



Výzkumné aktivity Centra CERIT-SC, spolupráce s partnery

Tomáš Rebok

Centrum CERIT-SC

Ústav výpočetní techniky Masarykovy univerzity

(rebok@ics.muni.cz)

CERIT-SC – výzkumné aktivity Centra

Hlavní výzkumné pilíře:

I. Výzkum v oblasti e-Infrastruktur

- **e-infrastrukturní/IT výzkum** – zaměřený na principy a technologie e-Infrastruktury a její optimalizaci

II. Výzkumné spolupráce s uživateli e-Infrastruktury (= partnerny)

- **kolaborativní výzkum** – synergický posun informatiky a spolupracujících věd

– **dostupnost odborníků/konzultantů** jak teoretického (Fakulta informatiky MU), tak praktického (CESNET) zaměření

– snaha o maximální **zapojení studentů**

- bakalářského -> **magisterského** -> **doktorského** studia
- nejen úzce zaměřená a dedikovaná pracovní síla, ale především
 - **výchova nových odborníků** v oblasti e-infrastruktur
 - **výchova erudovaných uživatelů** e-infrastruktury

CERIT-SC – výzkumné aktivity Centra

Formy výzkumu/spolupráce

I. Participace na projektech:

- **e-infrastrukturní/IT projekty** (úzká spolupráce s CESNET/MetaCentrum NGI)
 - projekty zaměřené na **vylepšování služeb a technologií e-infrastruktury**
 - *DataGrid, EGEE, EMI, EGI InSPIRE, EUAsiaGrid, CHAIN, Thalamos, ...*
 - **aktivní participace** (výzkumná i organizační – *EGI Council Chair*)
- **kolaborativní projekty**
 - participace a podpora **projektů spolupracujících věd** (výzkumných partnerů)
 - **návrh a vývoj nových metod, algoritmů a principů pro realizaci výzkumných infrastruktur a top-level výzkumu**
 - **výpočetní a úložné kapacity + know-how pro práci s nimi**
 - *ELIXIR-CZ, BBMRI, Thalamos, SDI4Apps, Onco-Steer, CzeCOS/ICOS, ...*
 - *KYPO, 3M SmartMeterů v cloudu, MeteoPredikce, ...*



CERIT-SC – výzkumné aktivity Centra

Formy výzkumu/spolupráce

II. Výzkumné aktivity:

- **e-infrastrukturní/IT výzkum** (*úzká spolupráce s CESNET/MetaCentrum NGI*)
 - výzkum a vývoj nástrojů, technologií a služeb pro oblast e-infrastruktur
- **kolaborativní výzkum**
 - výzkum **ve spolupráci s uživateli / výzkumnými partnery**
 - (týmy i jednotlivci)
- **často přechází v projektový výzkum/spolupráci**
- **příklady výzkumu/výzkumných spoluprací – viz dále**

e-Infrastrukturní/IT výzkum

Rozvrhový plánovač I.

Navržen a vyvinut nový plánovač nahrazující dosavadní frontový

- návrh realizován v rámci disertační práce
- **experimentální nasazení** od července 2014

Hlavní funkce:

- vytváří se plán (rozvrh) spouštění úloh
- možná predikce doby spuštění/čekání
- zaplňování „děr“ v rozvrhu vhodnými úlohami
- vyšší vytížení infrastruktury
- optimalizace rozvrhu vzhledem ke zvoleným kritériím (čekání, férovost, ...)

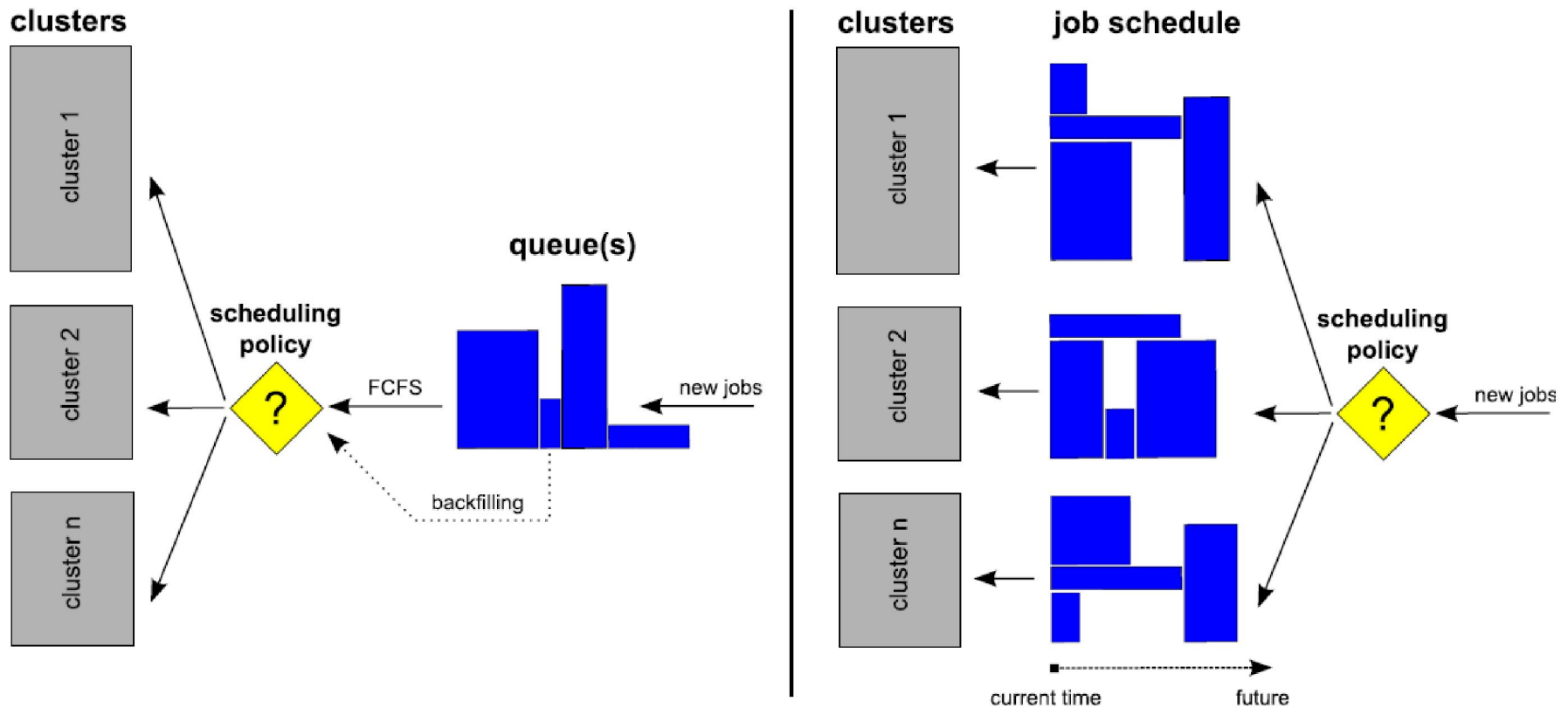
Dílčí výsledek práce: simulátor plánování/běhu úloh

