

Helix Nebula – The Science Cloud

Title: Analysis of the existing supply and demand side – Big Science - business processes for the procurement of IT infrastructure services

Editor: SAP - Burkhard Neidecker-Lutz, Julia Doll

Work Package: WP7

Submission Date: 19.03.2013

Distribution: public

Nature: report



Revision History		
<i>Date</i>	<i>Version</i>	<i>Contributor(s)</i>
30.12.2012	1	Burkhard Neidecker-Lutz, Julia Doll
24.01.2013	1.1	Burkhard Neidecker-Lutz, Julia Doll, Michael Symonds, Jurry de la Mar, Sergio Andreozzi
06.02.2013	1.2	Michael Symonds
08.03.2013	1.3	Bob Jones, Dante Gregorio
18.03.2013	1.4	Jurry de la Mar



Table of Contents

Executive Summary	4
1. Market Imperfections.....	7
1.1 Investment Costs	7
1.2 Switching Costs	8
1.3 Transparency.....	9
2. Procurement Process.....	11
2.1 The procurement process of public research organizations	11
2.1.1 Public procurement rules	11
2.1.2 Procurement process of public organizations in HN.....	12
2.1.3 Negotiations.....	13
2.1.4 Restricted tender.....	14
2.1.5 Open tender	15
2.2 The Procurement Process of PaaS and SaaS providers	16
3. Offer Creation Process	18
3.1 Offer Creation Process of PaaS/SaaS providers	18
3.2 Offer Creation Process of IaaS providers.....	19
3.3 New offers.....	20
3.4 Competitive offers.....	21
4 Budgeting, Accounting and Costing.....	23
4.1 Budgeting and Accounting process of public research organizations..	23

4.2	Budgeting, Accounting and Costing of PaaS/SaaS providers.....	24
4.3	Costing of IaaS Providers	25
5	Matching the Procurement and offer creation process.....	26
5.1	Negotiation.....	26
5.2	Restricted and open tender	26
5.3	Procurement process of SaaS/PaaS providers.....	28
6	Matching budgeting/accounting and costing.....	30
7	Theoretical & practical difficulties for decision making in cloud markets.	31
7.1	A Sample Cloud Application	31
7.2	Calculating Prices	33
7.3	Consequences for rational demand side decisions	36
7.4	Consequences for providers.....	37
8	Conclusion: Next steps and obstacles to overcome.....	39

Executive Summary

The description and analysis of business processes in this deliverable define the framework and restrictions for the definition of potential business models. In this context it was essential to uncover mismatches and to define next steps in order to combine procurement processes on the demand side (segment restricted to big science) and the offer creation processes of the supply side (commercial cloud service providers).

Further, accounting and costing processes of both sides and the process of budget provision on the demand side has been examined to find potential restrictions on future contract terms.

In the course of our investigations based on conducted surveys and interviews we found that Helix Nebula is a scenario shaped by market imperfections such as high switching costs, lack of transparency and risks as described in this deliverable. This and mismatches in the business processes hinder participants of this market to conduct rational decisions and move forward.

To specify some of the mismatches found, investment risks seem to be a rather minor obstacle, since it can easily be avoided through cooperation. They occur on the supply side if a major amount of resources are required without a long term commitment from the demand side. In order to share risks we found that suppliers offer a price reduction of up to 2/3 if the contract length reflects the depreciation time of resources.

In this context transparency is a more complex obstacle as it is restricted by facts such as complex comparability of prices. Even in the simplest case (same pricing metrics) different units offered and different resource types can make it almost impossible to compare providers.

Assuming it would be possible to compare prices; customers would still have to deal with high switching costs that can make a change to the most cost effective provider unprofitable. Hence, it also prevents providers from getting the demand they would be entitled to.

Looking at the mismatches of business processes the major challenge results from a very strict public procurement process following EU rules that EU research organizations have

to abide by. This induces further costs and effort on the demand and supply side and requires suppliers to optimally cope with it, since there is no workaround.

Most of the obstacles found can be at least diluted to a certain degree. For this, we considered a blue box (the blue box is a tool, potentially deployed by a broker, that could coordinate traffic such as service consumption or financial streams on the platform. Its exact functionality has to be defined from a requirement analysis) as one of the potential solutions to lower market imperfections as explained in this deliverable.

The business processes defined in the deliverable resulted out of the aggregated data received from interviews conducted. They represent an approximation of the actual processes of each party participating in the interviews. The processes of non-participating parties are not reflected in this investigation.

In the last section of the deliverable the cost and pricing model examination derived of completed surveys, shows potentials from cloud computing for the demand side. It describes the outcomes of applying different pricing models to the different needs of the demand side scenarios.

Many facts such as operational costs that involve third party information (e.g. individual discounts from suppliers) or that could reveal competitive advantages could not be shared. Hence, this information is not taken into account in the following.

Further, many answers given to interview questions cannot be openly shared in the deliverable. Hence, the data received had to be aggregated in order not to hand out individual information. Due to this, it does not reflect each individually and is rather to be interpreted as an average.

The participation in the survey and interviews of parties involved in Helix Nebula is depicted in the following table.

Parties	Participation Survey	Participation Interviews
CERN	Yes	Yes
EMBL	Yes	Yes
ESA	Yes	Yes
DLR	No	Yes
CNR	Yes	No
CNES	Yes	Yes
Logica	No	No
Atos	Yes	Yes
T-Systems	Yes	Yes
EGI	No	Yes
OpenNebula	No	No
Terradue	Yes	Yes
Interoute	Yes	No
SixSq.	Yes	No
CloudSigma	Yes	Yes
The Server Labs	Yes	Yes
Nextworks	Yes	No
IREA	Yes	No
CSA	Yes	Yes
EBI	Yes	Yes

1. Market Imperfections

Market Imperfections occur in markets where required information is not accessible to all parties at essential times. These markets are shaped by a higher degree of uncertainty. Hence, the decisions made in these markets bear higher risks of failure and hinder customers and providers from making rational decisions.

Uncertainty has a major influence on the decision making process of parties involved. It can shape the decisions regarding procurement and offer creation and sets the framework for these processes. Thus, it is essential to investigate the uncertainty and hence, the market imperfections that suppliers and consumers have to cope with, in the context of Helix Nebula, which will be described in the following paragraphs.

1.1 Investment Costs

In initial phases of markets investment costs to build up required resources play a major role and bear a risk of loss if they can't be balanced by profit over time. For the supply side this includes facets of uncertainty such as not knowing what commitment the demand side will be able or willing to make.

Clouds and the characteristic of on demand procurement enabling customers to avoid investment costs for rarely used resources on their side and the maintenance thereof, shift this risk from customers to providers. Beforehand, many providers have been in the outsourcing business renting out resources to customers. The contracts made were long term lowering investment risks for the supply side, since they at least outlasted depreciation times of resources.

