České vysoké učení technické v Praze
Fakulta elektrotechnická
Software Engineering & Networking

Projekt Fondu rozvoje sdružení CESNET-513/2014/1
HS: 13144 / 830 / 8301442C
Integrace datových služeb vědecko-výukové skupiny

Bezpečnost při zálohování a archivaci

Technická zpráva

Jan Kubr, Pavel Strnad, Ondřej Macek, Ondřej Votava

Červen 2015

Obsah

Úvod

Práce výzkumných skupin vyžaduje snadné sdílení dat mezi jednotlivými členy těchto skupin. Kromě sdílení jsou pro výzkumnou skupinu důležité i další aspekty jako je šifrování dat, zjednodušení přístupu uživatelů k těmto datům, možnost nastavení procesů data managementu k datům a integrace všech služeb do společného prostředí. Komerční aplikace však často nedokáží zcela uspokojit všechny požadavky, proto se v projektu zaměřujeme hlavně na tyto speciální požadavky. Výhodou použití úložišť CESNET je možnost spojit sdílení dat s dlouhodobou archivací.

Náplní této technické zprávy je popis požadavků kladených na zabezpečení zálohovaných a archivovaných dat. Na základě této zprávy budou definovány požadavky na integrační portál (viz kapitola Požadavky).

Uživatelé a základní scénáře užití systému

Uživatelé, kteří systém mohou využívat, musí mít přiřazenu alespoň jednu z následujících rolí.

- Anonymní uživatel
- Akademik (Přednášející, cvičící, výzkumník)
- Vlastník dokumentu

Činnost vědecko-výukové skupiny vyžaduje společnou tvorbu či přípravu dokumentů. Ať už se jedná o dokumenty využívané při výuce nebo při výzkumu, je vhodné, aby měla skupina nějaký nástroj pro jejich snadné sdílení či společnou přípravu. Při běžném provozu skupiny se setkáváme s těmito případy užití:

1. Sdílení dokumentů mezi členy skupiny
2. Sdílení dokumentů vybrané skupině studentů
3. Dlouhodobé veřejné vystavení dokumentů
4. Dlouhodobé neveřejné vystavení dokumentů
5. Časově omezené sdílení dokumentů
6. Archivace dokumentů po určité období
7. Zálohování společných dat skupiny
8. Zálohování soukromých dat uživatelů

Námi zvolené řešení musí řešit problémy, které souvisí s výše popsanými případy užití. Musí tedy správně pracovat s právy dokumentů a složek, ale také s uživatelskými rolemi.
Z pohledu souboru pak musí být uživatel schopen nastavit práva z následující množiny. V závorce se nachází čísla případu užití, pro které je tato vlastnost potřebná.

- Pouze pro čtení (1,2,3,4,5)
- Pro čtení i zápis (1,2,5)
- Čtení, zápis i smazání (1,2,5)

Tyto základní scénáře jsou řešeny v průběžné zprávě [1].

Pokročilé scénáře užití systému

Pro případ zálohování a archivace dat může být požadováno zabezpečení proti přečtení či změně dat neoprávněným uživatelem. V tomto případě se setkáváme s těmito případy užití:

1. Zálohování či archivace soukromých dat zabezpečená proti přečtení jiným uživatelem než vlastníkem.
2. Zálohování či archivace dat výzkumné skupiny zabezpečená proti přečtení jiným uživatelem než členem výzkumné skupiny.
3. Zálohování či archivace soukromých dat zabezpečená proti změně.
4. Sdílení dokumentů vybrané skupině uživatelů s vyšším zabezpečením.
5. Sdílení dokumentů u kterých musíme mít jistotu, že dokumenty nebyly změněny.

Z pohledu souboru pak musí být uživatel schopen nastavit kryptografické funkce z následující množiny. V závorce se nachází čísla případu užití, pro které je tato vlastnost potřebná.

- Zašifrování souboru takovým klíčem, který umožní dešifrování pouze vlastníkovi souboru. (1)
- Zašifrování souboru takovým klíčem, který umožní dešifrování definované skupině uživatelů. (2,4)
- Digitální podepsání souboru. (3,5)

Mezi příklady užití patří:

