comparison of objectives is possible except comparisons with former production periods. Additionally, with detailed planning resources can be planned much more accurately.

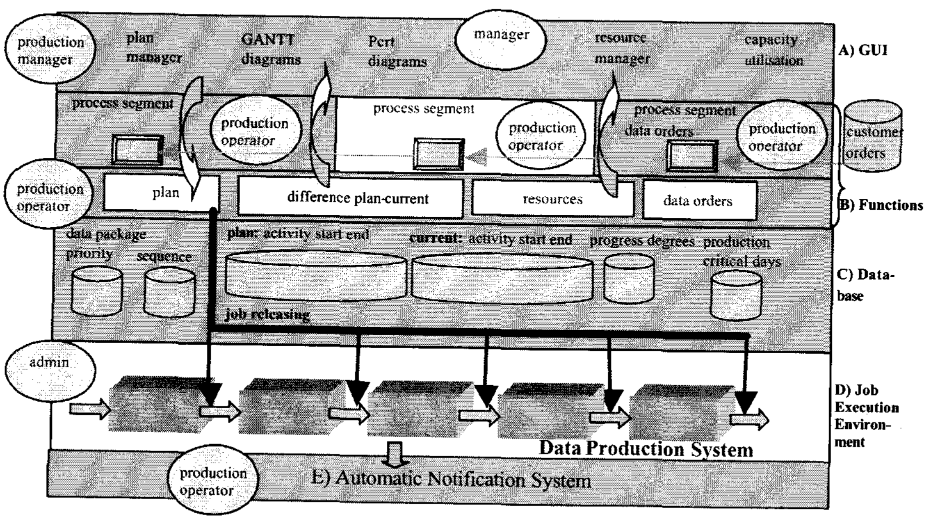
The System Architecture

Research of related work basically illustrates two tactics for supervising goods production: Production Planning Systems (economic based gross-planning) and Shop Floor Planning (detailed job planning) [5]. Both well-known technologies are role models for the following system architecture (see figure 2).

A) Management Information System (MIS) of the Supervisory Tool: This user interface offers detailed views into the production processes in form of a MIS and a planning tool. The MIS includes overviews like GANTT and Pert diagrams (e.g. interrelationships and critical paths). Resource management can be conducted. The detailed planning possibilities of Shop Floor Planning allow for the correlation of humans and machines to current orders. This enables the management to plan and to react directly on load and personnel situations. Thus, capacity utilisation can be displayed correctly. As it is not recommended to create fully automatic plans, (e.g. unforeseen events, necessary manual inspection of the plans, rearrangement of jobs, etc.), the production managers must have a tool to enable reactive planning.

B) Supervisory Tool Functions: The functional level can be divided into segmentoriented and centralized functions. 'Customer orders' are requested reports from 5 Anja Schanzenberger and D.R. Lawrence customers, and must be completely produced by a defined date. 'Data orders' are derived from customer orders, and contain due dates and information about data which has to be produced to fulfil the customer orders. If data orders are backwards propagated to each process segment, then they can be used to inform former process segments of what data has to be available when. Production operators are responsible for controlling the progress of data orders. Centralised functions are partially automated plan creations and calculation of production status, (difference between plan and current states). Differences are displayed as progress degrees. In addition to data orders, customer orders and data packages can have a progress degree. Resource functions are needed for planning human resources and PC loads When using this approach, it is essential that production (as often as possible) will be conducted as planned. Plan-creation must be carefully and accurately arranged to meet optimisation goals. After completing a job, the job execution environment must notify the supervisory tool about its state and duration times. The supervisory tool must then release the next job. This entire process demands coordination and communication between job execution environment and the supervisory tool. If delays or errors are recognised, staff must be notified to change the plan if necessary.

