

An economic decision model for business software application deployment on hybrid Cloud environments

Jörg Strebel¹, Alexander Stage²

*¹Institute of Information Systems and Management (IISM),
Information & Market Engineering (IM), Karlsruhe Institute of Technology*

*²Internet-based Information Systems (IBIS),
Department of Informatics, Technische Universität München*

1 Introduction and Motivation

Large-scale, virtualization-based distributed computing environments have emerged during the last couple of years providing scalable computing resources as a service. From a customer perspective, these environments offer an almost inexhaustible pool of resources that can be utilized at any time. Especially their ability to provision computational resources to virtual machines on demand renders them a promising hosting platform for permanently running enterprise applications. In the past, a main challenge in operating corporate data centers has been and still are resource demand fluctuations, recurring patterns at various time scales as well as trends and shifts in application resource demands (Urgaonkar et al. 2008; Gmach et al. 2008). In dedicated hosting environments these phenomena often required peak demand oriented capacity planning in order to prevent resource shortages, which in turn lead to low average resource utilization levels. Even in virtualized data centers that allow for workload consolidation, correlated bursts or regular peak demand periods require the acquisition and maintenance of large pools of physical servers. Enterprises that aim at lowering overall data center costs, which include (but are not restricted to) hardware acquisition, opportunity, energy, and labor costs for administration, can utilize this new breed of scalable application hosting platforms in order to achieve cost reductions. In order to compensate for peak demands or whenever the overall data center workload exceeds locally available resource capacities, workloads can be evicted from the local data centers by offloading virtual machines into public IaaS (Infrastructure-as-a-Service) clouds. However, the existence of various pricing and tariff models renders capacity planning and management decisions difficult, even in highly transparent infrastructure

markets. On a strategic level, decision makers are confronted with the problem of selecting economically beneficial providers and virtual machine tariffs, while on an operational level the workloads to evict and the assignment of appropriate virtual machine types needs to be done in a way that conforms with the tariff and provider selection outcomes. Both problems are highly interdependent and require sophisticated models for making efficient decisions. Consequently, we propose an integrated decision model for selecting virtual machine tariffs, types and workloads to evict from local data centers that aim at economic efficiency. Our model can be used to determine, given internal cost structures, an optimal mix of internal and external resources for a given planning period.

1.1 Related Work

The majority of related work in this field originates from research related to Grid computing; as IaaS Cloud Computing is an emerging computing paradigm, a significant body of research literature has yet to be established. Conceptually, both paradigms share the infrastructure focus (the characteristics of IaaS and Grid computing are explained in Weinhardt et al. (2009)). The research problem mentioned above can best be studied using the following three perspectives: research on TCO models, research on decision support tools and research on resource management. This paper focuses on a cost-oriented decision model only, as benefits resulting from the usage of IaaS resources are generally hard to quantify. One of the most important cost-oriented models both used in research and in real-life settings is the Total-Cost-of-Ownership (TCO) model; initially created by Gartner Inc. in the late 1980s to assess the cost structure of Client/Server computing (Silver 2007). One of the latest varieties of TCO models applies the approach to the field of IaaS; Leong (2009) proposes a TCO model for IaaS which will act as one of the inputs for the TCO model used in this paper. As a researcher in Grid computing cost, Opitz et al. (2008) tries to quantify the total cost of ownership (TCO) for grid computing resource providers in absolute cost figures from real-world scenarios and comes up with an estimate for the total cost of a CPU hour offered by a commercial resource provider; the model in Opitz et al. does not include any storage costs, which are of major importance in business settings. So, the necessary cost calculations for computing resource-consuming enterprises have to look different from the model in Opitz et al. for resource providers. Risch and Altmann (2008) analyzed a number of Grid computing scenarios using a cost-based approach; they showed that Grid computing is beneficial in scenarios, where either short and infrequent peaks have to be covered or where data backups have to be conducted or where lightly used resources have to be replaced. However, they recommend that each company performs its own cost analysis as the benefits are depending on the cost level of the in-house resources. Gray (2003) specifically deals with the decision when to outsource given the price ratios between the different computing resources. Generally, the business model behind Grid computing remains case-specific; he maintains

that business benefits are only realized for very CPU-intensive software applications. As a conclusion, it can be stated that the cost models for Grid computing found in the current research literature are rather incomplete, scenario-specific and not focused on the corporate decision-maker.

