

A multi-agent system architecture for coordination of just-in-time production and distribution*

PAUL DAVIDSSON and FREDRIK WERNSTEDT

Department of Software Engineering and Computer Science, Blekinge Institute of Technology, Soft Center, 372 25 Ronneby, Sweden; e-mail: paul.davidsson@bth.se, fredrik.wernstedt@bth.se

Abstract

A multi-agent system architecture for coordination of just-in-time production and distribution is presented. The problem to solve is twofold: first the right amount of resources at the right time should be produced, then these resources should be distributed to the right consumers. In order to solve the first problem, which is hard when the production and/or distribution time is relatively long, each consumer is equipped with an agent that makes predictions of future needs that it sends to a production agent. The second part of the problem is approached by forming clusters of consumers within which it is possible to redistribute resources fast and at a low cost in order to cope with discrepancies between predicted and actual consumption. Redistribution agents are introduced (one for each cluster) to manage the redistribution of resources. The suggested architecture is evaluated in a case study concerning management of district heating systems. Results from a simulation study show that the suggested approach makes it possible to control the trade-off between quality of service and degree of surplus production. We also compare the suggested approach to a reference control scheme (approximately corresponding to the current approach to district heating management), and conclude that it is possible to reduce the amount of resources produced while maintaining the quality of service. Finally, we describe a simulation experiment where the relation between the size of the clusters and the quality of service was studied.

1 Introduction

Agent-based computing has long been suggested as a promising technique for application domains that are distributed, complex and heterogeneous (Wooldridge, 2002; Weiss, 1999). One such application area is the coordination of production and distribution in supply chains. A supply chain is a collection of autonomous or semiautonomous suppliers, factories and distributors, through which raw materials are acquired, refined and delivered to customers. The supply network is often represented as a network similar to that in Figure 1, where resources flow (are transported) “downstream” from *raw material suppliers*, to *tier suppliers*, to *manufacturers*, to *distribution centres* and *retailers*, and finally to *customers*.

This is a somewhat simplified view, e.g. resources may also flow upstream and the network is rarely static with regard to the number of participants. The traditional goal of supply-chain management is cost optimisation with respect to fixed demand. This often involves a trade-off between different objectives, e.g. consider the trade-off between customer service, measured in delivery time, and transportation cost. Also, a typical supply chain faces uncertainty in terms of both supply and demand.

* This work has been financially supported by VINNOVA (the Swedish Agency for Innovation Systems). The authors also wish to acknowledge the valuable input from the ABSINTHE project members employed at Cetetherm AB.

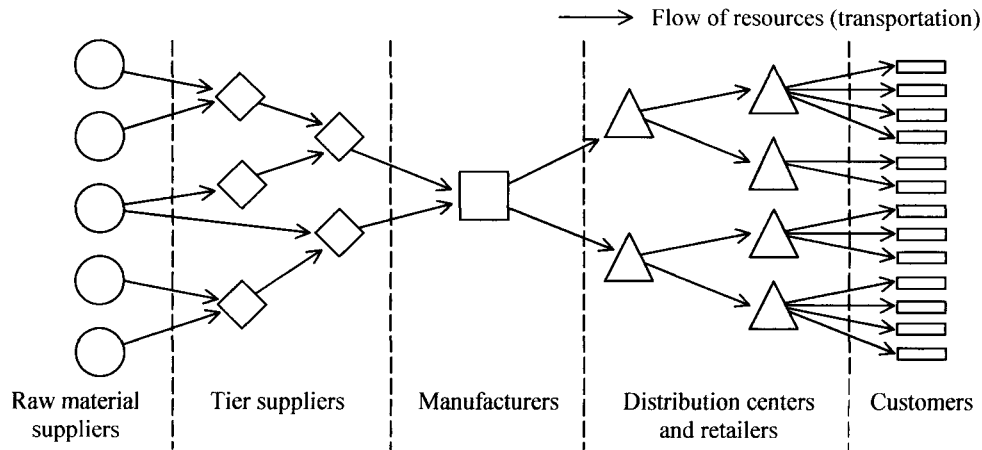


Figure 1 Example of a supply chain network

Thus one of the most common problems faced by managers concerning production decisions is to anticipate the future requirements of customers.

In this paper, we will focus on just-in-time (Ohno, 1988) production and distribution of resources where the production and/or distribution time is relatively long. By relatively long, we mean in relation to changes in customer demand, i.e. systems where it is impossible to immediately and dynamically compensate for such fluctuations by increasing or decreasing production. We suggest a Multi-Agent System (MAS) architecture for solving this problem and investigate its performance.

A characterisation of the general problem of coordination of just-in-time production and distribution is followed by a short survey on approaches to optimisation of supply-chain management. Then a general description of a MAS architecture that solves this problem is given. We briefly describe the principles of district heating systems, which is the domain that has been chosen for a case study, and how the MAS architecture has been adapted to this particular domain. Finally, some results and an analysis of simulation experiments are followed by conclusions and pointers to future work.

2 The problem of just-in-time production and distribution

The general just-in-time production and distribution problem that we study can be characterised as follows:

- there is a set of producers and a set of consumers of resources;
- it is possible to control how much resources are produced at a certain time;
- it is *not* possible to control the needs of the consumers;
- it is *not* known in advance exactly how much resources a particular consumer needs at a certain time;
- the production time, T_p , and/or distribution time, T_D , (between producers and consumers) is relatively long;
- the resources produced must be consumed relatively soon, i.e. they deteriorate fast, are expensive to store or/and there is a limited storage capacity; and
- it is possible to redistribute resources between consumers that are close in proximity relatively cheaply and fast.