Correspondingly, one way to deal with investment risks in on demand scenarios such as cloud computing is offering long term contracts. To make this offer attractive for the demand side, providers give large discounts for long term contracts such as 1 or 3 years reflecting the depreciation time of hardware.

On the demand side it encompasses uncertainties and risks of being locked-into long term contracts by the supply side that might not be competitive for the whole contract duration.

However, on the positive side these long term contracts provide a financial stability by giving certainty on future costs occurring, which can offer a simplifying effect on accounting. Further, the discounts provided lower costs enormously for consumers and hence, are supposed to be competitive for a long term. But since prices in the computer industry drastically shrink over time it still leaves customers in uncertainty.

The longest depreciation time and hence, investment risk for providers originates from the data center. This induced some of the suppliers to outsource their datacenters to other companies. Datacenters can then be shared with other data providers in form of co-location. Providers only have to carry the costs for the purchase and maintenance of their own resources inside the data center, which lowers long term uncertainties on their side.

For the initial investment providers have to estimate required resources trying to balance the trade-off between not having sufficient resources and having unused resources. The characteristic of the on demand business entail the uncertainty of demand. This estimation is complicated by the uncertainty of commitment from consumers.

However on the positive side, acquisition of infrastructure resources has become quite efficient so that lead times could be lowered to a few months or weeks. Resources are shipped in form of small data centers built into containers that can easily be transported. There are many solutions on the market from companies such as IBM, Sun, and Dell.

1.2 Switching Costs

Switching costs occurring between providers can hinder consumers to make rational decisions or moves at all times. On the one hand, they offer an entry barrier for providers outside of Helix Nebula. But on the other hand, they can prevent customers from switching from one Helix Nebula provider to the other, even though their current offer might be more competitive. Hence, customers might not be able to reap the benefits of special offers.

For the supply side, providers capturing most of the demand can find switching costs attractive. However, other providers within the Helix Nebula Consortium would find it more difficult to increase demand with competitive offers.

The costs of switching a provider can't be calculated exactly. The costs to be incorporated into the calculations change between different providers, consumers as well as use cases and their different requirements regarding technology and resources. Further, they change over time based on the effects of the learning curve.

Strictly speaking switching costs can only be correctly determined once the application ran on the corresponding system. Performance is complicated to be estimated for different applications on different resources, since the behavior is not always predictable.

Based on survey results, they are currently estimated to be around 70.000-250.000 euro. Assuming the correctness of the estimation, to induce a consumer to change between providers would require an offer that lowers costs for consumers around 70.000-250.000€ plus a margin that encourages them to accept the effort of switching.

1.3 Transparency

In order to make rational decisions, consumers have to be able to compare the offers provided at the market. However, pricing models and metrics utilized in cloud computing differ drastically from each other, which complicates their comparability. Further, different types of discounts applied increase the complexity of metrics implemented.

Even if the same pricing metric is implemented, the units offered can be machines with different configurations and capacities. Often metrics used are GB/hour or Ghz/hour which on different processing units or storage types deliver different performances or results. Again behavior of applications is hard to predict on different combinations, and can only be determined precisely once tested. Hence, information to make informed decision is not available without a major effort of test runs or benchmarking.

Currently to compare the attractiveness of solutions provided in the cloud computing market consumers have to exactly know the characteristics of their workloads to be processed and stored in the cloud. The different peculiarities have to be entered into each pricing model and calculated for the different discounts and contract terms. The complexity of this can be seen in the calculations conducted for major flagships in the last section of this deliverable.

Further, lock-ins created by contract durations, intensified by relatively high switching costs make the comparability of current offers with potential future offers almost impossible.

2. Procurement Process

Having described major uncertainties in the context of Helix Nebula, the following paragraphs will analyze the relevant business processes of parties involved.

The procurement process describes how the demand side communicates and negotiates with the supply side in order to acquire information regarding their solutions. It further determines the evaluation criteria applied to decide between different offers on the market.

In Helix Nebula there are two types of customers to be considered: major public research organizations and SaaS/PaaS providers, also consuming cloud services in order to develop and offer value adding services.

2.1 The procurement process of public research organizations

Public research organizations in the context of Helix Nebula are a European organization, which means they have to adhere to European public procurement rules. These organizations are usually funded through European states, so called “member states”.

2.1.1 Public procurement rules

In the selection process several constraints have to be considered: the distribution of contracts amongst “member states” should be balanced (“fair return”), further some require a balance between major corporations and small and medium enterprise. The process has to be transparent, impartial and non-discriminatory. The objective is to achieve an optimum use of the (monetary) resources. Hence, a supplier has to offer lowest possible overall cost or best value for money to get the offer.

One of the major objectives to be considered is the balanced distribution of purchases amongst member states. This means that a public organization endeavours to distribute purchases amongst all member states in relation to their share of monetary contributions to the organization. However, this rarely has an impact on a decision. Usually, it comes into

play if the best offer is from a supplier based in a well-balanced member state and the second best offer is from a supplier based in a poorly balanced member state and the price difference is within a maximum limit. In this case, the latter will be requested to align its price to the offer of the top candidate, but since prices mentioned in proposals are very tightly measured they rarely are adaptable.

Some organizations also require encouraging small and medium enterprises to engage in the procurement. For this, some procurement contracts are restricted to offers from small and medium enterprises or other large enterprises are inquired to team up with small and medium enterprises. Some organizations also split procurement contracts between two providers the top candidate and the second candidate in order to distribute the risks of system failure.

2.1.2 Procurement process of public organizations in HN

The procurement processes of public organizations involved in Helix Nebula are relatively similar, since they follow the same public procurement process and rules. They will be described in aggregated form as depicted in figure 1.

