

# Rozdělení přenosové rychlosti disku

Vladimír Třebický

10. května 2006

Pevné disky osobního počítače nepracují vždy stejně rychle. Rozdíly v rychlosti sekvenčního přístupu mají několik důvodů, důležitá je však vždy hladina redukce, pomocí které se na problém pohlíží.

V této práci se snažím podat informaci o statistických vlastnostech několika měřeních jednoho disku pomocí unixové utility `hdparm`<sup>1</sup>, která může mj. sloužit k měření rychlosti sekvenčního čtení. Tato po svém spuštění zamkne disk tak, aby nebylo možné k němu jinak přistupovat a zařídí si tak výhradní přístup. Takto pokryjeme nejvyšší hladinu redukce a totiž rozdíl v rychlosti díky sdílení čtecích hlav.

V nižších vrstvách redukce můžeme narazit např. na počáteční polohu čtecí hlavy, optimalizace přístupu k disku jádrem, nebo třeba NCQ<sup>2</sup>. Ty ale již nechám vesele ovlivňovat svoje data. Koneckonců, běžný uživatel toho z výše uvedeného může ovlivnit pouze minimum; proto mu bude níže uvedené spíš prospěšné.

**Data** K účelu sběru dat jsem napsal jednoduchý skript, který s prodlevou 10 s opakovaně spouští „`hdparm -t /dev/sda`“ a ukládá výslednou rychlost sekvenčního čtení.

Naměřené hodnoty jsou součástí přílohy [1]. Grafické znázornění rozložení najdete na Obr. 1. Jedná se o tisíc vzorků odebraných v době mezi 15:00 a 20:00. Čáry jsou interpolacemi četností výskytu v daném intervalu. Hodnoty použité pro vytvoření toho obrázku jsou potom v Tab. 1.

**Hypotéza** Pro účely semestrální práce vyjádřím hypotézu:

$H_0$  Naměřená data se *řídí* normálním rozdělením.

$H_1$  Naměřená data se *neřídí* normálním rozdělením.

**Testování** Porovnám naměřené hodnoty s odpovídajícím normálním rozdělením. Pomocí  $\chi^2$  testu dobré shody na hladině významnosti 5% se pokusím hypotézu vyvrátit.

