

Do Gridu snadno a rychle – prostředí VOCE

Daniel Kouřil, Jan Kmuníček, ÚVT MU

Pojem *Grid* se objevil v devadesátých letech minulého století jako nová forma distribuovaného zpracování informací a řešení složitých problémů. Gridy se zpočátku soustředily primárně na oblast náročných výpočtů, ale postupně pronikají i do jiných oblastí, které mohou profitovat z řešení, která byla vytvořena v rámci rozvoje Gridů. Za všechny lze jmenovat např. oblast digitálních knihoven, e-learningu nebo kvalitních videokonferenčních či multimediálních nástrojů. Gridy se také vždy pohybovaly na horní hranici technologických možností a přitahovaly pozornost řady odborníků jak z akademického prostředí tak z IT průmyslu. Více informací o Gridech, jejich možnostech a vývoji lze nalézt v článku „Gridy jako klíčový fenomén informačních technologií nového tisíciletí“ v předchozím čísle Zpravodaje.

Po řadě let intenzivního vývoje se v současné době Gridy dostávají do fáze, kdy se přibližují běžným uživatelům, kteří tak mohou začít využívat možnosti, které gridové prostředí nabízí. Vzhledem k enormnímu zájmu, který o gridové technologie panuje, a obrovskému úsilí a množství prostředků, které byly do jejich vývoje investovány v minulém desetiletí, nabízí současné Gridy velmi širokou škálu možností a mechanismů, což zároveň nutně vede k tomu, že současné gridové řešení jsou velmi komplexní a s pozvolnou učící křivkou. Zejména řadoví uživatelé, kteří nemají zkušenost z podobného distribuovaného prostředí, se mohou cítit zaskočení komplexností Gridů. Gridové prostředí je náročné i z hlediska administrativy, protože obsahuje množství služeb, často se složitou konfigurací, které nemají mimo gridový svět obdobu. Administrátoři proto potřebují hodně času a úsilí k instalaci aspoň základních služeb a zejména k jejich rutinnímu provozu.

Noví uživatelé, kteří se chtějí seznámit s Gridy a začít je používat pro řešení svých problémů, tak často naráží na velkou bariéru. V lepším případě mají dostupné nějaké gridové prostředí (zpravidla postavené pro některý existující gridový pro-

jekt), v horším případě začínají na zelené louce a snaží se o vybudování infrastruktury (nebo aspoň její části) vlastními silami. V obou případech je velmi často složitost a množství všech potřebných kroků odrazující a uživatelé se často vrací ke svým zaběhlým zvyklostem, které jim poskytují zajištěné zázemí, přestože pomocí gridových technologií by bylo často možné řešit jejich problémy efektivněji.

V současnosti se věnuje netriviální úsilí tomu, aby se Gridy staly přístupnější pro širší skupiny uživatelů. Ve zbytku tohoto článku poskytneme popis prostředí VOCE (*Virtual Organization for Central Europe*), které bylo vybudováno pro maximální usnadnění přístupu nových uživatelů ke gridovým technologiím a plně funkční gridové infrastruktuře. Prostor je budováno jako součást projektu EGEE¹, jehož hlavním cílem je vytvoření produkční panevropské gridové infrastruktury. VOCE je aktivita Středoevropské federace EGEE, kterou tvoří instituce z České republiky, Maďarska, Polska, Rakouska, Slovenska a Slovinska. Vývoj VOCE je koordinován sdružením CESNET, který v projektu zastupuje Českou republiku.

1 VOCE

Cílem VOCE je poskytovat všem uživatelům ze středoevropského regionu plně produkční gridové prostředí (tzv. *virtuální organizaci*). Vybudovaná infrastruktura VOCE obsahuje všechny potřebné gridové služby a je v maximální možné míře nezávislá na komponentách, které by poskytovala nějaká třetí strana mimo region. Chod jádra infrastruktury je zajišťován sdružením CESNET, výpočetní zdroje a úložiště dat poskytují všichni účastníci. V současnosti je k dispozici přes 500 procesorů a téměř 6 TB úložného prostoru. Přestože jsou tyto kapacity zpravidla sdíleny i s jinými projekty a nejsou dedikovány výhradně pro VOCE, mají tak uživatelé VOCE přístup k velmi výkonnému prostředí.