- usnadnění simulace budoucích plánovacích mechanismů



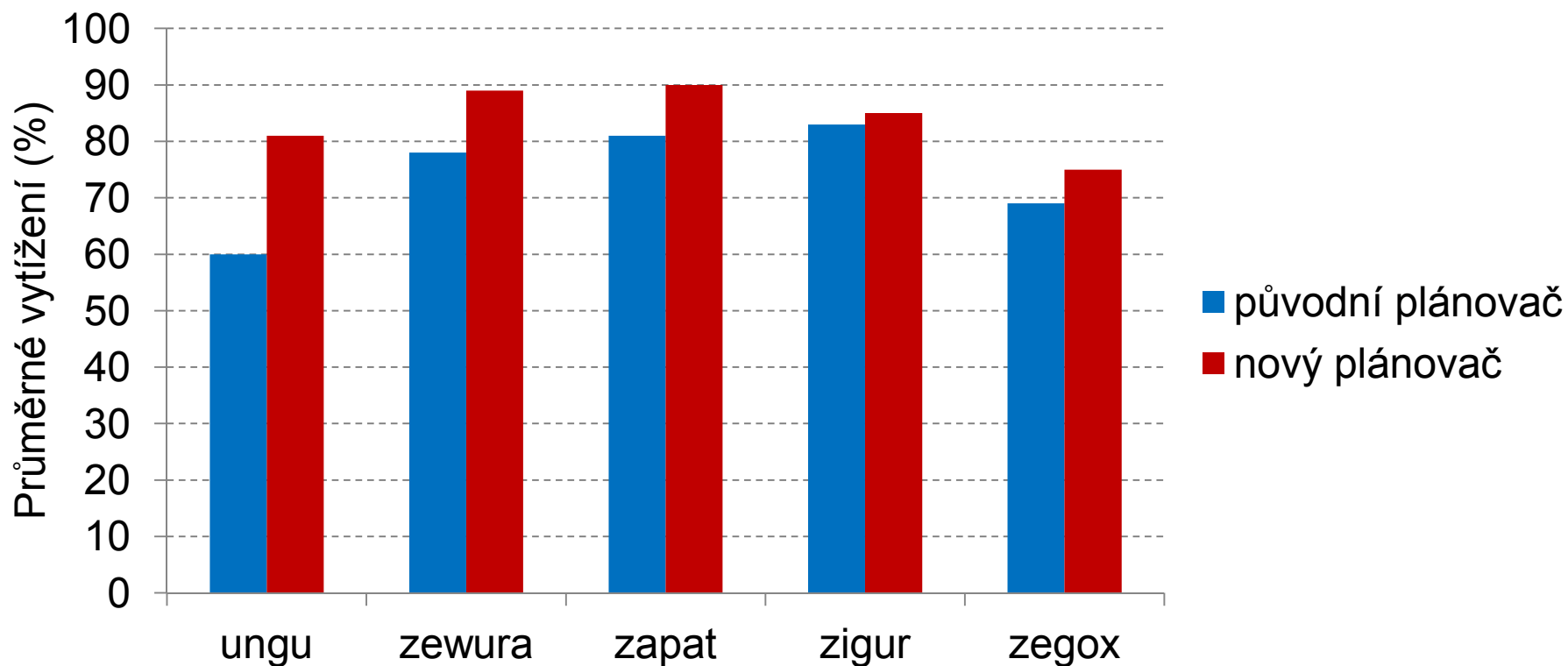
Rozvrhový plánovač II.

Frontový (vlevo) vs. Rozvrhový plánovač (vpravo)



Rozvrhový plánovač III.

Zlepšení vytížení strojů v CERIT-SC (data za rok 2014)



■ „Férové“ plánování I.

Férové plánování

- **cíl: zajištění rovnoměrného rozložení využití zdrojů v heterogenním prostředí výpočetního gridu**
- probíhající disertační práce

Fairshare:

- dynamická priorita založená na historické spotřebě zdrojů
 - s propočítaným časem priorita klesá (aktuální spotřeba zdrojů má vyšší váhu než historická)
 - navyšování priority při vykázání publikací s poděkováním MetaCentru/CERIT-SC

„Férové“ plánování II.

„Cena“ úlohy – dosavadní přístup:

- závislý na aktuálním stavu systému
- nezohledňuje heterogenitu infrastruktury
- jediná metrika: propočítaný čas

„Cena“ úlohy – nový přístup:

- výpočet se řídí výhradně **požadavky úlohy** (nezávislost na aktuálním stavu systému)
- úlohy požadující **cennější zdroje** „platí“ víc
- výpočet používá **dominantní zdroj** (CPU, paměť, ...)

$$cost = queue\ cost * \min(node\ cost) * \min(machine\ spec)$$

$$\max\left(\frac{requested\ cores}{machine\ cores}, \frac{requested\ memory}{machine\ memory}\right) * machine\ cost$$

■ Správa uživatelů/služeb – Perun I.

Perun

- systém pro **správu identit, skupin a přístupu na služby**
- integrovatelný do existujících prostředí, kde funguje jako **konsolidátor uživatelů a skupin**
- **pracuje s různými typy identit** – federovaná a sociální (Facebook, Google) identita, digitální certifikáty apod.
- vyvíjen v **úzké spolupráci Centra CERIT-SC se sdružením CESNET (NGI MetaCentrum)**



Správa uživatelů/služeb – Perun II.

- většina služeb české eInfrastruktury a je spravovaná systémem Perun
- systém je úspěšně nasazován i v cizině
 - Malaysia (Sifulan), Nigeria (NgREN), South Africa (SAGRID), Maroco, Italy (GARR), EGI - core service, ...

The screenshot displays the Perun web interface. At the top left is the Perun logo and the MetaCentrum instance name. The user is identified as Tomáš Rebok, a VO/Group/Facility Manager. The main navigation menu on the left includes options for VO manager, Group manager, Facility manager, and User. The central area shows the 'MetaCentrum' group details, including a tabbed interface with 'Groups' selected. A table lists various groups with their names and descriptions. The footer contains contact information and version details.

