

Centrum excellence IT4Innovations: superpočítání ve službách vědy

Ivo Vondrák et al.¹

Centrum

VŠB - Technická univerzita Ostrava

Abstrakt

Globálním cílem projektu IT4Innovations (IT4I) je vybudovat v České republice národní centrum excelentního výzkumu v oblasti informačních technologií. V rámci projektu bude vytvořeno výzkumné prostředí, včetně odpovídající infrastruktury, zaměřeného především na výzkum a vývoj výpočetních metod, souhrnně nazývaných jako High Performance Computing (HPC), s důrazem kladeným na jejich využitelnost v oblasti aplikovaných věd a rozvoji informační společnosti. V současné době je problematika výpočetně náročných úloh v České republice poměrně velmi roztržena a de facto neexistuje pracoviště, které by centrálně poskytovalo služby HPC výzkumné komunitě. Cílem projektu je tedy vytvořit na základně vysoce výkonné výpočetní infrastruktury, na úrovni TOP50 nejvýkonnějších superpočítačů, široké portfolio těchto služeb, počínaje výzkumem a vývojem nových metod a softwarové podpory paralelních algoritmů, přes návrh a implementaci výpočetně náročných úloh multidisciplinárního charakteru a poskytováním výpočetních kapacit konče.

1. Úvod

Záměrem realizace projektu IT4I a vybudování stejnojmenného Centra excellence IT4I je přispět ke zvýšení konkurenceschopnosti a životní úrovně obyvatelstva České republiky prostřednictvím výrazného rozvoje výzkumného a vývojového potenciálu v oblasti informačních technologií, výpočetní matematiky a s nimi souvisejících aplikovaných věd. V rámci projektu bude vytvořeno výzkumné prostředí, včetně odpovídající infrastruktury, zaměřené na rozvoj výpočetních metod jako takových - *IT jako předmět výzkumu*, tak především nástrojů pro jejich efektivní využití - *IT jako prostředek pro další výzkum či aplikační využití*. Výše uvedený globální cíl projektu je tak možno dále rozčlenit do těchto specifických cílů:

1. vytvořit *výzkumné zázemí*, které umožní vědecké bádání a realizaci výzkumu v oblasti výpočetních metod produkující mezinárodně uznávané výsledky;
2. vytvořit *inovativní prostředí*, které podpoří vývoj v předmětné oblasti, na němž má zájem participovat či o jehož výsledky má velký zájem aplikační sféra;
3. vybudovat *superpočítačové centrum*, které zajistí rozvoj kompetencí v oblasti řešení výpočetně náročných úloh.

Takto pojaté centrum rovněž vymezuje okruh cílových skupin, pro které je určeno. Z jedné strany se jedná o výzkumnou obec tvořenou uživateli z akademické sféry reprezentovanou univerzitami a pracovišti akademie věd. Z druhé strany pak subjekty z aplikační oblasti definované průmyslovými partnery, státní správou, samosprávou a jejich institucemi.

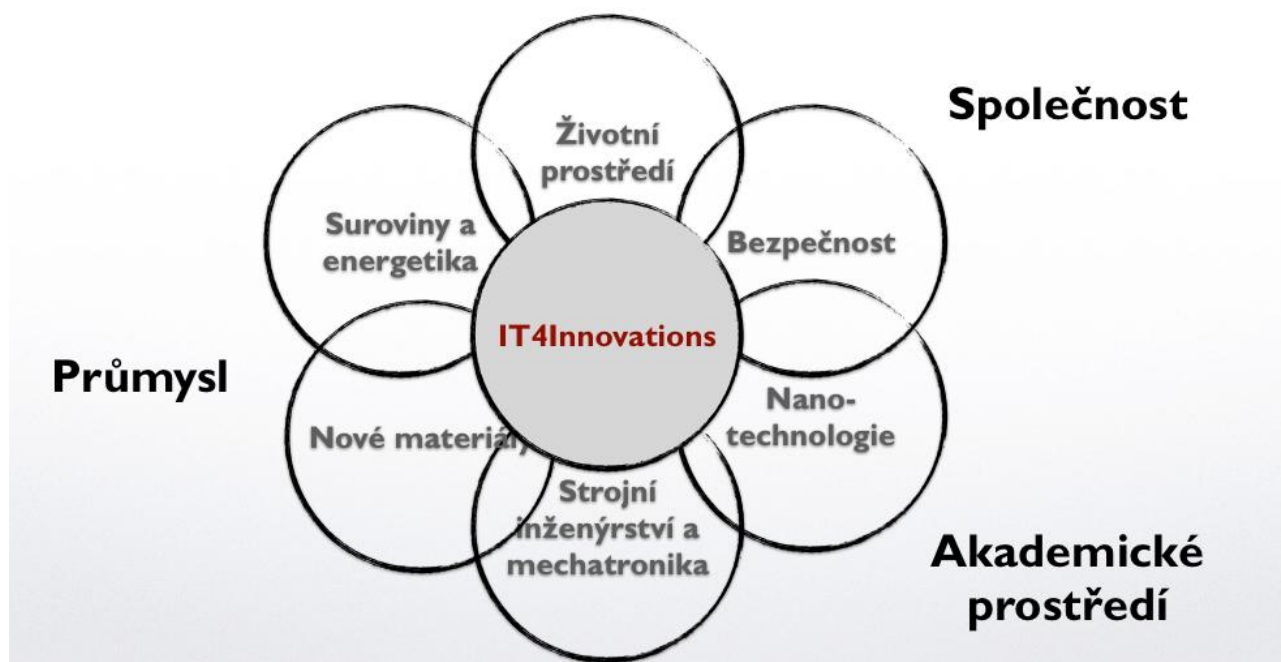
Nutnost zajistit potřebou kvalitu výzkumných pracovníků a vytvořit potřebnou tzv. kritickou masu lidských zdrojů pro přípravu, realizaci a následnou fázi udržitelnosti centra excellence vedla ke koncepci definovat celý projekt jako partnerský. Partnery projektu IT4I je tak pět výzkumných organizací, které do projektu vložily své výzkumné kapacity, jež jsou aktivní a dosahují odpovídajících výsledků v předmětné oblasti výzkumu: VŠB –

