

Hešování

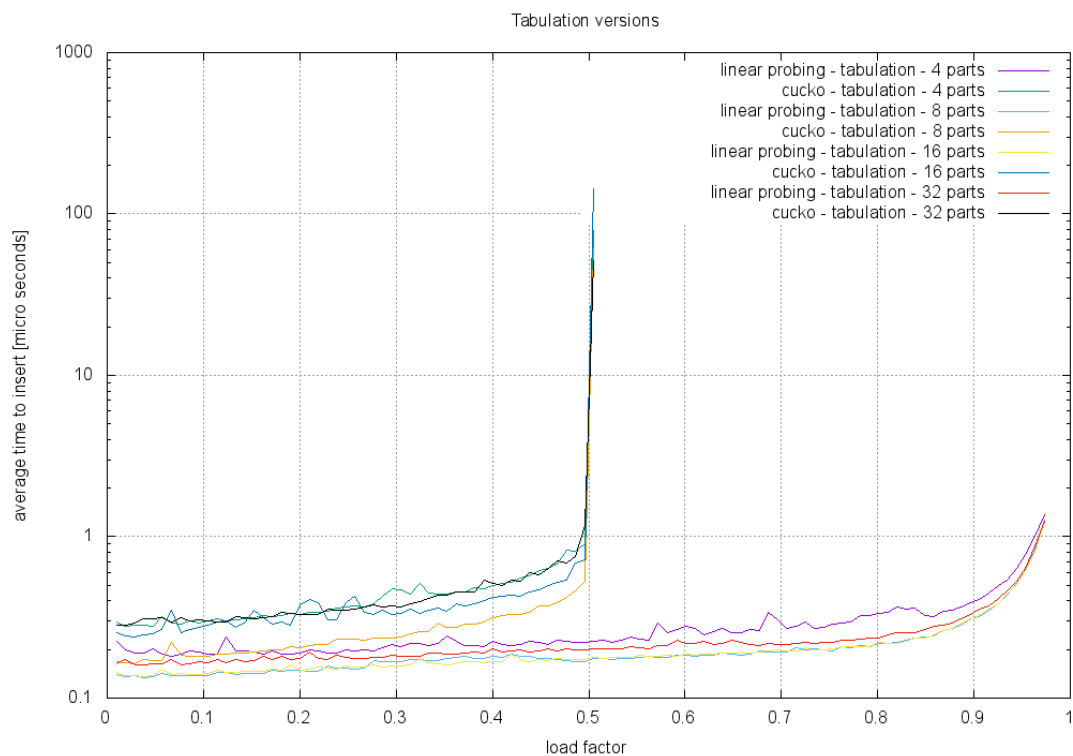
Ve všech následujících testech se používají klíče z univerza o velikosti 2^{64} .

Náhodný testovací případ

V tomto testu jsme vždy používali tabulky velikosti 2^{21} a veškeré měření jsou průměrováno přes 20000 vložení. Tabulky s lineárním přidáváním byly testovány až do faktoru zaplnění 0.98. Vkládání do kukaččí tabulky bylo přerušeno pokud se alespoň 5-krát za sebou nepovedl rehash.