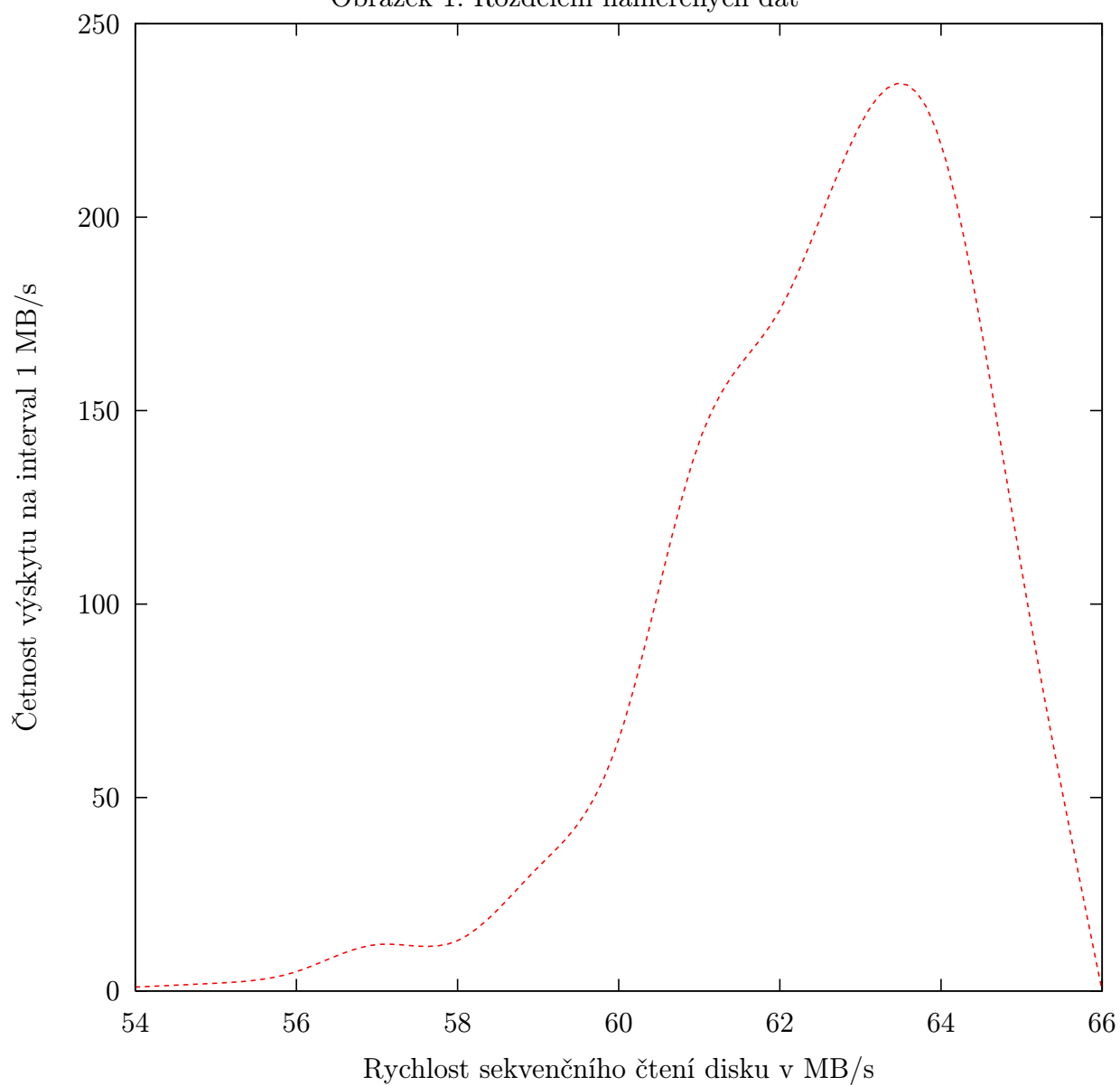
Postup volím následující:

---

<sup>1</sup>Viz. <http://sourceforge.net/projects/hdparm/>

<sup>2</sup>Native Command Queueing – technologie umožňující nelineární načítání dat z pevného disku, což může šetřit čas, který by byl ztracen při pohybech čtecí hlavy.

Obrázek 1: Rozdělení naměřených dat



Tabulka 1: Histogram.

$i$	Střed $\xi_i$	Četnost	$i$	Střed $\xi_i$	Četnost
1	54	1	7	60	65
2	55	2	8	61	142
3	56	5	9	62	176
4	57	12	10	63	224
5	58	13	11	64	219
6	59	32	12	65	109

1. Vypočítám odpovídající hranice intervalů z histogramu o granularitě 1 MB/s. Vzhledem k tomu, že výše uvedené rychlosti jsou zaokrouhlením přenosových rychlostí disku na celá čísla, jsou hranice intervalu vždy:

$$(\xi_i - 0.5; \xi_i + 0.5)$$

Kde  $\xi_i$  je střed intervalu  $i$ . Tedy:

$$\begin{aligned} h_i &= \xi_i - 0.5 \\ h_{i+1} &= \xi_i + 0.5 \\ &= \xi_{i+1} - 0.5 \end{aligned}$$

Kde  $h_i$  je dolní hranice intervalu se středem  $\xi_i$  a  $h_{i+1}$  je jeho horní hranice. Je to taktéž dolní hranice intervalu  $\xi_{i+1}$ , pokud takový interval existuje.

2. Určím parametry  $\mu$  a  $\sigma^2$  normálního rozdělení metodou maximální věrohodnosti:

$$\begin{aligned} \hat{\mu} &= \frac{1}{N} \sum_{i=1}^n \xi_i \cdot p_i \\ \hat{\sigma}^2 &= \frac{1}{N} \sum_{i=1}^n p_i \cdot (\xi_i - \hat{\mu})^2 \end{aligned}$$

Kde  $N$  je počet měření (1000),  $n$  je počet intervalů (12),  $\xi_i$  je střed intervalu  $i$  a  $p_i$  je jeho četnost. Pro svoje data obdržím  $\hat{\mu} = 62.46$  a  $\hat{\sigma}^2 = 3.41$ .