¹ Kolektiv řešitelů projektu IT4Innovations

Technická univerzita Ostrava, Ústav Geoniky Akademie věd ČR, Vysoké učení technické v Brně, Ostravská univerzita v Ostravě a Slezská univerzita v Opavě.

2. Výzkumná strategie

Definování výzkumné strategie projektu IT4I vychází z koncepce *“Triple Helix”* propojující akademickou oblast s průmyslem a veřejným sektorem a bylo založeno na kontextu vědecko-výzkumných aktivit realizovaných na partnerských pracovištích. Graficky lze tento výzkumný kontext vyjádřit následujícím obrázkem (obr. 1):



Obr. 1: Výzkumný kontext centra excelence IT4Innovations

Na základě takto definovaného rámce byly stanoveny tři navzájem se prolínající témata výzkumného zaměření celého centra označené následujícími kódovými názvy: IT4People, SC4Industry a Theory4IT.

2.1. IT for People (IT4People)

Klíčové pro budoucí rozvoj společnosti bude zajištění její bezpečnosti a poskytnutí pomoci ve stavech nouze a krizových situací. Přestože každým rokem Českou republiku a celou řadu dalších evropských zemí sužují povodně, požáry a ekologické problémy, které představují finanční ztráty v řádu desítek miliard korun, neexistuje v současné době efektivní nástroj, který by dokázal občany i zodpovědné orgány na takové situace připravit a včas zajistit jejich řešení. Cílem oblasti IT4People je tedy napomoci tuto situaci změnit cestou poskytnutí nástrojů pro simulaci a predikci vývoje výše uvedených krizových situací a napomoci orgánům a institucím zodpovědným za jejich management při jejich rozhodování. Je nutné vytvořit a dále rozvíjet výkonné nástroje, které předpoví např. oblast zatopení s několikahodinovým předstihem, dokáží včas identifikovat dopravní kolaps a zajistit přesměrování dopravy takových způsobem, aby nezpůsobil stejný kolaps v jiné lokalitě. Nástrojem k takovým řešením jsou rozsáhlé soubory heterogenních dat, která je nutné sbírat, ukládat do datových skladišť a efektivními algoritmy, které dokážou s takovými daty pracovat a poskytnout požadované informace nutné pro efektivní a kvalifikované rozhodování. Všechny tyto záležitosti vyžadují dlouhodobě orientovaný výzkum a vývoj soustředěný do oblastí jako jsou výpočetní matematika, inteligentní systémy, softwarové inženýrství, internetové technologie, distribuované aplikace apod..