Name	Description
members	Group containing VO members
2LFUK-Sopko	Skupina Dr. Bruno Sopka (2. LF UK, Motol)
amber	skupina amber
astro	astronomové, Ústav fyziky a astrofyziky PřF MU
ceitec	Central European Institute of Technology
CEITEC-BIOINF	Research group Bioinformatics - CEITEC
ceitec:strbio	Structural Biology
cpg	Cluster Physics Group - Ústav geoniky AV ČR
cybernet	Skupina katedry kybernetiky ZČU v Plzni
czdi	Centrum zpracování přirozeného jazyka FI MU

■ GPU akcelerátory

Výpočty na GPU kartách

- uplatnění pro širokou škálu aplikací (vyšší aritmetický výkon a paměťová propustnost)
- **výpočetní kernely** (= funkce spouštěné na GPU) pracují (vstup/výstup) s pomalou globální pamětí
 - kernely sdílející data tyto přenáší zbytečně
 - **fúze těchto kernelů** může výpočet významně urychlit
- navržena metoda a prototyp kompilátoru pro **automatické spojování těchto kernelů**

Kolaborativní výzkum

Rekonstrukce stromů I.

Rekonstrukce individuálních stromů z laserových skenů

- **partner:** *Centrum výzkumu globální změny AV ČR (CzechGlobe)*
- **cíl projektu:** návrh algoritmu pro rekonstrukci 3D modelů stromů
 - z mraku nasnímaných 3D bodů
 - strom nasnímán laserovým snímačem LiDAR
 - výstupem jsou souřadnice XYZ + intenzita odrazu
 - *očekávaný výstup:* 3D struktura popisující strom
 - identifikovat **základní strukturální prvky** (kmen a hlavní větve)
 - *primární zaměření:* smrky
- **hlavní problémy:** překryvy (mezery v datech)



Rekonstrukce stromů I.

Rekonstrukce individuálních stromů z laserových skenů

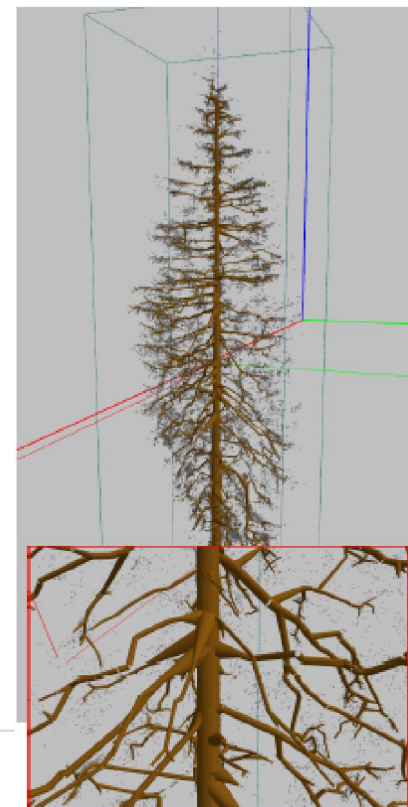
- **partner:** *Centrum výzkumu globální změny AV ČR (CzechGlobe)*
- **cíl projektu:** návrh algoritmu pro rekonstrukci **3D modelů stromů**
 - z mraku nasnímaných 3D bodů
 - strom nasnímán laserovým snímačem LiDAR
 - výstupem jsou souřadnice XYZ + intenzita odrazu
 - *očekávaný výstup:* 3D struktura popisující strom
 - identifikovat **základní strukturální prvky** (kmen a hlavní větve)
 - *primární zaměření:* smrky
- **hlavní problémy:** překryvy (mezery v datech)



Rekonstrukce stromů II.

Rekonstrukce stromů – nový přístup

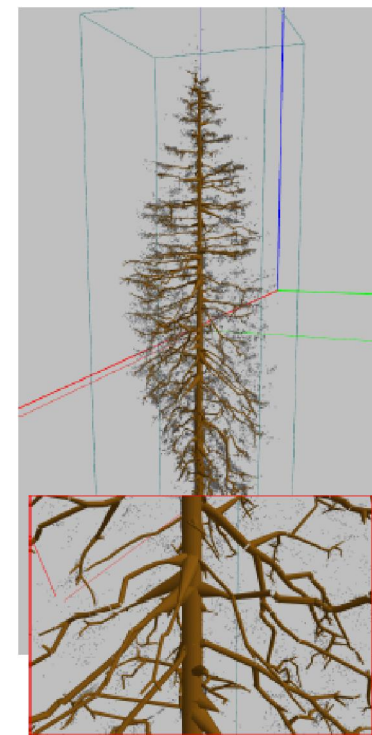
- v rámci DP navržena *inovativní metoda* rekonstrukce 3D modelů smrkových stromů
- **základní myšlenka metody:**
 - identifikace zřetelných a souvislých oblastí
 - rekonstrukce kmene
 - propojování identifikovaných souvislých oblastí
 - odhady tloušťky větví
 - závěrečné čištění modelu



Rekonstrukce stromů III.

Využitelnost rekonstruovaných modelů:

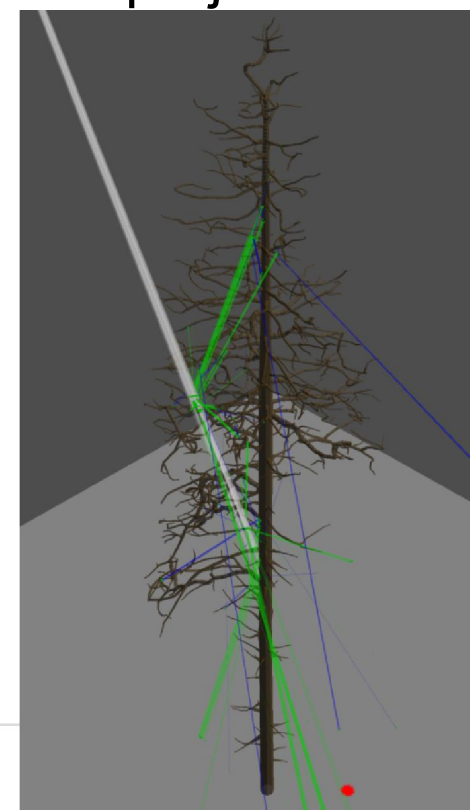
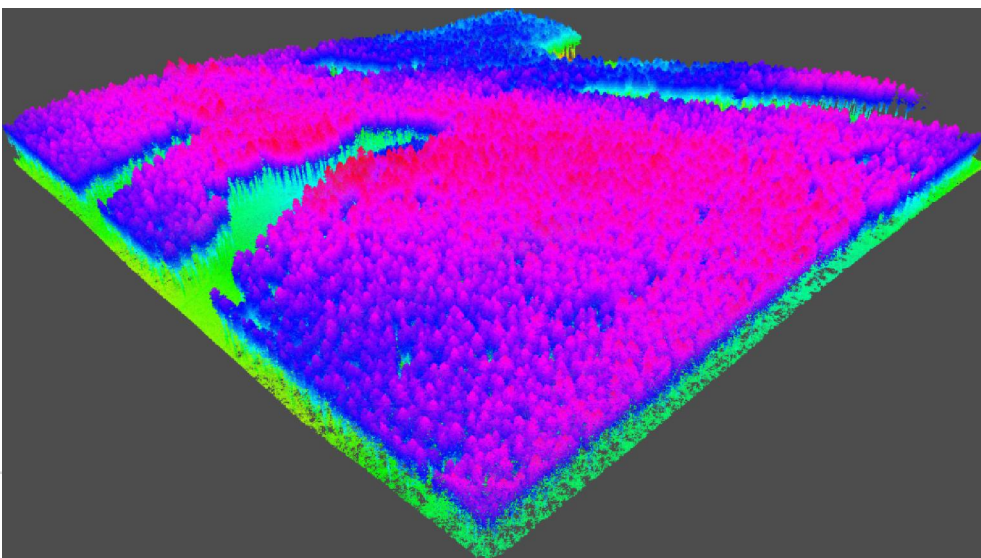
- **neinvazivní získávání statistických informací** o množství dřevité biomasy a o základní struktuře stromů
 - kmeny rekonstruovány s vysokou přesností
 - struktura větví rovněž poměrně věrná
- **parametrizované opatřování zelenou biomasou**
 - mladé větve + jehličky
 - návazný PhD výzkum
- **importování modelů do nástrojů** umožňujících analýzu šíření slunečního záření s využitím DART modelů
- **většina kroků algoritmu nezávislých na druhu stromu**
=> lze rozšířit i na další druhy



