Integrating Business Processes with Peer-to-Peer technology

Florian Kupsch[1], Dirk Werth[2]

[1] German Research Center for Artificial Intelligence (DFKI), Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany
kupsch@iwi.uni-sb.de

[2] German Research Center for Artificial Intelligence (DFKI), Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany
werth@iwi.uni-sb.de

Abstract. Since file sharing tools as eDonkey or Grokster are one of the favourite applications for a large number of internet users, the peer-to-peer paradigm experiences a new boom, as it enables very robust, scalable and fault-tolerant architectures. Indeed, the functionality of current applications is limited to quite simple operations such as searching and downloading music files or movies. Hence, to benefit economically from the characteristics of Peer-to-Peer technology, there is need of additional, innovative applications for business usage. A predestined scope of application is the context of Business Integration, where various distributed IT-systems act jointly by exchanging data records and the corresponding control flow. In this article, we present the concept of a Peer-to-Peer based integration architecture that solves the structural problems of traditional integration approaches. It allows a holistic integration of data, applications and business processes without the common insufficiencies of existing EAI solutions.

1 Introduction

Emerging the internet was an unpredictable development. All the same, important operating criteria such as reliability, local structures and robustness always came to the fore. As a result, users had the opportunity to get access to a highly available international data network that stays operative even in case of a breakdown of some single nodes.
One important aspect of globalization is concentration. As a logical consequence, international companies have to merge or to collaborate with each other to meet the requirements for a global distribution of their goods and services. In a networked economy, these enterprises are bound to unseal some parts of their IT-infrastructure to allow the engaged parties an exchange of product and accounting data as well as current status information in order to sustain the supply chain.

Furthermore, employees need improved business applications with enhanced functionalities to manage their all-day work. These applications replenish existing legacy systems by degrees. Long-ranging, these enhancements result in a heterogeneous network of computers and applications, as all those components require each other and have to be consolidated.

In the context of integrating distributed business application systems, the vision of redundancy and robustness was not resumed consequently: A changing business environment results in new technical developments and increasing demand of IT supported execution of business processes.

A first naive solution to provide an interaction of different systems is shown in figure 1. Apparently, an appropriate description for this point-to-point integration of single applications and platforms is described with the term *Spaghetti Integration.* This (n:n)-connection of resources is not able to be maintained properly, as it contains too many non-standardized interfaces. In the last resort, the number of interfaces that have to be implemented will arise quadratic.

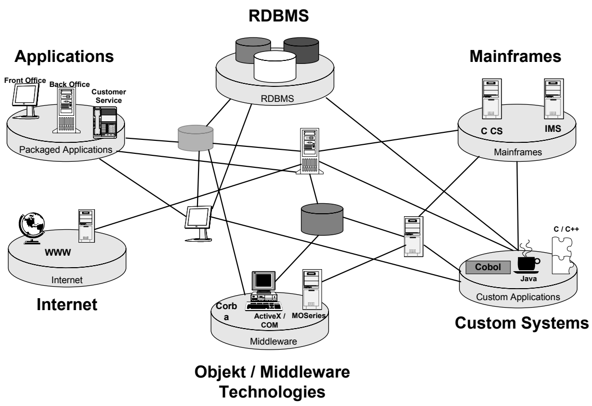

Fig. 1. Heterogeneous IT-Infrastructure

After having widely experienced those scenarios, most enterprises recognized that there has to be found another solution for a reliable integration of business applications.

Some years later, the efforts of bringing together single systems to one logical unit are paraphrased with the term *Enterprise Application Integration (EAI)* [1]. EAI means a bold venture, as the interaction of business applications is restricted due to several constraints:

- *different communication protocols* and interfaces,
- *syntactical differences* between data of the single applications
- *proprietary semantic* of particular system messages, and finally
- Business modelling methodology differs from the technical implementation. There is an essential difference between an integration at technical level and business process integration

However, the market for specific integration software is estimated to grow up to 30 billion US$ in 2004 [2].

In chapter 2, we mention why the existing state-of-the-art solutions for business integration have tremendous disadvantages. We will also propose an advanced Peer-to-Peer (P2P) integration architecture.

Certain applications based on P2P technology are already in use for non-commercial purposes. A prominent example is the exchange of music, movies and software via file-sharing networks where single peers communicate directly without detouring via any central servers. In the specific field of business integration, most development activities focus on 'old-school' solutions that rely on client/server architectures. All the same, some crucial advantages of P2P technology (performance, resilience, load-balancing, etc.) are well-known to many software architects. Chapter 3 gives an overview about existing P2P applications.