2.2. Supercomputing for Industry (SC4Industry)

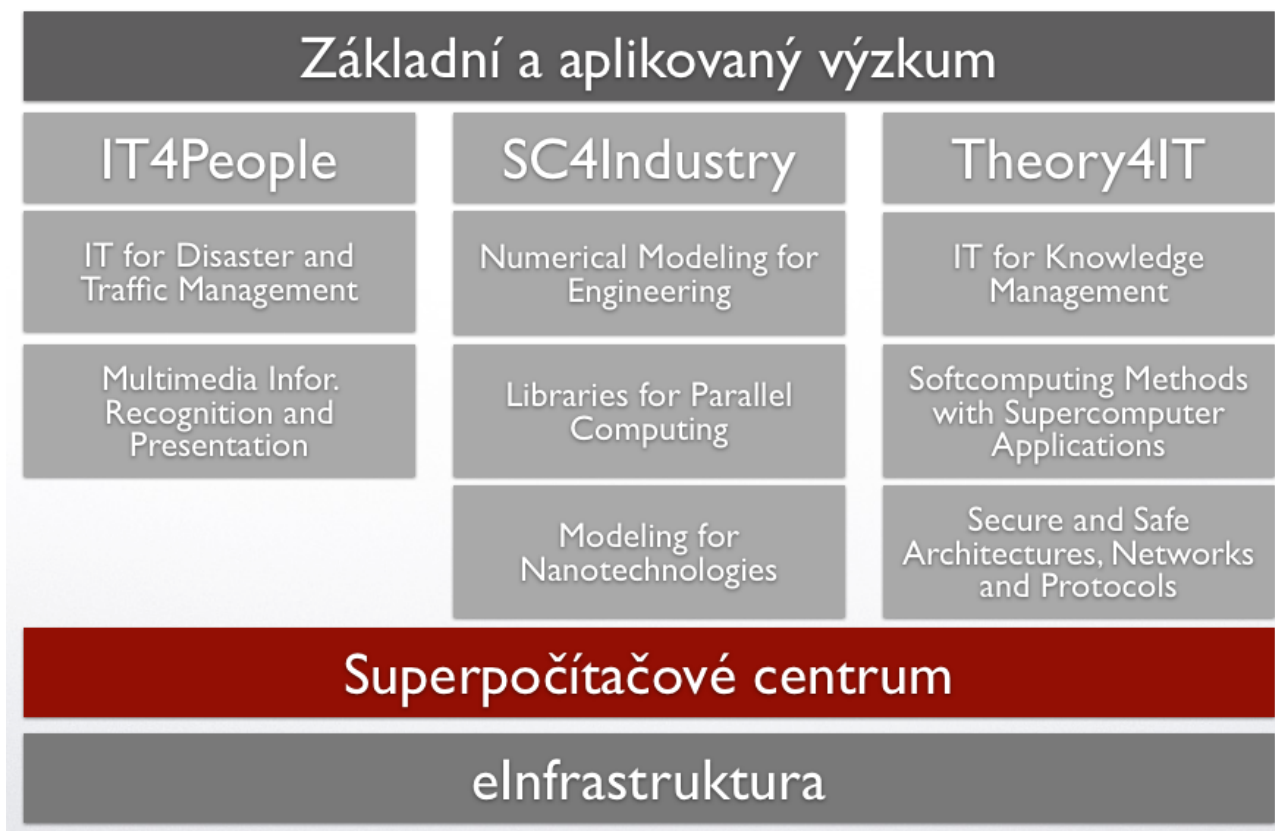
S rychlým rozvojem informačních technologií v posledních pár letech se otevřely zcela nové možnosti i v oblasti počítačového modelování komplexních fyzikálních a obecněji přírodních jevů. Tento pokrok nachází rozsáhlé uplatnění i v oblasti průmyslových aplikací. V současné době jsme svědky toho, jak kvantitativní přístupy k řešení problémů pronikají do běžné praxe a nahrazují tradiční časově a finančně náročné procesy. Počítačové modelování a simulace umožňují, v celé řadě případů, finančně a časově nákladné fyzikální experimenty nahradit systémy složitých soustav matematických rovnic a tyto efektivně řešit na výkonné výpočetní technice. Superpočítač se tak stává nenahraditelným laboratorním prostředkem realizace experimentů snad ve všech vědních disciplínách. Tento proces je však současně čím dál náročnější. Nelze od sebe jednoduše oddělit základní výzkum od aplikací, při nichž vznikají nové problémy. Jejich řešení je tak velmi často postaveno na významném průniku základního a aplikovaného výzkumu.

2.3. Theory for IT (IT4Theory)

Hlavním a zásadním přínosem části zaměřené na problematiku teoretického přístupu k výpočetním metodám a informačním technologiím je hledání nových vědeckých poznatků z pohledu základního výzkumu. Výsledky jsou jednak směřovány do rozvoje informatiky a informačních technologií jako takových, dále pak se ale také očekává důležitý impakt na ostatní části centra excelence. Výsledky v oblasti softcomputingu, přírodou inspirovaných algoritmů, znalostního managementu a dalších teoretických sekcí budou tvořit základ pro jejich využití ve výzkumu a vývoji koncentrovaného kolem dvou výše uvedených aplikačních oblastí. Za velmi důležitý až klíčový lze považovat i výstup výzkumu v oblasti formálních metod softwarového inženýrství, protože vývoj softwaru orientovaný na výše uvedené oblasti vyžaduje specifický přístup a definování adekvátních softwarových procesů, které musí zajistit požadovanou kvalitu. Předpokládané výstupy v podobě aplikací určených pro krizový management musí splňovat ty nejnáročnější kritéria kvality výsledného softwarového produktu.