- Přednášející vytvoří dokument, který chce studentům nasdílet. Vloží dokument do systému a v jeho vlastnostech nastaví sdílení pro všechny studenty předmětu.
- Cvičící si vytvoří doplňující dokument, který chce nasdílet jen svému cvičení. Vloží dokument do systému a ve vlastnostech nastaví sdílení pouze pro jeho cvičební paralelku.
- Výzkumník vytvořil text, který chce, aby si přečetli jeho kolegové a okomentovali mu jej. Vloží jej do systému a vytvoří novou skupinu, která obsahuje pouze jím zvolené kolegy, a té dokument nasdílí.
- Výzkumník vytvoří závěrečnou zprávu projektu, která musí být veřejná a dostupná minimálně 10 let. Proto dokument nahraje do systému, nastaví veřejné sdílení, zvolí si URI, které lze snadno opsat a nastaví dobu sdílení na 10 let.
- Akademik pravidelně zálohuje svůj pracovní počítač/server a zálohy ukládá do bezpečného úložiště, aby je mohl v případě nutnosti využít pro obnovení pracovního počítače/serveru.
- Akademik pravidelně zálohuje svůj pracovní počítač/server a zálohy ukládá do bezpečného úložiště, aby je mohl v případě nutnosti využít pro obnovení pracovního počítače/serveru. Data musí být čitelná pouze vlastníkovi. Není ani přípustné, aby data přečetl správce systému.
- Akademik vytvoří dokument, který chce nasdílet členům své výzkumné skupiny. Jelikož se jedná o dokument podléhající NDA, potřebuje mít jistotu, že dokument přečtou pouze členové definované skupiny.

- Správce rozpočtu vytvoří sdílené účetní podklady. Členové skupiny si musí být jistí, že podklady opravdu vytvořil správce rozpočtu a nikdo tyto podklady nezměnil.

Řešení zabezpečení pro archivaci dat

V rámci řešení projektu jsme pro zálohování a archivaci zvolili program Open Bacula. Tento program používáme jako univerzální řešení pro všechny podporované systémy. Z hlediska zapojení zálohovacího pracoviště je možné pro zálohování a archivaci využívat i nástroje přítomné v jednotlivých operačních systémech (např. Time Machine v OSX). Tyto nástroje mohou poskytovat vlastní přístupy pro zašifrování zálohy. Podrobněji tedy probereme možnosti zabezpečení dat v programu Open Bacula.

Možnosti Open Baculy v zabezpečení dat

Zabezpečení komunikace – Open Bacula nabízí zabezpečení přenosu dat pomocí TLS. V případě zabezpečení komunikace pomocí TLS je vlastní archiv zálohy nezašifrovaný a každý uživatel s přístupem k archivním médiím může data v záloze přečíst. Nastavení zabezpečené komunikace je popsáno v dokumentu http://www.bacula.org/7.0.x-manuals/en/main/Bacula_TLS_Communications_E.html

Zabezpečení archivu – Open Bacula nabízí možnost šifrování dat, případně jejich podepisování, na straně File daemona (klienta). Zašifrovaná/podepsaná data jsou poté odesílána Storage daemonu. Pokud tedy zvolíme šifrování na straně File daemona, jsou data zabezpečena jak při přenosu, tak i na archivních médiích. Pokud jsou data na straně File daemona digitálně podepsána, jsou jak při přenosu, tak na archivních médiích nezašifrovaná. V okamžiku obnovy digitálně podepsaných dat jsou nejprve zkontrolovány digitální podpisy a případné nesrovnalosti jsou reportovány.

Open Bacula využívá pro šifrování systém OpenSSL, takže je možné pro šifrování používat šifry v tomto systému obsažené.

Je důležité si uvědomit, že nastavení šifrování a klíčů je společné pro daného File daemona, takže lze nastavit šifrování nezávisle pro jeden počítač. Není možné nastavit různé úrovně šifrování a hesel pro různé uživatele na stejném počítači. Open Bacula šifruje pouze obsah souborů. Metadata (jména souborů, přístupová práva a vlastník souboru) zůstávají nezašifrovaná, stejně jako rozšířené atributy.

Nastavení šifrování a podepisování dat na straně File daemona je popsáno v dokumentu http://www.bacula.org/7.0.x-manuals/en/main/Data_Encryption.html

Požadavky vzniklé zabezpečením Open Baculy

Open Bacula pro využití zabezpečení vyžaduje změny v konfiguraci File daemona a Storage daemona. Vzniká zde požadavek na integrační portál:

- možnost nastavení zabezpečení v GUI integračního portálu,
  - nastavení bezpečné komunikace,
  - nastavení lokálního šifrování,

    - nastavení lokálního podepisování souborů,
- vznik skriptů, které vytvoří správné konfigurační soubory,
- nastavení zálohování klíčů.

Řešení zabezpečení pro archivaci dat

Pro archivaci dat jsme zvolili nástroj OpenKM. Pokud bychom nepožadovali pokročilé funkce pro archivaci dat poskytované systémem OpenKM, je možné ukládat data pomocí Open Baculy. Tento postup nabízí pouze nastavení doby uložení dokumentu před jeho smazáním. Tato doba se nastavuje společně pro celou zálohu a není možné ji nastavit pro jednotlivé soubory rozdílně. Pak je možné nastavit zabezpečení dat tak, jak bylo uvedeno v kapitole Možnosti Open Baculy v zabezpečení dat.

