

Multi-objective Service Oriented Network Provisioning in Ultra-scale Systems

Dragi Kimovski*, Sashko Ristov, Roland Mathá, and Radu Prodan

Distributed and Parallel Systems, Institute of Informatics,
University of Innsbruck, Austria.

*dragi@dps.uibk.ac.at

Abstract. The paradigm of ultra-scale computing has been recently pushed forward by the current trends in distributed computing. This novel architecture concept is focused towards a federation of multiple geographically distributed heterogeneous systems under a single system image, thus allowing efficient deployment and management of very complex architectures applications. To enable sustainable ultra-scale computing, there are multiple major challenges, which have to be tackled, such as, improved data distribution, increased systems scalability, enhanced fault tolerance, elastic resource management, low latency communication and etc. Regrettably, the current research initiatives in the area of ultra-scale computing are in a very early stage of research and are predominantly concentrated on the management of the computational and storage resources, thus leaving the networking aspects unexplored. In this paper we introduce a promising new paradigm for cluster-based Multi-objective service-oriented network provisioning for ultra-scale computing environments by unifying the management of the local communication resources and the external inter-domain network services under a single point of view. We explore the potentials for representing the local network resources within a single distributed or parallel system and combine them together with the external communication services.

Keywords: Inter-domain network provisioning, Multi-objective optimization, Machine learning

1 Introduction

In order to successfully handle the growth of data volume and maintain computational performance on large scales, it is essential for the emerging hardware and software systems to be re-evaluated to handle the foreseen challenges introduced by the large scale distributed environments. The emerging field of ultra-scale computing aims at tackling these rather ambitious challenges by paving the road for the development of highly distributed architectures, spawning over multiple administrative domains [1]. The paradigm of ultra-scale computing has been recently pushed forward by the current trends in distributed computing, and to some extent in high-performance computing (HPC), focused towards a

federation of multiple geographically distributed heterogeneous systems under a single system image, thus allowing efficient deployment and management of very complex architectures [2].

Unfortunately, supporting the evolution of the ultra-scale systems requires immense research activities focused towards development of domain-specific tools and architectures for enabling robust computing solutions through multi-domain cooperative approaches. To enable sustainable ultra-scale computing, there are multiple major challenges, which have to be tackled, such as, improved data distribution and data locality, increased systems scalability, enhanced fault tolerance and availability, elastic resource management, low latency inter-domain communication and etc.

Regrettably, all promising research initiatives in the area of distributed ultra-scale computing are in a very early stage of research and are predominantly concentrated on the inter-domain management of the computational and storage resources, thus leaving the networking aspects unexplored. As a result, multiple challenges in terms of description, allocation, operation and management of network services and resources, especially in heterogeneous distributed and parallel environments, have been neglected and remained unexplored till today.

In this paper we introduce a promising new paradigm for Multi-objective service-oriented network provisioning for ultra-scale computing environments by unifying the management of the local communication resources and the external inter-domain network services under a single point of view. We explore the potentials for representing the local network resources within a single distributed or parallel system and combine them together with the external communication services. The composition of the local resources and external services will result in the creation of an inter-domain communication environment called *communication super-service*. The introduction of such a paradigm will enable transparent deployment of highly adaptive service-based virtual networks, spawning across various domains and system architectures.

In order to enable efficient and low latency provisioning of network super-services we utilize algorithms and techniques from the field of multi-criteria optimization and clustering. More concretely, we have exploited current clustering techniques to initially divide the available network services and resources based on the user's preferences. Afterwards, multi-objective optimization and non-domination sorting algorithms, together with novel decision making strategy, are used to provide a set of "optimal" trade-off combination of network services and network resources. In what follows a detailed description and comprehensive evaluation of the proposed communication super-service provisioning is provided.

