

OVĚŘENÍ MODELU
OBJECT-BASED STORAGE DEVICE
PRO AFS

Závěrečná zpráva projektu

Michal Švamberg

Luboš Kejzlar

Západočeská univerzita v Plzni
Centrum informatizace a výpočetní techniky

e-mail: {svamberg,kejzlar}@civ.zcu.cz

2. června 2010

Závěrečná zpráva projektu Fondu rozvoje CESNET, z.s.p.o. pro rok 2010, vedená pod číslem 293/2009. Projekt je zařazen do oblasti I, tématického okruhu A. Řešitelem projektu je Ing. Michal Švamberg.

1 Postup a způsob řešení

Projekt byl řešen v souladu se zadáním. Pro potřeby testování byly zakoupeny dva servery, každý se 4 CPU jádry, 8 GB RAM, 2×1 Gb Ethernet a čtyřmi nezávislými rychlými disky. Hardwarová konfigurace serverů byla volena tak, aby každý z nich bezproblémově umožňoval paralelní běh čtyř virtuálních strojů bez jejich vzájemného ovlivňování a degradace výkonu.

V testovacím prostředí byly oba servery připojeny přímo do páteřního switchu dvěma 1 Gb linkami. Každý server hostoval čtyři virtuální stroje s 1 GB RAM, dedikovaným CPU a vyhrazeným pevným diskem. Do sítě byly virtuální stroje připojeny po dvojicích. Pomocí základních výkonnostních testů bylo ověřeno, že každý virtuální stroj disponuje dostatečným výkonem (CPU a I/O průchodnost), který není ovlivněn zatížením ostatních strojů.

Na všech virtuálních strojích byl nainstalován distribuovaný souborový systém OpenAFS s Object-based Storage Device rozšířením (OpenAFS-OSD). Vzhledem k možným nežádoucím interakcím byla testovací AFS buňka nakonfigurována jako zcela nezávislá a bez vazby na produkční prostředí. Tato změna oproti původnímu předpokladu vyžadovala doplnění infrastruktury dalším, tzv. databázovým serverem, zodpovědným za autentizaci a poskytování metadat nezbytných pro provoz AFS.

Původním záměrem zadání bylo implementovat dvě oddělené konfigurace OpenAFS a OpenAFS-OSD sloužící ke vzájemnému porovnávání výsledků prováděných testů. V průběhu řešení projektu jsme však ověřili, že implementace OpenAFS-OSD je i na nejnižší úrovni plně zpětně kompatibilní a dovoluje tedy bezproblémové porovnání obou variant v rámci jediné AFS buňky.

Vlastní realizace popsané serverové infrastruktury byla bohužel provázena řadou nepředvídaných organizačních a technických problémů. Prvním bylo zpoždění celé dodávky o více jak dva měsíce (deklarovaným důvodem dodavatele byl nedostatek serverových paměťových modulů na trhu). Následně jsme se potýkali s nekompatibilitou virtualizační technologie Xen s novými jádry Linuxu, jež jsme byli nuceni použít kvůli ovladačům síťových karet. To nakonec vyústilo v žádost o změnu projektu na nákup síťových karet a doplnění serverů o vzdálený management.

Řada problémů a nejasností se pochopitelně vyskytla také při vlastní konfiguraci testovací infrastruktury a jejích následných funkčních testech. Díky získaným kontaktům a prostředkům z projektu *Erasmus pro zaměstnance*, jsme měli neocenitelnou příležitost osobní konzultace a řešení problémů přímo s vývojáři ve výzkumném centru IPP Garching¹. V průběhu této týdenní stáže jsme odladili testovací infrastrukturu a ověřili základní funkčnost všech komponent rozšíření OpenAFS-OSD.

Pro výkonnostní testy byly využity tři veřejné počítačové učebny, vybavené celkem 50 ks „standardních“ PC stanic rozdílné hardwarové konfigurace. Na stanicích byl provozován OS Debian Linux ve 32 i 64-bitovém provedení s OpenAFS-OSD v rozdílné

¹Rechenzentrum der Max-Planck-Gesellschaft am Max-Planck-Institut für Plasmaphysik

konfiguraci (v závislosti na testu a HW konfiguraci). Testy zpětné kompatibility probíhaly také na systému ORIONXP se standardní instalací OpenAFS.

Každá stanice byla připojena do sítě 100 Mb linkou na switch, který má přímou 1 Gb linku do stejného páteřního switchu jako zbytek infrastruktury testovací buňky.

Popsané testovací prostředí bylo rozšířeno o dedikovanou vysoce výkonnou klientskou stanici připojenou 1 Gb linkou přímo do páteřního switchu.

Veškeré testy probíhaly mimo hlavní provozní dobu a o víkendech, což spolu s vhodně zvolenou síťovou topologií umožnilo prakticky eliminovat nežádoucí vliv zatížení sítě a dalších náhodných prvků na výsledky prováděných měření.