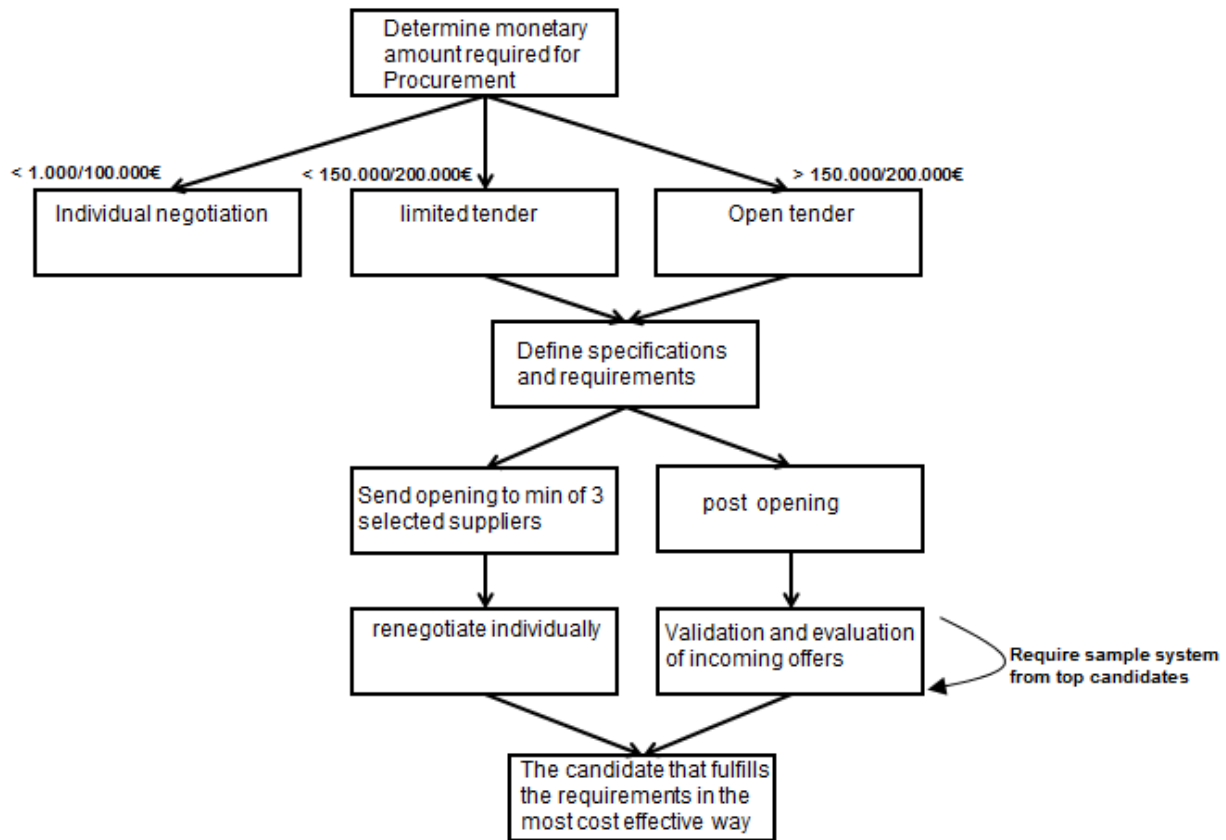


Figure 1: Public Procurement Process

There are three potential forms of procurement depending on the monetary amount required for the purchase of needed resources: negotiation, restricted tender, and open tender. These processes follow different procedures and rules.

2.1.3 Negotiations

Below a certain threshold (around 1.000€ up to 100.000€ depending on the organization) contracts can be negotiated and require a simple purchase order. It is the simplest form of the purchasing process and similar to procurement processes between private companies.

Some organizations directly start with restricted tender and skip the process of negotiations. All process types require a management approval, except insignificant monetary amounts or standard products.

Direct negotiations are possible for cases in which only one provider can be found on the market that delivers a certain kind of solution, which might require prior approval for budgets over a certain threshold. For larger monetary amounts this would require a limited tender. It is also a potential scenario for standard products that don't require many specifications and are sold at relatively low prices.

2.1.4 Restricted tender

Beyond this threshold up to a certain amount (around 150.000€ up to 200.000€ depending on the organization) a restricted tender process is mandatory. Usually companies specify requirements in written form. Subsequently, at least 3 companies have to be selected that are able to deliver. These companies are chosen in different manners. Usually organizations have an established survey process to qualify companies.

Restricted tender usually allows negotiations between the organization and the applicants. It enables clarifications of requirements in order to find the best solution. This process can be of advantage if organizations want very specific or unique solutions, which need clarification and negotiations, since there might be no standardized or established solution.

The organization will request the selected companies to tender usually in written form by sending required documents such as specifications, requirements and evaluation criteria. In certain cases a verbal form might be possible depending on the organization and situation. Written proposals, which are the norm, require more time on the supply side to specify an offer and to evaluate sent proposals. Verbal form can be considered as negotiations with high monetary amount approved as exceptional case e.g. in case of only one available supplier.

The decision criteria are the same as for all types of procurement processes: The supplier that fulfills the requirements in the most cost effective manner or best value for money gets the contract (depending on the organization). Usually weights for evaluation criteria such as quality or price are stated in the requirements of a tender.

This process takes a few months from the need arising to the delivery. Applicants need to be granted a certain amount of time in order to apply correctly in written form. Determining the specifications and evaluating sent applications takes up most of the time.

2.1.5 Open tender

For any larger monetary amounts a thorough open tender process has to be conducted. However, some of the public organizations only conduct negotiations and restricted tender processes in order to have a lean procurement process. Open tender processes require a couple of months to complete the process itself from order to selection of a supplier.

In open tender processes requirements, specifications and evaluation criteria are openly published to allow a broad range of competitors to apply. Open tenders in one field such as IT procurement, are usually only posted once a year or every 18 month by the public research organizations in order to lower effort and costs on procurement. Hence, for all procurements that exceed the budget threshold and fall into the open tender process have to define contract periods with a minimum of one year. The demand for the whole year has to be estimated and determined in the open tender.

Incoming needs for procurement are evaluated by the procurement department. If it requires an open tender process, the initial step is to prepare the specifications of an open tender. The exact requirements such as capacity needed are described thoroughly. Once approved this description is posted.

Suppliers have a certain timeframe (usually 6 to 8 weeks) to apply for the task described. They have to apply based on the exact requirements and specifications posted in the open tender. Providers have to propose a certain price for their offer made, which is one of the major criteria for final evaluation.

Subsequently, the applications will be collected by the public organizations and validated regarding their completeness. Some of the organizations also require a sample system from top candidates to be sent to the organization in order to be tested. Top candidates will be determined in an initial evaluation process.

The next step comprises the evaluation of the applications. In the final evaluation process usually the offer with the best value for money gets the contract. The ratio in which quality and price are weighed for evaluation is usually specified in the open tender. However, for some organizations it's the offer that fulfills all requirements in the most cost effective manner. In this case, any added value beyond what is stated in the requirement won't be taken into account for the evaluation.

This process does not include negotiations in the process. A final negotiation with the selected supplier is possible after the process. It has to be a clear definition of what is required. It can be quite advantageous for standardized or well established products/services, that don't necessitate further clarifications. However, for newer scenarios such as Helix Nebula it can be complex to define what is really required from the suppliers, when consumption is still uncertain. In some cases open tenders stay unanswered due to vagueness.

To conduct restricted and open tender processes – from the need arising to delivery - it takes up a similar amount of time which is usually around 6-9 months. There are different timeframes to be kept such as posting the opening for at least a few weeks (not for restricted tender since opening is sent only to selected few), and giving potential applicants a few weeks in order to write their offer submissions. The writing of specifications and the evaluation of applications in the end require the largest amount of time.

2.2 The Procurement Process of PaaS and SaaS providers

SaaS and PaaS in Helix Nebula are mostly small and medium enterprises. The advantage of small businesses in general is that based on their flexibility and agility, they can follow or lead new trends faster than large enterprises. In many contexts they are often referred to as the drivers of innovation. In Helix Nebula their role is to enrich the platform with innovative value adding services. In order to develop and offer these services, they must be enabled to consume cloud services on demand.

