Utilization of GRID technology in processing of medical information

Mgr. Tomáš Kulhánek

Supervisor: Ing. Milan Šárek, CSc.
(CESNET z.s.p.o.)

Consultant: Doc. MUDr. Jiří Kofránek, CSc.
(ÚPF 1.LFUK)

Problem statement

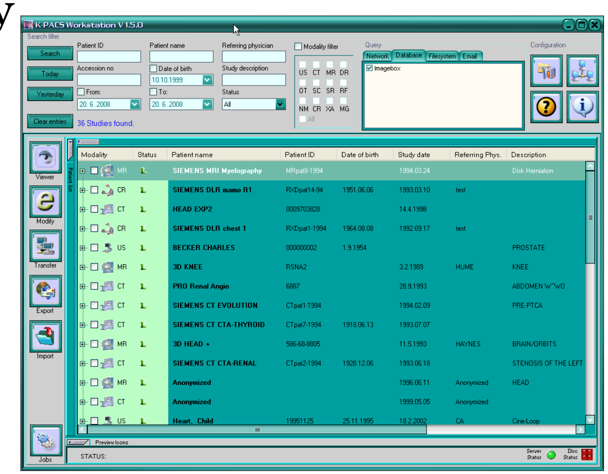
- Hypothesis:

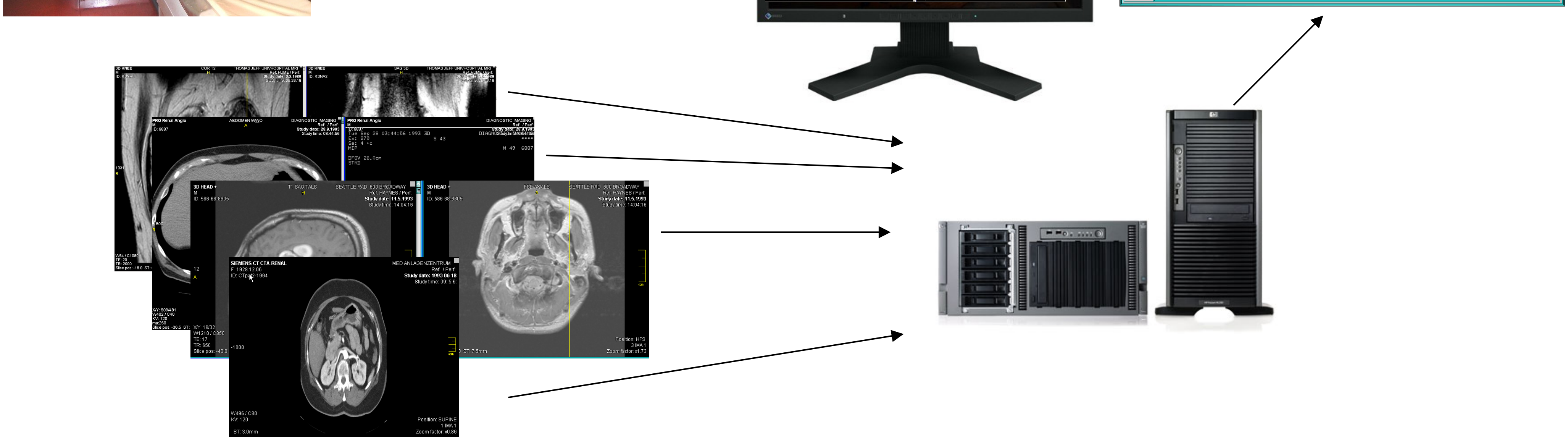
  The technologies related to grid-computing and cloud-computing may improve processing of medical information to perform demanding tasks which are almost impossible or may need onerous effort to achieve using classical local or institutional resources.

- Research questions:
  - Is it (grid-computing or cloud-computing) beneficial and how?
  - What are the limitations?
  - What can be the direction of future biomed. research?

Sharing medical images

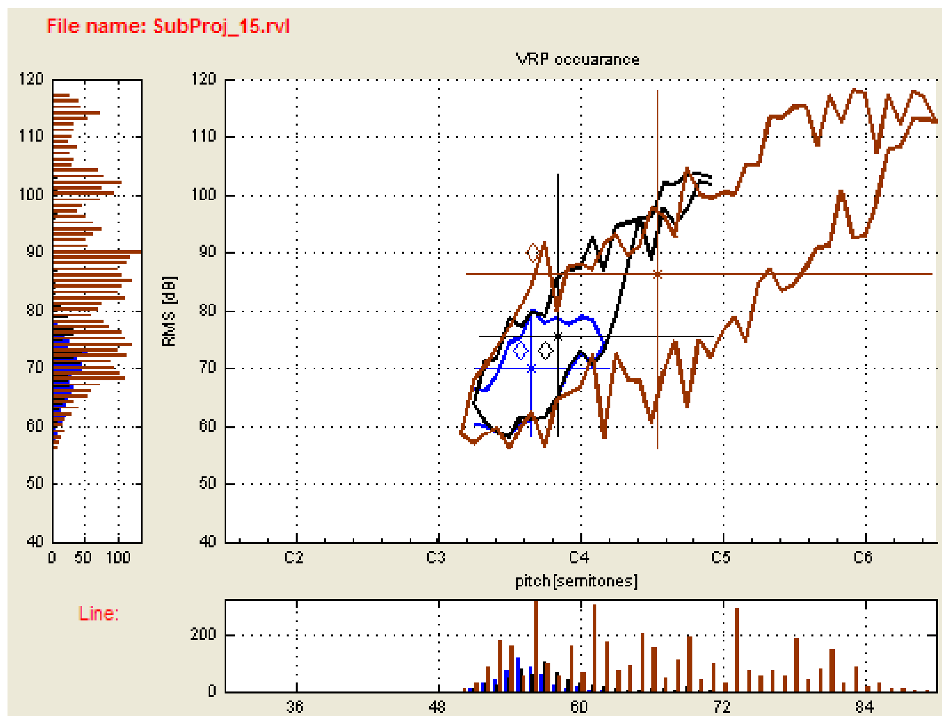
- Within hospital - PACS(Picture Archiving and Communication Systems) DICOM (Digital Imaging and Communications in Medicine
- Among hospitals - MEDIMED – based in Masaryk University
- EPACS – based in VFN, provided by ICZ



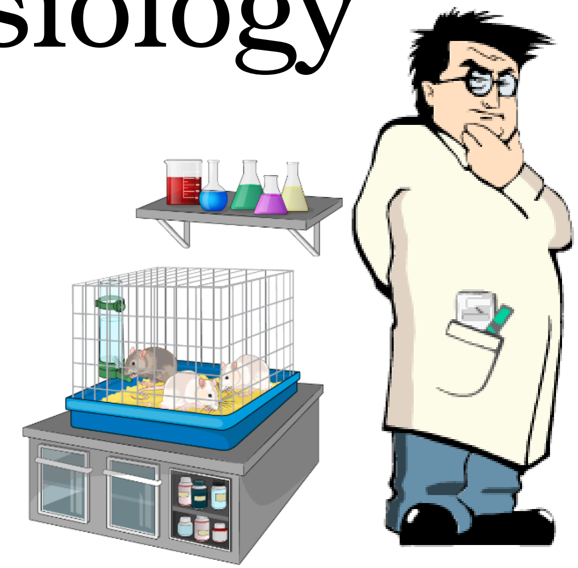


Voice Science

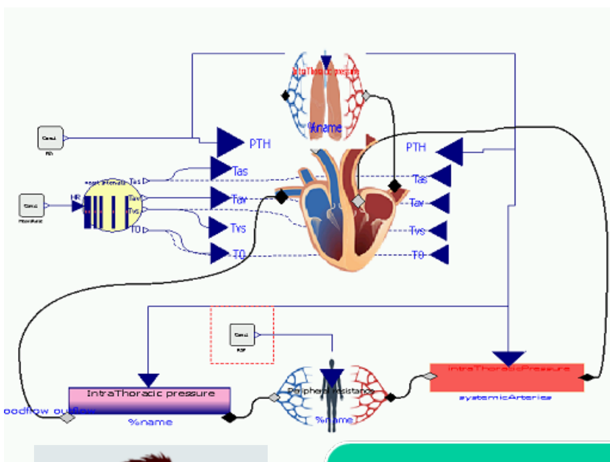
Voice analysis – phoniatric examination, voice professional training



Computational physiology



Experiment, measurement of experimental data



Model, simplified mathematical description of reality

Verification, Validation, compare model simulation with real data

Methods

Distributed computing

- Resources: computers, devices connected via computer network, shared access
- Users need sometimes resources.

- Manage fair use of resources over network.

