

Menej je niekedy viac: Mapreduce a Flume v paralelných výpočtoch (pozvaná prednáška)

Richard Kráľovič

Google Zürich, Brandschenkestrasse 110, 8002 Zürich Switzerland

V porovnaní so sekvenčnými výpočtami sú výpočty využívajúce paralelizmus výrazne zložitejšie. Z teoretického pohľadu je veľmi prirodzeným modelom paralelných výpočtov model PRAM, ktorý je analógiou modelu RAM často používaného v teórii zložitosti. Z praktického pohľadu je však zaručenie úplnej synchronizácie a uniformného prístupu k zdieľanej pamäti, garantovaných modelom PRAM, značne problematické. Pre realizáciu paralelných výpočtov sa preto často využívajú distribuované systémy s rôznou mierou asynchrónnosti.

Klasické systémy používané pre implementáciu paralelných výpočtov v distribuovaných systémoch, ako napr. MPI, poskytujú veľkú flexibilitu. To však so sebou prináša aj nevýhody. Používateľ sa totiž musí postarať o mnohé technické detaily, ako napr. správnu synchronizáciu výpočtu, odolnosť voči chybám, a pod., čo implementáciu paralelných algoritmov výrazne komplikuje.

Alternatívou k takýmto všeobecným systémom pre využitie paralelizmu sú systémy, ktoré kladú isté obmedzenia na komunikačnú štruktúru paralelného výpočtu. Sem patrí napr. systém MapReduce, určený na spracovávanie veľkého množstva dát. Používateľ tu stráca flexibilitu, čo znamená, že použitie takýchto systémov nie je vždy vhodné. Na druhej strane, používateľ sa môže sústrediť na samotné jadro logiky paralelného výpočtu; odolnosť voči chybám je k dispozícii bez jeho zásahu.

Z pohľadu užívateľa a pozostáva výpočet v MapReduce pozostáva dvoch fáz: Vo fáze Map môže užívateľom určená funkcia vytvoriť, pre každú vstupnú položku, niekoľko párov (kľúč, hodnota). Vo fáze Reduce je, pre každú použitú hodnotu kľúča, zavolaná užívateľom definovaná funkcia, ktorá spracuje všetky hodnoty prislúchajúce k danému kľúču.

Systém MapReduce je, aj napriek jeho jednoduchosti, použiteľný pre prekvapivo veľké množstvo problémov. V mnohých situáciách je však na riešenie problému potrebných niekoľko cyklov MapReduce a ich ručné plánovanie a menežovanie býva zväčša zdĺhavé. Na uľahčenie tejto situácie bol navrhnutý systém Flume, ktorý poskytuje užívateľovi komfortnú abstrakciu nad MapReduce. Flume poskytuje dátový typ pre paralelizovaný súbor dát, s ktorým potom môže užívateľ pracovať ako s jednoduchou premennou. Na súbore dát je možné spustiť paralelnú operáciu, ktorá aplikuje danú funkciu na každú položku jednotlivo. Okrem toho je možné preusporiadať súbor obsahujúci páry (kľúč, hodnota) podľa kľúča, podobne ako v systéme MapReduce. Knižnica Flume na základe vykonaných operácií navrhne, zoptimalizuje, a vykoná sériu MapReduce cyklov, ktoré realizujú užívateľom požadované operácie.

Richard Kráľovič absolvoval štúdium informatiky na FMFI UK v Bratislave. Doktorát získal na FMFI UK, kde sa venoval najmä oblasti distribuovaných výpočtov a na ETH Zürich, kde sa v pracovnej skupine Juraja Hromkoviča zaoberal stavovou zložitou konečných automatov, aproximátnymi algoritmi, online algoritmi a advice zložitou. Po skončení postdocu na ETH Zürich pracuje pre Google Zürich.