A widely used standard for exchanging data between enterprises is Electronic Data Interchange (EDI) and its sectoral implementations such as EDIFACT. Hence, these standards determine a relationship where only two parties are involved. As soon as an inter-organizational process stretches across more than two companies, conventional EDI operations will fail. With the appearance of *Web-Services*, a new paradigm was born. Now, it is technically possible to encapsulate functionality within services that are made available via internet. The interaction of these services is syntactically supported by the XML standard. Unfortunately, common Web-Services are stateless and therefore do not include any functionality that would be capable to control complex business processes. Furthermore, one needs a central Directory for the Universal Description, Discovery and Integration of Web Services (UDDI) where a catalogue of all offered services is stored. However, there are ambitions to enable Event-Driven architectures using Web Services [3].

At this time, P2P technology comes into play: Why should it not be possible to accomplish a P2P-based integration architecture that is not dependent on any central server and that allows a standardized exchange of both, business process data and control information?

The advantages would be evident: The abandonment of EAI-Servers that are hard to maintain is as fascinating as the easy customization of the single peers and the dynamic aggregation of new peers during run-time. Chapter 4 will introduce our idea of a P2P integration environment.

2 Business Integration and EAI

As already mentioned in the first chapter, an integration of the core IT-systems is essential to guarantee a frictionless handling of business processes. In an optimal scenario, the result would be a seamless IT infrastructure that is completely transparent for all participants: It appears as one single system, supporting a service-oriented interaction of all relevant business processes.

So far, the efforts to manage this integration are pooled with the keyword Enterprise Application Integration (EAI). They contain a set of technologies and concepts such as Middleware, ETL-Tools and EAI-Software that focus on a central planning and control of business application data in real-time. In figure 2, an example for embedding an EAI Server into the enterprise architecture is shown:

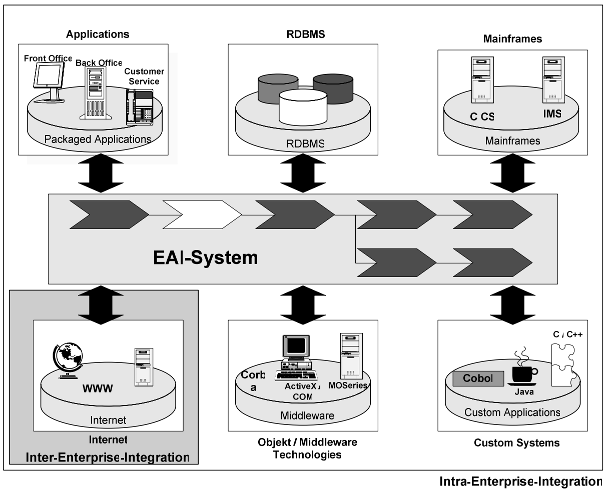

Fig. 2. EAI Scenario

In this scenario, a central EAI-Server manages the coordination of both, control and data flow, between the attached systems. In this manner, the number of required interfaces can be reduced, as there exists only one bidirectional connection from each system to the server.

Because of the central component (which is ordinarily called *Information Hub*), these approaches cause several problems that complicate a reliable integration of numerous systems. Major well-reported problems are the following:

- **Single point of failure**: The Information Hub is the central node between the different applications. In case of a breakdown, all business processes are affected, possibly even inoperable.
- **Bottleneck**: All network traffic has to be forwarded through the Information Hub. This results in an extremely high load of data that has to be handled. In a situation of peak load, the system performance will crash down dramatically.
- **Configuration Icebergs**: The EAI application must contain all relevant business transactions: The distribution of information has to be represented by formal rules for transformation and routing. Because of its complexity, interdependencies and lots of exceptions, the number of configuration rules increases exponentially. This may cause insufficient and fault-prone integration solutions.

We consider EAI systems as mission critical applications whose blackout is associated with a substantial business risk. In recent years, awareness of cost-intensive administration and insufficient management of complex business application systems by centralistic approaches has arisen. New fields of research such as *Autonomic*

Computing gave thought-provoking impulses to find better alternatives for an efficient management of business integration [4].