In the general problem of just-in-time production and distribution, the distribution time may be different for different producers-consumers pairs. Also, the resources may be of many different kinds.

In sum, the problem to solve has two main components:

1. to produce the right amount of resources at the right time, and
2. to distribute these resources to the right consumers.

The first part of the problem is further complicated by the fact that the resources will be needed by the consumers some time in the future (not at the time of production).

Below, we will focus on a particular application domain that has these characteristics, namely the control of district heating systems. However, there are a number of other application domains that share these characteristics, e.g. iron and steel production, car production and sales, and the production and distribution of oil and petrol as well as some types of provisions, e.g. milk. Obviously, the time for production, distribution and consumption varies between these domains.

3 Approaches to supply-chain optimisation

Optimisation of supply chains has long been a very active area of research (Clark & Scarf, 1958; Arntzen *et al.*, 1995; Cohen & Lee, 1988). The developed models are used to perform strategic (long-term) planning, tactical (short-term) planning, or operational scheduling. A typical supply chain faces uncertainty in terms of both supply and demand of commodities (Lee & Billington, 1995), e.g. caused by late deliveries, poor quality of incoming material or unforeseen demand variability. One of the most common problems faced by managers concerning production decisions is to anticipate the future demands of customers, which typically results in either overproduction or shortage of the commodity. Furthermore, demand uncertainty exists at each level in the supply chain, which can cause the predicted customer demand to be amplified at the production site (at each level some small amount is added to cover for uncertainty at lower levels), known as the bullwhip effect (Lee & Padmanabhan, 1997). The amplification of the demand appears at the first mini-fluctuations (see Figure 2) and, as pointed out by Lee *et al.* (1997), improved information exchange and management may provide an important remedy for these effects.

The management of supply chains is generally considered to lie between systems where the entire flow has a single owner and those with fully vertically integrated facilities, i.e. where one company operates or controls the entire chain. Thus when managing a supply chain the coordination between the participants, e.g. to avoid bottlenecks, is of primary importance. Other general issues regarding the optimisation of supply chains concerns the quality of the information that the decisions are to be based on. The information quality is known to have several dimensions (Mirsky, 1996):

- The *size* of the sample that the information represents. For instance, a forecast can be made for the total consumption of a single product, the consumption of that product by a country or by a customer and so on.
- The *time bucket*, corresponding e.g. to a year, a week, a day or an hour. Breaking the data down into smaller time buckets implies a higher level of detail but also makes forecasting more difficult.
- The *lead-times*. Information is often only valuable within the time frame so that one can react on the situation it represents.
- The *variability* of the information. For example, the production rate is rarely the same across multiple runs. An average rate might be what is used as input to the optimisation model.
- The *frequency* of information updates. The information fed to the model needs to be updated at the actual rate the information is changing in the real world.

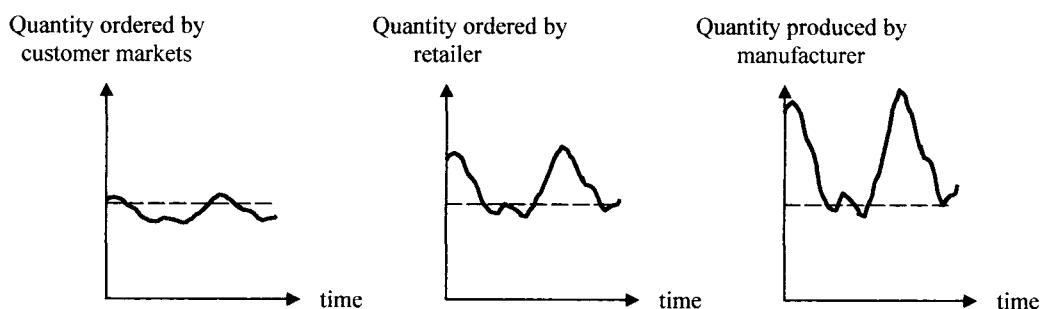


Figure 2 Example of the bullwhip effect increasing the amplitude of the fluctuations at the companies as the information travels upstream in the chain

The traditional approach to supply-chain modelling and analysing has been performed by using control theory models (differential equations), simulations (what-if scenarios) or operations research methods (optimisation theory, game theory and statistical analysis). However, these are static approaches (Kalakota *et al.*, 1995) and are not made for handling the dynamical characteristics of the supply chain (Swaminathan *et al.*, 1998). Moreover, the major part of the work has been focused on limited problem spaces, e.g. optimisation of either production or distribution, and treats these areas as separate systems. Focus now has turned towards the integration of the different areas of the supply chain and the importance of a holistic view of the supply chain has been stressed. For instance, Milgate (2001) argues that competition will evolve from a firm-versus-firm basis to a supply chain-versus-supply chain basis.

Recently, many interesting approaches to supply-chain management have emerged. Current agent-based approaches to supply-chain modelling and analysis typically address the complete chain of entities within the supply network. The general goal for these approaches is either to build an on-line instrumentation and control model (Fox *et al.*, 2000; Parunak *et al.*, 1998; Fan *et al.*, 2002) or to build a simulation model used for analysis or re-engineering purposes (Swaminathan *et al.*, 1998; Strader *et al.*, 1998).

However, these approaches have assumed ideal conditions, e.g. they generally do not consider the problem of forecasting demands and how to deal with discrepancies between predicted and actual demand, which is one of the main issues in real-world supply-chain management.