SaaS and PaaS providers in the context of Helix Nebula follow a procurement process which differs from the public procurement process. A characteristic of small companies is that many of the processes such as the procurement process are not overly standardized yet.

Some small companies have no time based restriction to conduct procurement it is conducted whenever a need arises. Usually, procurement is executed at a higher level, due to the flat hierarchy of small companies. There is no team that focuses solely on procurement. Further, there are no complex rules assigned to the procurement process, which reveals this process to be straightforward in comparison to the public procurement process.

Usually procurement follows a process similar to negotiations. Different offers are compared. Then management will decide or approve the procurement to be conducted in few hours or days.

However, PaaS and SaaS providers also have requirements for the procurement of cloud services that have to be taken into consideration for the establishment of potential business models.

There are two different types of procurement situations for SaaS/PaaS providers. Firstly, whenever a customer invokes a SaaS/PaaS service the required IaaS resources will be consumed on demand. They have to be available in constant quality in order for them to be able to deliver to their customers. Further, they have to be able to scale up on demand to satisfy new unplanned customers.

Secondly, SaaS and PaaS providers also require resources for the development of their own value adding services. However, this consumption is less restrictive, since there is not a tight time pressure on this process as it would occur for satisfying customer demand.

3. Offer Creation Process

The definition of offer creation in this context is the creation of an application in verbal or written form. This has to be adapted to the requirements of the procurement process types such as negotiation, restricted tender or open tender.

A verbal form can be required for negotiations and for exceptional cases of restricted tenders. Usually open tenders and restricted tenders require a written proposal; suppliers have to hand in a written offer to be evaluated by the research organization. For open tender this can further require a sample system to be sent in order to run performance tests.

In the context of Helix Nebula two types of providers can be found: IaaS providers which are mostly represented by major corporations or medium sized companies and PaaS/SaaS providers represented by small companies. IaaS providers offer the storage and processing resources, whereas PaaS/SaaS providers offer the value adding services.

3.1 Offer Creation Process of PaaS/SaaS providers

PaaS and SaaS providers maintain lean processes, offer high flexibility and agility, which means they can react quickly to requirements such as public procurement processes.

Most of the PaaS and SaaS providers in the context of Helix Nebula have dealt with tender processes before, since most of them are specialized service providers in the area of science.

Based on previous experience they developed templates for written offers that can quickly be filled out for each new tender process. They further have people with the needed background knowledge to proceed in this area.

The offer creation process is iterative. Written offers are defined by corresponding team members. Subsequently, the written offers will be reviewed by the board. If further improvement is required it will take another round to adapt the proposal and an additional review. Iterations stop once the approval of the board is granted.

The time required to create an offer for the different procurement processes such as negotiation, restricted and open tender are quite distinct. For negotiations it usually takes only 1-2 days to conduct required dialogues. The same applies for exceptional cases of restricted tenders following a negotiation process. Restricted tenders that require a written proposal can necessitate a whole week. To create an offer for an open tender it can require weeks or months.

Sometimes applications require companies to team up in order to be able to deliver the needed capacities. The coordinating company will take over the major part of the application, so that others such as small enterprises only have to put in a few days of effort.

There are usually no preferences between these three procurement processes since they offer a balanced trade-off between length of received commitment and costs.

The criteria to evaluate a decision to engage in an offer creation process are the following and correspond to the criteria of the IaaS providers as described in the next section:

- The size and purchasing power of the potential market has to be sufficient in order to achieve sustainable revenue streams
- The estimated time required until ROI commences, has to be short enough to justify investments
- The availability of required resources, and thus potential investment costs have to be reasonable
- The offer has to fit to the long term strategy of the company and its existing portfolio

If these criteria are not fulfilled the risks to commit to this offer outweigh its potential value.

3.2 Offer Creation Process of IaaS providers

For negotiation based procurements it only requires days or weeks to conduct the required communication. All major companies are proficient in these concepts.

Regarding written applications as required for restricted and open tender, some of the IaaS providers have established processes to cope with tender processes. They follow the deadlines specified in the tender process and hence, require 2-10 weeks for an application.

Teams keen on applying to open/restricted tender have to present requirements and potential profits to the management of the company. The management then decides whether it's a profitable opportunity. Subsequently, they would define a fixed bidding budget. The decision of the management is based on criteria similar to S/PaaS providers:

- Profitability
- Availability of required resources
- Fit to long term strategy and existing portfolio
- Chance of winning in relation to required effort for application

If these criteria are not met the risk associated with the proposal are perceived too high and hence, the proposal will be rejected. In the context of Helix Nebula some of the criteria mentioned are still uncertain. It is not sure how much will be sold via this platform, the demand side still has to state their initial demand commitment, hence revenues can't yet be estimated. Moreover, costs for initial investments are uncertain, not knowing how much will actually be purchased and between whom those costs will be shared or not.

Once the field of business is established and sales people are trained it takes a few weeks to create a standard offer. This comprises the customization of the solution and the onboarding of a new customer. For a certain contract budget (maximum depending on the size of the company) the contract can usually be handed out and signed by sales. For large companies this amount can be up to around 10 million. Above this limit several board approvals are required.

3.3 New offers

To propose a new offer to management business cases have to be conducted (depending on the size of the company more or less detailed), and in some cases it might require to conduct POCs. These have to be handed over to the decision makers. Based on the material provided, management evaluates the profitability of the proposal.

New offers can be extended or changed offers, innovative offers, new pricing models for an existing offer, or any combination thereof. All these have to be approved by higher level management. Usually it requires certain management meetings, which take place on a regular but rare basis in the case of large corporations.

Sometimes it can require multiple rounds of adapting the proposal and business cases until the proposal is either approved or rejected. It has to be investigated if new offers are required in Helix Nebula and hence, if it is needed to go through this processes.

3.4 Competitive offers

To create a competitive value proposition it requires a competitor analysis especially of different solutions considered by target customers. The flagships have different solutions in place some outsourced parts of their computing or storage resources to grid or cloud computing platforms others have private data centers.

There are non-commercial cloud providers focusing on provisioning to science, but current solutions seem to lack convenience. Organizations must contribute and manage their own computing resources or go through a selection process to be allocated CPU time. These providers solely rely on the funding of public institutions or contributions of their members.

There are commercial IaaS providers at the market, that some of the demand side organization already utilize. Offers such as Amazons, provide attractive pricing models. However, they are not customized to science and don't offer cross-disciplinary collaboration or innovative services tailored to the needs of science. Further, it's a question of trust to put confidential data onto these platforms that are open to the public.

In this document we refer to Amazon as an impartial (in the context of Helix Nebula consortium) major player in the cloud business. However, it has to be noted that the market, in which Amazon operates is quite different to the distinct requirements of the big science market. Hence, the assumption that Amazon's business model could be simply be

“copied” is wrong. In this context Amazon is used as an example to get a “feeling” of what are currently implemented standards in this industry.