The decision-support perspective is the second important perspective on the potential of Grid computing. Kenyon and Cheliotis (2004) addressed the area of Grid resource commercialization. They conceive Grid resources as commodities and apply financial instruments for conventional commodities like gas or electricity to those resources. Within the scope of their analysis, they identified the necessity for decision support, when Grid users buy or sell Grid resources on a Grid marketplace. However, the need for such elaborated decision support models will only arise in a real Grid resource market.

In the area of resource management, Rolia et al. (2003) suggest a resource-management framework for automatic software application placement in the data center using Grid-computing principles like resource allocation and scheduling. Their main focus lies on the optimization of in-house data-center resources; they do not address the question under which conditions to use external resources. Their optimization approach minimizes the number of CPUs and does not consider actual cost factors from an enterprise environment.

Literature on outsourcing decision-support such as the reviews of Gonzalez et al. (2006) and Dibbern et al. (2004) show that the question of what to outsource has mostly been analyzed conceptually or in a positivist fashion so far, but not through mathematical modeling, even though cost is universally recognized as a dominant criterion for outsourcing (Gottfredson et al. 2005).

In the following section, a cost model, an optimization model and a regression model are developed and their underlying assumptions are stated. In Sec. 3, the models are instantiated and experiments are set up to determine the driving factors for hybrid IaaS usage in a cost-optimal fashion; the solution is then discussed and future research directions are given in the last section.

2 Methods

In this section, an optimization model for cost-optimal assignment of virtual machines to computing resources and a linear regression model for analyzing statistical properties of those assignments are presented. Both are fundamental for analyzing the question of IaaS computing use cases.

2.1 Optimization Model

Current infrastructure providers offer virtual machines on demand that can be entitled with any of a predefined set of resource bundles. We denote a specific resource bundle as a virtual machine type and define, for each resource $e \in E$, the

maximum amount of guaranteed resources available to a virtual machine at any point in time (e.g. the number of CPU cores or the amount of RAM assigned to that particular virtual machine). Each type of virtual machine is associated with a tariff. Customers can decide to make initial investments and buy virtual machine time contingents. A contingent is valid for a certain time period (the discount period) and consist of an initial capacity s^d of computing time and a certain hourly tariff c^m . The initial capacity is charged as a lump sum, afterwards the computing fees will be hourly billed. We also factor in extra capacity as a percentage of the overall, aggregated capacity to allow for some failover capacity or hot stand-by resources. The model furthermore assumes, that any application can take advantage of a scale-up approach, i.e. it can make use of additional IT resources once they become available.

Table 1: Model parameters

I	Set of available servers in the enterprise data center. s_i is the computing capacity of server type $i \in I$.
J	Set of all hosted business applications. It is assumed that each application runs in its own virtualized environment on a separate software stack and operating system. p_{jt} is capacity demand of virtual machine $j \in J$ at time period $t \in T$. n_{jt}^i are the inbound networking requirements, n_{jt}^o are the outbound networking requirements and w_{jt}^s are the storage requirements of virtual machine $j \in J$ at time period $t \in T$.
E	Set of all resource bundles available from the IaaS provider. s_e is the computing capacity of server type $e \in E$.
T	Set of all time slots $T = \{t_1 \dots t_n\}$, ($t \in T$) of equal length in the planning period
K	Set of all discount periods $K = \{k_1 \dots k_l \mid k_x = \{t_{xd} \dots t_{x(d+1)}\}$, where d is the number of time slots in which a contingent is valid and where l the number of discount periods.
α	Ratio of failover capacity to regular capacity in the corporate data center. The ratio depends on the criticality of the involved application system. Production-grade ERP systems are usually secured by a high-availability setup in which $\alpha = 1$.

The CPU, networking and storage requirements can be estimated from historical real-world system traces coming from performance monitoring systems (c.f. Rolia et al. (2003)). Table 2 and Table 3 describe the cost model used throughout the paper. It is based on work of Leong (2009) and Gray (2003), but does not feature LAN costs, which are magnitudes cheaper than Internet transports (Gray 2003).