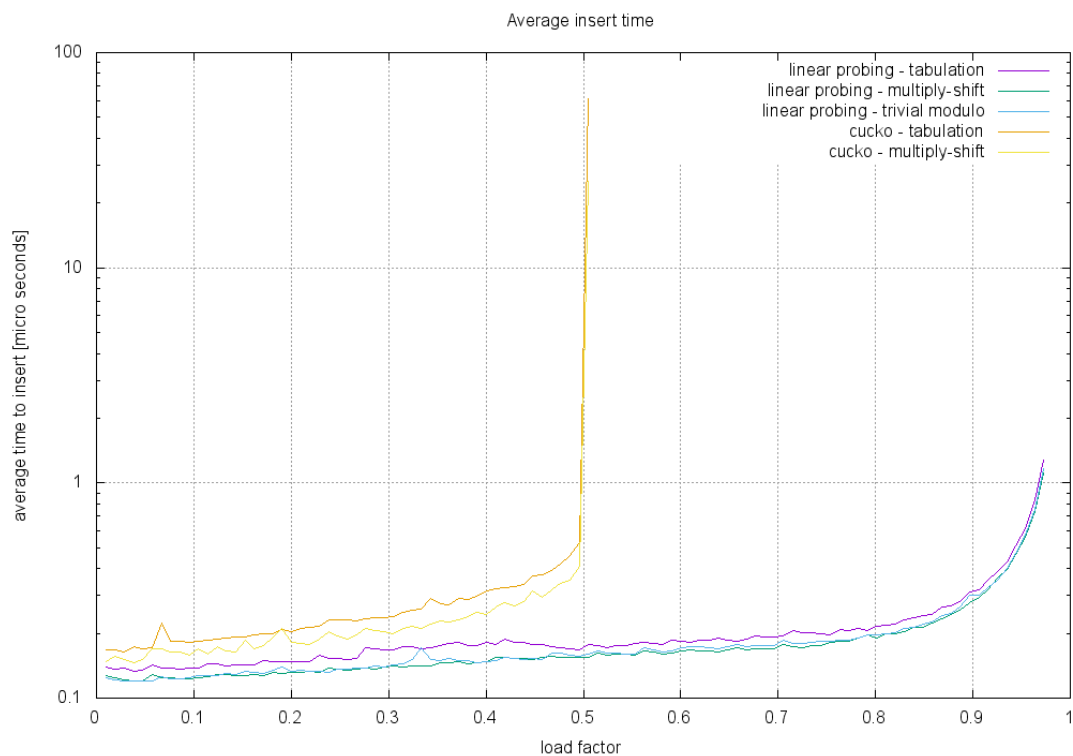
Výběr vhodné konstanty c pro tabulační hešování

Jak ukazuje následující graf nejlépe vychází dělení klíče na 8 či 16 částí. Následující grafy ukazují čas na celé vložení prvku, nikoliv pouze hešování. Nicméně, jelikož vkládáme náhodná data je vliv špatné heše resp. počet kroků při vkládání omezen a rozdíly v časech by měly reflektovat dobu výpočtu heše.



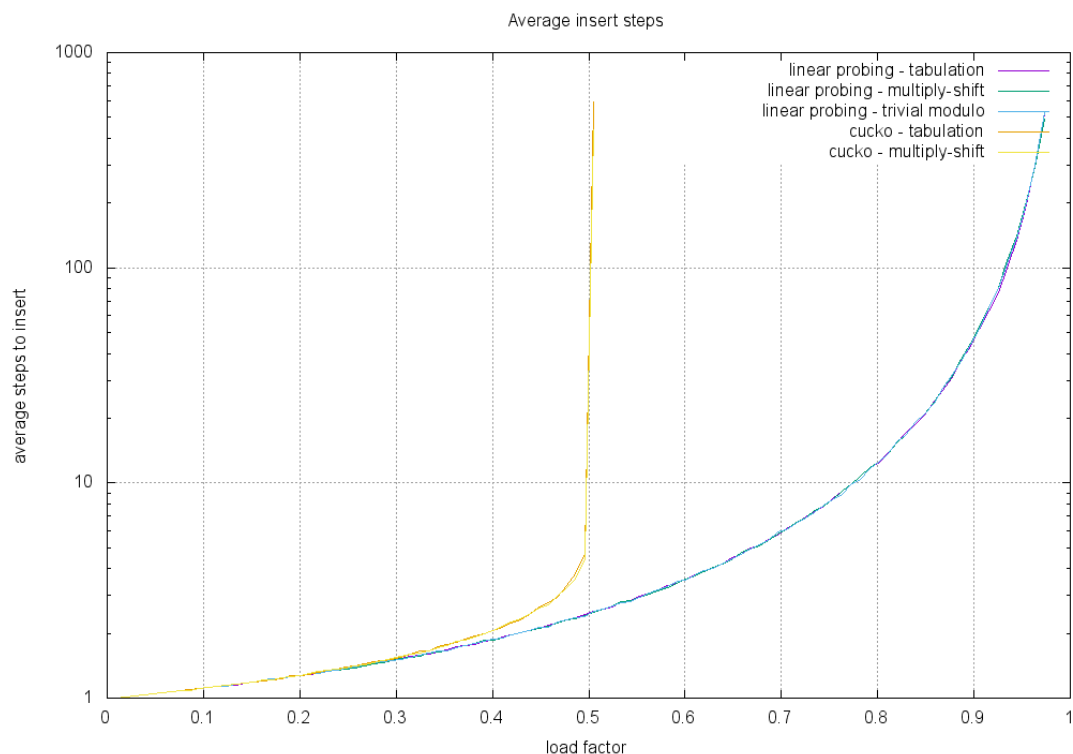
Porovnání průměrné doby na vložení do tabulky

Z grafu je vidět, že naivní a multiplikativní hešovací funkce potřebuje průměrně nejmenší dobu na insert. Je to způsobeno tím, že do tabulky vkládáme náhodná data. Proto doba reflektuje dobu výpočtu heše, která je pro tabulační heš nej-složitější.