2 Related work

Recently, promising new research initiatives have been started in the European research community, focused towards solving the issues that prevent efficient management of the network resources in an inter-domain environment. One

of these initiatives is the SSICLOPS project, which aims at developing novel techniques for management of software-defined networks within federated Cloud infrastructures [3]. Furthermore, the BEACON research project [4] targets a virtualization layer on top of heterogeneous underlying physical networks, computing and storage infrastructures, providing automated federation of applications across different Clouds and data centers. Significant research progress has also been reported in the literature. The authors in [6] introduced a novel approach for designing Cloud systems, developed around the notion of robust virtual network infrastructure capable of specifying complex interconnection topologies. A promising architectural solution for Cloud service provisioning was proposed in [7] relying on service-based IP network virtualization. Within this research initiative, various management schemes have been designed and implemented, such as novel resource description and abstraction mechanisms, complex virtual network request methods, and a resource broker mechanism called “Marketplace”. The work in [8] proposes an OpenFlow service based network virtualization framework for supporting Cloud infrastructures and presents promising new network abstraction methods for virtualization of the physical infrastructure. Furthermore, the innovative virtual network provisioning approach in [9] comprises an elasticity-aware abstraction model and virtual network service provisioning method that allows for elastic network scaling in relation to the communication load in the data center or Cloud infrastructure. Lastly, the authors in [10] proposed an adaptive virtual resource provisioning method capable of adapting in response to the demand for virtual network service requests, extended to support fault-tolerance embedding and provisioning algorithm.

In spite of these important advances, the management and utilization of the network resources are still in an early stage of research. Currently, even within a single Cloud environment or multi-cluster infrastructure, the guarantees on the Quality-of-Service (QoS) on the communication infrastructure are limited. For example, in the current Cloud architectures, only minimal bandwidth is assured per Virtual Machine (VM), without considering the communication latency [11]. Moreover, the current research advances have been only focused towards overcoming the barriers that limit the efficient utilization of the interconnection resources in Cloud environment, which neglect the requirements of the high-performance community for low latency communication between heterogeneous distributed and parallel systems.

3 Background

3.1 Multi-objective optimization

In this work we utilize multiple concepts from the area of multi-objective optimization to enable efficient network services provisioning in ultra-scale systems. In general, optimization is a process of identifying one or multiple solutions, which correspond to the extreme values of two or more objective functions within given constraints set. In the cases in which the optimization task utilizes only a single objective function it results in a single optimal solution. Moreover, the

optimization can also consider multiple conflicting objectives simultaneously. In those circumstances, the process will usually result in a set of optimal trade-off solutions, so-called Pareto solutions. The task of finding the optimal set of Pareto solutions is known in the literature as a multi-objective optimization [12].

The multi-objective optimization problem usually involves two or more objective functions which have to be either minimized or maximized. The problem of optimization can be formulated as: $\min/\max(f_1(Y), f_2(Y), \dots, f_n(Y))$, where $n \geq 2$ is the number of objectives functions f that we want to minimize or maximize, while $Y = (y_1, y_2, \dots, y_k)$ is a region enclosing the set of feasible decision vectors.

Even though the above formulation of the multi-objective optimization is without any constraints, this is hardly the case when real-life optimization problems are being considered. The real-life problems are typically constrained by some bounds, which divide the search space into two regions, namely feasible and infeasible region.

3.2 Clustering

In the Big Data era the vital tool for dealing with large data-sets is the concept of classification or grouping of data objects into a set of categories or clusters. The classification of the objects is conducted based on the similarity or dissimilarity of multiple features that describe them. Those differences are usually generalized as proximity in accordance to certain standards or rules. Essentially, the classification methods can be divided into two categories, namely supervised and unsupervised [5]. In supervised classification, the features' mapping from a set of input data vectors is classified to a finite set of discrete labeled classes and it is modeled in terms of some mathematical function. On the other hand, in unsupervised classification, called clustering, no labeled data-sets are available. The aim of the clustering is to separate a finite unlabeled data-sets into a finite and discrete set of clusters. For the purposes of our work, we utilize distance and similarity based clustering algorithms, such as k-means, which allow for low-latency coarse-grained clustering [16].

4 System architecture

To tackle the issues that limit the possibilities for transparent inter-domain communication we present a use-case scenario for the proposed multi-objective provisioning environment. Furthermore, based on the use-case scenario the top-level view architecture of the proposed system is provided.