3. Výzkumné programy

Na základě takto stanovené výzkumné strategie byly následně definovány výzkumné programy, které tvoří základ tvůrčí činnosti centra excelence IT4I. Tyto lze opět graficky vyjádřit na následujícím obrázku (obr. 2):



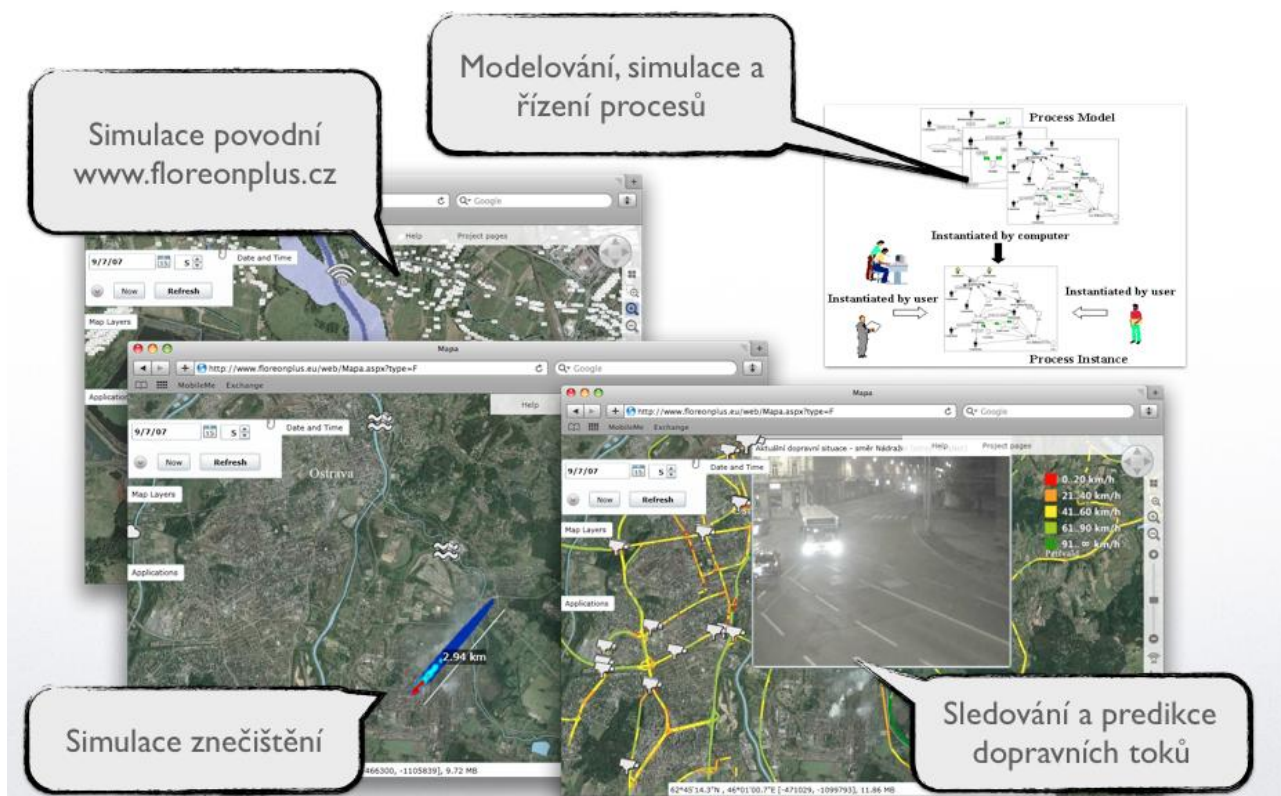
Obr. 2: Výzkumné programy centra excellence IT4I

Důležitým atributem projektu IT4I je i jeho začlenění do širšího kontextu budování národních infrastruktur pro vědu a výzkum v České republice a EU. Je zjevné, že superpočítání se nemůže obejít bez stabilní vysokorychlostní komunikační sítě a nad nimi vystavěných služeb využívajících výpočetní a úložné gridy. Tyto potřeby vedly k vytvoření strategického partnerství tří významných projektů z oblasti IT financovaných především z Evropských strukturálních fondů, kterými jsou kromě IT4I také CESNET a CERIT-SC. Takto definovaná spolupráce umožňuje vědecké komunitě efektivně využít drahých výkonných výpočetních prostředků a do budoucna i sjednotit vynakládané úsilí nutné k udržení a dalšímu rozvoji takové národní e-Infrastruktury.