Grid computing

- Distributed computing,
  - shared resources among universities, research institutions, companies
  - parallelism, federated access
- Service-grid: services by grid-middleware, shared access, CESNET (METACENTRUM), EGI
- Desktop-grid, volunteer computing: BOINC framework, SETI@home, Rosetta@home …

Cloud computing

- Virtualization
- Providers of infrastructure: Amazon EC2, Google Cloud, …,

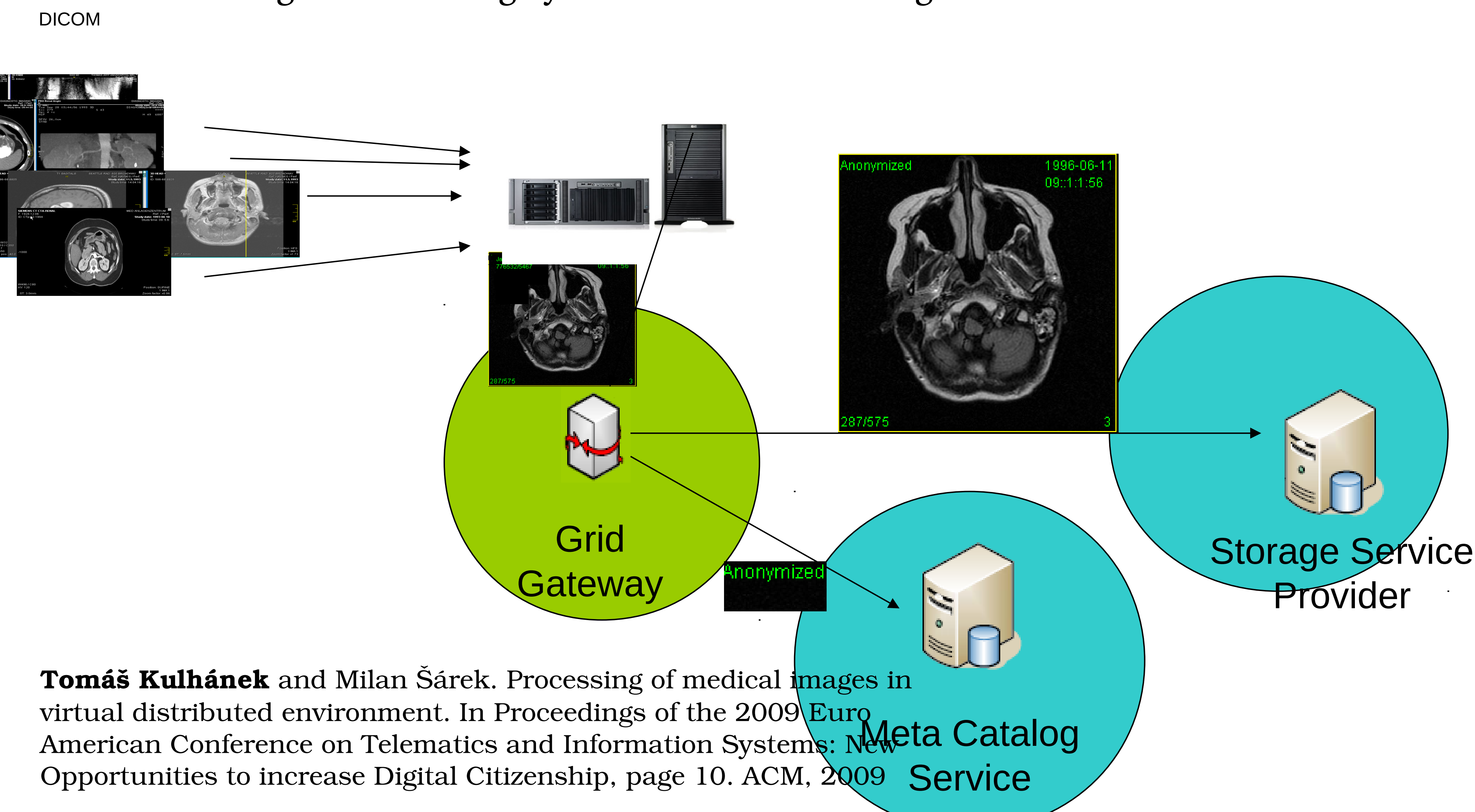
  physical infrastructure – scientific grid
  - Ability to execute virtual machines: METACloud, EGIFedCloud

- Providers of services: Facebook, Gmail, GoogleDocs, …

Results

Sharing medical images

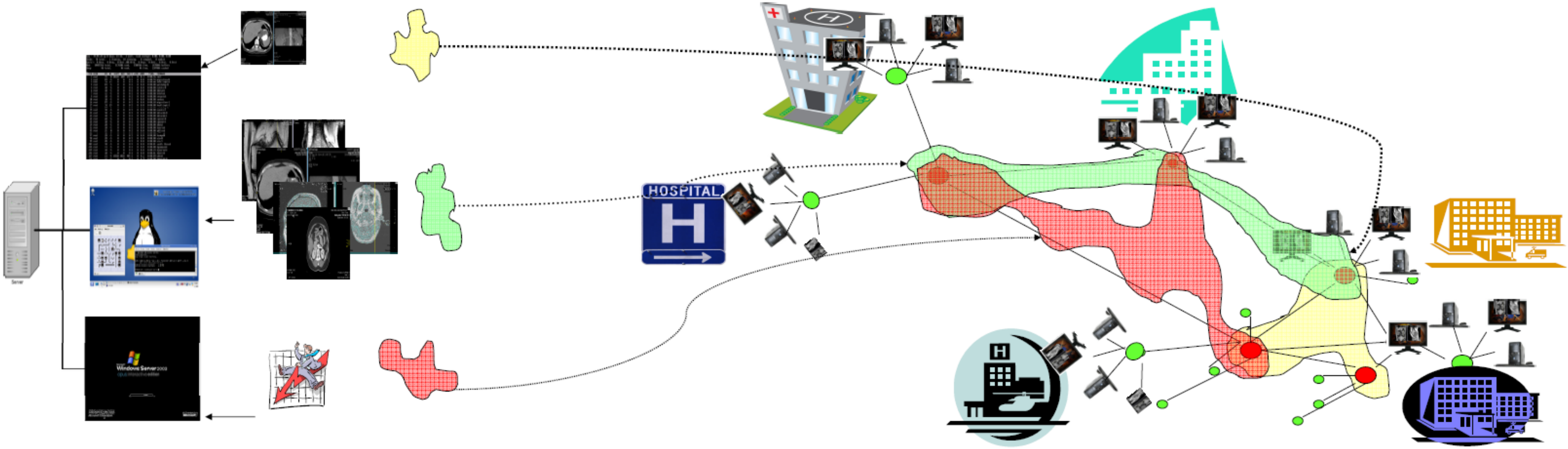
- Integrated existing system MEDIMED with grid based Globus MEDICUS



Tomáš Kulhánek and Milan Šárek. Processing of medical images in virtual distributed environment. In Proceedings of the 2009 Euro American Conference on Telematics and Information Systems: New Opportunities to increase Digital Citizenship, page 10. ACM, 2009

Virtual infrastructure

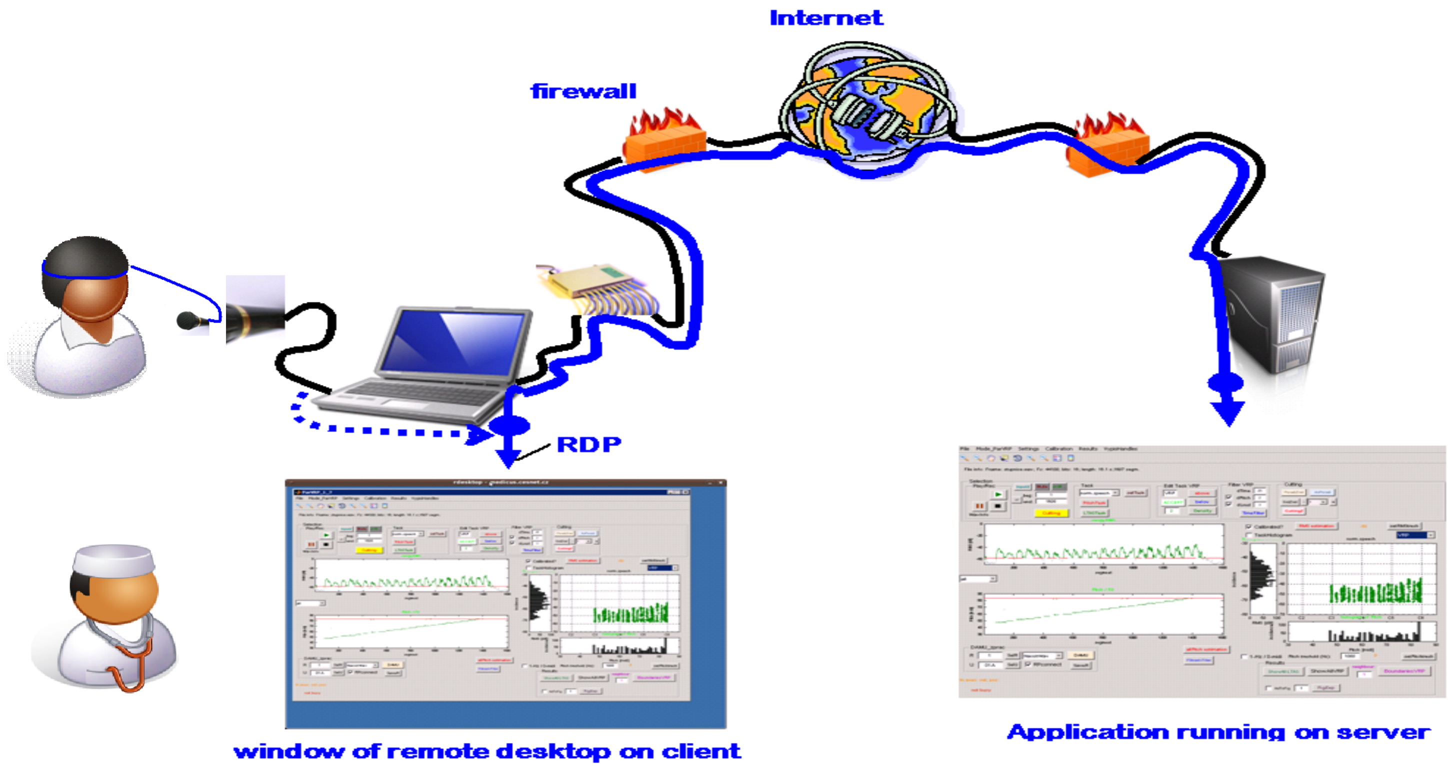
- Virtualization – consolidating computer servers
- Connected servers in CESNET, Central Military Hospital and First Faculty of Medicine
- Sharing HW for multiple application



Tomáš Kulhánek. Infrastructure for data storage and computation in biomedical research. European Journal for Biomedical Informatics (EJBI), 6(1):55–58, 2010.

