

Predmet: Distribuované programové systémy

Téma: **Distribuované vyhľadávanie**

Autori: Bc. Milan Laslop

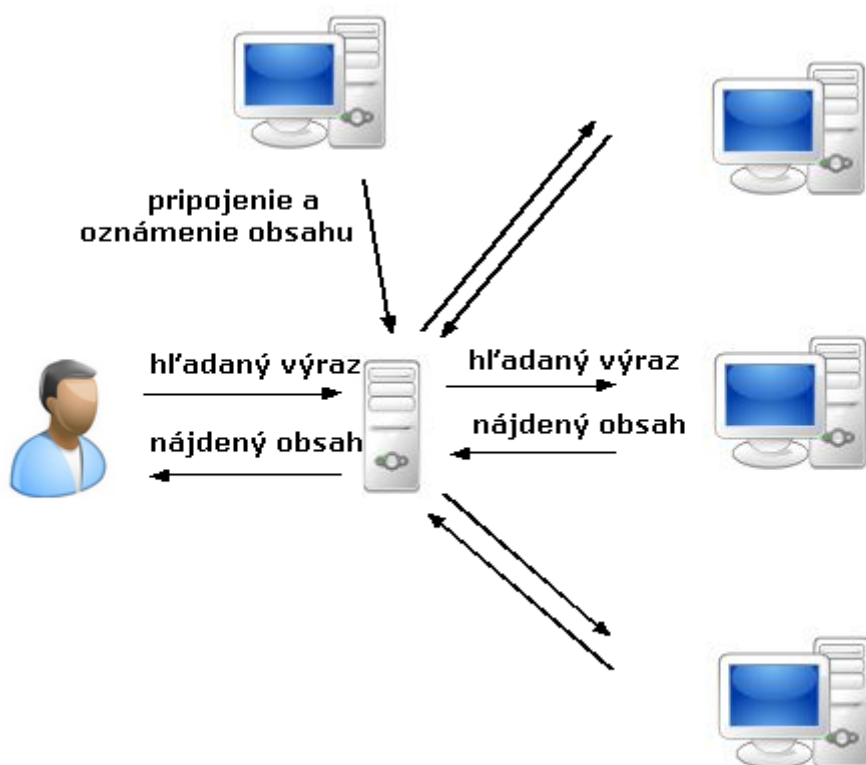
Bc. Tomáš Hurban

Popis navrhovanej aplikácie:

Systém pozostáva z 3 častí:

- *Klient* – chce vyhľadávať zadané slová alebo texty. Komunikuje len so serverom.
- *Poskytovatelia vyhľadávania* – majú pridelenú určitú časť obsahu, v ktorom vyhľadávajú. Komunikujú so serverom od ktorého získavajú hľadané výrazy a po prehľadaní obsahu mu vrátia výsledky vyhľadávania.
- *Server* – komunikuje s klientom aj s poskytovateľmi vyhľadávania. Priamo na serveri sa žiadny obsah nenachádza, zabezpečuje vyberanie vhodných poskytovateľov, ktorým posíla požiadavky na vyhľadávanie a po vrátení jednotlivých výsledkov vyhľadávania z týchto vyskladá kompletnú odpoveď, ktorú pošle klientovi.

Komunikácia medzi jednotlivými časťami:



- Inicializácia poskytovateľa:
 1. Pri pripojení poskytovateľa vyhľadávania k serveru, poskytovateľ oznámi serveru obsah lokálnej časti obsahu, v ktorom môže vyhľadávať. Server mu oznámi, ktoré časti obsahu by ešte mal mať a od ktorých iných poskytovateľov si ich má vypýtať.
 2. Poskytovateľ vyhľadávania získa časti obsahu od iných poskytovateľov.
- Vyhľadávanie:
 1. Klient zadá hľadaný výraz.
 2. Server pošle požiadavky jednotlivým pripojeným poskytovateľom vyhľadávania na základe obsahu v ktorom môžu vyhľadávať.
 3. Poskytovatelia obsahu vrátia nájdené texty.
 4. Server tieto čiastočné výsledky poskladá do finálnej odpovede klientovi.

Distribúované problémy, ktoré bude riešiť aplikácia:

- *Riešenie výpadku poskytovateľa* – ak sa poskytovateľ odpojí od servera počas vyhľadávania alebo nevráti nájdený obsah do určitého času.
- *Optimalizované vyhľadávanie* – server si pamätá, ktorému poskytovateľovi koľko trvalo vyhľadávanie v danej časti a snaží sa vyberať najrýchlejších poskytovateľov.
- *Distribúvanie obsahu* – server oznamuje poskytovateľom vyhľadávania po ich pripojení, ktoré časti obsahu majú mať a tým zabezpečuje rovnomerné distribúovanie a zálohovanie obsahu. Navyše treba riešiť inicializáciu celého systému (musí napríklad existovať jeden poskytovateľ, ktorý na začiatku má celý obsah, v ktorom sa bude hľadať, a tak sa tento obsah dá od neho distribuovať ďalším poskytovateľom, ktorí môžu vzniknúť neskôr).

Dáta uložené na serveri:

Server nemá uloženú žiadnu časť obsahu, v ktorom sa hľadá, má uložené len informácie o častiach obsahu a poskytovateľoch vyhľadávania:

- pripojení poskytovateľa (ku každému jeho ID, časti obsahu, ktoré má a ako dlho mu trvalo vyhľadávanie v každej časti)
- stav pripojených poskytovateľov – inicializácia, pripravený na vyhľadávanie, práve vyhľadáva, odpojený
- všetky existujúce časti obsahu (ku každej časti obsahu zoznam poskytovateľov, ktorí ju majú)