The use-case of the proposed environment can be identified in the field of distributed ultra-scale computing. More concretely, the future large-scale systems have been foreseen as a heterogeneous fusion of the tightly coupled HPC systems and loosely coupled Cloud infrastructures, interconnected by external network infrastructures provided on the network-as-a-service basis. The heterogeneity of

such platforms can pose many challenges for efficient and low latency communication between processes located in different domains and systems. For example, let us assume that we have distributed application located at two distant geographical locations, where both computing systems are of different architecture. The current state-of-the-art technology will only allow for a high-level protocol, such as TCP, to be used over the shared Internet network in order to provide a communication channel between the application components. This may induce high latency and low communication bandwidth. Contrary, the proposed architecture aims at utilizing the high bandwidth communication systems, based on the network-as-a-service paradigm, and combine them with the local network resources to achieve better communication performance.

In relation to the use-cases, we envision the proposed system, depicted in Figure 1, as a full environment capable of providing a universal backbone for super-service network provisioning. Essentially, the environment allows for the network service providers, together with local system administrators, to register the offered services, including the functional parameters, to a specific database. After the proper description of the available services, they are clustered in multiple classes based on the functional parameters, such as latency and bandwidth. The given set of services is then provided to the multi-objective service composition module, which explores for an optimized combination of services and resources that meet the performance requirements. Subsequently, this mapping, or more concretely combination of services is provisioned to the application that required it.

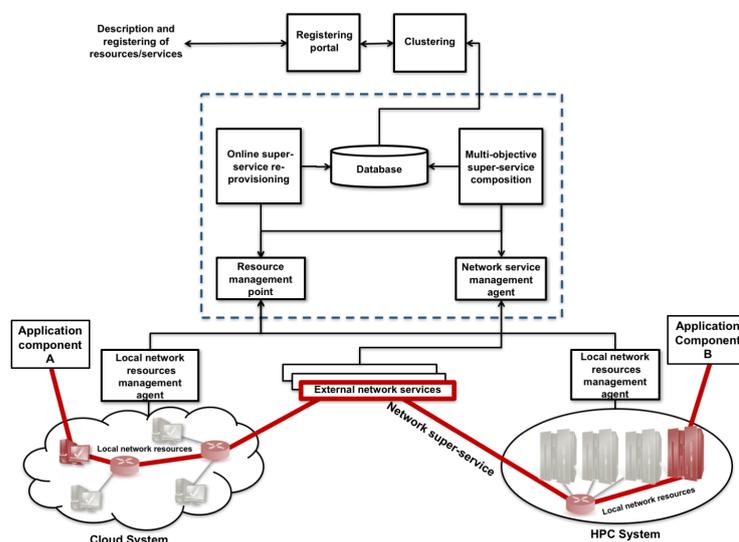


Fig. 1. Top level view of the multi-objective super-service provisioning architecture

Furthermore, based on the previous usage data, a separate re-provisioning module continuously gazes for degraded communication performance. Based on the provided data, this module utilizes the same multi-objective core algorithm to re-provision the given super-service if some faults are imminent or there are many QoS violations.

4.1 Multi-objective super-service provisioning

The system design of the proposed super-service provisioning environment is modular in nature, encapsulating variety of different components which interact by exchanging structured information on the available network services and resources. Each component in the system provides specific functions, which are essential for the normal functioning of the provided network resources. The core of the proposed environment is based upon multi-objective optimization module, capable of composing various network resources and service into a compound network super-service.