■ Rekonstrukce lesů I.

Rekonstrukce lesních porostů z full-wave LiDAR skenů

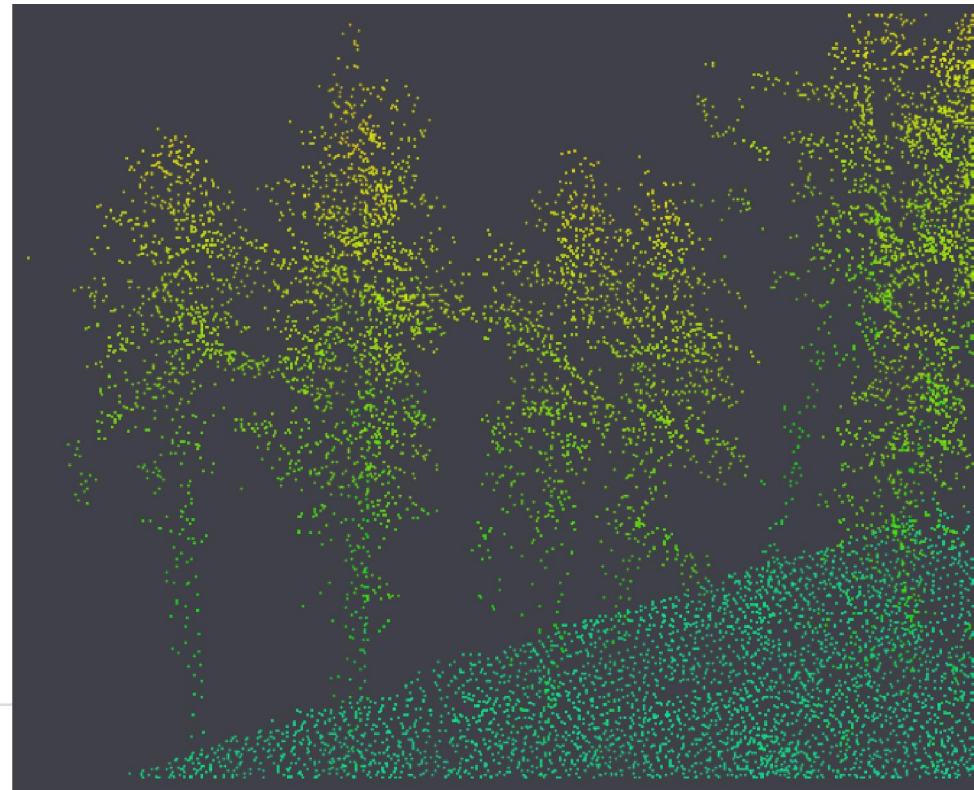
- „s jídlem roste chut“ 😊
- návazná PhD práce, příprava budoucího společného projektu
- **cíl: co nejvěrnější 3D rekonstrukce celých lesních porostů z leteckých full-wave LiDARových skenů**
 - možné využití hyperspektrálních skenů, termálních skenů, in-situ měření, ...



■ Rekonstrukce lesů II.

Rekonstrukce lesních porostů z full-wave LiDAR skenů

- skeny získávány leteckým snímáním
- **diametrálně odlišný problém** – extrémní množství bodů, které jsou však *mnohem řidší*
 - nastíněné algoritmy pro přesné rekonstrukce jednotlivých stromů **nelze aplikovat**
 - nutno revidovat i metody pro **vizualizaci a uložení dat/modelů**