Jednotlivé výzkumné programy, tak jak byli definovány jejich realizačními týmy, lze pak stručně shrnout do následujícího popisu.

3.1. IT for Disaster and Traffic Management

Hlavním cílem tohoto výzkumného programu je vývoj komplexním softwarových nástrojů pro potřeby podpory rozhodování v krizových situacích způsobených nepříznivými jevy jako jsou povodně, znečištění, ekologické katastrofy či dopravní krizové situace. Smyslem je propojit heterogenní systémy poskytující data o všech těchto událostech a využít výkonných výpočetních prostředků k simulaci celé řady variant vývoje krizové situace. Na základě těchto predikcí je pak nutné plánovat a řídit činnost integrovaných záchranných center směrem k minimalizaci ztrát na životech a majetku. Tento výzkumný program navazuje na systém Floreon+ (obr. 3), který je na VŠB-TU Ostrava vyvíjen od roku 2006 a který je v dnešní době implementován ve formě poloprovozu v rámci IBC (Integrované bezpečnostní centrum) Ostrava.



Obr. 3: Výstupy systému Floreon+ pro podporu řízení krizových situací

3.2. Multimedia Information Recognition and Presentation

Výzkumný program zaměřený na oblast rozpoznávání a prezentaci informací z multimediálních dat se zabývá zejména zpracováním multimediálních dat s ohledem na jejich vlastnosti na úrovni signálů pořízených nejrůznějšími senzory, a tím umožňuje multimodální rozpoznávání a pořizování znalostí. Důraz je kladen zejména na analýzu obrazu, videa a řeči, zobrazování, vizualizaci a modelování včetně modelování lidských tkání, akceleraci ve specializovaném hardware a techniky sémantického webu, formálních jazyků a gramatik. Výsledkem programu budou nové postupy a znalosti ve výše uvedených oblastech, které nejen pomohou zlepšit parametry stávajících aplikací, ale měly by umožnit i tvorbu zcela nových aplikací, které dosud nejsou realizovatelné jak z technických, tak i z ekonomických důvodů.

3.3. Numerical Modeling for Engineering

Numerické modelování pro řešení inženýrských problémů je zaměřeno na rozvoj efektivních matematických metod pro řešení problémů, které jsou v současné době díky svému rozsahu, nehladkosti nebo nelinearitě na hranici řešitelnosti, a následně jejich zapojením do řešení vybraných komplexních vědeckých a inženýrských problémů. Důraz je kladen především na vývoj algoritmů, které dokážou efektivně využít moderní superpočítače a jsou v jistém smyslu optimální, čili škálovatelné (obr.4). Budou to zejména algoritmy, které mají pro některé třídy úloh asymptoticky lineární složitost, přičemž doba výpočtu je přibližně nepřímo úměrná počtu procesorů. Výzkum vychází z rozvoje efektivních metod pro řešení základních přímých úloh se speciální strukturou, včetně vnějších úloh, které budou dále využity k řešení multifyzikálních úloh, multiškálových úloh, inverzních úloh identifikace fyzikálních parametrů a k řešení úloh optimálního navrhování. Výzkumná činnost zahrnuje všechny fáze matematického modelování, od výběru problémů a jejich korektní matematické formulace, přes diskretizaci a návrh efektivních

algoritmů, studium výpočetní náročnosti, analýzu spolehlivosti výsledků (aposteriorní odhady) až po implementaci nových algoritmů.



Obr. 4: Škálovatelné paralelní řešiče postavené na doménové dekompozici

3.4. Libraries for Parallel Computing

Kromě standardních softwarových balíčků ať už komerčního či nekomerčního charakteru, které budou k dispozici na superpočítačovém centru pro řešení praktických problémů, bude možné využívat i nově vyvinuté softwarové balíky, které budou výsledkem výzkumu programu Numerické modelování pro řešení inženýrských problémů a které budou využívat nejnovějších poznatků v této oblasti. Jako nejzajímavější se v této části řešení projektu jeví vzájemné propojení vlastních softwarových nástrojů se standardně využívanými komerčními a nekomerčními balíky.

3.5. Modeling for Nanotechnologies

Jedna z klíčových aplikačních oblastí využití superpočítače IT4I jsou také nanotechnologie. Výzkumný program pro modelování v oblasti nanotechnologií je založen na vytváření matematicko-fyzikálních modelů jak v základním, tak také v aplikovaném výzkumu a aplikace superpočítačových přístupů pro efektivní řešení těchto modelů. Konkrétně se jedná o výpočtové modely směřované do oblasti vývoje nových magnetofotonických nanostruktur, specifikace nových přístupů k nedestruktivní bezkontaktní metrologii a diagnostice, realizace nových farmaceutických forem a návrhu nových progresivních nanokompozitů.