Porovnání průměrného počtu kroků na vložení do tabulky

V následujícím grafu je vidět, že pokud vkládáme náhodná data, tak mezi použitými hešovacími funkcemi skutečně není rozdíl v kvalitě heše. Počty kroků jsou pro všechny typy heší srovnatelné a to v obou typech tabulek. V obou grafech je také vidět, že pro kukaččí tabulku odpovídá teoretický faktor zaplnění 0.5.



Sekvenční testovací případ

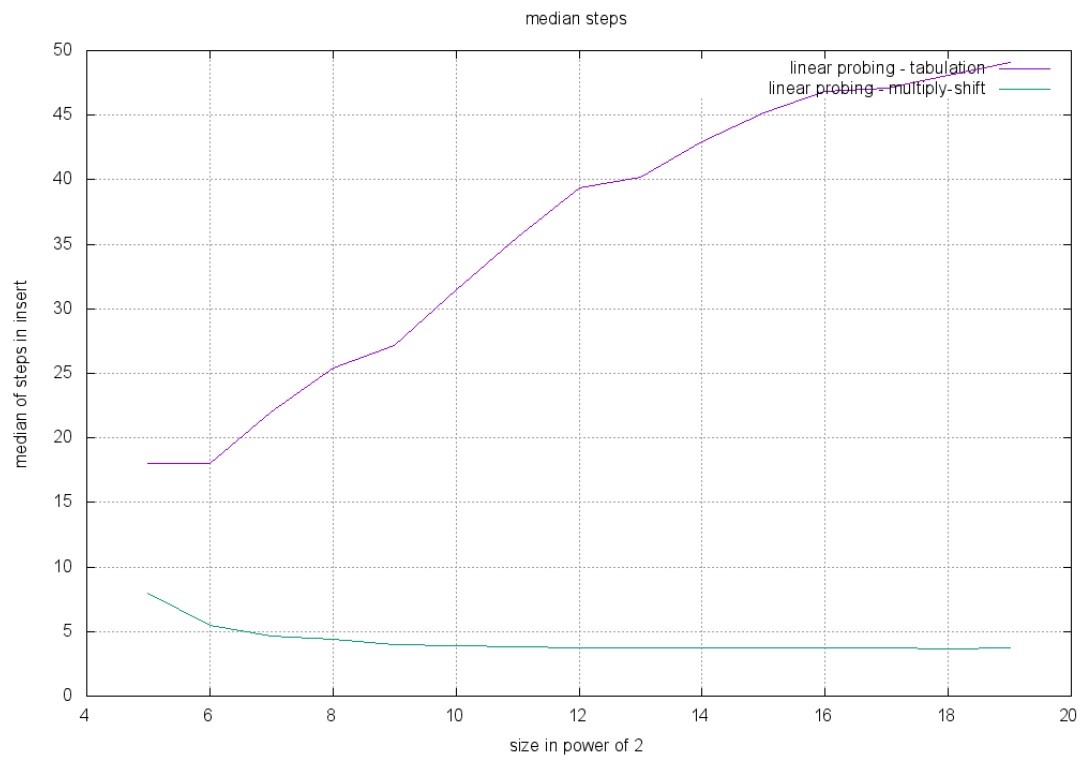
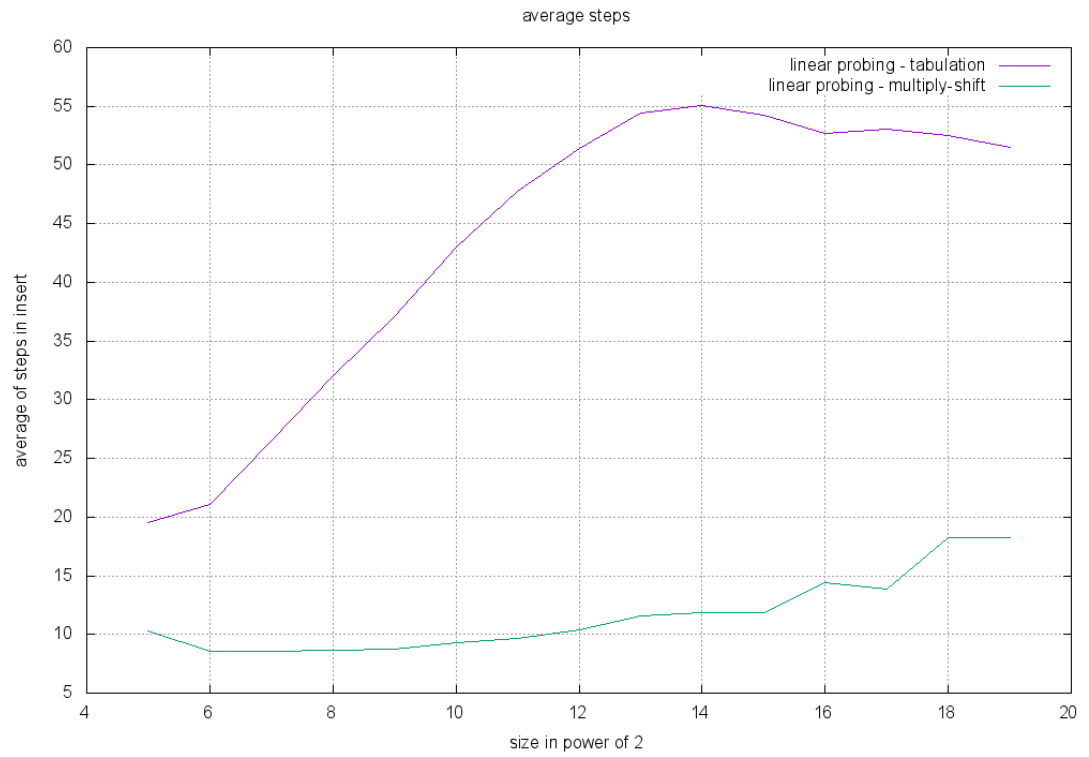
V tomto případě byla použita tabulační hešovací funkce dělicí klíč na 16 částí. Zkoušeli jsme i dělení na 8 částí, které dopadlo obdobně. Pro první verzi jsme prováděli testy pro větší tabulky až 2^{19} , proto je zde uvádíme.