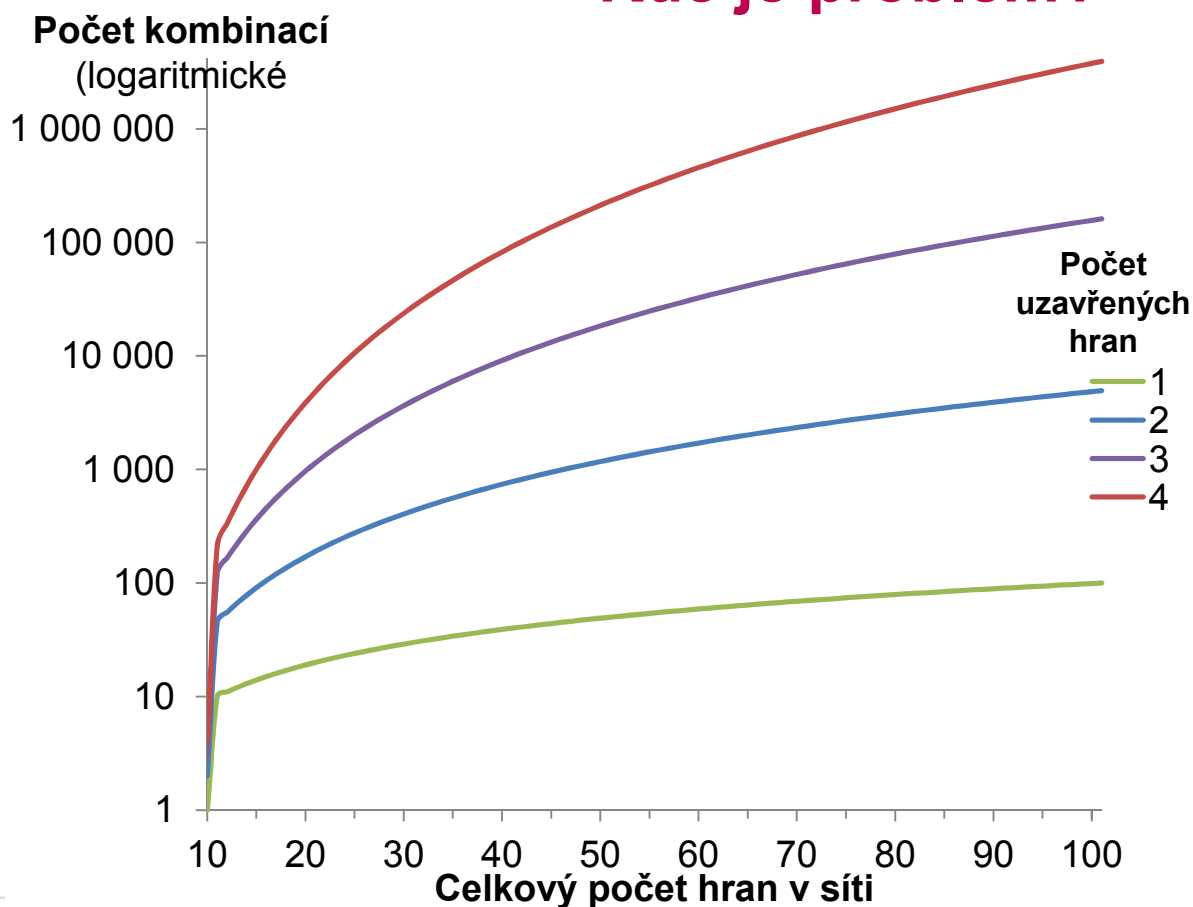
Identifikace problémových uzavírek I.

Hledání problematických uzavírek v silniční síti ČR

- **partner:** *Centrum Dopravního Výzkumu v.v.i., Olomouc*
- **cíl projektu:** nalezení metody pro identifikaci problémových uzavírek v silniční síti ČR (aktuálně Zlínského kraje)
 - identifikace uzavírek vedoucích (dle definovaných ohodnocovacích funkcí) k problémům v dopravě
 - převedený problém: **nalezení všech rozpadů grafu**
 - zjednodušený problém: **nalezení všech rozpadů grafu generovaných N hranami**
- **hlavní problémy:** výpočetní náročnost (NP-těžký problém)
 - přístup „hrubou silou“ selhával již při uzavření 3 hran

Identifikace problémových uzavírek II.

Kde je problém?



Sít' Zlínského kraje

724 uzlů

974 hran

1. 974

2. 473 851

3. 153 527 724

4. 37 268 855 001

5. 7 230 157 870 194

...

Identifikace problémových uzavírek III.

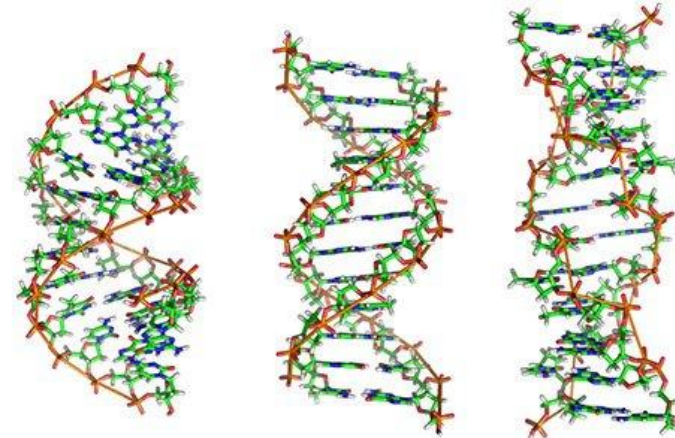
Hledání problematických uzavírek v silniční síti ČR

- v rámci DP navrženy 2 algoritmy založené na teorii grafů
 - za významné podpory grafových odborníků FI MU
- širší možnosti **konfigurovatelnosti výpočtu** (hrany vs. komponenty)
- lze upočítat všechny **rozpady generované až 4 hranami** (cca 1 den výpočtu) nebo **více hranami s omezením počtu komponent**
 - původně pouze 2 hrany!
- stále existuje **prostor pro vylepšení**

Korekce chyb a skládání genomu

Sekvenování *Trifolium pratense* (Jetel luční)

- partner: *Ústav experimentální biologie PŘF MU*
- cíl: optimalizace dostupných nástrojů pro skládání a opravy chyb v DNA kódech
 - *analýzy DNA (nejen) jetele vedou k výpočetně náročným problémům*
 - 50 GB vstup => **cca 500 GB potřebné paměti** (aplikace Echo)
 - existují **větší vstupy**
- v rámci DP **paralelizováno a optimalizováno až na cca 50% využití paměti**