Remote analysis of human voice

- Customized RDP protocol with lossless soundrecording redirection, application deployed on virtual server, user connects with remote desktop



Tomáš Kulhánek, Marek Frič, and Milan Šárek. Remote Analysis of Human Voice Lossless Sound Recording Redirection. Analysis of Biomedical Signals and Images. Proceedings of 20th International EURASIP Conference (BIOSIGNAL), pages 394–397, 2010.

Computational physiology

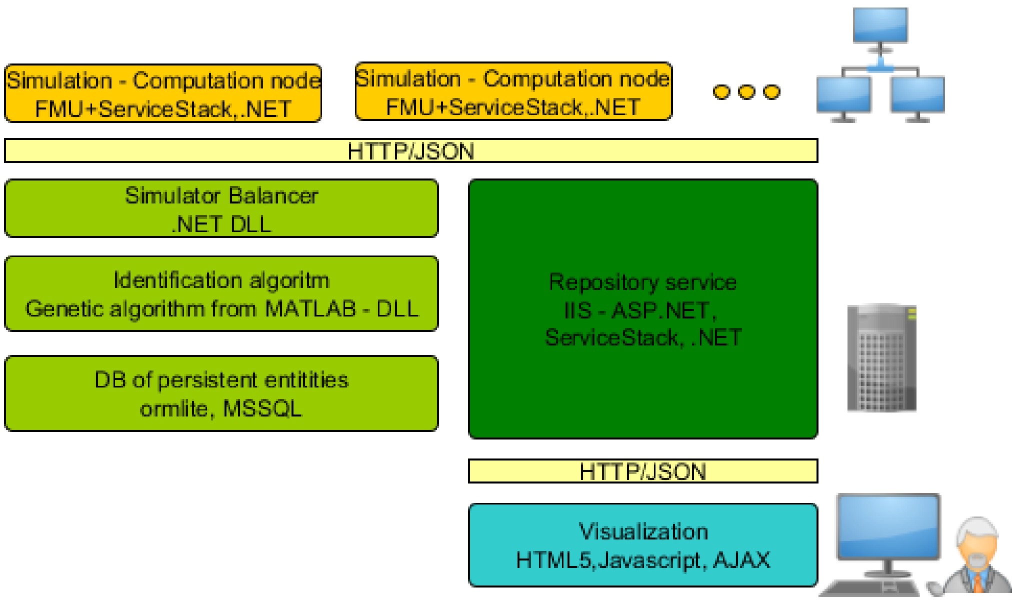
- **System for parameter estimation**

Modelica model => FMI Executable (commercial Dymola tool)

Web portal and web services => .NET, REST API (ServiceStack,SignalR)

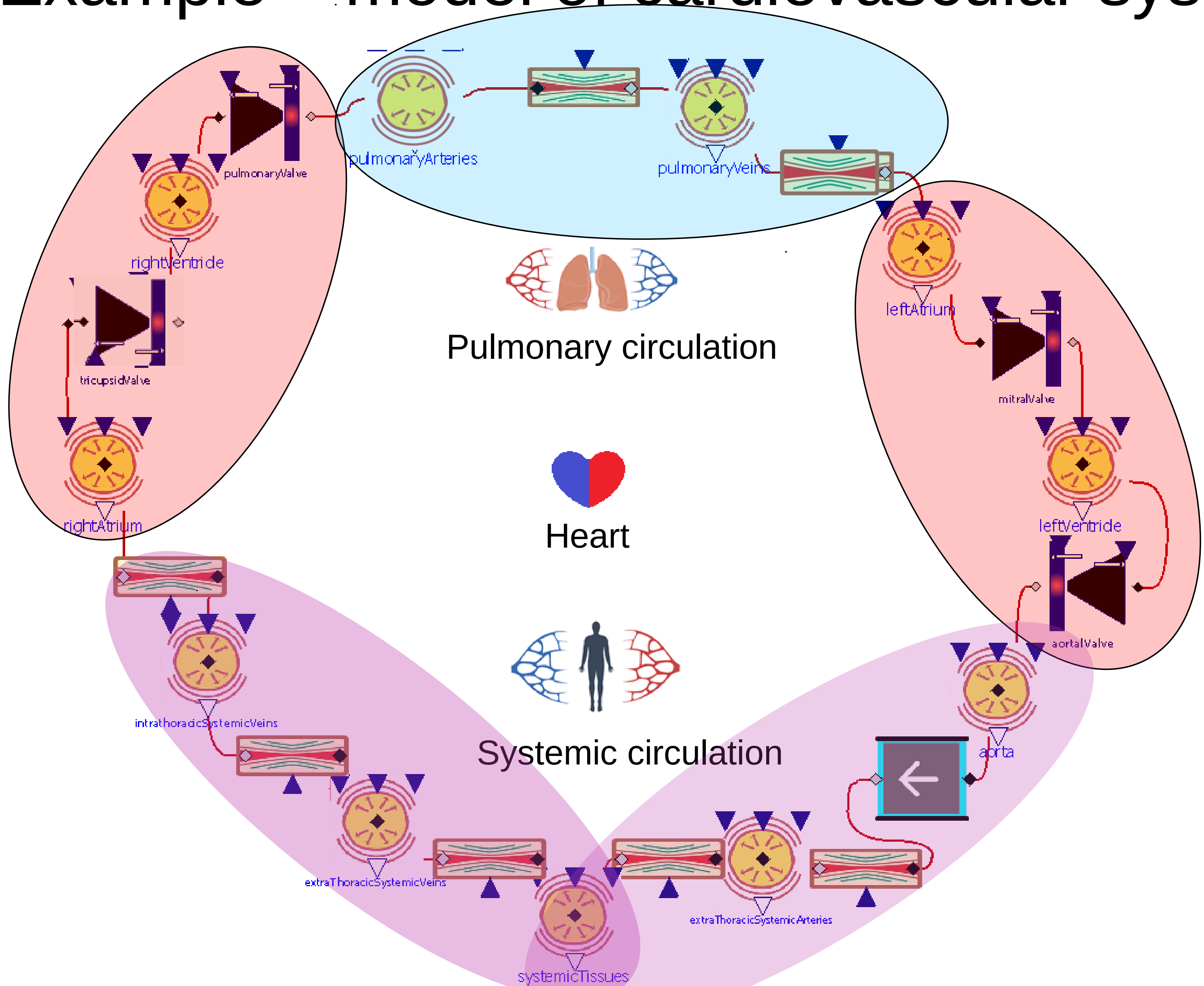
Non-linear mathematical models => Global optimization methods => Genetic algorithm

Each iteration produces independent tasks => we tried Desktop grid (BOINC), local cluster, EGI cloud (CESNET NGI)

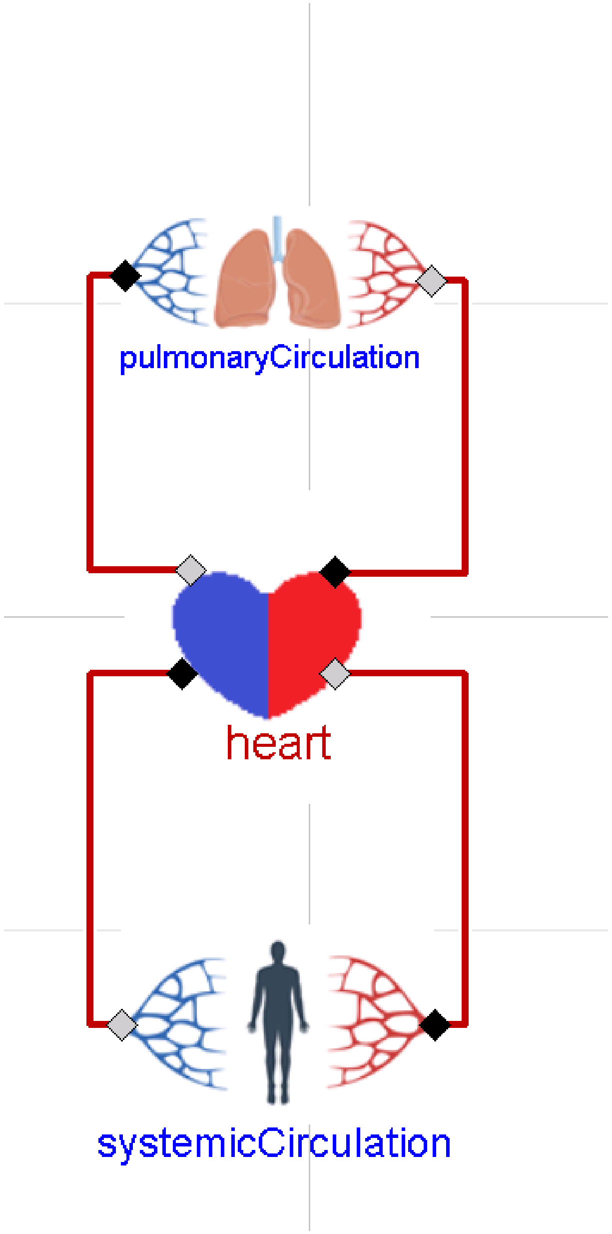


Tomáš Kulhánek, Marek Mateják, Jan Šilar, and Jiří Kofránek. *Parameter estimation of complex mathematical models of human physiology using remote simulation distributed in scientific cloud.* In Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on, pages 712–715, 2014. doi:10.1109/BHI.2014.6864463.