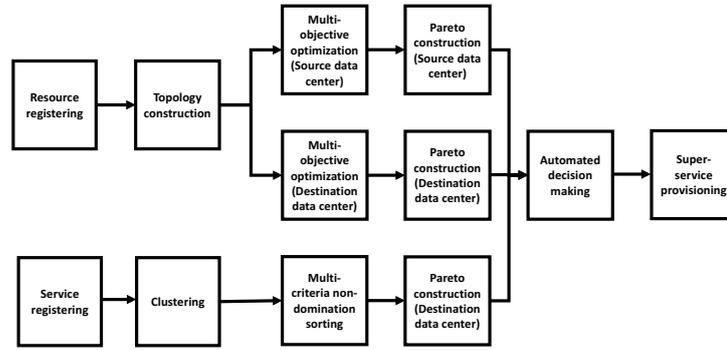


Fig. 2. Multi-objective super-service provisioning model

The process of network service and resource composition, depicted on Figure 2, is divided into two distinctive stages: resource provisioning (see Section 4.2) and service provisioning (see Section 4.3). The two stages are conducted in parallel and separate Pareto fronts are constructed for the network services and source/destination network resources. The automated decision making considers the independent Pareto fronts, together with the user's a-priory preferences, to complete the process of super-service provisioning.

4.2 Network resource provisioning

During this process the registered network resources are fetched from the database and a set of starting "candidate" routes is constructed, both for the source and destination data-centers or clusters. This information is then used as an input for

two separate multi-objective optimization processes based on the NSGA-II algorithm [15]. The NSGA-II is an evolutionary optimization algorithm, therefore it requires proper representation of the routes that will be optimized as individuals in the algorithm's population. For our purposes we represent the routes as a directed vectors containing all channels, switches and router through which the data will be send within the source or destination system.

Each of these processes is focused on finding a set of optimal "trade-off" routes in relation to the communication bandwidth and latency within the local source and destination systems respectively. The optimization processes result in two separate Pareto fronts, which are later used during the decision making (see Section 4.4). Every solution in the Pareto represents a possible internal route through which a virtual channel can be established within the source and destination large scale computational centers.

4.3 Network service clustering and provisioning

In parallel to the previous described stage, the registered network services are clustered each time a new service has been added to the database by utilizing k-means clustering technique [16]. This technique has been selected primary because it requires low computational resources for small number of clusters. In our case, we create three different clusters of services in relation with the following objectives: communication latency and network bandwidth. This allows us to initially prioritize the services in relation to the given objectives, thus reducing the execution time of the computational costly non-domination sorting algorithms.

In relation to the clustering objectives, we divide the registered network service in the following categories:

- **High-bandwidth/low-latency:** this cluster encompasses the services which provide the best communication latency and bandwidth rates. Usually these network services induce higher financial costs.
- **Medium-to-low-bandwidth/low-latency:** this cluster provides low latency, similar to the previous one, however with reduced bandwidth. The network services belonging to this cluster are usually more cost effective, compared to the first cluster.
- **Low-bandwidth/low-latency:** the last cluster encompasses the network services that are not capable of providing sufficiently high bandwidth or low latency. The services belonging to this cluster are discarded and not used during the following process of non-domination sorting.

Afterwards, the feasible clusters are sorted based on non-domination multiple-criteria sorting algorithm, resulting in a separate Pareto fronts for each of the clusters. In the case of the non-domination sorting we consider three objectives: communication latency, network bandwidth and financial costs. To be more concrete, the utilization of the clustering techniques allows for the services to be classified in a coarse-grained manner, while the non-domination sorting enables fine-grained selection of the most optimal network services.

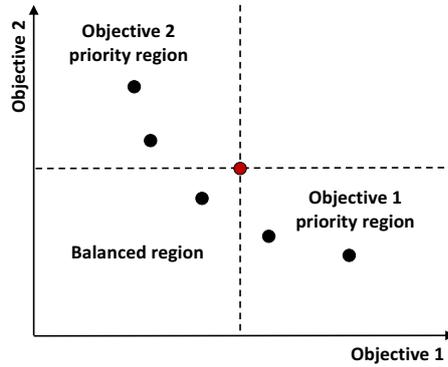


Fig. 3. Automated decision making

The process of service clustering and sorting is only conducted when new service has been added or the functional parameters of some service have been changed. The constructed Pareto front during this process is later utilized for the automated decision making (see Section 4.4).