Figure 2: Supervisory tool as Shop Floor Planning combined with Job Scheduling



C) The Supervisory Tool's Database: The control information storage is a centralised database where world-wide access can be provided using web technology. Tables are needed for the plan, current states, the data packages and their priorities. These priorities are essentially important to separate high prior jobs from other jobs, and thus effectively enable demand driven production. The flow sequence (workflow definition) is necessary to determine the sequence of events and for job releasing. For calculations of the difference, two pools for planned and current activities are available. Production critical days (days with high work load) can also be monitored and included in planning decisions. manager

D) Job Execution Environment: Administrators are responsible for controlling the job execution environment, and within it, all production PCs have to be readily available to ensure on-time production. Executing jobs, working with job priorities and additionally a load balancing has to be conducted from this system. If new data package arrive, a job plan template (job plan = sequence of more production steps) is copied to enable production. After every job execution, the release notification of the supervisory tool must be awaited. To ensure on-time deliveries it is always necessary to avoid delays in production. Nevertheless, job realising offers the chance to work exact as planned (plan was created under strict optimisation rules).

E) Automatic Notification System: Early world-wide notification is one of the main goals for a supervisory tool. This means such a system (e.g. based on e-mails) is a good choice for proactive error handling and thus for production optimisation.

Beside all of the positive factors contained in this system architecture, there are some disadvantages. In short, criticisms of the closely coupled approach, is the high level of planning effort, the problem to determine strong planning algorithms, (e.g. heuristics, artificial intelligence or operational research algorithms, etc.) [4], a slow down in production, caused by the need for releasing each job and usually an extraordinarily high effort to control workflows (see figure 5).

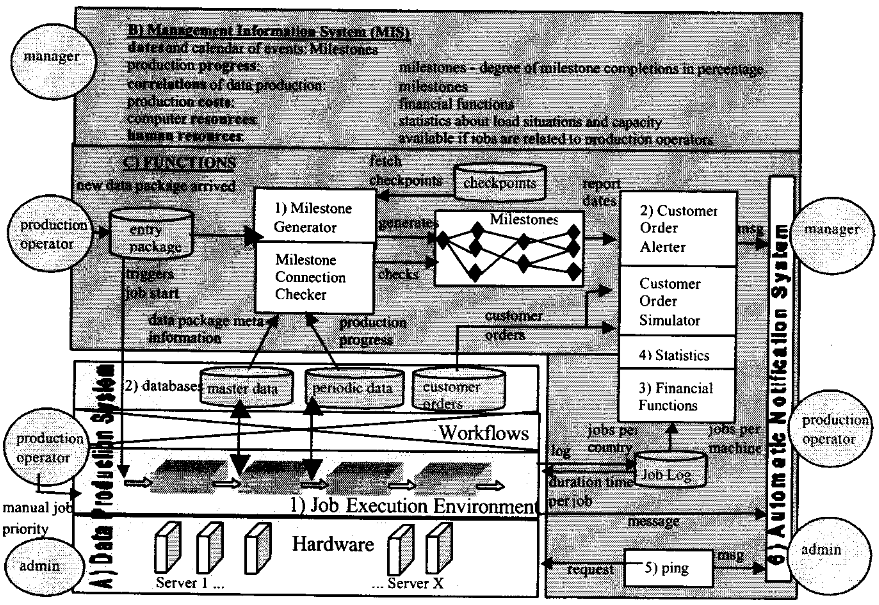
2.2 Loosely coupled Approach – The Milestone Architecture

The loosely coupled approach (see figure 3) is yet another suggestion to automatically supervise data production. One of the most appealing ideas was to discard plans totally. Not following a plan means there is no need for job releases, and also no need to control the whole workflow. Thus, some of the most significant problems (recognized in the closely coupled approach) can be avoided. However, the absence of plans inhibits an ideal optimisation of data production, due to the fact that jobs would usually run whenever they came in. Thus, other supervisory mechanisms must be invented to fulfil the requirement for a high-quality production control. Proven traditional project management delivers the guiding ideas for the loosely coupled concept.

A) The Data Production System: The Data Production System consists of production servers, an environment to execute the production processes (JEE) and databases containing all production data. Noticeable is that no workflow level exists within this loosely coupled method. Forwarding of data and executing the next jobs has to be provided totally within the JEE. The JEE works without any connections to the supervisory tool.

