Dear Professor Tvrdik,

below please find my review report of the PhD dissertation of Ing. Martin Štava entitled "Overlapping Non-dedicated Clusters Architecture".

The main motivation of the thesis is the desire to make use of unutilized resources in a distributed systems environment by enabling parallel execution and load distribution for complex computing tasks. Most current approaches in this field, however, are completely centralized with all known deficiencies such as a single point of failure and a single bottleneck. Moreover, due to the assumption of a closed-world local area network environment, most existing approaches also ignore security requirements.

The goal of the given thesis is to overcome these limitations by designing and implementing a decentralized, peer-to-peer based clustering architecture that also takes security as a main goal into account. For validation purposes, the provisioning of flexible scheduling mechanisms within a specific framework has also been an important goal. As a basis for the overall project, the so-called Clondike system has been used that was developed within the same institution.

With respect to the current state of the art and the research literature, the dissertation is rather up-to-date and pursues important goals that have not been thoroughly addressed by existing work. Moreover, from a practical point of view, the overall topic is also very important due to the huge potential in clustering underutilized resources on a broad scale.

In chapter 1, the main goals of the thesis are being motivated in more detail, and a clear distinction from existing approaches is outlined. Important terms such as clustering and grid computing are defined and illustrated. The background in the area of operating systems, networking and distributed systems, and cloud computing is then discussed in more detail in chapter 2. Actually, this discussion is relatively basic and general and could have been shortened a bit by referring to adequate textbooks covering these topics already. As opposed to
that, the subsequent discussion of directly related work in the same chapter is very valuable and very well-elaborated; system approaches such as the mentioned Clondike system, Mosix, OpenSSI, Now, Globus, OurGrid, Plan 9, and several other solutions are addressed in detail and are clearly compared and distinguished from the own approach of the author. Overall, this represents a very good and original contribution regarding the structured discussion of the current research state.

Based on the given motivation and analysis, the own cluster system architecture is then outlined in chapter 3, the main conceptual chapter of the thesis. Important aspects include the enhanced flexibility of clustering, the support for more generic, non-dedicated clusters with possible overlapping for redundancy, the completely decentralized structure, process migration support, caching support, and also a dedicated security scheme. Moreover, a combination of non-preemptive and preemptive scheduling is supported, and a specific classification and prediction approach for scheduling control is introduced. Altogether, the broad set of developed concepts with their sophisticated technical details represents a very important and valuable systems-level scientific contribution of the author.

Chapters 4 and 5 cover the implementation and validation of the approach. The description of dedicated kernel-level components (patch and selected modules) and higher-level user-mode components with the so-called director, the communication channels, and the migration mechanisms is very convincing and addresses the important technical details at a sufficient level of implementation. Within the experimental evaluation, typical application problems such as a Mandelbrot set calculation and a kernel compilation are applied, and scalability in terms of the problem size and the involved machines is addressed in particular. By thoroughly providing a comparison with the potential ideal scaling behavior, the reader of the thesis can understand and judge the made contributions very well and can easily compare them with the theoretical potential. Finally, the major results in terms of system implementation, security aspects, and scheduling framework are summarized in chapter 6, together with an outlook to future work.

Altogether, the structure and organization of the dissertation is as expected and conforms to the typical scientific approach for systems theses. The originally set objectives have been fully achieved, with own scientific concepts, and with their full implementation and validation. The scientific methods applied are mainly based on architectural design of software components with an informal description of their behavior, on systems-level implementation and on systematic experimental evaluation based on measurements. While this represents a usual and acceptable approach, it should also be mentioned that formal techniques, such as formal behavior specification by automata or formal specification of security protocols, have not been applied by the author. For the given solution, this was not necessary indeed, but more advanced correctness proves, e.g. of the fulfillment of security protection goals, would require such more rigorous techniques.