4.4 Automated decision making

The basic prerequisite of the multi-objective super-service provisioning is the implementation of Automated Decision Making (ADM). Due to the basic requirements for low provisioning latency it is essential to enable efficient decision making techniques. To this end, we have implemented a simple and computationally efficient a-priori ADM procedure, which takes into consideration the user's preferences. During the process of automated decision making, the Pareto fronts from the source and destination routes and networks services are considered independently. Consequently, from the three Pareto fronts separate solutions are selected and are then joined together to provide the final solution on how to provision the super-service. The proposed ADM process assumes that all solutions in the Pareto front belong to the same cluster. Based on this assumption, we find the centroid of the Pareto front. Afterwards, we map the centroid to the objectives' axis. This allow us to divide the objective space and the Pareto solutions into distinctive regions. More concretely if a solution is located within the parallels of the centroid it is considered that it belongs in the "balanced" region. The solutions which are within the centroid's parallel in one objective dimension, but not in the other are consider to belong to the "objective's priority" region. For illustration, Figure 3 shows the division of the solutions in two-dimensional space in relation to the centroid of the Pareto front.

In order to perform the final decision, the ADM relies on the user's preferences, i.e. which objective function should be given priority. In the case of our implementations, this could be communication latency, network bandwidth or service cost. If the user gives a strong priority towards a single objective,

then only the solutions in the given "objective's priority" region are considered. Afterwards, within this region we measure the distance of every solution to the centroid. Based on the distance, in the preferred objective dimension, we sort and weight the solutions and select the one closest to the objective weight preferred by the user.

5 Experimental evaluation

The experimental evaluation of the proposed concept of super-service provisioning was conducted based on a monitoring data-sets provided by RIPE NCC [17] and CEDEXIS [18]. The data-sets include comprehensive information on the response time, communication bandwidth and communication latency of multiple Cloud service providers from around the world. With respect to the implementation of the super-service provisioning algorithm, we have utilized the jMetal framework [13] for the purposes of multi-objective optimization and the Waikato environment for knowledge analysis [14] for the clustering.

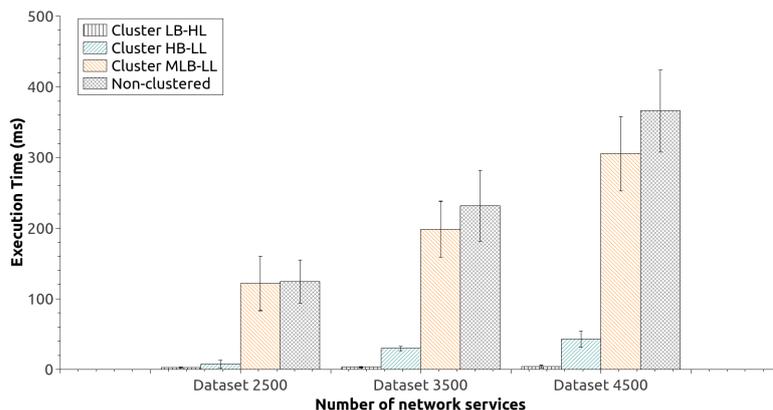


Fig. 4. Scalability of the clustered non-domination service sorting

As previously described, the process of super-service selection is conducted in two independent stages, therefore requiring distinctive set of evaluation experiments. The service non-domination sorting and Pareto construction has been evaluated on the basis of the degree of scalability for various cluster-sizes, while the behavior of the source and destination route multi-objective provisioning algorithm has been examined from multiple aspects, including solutions quality, scalability and computational performance.

To begin with, we evaluated the scalability and computational performance of the non-domination service sorting and selection algorithm by considering three distinctive data-sets with varying sizes from 2500 to 4500 network services. The

data-sets were clustered in three categories: high-bandwidth/low-latency with relative size of 12%, medium-to-low-bandwidth/low-latency with relative size of 85% and low-bandwidth/high-latency with relative size of 3%. The clustering time for all data-sets was below 2ms and was included in the total service sorting time. Figure 4 shows the correlation between the average sorting time for the full sets of non-clustered and clustered network services. It is evident that the clustering reduces the execution time of the multi-objective non-domination sorting, compared to the non-clustered datasets, from 20% to more than 1100% in the cases when small clusters have been created.