Bylo prováděno vždy 100000 pokusů pro každou velikost tabulky, kde každý pokus je průměrný počet kroků na vložení pro faktor zaplnění v intervalu od 0.89 do 0.91.

Průměr a medián

Z průměru a mediánu je výsledek ve prospěch multiplikativní hešovací funkce.

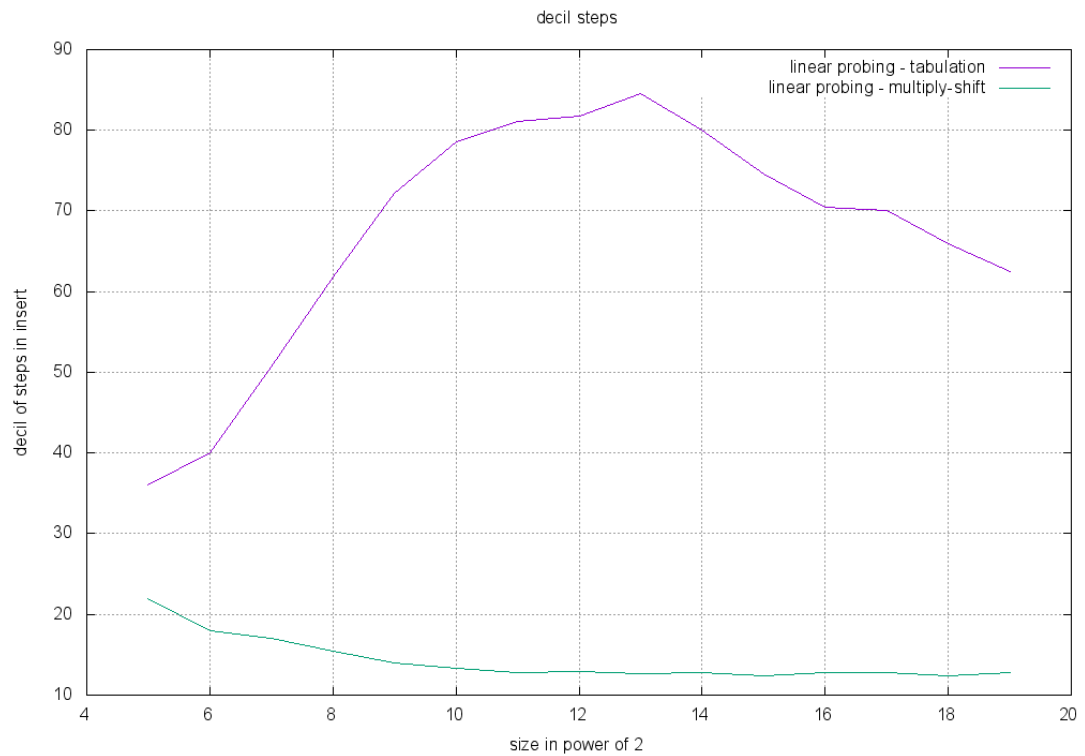
Z rozdílu mezi průměrem a medianem je vidět, že pro tabulky přibližně od velikosti 2^{14} jsou jednotlivé testy nevybalancované. Zdá se, že v některých pokusech je u tabulační funkce dosaženo velmi malého počtu kroků. A naopak u multiplikativní funkce v některých pokusech jsou počty kroků velmi velké.