3.1 *Just-in-time strategies*

Driven by rapid changes in customer demands, manufacturing strategies has moved from large mass production to small batch production (Anwar & Nagi, 1996). Large errors in prediction of customer demand lead to large discrepancies between production and actual demand. This results in higher inventory costs, i.e. larger buffers are needed and/or the worth, or quality, of the products in the buffers decreases over time (deterioration). To deal with this problem, Just-In-Time (JIT) strategies have been developed. However, one of the most common sources of nonlinearity is capacity limits which means that the tendency of JIT strategies to shrink buffer sizes increases the frequency with which a system is likely to exceed a buffer capacity, resulting in a nonlinear behavior of the system (Parunak & VanderBok, 1997). For example, when a retailer attempts to take commodities from an empty storage at the distribution centre or when a retailer tries to push commodities into a full buffer at the consumer markets, operation needs to stop until the capacity has been restored. Furthermore, the distribution network must be streamlined to provide for quick transportation of commodities as well as stable, predictable production schedules and the ability to respond quickly to small fluctuations in demand.

The major concerns in this type of domain are how to cope with the uncertainty in production and demand and how to cope with the temporal constraints imposed by the relatively long production and/or distribution times that are present in the JIT domains that we are considering.

The approaches mentioned above to optimise the supply chain tries to optimise the normal running condition of the system. However, in a JIT domain faulty predictions or erroneous inventory philosophies are likely to cause resource shortages and require means to deal with them. In such an environment other approaches that balance the actually distributed resources, preferably in an automatic way, are necessary.

3.2 *A general multi-agent system architecture for JIT production and distribution*

There are a number of different approaches to solve the just-in-time production and distribution problem outlined above. The most basic approach (and probably the most used) is strictly centralised, where the producer, based on experience, makes predictions of how much resources to produce in order to satisfy the total demand of the consumers. These resources are then distributed directly on demand from the consumers. However, as indicated earlier, this is not possible for domains with relatively long distribution times.

A bit more sophisticated is the approach where each consumer makes predictions of future consumption and informs the producer about these predictions. Since local predictions typically are more informed than global predictions, this approach should give better results. The MAS architecture we suggest below partly builds upon this insight but also introduces a means for automatic redistribution of resources. In order to solve the problem of producing the right amount of resources at the right time, each consumer is equipped with an agent that makes predictions of future needs that it sends to a production agent through a redistribution agent.

The other problem, to distribute the produced resources to the right consumer, is approached by forming clusters of consumers within which it is possible to redistribute resources fast and at a low cost. This usually means that the consumers within a cluster are located close to each other. In this way it is possible to cope with the discrepancies between predicted and actual consumption, which happens for instance when the demand of a consumer changes while the resources are delivered. Without a redistribution mechanism, the consumer would be faced with either a lack or an excessive amount of resources, leaving the system in an undesired state. An important requirement for a redistribution mechanism of this kind is that it can deal with shortages in a way that is fair to the consumers. We have chosen a semi-distributed approach since a completely centralised approach may result in severe communication problems without achieving greater fairness. It would use the same number of messages as the semi-distributed approach, but with a possible communication bottleneck at the central computer. Also, each message would need to travel a longer route, which would increase the total network load. A completely distributed approach, on the other hand, would result in a larger number of messages being sent than the semi-distributed approach without achieving greater fairness.

Based on the above insights, we used the Gaia methodology (Wooldridge *et al.*, 2000) to design the MAS. This led us to an architecture that has the following three types of agent:

- *Consumer agents* (one for each consumer), which continuously (i) make predictions of future consumption by the corresponding consumer and (ii) monitor the actual consumption, and send information about this to their redistribution agent.
- *Redistribution agents* (one for each cluster of consumers), which continuously (i) make predictions for the cluster and send these to the producer agent and (ii) monitor the consumption of resources of the consumers in the cluster. If some consumers use more resources than predicted, it redistributes resources within the cluster. If this is not possible, i.e. the total consumption in the cluster is larger than predicted, it will redistribute the resources available within the cluster according to some criteria, such as fairness or priority, or it may take some other action, depending on the application.
- *Producer agents* (one for each producer; however, we will here only regard systems with one producer), which receive predictions of consumption and monitor the actual consumption of consumers through the information received from the redistribution agents. If necessary, e.g. if the producer cannot produce the amount of resources demanded by the consumers, the producer agent may notify the consumers about this (via the redistribution agents).

The suggested approach makes use of two types of time interval: (i) prediction intervals and (ii) redistribution intervals. A prediction interval is larger than a redistribution interval, i.e. during each prediction interval there are a number of redistribution intervals. Each consumer agent produces one prediction for each prediction interval and sends this to its redistribution agent, who sums the predictions of all consumer agents belonging to the cluster and informs the producer agent about this. The predictions made by the consumer agent must reach the producer at least $T_p + T_D$ before the resources are actually consumed (T_D is individual for each consumer). Typically, there is also a production planning time that also should be taken into account (i.e. included in T_p).

4 Case study: monitoring and control of district heating systems

This case is taken from a current collaboration project with Cetetherm AB, one of the world-leading producers of district-heating substations. The technological objective is to improve the monitoring and control of district heating networks through the use of agent technology. One of the means for

achieving this is to increase the knowledge about the current and future state of the network at the producer side. In current district heating networks, the operators usually have very little knowledge about the actual state of the network and even less knowledge about future states. Instead, the operators control the network by using rules of thumb based on the outside temperature and what time of the day it is. In contrast, our approach makes use of local (and therefore more informed) predictions of future needs. Another objective of this project is to better deal with situations where there is a major shortage of heat in the district heating network. The approach that we will investigate is based on the idea of imposing minor, for the consumers unnoticeable, restrictions on a set of substations in the network.