3.6. IT for Knowledge Management

Zpracování rozsáhlých souborů dat a jejich analýza je nepochybně jedna z klíčových oblastí soudobého výzkumu v IT. Nejinak je tomu i v případě IT4I. Paralelizace algoritmů používaných v oblasti získávání znalostí a vytěžování z dat je nepochybně velkou výzvou do budoucna. Stále narůstající množství dat, která získáváme z různých zdrojů, je čím dál

tím více potřeba vhodně strukturovat a ukládat, efektivně v nich vyhledávat a analyzovat za účelem získání dalších informací a poznatků. Konkrétně se výzkumný tým bude orientovat především na oblast zpracování rozsáhlých kolekcí slabě strukturovaných dat, extrakce znalostí z dat, využití metod redukce dimenze pro vysoce dimenzionální data, využití metod lineární algebry, neuronových sítí, statistických metod a shlukovacích metod pro analýzu dat.

3.7. Softcomputing Methods with Supercomputer Applications

Primárním cílem tohoto výzkumného programu je vývoj matematických metod a nástrojů postavených např. na matematické teorii fuzzy systémů s jejich následnou implementací na paralelních výpočetních prostředcích. Softcomputing tak lze v kontextu centra excelence IT4I vnímat jako rozšíření tradičních superpočítačových metod, které do budoucna umožní řešit extrémně komplikované problémy s neurčitou nebo omezeně definovanou strukturou. Výstupy z této oblasti výzkumu by měly být směřovány mimo jiné i do oblasti krizového řízení, kde by měly tvořit alternativní přístupy k modelům postavených na klasických metodách vyžadujících exaktně definované vstupy.

3.8. Secure and Safe Architectures, Networks and Protocols

Hlavním záměrem a cílem je zlepšení bezpečnosti a spolehlivosti vyvíjených IT aplikací. Jejich úmyslné či neúmyslné zneužití představuje jednu z hlavních hrozeb pro moderní společnost. V kontextu zneužití znalostí a s nimi i informačních technologií se požadavek na bezpečnost systému stává stejně důležitým jako je funkčnost, cena či udržitelnost. Výzkum se tedy bude ubírat cestou modelování, návrhu, optimalizace, diagnostiky, validace a verifikace vybraných HW/SW systémů, které vyžadují vysokou úroveň zabezpečení a spolehlivosti. Je zjevné, že právě superpočítačové technologie v této oblasti sehrávají postupně stále důležitější roli.

4. Technické parametry superpočítačového centra

Vlastní budování superpočítačového centra je rozloženo do dvou etap, kde v první etapě se předpokládá instalace menšího výpočetního clusteru, který se svými 256 výpočetními uzly architektury x86 bude dosahovat celkového špičkového výkonu až 70 TFLOP/s a disponovat celkem 12 TB paměti RAM. Doplnovat jej bude systém úložných kapacit tvořený několika úrovněmi úložných technologií, který dosáhne svou kapacitou přibližně 1 PB. Budou použity jak rychlé, tak i vysokokapacitní diskové prvky, aby byla dosažena požadovaná velikost a propustnost celého úložiště. Pro výzkum v oblasti využití GPU akcelerace bude pořízen menší 50 uzlový cluster architektury x86 s 2,4 TB paměti RAM, plně osazen GPGPU akcelerátory s předpokládaným výkonem až 35 TFLOP/s. U obou těchto výpočetních clusterů bude použita technologie Infiniband (40 Gb/s) na zajištění rychlé komunikace mezi jednotlivými uzly. Protože se už počátku provozu předpokládají výpočty v oblasti modelování globálních změn klimatu projektu CzechGlobe, bude systém doplněn o další úložnou kapacitu až do velikosti 2 PB, tvořenou kombinací diskových polí a páskových knihoven s důrazem kladeným na úložnou kapacitu před vysokou propustností.

V druhé polovině roku 2014 bude instalován tzv. "velký" cluster o výpočetním výkonu 1PFLOP/s s více než 30 tisíci jádry o celkové paměťové kapacitě 40 TB RAM. Tento systém bude doplněn o dalších 3.5 PB úložných kapacit. Dále bude pořízen výpočetní systém se sdílenou pamětí (SMP resp. NUMA) o velikosti minimálně 2 TB a s 256 výpočetními jádry o předpokládaném výkonu 6 TFLOP/s. SMP/NUMA systém bude vybaven vlastním diskovým subsystémem o velikosti 20 TB pro potřeby dočasného úložiště výpočetních dat. Bude posílena síťová konektivita do CESNETu na alespoň 40 Gb/s. Pro potřeby vizualizace, především preprocessing a postprocessing výpočetních