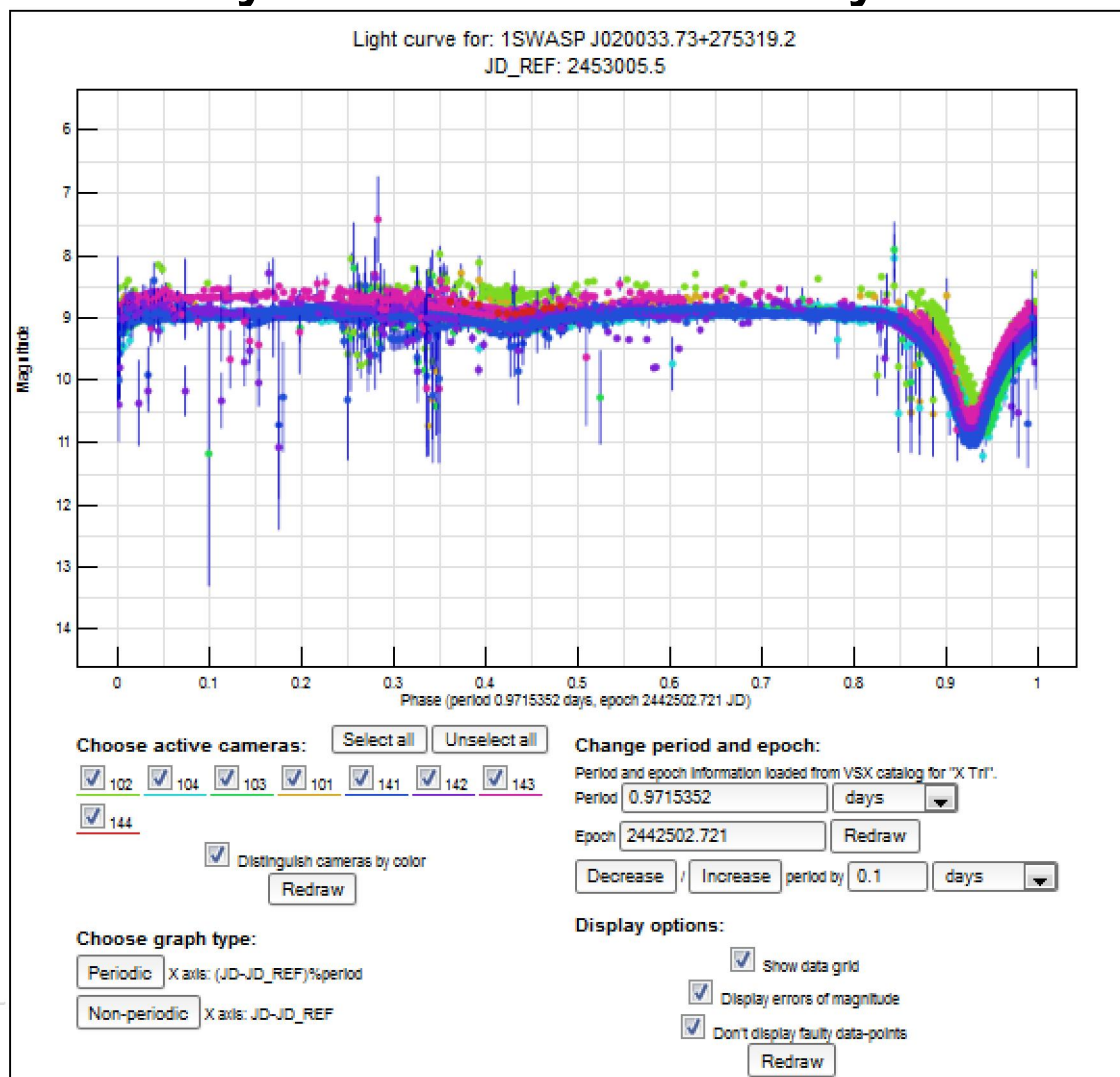


Fotometrický archiv astronomických snímků I.

Fotometrický archiv astronomických snímků

- **partner:** *Ústav teoretické fyziky a astrofyziky PŘF MU*
- **cíl projektu:** vytvoření a provoz portálu pro získávání dat o světelnosti proměnných hvězd (projekt SuperWASP)
 - databáze cca 18 miliónů hvězd
- **dosažené výsledky:**
 - portál v produkčním režimu: <http://wasp.cerit-sc.cz>
 - rozšířen o vykreslení grafu světelné křivky (DP práce)
 - provoz systému pro detekci hvězd v hvězdokupě:
<http://clusterix.cerit-sc.cz/>
 - archiv CCD snímků: <http://wasp.cerit-sc.cz/paw/>

Fotometrický archív astronomických snímků II.

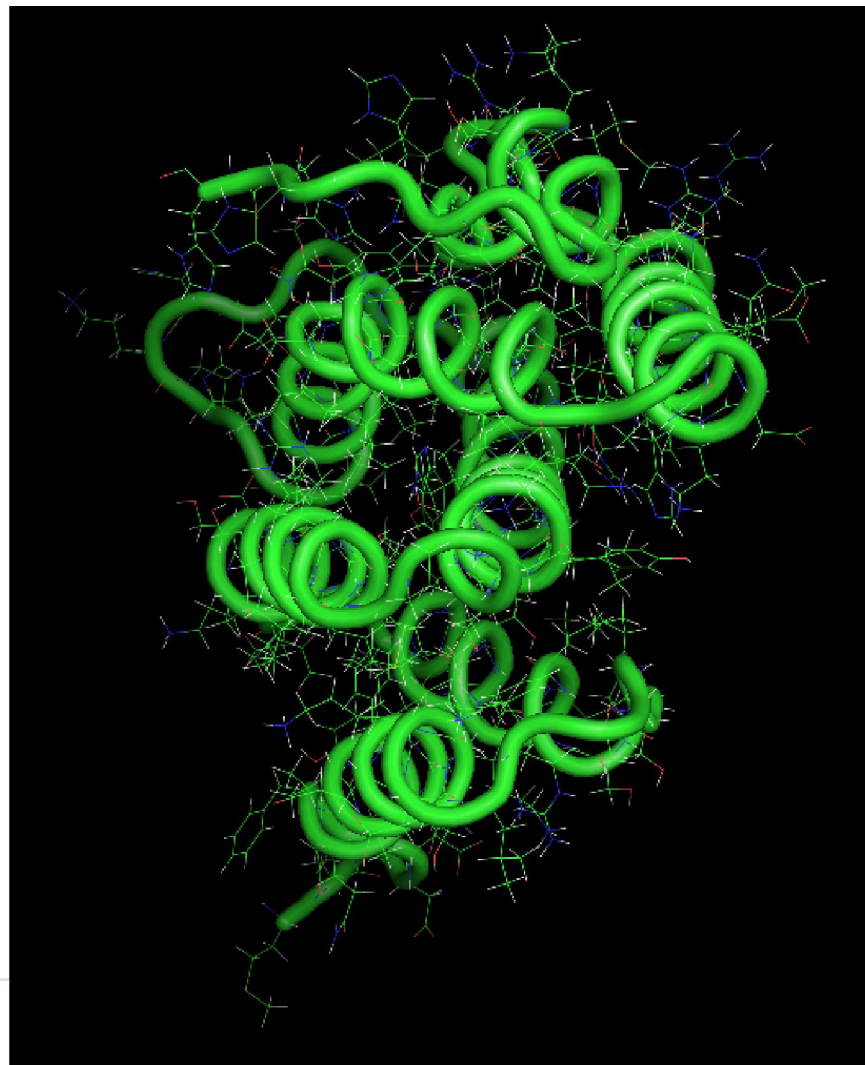
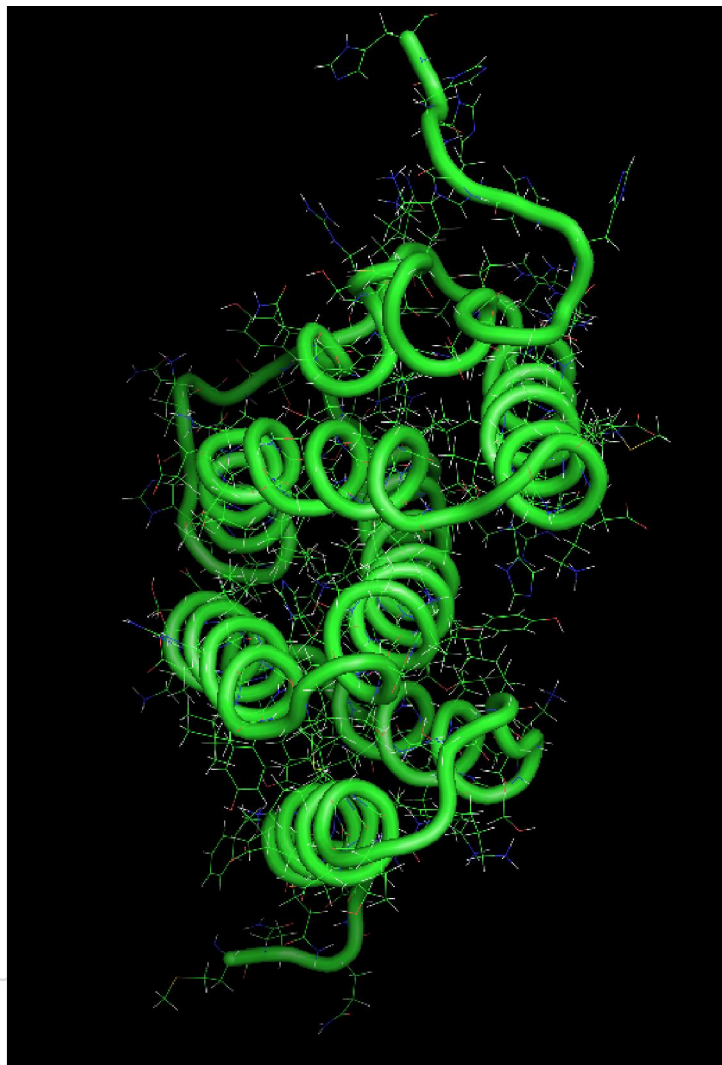