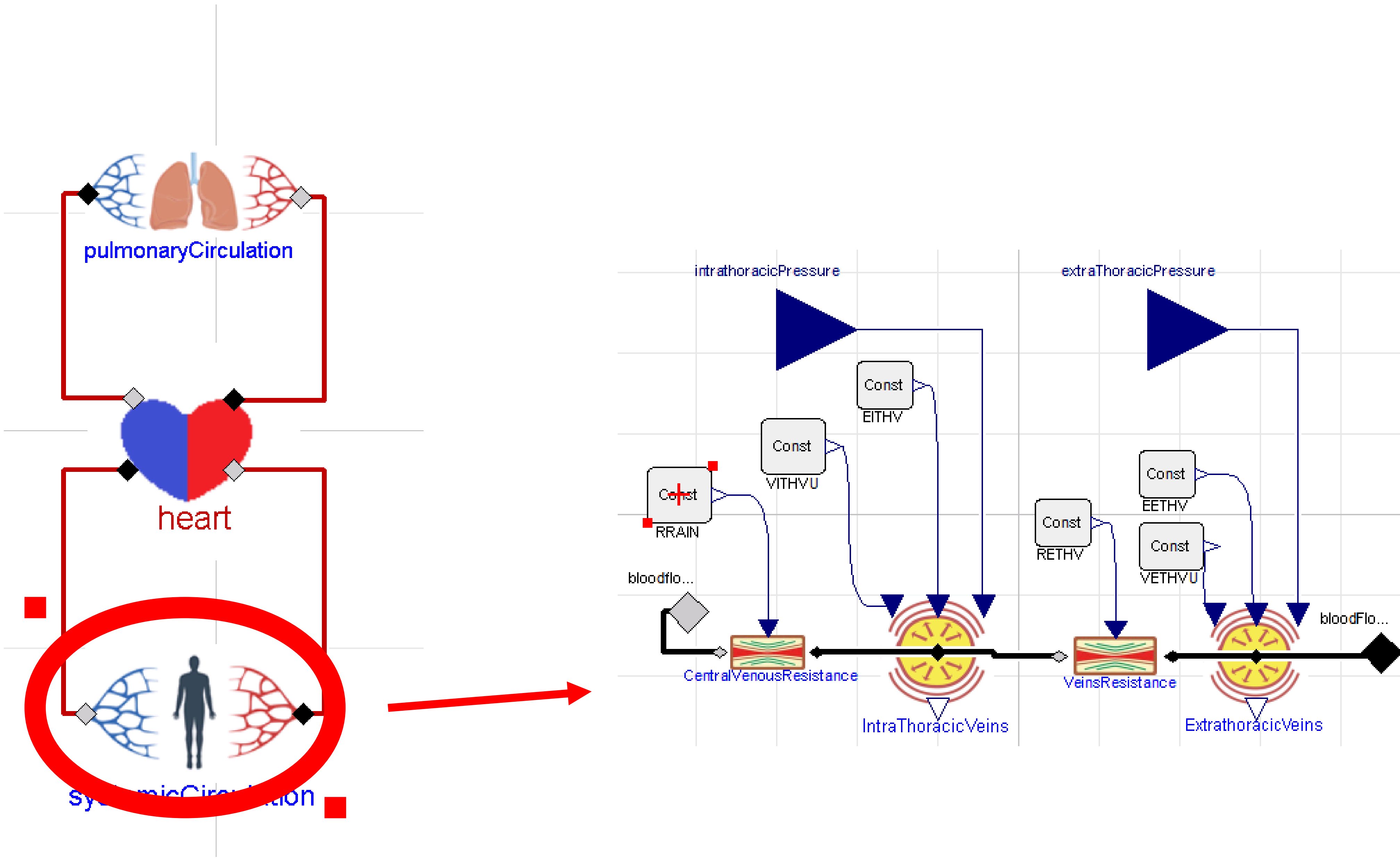
Example – model of cardiovascular system



Example – model of cardiovascular system



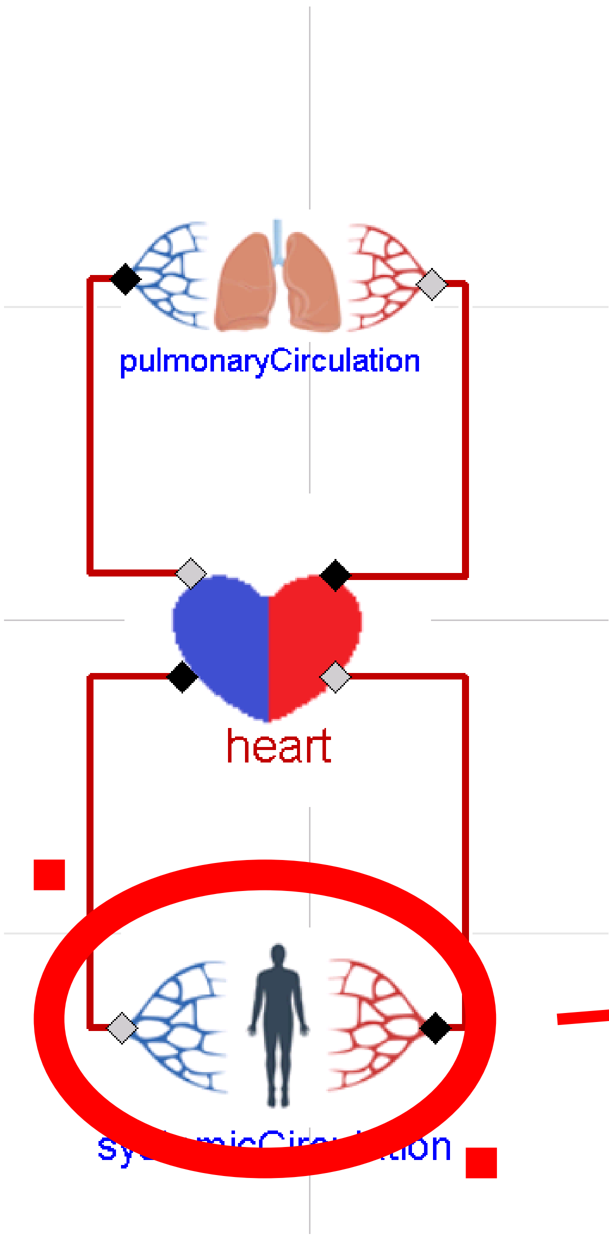
Example – model of cardiovascular system