The process of source and destination network resource provisioning has higher computational complexity, compared to the service non-domination sorting, therefore requiring more comprehensive experimental evaluation. Figure 5 shows the correlation between the execution time of the multi-objective optimization algorithm and the length of the network route, for two different population sizes and evaluation limits. It can be easily observed that the optimization algorithm scales very good for different route lengths, with latencies ranging from 70 to 100 ms for routes with 30 hops. Furthermore, Figure 6 provides detailed information on the quality of the provided routes for different optimization parameters. The Hypervolume indicator was used to represent the quality of the Pareto set of solutions provided by the multi-objective algorithm. The presented results show that the quality of the Pareto routes decreases by up to 10% with the increase of the number of hops, which can be considered as satisfactory.

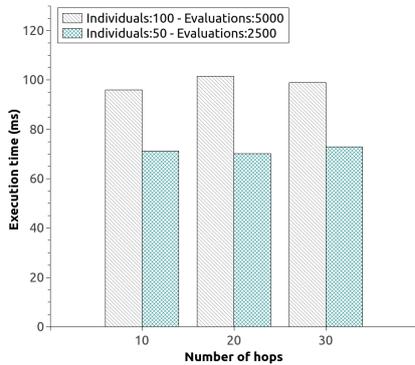


Fig. 5. Scalability of the resource provisioning multi-objective algorithm

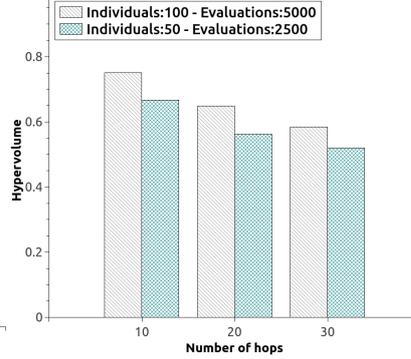


Fig. 6. Solutions quality of the resource provisioning multi-objective algorithm

In order to prevent the decrease of the quality of the Pareto front, in the cases when the number of hops in the route is higher, the number of individuals or evaluations for the multi-objective algorithm can be increased. Unfortunately, this can increase the execution time exponentially, which is not adequate for low-latency processes. Figure 7 shows the relation between the number

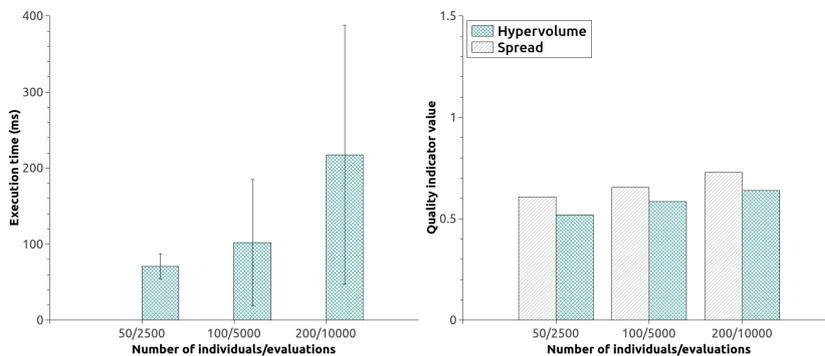


Fig. 7. Correlation between the execution time and the number of individuals/evaluations for fixed route size

Fig. 8. Correlation between the solution quality and the number of individuals/evaluations for fixed route size

of individuals/evaluations and the execution time for a fixed route of 30 hops. Furthermore, Figure 8 provides comparison between the number of individuals/evaluations and the quality of solutions. The quality of the solutions in this case was evaluated based on two indicators: hypervolume and spread. Overall, it can be determined, that in the cases when higher quality routes are required, higher number of individuals can be used with a penalty on the execution time.

6 Conclusion

This paper introduces a promising new paradigm of super-service provisioning for ultra-scale computing environments by unifying the management of the local communication resources and the external inter-domain network services under a single point of view. The research work has resulted in a development of an efficient technique for network services clustering and non-domination sorting, multi-objective network resource provisioning and a-priori automated decision making.