Important public science organizations such as the flagships of Helix Nebula receive major discounts on resources. However, cloud computing avoids large investment costs and costs for the maintenance of these resources.

If pricing is not a competitive factor based on the competition on the market, then the value proposition has to offer a distinct non-monetary benefit to the demand side. As a promise of on demand solutions, the offer should still reduce costs in comparison to on premise solutions implemented.

4 Budgeting, Accounting and Costing

The final business processes to investigate are budgeting, accounting and costing. They define how and when businesses estimate and calculate their costs and profits. Further, they define how to handle surpluses or deficits occurring when estimations are contrasted with actual numbers.

For the demand side budgeting and accounting regarding procurement are investigated. For the supply side the costing process in regards to provision of the corresponding resources or services is examined.

4.1 Budgeting and Accounting process of public research organizations

For all public organizations involved budgeting is conducted in a 5 year cycle as depicted in figure 2. Then each year a refinement budgeting cycle is executed, to determine the budget for the corresponding year. The budget is not necessarily split up equally between 5 years.

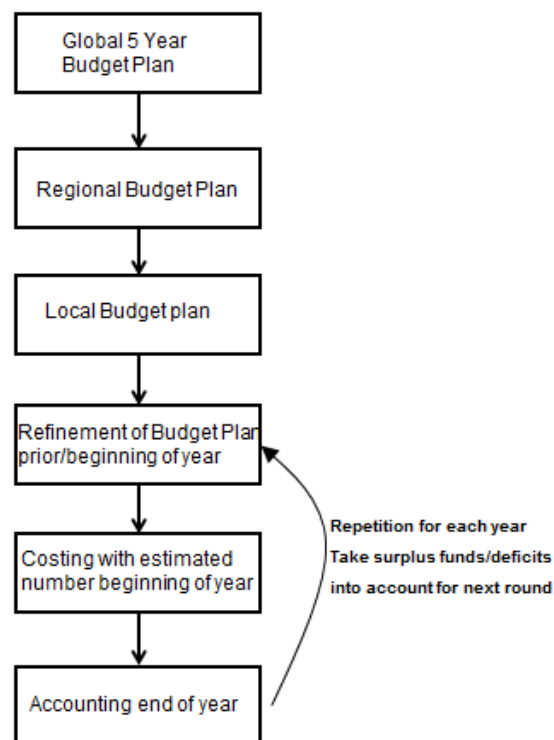


Figure 2: Public Budgeting & Accounting Process

Budgeting is calculated for the global organization, which is then broken down to regional areas. From there the budget is separated onto different departments and teams. This means that the budget might be assigned to teams not before early in the year.

Teams still have the possibility to inquire more budget from the financial department than assigned to them, if an urgent need arises. Surplus funds or deficits are taken into account with the next years budgeting. However, often the budget to be spent on IT procurement is fix, so that cheaper cloud services allow a higher consumption. Further, expensive solutions can't be considered with a restricted budget.

Accounting is usually conducted once a year. Costs for required procurements are estimated a priori when the demand arises for certain IT resources. In the end of the year the planned budget and costs are contrasted with actual expenses.

4.2 Budgeting, Accounting and Costing of PaaS/SaaS providers

Similar to their procurement process SaaS and PaaS providers have no overly standardized process for accounting, budgeting and costing. There is no team solely dedicated to these tasks. Due to their flat hierarchy, these processes are conducted on a high level, which attempts to involve financial and corresponding project experts if available on lower levels.

Accounting, Costing and Budgeting is an iterative process that is calculated on a project basis. Accounting and Costing is done whenever costs in the form of incoming invoices or similar occur. Respectively, Budgeting depends on arising proposals for investment, which have to be decided by management.

Surplus is usually reinvested in future projects. Small companies try to avoid deficits, by closely working with the customers and regularly exchanging feedback on the expected and actual progress of a project.

In the end the project based costs and profits will be summed up into one major calculation at the end of the year.

4.3 Costing of IaaS Providers

Major Corporations have an established costing and accounting process that is done time based, usually monthly or quarterly. Moreover the costing of projects is estimated and then contrasted with actual costs and profits at the end of the projects.

Major corporations usually prefer long term projects, since they deal with higher profits and profit margins than small enterprises.

5 Matching the Procurement and offer creation process

Public research organizations have three forms of procurement processes that providers have to be able to cope with: negotiation, restricted tender, open tender. The form chosen usually depends on the budget required.

5.1 Negotiation

For procurements with monetary amounts below 1.000/100.000€ (depending on the organization) no management approval is required for standardized products. In this budget range the full potential of cloud computing such as on demand, spot market, heavy utilization and other forms of procurement seem feasible.

The procurement decisions in this case can be made by the corresponding line of business and will allow quick and individual decisions. Customers can decide for on demand resources or spot markets in case of urgent needs arising, or they can invest in contracts with shorter periods such as 3-9 months. Long term commitments should be covered by tender processes.

The negotiation process still leaves the risk of investment on the supplier side with unknown long term sales and commitment. However, if the demand can be satisfied with existing resources of the suppliers, it won't induce any risks since it does not necessitate investment in resources. Thus, only large investments that are needed on demand without long term commitment will induce risks on the supply side in this case.

Depending on the strategic purpose of the cloud computing in the context of Helix Nebula such as covering demand peaks or covering frequent long term demand, the budget area of negotiation might quickly be exceeded, which would require a tender process.

5.2 Restricted and open tender

If the yearly demand is expected to exceed a budget of 1.000/100.000€, the demand throughout the whole year has to be estimated and an open or restricted tender has to be conducted. The demand can't be split into smaller monthly contracts for 50.000€ which would allow negotiations.

Further, restricted and open tender publications are usually run once a year and hence require a minimum length of 1 year, which requires a planning for required IT resources months prior to the actual usage. The procurement process in Helix Nebula has to incorporate these time spans to enable a constant delivery to the demand side. Further the tender process incurs costs and effort weeks up to months on both sides.

On the face of it, this excludes cloud purchasing concepts such as spot markets and on demand procurement from the portfolio of options. The procurement would have to be known and defined in advance.

To integrate on demand and spot market into the scenario of a tender process it would require the demand side to commit to an upfront budget for the period of at least one year. This budget can then be booked for on demand usage throughout the year.

In the case of spot markets it would require an agreement of customers to hand over control to the provider to decide when and what kind of resources are utilized throughout the year. There could be a maximum price in the contract stating the acceptable price range for spot instances in order to prevent customers from high expenses.

In the beginning of Helix Nebula defining an open tender might induce difficulties, since many specifications and requirements are still uncertain. However, open tender leaves no room for negotiations, it has to be clearly stated what is necessitated by the organizations. For later phases of Helix Nebula this issue will diminish, once exact needs and requirements crystallize.

Some of the providers have established processes others lack standardized processes to cope with creating open tender applications, which requires a working knowledge of the process steps and rules. The development of a standardized process requires the assignment of employees to this task, the definition of process steps and templates which will induce further costs and effort for these companies.