3 Recent P2P-applications

As denoted at the end of the last chapter, an important field of research consists in *Autonomic Computing*. However, the underlying principles are not new, but became popular by the introduction of file sharing systems, and thus have found their breakthrough.

P2P means a networked structure where the participants (herein called *peers*) interact and share resources directly and equitable, i.e. “sharing of computer resources and services by direct exchange between systems”. In contrast to client/server architectures, P2P networks do not include any kind of hierarchical structure. In the following, the term peer describes a participant within a P2P network.

In principle, single peers are independent from certain hardware platforms. This comprises a wide range of supposable systems, ranging from PDAs via desktop computers up to mainframes [5]. They are all characterized by the following properties [6]:

- **Client and server functionality:** Every peer is able to receive data from other peers as well as to provide data for others.
- **Direct exchange:** There is no central coordinating instance controlling the communication between the peers.
- **Autonomy:** It is in charge of each single peer at which time it provides which service, data or output to the network.

Especially the property of autonomy is of special importance for the integrated use of mobile computers such as notebooks, PDAs or mobile smart phones, as these per se can not be available to the network permanently.

Nowadays, P2P technology is used for non-commercial, private or academic purposes. Well-known examples are file-sharing applications (e.g. Napster, eDonkey, Kazaa) or grid computing projects (e.g. seti@home).

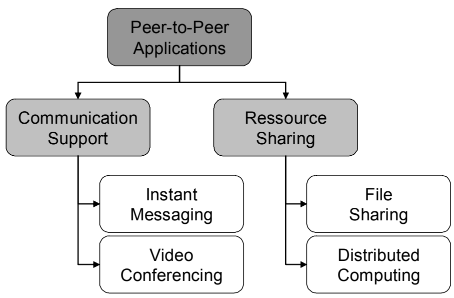

Fig. 3. Typical P2P applications

Current applications that are usable within business integration scenarios are seldom [7]. In particular, existing approaches can be categorized in two main clusters as shown in figure 3

Communication support is the most common kind of applications for P2P networks within enterprises. The best known representatives are Microsoft's Netmeeting (for video conferencing) as well as ICQ, AOL Instant Messenger or the Microsoft Instant Messanger for sending and receiving instant messages. These applications are characterized by an open and highly dynamic number of participants.

Instant messaging applications are used for the direct, real-time communication between peers. Thus, they allow the determination who of the known participants is online at a particular time. This accelerates business communication processes, as it ensures a synchronous, text-based communication that enables a direct reaction on incoming messages and accordingly business events. For example, the retailer *Land's End* uses instant messages for after-sales service and support. The probability of selling goods was enhanced by 67%, compared to customers not being supported by this instrument [8].

Video conferencing means the real-time transmission of video and audio data between two or more participants. It mainly serves as an appliance to support cooperation and collaboration processes. Thus, natural communication trough wide distances is possible. Currently, real meetings can be substituted by the use of videoconferencing. Simultaneously, time and traveling costs can be reduced significantly. Moreover, video conferencing systems are deployed for personal education purposes (classroom function) [9].

Resource sharing can benefit from the exponential development of computer performance in the last decades. In parallel, the price for computation power has decreased in at least the same dimension. Subsequently, the available capacity of today's clients is only rarely used, as most of the work load is assigned to server(s). P2P can profit from these idle resources and hereby achieve a drastic cost reduction, combined with other competitive advantages. The sharing of resources is implemented in two concepts:

File sharing provides shared access to any files that are stored locally. It consists of special, efficient mechanisms for searching as well as algorithms for a non-central storage [10].

Compared to central data storage applications, this kind of data access eliminates the single point of failure, transfers data to low-cost mass storages and levels off the peaks in network traffic. The main problem is to ensure data consistency within the network as well as a 24/7 availability. After all, only those applications that do not affect mission-critical cases found their way in the business context.

Distributed computing is used for complex business computing tasks such as product development, simulation, financial forecasting or data mining tasks. The main resource for these time-consuming processes is computation power. Distributed computing can substitute powerful single computer systems with a network of peers by decomposing (dividing) the problem in small sub-problems and spreading them within the network. Hereby in the ideal case, the computation power of host systems can be reached with a fractional amount of the costs being calculated originally. The costs even can be minimized by using idle time of the peers to calculate single computation tasks. Thus such computations can be even processed by small or medium-sized enterprises (SME) that usually do not have the capability to invest in powerful host systems [11]. Even world-class enterprises use distributed computing for this purpose: Intel introduced distributed computing in 1990 for the development of new microprocessors [12].

