

## Projekt GridLab a MU

Luděk Matyska, FI a ÚVT MU

V loňském roce se Ústavu výpočetní techniky podařilo stát se jedním ze zakládajících členů projektu 5. Rámcového programu EU s názvem *Grid-Lab*. Projekt úspěšně prošel jak vlastním výběrovým řízením, tak i fází přípravy finálního projektu – to bylo zvláště obtížné vzhledem k tomu, že ze strany EU došlo k téměř 40% krácení původně požadovaných prostředků – a oficiálně byl zahájen k 1. lednu 2002.

Základním cílem tohoto projektu je vytvořit soupravu nástrojů (programy, knihovny, ...), které umožní snazší tvorbu aplikací schopných využívat vysoce heterogenní výpočetní prostředí gridů. Tento *Grid Application Toolkit* (GAT) vzniká v přímé návaznosti na požadavky konkrétních aplikací a jeho vlastnosti jsou ověřovány na Gridu, který je postaven pro účely tohoto projektu.

### Základní principy

Projekt GridLab je zaměřen na podporu aplikací s mimořádnými nároky na výpočetní výkon. Prototypovou aplikaci tvoří simulace astrofyzikálních jevů, generujících silné gravitační vlny, které by bylo možno zachytit pozemskými detektory (např. kolize černých děr nebo neutronových hvězd, gravitační kolaps hvězd, signály pulzarů apod.). Požadovaný výkon se pohybuje v řádu stovek TeraFlops ( $100 \times 10^{12}$  operací v pohyblivé řádové čarce za sekundu) a je požadován hodiny až dny, což přesahuje možnosti dnešních nejvýkonnějších superpočítačů. Jediným možným řešením je logické propojení více počítačů dohromady a složení jejich výkonu (fyzické propojení zajišťuje vysokorychlostní síť). To je však možné pouze tehdy, je-li k dispozici odpovídající programové vybavení, které dokáže efektivně takto propojené počítače využít. Tvorba takovýchto programů je doposud velmi složitou a ve své podstatě zatím neautomatizovatelnou činností, a proto je takovýchto aplikačních programů velmi málo. Jedním z možných řešení tohoto problému je vytvoření speciální programové vrstvy (middleware), která programátorovi usnadní použití gridového prostředí bez toho,

aby detailně znal strukturu a požadavky jednotlivých počítačů, které konkrétní Grid tvoří, či se detailně musel zabývat řešením všech problémů, spojených např. s rozpoznáním výpadku či nedostupnosti nějakého výpočetního uzlu a jeho náhrady. Právě to je obsahem prostředí GAT, vyvíjeného v rámci projektu.

Kromě orientace na podporu vysoce náročných výpočtů se projekt GridLab snaží o úzké a okamžité propojení vývoje nové programové vrstvy, požadavků aplikací a konkrétního Gridu, na němž je možno okamžitě ověřovat výsledky vývoje na konkrétních výpočtech. Požadavky na GAT jsou generovány a modifikovány podle potřeb aplikací, současně je ověřováno, zda jejich implementace jsou skutečně funkční na dostatečně rozsáhlém a heterogenním Gridu, tvořeném několika desítkami systémů zcela odlišné architektury i výkonu.

Další významnou součástí projektu GridLab je podpora *adaptability* a *mobility*. Jak bylo uvedeno již dříve, reálný cílový Grid je velmi heterogenní a každá konkrétní aplikace se musí být schopna s touto různorodostí vypočítat. Je nutné vzít do úvahy nejen výkon a kapacitu výpočetních uzlů, ale i kvalitu a propustnost propojovací sítě. Rovněž je třeba počítat s tím, že bude docházet k výpadkům uzlů, které přitom nesmí vést ke zhroucení celé aplikace. Na druhé straně dlouhodobé výpočty může být třeba sledovat a řídit z více míst, navíc vědci poměrně značně cestují a přitom by rádi udrželi kontakt s rozsáhlou simulací, která na Gridu běží. Podpora mobility tak zahrnuje jak možnost (případně i současného) sledování (a řízení) simulace z více míst, tak i upozornění na důležité jevy např. formou zaslání SMS zprávy na mobilní telefon či pager.