4.1 Introduction to district heating

The basic idea behind district heating is to use cheap local heat production plants to produce hot water (in some countries steam is used instead of water). This water is then distributed through pipes to the customers where it may be used for heating both tap water and the building. At the customer side, substations are used to exchange heat from the primary flow of the distribution pipes to the secondary flow of the building (see Figure 3). District cooling makes use of the same principles, but uses cold water.

The main problem for the control engineer of a district heating system is to decide the correct amount of heat to produce. As the distribution time from the heat production plant to the customers is large (up to 24 hours in large networks), the control engineer has to estimate the future heat consumption of the customers. This estimation is in most current systems mainly based on the experience of the control engineer together with some environmental and statistical information, e.g. outdoor temperature (Arvastson, 2001). However, an optimal policy for controlling the district heating network cannot only rely on the value of the outside temperature. In order to ensure satisfactory heat supply the tendency has been to produce more heat than necessary and hence an important waste of energy (Arvastson, 2001; Canu *et al.*, 1994).

Furthermore, if consumption is higher than a special threshold an additional heat source needs to be started and connected to the system. Typically it is much more expensive to produce heat using such heat production plants. So another important problem is to decide whether or not this additional source has to be started. This decision needs sometimes to be taken one day before it is actually up and running, depending on starting-up procedures (Canu *et al.*, 1994).

The heat consumption in a district heating network (see Figure 4) is dependent on several physical and social factors, e.g. the weather and hot-water consumption (Werner, 1997). Tap water consumption is very “bursty” even in a large building, and therefore very difficult to predict, whereas the radiator water consumption is “smoother”, mainly depending on the outside temperature and therefore relatively easy to predict (Arvastson, 2001).

We argue that the district heating problem is an instantiation of the abstract problem description of just-in-time production and distribution. It has some particular features that are worth mentioning. For

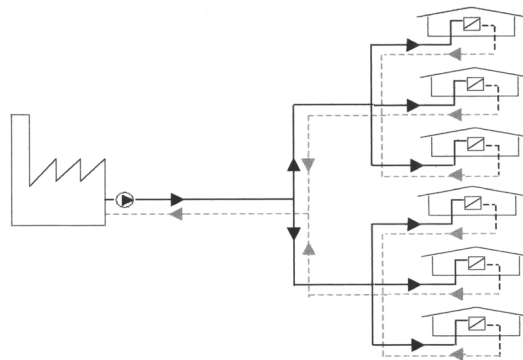


Figure 3 A simple district heating system containing one heat production plant and six substations. The secondary flow (within a building) is separated from the primary flow by a heat exchanger in the substation

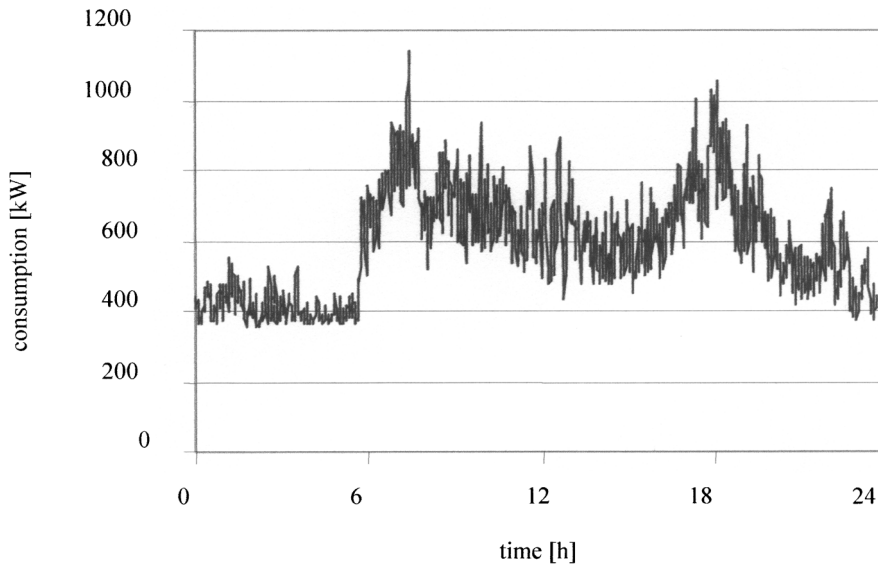


Figure 4 Example of the total consumption of 500 apartments during one day

instance, redistribution within a cluster requires no time and has no cost. Also, there is a buffer of resources, i.e. the produced heat that is not used stays in the distribution pipes for some time.

4.2 A multi-agent system for district heating monitoring and control

The MAS implemented to monitor and control the district heating system has the architecture described above. Each substation has its own consumer agent. It makes a new prediction of the heat to be consumed during a future prediction interval by the corresponding substation every 10 minutes. The prediction is made at least $T_D + T_P$ in advance and is based on the average consumption during the corresponding time period of the last 5 days. In addition, it continuously monitors the heat consumption and sends each minute a report to the redistribution agent of the cluster it belongs to.

Each cluster of substations has a redistribution agent whose task is to ensure that the substations in the cluster do not use more heat in total than predicted. It monitors the consumption of the substations belonging to the cluster (by summing the consumption reports it receives each minute). If the total consumption is higher than predicted (i.e. it is not possible to “redistribute” the heat between the substations in the cluster) and the buffer is not large enough to compensate for this, the redistribution agent will impose restrictions on the individual consumer agents in the following way: no substation is allowed to use any radiator water during the next one-minute interval. If this is not enough to compensate for the additional consumption, the redistribution agent will also impose restrictions on tap-water consumption.

This redistribution strategy (and all other strategies based on radiator-water restrictions) will lead to a radiator water deficit that needs to be compensated (otherwise the temperature in the apartments will fall). In order to do this, each substation will individually compensate by using more radiator water than predicted when the tap-water consumption is less than predicted. Compensation for tap-water deficits work in a similar way, but on the cluster level.