Zájem VOCE se primárně soustředí na dvě skupiny uživatelů. První skupinou jsou noví uživatelé, kteří nemají žádnou zkušenost s Gridy

¹<http://www.eu-egee.org/>

a chtějí rychle proniknout do základů gridových technologií bez toho, aby sami museli konfigurovat a udržovat vlastní infrastrukturu, domlouvat využití zdrojů s jinými institucemi apod. Další cílovou skupinou jsou uživatelé, kteří již mají nějakou předešlou zkušenost s Gridy a mají také konkrétní aplikace, které by rádi v Gridu vyzkoušeli, ale zároveň nemají kapacity (časové, finanční, lidské) na vybudování vlastní infrastruktury. Těmto uživatelským skupinám VOCE umožňuje, aby mohly rychle vyzkoušet, jak zapojit své stávající aplikace do gridového prostředí. VOCE nabízí všem těmto zájemcům snadnou proceduru pro získání účtu a rychlého připojení do plně funkčního gridového prostředí. Díky tomu, že VOCE provádí outsourcing celého provozu a administrativy gridové infrastruktury, se uživatelé mohou plně soustředit na řešení svých problémů, aniž by byli zatěžováni problémy, které se týkají správy.

Infrastrukturu VOCE lze vedle vyzkoušení Gridů použít i pro plně produkční provoz. Uživatelé, kteří se již seznámili s gridovým prostředím a mají své aplikace připraveny pro běh v Gridu, mohou použít VOCE pro jejich rutinní provoz. VOCE tak umožňuje uživatelům hladký přechod od prvního seznamování se s gridovými nástroji až po produkční použití, vše ve stejném prostředí bez nutnosti dalšího přeškolení. V případě, že prostředí VOCE přestane některé skupině stačit a bude vyžadovat více zdrojů, příp. větší garanci jejich dostupnosti, předpokládáme, že si vybuduje vlastní infrastrukturu. Infrastruktura VOCE je založena na nejrozšířenějším gridovém middleware dneška, je proto velmi pravděpodobné, že nově ustavené prostředí bude poskytovat totožnou funkcionalitu jako VOCE a pro uživatele bude velmi snadné začít pracovat v nové infrastruktuře na základě znalostí získaných ve VOCE.

Přestože VOCE poskytuje platformu, kde si uživatelé mohou vyzkoušet možnosti Gridů, liší se od jiných testovacích infrastruktur v tom, že zajišťuje alespoň minimální úroveň bezpečnosti. Současná prostředí, která se používají pro demonstraci a testování Gridů, umožňují v podstatě anonymní přístup, kdy v zájmu jednoduchosti povolují přístup uživatelům, jejichž iden-

tita není důvěryhodně ověřitelná (zpravidla se přístup povoluje na základě e-mailové komunikace). Tento stav samozřejmě vede k tomu, že počet dostupných zdrojů a jejich rozšiřitelnost je omezená. Vlastníci prostředků totiž nemají zájem vkládat své zdroje do infrastruktury, která nezajistí, že případně problémový uživatel je jednoduše dohledatelný. VOCE naproti tomu striktně vyžaduje, aby registrovaní uživatelé měli certifikát vydaný některou akreditovanou gridovou certifikační autoritou. Přísné požadavky na akreditaci certifikačních autorit zajišťují, že uživatelé jsou spolehlivě identifikovatelní v průběhu jejich pohybu po Gridu.

Na rozdíl od většiny současných virtuálních organizací, není VOCE spjato s žádnou konkrétní aplikační skupinou. Naopak VOCE se snaží přitáhnout různé aplikace a podporovat široké portfolio aplikačních uživatelů. Vývojáři VOCE jsou schopni podat pomocnou ruku v přenosu aplikací do prostředí Gridu nebo také pomoci zprostředkovat kontakty na jiné podobně zaměřené aplikační skupiny v gridovém světě.

V současné době má VOCE propracovaný systém podpory, který slouží pro efektivní řešení uživatelských problémů. VOCE disponuje experty z různých oblastí gridových technologií a je schopno poskytovat svým uživatelům plnou podporu.

2 VOCE z pohledu uživatele

VOCE je otevřeno všem uživatelům z akademického prostředí ČR. Pravidla pro registraci jsou uvedena na portálu VOCE², v této kapitole uvádíme jejich shrnutí a popis vlastností, které VOCE nabízí svým uživatelům.

Základním předpokladem pro získání členství ve VOCE je vlastnictví certifikátu veřejného klíče, který vydala certifikační autorita sdružení CESNET. Tato autorita je uznávaná v celosvětovém gridovém prostředí a držitelé jejich certifikátů mají tedy usnadněný přístup ke spolupráci se zahraničními partnery. Více informací o certifikátech lze nalézt např. v [1].

Pro získání certifikátu je nutná osobní návštěva pracoviště Registrační autority CESNET CA, která

²<http://egee.cesnet.cz/cs/voce/>

ověří patřičné doklady. Pro pracovníky a studenty MU je k dispozici pobočka na Superpočítačovém centru ÚVT³.