Over all, the systems-level contribution to computer science in the areas of operating systems, clustering and cloud computing, and distributed systems is definitely very important and valuable. Several new mechanisms for process clustering and scheduling, for decentralized protection according to security goals were developed, together with a full practical val-
idation. The results were published at several international conferences and gained significant external visibility this way.

As usual for such a complex thesis topic, not all questions and issues could be addressed to the fullest extent. A list of typical and interesting topics for further discussion, for example also during the thesis defense, includes the following aspects:

- The decentralized security architecture without a trust center or any other central authority also comes with major disadvantages and risks. In the outlook, the author mentions reputation-based systems as a possible further-ranging solution. It should be discussed how such a solution could be integrated with the own approach, or also whether a conventional PKI (public key infrastructure) could be integrated for a slightly less decentralized, but even more secure approach.

- The issued certificates that help to implement access control and delegation of access should also be subject of protection themselves. It should be discussed to which degree such certificates could potentially be forged and whether an attacker might find ways to artificially introduce malicious certificates into the system.

- Network security is actually being discussed within the thesis (e.g. on page 51), but the overall solution then simply resorts to the use of existing techniques such as IPsec channels and OpenVPN. It should be discussed whether more sophisticated techniques could be applied that would for example support a whole range of different security protocols at different layers (e.g. including the transport or application layer).

- The mechanisms of classification and prediction in the area of process scheduling (page 58), together with the possible migration strategies found in the implementation chapter, seem to be very interesting and complex. The author should further discuss the details of these approaches, especially with respect to the potential of self-learning scheduling systems.

- From an implementation-level point of view, the actual implementation of multi-threaded code migration (page 67) seems to be very interesting. The author should provide further insights into the techniques employed at implementation level.

- From a validation point of view, the performance of the security rule evaluations (page 71) seems to be a critical issue when many different rules are being applied. The author might discuss the limitations and possible improvements in more detail.

- Similarly, the possibilities to improve scalability in general, as point out on page 87 for example, should also be further discussed as they will provide important hints towards future research work in this domain.

- Finally, it has to be remarked that performance was validated very thoroughly, but security (from an attacker point of view) was not actually part of the validation. The
author should at least outline how this field could be addressed (for example based on dedicated attacker models with associated test cases, based on formal proofs of security protocols, or based on other means).

Over all, the dissertation represents an important scientific contribution, including a significant number of novel concepts and a full practical validation and implementation. The thesis is well-written and has also lead to a number of important scientific publications. Based on the overall contributions judged in more detail above, the thesis is considered to be a very good piece of work, typically reflected by the grade “magna cum laude”. The author of the dissertation definitely proved the ability to conduct research and achieve scientific results. In accordance with par. 47, letter (4) of the Law Nr. 111/1998 (The Higher Education Act), I do recommend the thesis for the presentation and defense with the aim of receiving the Ph.D. degree.

Dresden, March 6, 2013

Prof. Dr. Alexander Schill

Prof. Dr. Alexander Schill
The thesis studies the area of clustering architectures, especially the security of decentralized peer-to-peer solutions. Based on the Clondike project, the author have proposed and implemented extensions of the framework targeting the security issues in the P2P cloud environment. Moreover, some other modules of the original Clondike system were improved by the author - the migration support, the scheduling membership management, caching etc. The evaluation presented in the thesis shows that the overhead of the proposed system imposed by a more complex architecture is appropriate and acceptable.

Since most of the relevant cluster and grid frameworks does not support such level of security, this area is intensively researched. The author compares his proposed solution to other relevant contemporary systems. The presentation is well ordered with concepts being introduced in a logical order. The thesis is well written and the context is reasonably structured. The scientific results were reviewed and published on international conferences.