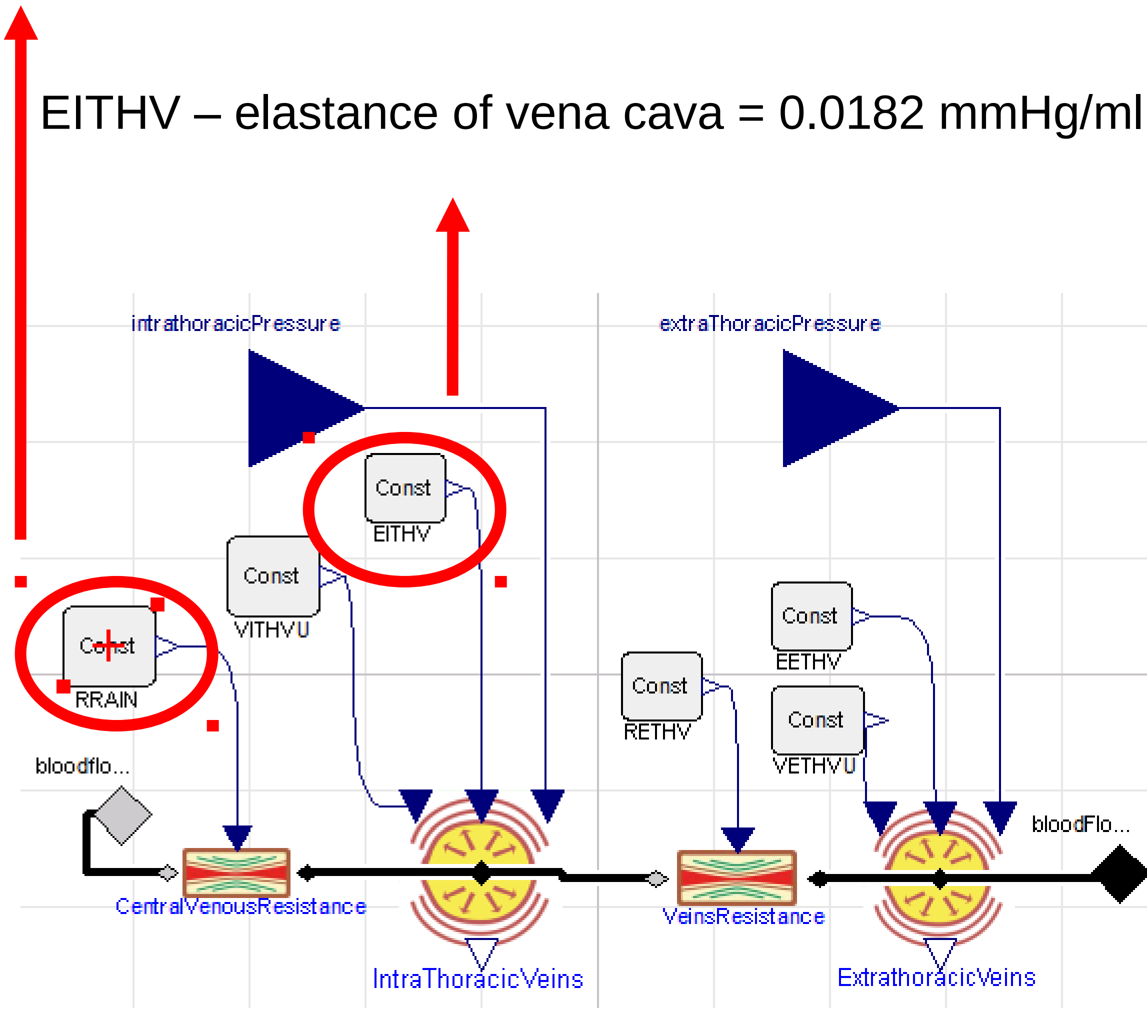


Example – model of cardiovascular system

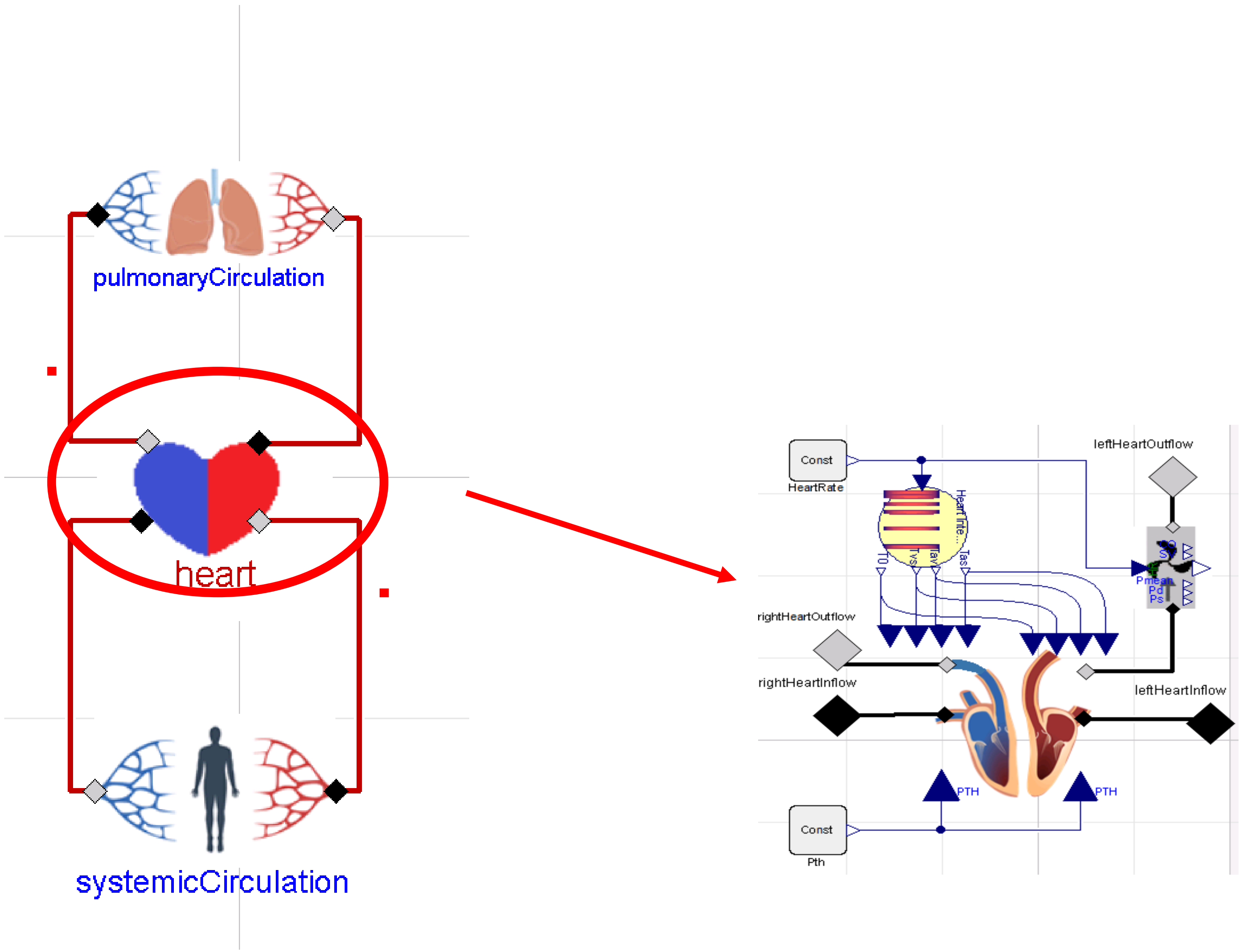
RRAIN – resistance of vena cava=0.003 mmHg*s/ml

EITHV – elastance of vena cava = 0.0182 mmHg/ml



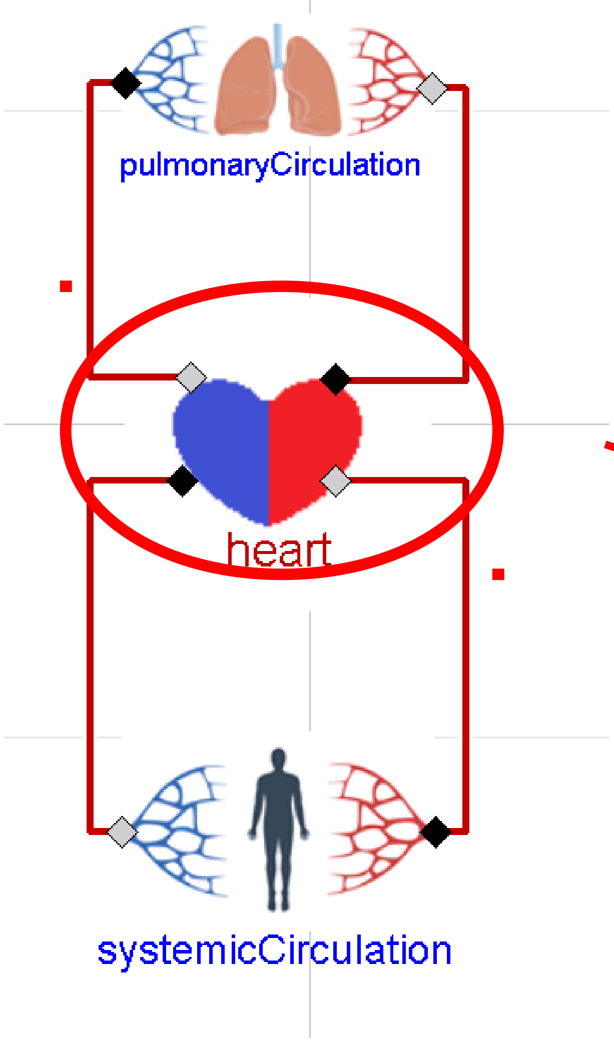


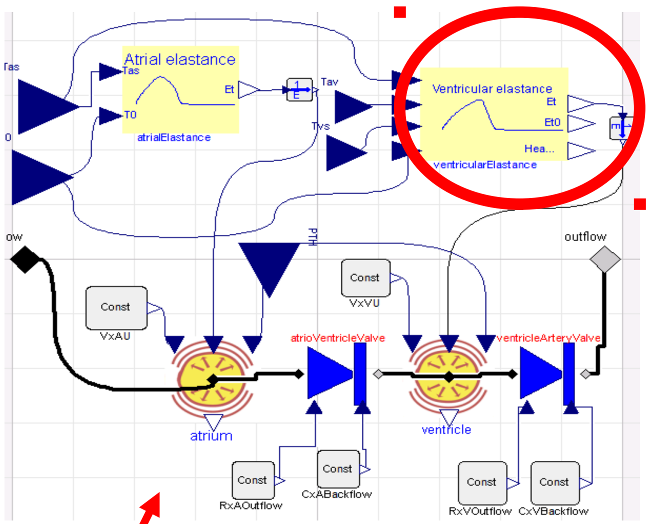
Example – model of cardiovascular system