úloh, bude vybudováno vizualizační centrum, které bude obsahovat pracovní stanice s možností zpracovávat velké objemy dat projekčním zařízením. Vše by mělo být dostupné v roce 2015 v nově vybudovaných prostorách umístěných v areálu VŠB - Technické univerzity Ostrava (obr. 5):



Obr. 5: Vizualizace budovy superpočítačového centra IT4I

6. Využití superpočítačového centra a vazby na další projekty

Dle výše uvedeného plánovaného zprovoznění klíčové infrastruktury projektu budou moci být od roku 2013 výpočetní prostředky superpočítače IT4I využívány také externími zájemci. Realizační tým IT4I v současné době již rozpracovává základní pravidla pro přidělování těchto prostředků. Výpočetní kapacity superpočítačového centra IT4I budou rozdělovány ve třech základních režimech:

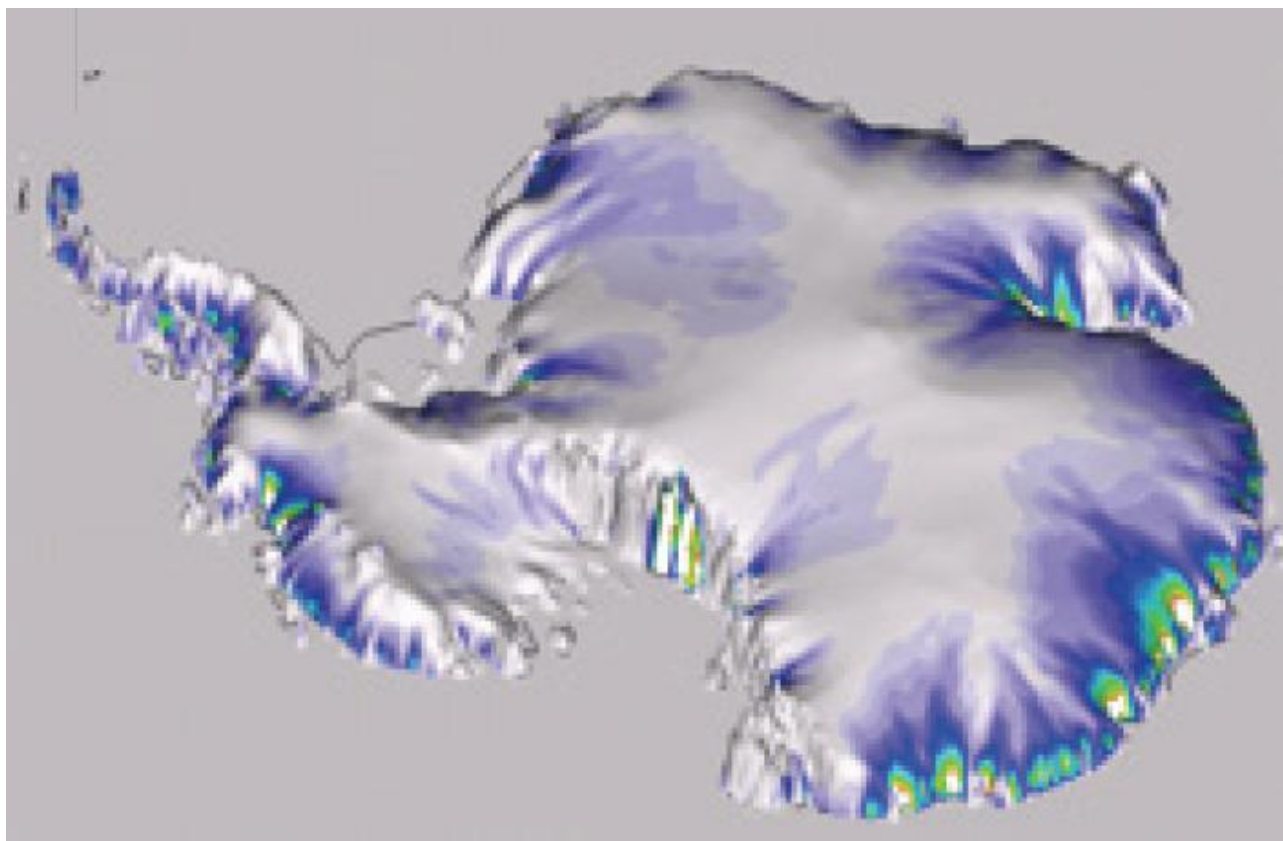
1. Výpočetní zdroje určené pro realizaci *vlastních výzkumných programů*;
2. Otevřený přístup (*Open Access*) k využití výpočetních kapacit ze strany jiných uživatelů;
3. *Vyhrazený specifický přístup* k využití výpočetních kapacit superpočítače ze strany jiných výzkumných infrastruktur národního významu.