1. Job Execution Environment (J EE): Main focus of the JEE is a well-controlled and reliable job execution. The throughput can only be optimised by working with priorities and load-balancing. However, human intervention is supported through allocation of manual priorities to data packages to enable a faster flow through the production system. For example, the JEE can consist of message queues where jobs are inserted to wait for their execution. Duration times of jobs have to be logged to support clues for future production cycles and evaluation of load situations. Moreover, in case of components which need interactions, the assignment of human resources could be gathered and thus considered. The JEE has also to cope with controlling the health of the production system (e.g. recognising hanging components). Errors in job execution have to be reported immediately to the automatic notification system to inform production operators early.

Figure 3: Milestone Architecture



2. Data Production Databases: The following data types are needed for data production in general.

- *Master data:* Reference data for information about the observed data.
- *Periodic production data:* All observed data delivered periodically.
- *Customer orders:* Due dates of customer orders are the most important ele ments to meet for avoiding penalties and unsatisfied customers in data production.

B) Management Information System (MIS) of the Supervisory Tool: One main goal of automated supervision is to provide the management with substantial information about production success. A commonly accepted measurement is project management. Key points to be controlled with it are: dates, costs and resources. Without exception, supervision of data production systems follow the same rules, but with respect to a periodical business. Aim of the MIS is not to provide information of each single job and the workflows, but to give aggregated production overview. The clever trick to get aggregated overview is just to provide a survey of milestones with progress degrees. Production progress can then be measured and displayed. Gathering a history of milestones allows a comparison of former production periods with current production. This comparison has not the same force as comparing a production plan to actual production as shown in section 2.1, but nevertheless offers an overview of 8 Anja Schanzenberger and D.R. Lawrence historical operational development regarding duration times and throughput. Moreover, a look-ahead of milestones can predict the production situation in the near future. Other important benchmark data are costs, which can be calculated exactly by assessing the job log. Thus all jobs have to be logged, analysed and summarised through appropriate financial functions. The history of jobs is also the basis for statistics about resources, load situations, capacities, bottlenecks or production critical days. Summarised, the MIS offers the management a well-founded base for decisions.

C) Supervisory Tool Functions

1. Milestone Module: A popular and well established tool for project management are milestones. Traditional milestones are activities with no duration time, but with a due date. Milestones for data production have to be enriched with data content information and a progress degree. Data content can be distinctly differentiated through its primary keys. Content-oriented milestones thus deliver status information about production. To simplify the implementation, checkpoints are introduced. They represent the different points of interest in the production process and are templates for the content-oriented milestones. All production data must pass

these checkpoints. The checkpoints as well as the milestones have predecessor-successor relationships. They form a directed acyclic graph and can thus easily be mapped and stored in a database. This enables displaying content-based relationships based on time -oriented milestones states. Current production situations ('what' is produced now, and how is the progress), can be easily controlled. Information about 'what' is produced 'when' can be estimated by creating milestones with look-ahead. The look-ahead can be created if the times between related milestones are gathered. The milestone history can then be used to estimate the milestones due dates in future. Milestones can be created automatically if the event of a new data packages arrives at the systems entrance is recognised. A look-ahead can be created if planned dates of arriving data packages are gathered.

primary keys). At Checkpoint CP0, two milestones are shown. Since Checkpoint 0 has two dimensions (retailer and delivery period), one milestone could be, for example, 'Dixons Jan2004' and the other 'Marc&Spencers Jan2004'. Both would then have one common predecessor at checkpoint level CP1 since both retailers would have provided data for the category 'Color TVs' and for the delivery period 'Jan2004'. If it were now assumed that the category 'Color TVs' must be reported bimonthly, the reporting period, a dimension of checkpoint 2, would then be 'Jan-Feb2004'. Then, after extrapolation in CP3, the end product 'statistical report over Color TVs' would be delivered to a Customer (e.g. 9 Anja Schanzenberger and D.R. Lawrence 'Sony') based on the reporting period 'Jan-Feb2004'. Since all milestones have due dates and progress states, production operators can be informed about content,