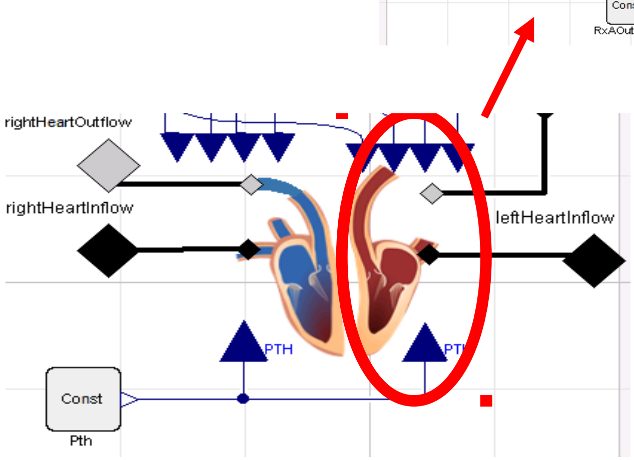


Example – model of cardiovascular system

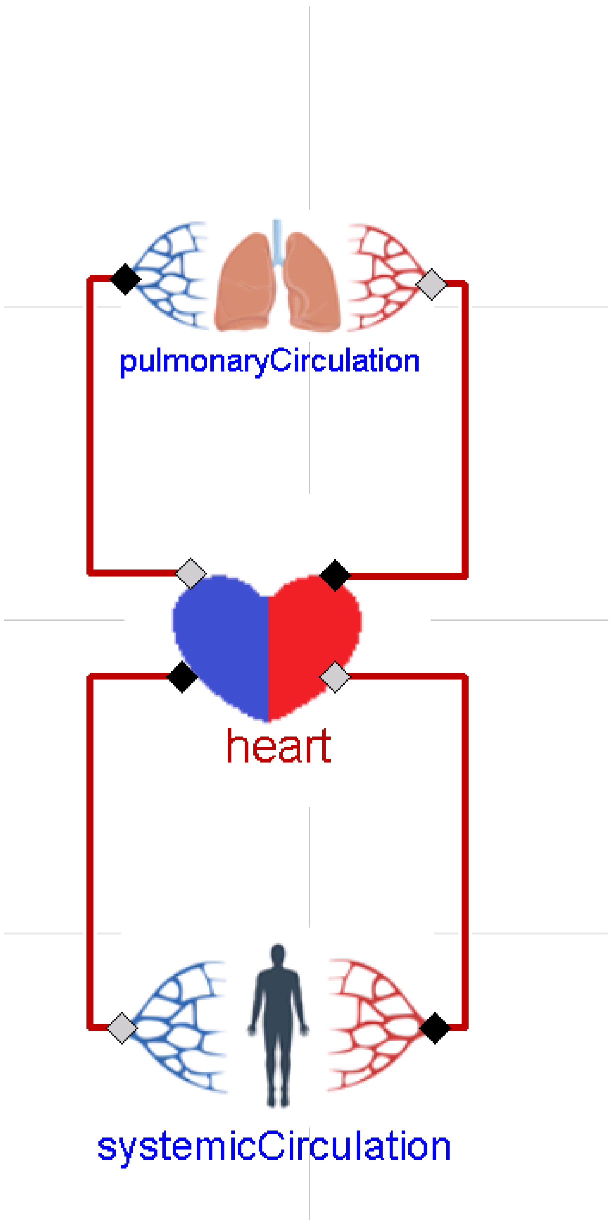
Ventricula Elastance EMAX = 4 mmHg/ml







Example – model of cardiovascular system



Ventricula Elastance EMAX = 4 mmHg/ml

RRAIN – resistance of vena cava=0.003 mmHg*s/ml

EITHV – elastance of vena cava = 0.0182 mmHg/ml

…

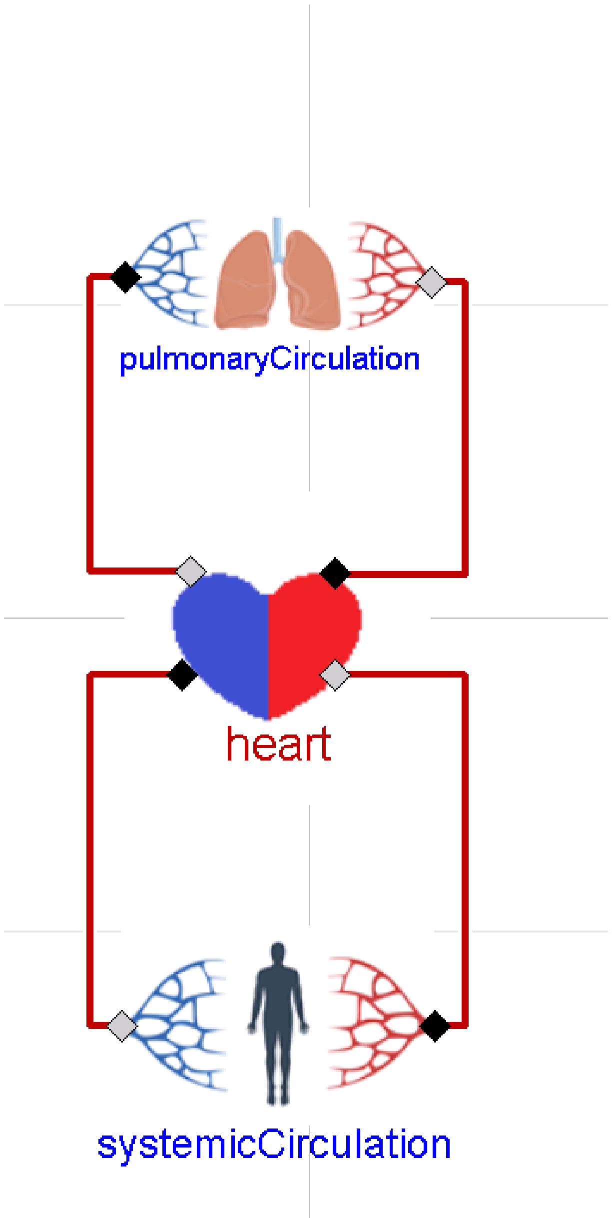
[7] J. A. Goodwin, W. L. van Meurs, C. D. Sá Couto, J. E. W. Beneken, S. A. Graves, A model for educational simulation of infant cardiovascular physiology., Anesthesia and analgesia 99 (6) (2004) 1655–1664, table of contents. doi:10.1213/01.ANE.0000134797.52793.AF.

[8] C. D. Sá Couto, W. L. van Meurs, J. A. Goodwin, P. Andriessen, A Model for Educational Simulation of Neonatal Cardiovascular Pathophysiology, Simulation in Healthcare 1 (Inaugural) (2006) 4–12.

[9] W. van Meurs, Modeling and Simulation in Biomedical Engineering: Applications in Cardiorespiratory Physiology, McGraw-Hill Professional, 2011.

Parameters of the model

Example – model of cardiovascular system



Ventricula Elastance EMAX = ?

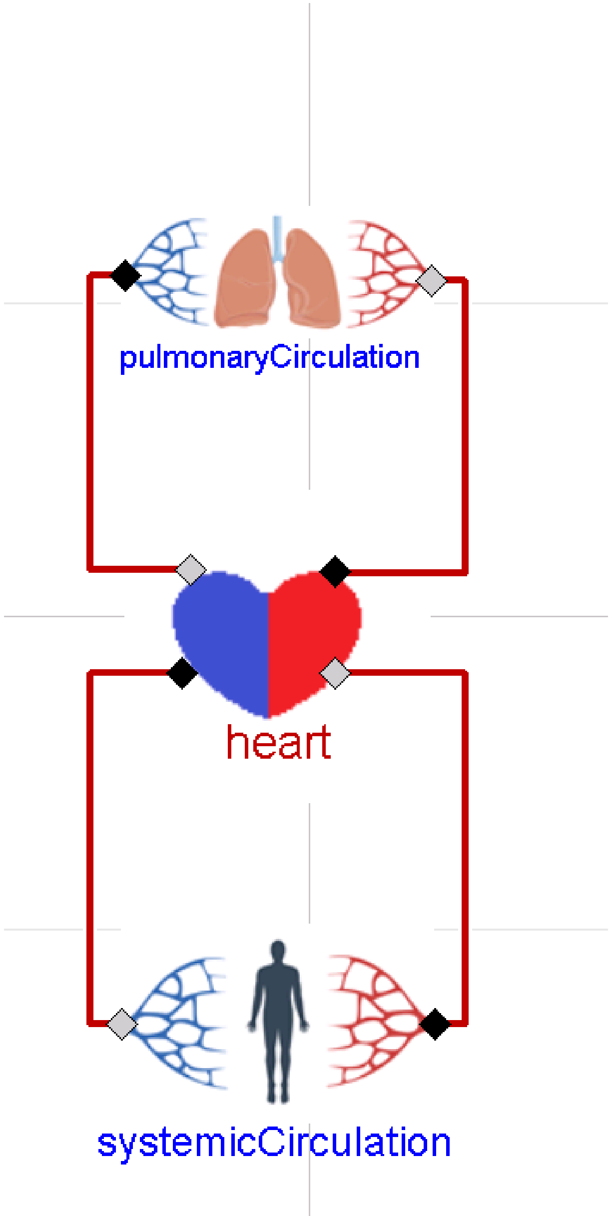
RRAIN – resistance of vena cava=?

EITHV – elastance of vena cava = ?

…

What are the values of the parameters, which fits to concrete human?

Example – model of cardiovascular system



Ventricula Elastance EMAX = ?

RRAIN – resistance of vena cava=?

EITHV – elastance of vena cava = ?

...

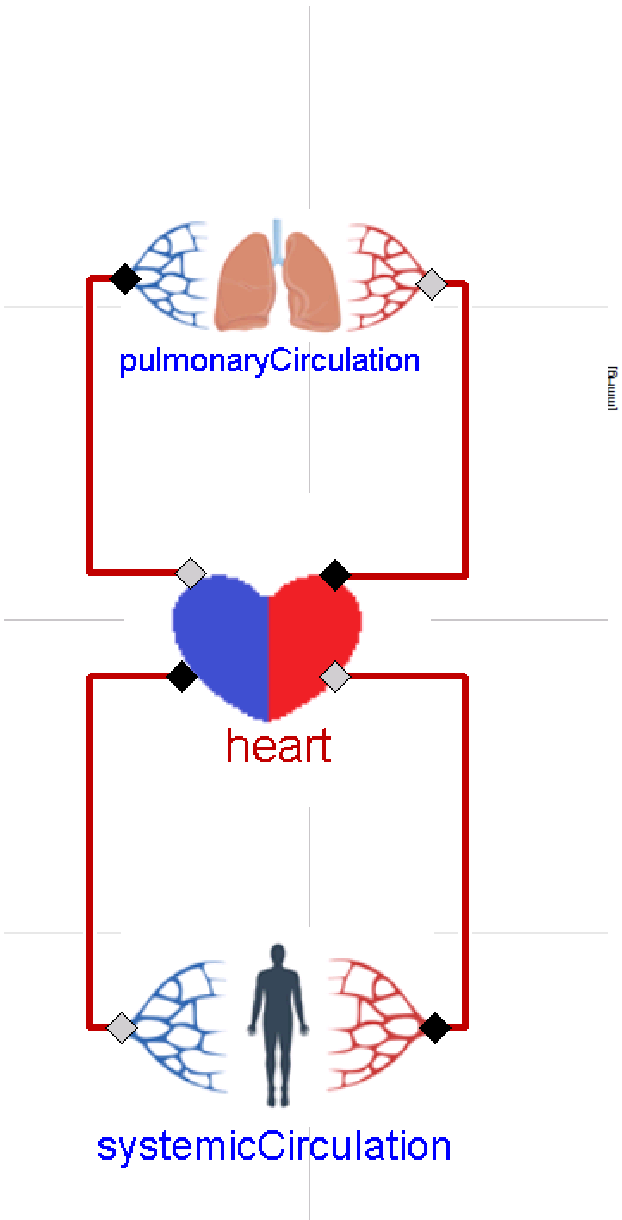
What are the values of the parameters, which fits to concrete human?