Another problem that has to be solved by the redistribution agent is how to cope with the “bursty” consumption of tap water without commanding unnecessary restrictions. The approach used here is to let the cluster use more tap water than predicted (measured in minute averages) in the beginning of the 10-minute period and then gradually lower the allowed average consumption towards the predicted average consumption.

The producer agent receives predictions of consumption and monitors the actual consumption through information passed on by the redistribution agents. The producer agent may impose

restrictions of consumption on the substations (via the redistribution agents) if necessary, i.e. if the producer is not able to produce all the desired (predicted) heat.

The interaction patterns between the agents (and the substations and production units) in the system is summarised in the message sequence chart in Figure 5. The interaction cycle starts with all consumer agents creating and sending a prediction message to their designated redistribution agent at $t_0 - (T_p + T_d^i)$, i.e. the time required to produce and distribute the resources to substation i . When the redistribution agent has received predictions from all members of the cluster that it manages, it creates and sends a cluster prediction message to the producer agent. The producer agent collects all cluster prediction messages and informs the production source about the requested amount and at what time the resources are to be delivered to the clusters. The production source responds with the effect that it will produce (which may be lower than the total demand) and, depending on the amount, the redistribution agents computes potential future restrictions that it informs the consumer agent about, which in turn informs the substation.

At time t_0 the consumption interval starts with all consumer agents being informed by the substations of the current consumption. The consumer agent informs the redistribution agent about the consumption. Upon completion of the collection of consumption messages the redistribution agent sends a cluster consumption message to the producer agent and calculates if there is need for additional effect, depending on the demand in relation to the previously produced effect. If additional need exists the redistribution agent sends an additional request message to the producer agent. The producer agent informs the production agent of the total consumption and potentially about additional need. If there is a need for additional effect the production source responds with the effect available (depending on the current state at the pump there exists a small buffer in raising the pressure, an effect which almost immediately will be available at the cluster). When the redistribution agent receives the message of available effect it calculates (if necessary) the restrictions and informs the consumer agents about these. The agent communication language used is FIPA (Foundation for Intelligent Physical Agents)-compliant (Bellifemine *et al.*, 2001).

4.3 Simulation experiments

The MAS described above has been evaluated in a series of simulation experiments. In this chapter, we describe the simulation software, the experiment design and the results of the experiments. We have assumed that the distribution time from the producer to the consumers is 1 hour and that there is a single production source.

4.3.1 The simulation software

The MAS as well as the simulation environment was implemented in JADE (Java Agent DEvelopment framework) (Bellifemine *et al.*, 2001). Thus we used an agent-based approach that was time-driven, i.e. where the simulated time is advanced in constant time steps, to simulate the environment. Each simulated entity was implemented as a separate agent. Figure 6 shows the different parts of the simulation software.

A simulator for hot-water usage, implemented by Arvastson and Wollerstrand (1997) and based on detailed field measurements performed in Sweden by Holmberg (1981), was used to generate consumption values. The predictions of consumption for each substation were calculated as the average over five generated consumption sequences (of course, different from those used for the simulated consumption). This may be thought of as corresponding to averaging the consumption for that particular 10-minute interval for that particular substation of the last 5 days. This results in a discrepancy between “predicted” and “actual” consumption, which the redistribution agent needs to handle; for an example of this see Figure 7.

4.3.2 Experiment design

Two series of simulation experiments have been made. In each experiment there were one producer agent and one redistribution agent (since a cluster is independent of other clusters, we do not need to