Figure 4: Simplified example of content-oriented milestones

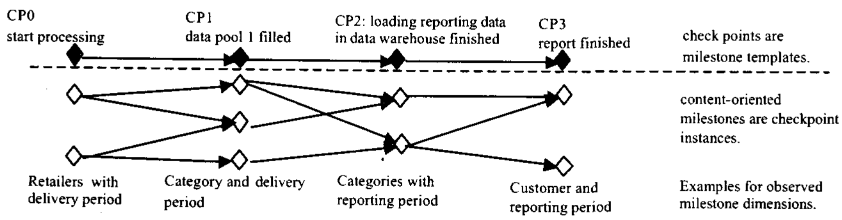


Figure 4 shows the checkpoints and their instantiated milestones. Each milestone consists of observed dimensions (i.e. its delays, and the progress in their data production.

2. Customer Order Tools: Warnings of

current upcoming customer orders given from an alerter could help to identify production critical data. Recognising delayed or critical data early allows timely to increase the priorities of the implicated jobs to fasten the production process. Often it is difficult for management to answer questions about the customer order coverage. A simulator based on information of former production cycles can be set up for this estimation.

3. Financial Functions: For calculating production costs financial functions based on the job log are useful. Costs can be divided per participating segments, departments or countries to fulfil the requirements of a smart accounting. Additionally, taking job priorities and quantity of jobs into account for cost calculation increases the fairness of accounting and the explanatory power of the calculated total.

4. Statistics: A job log is the history of all processed jobs. The benefit of a job log is the possibility to analyse the production process in detail. For example, especially Business Process Management System vendors offer often functionalities to query exactly those process logs [8]. Identified benefits are the possibilities to track operations, to detect inefficiencies and thus to gain business insights. For surveying production in detail, those analyses are essentially important.

5. Hardware Surveillance: Surveillance of system health is essential to keep production processes alive. Hardware surveillance means to control PCs (ping) and network reliability.

6. Automatic Notification System: Staff has to be informed immediately about errors or delays in production. Thus, a dedicated notification system has to support proactive error handling and production optimisation. The world –wide distribution of messages has to be guaranteed (e.g. implementation as e–mail system).

Ⅲ. Early Experiments and Implementation Results

This report deals with the adoption and customisation of data production rather than with the usage and comparison of the role models since these are discussed in detail in the literature PPS [5], WFMS [1], Project Management [6], etc. Until now, there has been no detailed investigation into the consequences and effects of supervising data production. One must cope with problems such as changing primary keys, immense volumes of distributed data and jobs, tracking the manifold aggregations and separations, and the frequent instability of the data packages. All these problems are not fully dealt with if one uses the traditional role models.

Due to the immenseness of data production systems, the enormous effort involved in establishing automated supervision, and the restrictions imposed on the resources available for implementation, the decision was made to concentrate only on one of the architectural approaches rather than prototyping them both. Today, GfK Marketing Services, have started to implement the loosely-coupled approach due to their large data production system with countless deviations. The initial results are discussed as follows:

- Milestones are created automatically when new data packages arrive (i.e. semi automation). As data entrances are distributed, messages associated with data package arrivals are sent to the centrally located milestone database. Thus, the supervision information is centralised in order to give all participants the same state information (e.g. important for gaining international reports).
- GANTT or Pert diagrams were deemed to be insufficient in the case of low data aggregation because the numerous correlations that exist between the data packages prohibit a quick and accurate overview from being produced. High data aggregations do not lend themselves to exactness and explanatory power (e.g. a traffic-light is always yellow, but never green or red).
- An interactive management system for milestones was generated. The interface displays a list of milestones all belonging to a single checkpoint-level. The trick to keep the overview is, only one milestone can be selected at any one time and the predecessors and successors of the selected milestone are shown in accompanying lists. These predecessors and successors can be, in turn, selected, so that their corresponding predecessors and successors can be viewed, and so on. This enables very fast navigation throughout the milestone chain. work without planned due dates. Thus, each Milestone has now a planned due date and a history due date (which is the average of the last X pre -periods). In order to determine planned due dates, a Due Date Editor was developed, whereby due date rules that are specific to particular milestones can be managed.
- Since data production is ever changing (e.g. deviations), generating connections between milestones cannot simply be done once. Thus, milestone connections were divided into 'Planned' (these are based on the connections in the last X pre-periods) and 'latest' connections, which are the actual connections.
- The arrival of new data packages (which generate milestones), milestone

