

MASARYKOVA UNIVERZITA

FAKULTA INFORMATIKY



# Fedora a Hydra

PV070 DIGITÁLNÍ KNIHOVNY

**Peter Dubovský**

**UČO 395947**

Brno, 8. December 2015

# 1. Úvod

Jedným z vážnych problémov, s ktorými sa stretávajú veľké inštitúcie je široká škála rôznych technológií, ktoré síce umožňujú funkčnú špecializáciu (digitálne knižnice, archív záverečných prác, ...) pre konkrétny typ média (video, fotky, text, ...) avšak dlhodobá správa veľkého množstva takýchto rôznorodých systémov je nie len drahá, ale aj neudržateľná kvôli nutnej správe či prevádzke. Naskytlo sa však riešenie v podobe open-source technológií menom Fedora a Hydra. Fedora je robustný, modulárny úložný systém (univerzálny repozitár), ktorý je určený predovšetkým pre správu a šírenie digitálneho obsahu. Hydra je flexibilný aplikačný framework, ktorý dopĺňa technológiu Fedora. Dokopy ide o robustné úložisko, ktoré obsahuje aplikácie a pracovné toky šité na mieru. Teda ide o technológiu pre vytváranie aplikácií nad systémom Fedora.

## 2. Vývoj Fedory a Hydry

Fedora bola pôvodne vyvinutá vedcami menom Sandy Payette a Carl Lagoze z Cornell University v roku 1997<sup>1</sup> ako architektúra pre ukladanie, správu a prístup k digitálnemu obsahu vo forme digitálnych objektov. Ako inšpirácia im poslúžil Rámec pre Distribuované služby digitálnych objektov (A Framework for Distributed Digital Object Services), ktorý vytvorili Robert Kahn a Robert Wilensky. V roku 2000 Thorton Staples a Rosser Wayland v knižnici na University of Virginia vytvorili prototyp systému digitálnej knižnice, ktorý využíval Fedora architektúru. O tri roky neskôr vyšla prvá, verejnosti dostupná, verzia open source softvéru Fedora (Fedora 1.0). V decembri roku 2014 vyšla už štvrtá verzia softvéru Fedora a pravidelne vychádzajú aktualizované vydania.<sup>2</sup>

Projekt Hydra začal o dosť neskôr ako projekt Fedora. V roku 2009 sa 3 veľké americké univerzity, Hull, Stanford a Virginia, spolu s Fedora Commons spolu stretli a uvažovali, či je možné vytvorenie aplikačného rozhrania pre doplnenie Fedory, a na ktorého vývoji by sa všetci zúčastnení podieľali. Takýmto spôsobom vzniklo i motto projektu: Ak chceš ísť rýchlo, choď samostatne, ak však chceš dôjsť ďaleko, choď v skupine. Od tej doby

---

<sup>1</sup> Fedora-commons.org. *History* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.fedora-commons.org/about/history>

<sup>2</sup> Fedorarepository.org. *About Fedora* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://fedorarepository.org/about>

vznikla veľká dynamická a udržateľná komunita okolo projektu, ktorá neustále prispieva k zdokonaleniu Hydry.<sup>3</sup>

### 3. Ciele Fedory a Hydry

Cieľom Fedory bolo vytvoriť originálnu architektúru pre flexibilné a rozšíriteľné úložisko digitálnych objektov. Také, ktoré by mohlo obsahovať obrovské množstvo rôznorodého digitálneho obsahu. Takisto dôležitým atribútom takéhoto repozitára by malo byť jeho dlhodobé zachovanie, udržanie, s čím je spojený aj ničím nerušený prístup do samotného úložiska.

Cieľom Hydry je skombinovať silu úložiska určeného pre veľké podniky, a teda správu a uchovávanie digitálnych dát s na mieru šitými rozhraniami, pracovnými tokmi a prístupovými systémami špecifickými pre rôzne typy obsahu. Môžeme si pod nimi predstaviť obrovské množstvo článkov, obrázkov, kníh, videí, fotografií atď.<sup>3</sup> Jednou z ciest ako by niečo také bolo možné sa stalo presvedčenie v otvorenosť a transparentnosť vo všetkom čoho sa týka: kódu, dizajnu, diskusií.

