

## 1st Speech on the Grid VT project teleconference

Date and time: 23.4. 2012 at 11:00

Participants:

EGI: Gergely Sipos

UI SAV (Slovakia): Ladislav Hluchý, Milan Rusko, Jolana Sebestyenová, Peter Kurdel, Marian Trnka, Marian Ritomský

TU Košice (Slovakia): Jozef Juhár, Matúš Pleva

IDIAP (Switzerland): Milos Cernak

CSC (Finland): Ville Savolainen

Claire Devereux (STFC RAL, NGI Manager, Great Britain) was not able to take part in the EVO call because of a medical appointment.

Program:

1. Greeting and introduction (Ladislav Hluchý)
2. Speech on the Grid project presentation (Milan Rusko)
3. Reactions of Gergely Sipos
4. Reactions of Ville Savolainen
5. Reactions of Milos Cernak
6. Reactions of Jozef Juhar
7. Conclusion (Ladislav Hluchý)

**Ladislav Hluchý** (the Slovak NGI and director of UI SAV) welcomed the participants and shortly explained the main aims of the VT projects.

**Milan Rusko** presented the main ideas of the SPEED project, the slides can be found at ([http://www.slovakgrid.sk/downloads/MR\\_SPEED\\_Presentation.ppt](http://www.slovakgrid.sk/downloads/MR_SPEED_Presentation.ppt))

**Gergely Sipos** (EGI) has reacted to the presentation. As the sound quality was very poor he also sent the following notes and questions in written form:

*Finding more members*

- Which institutes are the most active and advanced on this research field? Contact those, or send me their name and I connect them to the local National Grid user support team and they can discuss involvement in our project.

- Who develops speech processing algorithms/applications in Europe that should be used in the model which is on slide 4?

*Requirements*

- Large data storage; secure data transfer (the data is expensive) Can you name any database and/or app where this is relevant?

- Do you have components that could be used to implement the model which is on slide 4? (real DB, real model; real decoder; real pre-processor, ...) The VT should collaborate on collecting components for the model, port/connect these components with the European Grid Infrastructure, and provide this as a service for speech processing communities in Europe.

*The presented model*

*- Is it supported by Slovakia and by Switzerland? How generic is this model in speech processing sciences?(Milos answered that it's a generic model in this field.)*

*We should hear about the goals and priorities of the British and Irish members of the VT who did not attend the telcon today.*

*For future teleconferences we can use the EGI Webex system, that should be more robust and usable than EVO.*

**Ville Savolainen** has introduced himself as the Development Manager at CSC - IT Center for Science Ltd., Helsinki. He informed the participants that CSC is a computing service provider for universities and he offered passing the information on our project and call for participation to the speech processing groups at these universities.

**Milos Cernak** from IDIAP also tried to join discussion, however there was a problem with dropouts in his audio, so he later sent the summary of his presentation in written form:

*I represent Idiap research institute in this VT. Idiap has a long term experience with speech processing (both recognition and synthesis) and can be a vital partner in ongoing discussions about feasibility of using Grid technology for speech processing. Moreover, our spin-off Koemei is also very well advanced in this grid area, and is building some of its speech activities around cloud computing.*

*By the way, how is EGI related to cloud computing, the service adopted by many current speech technology companies?*

*The presented model was created at Slovak Academy of Sciences and I can confirm that is it a generic model in this field. I believe that every academic institute that is working in the field has some of the mentioned components. It is worth to note that most of them are using open source generic tool-kits. Contrary to academic institutions, most of companies use in-house implementation of the components.*

*Our (Idiap) goal in the project is to investigate possibility of using huge grid in speech technology, specifically definition of new areas in research in development that push technology forward. It could be holistic optimization as proposed, or something completely new that can arise from partners' discussion. The SPEED looks interesting, although probably very challenging to put together.*

**Jozef Juhár** from Technical University of Košice has explained, that as he was not an expert in GRIDs. He is looking at the project from the point of view of practical use of computing capacity for speeding up the processes, which are the most time consuming in the speech research area.

He said, that the task as it was presented by Milan was really complex. He therefore proposed to start from dividing the whole task into smaller subtasks and concentrate only to some of them at the beginning of the project.

As the first tasks to focus on, I would recommend the optimization of off-line training and testing procedures (acoustic an language modeling, building finite state transducers and recognition net.

### **Milan Rusko's reaction to the teleconference, answers, deductions and conclusion**

Dear participants, let me first thank you all for joining the conference and for being so tolerant to the connection problems.

I will try to answer all questions and to give some proposals for the future stage of the project.

At first let me please point out that the main goal of the project is to make the GRID computing an everyday tool of research and development in speech processing. The scheme on the slide 4 serves only for demonstration of complexity of the problems solved in this area. This scheme should also serve as a graphical version of classification of different subsystems that need to be optimized and coordinated. It can be also useful for referring to some parts of the system during discussions via email, skype etc. Despite of the complexity of this scheme and the task of holistic optimization that it represents, this is only one type of issues in which the high computing power is necessary to make the solution possible.

We hope that further members of the consortium will bring further pilot tasks and help the project to get more insight to the needs of the speech processing community.

Gergely, we will provide with a list of the institutes and research labs that are the most active and advanced on this research field. We would be very glad if you could ask NGIs to contact them and check their possible willingness to join the project.

*“Requirements - Large data storage; secure data transfer (the data is expensive) Can you name any database and/or app where this is relevant?”*

Collecting and processing the speech and text (or even video) databases which are inevitable for modern speech recognizers and synthesizers research and development is the most time and money consuming part of the whole system and therefore the databases are expensive. Some of them are sold by LDC and ELRA and one can check prices on their pages, but often even this offer is insufficient for many purposes and the databases need to be recorded and annotated which requires a lot of expert work. Therefore they are expensive and they represent an advantage in competition.

*- Do you have components that could be used to implement the model which is on slide 4? (real DB, real model; real decoder; real pre-processor, ...) The VT should collaborate on collecting components for the model, port/connect these components with the European Grid Infrastructure, and provide this as a service for speech processing communities in Europe.*

Yes, probably all of the partners working in speech processing research have some of these components, but not too many of them want to make them freely available. Therefore we propose using resources freely available for research purposes (e.g. under GNU license). To prevent language problems we propose to start with some freely available English speech database and freely available English texts database. The well known Hidden Markov Toolkit HTK < <http://htk.eng.cam.ac.uk/>> should be used for modeling, decoding and testing in the starting phase of the project if common experiments will have to be done.

*We should hear about the goals and priorities of the British and Irish members of the VT who did not attend the telcon today.*

Yes, we are very interested in the opinions and concepts, and we hope they will join our discussions.

*For future teleconferences we can use the EGI Webex system, that should be more robust and usable than EVO.*

We can try it if you, Gergely recommend this system.