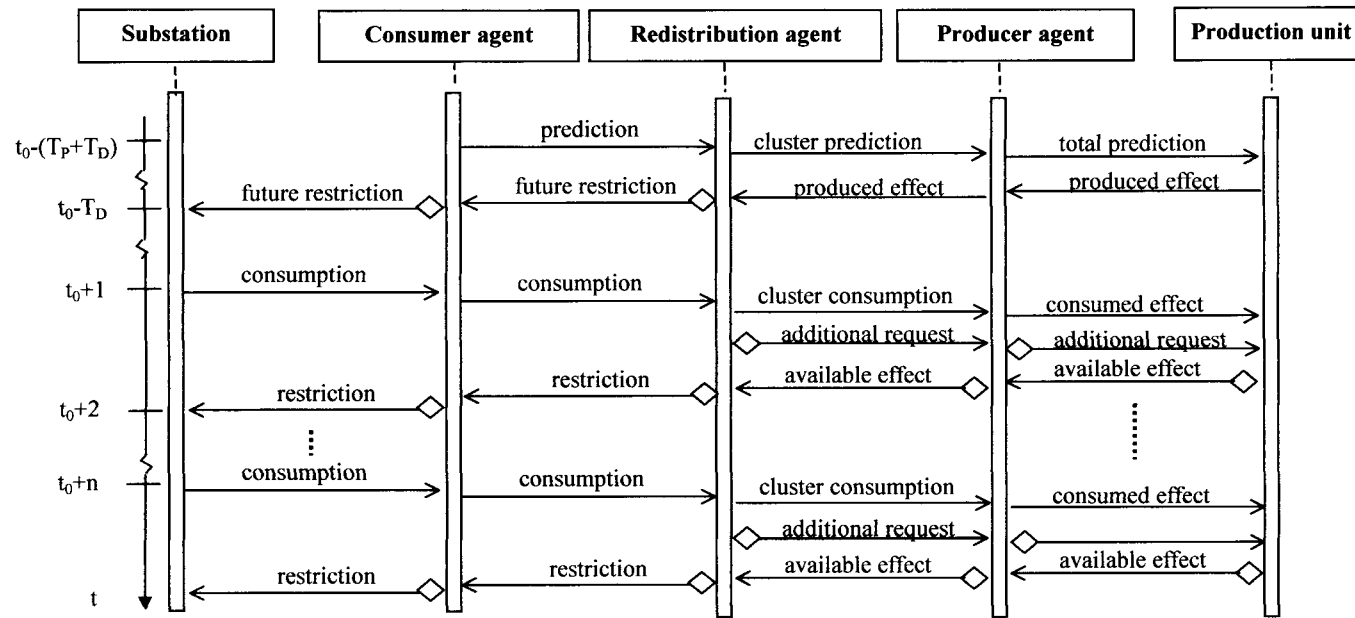


Figure 5 A message sequence chart for the interaction between the actors of the system (pseudo-AUML): T_p = production time, T_d = distribution time, and t_0 = the start time of the actual consumption interval

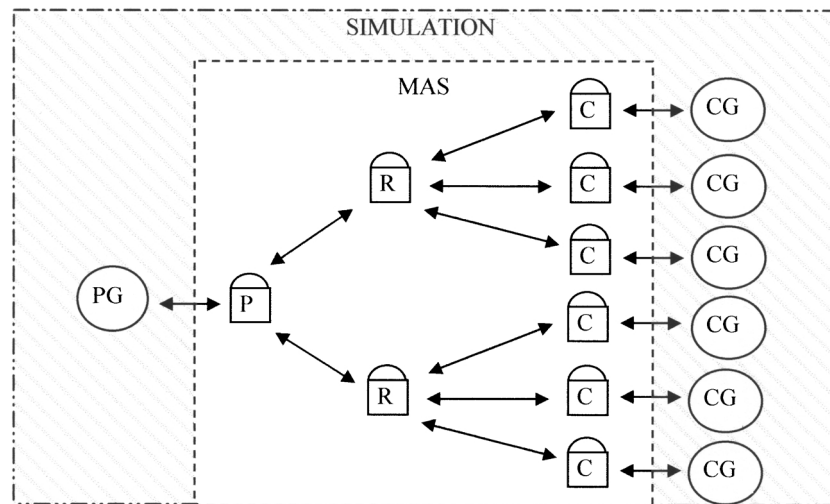


Figure 6 The MAS and Simulation software parts: R=redistribution agent, C=consumer agent, P=producer agent, CG=consumption generator and PG=production generator

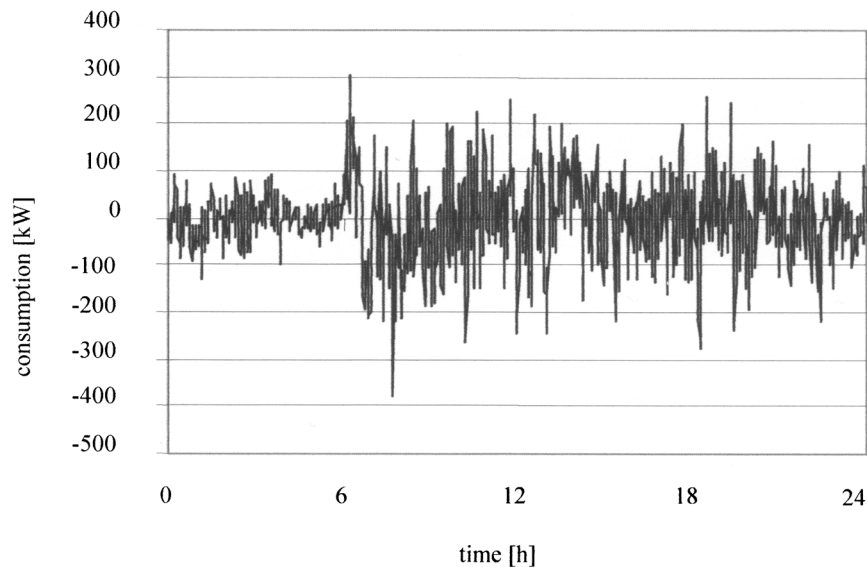


Figure 7 Example of the difference between two consumption sequences over 24 hours

simulate more than one cluster). In each experiment the district heating system was simulated for 24 hours. Each experiment was run over 5 different days (different series of consumption values) and the averages are shown in the diagrams below.

In the first series of experiments, the redistribution agent was managing a cluster consisting of 10 consumer agents (substations), where 5 of the consumer agents were serving a building with 40 apartments and 5 consumer agents serving a building with 60 apartments. We ran experiments on different degrees of surplus production (from 0% surplus production, in steps of 1%, to 5%), where surplus production is defined with respect to the predicted consumption. For example, if the predicted total consumption is 200 kW and the surplus production is 2%, 204 kW is produced. The quality of service was measured in terms of the number of restrictions issued by the redistribution agent.

We also compared the suggested approach to a reference control scheme, which we believe is a very optimistic approximation of the current production strategy for district heating. Also, here, different degrees of surplus production were tested. The production of the reference control scheme is shown in Figure 8.