states (e.g. complete or incomplete) and the connections between milestones are checked at regular intervals, in order to keep information up-to-date.

- So as to reduce the number of emails sent via the automatic notification tool, responsibilities were clearly set and how one reacts to an error message was clearly defined.

Ⅳ. Conclusion - Comparing both Approaches

The complexity and possible varieties of data production systems has a likely consequence that there will never be only one single correct solution for automated supervision. In addition, industry is multi-faceted. The study of all role models, the study of automated supervision for data production systems and early experiments made with the discussed approaches shows that theoretical research into an informative categorisation of high-quality approaches is very useful (see figure 5).

Figure 5: Categorisation of automated Supervision approaches

Category	Closely Coupled Approach – Shop Floor Planning Architecture	Loosely Coupled Approach – Milestone Architecture
basis (role-model)	Production planning systems - Supervisory tool is strongly coupled to Job Execution Environment	Project management - Supervisory tool is NOT coupled to Job Execution Environment
level of supervision	lowest level of surveillance (job level); Needs a lot of control data; Delivers exact data for surveillance;	Aggregated level of surveillance (milestones): detailed job information is not used for heavy surveillance; Needs less control data than closely coupled approach; Delivers inexact data for surveillance because of aggregation;
degree of IT-aided supervision		
degree of planning	(high) Creation of production plans; Comparing plan with actual production; Comparing former production periods; Backwards propagated Data orders;	(medium) Creating a look-ahead of milestones; Comparing former production periods Look-ahead of Milestones
degree of monitoring	Planned activities - current activities and history; Progress degrees of activities;	Job log; milestone history (including progress degrees);
degree of controlling	Planning tool (part manual, part automated planning); Planning and replanning of activities; Deadline surveillance: data orders, plan of activities and customer order alerter; Setting breakpoints depends on support in the Job Execution Environment; Resource management Capacity management	Automated look-ahead creation of milestones; Changing job priorities (within the JEE); Deadline surveillance: milestone due dates and customer order alerter; Setting breakpoints depends on support in the Job Execution Environment; Resource monitoring Capacity monitoring
degree of optimisation	(very high) Every single job can be optimal planned; no bottlenecks;	(medium) Jobs are processed with priorities; sometimes bottlenecks;
manual	A lot of work time for reactive planning must be invested;	Adjustment of job priorities: less work time must be invested;
automated	(less) Problem with good optimisation algorithms (can be a NP-hard Problem)	(high) Jobs are processed with priorities
Reliability and stability		
of production	Delays through resource conflicts - reactive planning is needed; Automated releasing jobs can cause delays; High communication effort between supervisory tool and data production system;	No delays in production due to supervision;
of supervision	Exhausting manual planning may cause delays	No delays in production due to supervision;

Effort and expenditure		
to conduct supervision	A planner is permanently occupied with planning; Future and past resource bottlenecks are identifiable through provable plans; Just interventions if problems occur; Permanent checking management overviews;	A production operator can change job priorities if needed exceptional; Only past resource problems are identifiable, due to check functions only assess current situations (no plans); Just interventions if problems occur; Permanent checking management overviews;
of concept implementation	A plan manager software is needed: effort high; workflows have to be considered: effort high;	No planning software is needed: effort low; No workflows: effort low;
GUIs	All GUIs must be aggregated views; A drill-down to job level is possible;	Less aggregation effort due to milestones are aggregated; Drill-down to job level is not possible;
Contingent and type of control		
manual control	Replanning; Reaction on automatic notifications; Setting breakpoints if supported;	Changing job priorities exceptional; Reaction on automatic notifications; Setting breakpoints if supported;
automated control	Plan creation; Comparison plan-current production; Comparison former production periods; Statistics;	Creation of Milestones with look-ahead; Anticipating job aging through changing job priorities; Comparison former production periods; Statistics;
Support for organisational levels		
management	Plan, Gantt-Pert diagrams, resource and capacity manager, customer order alerter and simulator, financial functions, statistics;	Milestones, resource and capacity monitoring, customer order alerter and simulator, financial functions and statistics (logged by JEE);
production operators	Plan, automatic notification	Milestones, automatic notification
administrators	Ping	Ping (logged by JEE)