Nevertheless, I have several minor remarks and questions for the defense:

- The comparison of the proposed system to its relevant counterparts is based on the top-level ideas only. The evaluation and performance benchmarks cover only the proposed system. More detailed comparison (both at the functional point of view and the performance) would be very valuable.
- The author mentioned a proposal and an implementation of the migration mechanisms several times. Nevertheless, the citations reference master theses of other authors. What is the exact contribution of the author in this area?
- What is the policy when new nodes join a particular cluster? Is there some form of Trust Management or similar principle being used?
- Was the implemented system used in a real-world (or close to real-world) environment for other tasks than just the prototype evaluation?
The author of the dissertation proved the ability to conduct research and achieve scientific results. In accordance with par. 47, letter (4) of the Law Nr. 111/1998 (The Higher Education Act) I do recommend the thesis for the thesis for the presentation and defense with the aim of receiving the Ph.D. degree.

Prague, March 20, 2013

RNDr. Filip Zavoral, PhD.
Department of Software Engineering
Faculty of Mathematics and Physics, Charles University in Prague
Review of the PhD Thesis

Overlapping Non-dedicated Clusters Architecture

submitted by

Mgr. Martin Štáva

The doctoral thesis presented by Mgr. Martin Štáva is focused on scheduling of the application processes for a cluster of computation processors. The thesis incorporates the new approach to the cluster management - designed, developed and tested by Mgr. Martin Štáva himself.

Contribution of the thesis text is primarily presentation of the new overlapping non-dedicated clusters (ONDC) architecture, description of its implementation at the Clondike Linux-based cluster, and the results of the ONDC scheduling for two selected multi-process applications.

ONDC supports more concurrent applications running in the cluster controlled from different processors. Moreover, it supports a computing priority of local user applications and it supports security.

Chapter 3 presents description of the overlapping non-dedicated cluster (ONDC) that is based on the idea of several overlapping placements of more groups of processes. Each of these groups consists of one manager and a set of workers. Moreover, all of these groups have to take into account a possibility of external processes with higher priority and of processes trustfulness.

All support technologies required (filesystem, caching, security, scheduling and network interconnection) in the ONDC are described in short subsections of the Chapter 3. Probably the most important for the process scheduling is the part 3.5. (Scheduling) that presents the main difference between ONDC and scheduling in traditional clusters.

Chapter 4 presents the structure of ONDC processes (management and workers) and their support implementation in the Clondike Linux-based cluster that serves as a basis for ONDC core and detached processes.

Communication channels, migration support, filesystem integration, scheduling and cashing are described mainly in textual form and simple corresponding figures.

The last part of the chapter 4 mentions requirements to support clusters with different hardware architecture (e.g. 32 and 64 versions of processors). Is this support implemented in the ONDC?

Finally, Chapter 5 presents results of experiments with two applications in the ONDC cluster. The first one is calculation of Mandelbrot set, the second is the compilation of the standard Linux kernel.

The Mandelbrot set experiments compare computation on a single processor, single computation with interrupts on more processors and multiple computation on more processors.

The OS kernel compilation has been tested on the homogeneous Clondike cluster, and on the homogeneous cluster with four different machines.
In the following paragraph I will present some notes related to the main goal of the thesis, which could be clarified during the defence.

- Can more specific relation of process scheduling requirements (mentioned in the pg. 34) and implementation of ONDC in Clondike cluster be presented?
- Any description of the distribution of Mandelbrot set algorithm into several cooperating processes, which has been used in the experiment, would be reasonable. Are the computation requirements of these processes same or different? Moreover, how the distribution of the algorithm is related with the number of processors in experiments? Is the algorithm consisting of 4*25 processes (Fig. 5.3) or more of them?
- Has been the number of tasks for every computation in the single and in the multiple computation evaluation the same?
- I can understand usage of interpreted language Ruby in testing due to its simplicity. However, a compiled version of the model would provide better information.

The thesis shows authors' capability to formulate a scientific problem, find a convenient solution, and prove it in practice. However, I would like to see a more precise formal description of the ONDC method and of the test based on Mandelbrot set generator during defence of the thesis.

I can recommend the thesis, in the interpretation of the Law 111/98 in the Digest, for the defence with the goal to be awarded by the PhD degree in Informatics at FIT CTU in Prague.

Prague, May 7th, 2013

[Signature]