### 4. Popis Fedory a Hydry

#### 4.1. Fedora

Pod názvom Fedora si veľa ľudí, najmä v informatickej oblasti, predstaví operačný systém linux. V našom prípade sa však jedná o robustný, modulárny a open source úložný systém, ktorý je určený predovšetkým pre správu a šírenie digitálneho obsahu. Fedora je skratka od Flexible Extensible Digital Object Repository Architecture (vo voľnom preklade Flexibilná, rozšíriteľná úschovňa digitálnych objektov). Je vhodná pre digitálne knižnice a digitálne archívy a to ako na pre prístup k dátam, tak i pre ich uchovávanie. Takisto sa často využíva na poskytovanie špecializovaného prístupu ku komplexným, rozsiahlym zbierkam rôznych materiálov, či už ide o historické, kultúrne materiály, či vedecké údaje. Fedora má silnú užívateľskú základňu, ktorú tvoria univerzity, výskumné inštitúcie,

---

<sup>3</sup> OWENS, Trevor. *Hydra's Open Source Approach: An Interview with Tom Cramer* [online]. Publikované 13. 05. 2013. [cit. 08. 12. 2015]. Dostupné z: <http://blogs.loc.gov/digitalpreservation/2013/05/hydras-open-source-approach-an-interview-with-tom-cramer/>

univerzitné knižnice, národné knižnice, vládne agentúry i organizácie zaoberajúce sa akademickým či kultúrnym dedičstvom.<sup>4</sup>

Softvér Fedora umožňuje dlhodobý prístup k digitálnym zdrojom. Sila softvéru spočíva v jeho flexibilita. Znamená to, že ho možno použiť pre podporu všetkých typov digitálneho obsahu. Ponúka aj rozšíriteľnosť (použiteľná pre milióny objektov, malých aj veľkých) a životnosť (všetky informácie sú udržiavané v súboroch bez závislosti na softvéri, z ktorých sa môže celé úložisko vybudovať kedykoľvek). Fedora takisto poskytuje možnosť vyjadrenia bohatej sady vzťahov medzi digitálnymi zdrojmi. Na dotazovanie v úložisku, ale aj mimo neho sa používa sémantický SPARQL dotazovací jazyk. Ďalšou výhodou je interoperabilita, teda súčinnosť s ostatnými aplikáciami a službami. Existuje už mnoho príkladov, kde sa Fedora používa. Napríklad pre e-výskum, správu dokumentov, digitálne knižnice, archívy, digitálne uchovávanie, inštitucionálne úložiská a veľa ďalších.<sup>2</sup>