### Předpokládané výsledky

Mezi hlavní plánované výsledky projektu GridLab patří:

1. Vývoj GAT. Jeho součástí bude celá řada nezávislých modulů, podporujících nejdůležitější součásti gridových programů, zejména: sledování výkonu aplikací i celého Gridu, vyhledání, plánování a alokace zdrojů, informační služby, bezpečnost, upozornění na

události a podpora společné práce (collaboration) včetně vzdáleného řízení simulace, správa rozsáhlých objemů dat a vzdálená vizualizace.

2. Nové gridové aplikace, resp. převod stávajících aplikací do nového prostředí (s využitím GAT) tak, aby mohly plně využít všech předností gridové infrastruktury.
3. Vývoj konkrétního Gridu, včetně prostředků na jeho sledování a správu. Tento Grid bude současně použit jako výpočetní prostředí pro aplikace zmíněné v předchozím odstavci.

## Zapojení MU

Masarykova univerzita je zapojena především do třetí z výše zmíněných oblastí, tedy do vývoje a správy Gridu. Pracovníci ÚVT MU jednak odpovídají za celkovou správu Gridu, jednak vyvíjejí nové nástroje a prostředky pro sledování stavu Gridu a služeb jím poskytovaných. Je naplánován vývoj konkrétní aplikace, jejímž úkolem bude použití všech nebo konkrétní podmnožiny služeb poskytovaných GAT a sledování jejich stavu. Výsledky budou dále zpracovávány tak, aby poskytly co nejpřesnější obraz stavu celého Gridu a současně je bylo možno použít pro odstranění nalezených nedostatků (identifikace nefunkčních uzlů či služeb, upozornění příslušného odpovědného správce atd.). Celý systém sledování a správy Gridu by přitom měl v maximální možné míře využívat GAT a tím poskytovat obraz co nejbližší tomu, který „vidí“ vlastní aplikace.

ÚVT MU se dále podílí na pracích souvisejících s bezpečností gridového prostředí (v současné době spravuje certifikační autoritu projektu GridLab), na zajištění základních informačních služeb a na vývoji administrativního portálu, který by měl shrnovat informace o stavu Gridu a požadavcích na nápravu chybných stavů.

Nedílnou součástí odpovědnosti ÚVT MU je i správa uživatelských účtů, certifikátů atd.

## Partneři a rozpočet

Koordinačním pracovištěm celého projektu je Síťové a superpočítačové centrum v Poznani, další

spolupracující akademické organizace jsou uvedeny dále:

- Albert Einstein Institute (část Max Planck Institute zur Foerderung der Wissenschaften), Německo
- Konrad-Zuse Zentrum für Informationstechnik Berlin, Německo
- Vereniging voor Christelijk Wetenschapelijk Ondervijs (Vrije University), Holandsko
- Masarykova univerzita v Brně, Česká republika
- University of Wales, Cardiff, Velká Británie
- Università degli Studi di Lecce, Itálie
- Institute of Communication and Computer Systems of the National Technical University of Athens, Řecko
- Magyar Tudományos Akadémia Számítástechnika és Automatizálási Kutatóintézet (MTA SZTAKI), Maďarsko

Kromě nich se na řešení projektu podílí i následující komerční organizace:

- Gridware GmbH, Německo
- Compaq Computer France, Francie

GridLab je jedním z mála EU projektů, na jejichž řešení se přímo podílí (a jsou placeni z prostředků EU) i partneři za USA:

- University of Wisconsin
- Argonne National Laboratory (ANL)
- Information Sciences Institute, součást University of Southern California (ISI)

Projekt je plánován na 3 roky s celkovým rozpočtem více jak 5 milionů Euro. Jedná se tak o finančně největší projekt, na jehož řešení se podílí pracoviště Masarykovy univerzity. Výsledky projektu by se kromě jiného měly využít i v rámci *META Centra*, zejména pak při jeho správě. Předpokládáme, že zkušenosti z testování stavu Gridu či administrativní portál budou takto přímo využity.

Dlouhodobějším přínosem je know-how v oblasti vývoje Gridových aplikací, zejména pak s využitím vyvíjeného GAT. V současné době se pracuje na převodu jedné aplikace z oblasti výpočetní chemie do gridového prostředí. Zde chceme postupně nabízet spolupráci těm, kteří by své aplikace chtěli převést do gridového prostředí. □