Zájemci o zapojení do VOCE musí vyplnit registrační formulář na webu, kde uvedou základní kontaktní informace. Pro přístup k formuláři se vyžaduje autentizace pomocí uživatelského certifikátu. Po vyplnění elektronické přihlášky je nejvýše do tří kalendářních dnů zřízen účet uživatele ve VOCE a lze začít s použitím Gridů. VOCE je od počátku spjato s projektem *META Centrum*⁴, který nabízí infrastrukturu pro náročné výpočty v České republice. Českým uživatelům proto doporučujeme požádat o členství v *META Centru* a v rámci této žádosti uvést také zájem o zapojení do VOCE. Uživatelům se tak otevře bohatší prostor pro realizaci výpočtů a získají přístup k více zdrojům. Přihlášení do *META centra* však není nutné pro členství ve VOCE.

Pro přístup ke gridovému prostředí VOCE je vyhrazen zvláštní počítač (tzv. *User Interface - UI*), kde je nainstalován veškerý software potřebný k použití gridových prostředků. Zřízení účtu na tomto stroji je také součástí registrační procedury VOCE. Od okamžiku, kdy je uživateli aktivován účet ve VOCE, je možné přihlásit se na tento stroj (skurut4.cesnet.cz) pomocí protokolu ssh a otevřít tak bránu do gridového světa. Z UI se zadávají úlohy, které jsou následně spouštěny na některém z volných strojů v Gridu bez ohledu na to, zda stroj je v Praze, Krakově nebo Košicích. Stejně tak lze průběh úlohy monitorovat a sledovat, jak prochází jednotlivými gridovými komponentami až do okamžiku, kdy skončí a uživatel si stáhne výsledek úlohy. UI dále obsahuje příkazy pro přístup k datovým úložištím, které umožňují snadno využít velkou diskovou kapacitu, která jsou ve VOCE dostupná.

Základní operace pro využití gridové infrastruktury jsou popsány v dokumentaci dostupné na portálu VOCE. Příprava a spuštění jednoduchých úloh je tak otázkou krátkého času. Podobně přístup k nabízeným úložištím je jednoduchý a základní operace lze zvládnout poměrně rychle. I uživatel, který neměl žádné znalosti gridových

technologií, si může osvojit základní gridové operace během velmi krátké doby.

Pro přístup ke zdrojům VOCE lze také použít specializované webové portály, které skrývají technické detaily a nabízí snadnější přístup k infrastruktuře. Pro VOCE jsou dostupné portály GILDA a P-GRADE. První byl vytvořen primárně pro účely demonstrací a byl vyvíjen s důrazem na použití ve výuce. Druhý nabízí sofistikovanější prostředky, zejména řízení workflow, tj. vztahů mezi více úlohami, které tvoří jeden celek.

Po zvládnutí základů použití Gridů se mohou uživatelé pustit do přenášení svého aplikačního portfolia. Mohou buď využít zmíněné portály nebo hledat vlastní řešení pomocí nativních gridových příkazů. Pro VOCE je také k dispozici unikátní systém Charon⁵, který vznikl na Národním centru pro výzkum biomolekul na Přírodovědecké fakultě MU. Nejdříve byl vytvořen pro účely *META Centra*, později byl upraven i pro VOCE. Charon tvoří vrstvu na pomezí mezi webovými portály a čistými prostředky gridového middleware. Skrývá řadu detailů nižších úrovní, takže uživatel se nemusí učit často komplexní syntax příkazů ani pronikat do detailů spletité gridové architektury. Poskytuje jednoduché řádkově orientované rozhraní, které je dostatečně silné na práci s širokou škálou úloh. Výhodou v našem prostředí je to, že Charon lze používat jak v *META Centru* tak ve VOCE, takže uživatelé mohou používat úplně stejné příkazy pro práci v těchto, jinak velmi odlišných, infrastrukturách.

I když na první pohled složité, gridové prostředí dneška je poměrně vstřícné a nabízí velké množství nástrojů pro řešení úloh, které jsou náročné na čas, síťovou kapacitu, úložné prostory nebo schopnosti administrátorů. VOCE bylo vytvořeno pro podporu řešení těchto problémů a je otevřeno všem uživatelům, kteří potřebují bohaté prostředí, které Gridy nabízí.

Literatura

- [1] D. Kouřil. „Bezpečnost v distribuovaném prostředí.“ *Zpravodaj ÚVT MU*. 2005, roč. 15, č. 4, s. 2–6. □

³Případnou návštěvu doporučujeme domluvit předem.

⁴<http://meta.cesnet.cz/>

⁵<http://egee.cesnet.cz/cs/voce/Charon.html>