## 4.2. Hydra

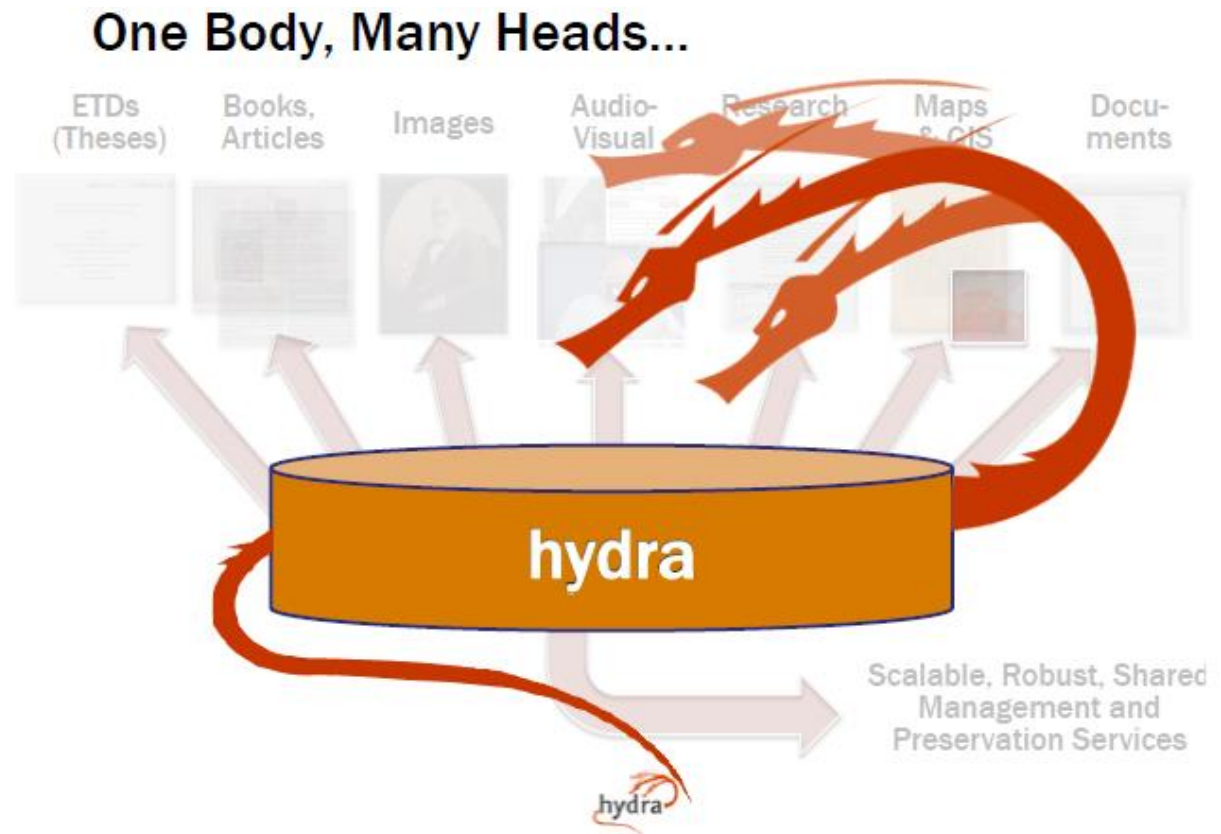
Hydra je flexibilný aplikačný rámec (framework), ktorý je postavený nad robustným úložiskom Fedora, a poskytuje mu aplikácie a pracovné toky šité na mieru. Ďalej v texte sa o nich budeme rozprávať ako o hlavách. Svoje meno dostala podľa bájnej mytologickej kreatúry, ktorej výnimočnosť spočívala v zjave. Mala jedno telo, z ktorého vychádzal veľký počet hláv a práve tento fakt, atribút sa stal jednou z vlastností Hydry (Jedno telo, veľa hláv). Ide o balíčky postavené na spolupráci, ktoré môžu byť upravené a adaptované tak, aby vyhovoval konkrétnym potrebám.

Veľkou výhodou Hydry je jej silná a početná komunita vývojárov a osvojiteľov, ktorá neustále rozširuje a posilňuje jej Jadro. Táto výhoda existuje aj vďaka faktu, že ide o open architecture (otvorenú architektúru), teda voľne dostupnú pre verejnosť. Ide teda o distribuované a skupinové úsilie už od začiatku projektu.<sup>5</sup>

---

<sup>4</sup> Fedora-commons.org. *About Fedora* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.fedora-commons.org/about>

<sup>2</sup> Fedorarepository.org. *About Fedora* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://fedorarepository.org/about>  
5 DURASPACE. 9 25-12 *DuraSpace Hot Topics, Slides, Introduction to Hydra* [online]. Publikované 25. 09. 2012 .[cit. 08. 12. 2015]. Dostupné z: <http://www.slideshare.net/DuraSpace/9-2512-duraspace-hot-topics-slides-introduction-to-hydra>



Obrázok reprezentujúci technológiu Hydra a jej hlavy. Dostupné na <http://www.slideshare.net/Tom-Cramer/hydra-for-cni-spring-2014-33003596>

### 4.3. Komponenty Hydry

Hydra je vybudovaná na 4 hlavných komponentoch.

Prvým komponentom je Fedora Commons software (Fedora), ktorá poskytuje robustné a trvanlivé úložisko a správu digitálnych objektov.

Ďalším komponentom je Solr. Solr poskytuje indexy (indexáciu), vďaka ktorým je možný rýchly prístup k informáciám o zdrojoch. Solr môže byť využitý ako lingua franca, teda obsah z akéhokoľvek zdroja, ktorý môže byť generovaný ako Solr index (napríklad OPAC alebo úložisko metadátových záznamov s rôznou schémou) a potenciálne pridaný do vyhľadávacieho prostredia Hydry.

Ďalším komponentom je Blacklight, čo je Ruby on rails plugin, ktorý sa nachádza nad Solr, a ktorý poskytuje vyhľadávanie, prezeranie a prispôbené pohľady na objekty.