Table 2: Cost model for enterprise resources

Cost factor	Explanation
Server Expenses	The regular payments per period for internal server hardware could either be the server rent or server depreciation costs. The server expenditures do not depend on the actual server utilization; if the server is used within T , its expenditures have to be paid for the complete time interval T . Potential hardware replacements and other repairs are included in the hardware cost (the hardware cost also includes cost factors like electrical power, cooling, hardware handling, operating system license, virtualization software licenses, server operations and data center facilities and other software fees). Cost c^{sm}_i for running a server $i \in I$ for one period of time in the enterprise data center.
Storage	The storage costs include all surrounding activities like backup and recovery, device management etc. Cost: c^o_t for one period.

Table 3: Cost model for IaaS resources

Cost factor	Explanation
Virtual machine CPU hour	Cost c^{sm} for running a virtual machine in a time contingent under the hourly tariff. The IaaS provider offers the OS licenses along with the server. The software cost is included in the server rent. It is assumed that support and management service charges are already included in the virtual machine charges.
Time Contingent fee	Cost c^{disc} cost of buying a contingent.
Cloud storage	Cost c^{stor}_t for using a unit of storage for one time period.
Internet bandwidth	Internet transports are fundamental; only the variable cost of Internet bandwidth is considered here, as the Internet connection will most likely be used by other IT systems in the enterprise. Cost: c^{wan}
Data transfer outbound	Cost of serving data from the IaaS provider. Cost: c^{out}
Data transfer inbound	Cost of uploading data to the IaaS provider. Cost: c^{in}

The task of the optimization model consists of assigning each software application in each period to a resource from either the internal or the external pool. The model is essentially a mixed-integer programming problem (MIP). The decision variables of the optimization model are as follows: if x_{jt} is set to 1, then application j

will run on the internal server i in period t ; x_{ijt} is defined in the analog way for external servers e in period t . z_i records if the internal server i was used at all in T (e.g. $z_i = 1$ means that internal server 1 was used at least once). This information is required to accurately calculate the hardware costs. y_k records if any external resource was used in discount period k ; if that is the case, then the external compute fees for that discount period are factored in the total cost. o aggregates the required size of the internal storage infrastructure. c^{cc}_k stores the total computing cost in discount period k . The objective function combines the cost factors and the decision variables in one function, which is solved to minimize total costs; table 4 shows the summands of the objective function.

Table 4: Objective function summands

Cost factor	Cost coeff.	Cost component
Server expenditures	c_{it}^{hw}	$(1 + \alpha) \sum_{\forall i \in I} z_i \sum_{\forall t \in T} c_{it}^{hw}$
Internal Storage	c_t^o	$\sum_{\forall t \in T} c_t^o * o$
WAN usage	c^{wan}	$c^{wan} \sum_{\forall t \in T} \sum_{\forall j \in J} (n_{jt}^I + n_{jt}^O) \sum_{\forall e \in E} x_{ejt}$
Data transfer to enterprise	c^{dout}	$c^{dout} \sum_{\forall t \in T} \sum_{\forall j \in J} n_{jt}^O \sum_{\forall e \in E} x_{ejt}$
Data transfer to the external pool	c^{din}	$c^{din} \sum_{\forall t \in T} \sum_{\forall j \in J} n_{jt}^I \sum_{\forall e \in E} x_{ejt}$
External Compute fees	c_k^{cc}	$\sum_{\forall k \in K} c_k^{cc}$
External storage fees	c_t^{stor}	$\sum_{\forall t \in T} c_t^{stor} \sum_{\forall j \in J} n_{jt}^{st} \sum_{\forall e \in E} x_{ejt}$

The constraints of the optimization model are as follows: one application is assigned to exactly one server per period. It is neither possible to run one application on several servers nor to run several application instances during the same period. Each IaaS resource can only run at most a single application instance; hence it is not possible to consolidate virtual workloads on a IaaS resource, as they are already virtualized. The combined load per period placed on each internal or external server is at most the maximum capacity per period of that server; z_i is set as soon as server i is used at least once in T . The internal storage requirements mandate that o is set to the minimal amount of internal storage required across all periods. The total computing cost c^{cc}_k in discount period k is at least as big as the one-time fee for the time contingent and it is at least as big as the amount needed to cover the sum of the hourly compute fees in discount period k . As a result, the minimum of the two cost factors, time contingent cost and hourly computing cost, is chosen for each discount period.