Také při přípravě klientské části vznikly problémy se vzájemnou kompatibilitou verzí OpenAFS-OSD a jádra OS². Pro odstranění problémů bylo nezbytné provést speciální úpravy kódu jádra i OpenAFS-OSD.

Kromě výkonnostních testů byla v průběhu řešení projektu ověřena funkčnost a reálná využitelnost nově implementovaných funkcí, zejména RO/RW mirror a striping.

Pro simulaci základních vlastností HSM byly použity dva levné vysokokapacitní disky pořízené z projektu. V rámci simulace vznikl i zajímavý koncept HSM řešení včetně možné integrace zálohovacího systému Bacula (využívaného provozně na ZČU).

Položka rozpočtu určená na zahraniční cestovné byla použita na uhrazení pobytu řešitelů na konferenci „European AFS Workshop 2009“ pořádané v Římě. Účast na této konferenci měla zásadní vliv na úspěšné řešení projektu, neboť se nám podařilo získat důležité kontakty na vývojáře OpenAFS-OSD a podrobně s nimi diskutovat uvedenou problematiku.

Na realizaci celého projektu se průběžně podílel student Martin Strbačka, kterému byla za jeho aktivní přístup vyplacena odměna 30 000,- Kč.

2 Dosažené cíle

Díky realizaci projektu vzniklo kvalitní testovacího prostředí umožňující administrátorům prakticky vyzkoušet kompletní instalaci, konfiguraci a správu všech klíčových komponent rozšíření OpenAFS-OSD.

Byla ověřena stabilita, možnost vzájemné bezproblémová spolupráce a plná zpětná kompatibilita se standardní distribucí OpenAFS a to jak na straně klientů tak i serverů.

Prakticky byla potvrzena použitelnost nových funkcí jako je RO/RW mirror, stripping či HSM archiver. OpenAFS-OSD umožňuje obecně lepší distribuci dat mezi jednotlivé fileservery a v neposlední řadě taktéž ruší dřívější koncepční omezení „nedělitelnosti“ volumů³. Bylo prokázáno, že s pomocí těchto vlastností lze dosáhnout podstatného zvýšení propustnosti (2-4×) a dostupnosti celého souborového systému.

Během pobytu na vývojářském pracovišti v Garchingu jsme se aktivně podíleli na odstranění řady závažných chyb, které se objevily v souvislosti s implementací našeho testovacího prostředí. Postranním výstupem tohoto procesu byla i příprava a vydání oficiálních balíčků OpenAFS-OSD pro distribuci Debian.

V průběhu řešení projektu jsme získali mnoho neocenitelných osobních kontaktů z řad významných uživatelů, klíčových vývojářů a členů „rady starších“, určující hlavní koncepční směry dalšího vývoje OpenAFS. Jedná se o neformální vazby, které nám

²Vzhledem k nezbytnosti zachovat rutinní provoz učebny nebylo možné měnit verzi jádra.

³Tím je možné uložit větší data do jednoho volumu než je velikost největšího diskového oddílu.

i případným dalším zájemcům, umožní aktivně ovlivňovat vývojové priority, sdílet cenné zkušenosti a daleko efektivněji řešit vzniklé problémy.

Z účasti na konferenci „European AFS Workshop 2009⁴“ vyplynulo neformální ocenění významu uživatelské komunity AFS v ČR a získání pořadatelství konference pro následující rok⁵.

3 Zdůvodnění změn v projektu

V projektu byla provedena změna financování, kdy nevyčerpaná částka určená pro zahraniční vložné byla přesunuta na nákup síťových karet a management modulů pro vzdálenou správu. Převod prostředků z kapitoly *Ostatní služby* byl schválen Radou Fondu rozvoje sdružení CESNET dne 28.1.2010 v maximální výši 10 000,- Kč.

4 Konkrétní výstupy

Všechny technické materiály a kontakty související s projektem jsou dostupné prostřednictvím WWW stránek CIV⁶.

V době psaní závěrečné zprávy bohužel nebyly ještě dokončeny všechny relevantní testy. Po jejich dokončení a vyhodnocení (konec června) bude publikována závěrečná technická zpráva popisující strukturu a vlastnosti OpenAFS-OSD, detaily pilotní implementace a výsledky provedených měření.

V rámci projektu byla uvolněna oficiální distribuce balíčků OpenAFS-OSD pro Debian.

Vzhledem k technickým problémům s přípravou testovacího prostředí jsme nestihli dosažené výsledky prezentovat na konferencích před ukončením projektu.

Hlavní souhrnná prezentace proběhne v září na konferenci European AFS & Kerberos Conference 2010⁷, pořádané v součinnosti se sdružením CESNET. Další nabídky na prezentaci byly odeslány organizátorům konferencí EurOpen.CZ⁸ a LinuxAlt⁹ pořádaných v listopadu.