Decil

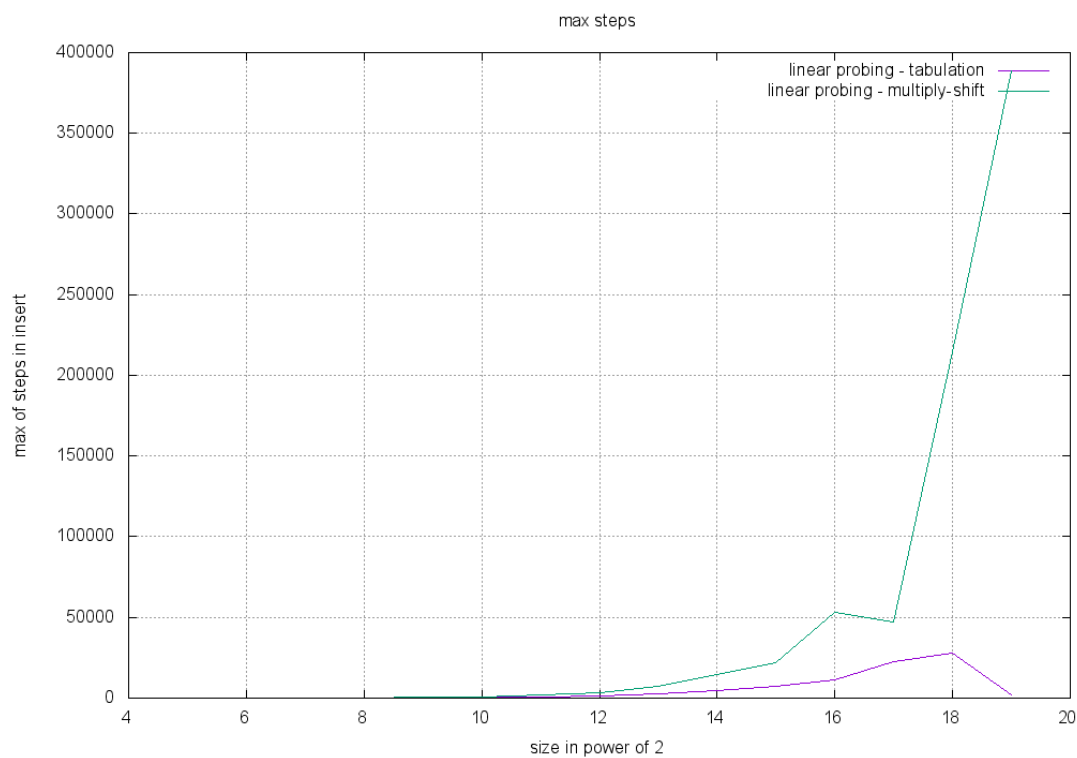
Z decilu je vidět, že 90% pokusů u multiplikativní funkce dává dobré výsledky. Nárůst počtu kroků tedy je nejvýše v 10% případů.

Pro tabulační funkci dochází pro velké tabulky, ke snižování počtů špatných pokusů.



Maximum

Zde se potvrzuje, že v případě multiplikativní funkce může dojít k hodně špatným výsledkům. Naopak tabulační funkce dává mnohem menší negativní extrém.



Minimum

Zde je situace opět opačná, multiplikativní funkce dokáže dávat mnohem lepší pozitivní hodnoty a to nezávisle na velikosti tabulky. Pro tabulační funkci se nejlepší hodnoty zhoršují s velikostí tabulky.