4 Our P2P Integration approach

Assigning the peer-to-peer paradigm to the context of business integration means that all systems and components of an enterprise work in a self-organizing manner. Thereby, administration and integration costs can be reduced. [13] A basis for this approach is the non-centralized architecture of a peer-to-peer system which can offer the following advantages:

- **Non-central topology**: The complete IT-infrastructure is defined by a variable number of flat (non-hierarchical) peers. Every peer offers (and receives) at least one service (e.g. generate an order, create an invoice, check consistency of data, dispose payment, etc.), where several peers may contain the same services.
- **Reliability**: There is no central component that may cause problems. If a peer breaks down, another peer with similar functionality can replace the broken peer. If a peer with unique services crashes, only those business process instances are affected that require that service.
- **Scalable performance**: The performance of the network can be enhanced nearly unlimited by appending additional peers. Already existing components do not have to be replaced.
- **Easy configuration**: It is no longer necessary to customize the whole EAI system by central transformation rules. Every peer only contains the business knowledge it requires to accomplish its functionality. The configuration of the complete architecture results from the sum of the configuration of the single peers.
- **Adaptive self-configuration**: By implementing intelligent search mechanisms, a peer can find the next service in the process chain by a broadcast into the network. If another peer is able to offer the desired service, it responds. From now on, this peer is part of the process chain and can accept tasks from other peers as well as delegating services to any peers.

While the advantages mentioned above are mainly of a technical nature, these features will not be sufficient to manage the complete field of business integration. There is also a high demand of adequate logical representations of business processes to provide an essential process-oriented view that focuses also economical aspects. The vision of both, distributed business processes that are associated with distributed IT-systems, allows an optimization of business processes as well as improving IT applications without interacting each other.

In our research project *Peer-to-Peer Enterprise Environment (P2E2)*, we develop an integration architecture that focuses typical business integration scenarios. As a proof of concept, we will also implement and evaluate a prototypic system. For achieving this ambitious goal, we collaborate with the working group for Databases and Information Systems of the Max-Planck-Institute in Saarbruecken (Germany) and several partners of the software industry. In Figure 4, a P2P scenario is shown that gives a first impression of our intention:

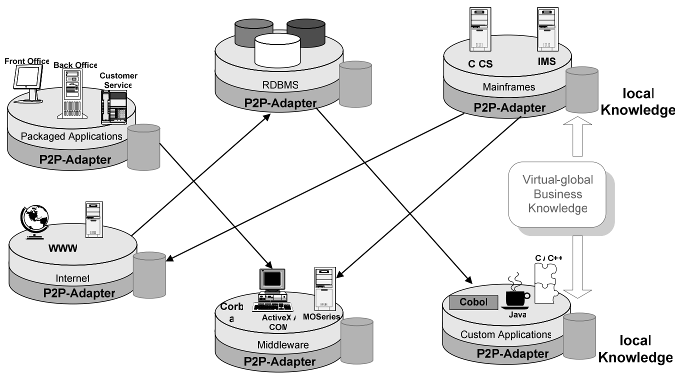

Fig. 4. P2E2 integration scenario

Every application system (AS) that participates in the whole business process is encapsulated by a P2P-Adapter. Every Adapter enhances the functionality offered by the single components with additional Web Services (WS) that allow a composition of very complex services by a dynamic interaction of different adapters. A peer can initiate business processes, embed local processes in the application flow and even get embedded by other peers. It only has its own *local business knowledge*, but can also acquire *global business knowledge* by interacting with other peers. In this way, a comprehensive management of meta-data can be achieved without requiring centralistic client/server architectures.

To implement the P2E2 Adapter, a technical architecture was specified. It consists of four detailed interface specifications, whose interrelation is illustrated in Figure 6.

The *Configuration Interface* reads the configuration data provided by the P2E2 Configurator. The Business API (*BAPI*) is an automated interface for executable Business Process models that can be created with standard BPM tools such as the ARIS toolset by IDS Scheer AG. To find the following Service that provided the next step in a process chain, a *Search API* is implemented that offers efficient P2P-Search algorithms. The completed API consists in the monitoring and controlling component; it covers the whole BPM-Lifecycle and allows to specify Controlling queries to measure the performance of the integration environment.