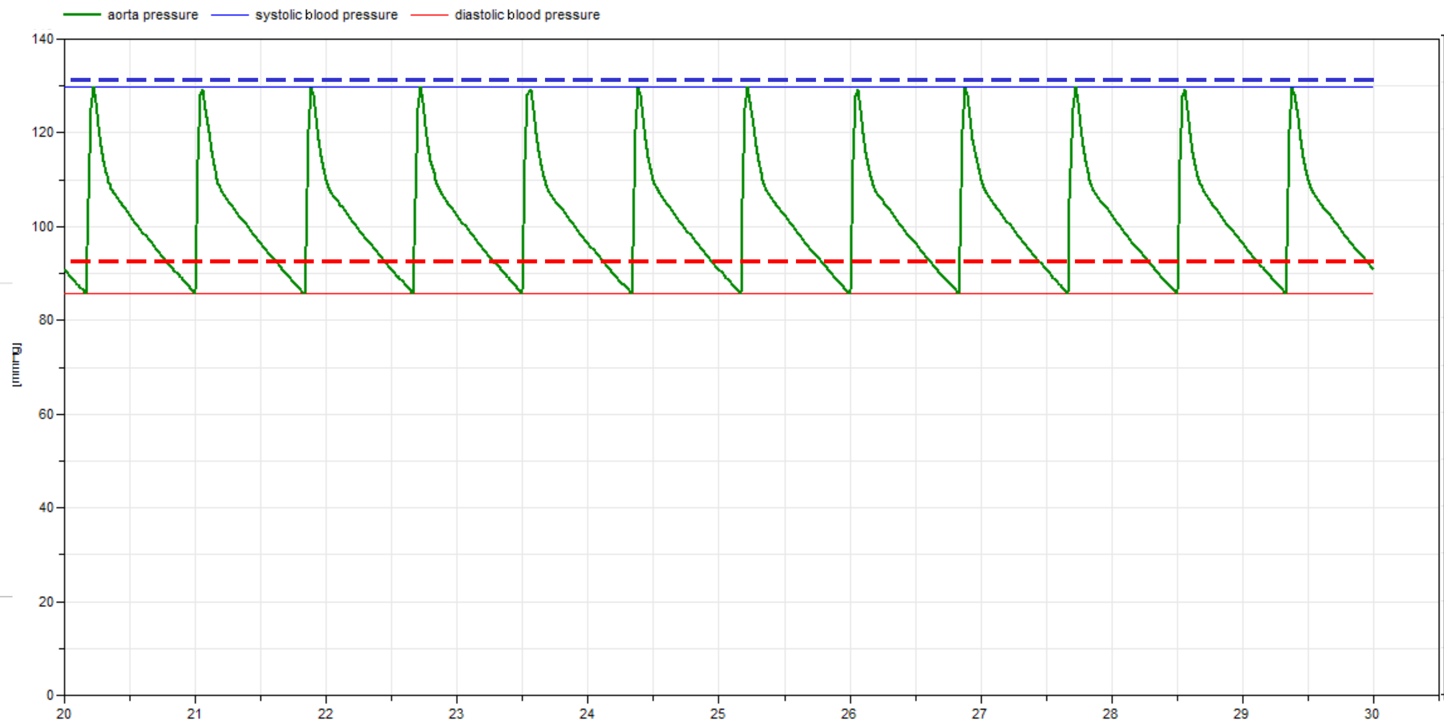
- Hard(invasive) to measure directly
-Measure other model variable
-compute parameter, fit simulation with measured data

„parameter identification“

„Model calibration“

Example – model of cardiovascular system





$$f(\vec{p}) = \sum_{i=1}^{n} (M(t_i, \vec{p}) - d(t_i))^2 \rightarrow min$$

Repeat until simulation gives similar result as the measured data

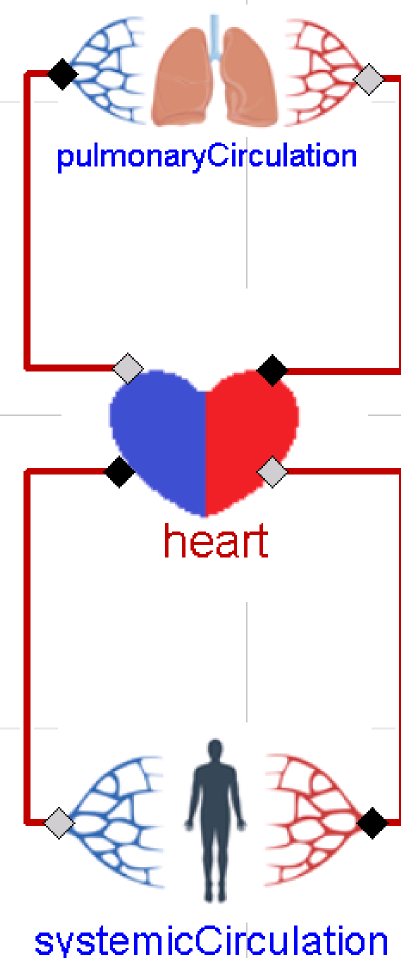
Computational physiology - results

Model of hemodynamics of cardiovascular system

Meurs et al. (2006-2011), Burkhoff et al.(1997-2013) 7), Fernandez de Canete et al. (2013), …

5 hours in local cluster
30 minutes in EGI cloud (30 CPU)

heart.heart.leftHeart.ventricularElastance.factor	0.276715208845232
pulmonaryCirculation.EPA.factor	2.42977968242538
pulmonaryCirculation.RPP.factor	0.326200900501185
pulmonaryCirculation.RLAIN.factor	7.03415309534463
heart.heart.leftHeart.RxAOutflow.factor	0.0579461745835701
heart.heart.leftHeart.RxVOutflow.factor	0.115592156812037
heart.heart.rightHeart.ventricularElastance.factor	0.00138825200390164
heart.heart.rightHeart.RxVOutflow.factor	0.0339999571366457
systemicCirculation.systemicVeins.RETHV.factor	21.5099238722829
systemicCirculation.systemicVeins.EITHV.factor	0.183594380817957
systemicCirculation.systemicVeins.VITHVU.factor	0.00115139326808882
systemicCirculation.systemicVeins.RRAIN.factor	29.4603401420539
systemicCirculation.systemicPeripheralVessels.RTA.factor	2.58765141868037
systemicCirculation.systemicPeripheralVessels.RSP.factor	0.744428239928714
systemicCirculation.systemicArteries.EETHA.factor	0.318031878463708
systemicCirculation.systemicArteries.RETHA.factor	0.0792556174368193
systemicCirculation.systemicArteries.EITHA.factor	0.416618612811922



Computational physiology - results

Model of saturation O2 in Hemoglobin

- Data Imai (1972), Roughton(1967), …
- Theory:

$$\boldsymbol{sO_2} = \frac{\mathrm{K_4}\left(\mathrm{K_3}\left(\mathrm{K_2}(x)+2x^2\right)+3x^3\right)+4x^4}{4\left(\mathrm{K_4}\left(\mathrm{K_3}\left(\mathrm{K_2}(\mathrm{K_1}+x)+x^2\right)+x^3\right)+x^4\right)}$$

$$\boldsymbol{x} = \frac{fn_D}{fn_O} \times \boldsymbol{pO_2}$$

…

- Find K1 … K4 to fit the data

K1=K2 = 0.1301628020369
K3 = 0.744154273954473
K4 = 32.4804162428542

200 000 simulations in local cluster in 20 minutes

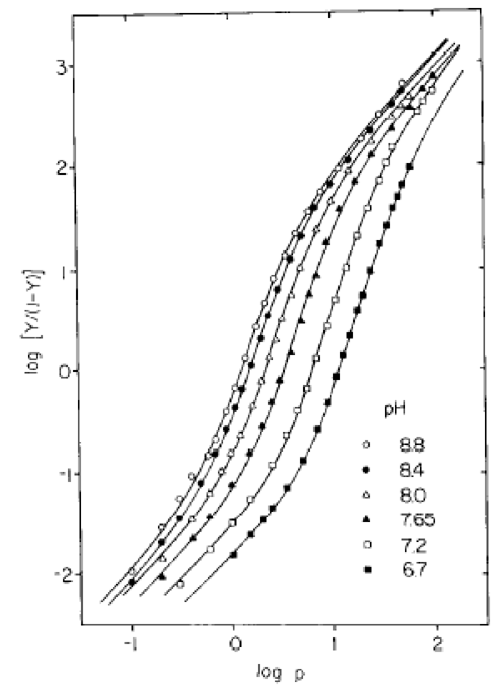
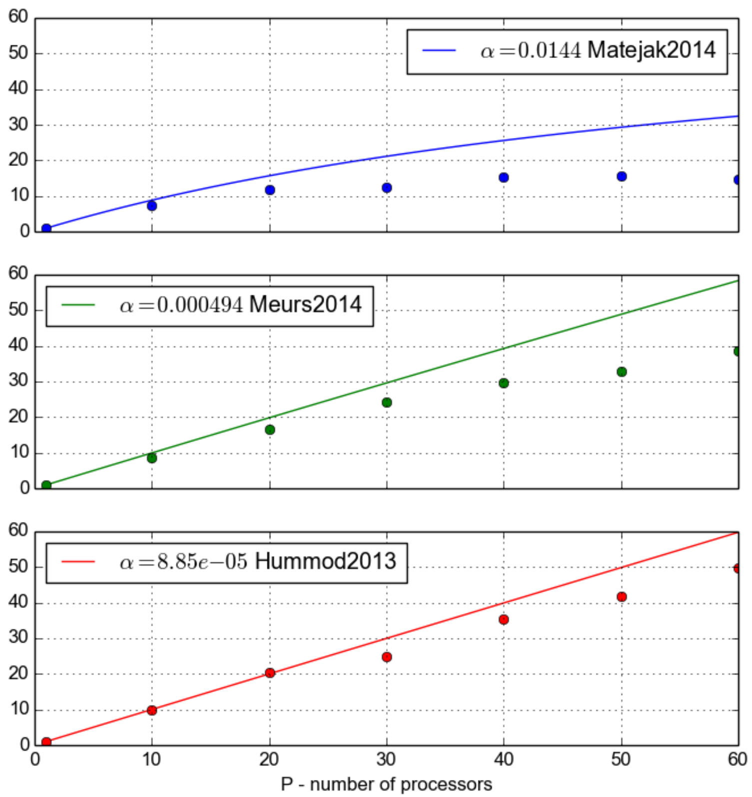


FIG. 2. Hill plots of oxygen binding by hemoglobin. Y, fractional oxygen saturation of hemoglobin; p, partial oxygen pressure in mm Hg. Heme concentration, 60 μM; 20°, in 0.1 M phosphate buffer. Points were experimentally obtained. Curves were calculated using estimated values of the Adair constants which are listed in Table I. Some of the data are not presented in this figure in order to avoid overcrowding.