Informační seminář pro zaměstnance CIV proběhl v květnu, pro další zájemce z řad studentů a zaměstnanců ZČU bude uspořádán počátkem nového školního roku.

5 Přínosy projektu

Projekt umožnil vytvořit dosud chybějící kvalitní testovací prostředí umožňující pilotní provoz plnohodnotné instalace OpenAFS-OSD včetně jeho postupné migrace do produkčního prostředí.

Získané osobní kontakty přispěly nejen k úspěšnému vyřešení projektu, ale budou i nadále sloužit jako neocenitelný zdroj informací a prostředek pro vzájemnou výměnu

⁴<http://www.dia.uniroma3.it/afscon09/>

⁵<http://afs2010.civ.zcu.cz>

⁶http://support.zcu.cz/index.php/CIV:Granty/Cesnet_293_2009

⁷<http://afs2010.civ.zcu.cz>

⁸<http://www.europen.cz>

⁹<http://www.linuxalt.cz>

zkušeností. S několika zahraničními odbornými skupinami byla navázána dlouhodobější spolupráce nejen v oblasti OpenAFS, ale i souvisejících projektech (např. iRODS).

V průběhu projektu došlo ke značnému zviditelnění akademické uživatelské komunity v ČR a zvýšení jejího kreditu.

Díky projektu došlo k podstatnému rozšíření odborné kvalifikace řešitelů i spolupracujícího studenta.

6 Tisková zpráva

Ve spolupráci Západočeské univerzity v Plzni a Fondu rozvoje CESNET, z.s.p.o bylo úspěšně vybudováno testovací prostředí pro implementaci distribuovaného souborového systému OpenAFS s rozšířením Object-based Storage Device. Výsledky testů potvrdily reálnou využitelnost nových vlastností v produkčním prostředí a naznačily možnosti výrazné zvýšení celkového výkonu a dostupnosti systému AFS. Projekt umožnil rozšíření odborné kvalifikace řešitelů a spolupracujícího studenta. Na projekt navazuje příprava mezinárodní konference *European AFS & Kerberos Conference 2010*¹⁰ pořádané ve dnech 13. až 15. září 2010 na Západočeské univerzitě v Plzni.

7 Výkaz hospodaření s prostředky

Na projektu se podílely dva subjekty a to Fond rozvoje CESNET a Západočeská univerzita v Plzni. Z prostředků Fondu rozvoje CESNET byly hrazeny dva servery včetně dokupovaných karet dle schválené změny projektu, ostatní položky hradila Západočeská univerzita.

Zaúčtování provedlo ekonomické oddělení Západočeské univerzity v Plzni dle platných předpisů. Doklady o účtování jsou k nahlédnutí taktéž na ekonomickém oddělení ZČU.

Náklady	Hodnota	Hrazeno
2× server	170 598,-	FR CESNET
2× síťová a management karta	9 960,-	FR CESNET
2× HDD pro simulaci HSM	7 550,-	ZČU
zahraniční cestovné a vložné	34 658,-	ZČU
tuzemské cestovné a vložné	11 711,-	ZČU
odměny studentských řešitelů	30 000,-	ZČU
Celkem	264 477,-	

Tabulka 1: Tabulka nákladů v Kč včetně DPH.

Celkové náklady Fondu rozvoje CESNET tvořily 180 558,- Kč, náklady na straně ZČU 83 919,- Kč. Celková spoluúčast ZČU na projektu představuje 32%. Vzhledem k plánovaným nákladům ZČU na prezentaci výsledků bude spoluúčast ještě narůstat.

¹⁰<http://afs2010.civ.zcu.cz>

7.1 Zdůvodnění změn v hospodaření s prostředky

Nebyly čerpány prostředky z kapitoly „*Ostatní služby*“ určené na zaplacení vložného na mezinárodní konferenci. Oproti minulým ročníkům konference bylo tentokrát vložné zrušeno. Na základě schválené žádosti o změnu projektu bylo převedeno 10 000,- Kč do položky „*Dlouhodobý hmotný a nehmotný majetek*“.

Odkazy

- [1] Stránka s elektronickými zdroji projektu
http://support.zcu.cz/index.php/CIV:Granty/Cesnet_293_2009
- [2] Stránka projektu OpenAFS
<http://www.openafs.org>
- [3] Stránka projektu OpenAFS-OSD
<http://pfanne.rzg.mpg.de/trac/openAFS-OSD>
- [4] European AFS & Kerberos Conference 2010
<http://afs2010.civ.zcu.cz>
- [5] Stránka neziskové organizace EurOpen.CZ na podporu svobodného softwaru
<http://www.europen.cz>
- [6] Stránky sdružení LinuxAlt
<http://www.linuxalt.cz>