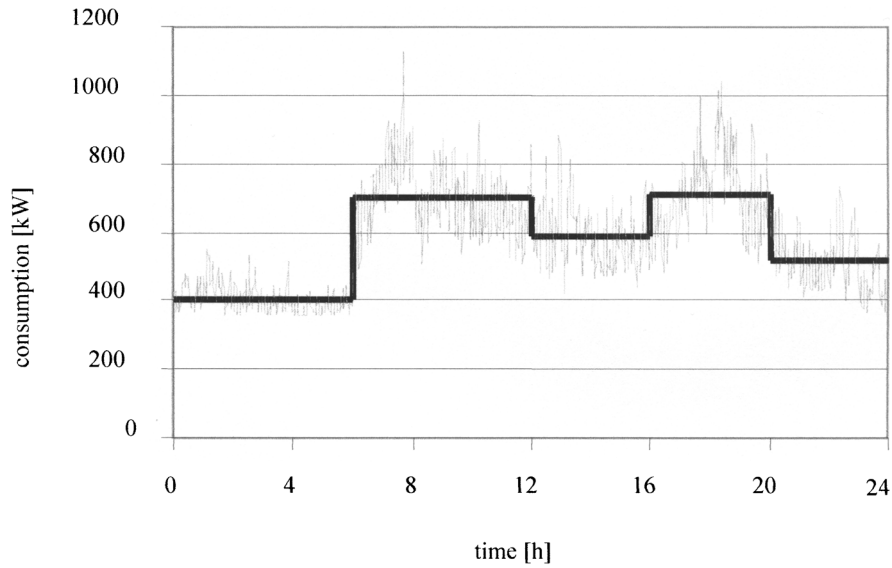


Figure 8 The amount of hot water produced by the reference control scheme (indicated by the thick line, 0% surplus production), compared to the 5-day average consumption

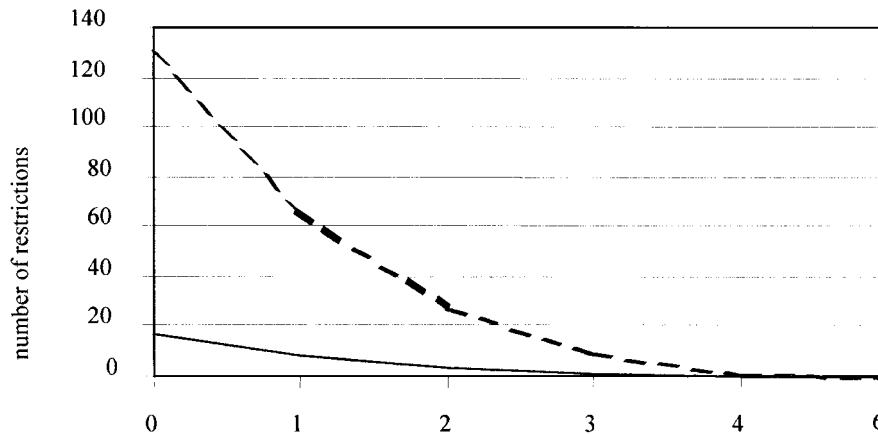


Figure 9 Trade-off between quality of service and surplus production. The dashed line corresponds to the number of restrictions for the radiator water and the other line to the number of restrictions for the tap water

In the second series of experiments, we varied the size of the clusters in order to study how the cluster size affects the quality of service. The number of consumer agents in the cluster that we studied were 2, 4, 8 and 16, where each consumer agent served a building with 40 apartments. In this series of experiments there was no surplus production.

4.3.3 Experimental results

We found that the multi-agent system described above performs well, coping with faulty predictions (even though the discrepancy between the predicted and the actual consumption is sometimes quite large).

Figure 9 shows the total number of restrictions to tap water and the number of restrictions to water for household heating (radiator) during one day for different degrees of surplus production. We see that there is a clear trade-off between the quality of service (number of restrictions) and the amount of surplus production and that there are almost no restrictions of any kind when 4% more hot water than the predicted consumption is produced.

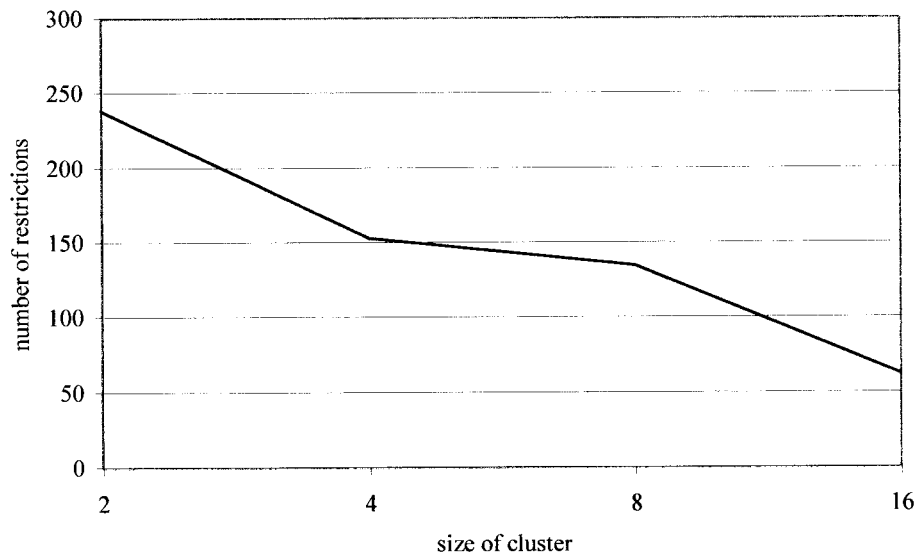


Figure 10 The quality of service as a function of cluster size

Using the reference control scheme, 7% surplus production was needed to achieve this level of quality of service. It should be noted that it is in current practice very difficult to find this minimum degree of surplus prediction. This is because the operators of district heating networks typically do not have exact information on either the quality of service delivered to the customer or the amount of the actual degree of surplus production.

In Figure 10 we see that there is a reduction in the number of restrictions when the size of the cluster increases. In this experiment there was no surplus production. However, it should be noted that the upper limit of the cluster size is determined by the structure of the district heating network. If the actual distance between the substations within a cluster is too large, the assumption of free and fast redistribution does not hold.