The presented paradigm has been evaluated based on a real-life monitoring data-sets. As our research deals with the utilization of clustering algorithms for reducing the complexity of multi-objective optimization problems, we present an experimental results that demonstrate the ability of our approach to provide an adequate super-service provisioning in inter-domain systems. The initial results confirm the scalability of the implemented algorithms and highlight the benefits arising from utilizing clustering for multi-objective non-domination sorting and optimization.

Acknowledgments

This work is being accomplished as a part of project *ENTICE: "dEcentralised repositories for traNsparent and efficienT vIrtual maChine opErations"*, funded

by the European Union's Horizon 2020 research and innovation programme under grant agreement No 644179.

References

1. M. Mihajlovic, L. Bongo, R. Ciegis, N. Frasheri, D. Kimovski, P. Kropf, S. Margenov, M. Neytcheva, T. Rauber, G. Runger, R. Trobec, R. Wuyts, R. Wyrzykowski, J. Gong, "Applications for ultra-scale computing", *Supercomputing Frontiers And Innovations*, Vol 2/1, 2015.
2. A. Celesti, F. Tusa, M. Villari, and A. Puliafito, "How to enhance cloud architectures to enable cross-federation", *IEEE CLOUD*, 2010.
3. M. Vincenzo, L. Rizzo, G. Lettieri, "Flexible virtual machine networking using netmap passthrough" *IEEE Symposium on Local and Metropolitan Area Networks*, 2016.
4. R. S. Montero, P. Massonet, M. Villari, G. Merlino, A. Celesti, A. Levin, D. Whigham, "BEACON: A Cloud Network Federation Framework" . *Advances in Service-Oriented and Cloud Computing: Workshops of ESOC*, 2015.
5. R. Xu, D. Wunsch, "Survey of clustering algorithms", *IEEE Transactions on neural networks*, 16(3), 2005.
6. W. Araujo, L. Z. Granville, F. Schneider, D. Dudkowski, M. Brunner, "Rethinking cloud platforms: Network-aware flexible resource allocation in iaas clouds", *IFIP/IEEE Symposium on Integrated Network Management*, 2013.
7. P. Bo, A. Hammad, R. Nejabati, S. Azodolmolky, D. Simeonidou, V. Reijs, "A network virtualization framework for IP infrastructure provisioning", *IEEE Conference on Cloud Computing Technology and Science*, 2011.
8. M. Jon, E. Jacob, D. Sanchez, Y. Demchenko, "An OpenFlow based network virtualization framework for the cloud" *IEEE Conference on Cloud Computing Technology and Science*, 2011.
9. S. Meng, K. Xu, F. Li, K. Yang, L. Zhu, L. Guan, "Elastic and Efficient Virtual Network Provisioning for Cloud-Based Multi-tier Applications", *Conference in Parallel Processing - ICPP*, 2015.
10. H. Ines, W. Louati, D. Zeglache, P. Papadimitriou, L. Mathy, "Adaptive virtual network provisioning", *ACM SIGCOMM workshop on Virtualized infrastructure systems and architectures*, 2010.
11. M. Jeffrey and L. Popa, "What we talk about when we talk about cloud network performance" *ACM SIGCOMM Computer Communication Review* 42.5, 2012.
12. Branke, Jürgen, et al., eds. *Multiobjective optimization: interactive and evolutionary approaches*. Vol. 5252. Springer, 2008.
13. J. J. Durillo, A. J. Nebro, "jMetal: A Java framework for multi-objective optimization", *Advances in Engineering Software*, 42(10), 2011.
14. S. R. Garner, "Weka: The waikato environment for knowledge analysis", *New Zealand computer science research students conference*, 1995.
15. K. Deb, A. Pratap, S. Agarwal, T. A. M. TMeyarivan, "A fast and elitist multi-objective genetic algorithm: NSGA-II", *Evolutionary Computation, IEEE Transactions on*, 6(2), 2002 .
16. T. Kanungo, D. M Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation", *IEEE transactions on pattern analysis and machine intelligence*, 24(7), 2002.
17. <https://atlas.ripe.net/>
18. <https://www.cedexis.com/>