2.2 Linear Regression Model

The main purpose of developing this regression model is the assessment of the internal validity of the aforementioned optimization model; in case the optimization model is internally valid, the regression model will prove obvious relationships between the variables of the optimization model. Moreover, it can be helpful in finding previously unknown relationships. The basis for the regression model is the experimental data coming from the optimization model runs. Each of those runs can be considered an experiment in which a certain number of business applications are assigned to internal and external resources per time slot in a cost optimal fashion. Hence, the regression model is constructed in several steps. First we define which experiments to execute. The total number of possible experiments is the cardinality of the power set \mathcal{Q} over J with $|\mathcal{Q}| = 2^{|J|}$. It is obvious that, even for a small number of applications, only a sample of the entire population can actually be analyzed in experiments. In the following, $J_a \in \mathcal{Q}$ ($n_a = |J_a|$) will designate the set of applications for one experiment. The total number of experiments and the choice of experiments depends on the purpose of the analysis and the required statistical confidence level. As a second step, the sum per time slot of all load traces in J_a for the same resource group is calculated $p_t = \sum_{j \in J_a} p_{jt}$, and analog

for n_{jt}^o ; n_{jt}^{sto} . This step leads to three unified load traces p_t , n_t^o , n_t^{sto} , in which the summation should have leveled out any negative correlations among the traces. The third step consists of calculating two characteristic measures for each of those three unified time series, the arithmetic mean and the empiric coefficient of variation; those values will be the independent variables in the regression model and are called \bar{p} , \bar{n}^o , \bar{n}^{sto} for the averages and s_p ; s_n^o ; s_n^{sto} for the standard deviations. For

example: $\bar{p} = \frac{\sum_{t \in T} p_t}{T * n_a}$. The averages are additionally divided by n_a ; otherwise, the

averages and the coefficient of variation would automatically become larger, as n_a increases and this would negatively affect the regression model. In the fourth step, the cost optimization model computes the resource assignments of applications to internal and external servers; only those applications $j \in J_a$ will be used for the optimization run. After the run, the decision variables contain the values which IaaS servers have been used. This information is used to determine \hat{Y} , the dependent variable which counts the average number of deployed Cloud instances per

time slot: $\hat{Y} = \frac{\sum_{\forall t \in T} \sum_{\forall e \in E} \sum_{\forall j \in J_a} x_{ejt}}{T}$. In the fifth step, an analysis of the correlation

between the independent variables and \hat{Y} , the dependent variable can be conducted. The relationship will then describe, how those statistical properties are

related to IaaS usage and will help answering the question under what conditions IaaS usage is beneficial. The linear regression model thus has the following basic formulation:

$$\hat{Y} = b_0 + b_1 * n_a + b_2 * \bar{p} + b_3 * s_p + b_4 * \bar{n}^{sto} + b_5 * s_{n^{sto}} + b_6 * \bar{n}^o + b_7 * s_{n^o}$$

$b_{1...7}$ are the regression coefficients. Inbound network loads n^i are not part of the model as they are assumed to be strongly positively correlated to their counterparts n^o .

3 Results

3.1 Scenario definition

The values of the enterprise cost factors were estimated from literature sources. The values used are similar to those suggested in Hamilton (2008) and Armbrust et al. (2009). J , the set of all hosted business applications, consists of 18 real-world SAP ERP systems in this scenario. Their resource consumption $(p_{jt}, n_{jt}^i, n_{jt}^o, n_{jt}^{sto})$ were measured as part of the performance monitoring ongoing in corporate data centers. The load traces used here were collected from Mid-February until end of April 2009 ($n=65$ days); their temporal resolution is one day, i.e. the load traces contain the maximum of each day's resource consumption. α is 0, as the ERP systems in this sample were development systems with no safety capacity in place. In this scenario, GoGrid¹ was used as a provider of resource bundles E . They offer four differently sized time contingents on a monthly basis. In each of the following optimization runs, the most efficient tariff was chosen. In total, 31 experiments were conducted (6 experiments for each level of n_a with $n_a \in \{3;6;9;12;15\}$ and one experiment with $n_a = 18$); as each experiment required 4 runs of the optimization model, a total of 124 runs were executed. Additionally, the optimization model was parameterized in another set of experiments to just utilize enterprise resources; another sequence of 31 experiments was conducted in that fashion. The optimization model was computed using the mixed-integer linear programming solver SCIP (Achterberg 2007).

3.2 Experimental results

When comparing the total cost of the experiments using IaaS computing with the total cost of the experiments using only enterprise resources, the Cloud-based experiments were on average 28% cheaper than those in the enterprise datacenter. However, the standard deviation of those average savings lies at 15%, which indicates a high volatility in the saving potential.