For the supply side it mostly restricts the offer to whatever the demand side specifies in the open tender. In the case where the most cost effective offer wins the contract the major

variable left for the supply side is the price to be set, which might reduce the market to price competition.

However, there is another form of public procurement contracts referred to as “framework contracts”, which can lower costs and effort on both sides. It is a contract usually set up for a maximum of 4 years amongst several parties.

To conclude a framework contract also requires a tender process, which depending on the monetary amount, follows a restricted or open approach. Any contract concluded later on as part of the framework contract between its parties is based on negotiation. Hence, costs and effort could be drastically lowered for both demand and supply side.

The framework contract specifies the terms of contracting for the 4 years duration. This allows shortening time spans from the need arising to the delivery of goods and might naturally include on demand and spot instances. Moreover, since it is a long term contract, suppliers tend to lower their prices for these kinds of contracts, due to savings in acquisition and administrative costs. The demand side benefit from a long-term supply guarantee without a commitment to procure a certain amount.

5.3 Procurement process of SaaS/PaaS providers

The procurement process of the SaaS and PaaS providers is far less complex than the public procurement process. It facilitates any form of cloud computing: spot markets, on demand, heavy utilization etc. Basically, any type of contract and terms can be supported.

However the demand scenario of the SaaS/PaaS providers also has implications on the potential market to be created in Helix Nebula in regards to financial streams.

SaaS and PaaS providers usually consume cloud services, when their own services are consumed by customers (except for resources required for service development). This would require revenue streams between IaaS, SaaS and PaaS providers. These could easily be avoided through cost dynamic processes.

There are approaches in the cloud computing market such as Amazon “devpay” that minimize the amount of revenue streams throughout the platform. Solely the revenue stream of the customers is directed onto the cloud platform. This stream is then split up

between SaaS, PaaS and IaaS providers. Hence, there are minimal revenue streams required in between the different providers.

This solution would require a financial broker to coordinate payments and could be considered an extension of the blue box. This reduces costs in advance and hence risks of investing in unused resources for the PaaS and SaaS providers. Any expenses will be covered by the payment of the end customer.

In the context of the blue box there are potential new roles to be investigated. The broker might also take over the application process for open tenders. This would require defining how demand could be split between the providers, which could prevent the market from price competition.

The blue box could provide a common interface to customers, which conceals the differences between providers' actual interfaces. Moreover, it might diminish switching costs, potentially unify payments and establish transparency through a supplementary marketplace. In order to achieve this, there are many requirements to be implemented such as a mandatory upfront benchmarking process that will evaluate offers of the supply side and capacities needed during runtime from the different demand side applications. It would necessitate a requirement analysis to figure out details of the functionality.

6 Matching budgeting/accounting and costing

Cloud Computing requires new kind of contracts and payments such as on demand payment or long term contracts with an agreed upon budget that is deducted on demand. Hence, it might have an influence on financial business processes or might be restricted by them.

Public organizations follow a budget cycle of 5 years with a refinement cycle each year. Any budget beyond this timeframe is considered uncertain for most of the organizations. This means that contract periods larger than 5 years are generally infeasible. Depending on when in the 5 year cycle contracts are signed the maximum length has to be shortened. For example contracts signed in the second year of the cycle can be maximum 3-4 years.

For public research organizations shorter contracts enable them to reassess competition and utilize their budget in the optimal way. The reassessment needed in form of a tender process taking in account the effort and costs often seems more beneficial than to commit to long term contracts.

However, there were no preferences found from a financial business process perspective to the length of a contract. Typically offered contract periods are 3, 6 or 9 months and 1 or 3 years. For most customers it is a simple trade-off between flexibility and costs, which has to be decided based on its strategic purpose.

Small companies calculate financial processes project based whereas large companies calculate them time based which can be monthly or quarterly. Each of them conducts a major yearly calculation, which fits with the public organizations.

Contracts can comprise a fixed price or variable prices as required for cloud computing concepts such as spot markets. With variable pricing it is important to implement transparency for financial business processes e.g. through frequent payments and invoices or through other monitoring tools. In long term contracts with variable pricing, a ceiling price should be determined in the case of spot markets etc. in order to avoid unexpected costs.

7 Theoretical & practical difficulties for decision making in cloud markets

It is naïve to assume that the market imperfections in cloud markets for IaaS are due solely to parties intentionally hiding information from one another at decision time. The whole value proposition of offering long term investment goods such as ICT infrastructure is a mixture of the underlying functionality and the assignment of commercial risk to the parties under various degrees of uncertainty. In addition the order in which the necessary information for rational decisions can be discovered by the parties make this a non-trivial game. In order to illustrate the factors and to put some quantitative boundaries on the relative magnitude of the various factors on cost and pricing, we'll discuss a non-trivial example from the current flagships. What matters for discussion are not the absolute numbers (as these are fictional for reasons that will become clear below, despite the usage of real prices quoted by IaaS suppliers) but rather to illustrate the relative influences.

7.1 A Sample Cloud Application

Our model application is roughly patterned after the EMBL genome sequencing application in characteristics. Like many other demanding applications, rather than directly consuming individual IaaS resources it assembles a structured, multi-tiered complex processing platform and pipeline of steps to achieve the complete process. In the case of our model application, there are two major phases, a highly parallelizable phase that can be fanned out to independent worker nodes in a batch fashion followed by a parallelizable, but highly coupled processing phase that assembles the partial preprocessed results into the final assembly on a large SMP node. In order to coordinate the work and distribute the intermediate and final results across the assembled cluster, a set of machines is assembled to implement a distributed cluster file system and a batch control for the worker nodes in the first phase. Overall there are 19 processing steps in the workflow, some parallel, some sequential. A representative resource assignment and consumption pattern is depicted in Table 1:

Stage	Name	RAM/ node	Cores/Node	Number of Nodes	Number of jobs	In Parallel	CPU time! h	Approx. real time! h	Instance Hours
		Suggested	suggested	suggested	required		required	required	
1	preprocessing	32	8	25	23	Yes	1	1	23
2	index raw	32	8	25	23	Yes	2	2	46
3	merge raw	32	8	25	11	Yes	2	2	22
4	merge raw1	32	8	25	6	Yes	3	3	18
5	merge raw3	32	8	25	3	Yes	3	3	9
6	merge raw4	32	8	25	2	Yes	5	5	10
7	merge raw 5	32	8	25	1	No	10	8	8
8	error correction	32	8	25	23	Yes	12	2	46
9	indexing correcting	32	8	25	23	Yes	3	4	92
10	merge corrected	32	8	25	11	Yes	2	2	22
11	merge corrected1	32	8	25	6	Yes	3	3	18
12	merge corrected	32	8	25	3	Yes	6	5	15
13	merge corrected	32	8	25	1	No	12	8	8
14	merge corrected	32	8	25	1	No	22	13	13
									350
15	filter	32	8	1	1	No	48	15	15
16	fm-merge	32	8	1	1	No	82	30	30
17	index	32	8	1	1	No	7	5	5
18	overlap	32	8	1	1	No	76	20	20
19	assemble	32	8	1	1	No	170	170	170
									240

Table 1: Job structure and resource consumption

The structure of the job and the resources likely were assigned initially due to the properties and available resources of in-house systems on which the application was developed, as is exemplified by the strict homogeneity of the node sizes and CPU allocations listed, while the actual resource consumption in some stages differs radically. In order to even to begin to calculate the price of running such a pipeline on a given cloud

infrastructure, one needs to translate the raw resources consumed (CPU hours, etc.) into the lifetimes of resources being billed for (instance hours), as all current IaaS providers do charge for the time a given (virtual) resource is being rented out, regardless of the actual load incurred.