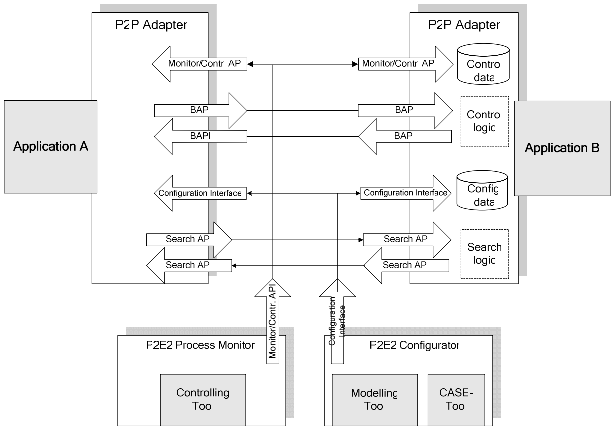

Fig. 6. Architecture of a P2E2 Adapter

To ensure the efficient use of the procedure models and methods, we will conceptually and technologically develop an integrated support tool that includes an *Adapter Development Kit* as well as some important pre-customized standard adapters and an *Adapter Management Tool*.

The problem faced above is complex and versatile. Therefore, a highly structured and planned proceeding will be necessary. In contrast to other approaches, we follow a 'meet in the middle' strategy, analysing the problem space and creating solving concepts both from the business oriented as well as from the system/technology oriented direction. This ensures that the conceptual solutions are suitable for the business problems targeted and that they are realizable with today's state-of-the-art technologies.

The business oriented approach will evaluate the requirements of enterprises within internal or collaborative business integration scenarios. The main reference object here is the (abstract) business process that has to be supported. Thus, the business oriented conceptual solution has to provide mechanisms and techniques how to interconnect independent business processes using the P2P paradigm. The main challenge is the lack of a central coordination instance. As a logical consequence, appropriate business process negotiation techniques have to be developed.

Looking at the system oriented problems, the main question is how to find a mapping between heterogeneous application systems in conformity with the business processes and rules to be supported. This does not only mean to connect interfaces, but also requires to find *reasonable matches* within concrete contexts, as well as to handle a reliable control of the interaction.

Peer-to-peer technologies have been proven to be very flexible and robust. Hence, new methods and algorithms for a distributed interface and interaction management will be created using the P2P paradigm. Our idea is to enable an auto-configuration of the interaction between two independent application systems that succeeds to predefined business processes and that is also compliant to existing, constraining business rules.

Finally, only developing the two solving concepts described above is not sufficient, as there are strong interdependencies between business- and system-layer. The formation of business process chains within business integration use cases is always limited by the capabilities of the existing applications supporting the business processes. On the other hand, applications themselves can only be interconnected in a way that the combination of systems is realizing a predefined business process.
Hence, a relationship between the two partial solutions has to be found and specified. The combination of these three concepts will be a conceptual methodology for a dynamic binding of business processes to the behaviour of the distributed environment via a (semi)automatic reconfiguration in case of need.
The technological *P2E2* results will consist of a family of congeneric components (adapters) that can be linked to application systems. These components will provide a non-central IT-support for interactive business processes within distributed business integration environments. Interactive business processes are business processes that are structurally influenced by humans (user interaction). Therefore supporting interactive business processes have to provide operative user interfaces (front-end) whereas non-interactive ones only need a system-to-system connection (EDI). The adapters themselves will not have to rely on a central control unit. It mainly addresses those business cases where a central integration approach is not reasonable, not desired or not realizable due to organizational or technical restrictions. Moreover, *P2E2* extends the focus of classic EAI solutions by the aspect of user's interaction and influence. This also reflects the business level view where such a distinction is not reasonable. Expected properties of the *P2E2* implementation are robustness, fault tolerance, adaptability, dynamic behaviour and scalability.
For business usage, a coherent computing of business process is necessary, but not sufficient. For the purpose of analyzing and optimizing processes continuously, a detailed history of key performance indicators (KPI) is required. This data is in particular used on tactical, dispositive level to manage the performance of critical core processes. Within business integration scenarios with several systems involved, the gathering of these figures by measuring the processes is difficult and complex. *P2E2* will cope with this problem domain by integrating special measurement and control instruments within each peer. Thus, an integrated and consistent business process controlling and measurement within distributed environments will be enabled. Hence, *P2E2* will develop the measuring methods, implement them in the adapters as well as create a reporting and analyzing tool for an efficient process performance measurement.