Marek Mateják, **Tomáš Kulhánek**, and Stanislav Matoušek. Adair-Based Hemoglobin Equilibrium with Oxygen, Carbon Dioxide and Hydrogen Ion Activity. Scandinavian journal of clinical and laboratory investigation, 2014. doi:10.3109/00365513.2014.984320.
IF(2013):2.009

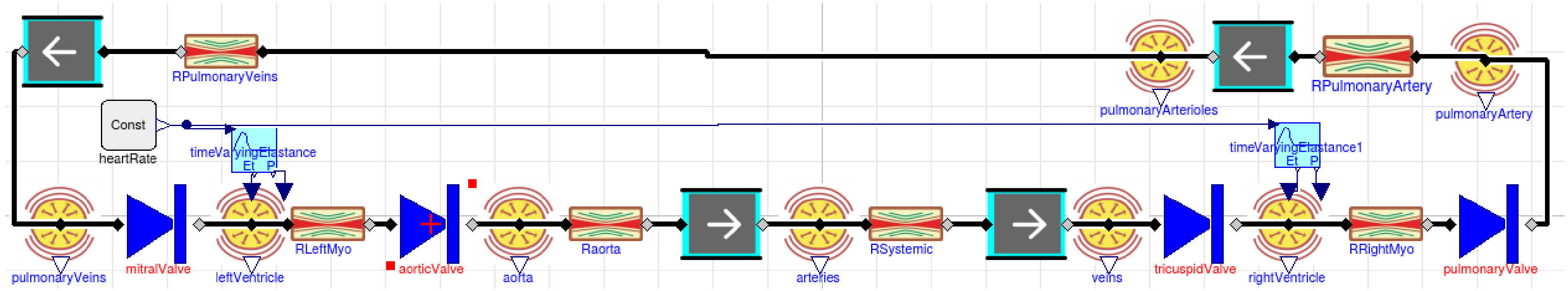
Computational physiology – results on cloud



- Theoretical speedup (Amdahl law) vs. Practical speedup on 60 CPU
- Complex – Hummod2013 – 49.8x time reduced from 4 days to 2 hours
- Medium – Meurs2014 – 38.7x
- Simple – Matejak2014 – 14.7x

Medium and complex simulation of models benefits highly from parallel parameter estimation, simple models outperforms on local HPC server

Computational physiology – modeling methodology for pulsatile CVS models



Tomáš Kulhánek, Jiří Kofránek, and Marek Mateják. Modeling of short-term mechanism of arterial pressure control in the cardiovascular system: Object oriented and acausal approach. Computers in Biology and Medicine, 54(1):137–144, 2014. **IF(2013):1.475**

Tomáš Kulhánek, Martin Tribula, Jiří Kofránek, and Marek Mateják. Simple models of the cardiovascular system for educational and research purposes. MEFANET Journal, 2(2):56–63, 2014.

Problem statement

- Research questions:
  - *Is it (grid-computing or cloud-computing) beneficial and how?*

    For sharing medical images – provides robust alternative to current systems of medical image exchange.

    For voice science – ability to connect users from broader geographically dispersed area.

    For computational biology – significant speedup gain in parameter studies – parameter estimation for medium and complex models which are hard to perform locally. For simple models simulation on local server outperforms cloud deployment.

Problem statement

- Research questions:
  - *What are the limitations?*

    Cost of porting and integration effort.

    - Grid-computing – high → reuse existing application
    - Cloud-computing – medium → virtual environment for already familiar application

    Communication overhead.

    - For simple models and small sized tasks degrades possible speedup

Problem statement

- Research questions:
  - *What can be the direction of future biomed. Research and grid-computing or cloud-computing research?*

    Web services with high added values

    - long data preservation, database of physiologically relevant parameters (http://www.physiovalues.org)
    - Web application for in-silico simulatio

    In silico trails to improve drug discovery and therapy testing. (Avicena, VPH)

    Deep parameter study with volunteer computing project – Physiome@home

    Lumped parameter models → 3D models

Summary

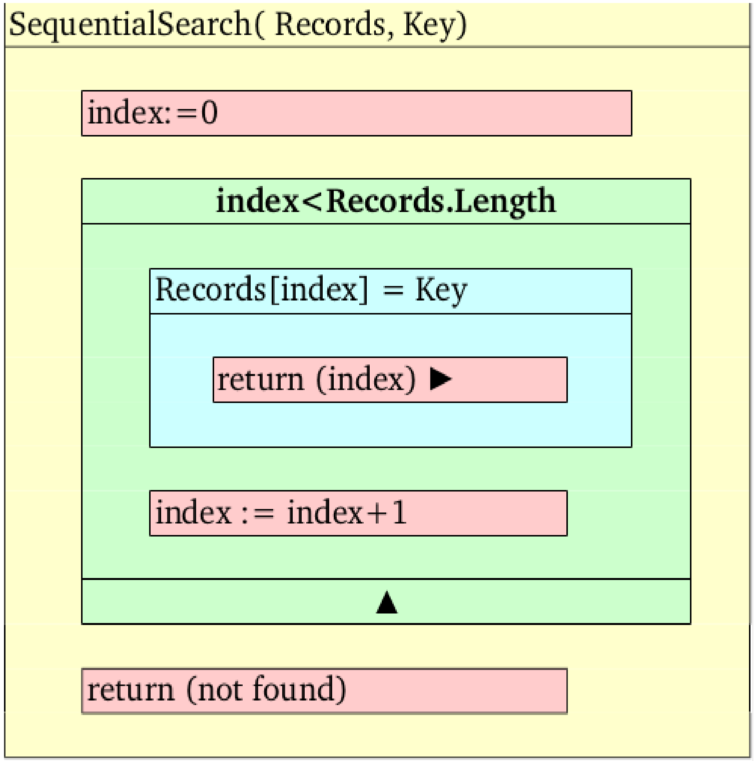
- Studied several grid-computing and clod-computing approaches in several biomedical use-cases
- Identified specific field which will benefit from cloud-computing mostly (parameter study of physiological models)
- Improved modeling methodology to build complex models of pulsatile CVS.

Thank you for your attention

Algorithm

- Order of instructions to solve task.
- Example: sequential search

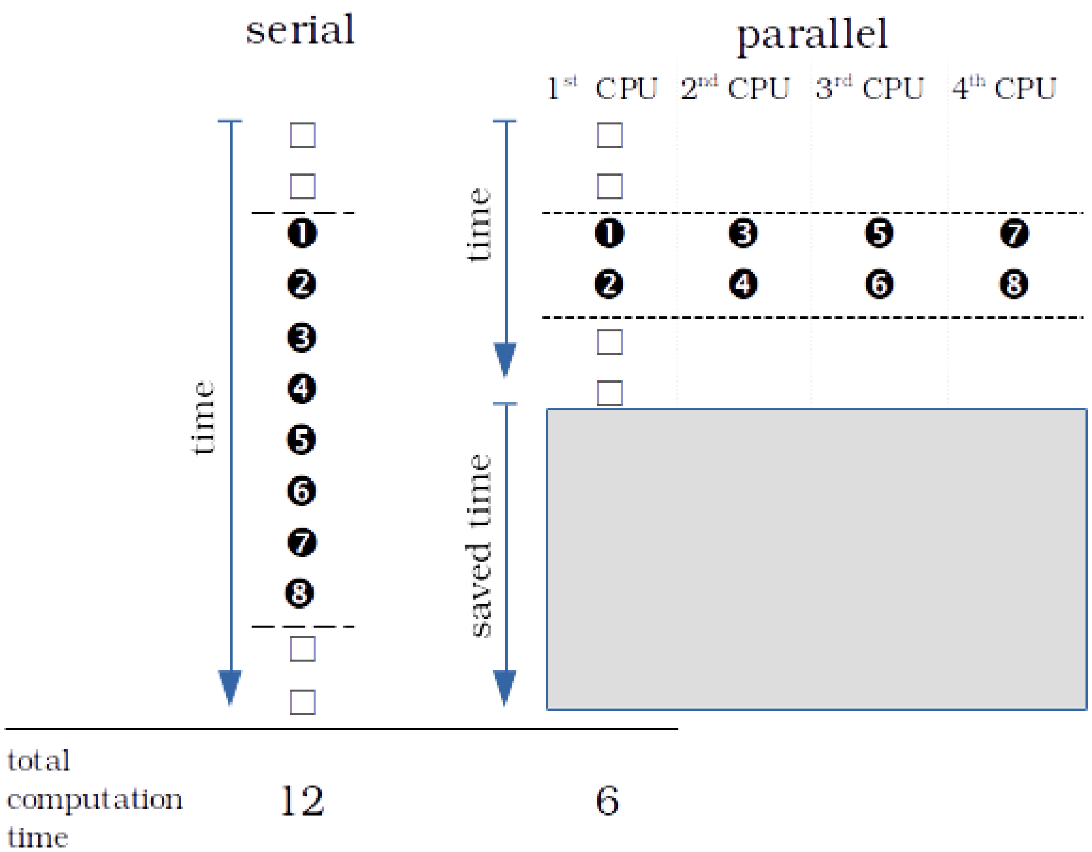
```
int function SequentialSearch(Array Records, int Key)
{


    for (int index=0;i<Records.Length;i++)
    {
        if (Records[index]=Key)
        {
            return (index)
        }



    }

    return(-1)
}
```



Efectivity, time complexity

- Sequential search $T = O(n)$
- Depth first search, exponential complexity $T = O(k^n)$
-

Paralelism



serial
parallel
1st CPU
2nd CPU
3rd CPU
4th CPU
time
time
saved time
total computation time
12
6

Genetic algorithm