7.2 Calculating Prices

So the start of our pricing exercise translates from resources to occupancy and then sums this up over the major lifetime phases of the infrastructure. Sub phases 1-14 require the parallel batch processing system, here scaled to 25 worker nodes and Sub phases 15-19 are executed on a single large machine. The underlying cluster file system (Gluster) needs to be available through both major phases, leading to the following structure of billable items:

Billable Items
Create 4 Server Glusterfs
GlusterFS lifetime (249 hours)
3000 Gbyte permanent storage
Upload 400 Gbyte Genomics data
Create 50 node Compute Cluster
Use 350 Instance Hours
Switch to single host
Use 240 Instance Hours
Download 100 Gbyte result

The next step in order to figure out the price is to match this to billable items in a cloud provider. For our example we map this against Amazon Web Services (as an impartial - in the context of Helix Nebula Consortium - market leading offer), selecting “Standard Medium” EC2 instances for the Gluster FS file servers and “2nd Gen M3 Double Extra Large” for all other node types. The storage is assumed to be highly request intensive; hence we select the Elastic Block Storage service. For ease of calculation, we’ll ignore the per-request fees as we got no numbers from EMBL to guide us here, and use the networking and other costs for the cheapest Amazon zone (i.e. data center location) available, US East Coast.

Mapping that against the Amazon pricing tables as of October 2012 (they have since changed) in a pure “on demand” fashion, this comes out to

Amazon On Demand	Per Unit €	Total €
Create 4 Server Glusterfs		
GlusterFS lifetime (hours)	0,10	101,69
3000 Gbyte permanent storage	0,08	81,49
Upload 400 Gbyte Genomics data	0,00	0,00
Create 50 node Compute Cluster		
Use 350 Instance Hours	0,91	318,87
Switch to single host		
Use 240 Instance Hours	0,91	218,66
Download 100 Gbyte result	0,09	9,42
Total Cost		730,13

While one may be tempted to take this as “the price”, there is a bewildering array of choices to be made, some of which require coding changes, some of which are pure configuration and business choices. Rather than exhaustively enumerating them, let us illustrate a few choices. The 50 node cluster in the first major phase is a batch processing system that can deal with worker nodes popping in and out of existence easily. Therefore one can take advantage of one of the “shift the risk from provider back to the consumer” pricing options that providers offer: spot markets.

There are different flavors of this, the Amazon system basically offers an auction, where machines currently idling are auctioned to bidders from highest to lowest bidder until the pool is exhausted. When higher bidders enter the fray, machines assigned to lower bidders are revoked without warning, requiring code to deal with frequent and violent death of worker nodes.

The business logic on the provider side for these auction systems is straightforward: They have to have enough capacity to fulfill all their SLAs towards their customers, hence need to

overprovide to a certain degree. In Amazon's case, they offer three distinct levels of availability of machines to their customers:

1. Reserved instances, which they guarantee can start whenever requested and remain available throughout
2. On Demand instances, which they do not guarantee can start, but remain available once initiated (absent errors)
3. Spot instances, which they do not guarantee can start and to be killed whenever competing bid prices exceed the owners bid

This allows Amazon to implement load shedding to accommodate their customers in SLA levels 1 and 2 by adjusting the threshold bid price against the known outstanding bids and reallocating resources accordingly. Other auction systems exist, with different logic for the demand side but a similar rationale on the provider side.

In our example utilizing spot instances for the worker nodes of the initial phase and sticking to on-demand instances for the rest reduces the overall price for our job to 539 €, a 27% percent reduction, assuming a (empirically plausible) average spot instance price of 40% of the full on demand price and negligible amount of wasted effort due to randomly killed worker nodes (which depends on the granularity of individual batch jobs).

Further optimization can be attempted and realized by noticing that there are very long running pieces in this workflow that not only benefit from a high availability of the resources but also force a high duty cycle for some of the instances over longer periods of time, where some of the uncertainty about continued resource utilization can be removed by the demand side by reserving instances ahead of time.

Amazon offers two types of reservation, one that removes part of the uncertainty for the provider (reserve for 1 year, pay some up front, pay a reduced rate for any actual resources consumed) and another that completely shifts the risk to the demand side (reserve for 3 years, pay a reduced rate for all resources consumable, regardless of actual use).

What is most illuminating about the tradeoffs faced by providers is the pricing structure that providers offer for these different SLA levels, as it gives direct inside into the actual cost of uncertainty incurred by the provider, assuming that the price curves mostly reflect

costs plus some margin. The providers in Helix Nebula were unable to share their actual cost structures with us, partially for reasons of confidentiality but also because it would be difficult to tease apart some of the cost elements sufficiently. We therefore resort to publicly available price information from providers inside and outside of Helix Nebula as a proxy for costs.

Averaging such a discount calculation across all instance types Amazon offers and then calculating the price for a fully utilized machine gives a 37% reduction for the partial shifting of risk for one year up to a whopping 66% price reduction for the all-the-risk-removed option over 3 years. Similar ranges can be found in the pricing structures of other providers such as CloudSigma (CloudSigma is “randomly” picked as a representative of the Helix Nebula Consortium in order to indicate that the consortium offers similar value) looking at either risk reduction through multiyear contracts or at volume discounts given to resellers.

For pricing the cost of running a set of our sample workloads on Amazon over the course of a year allows a further reduction of the cost-to-run for a single job to 430-445€, depending on overall number of jobs and arrival rate.

Thus far, we’ve encountered a price variability of 1 to 1.7 on a single variation of provider infrastructure, without even starting to take into account variability due to different split of instance types or the substantial variability encountered on the same instance type over time or underlying physical infrastructure variation¹, which introduces a similar variation of actual resource consumption and hence price.

7.3 Consequences for rational demand side decisions

As our example calculation demonstrates, there is a considerable difficulty in calculating the price/performance ratio of a given provider offering. Differences in properties of the actual virtualized infrastructure of different providers are likely to dwarf marginal

¹ Up to 60% variability was found by one study (“Exploiting Hardware Heterogeneity within the same instance type of Amazon EC2”, Zhounghong Ou et.al, Hotcloud '12, Boston, MA)

differences in prices for superficially similar resources. Buyers of IaaS services cannot reliably compare the relative cost effectiveness of different offers without actually trying out representative benchmarks of their workloads on a given provider.