5 Conclusions and future work

Simulation results show that the suggested approach to coordination of just-in-time production and distribution makes it possible to control the trade-off between quality of service (measured in terms of the number of restrictions) and degree of surplus production of resources. Compared to a reference control scheme, it is possible to reduce the amount of resources produced, from approximately 7% to 4% surplus production, while at the same time maintaining the quality of service. The simulation results also indicate that the larger the cluster size, the higher the quality of service that can be achieved. However, the number of restrictions decreases only logarithmically as the cluster size increases. Also, there are often other factors, such as distance between consumers, that limits the number of consumers that can belong to the same cluster

Future work includes

- trying out other restriction policies than fairness, based for example on priority;
- improving the prediction mechanism, for instance by using neural networks (Canu *et al.*, 1994; Malmström *et al.*, 1996);
- performing field studies to determine the current heat production philosophies in district heating systems, in order to define a more realistic reference control scheme;
- evaluating the generality of the results by applying the suggested MAS architecture in other just-in-time domains;
- improving the simulation environment by simulating the water flow; and
- testing the approach in actual field tests.

References

- Anwar, MF and Nagi, R, 1998, "Integrated scheduling of material handling and manufacturing activities for just-in-time production of complex assemblies" *International Journal of Production Research* **36**(3) 653–681.
- Arntzen, BC, Brown, GG, Harrison, TP and Trafton, LL, 1995, "Global supply chain management at Digital Equipment Corporation" *Interfaces* **25**(1) 69–93.
- Arvastson, L, 2001, "Stochastic modeling and operational optimization in district heating systems" doctoral thesis, Lund Institute of Technology, Sweden.
- Arvastson, L and Wollerstrand, J, 1997, "On sizing of domestic hot water heaters of instantaneous type" *Proceedings of the 6th International Symposium on Automation of District Heating Systems* 17–33.
- Bellifemine, F, Poggi, A and Rimassa, G, 2001, "Developing multi-agent systems with a FIPA-compliant agent framework" *Software: Practice and Experience* **31**(2) 103–128.
- Canu, S, Duran, M and Ding, X, 1994, "District heating forecast using artificial neural networks" *International Journal of Engineering* **2**(4) 767–773.
- Clark, AJ, and Scarf, H, 1958, "Optimal policies for a multiechelon inventory problem" *Management Science* **6** 475–490.
- Cohen, MA and Lee, HL, 1988, "Strategic analysis of integrated production-distribution systems: models and methods" *Journal of Operation Research* **36**(2) 216–228.
- Fan, M, Stalleart, J and Whinston, AB, 2002, Decentralised mechanism design for supply chain organisations using an auction market *Information Systems Research*, forthcoming.
- Fox, MS, Barbuceanu, M and Teigen, R, 2000, "Agent-oriented supply-chain management" *The International Journal of Flexible Manufacturing Systems* **12** 165–188.
- Hinkkanen, A, Kalakota, R, Saengcharoenrat, P, Stallaert, J and Whinston, AB, 1997, "Distributed decision support systems for real time supply chain management using agent technologies" *Readings in Electronic Commerce*, Addison Wesley, 275–291.
- Holmberg, S, 1981, *Norrköpingsprojektet – en detaljerad pilotundersökning av hushållens vattenförbrukning M81:5*, Departement of Heating and Ventilation Technology, Royal Institute of Technology, Stockholm, Sweden.
- Lee, HL and Billington, C, 1995, "The evolution of supply chain management models and practice at Hewlett-Packard" *Interfaces* **25**(5) 43–63.
- Lee, HL, Padmanabhan, V and Whang, S, 1997, "Information distortion in a supply chain: the bullwhip effect" *Management Science* **43**(4) 546–558.
- Malmström, B, Ernfors, P, Nilsson, D and Vallgren, H, 1996, "Short-term forecasts of district heating load and outdoor temperature by use of on-line-connected computers" *Värmeforsk* 1996:589 Technical Report.
- Milgate, M, 2001, "Supply chain complexity and delivery performance: an international exploratory study" *Supply Chain Management* **6**(3) 106–118.
- Mirsky, MJ, 1996, "Direction of optimization technologies" *APICS International Conference* [<http://www.supply-chain-systems.com>].
- Ohno, T, 1988, *Toyota Production System* Productivity P.
- Parunak, HVD, Savit, R and Riolo, RL, 1998, "Agent-based modeling vs. equation-based modeling: a case study and users' guide" in *Multi-Agent Systems and Agent-Based Simulation* LNAI Vol. 1534 Springer Verlag.
- Parunak, HVD and VanderBok, RS, 1997, "Managing emergent behaviour in distributed control systems" presented at *ISA-Tech*. [<http://www.isa.org>].
- Strader, TJ, Lin, FR and Shaw, M, 1998, "Simulation of order fulfillment in divergent assembly supply chains" *Journal of Artificial Societies and Social Simulation* **1**(2).
- Swaminathan, J., Smith, S and Sadeh, N, 1998, "Modeling supply chain dynamics: a multiagent approach" *Decision Sciences Journal* **29**(3) 607–632.
- Weiss, G, 1999, *Multi-Agent Systems* MIT Press.
- Werner, S, 1997, "Dynamiska värmelaster från fiktiva värmebehov. Fjärrvärmeföreningen" *FOU* 1997:10 Technical Report ISSN 1402-5191.
- Wooldridge, M, 2002, *An Introduction to Multiagent Systems* John Wiley & Sons.
- Wooldridge, M, Jennings, NJ and Kinny, D, 2000, "The Gaia methodology for agent-oriented analysis and design" *Journal of Autonomous Agents and Multi-Agent Systems* **3**(3) 285–312.