Možnosti OpenKM v zabezpečení dat

OpenKM nabízí služby rozšíření cryptography. Toto rozšíření umožňuje provádět s dokumentem následující operace:

- Add new encrypted document – lokální zašifrování dokumentu pomocí hesla a nahrání na server,
- Encrypt document – zašifrování nešifrovaného souboru a jeho uložení na vzdálené úložiště,
- Edit encrypted document – stažení zašifrovaného dokumentu na lokální úložiště a jeho rozšifrování a editace,
- Checkin encrypted document – zašifrování souboru na lokálním úložišti (používá se po Edit encrypted document, před odesláním na vzdálené úložiště),
- Download decrypted document – stažení zašifrovaného dokumentu a jeho lokální rozšifrování pomocí hesla,
- Decrypt document – rozšifrování zašifrovaného souboru a jeho uložení na vzdálené úložiště.

Všechny operace se provádí se symetrickým klíčem a OpenKM nenabízí žádnou správu klíčů. Rozšíření cryptography není dostupné v komunitní verzi OpenKM.
http://wiki.openkm.com/index.php/Document_encryption

Požadavky vzniklé zabezpečením OpenKM

Jelikož OpenKM provádí šifrování pouze na úrovni GUI OpenKM, nevznikají žádné nové požadavky na integrační portál.

Řešení zabezpečení pro sdílení dat

V rámci řešení projektu jsme pro sdílení dat zvolili systém Alfresco. V následující kapitole jsou shrnuty možnosti zabezpečení v systému Alfresco.

Možnosti Alfresca v zabezpečení dat

Systém Alfresco umožňuje využití Encrypted Content Store. Jedná se o zašifrované úložiště, do kterého jsou ukládána vlastní data. Data jsou na úložišti šifrována šifrou se symetrickým klíčem. Klíč pro symetrickou šifru je pro každý soubor jiný a je zašifrován asymetrickou šifrou.

Při použití Encrypted Content Store nelze současně využívat jiné úložiště. Encrypted Content Store vyžaduje vlastní licenci.
http://docs.alfresco.com/5.0/concepts/encrypted-cs-home.html

Další možností pro zabezpečení dat je add-on modul Alfresco Encryption Module určený pro komunitní 4.0 verzi Alfresca. Tento modul umožňuje šifrovat/dešifrovat jednotlivé soubory pomocí symetrické šifry AES.
https://addons.alfresco.com/addons/alfresco-encryption-module

Poslední možností je šifrování a podepisování pdf dokumentů pomocí alfresco-pdf-toolkit.
https://github.com/ntmcminn/alfresco-pdf-toolkit

Požadavky vzniklé zabezpečením Alfresca

Jelikož Alfresco provádí šifrování pouze na úrovni GUI Alfresca, nebo je přímo implementováno v jádru Alfresca při použití Encrypted Content Store, nevznikají žádné nové požadavky na integrační portál.

Shrnutí

Zvolené produkty nás omezují ve splnění pokročilých případů užití, které jsme pro systém stanovili.
Možnosti splnění pokročilých požadavků jsou uvedeny v následující tabulce.

Případ užití	Možnost splnění	Požadavky na integrační portál
Zálohování či archivace soukromých dat zabezpečená proti přečtení jiným uživatelem než vlastníkem.	Zálohování OMEZENĚ Archivace OMEZENĚ	Nastavení Open Baculy
Zálohování či archivace dat výzkumné skupiny zabezpečená proti přečtení jiným uživatelem než členem výzkumné skupiny.	Zálohování OMEZENĚ Archivace OMEZENĚ	Nastavení Open Baculy
Zálohování či archivace soukromých dat zabezpečená proti změně.	Zálohování ANO Archivace OMEZENĚ	Nastavení Open Baculy
Sdílení dokumentů vybrané skupině uživatelů s vyšším zabezpečením.	OMEZENĚ	NEJSOU
Sdílení dokumentů u kterých musíme mít jistotu, že dokumenty nebyly změněny.	NE	NEJSOU

Z výše uvedeného je zřejmé, že bez velkého zásahu do zvolených nástrojů je nemožné splnit pokročilé požadavky. Zaměříme se tedy na požadavky, které je možné splnit se stávajícími nástroji.

Seznam použité literatury

[1] J. Kubr, O. Macek, P. Strnad a O. Votava, *Integrace datových služeb vědecko-výukové skupiny, Průběžná zpráva projektu,* ČVUT v Praze, 2014.