The summarised conclusions which can be derived from the described research results can be found in (figure 6).

Figure 6: Comparison of automated supervision for data production

Conclusion	Closely Coupled Approach - Shop Floor Planning Architecture	Loosely Coupled Approach - Milestone Architecture
Main advantages	Very high optimisation potential; Difference plan – current production; Future resource problems identifiable; Oriented at proven production planning systems;	Not workflow oriented – easy to implement; freely running job execution environment; Oriented at proven project management
Main disadvantages	workflow orientation to strong: supervisory tool must have knowledge about the production sequence; critic on excessive planning; implementation and execution effort for data production is high (e.g. finding robust planning algorithms and communication between supervisory tool and data production system)	Optimisation potential not as big as with close coupled approach; Future resource problems not identifiable;
When recommended	For small to medium-sized data production systems, where not too much data packages are awaited with less deviations; For data production with strongly restricted resources;	For large-sized data production systems, where much data packages are awaited with many deviations; Data Production were optional data deliveries are allowed. For data production with tolerably restricted resources;

Finally, as discussed in section 3, it is not possible to implement all approaches and ideas in the near future. Nevertheless, future work for this research will continue to be prototyped using the loosely-coupled approach. The experiments made with the prototype are still in their infancy. Currently, the initial users are becoming familiar with the new supervisory tools. Their experiences with the system will ultimately affect the initial concepts. Through gathering information related to their experiences, the outcomes can be refined, evaluated and categorised for this most promising approach. Based on the milestone database it is expected to provide additional and more aggregated production meta information for higher management, to calculate and evaluate the return of investment of data production.

References

1. Frank Leymann, 2000, Dieter Roller, 'Production Workflow – Concepts and Techniques', Prentice Hall, New Jersey.
2. Anja Schanzenberger, D.R.Lawrence, 2003, 'A Web Service Based Approach to Monitor and Control a Distributed Component Execution Environment', EuroMedia-WEBTEC Conference 2003, Plymouth Apr. 2003.
3. Hilmar Schuschel, Mathias Weske, 2003, 'Integrated Workflow Planning and Coordination', DEXA 2003, LNCS 2736, Ed. V.Markik, W.Retschitzegger, O.Stepankova, Springer Verlag, Heidelberg.
4. Peter Brucker, 2001, 'Scheduling Algorithms', Third Edition, Springer Verlag Berlin Heidelberg New York.
5. K. Kurbel, 2003, ,Produktionsplanung und -steuerung, Methodische Grundlagen von PPS-Systemen und Erweiterungen', 5.Ed., Oldenburg Verlag München Wien.
6. Walker Royce, 1998, 'Software Project Management – A unified framework', Addison Wesley Professional.
7. UC4, 2003, ' UC4 global – job scheduling system' , [Online], http://www.uc4.com, [2.Nov.2003].
8. F. Casati, M. Shan, 2002, 'Semantic Analysis of Business Process Executions', LNCS 2287, Ed. C.S. Jensen et al., EDBT 2002, Springer Verlag, Heidelberg.
9. GfK Marketing Services, 2004, [Online], http://www.gfkms.com, [2004, Jun., 15].
10. Lusti, Markus, 2001, Data Warehousing und Data Mining, Sec. Ed., Springer Verlag Berlin, Heidelberg, New York, Springer-Lehrbuch.