

Grafové dotazy sloužící k vyhledávání vzorů v grafech dovolují flexibilní zkoumání grafů podobně jako SQL relačním datům. V této práci navrhujeme a vylepšujeme klíčové komponenty distribuovaného grafového dotazovacího systému běžícího pouze v hlavní paměti. Zaprvé jsme optimalizovali vyhledávání vzorů, které používá distribuované asynchronní vyhledávání do hloubky (DFS) za pomoci prohledávání do šířky (BFS). Chytrou kombinací obou přístupů jsme využili jejich předností. DFS umožňuje striktně omezit spotřebovanou paměť a BFS zase umožňuje dosahovat lepších výkonů díky lepší paralelizovatelnosti, vyvažování zátěže a lepší lokalitě přístupů. Zadruhé jsme představili originální algoritmus pro distribuované vyhledávání dosažitelných cest za pomoci regulárních výrazů (anglicky RPQ). Tyto dotazy dovolují vyhledávat cesty libovolné délky za pomoci syntaxe podobné regulárním jazykům. Náš návrh zachovává vlastnosti DFS algoritmu, nad kterým je algoritmus postaven. Dovoluje efektivně kontrolovat spotřebu paměti během vyhledávání, a taktéž dosahuje skvělého výkonu a škálovatelnosti. Zatřetí jsme vylepšili plánování dotazů, což je jedna z nejdůležitějších součástí každého dotazovacího systému, jelikož velkou měrou ovlivňuje jeho výkon. Ovšem vybrat „nejlepší“ plán je velmi složité, jelikož výkon systému, speciálně těch distribuovaných, ovlivňuje mnoho aspektů. Představili jsme tedy jednoduchý mechanismus pro sbírání informací o běhu dotazu, který následně pomůže s vybráním toho nejlepšího plánu pro dané běhové prostředí.