As an immediate consequence, this means that almost the entire switching cost for moving an application to a new provider has to be incurred by the demand side before reliable comparisons can be made. This underlines the importance of a true federated API that would allow to run unmodified applications in a benchmark mode to eliminate this obstacle. While this would still require applications to be written eventually to the federated API, it would probably be sufficient for benchmark purposes to also offer an API that transparently allows deploying applications written strictly for in-house data centers in a cloud. For an example of what that capability looks like technically, see the offering by Verizons CloudSwitch (<http://www.cloudswitch.com>). With solutions like Cloudswitch, it becomes possible to just run an unmodified in-house application (same network configuration, etc.) on the provider infrastructure in order to assess the performance it would have in that environment.

Other factors such as variability in production as well as the natural variability of workloads with varying data inputs blur the picture somewhat more, but could be addressed by provider side measurement infrastructure that gave feedback on how to choose better matches within the given providers infrastructure.

The other significant factor throughput oriented, less deadline critical applications (as is often the case in eScience) is the trade-off between complete flexibility and reserved capacity, i.e. utilization risk shifting. The massive discount range offered (almost 1:3 on Amazon, 1:2 on CloudSigma, comparable on other providers) makes long-range workload planning beyond the normal scheduling capabilities of typical batch processing systems used in eScience an essential component of cost-effective use of the cloud.

7.4 Consequences for providers

Providers need to make a conscious decision on how to balance their systems (both actual landscape and pricing models & schedules) between different user groups. If they want to be cost-attractive to eScience users, it could be helpful to integrate their spot-market

systems with popular middleware suites on the eScience application side that makes appropriate use of the intended load-balancing and shedding.

For pure throughput, deadline insensitive workloads the lower end of the achievable margin is bracketed by the costs of running their own datacentre for large eScience customers and the alternative of “somewhat on demand” additional capacity offered by companies providing modular datacentres (i.e. datacentres in a container), that can offer exclusive and relatively timely access to additional resources when there is a longer term commitment to resource consumption from the demand side.

It is not clear to us yet whether there is a (sufficiently large) gap between those two extremes that would allow the envisioned Helix Nebula ecosystem to remain viable from just the large eScience users currently participating. More diverse user groups, that are either naturally distributed (hence no logical place to put resources inside any single one of them) or that are too small by entity (and hence cannot afford their own resources the way outlined above) would give an advantage to cloud based solutions.

If there is to be federated ecosystem of providers to be established, efforts to standardize some of the performance characterization of complex landscapes for comparison efforts as well as mechanisms to be able to tailor resource types more smoothly to more closely match offerings from another provider (such as the incrementally infinitely variable instance types offered by CloudSigma) would be helpful.

8 Conclusion: Next steps and obstacles to overcome

For some of the obstacles found, process or technical measures can be taken to overcome them, others are based on intrinsic difficulties in prediction and can't be changed or have their roots in competitive considerations which participants are not willing to change.

On a basic level there are two opposing needs from the two sides involved. The demand side requires flexibility and scalability, which the supply side can satisfy for "smaller" demands. However, for large amounts the supply side would need a long term commitment in order to reduce investment risk, which creates a feeling of being "locked-in" on the demand side. However, these are obstacles that should still allow elaborating a compromise that is satisfactory for both sides. In the previous chapter we found that prices are reduced up to 2/3 in order to shift the risk from suppliers to demand side.

On a higher level, parties involved in Helix Nebula have to face switching costs. This might hinder customers to switch to the most competitive offer and prevent suppliers from the demand they would be entitled to. This obstacle requires more effort in form of the implementation of the previously mentioned upfront benchmarking through the blue box. Further, it requires the agreement of participants would have to accept the upfront benchmarking effort and its costs as minor onboarding costs in order to facilitate fairness of competition and rational decision making on the platform. A common API is supposed to avoid any code changes to be necessary in order to be enabled to switch to another provider. However, there is more to be considered when talking about switching costs, such as the movement of data from one supplier to the next, which can't be fixed with a common API.

If performance can be estimated reasonably accurately through benchmarking, it will also facilitate an approximate comparison of pricing. The utilization of different pricing metrics, discounts or resource types offered can't be avoided, due to different business models chosen by the providers. Different workloads fit better to different units (e.g. offered package sizes of small / medium / large) in order to not pay for unused resources and can make a cheaper metric less profitable. Hence, it will still require some calculation effort to compare different providers and won't establish full transparency without some initial investment from the demand side.

For price calculations as seen in the previous chapter it will always require exact knowledge of the applications and workloads in order to define optimum usage and combination of the different cloud computing models offered. For example parts of the application that require high availability reserved instances can lower costs and for parts that can be stopped and started arbitrarily that are less time critical spot instances can reduce prices drastically. However, it might be simplified through the blue box's benchmarking.

The tender process, which induces costs and efforts on both sides, could at least be lowered to a single upfront tender process through the application of framework contracts. Then following contracts can be based on negotiation. Either way suppliers will have to establish suitable offer creation processes if not implemented so far. As described there are enough possibilities to incorporate any kind of cloud computing models into the different process types. The only major coordination effort is to align the initial time spans required for tender processes and match it with the offer creation process on the supply side. How the specifications and requirements of the actual tender will look like is basically up to the demand side.

The criterion implemented should be "bestvalue for money" rather than "most cost effective offer that fulfills requirements" since it focuses competition on price and value. A pure price competition usually reduces "healthy" competition to a few or one survivors (mainly larger corporations), which in the end might rise prices to a higher level than before. The objective should be to keep competition also focused on value in order to foster the emergence of optimum value solutions.

In a last step the requirements for the blue box have to be analyzed and defined in order to establish transparency and diminish switching costs. Further, an integration of a financial broker - with the objective to avoid additional revenue streams between different types of suppliers - has to be investigated. This will induce major costs and effort initially, but drastically lower market imperfections in the long term.

In the table below potential steps to be taken with targeted outcomes are defined with assigned responsibilities. The obstacles that are encountered in each step are listed in order to raise awareness of their existence to facilitate overcoming them.

Actions	Who	Obstacles	Outcome
Agreement on initial capacity and budget	Demand and Supply Side	<ul style="list-style-type: none"> • Can investment risks be lowered? • Trade-off between costs and commitment (opposing needs) 	Determines the budget range -> type of procurement process
A: negotiation	Demand and Supply Side	Risk if major investment required without long term commitment	Determine contract type satisfactory for both sides
B: open/restricted tender specifications (legal background)	Demand Side	<ul style="list-style-type: none"> • Is it possible to privilege HN members? • Can price competition be avoided? • Establishment of offer creation process • Effort and costs on both sides 	Open/restricted tender specifications
Blue box and broker roles	SAP and others	<ul style="list-style-type: none"> • Transparency • Switching cost • Complex financial streams • Tendering 	Lower market imperfections