

Martin
Kučera

Diplomová práce

Inženýrská informatika
Softwarové inženýrství
2016/2017

Vedoucí práce:
Luboš Matějka

Simulátor distribuovaného souborového systému

Abstrakt

Distribuované souborové systémy (DFS) představují souborový systém přístupný pomocí počítačové sítě. Existují různé modely zapojení DFS a různé strategie pro uložení dat a výběr cesty při přenosu dat mezi klientem a serverem. Cílem této práce je vytvořit nástroj, který umožní simulaci DFS a jejich různých vlastností.

Úvod

Cílem této práce bylo prostudovat a navrhnout možnosti využití hierarchického modelu uložení dat v DFS a dále vytvořit simulátor, který dovolí modelování libovolných zapojení DFS. Nad zvoleným modelem zapojení následně umožní provést měření vybraných metod pro výběr cesty k přenosu dat mezi klientem a serverem bez hierarchického modelu a s ním.

Východiska, analytická část

Existuje mnoho způsobů uložení dat a přístupu k nim. Jedním z těchto způsobů je distribuovaný souborový systém, který přináší některé výhody oproti lokálnímu úložišti a to např. vyšší odolnost systému, vyšší dostupnost a možnost přistupovat k datům z libovolného zařízení pomocí počítačové sítě.

Existuje mnoho různých implementací distribuovaných souborových systémů, které se liší svými vlastnostmi. Jedním z těchto systémů je KivFS, na jehož vývoji se podíli studenti a vyučující Katedry informatiky a výpočetní techniky. KivFS má některé unikátní vlastnosti, jako je např. dynamické routování, používané při výběru cesty pro přenos dat a aktuálně zkoumané možnosti využití hierarchického modelu uložení dat za účelem zvýšení propustnosti. Skutečné přínosy podobných vlastností nelze dopředu jednoduše stanovit.

Testování a měření výkonu na reálném systému může být obtížné a to jak z časových nároků na implementaci tak z důvodu působení rušivých faktorů během měření. Faktory jako saturovaná linka nebo vytížení disku cizím procesem mohou nepříznivně ovlivnit naměřené výsledky.

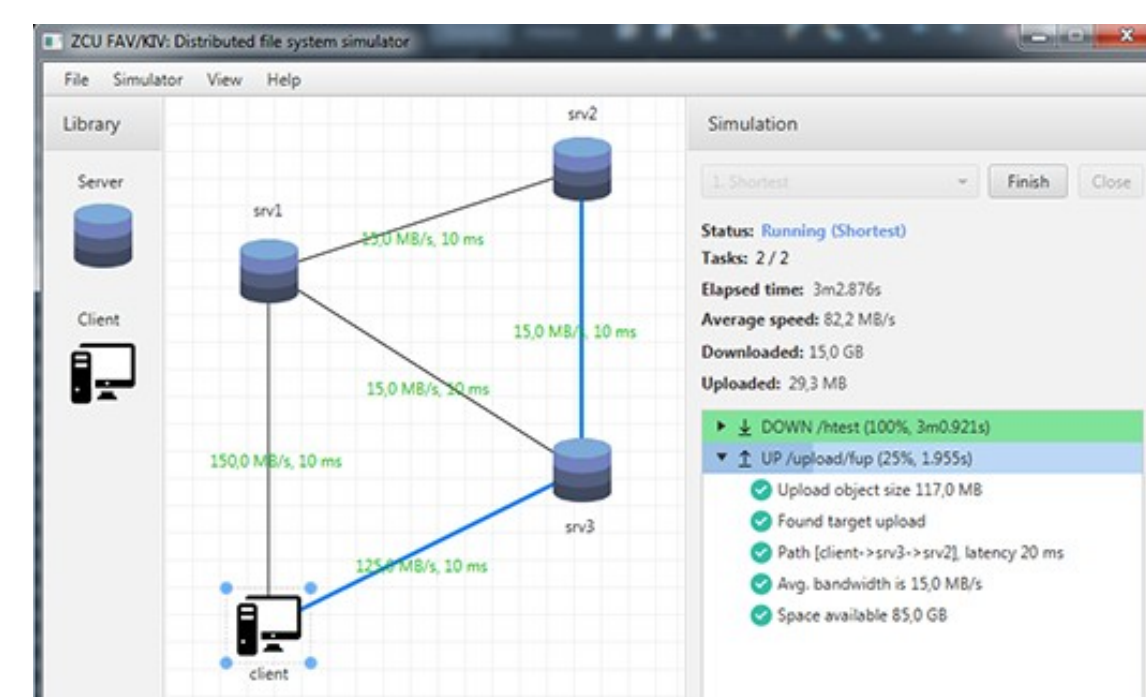
Simulační nástroje umožňují ověřit a porovnat klíčové vlastnosti systémů bez nutnosti funkčnost nejdříve implementovat a poté testovat na reálném systému. Simulace umožňuje testování výkonu bez vlivu nežádoucích rušivých faktorů a dovolí provedená měření opakovat pro porovnání výsledků.

Hlavní aspekty realizace

Výsledkem práce je multiplatformní simulátor, který umožňuje:

- modelovat libovolné zapojení DFS,
- simulovat najednou více metod pro výběr cesty,
- simulovat hierarchický model uložení dat,
- zobrazit a porovnat výsledky simulace (tabulka, graf a log operací),
- snadné rozšíření simulátoru o další funkcionality.

Pro usnadnění práce při modelování je k dispozici grafické rozhraní, pomocí kterého lze modelovat a konfigurovat celý systém.



Grafické rozhraní simulátoru

Základními komponentami při modelování jsou server, klient a spojení mezi nimi. Konfigurace serveru umožňuje definovat libovolná úložiště (kapacita, rychlost) a adresářovou strukturu serveru (složky, soubory). Během simulace je dale možné simulovat zatížení spojení pomocí zátěžové charakteristiky, která definuje propustnost spojení v závislosti na čase.

Simulace probíhá podle scénáře, který definuje jednotlivé požadavky klienta a ty jsou následně postupně zpracovávány. Během jednoho běhu lze simulovat více metod výběru cesty najednou a jejich výsledky lze po doběhnutí simulace porovnat pomocí tabulky s výsledky nebo grafem s průměrnou přenosovou rychlostí.

Dosažené výsledky

V úvodu práce byly zváženy možnosti využití hierarchického modelu uložení dat za účelem zvýšení propustnosti distribuovaného souborového systému spolu s možnými nevýhodami, které se mohou projevit při jeho použití.

Pomocí vytvořeného simulátoru bylo následně provedeno měření několika implementovaných metod pro výběr cesty a měření vlivu hierarchického modelu uložení dat na výslednou propustnost systému. Z experimentálních měření bylo vyvozeno, že existují případy, ve kterých hierarchický model uložení dat umožní zvýšení propustnosti systému.

Závěr

Byl vytvořen simulátor, který lze použít při výzkumu distribuovaných souborových systémů a to jak při návrhu nové funkcionality tak pro ověření vlastností před nasazením na reálný systém.

Simulátor byl navržen tak, aby šel v budoucnosti snadno rozšířit a to například o další metody výběru cesty nebo jiné strategie pro hierarchickou migraci dat.