Výpočetní chemie a biochemie I.

Výpočet konformace molekul z řídkých NMR dat

- **partner:** *Středoevropský technologický institut (CEITEC)*
- **cíl projektu:** kombinované výpočetní zpracování výstupů několika nezávislých experimentálních metod (vedoucí ke zjištění tvaru molekuly určitého vzorku)
 - kombinace výstupů **molekulové dynamiky, NMR a SAXS** metod
 - existuje vyzrálý (i komerční) SW, avšak **složitý na použití**
 - náchylnost k chybám (při formulaci zadání)
 - složitost při kombinaci dat z různých zdrojů
 - **vlastní vývoj kombinovaných výpočetních metod** (rozšíření existujících nástrojů)
 - obohacení SW pro zpracování NMR o simulaci molekulové dynamiky
 - snaha vystačit s výsledky časově i finančně méně náročných variant exper.
 - aktuální výsledky ukazují na **mnohem realističtější geometrie rekonstruovaných molekul**
 - **prototypová implementace** ve stadiu vyhodnocení

Výpočetní chemie a biochemie II.

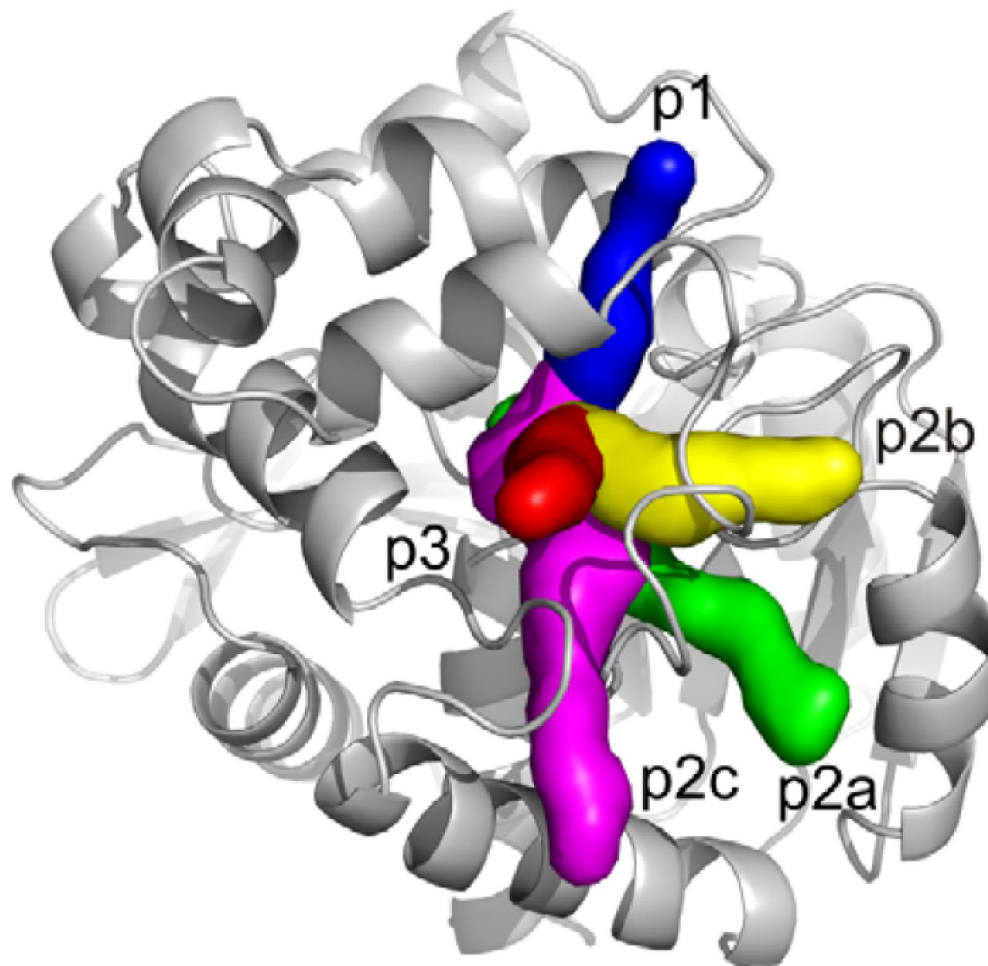


Výpočetní chemie a biochemie III.

Analýza transportních cest v proteinech

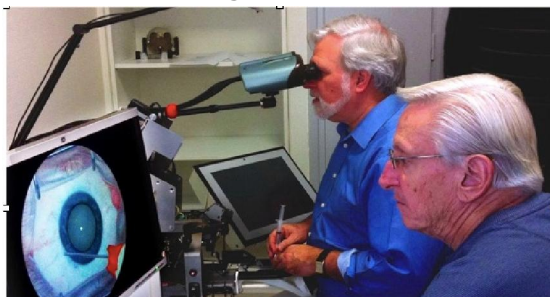
- partner: *Loschmidt Laboratories MU*
- cíl projektu: analýza možností transportu molekul ligandu (např. léčivo) na aktivní místa proteinů
 - tj. zajištění nejen kýženého účinku molekuly na protein, ale zejména ověření možností transportu této molekuly k aktivním místům proteinů
 - v současné době jsou metody analýzy transportu buď **nepřesné** nebo **velmi výpočetně náročné** (molekulová dynamika)
 - snaha o nalezení metody pro **analýzu energie nutné na průchod ligandu do proteinu** (vyhodnocení průchodnosti „tunelu“) **méně náročným způsobem**
 - zejména se zajištěním věrohodných/přesných výsledků
 - implementace ve stádiu prototypu, zatím bez plné automatizace

Výpočetní chemie a biochemie IV.