Pro druhý a třetí režim bude k dispozici ne méně než 50% všech výpočetních prostředků. Ty budou zejména „vyhrazeny“ českým výzkumným týmům, zároveň však část těchto prostředků bude sloužit pro účely zapojení se do evropské výzkumné infrastruktury v rámci ESFRI. U *otevřeného přístupu* předpokládáme rozdělení zájemců do třech základních skupin dle velikosti požadovaných výpočetních prostředků, a to na základě tzv. účtovacích jednotek, kdy 1 účtovací jednotka se rovná jedné hodině reálného času přidělení jednoho výpočetního elementu (jádra):

- *Malé alokace* (do několika tisíc účtovacích jednotek/měsíc) pro okamžité krátkodobé výpočty menšího rozsahu, snadno dosažitelné i pro jednotlivce včetně Ph.D. studentů;
- *Střední alokace* (desítky tisíc účtovacích jednotek/měsíc) budou určeny především pro pravidelné využívání uživatelskými komunitami za účelem plnění jejich výzkumných cílů;
- *Velké alokace* (v řádu stovek tisíc účtovacích jednotek/měsíc) pro projekty využívající významnou část celkového výpočetního výkonu po delší období.

V případě velkých alokací se předpokládá podávání žádostí v termínech vyhlašovaných 2x ročně. Jejich vyhodnocení a oznámení výsledku by nemělo být delší než 1 měsíc. V případě malých a středních alokací budou aplikovány flexibilnější režimy (aktuálně jsou zvažovány průběžné, resp. tříměsíční intervaly). U vyhrazeného přístupu je plánován odlišný režim, který bude spočívat na smluvním zakotvení podmínek dlouhodobější spolupráce. V současné době dostává tento režim konkrétní podoby zejména při spolupráci s Centrem výzkumu globální změny AV ČR, v.v.i. - projekt CzechGlobe.

Systematická práce a výsledky v oblasti paralelního programování dostaly nedávno nový impuls v podobě zapojení do prestižního partnerství evropských superpočítačových center sdružených v rámci infrastruktury PRACE RI (*Partnership for Advanced Computing in Europe Research Infrastructure*). V rámci projektu PRACE se výzkumný tým VŠB-TU Ostrava podílí i na vývoji zcela nových paralelních algoritmů pro řešení problému tání ledovců (obr. 6). Tyto algoritmy jsou implementovány do OpenSource softwaru Elmer, který je primárně vyvíjen ve finském superpočítačovém centru CSC. Kromě toho centrum IT4I je také kontaktním místem PRACE pro Českou republiku. Pomáhá zájemcům z řad českých výzkumníků využívat výpočetní čas v evropských superpočítačových centrech, získat přístup k dalším výpočetním službám a prostřednictvím vzdělávacích a informačních aktivit rozšiřovat počet uživatelů těchto technologií v ČR.



Obr. 6: Simulace odtávání ledovců

V rámci agenturou TAČR vyhlášeném programu Center kompetence IT4I získalo projekt RODOS (ROzvoj DOpravních Systémů), jehož primárním cílem je vývoj nových metod pro potřeby modelování, monitorování a řízení mobility. V rámci strategické výzkumné agendy takového centra kompetence budou mimo jiné řešeny i úlohy spojené s efektivním plánováním rozvoje dopravní infrastruktury. *Dlouhodobým strategickým cílem centra je vytvoření komplexní informační nadstavby nad dopravou s využitím nových nástrojů dopravní informatiky a její integrace do stávajících telematických systémů.*

7. Závěr

Z pohledu tzv. velkých projektů Operačního programu Výzkum a vývoj pro inovace je projekt Centra excelence IT4Innovations jediný, který je tématicky orientován do oblasti informačních technologií. Objem finančních prostředků ve výši 1,83 mld Kč představuje jednu z největších investic do výzkumných kapacit v této vědní oblasti. Cílem ale není jen vybudovat výkonné superpočítačové centrum velikosti Top50, ale především vytvořit výzkumné zázemí pro cca 200 kmenových zaměstnanců centra. Právě nutnost zajistit tuto kritickou masu výzkumných pracovníků vedla ke snaze vzájemně propojit několik akademických pracovišť do jediného centra. To ale v žádném případě neznamená, že je tímto projekt uzavřen pro další partnery. Je tomu přesně naopak. Jeho finanční udržitelnost v letech 2015-2020 bude vyžadovat další napojení na výzkumná pracoviště v České republice, v EU a ve světě obecně, protože tato finanční udržitelnost je přímo podmíněna udržitelností lidských zdrojů a vědeckých kapacit v již tak nepřilíš optimistických podmínkách financování vědy, výzkumu a inovací v České republice.