Posledným dôležitým komponentom je HydraHead (Hydria hlava, hlava Hydry), gem (drahokam), ktorý spolupracuje s ActiveFedora, čo je ďalší gem, ktorý poskytuje akcie vytvor, aktualizuj a zmaž (create, update, delete) objektom v úložisku, ako aj podporu

rôznych akcií týkajúcich sa správy obsahu akými sú napríklad nahratie súboru, úprava metadát či zmena oprávnení.<sup>6</sup>

#### **4.4. Konkrétne prípady využitia**

Mnoho univerzít už využíva technológiu Hydra. Jednou z nich je napríklad Northwestern University, ktorá pre ňu našla využitie hneď v dvoch prípadoch. V prvom ide o repozitár obrázkov, ktorý poskytuje rozsiahlu funkcionálnu pre zbierky obrázkov. Hlava Hydry umožňuje užívateľom manipulovať s nimi (orezanie, otočenie a podobne) a ukladať ich rôznymi spôsobmi, ako napríklad vybudovať osobnú zbierku pre použitie v prednáškach.<sup>7</sup> V druhom prípade sa Northwestern University rozhodla spolupracovať s Indiana University, z čoho vznikol projekt Avalon (Avalon Project). Avalon poskytuje silné úložisko pre použitie s video materiálom.<sup>8</sup>

## **5. Záver**

Od roku 1997 kedy vznikala Fedora cez rok 2009, ktorý bol rokom Hydry tieto technológie prešli dlhú cestu. Dneska však vidíme, že majú budúcnosť. Fedora už je vo verzii 4 a stále sa pridávajú nové aktualizácie.<sup>1</sup> Hydra je v súčasnosti využívaná mnohými inštitúciami akými sú napríklad Stanford University, University of Virginia, University of Hull, University of Notre Dame, na účely typu: inštitučný repozitár, kolekcia obrázkov, knižnica médií, elektronické úložisko pre dizertačné a diplomové práce atď.<sup>9</sup> Toto všetko sa stalo realitou vďaka kooperácii zakladajúcich univerzít, prispievateľov, ďalších organizácií, teda komunity, ktorá okolo projektu Fedora i Hydra vznikla a aj v súčasnosti prispieva a zdokonaľuje tieto technológie.

---

<sup>6</sup> Projecthydra.org. *Technical* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://projecthydra.org/technical-2/>

<sup>7</sup> Northwestern.edu. [online]. [cit. 08. 12. 2015]. Dostupné z: <https://images.library.northwestern.edu/>

<sup>8</sup> <http://www.avalonmediasystem.org/>

<sup>1</sup> Fedora-commons.org. *History* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.fedora-commons.org/about/history>

<sup>9</sup> Projecthydra.org. Applications and demos [online]. [cit. 08. 12. 2015]. Dostupné z: <http://projecthydra.org/apps-demos-2-2/>

## 6. Zdroje

Fedora-commons.org. *History* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.fedora-commons.org/about/history>

Fedorarepository.org. *About Fedora* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://fedorarepository.org/about>

OWENS, Trevor. *Hydra's Open Source Approach: An Interview with Tom Cramer* [online]. Publikované 13. 05. 2013. [cit. 08. 12. 2015]. Dostupné z: <http://blogs.loc.gov/digitalpreservation/2013/05/hydras-open-source-approach-an-interview-with-tom-cramer/>

Fedora-commons.org. *About Fedora* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.fedora-commons.org/about>

DURASPACE. *9 25-12 DuraSpace Hot Topics, Slides, Introduction to Hydra* [online]. Publikované 25. 09. 2012. [cit. 08. 12. 2015]. Dostupné z: <http://www.slideshare.net/DuraSpace/9-2512-duraspace-hot-topics-slides-introduction-to-hydra>

OWENS, Trevor. *Hydra for CNI 2014 Meeting* [online]. Publikované 01. 04. 2014. [cit. 08. 12. 2015]. Dostupné z: <http://www.slideshare.net/Tom-Cramer/hydra-for-cni-spring-2014-33003596>

Projecthydra.org. *Technical* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://projecthydra.org/technical-2/>

Northwestern.edu. [online]. [cit. 08. 12. 2015]. Dostupné z: <https://images.library.northwestern.edu/>

Avalon Partners. [online]. [cit. 08. 12. 2015]. Dostupné z: <http://www.avalonmediasystem.org/>

Projecthydra.org. *Applications and demos* [online]. [cit. 08. 12. 2015]. Dostupné z: <http://projecthydra.org/apps-demos-2-2/>