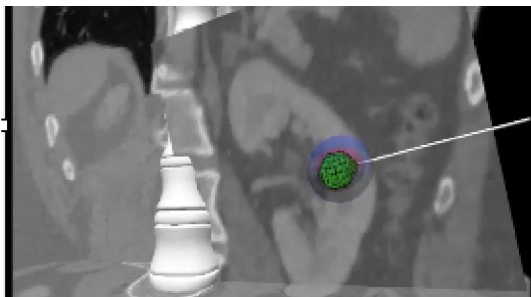


Modelování měkkých tkání v reálném čase I.

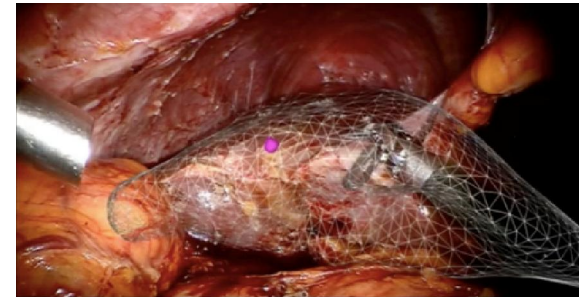
- Využití biomechanických modelů vytvořených z pre-operativních dat pacientů (CT, MRI) pro aplikace v medicíně
 - reálný čas [25Hz] nebo dokonce hmatová (haptická) interakce [$>500\text{Hz}$]



Simulátor operace kataraktu
MSICS



Kryoablace: plánování
umístění elektrody



Laparoskopie: vizualizace
vnitřních struktur

Chirurgické trenažéry

Pre-operativní plánování

Navigace během operace

2010

2014

2018

Simulace vyžadují kombinaci různých reprezentací objektů:

- **geometrie:** detekce kolizí, vizualizace, metriky pro verifikaci a validaci
- **fyzika:** realistické chování objektů, deformace, interakce mezi objekty

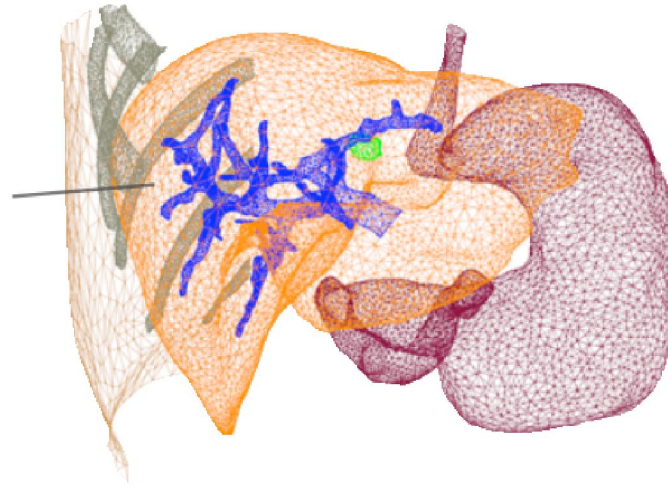
Modelování měkkých tkání v reálném čase II.

Nasazení v lékařské praxi

řešení reálných problémů,
metriky pro vyhodnocení
benefitu, robustnost,
kompatibilita s normami

Modelování interakcí

Modelování elastických
kontaktů
Simulace řezání, šití, vpichu
jehly
Haptická interakce



Mechanické a fyzikální modelování

metoda konečných prvků, mesh-
less metody, ale také
elektrofyzologie, heat-transfer

Numerické metody řešení

přímé a iterativní solvery, paralelní a
akcelerované algoritmy (např. GP-
GPU), interpolační metody a
generování sítí

Validace a verifikace modelů

správné řešení rovnic (porovnání se
standardním software), řešení
správných rovnic (porovnání s
realitou, experiment)

- **mezinárodní spolupráce** s instituty (IHU Strasbourg, INRIA France) a univerzitami (University of British Columbia, Koç University, Istanbul)
- **hledání partnerů pro spolupráci v rámci ČR** (biomechanické modelování, experimenty, kliničtí partneři)

Další spolupráce ...

- **Virtuální mikroskop, patologické atlasy**
 - *partner: LF MU*
- **Biobanka klinických vzorků (BBMRI_CZ)**
 - *partner: Masarykův onkologický ústav, Recamo*
- **Modely šíření epileptického záchvatu a dalších dějů v mozku**
 - *partner: LF MU, ÚPT AV, CEITEC*
- **Bioinformatická analýza dat z hmotnostního spektrometru**
 - *partner: Ústav experimentální biologie PřF MU*
- **Optimalizace Ansys výpočtu proudění čtyřstupňovou, dvouhřídelovou plynovou turbínou s chlazením lopatek**
 - *partner: SVS FEM*
- **3.5 miliónu „smartmeterů“ v cloudu**
 - *partner: Skupina ČEZ, MycroftMind*
- **Platforma pro poskytování specializovaných meteopredikcí pro oblast energetiky**
 - *partner: CzechGlobe, NESS, MycroftMind*
- ...

Závěrem

- Centrum CERIT-SC **nechce být** jen poskytovatelem HW zdrojů či dodavatel inforatické síly
 - výpočetní a úložné kapacity jsou **pouze nástrojem**
- **chceme být výzkumný partner:**
 - máme co nabídnout:
 - *netriviální/unikátní výpočetní a úložné zdroje nezbytné pro realizaci náročných výzkumných cílů*
 - *netriviální IT know-how + zázemí univerzity (doktorská škola)*
 - *zkušenosti s velkými národními/evropskými projekty*
 - společným výzkumem chceme pomoci s **překonáním stávajících limitů výzkumu** našich partnerů
 - a tím otevřít další možnosti „**top-level**“ výzkumu



EUROPEAN UNION

EUROPEAN REGIONAL
DEVELOPMENT FUND
INVESTING IN YOUR FUTURE



OP Research and
Development for Innovation

Projekt CERIT Scientific Cloud (reg. no. CZ.1.05/3.2.00/08.0144) byl podporován operačním programem *Výzkum a vývoj pro inovace*, 3 prioritní osy, podoblasti 2.3 *Informační infrastruktura pro výzkum a vývoj*.

<http://www.cerit-sc.cz>