¹ <http://www.gogrid.com>

The regression model was then applied to the results from the Cloud-based optimization runs to analyze the determinants of those savings. For assessing the regression model, the adjusted coefficient of determination R^2 , the F statistic (as a measure of goodness-of-fit) and the standard error have been calculated. The adjusted R^2 is 0,915 which indicates that the linear model accounts for almost all the variance in the sample. The F statistic is 47,25, which means that we have to reject the null hypotheses, that there is no relationship between the independent and the dependent variables. The level of significance is well below 5%.

Table 5 shows the correlation coefficients and their statistical properties. The p-value on the coefficients shows the level of significance of this variable, as determined by a two-tailed t-test. b_1 , b_2 , b_4 are significant at least on a 5% level, b_5 is significant on a 10% level, whereas b_3 , b_6 , b_7 are not significant for any reasonable level. The beta values are standardized correlation coefficients and show the relative importance of every coefficient; the higher the beta value, the stronger the impact on the dependent variable. The use of a linear model can be justified as the assumptions of the linear regression model are satisfied (based on Backhaus et al. (2006)): the number of correlation coefficients is smaller than the number of observations, the regressors are error-free with high probability as they were measured directly on the affected servers using performance monitoring software. Autocorrelation and heteroscedasticity among the error terms can be neglected as a visual inspection of the residual plot showed no signs of those effects. Hence, the internal validity of the model can be assumed.

Table 5: Regression model

Variables	Coefficients	Coefficient values	t Stat	P-value	Beta Value
Intercept	b0	-0,7170	-1,3953	0,1762	
n_a	b1	0,3772	5,1855	0,0000	0,8159
\bar{p}	b2	-1,4111	-4,8447	0,0001	-0,5620
s_p	b3	-0,2738	-1,3997	0,1750	-0,1215
\bar{n}^{sto}	b4	0,0074	4,9895	0,0000	0,5474
s_n^{sto}	b5	6,5614	1,7964	0,0856	0,2265
\bar{n}^o	b6	0,0058	0,2595	0,7976	0,0156
s_n^o	b7	-0,0105	-0,4835	0,6333	-0,0464

4 Discussion

Based on the experimental results described above, a first observation seems noteworthy: n_a , the number of applications available for Cloud computing, positively impacts IaaS usage more than any other variable. This leads to the hypothesis

that the cost-effectiveness of Cloud computing rises with the amount of Cloud-ready business applications; this outcome seems logical, given that Cloud providers offer time contingents where bigger contingents are associated with lower marginal cost for IT resources. Interestingly, \bar{p} , the variable with the second largest impact, has a negative effect on IaaS usage. This means that the more business applications become Cloud-ready, the less cost-effective it is to deploy them in the Cloud. This outcome could be explained with the higher consolidation potential that a growing number of virtualized applications possesses. The hypothesis here is that, the more applications are available for internal server assignment, the better those internal servers can be utilized. The amount of storage \bar{n}^{sto} is the variable which exerts the third biggest impact on IaaS usage and which is positively correlated to it. This outcome seems sensible and as modern ERP systems usually have rather demanding storage requirements compared to other applications and as the Cloud storage cost is assumed to be smaller than the storage cost in the corporate data center (see 3.1 for the cost values). s_p is not statistically significant, so it can be concluded that it does not exert any measurable effect on IaaS usage based on the data in this sample. This outcome suggests the rather counterintuitive hypothesis, that Cloud use cases which offload peak IT resource demands into the Cloud may not be beneficial cost-wise for the client; this notion is further supported by the negative sign of the correlation coefficient of this variable. However, further research is needed on this hypothesis, as s_p might simply not be an adequate measure for CPU load variability.

4.1 Limitations on the Research Design and Material

The results presented above abstract from the differences in the technical properties of IaaS resources and corporate data center resources and solely focuses on the cost aspect. Switching costs (for transferring applications between the Cloud and the corporate data center) are not included in the optimization model because it is assumed that the technical incompatibilities will be neglectable in the future; currently switching costs can still be assumed to be considerable. The effect of switching costs in the optimization model on IaaS usage will likely be negative; switching costs increase the total cost of Cloud usage and hence, the necessary critical mass of applications for a beneficial IaaS usage will be higher.