5 Conclusions

The sustainable success of peer-to-peer technology within the non-commercial sector (file sharing and instant messaging) advises it also for the business sector. Within enterprises, there were several potential application cases for the productive use of P2P technology.
Especially for business integration scenarios, P2P seems to be an interesting alternative, in particular those with are characterized by non-hierarchical, flat organization

structures. Thus, the integration of application systems using peer-to-peer technology is obvious.

Indeed, only integrating applications is not sufficient for efficient usage in business environments. In order to effectively and efficiently run a business, the respective business process is the central object. Therefore, an efficient support and integration of operative business processes is the primary objective.

P2E2 addresses this aim by creating a family of generic components that provide a general connectivity between business application systems and that act with client and server functionality. Hence, the classical separation between client and server will be resolved in favor of functional ad-hoc decisions. The coordination and control between the single adapters will not be managed by a central unit, but case-based negotiated between the adapters and afterwards transferred as needed. Thus, there is no functional dependency of the whole functionality network from only one edge.

Within a consortium of 6 partners, covering basic academic research, applied research, software industry and IT-consultancy, *P2E2* deals with that problem domain for a total of 24 months. During the project lifetime, several showcases realizing real business scenarios are built up and deployed in order to prove the practical use of the project results.

Finally, the use of P2P technologies within business integration scenarios marks a young area of research and development within the field of business information systems. It will remain an interesting object for actual research. *P2E2* has made a first step into this domain trying to solve some of the common problems, leaving some others open and maybe also rising some new ones. The way to consistent, operative environments that integrate enterprises and enterprise application trough peer-to-peer technology will be long, but with P2E2 we have entered the road.

References

1. Linthicum, D.: Enterprise Application Integration. Boston (2000)
2. Meta Group: E-Business und Enterprise Application Integration. (2001)
3. Smith, D.: Web Services Enable Service-Oriented and Event-Driven Architectures. In: Business Integration Journal, (May 2004) 12-14
4. Kephart, J.; Chess, D.: The vision of Autonomic Computing. IEEE Computer, (2003)
5. Intel Corporation: P2P Filesharing at work in the Enterprise. http://www.intel.com/deutsch/ebusiness/pdf/ wp011301_sum.pdf, online: 2004-01-21
6. Schoder; D., Fischbach, K.: Peer-to-Peer: Anwendungsbereiche und Herausforderungen. In: Schoder, D.; Fischbach, K.; Teichmann, R. (Ed.): Peer-to-peer: ökonomische, technische und juristische Perspektiven. (Springer) Berlin et al. (2002) 3-21
7. Saini, A.: Peer to Peer Distributed Business Processes. http://www.dmreview.com/whitepaper/WID463.pdf, online: 2004-06-10
8. Tischelle, G.; Swanson, S.: Not just Kid Stuff. http://www.informationweek.com/story/showArticle.jhtml?articleID=6506155, online: 2004-06-14
9. ViDe Video Conferencing Cookbook, www.videnet.gatech.edu/cookbook/ , online: 2004-04-02
10. Schoder, D.; Fischbach, K.: Peer-to-Peer : Anwendungsbereiche und Herausforderungen. In: Schoder, D.; Fischbach, K.; Teichmann, R. (Ed.).: Peer-to-peer: ökonomische, technische und juristische Perspektiven. (Springer) Berlin et al. (2002) 3-21, 6

11. Intel Corporation: Peer-to-Peer Computing in the Enterprise: Implications of IT and Business Decision Makers, http://www.intel.com/deutsch/ebusiness/pdf/ wp020702_sum.pdf, online: 2003-02-12, 3
12. Intel Corporation: Peer-to-Peer: Rechenleistung auf breitester Basis. http://www.intel.com/deutsch/eBusiness/ products/peertopeer/ar011102_p.htm, online: 2004-03-01
13. Leymann, F.: Web Services: Distributed Applications Without Limits. In: Weikum, G.; Schöning, H.; Rahm, E. (Ed.), Proceedings of the 10. Conference of Databases for Business, Technology und Web (BTW 2003). LNI 26, Gesellschaft für Informatik, (2003)
14. Zieger, A.: What Can P2P Apps Do for Enterprise Users. http://e-serv.ebizq.net /p2p/zieger_1a.html, online: 2003-04-17