The optimization model also currently does not allow for several competing tariffs per time slot in one model; only one tariff per time slot can currently be included. There is no linear formulation of the resulting multi-tariff optimization problem, but linearity is required for certain desired properties of the solution approach (availability of efficient solvers, feasibility of global optimization approach). The external validity of the regression model must be considered rather low because of the small number of experiments; however, it can be argued that

the results obtained show a clear first glimpse of the influence factors and their effects on IaaS usage.

4.2 Conclusion

This paper suggests a novel model of optimizing the cost of IT resource usage for enterprises and demonstrates a way to analyze the optimization model outcomes. The resulting method is helpful for both optimizing the internal and the external deployment of an application; it can be set up using data that is readily available in the enterprise (system traces, internal cost figures). A business scenario demonstrates the usefulness of the method. However, as the scenario shows, using external resources is mainly beneficial for scenarios in which a great number of applications with low computing requirements and high storage requirements need to be outsourced. A careful analysis of a larger number of business scenarios and a larger number of IT resource providers has to be conducted in order to verify those findings.

References

- Achterberg T (2007) Constraint Integer Programming. Dissertation, Technische Universität Berlin. <http://opus.kobv.de/tuberlin/volltexte/2007/1611/>. Accessed on 2009-11-15.
- Armbrust M, Fox A, Joseph AD, Katz RH, Konwinski A, Lee G, Patterson DA, Rabkin A, Stoica I, Zaharia M (2009) Above the Clouds: A Berkeley View of Cloud Computing. Technical Report UCB/EECS-2009-28, University of California at Berkeley.
- Backhaus K, Erichson , Plinke W, Weiber R (2006) Multivariate Analysemethoden (11th Edition). Springer Verlag, Berlin.
- Dibbern J, Goles T, Hirschheim R, Bandula J (2004) Information Systems Outsourcing: A Survey and Analysis of the Literature. *Database for Advances in Information Systems* 35(4):6—102.
- Gmach, D, Krompass S, Scholz A, Wimmer , Kemper A (2008) Adaptive quality of service management for enterprise services. *ACM Transactions on the Web (TWEB)* 2(1):1—46.
- Gonzalez R., Gasco J, Llopis J (2006) Information systems outsourcing: A literature analysis. *Information & Management* 43(7):821—834.
- Gottfredson M, Puryear R, Phillips S (2005) Strategic sourcing: From periphery to the core. *Harvard Business Review* 83(2): 132—139.

- Gray J (2003) Distributed Computing Economics.
http://research.microsoft.com/research/pubs/view.aspx?tr_id=655. Accessed on 2008-10-05.
- Hamilton J (2008) Internet-Scale Service Efficiency. In: Large-Scale Distributed Systems and Middleware (LADIS 2008), Yorktown Heights.
- Kenyon C, Cheliotis G (2004) Grid Resource Commercialization: Economic Engineering and Delivery Scenarios. In: Nabrzyski J, Schopf JM, Weglarz J (Eds.) Grid Resource Management: State of the Art and Future Trends (1st Edition). Kluwer Academic Publishers, Norwell.
- Leong L (2009) Toolkit: Estimating the Cost of Cloud Infrastructure. Research report G00165397, Gartner Inc., Stamford.
- Opitz A, König H, Szamlewska S (2008) What Does Grid Computing Cost? *Journal of Grid Computing* 6(4):385—397.
- Risch M, Altmann J (2008) Cost Analysis of Current Grids and Its Implications for Future Grid Markets. In: Altmann J, Neumann D, Fahringer T (Eds.) Grid Economics and Business Models 5th International Workshop, GECON 2008, Las Palmas de Gran Canaria.
- Rolia, J, Andrzejak A, Arlitt M (2003) Automating Enterprise Application Placement in Resource Utilities. In: Brunner M, Keller A (Eds.) Self-Managing Distributed Systems - 14th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, Berlin.
- Silver MA (2007) Gartner PC TCO: The Next Generation. Research report G00153432, Gartner Inc., Stamford.
- Urgaonkar B, Shenoy P, Chandra A, Goyal P, Wood T (2008) Agile dynamic provisioning of multi-tier Internet applications. *ACM Trans. Auton. Adapt. Syst.* 3(1): 1—39.
- Weinhardt C, Anandasivam A, Blau B, Borissov N, Meinel T, Michalk W, Stöber J (2009) Cloud-Computing. *Wirtschaftsinformatik* 51(5):453—462.