3. Ze získaných hodnot vypočtu pravděpodobnosti a normalizuji počtem měření.

$$T_i = p_i \cdot N$$

Pravděpodobnost  $p_i$  je hodnota normálního rozdělení  $N(\mu, \sigma^2)$  pro interval  $i$ . K jeho získání nám nejlépe poslouží distribuční funkce normálního rozdělení. Konkrétně po jednoduchých úpravách:

$$p_i = \Phi\left(\frac{h_{i+1} - \hat{\mu}}{\hat{\sigma}}\right) - \Phi\left(\frac{h_i - \hat{\mu}}{\hat{\sigma}}\right)$$

Kde  $\Phi$  je funkce normovaného normálního rozdělení a  $h_i$  je hranice intervalu.

4. Ze získaných hodnot určím hodnotu  $\chi^2$ :

$$\chi^2 = \sum_{i=1}^n \frac{(n_i - T_i)^2}{T_i} = 362.58$$

5. Určím odpovídající stupeň volnosti:

$$\nu = n - 1 - q$$

Pro  $n = 12$  počet intervalů a  $q = 2$  odhadovaných parametrů je:

$$\nu = 12 - 1 - 2 = 9$$

6. Podle hladiny významnosti  $\alpha = 5\%$  a stupně volnosti  $\nu = 9$  najdu odpovídající kvantil:

$$\chi_{1-\alpha}^2(\nu) = \chi_{0.95}^2(9) = 16.92$$

Podle zamítacího pravidla  $\chi^2$  testu dobré shody na hladině významnosti  $\alpha = 5\%$  a podle obdržené hodnoty  $\chi^2 = 362.58$  zamítám hypotézu  $H_0$  a přecházím k hypotéze  $H_1$ :

$$\chi_{1-\alpha}^2(\nu) < \chi^2$$

$$16.92 < 362.58$$

**Závěr** Na základě naměřených dat, jejichž vypovídací hodnota je nastíněna v úvodu, a na základě  $\chi^2$  testu dobré shody na hladině významnosti 5% zamítám hypotézu  $H_0$  a prohlašuji, že rozdělení přenosových rychlostí sekvenčního čtení disku se *neřídí* normálním rozdělením.

Mezivýpočty jsou součástí Tab. 2 a Tab. 3.

Tabulka 2: Výpočet střední hodnoty a rozptylu.

$\xi_i$	$n_i$	$\xi_i \cdot n_i$	$(\xi_i - \hat{\mu})^2$	$(\xi_i - \hat{\mu})^2 \cdot n_i$
54	1	54	71,52	71,52
55	2	110	55,60	111,21
56	5	280	41,69	208,46
57	12	684	29,77	357,34
58	13	754	19,86	258,24
59	32	1888	11,95	382,42
60	65	3900	6,03	392,39
61	142	8662	2,12	301,44
62	176	10912	0,20	36,75
63	224	14112	0,29	66,04
64	219	14016	2,38	521,40
65	109	7085	6,46	704,88
$\hat{\mu} = \frac{1}{1000} \Sigma = 62.45$			$\hat{\sigma} = \frac{1}{1000} \Sigma = 3.41$	

Tabulka 3: Výpočet teoretické pravděpodobnosti a  $\chi^2$ .

$\xi_i$	$n_i$	$\Phi\left(\frac{h_i - \hat{\mu}}{\hat{\sigma}}\right)$	$p_i$	$T_i = n \cdot p_i$	$\frac{(n_i - T_i)^2}{T_i}$
54	1	0,00426	0,00563	5,63	3,80
55	2	0,00989	0,01078	10,78	7,15
56	5	0,02067	0,01938	19,38	10,66
57	12	0,04005	0,03347	33,47	13,77
58	13	0,07352	0,04949	49,49	26,90
59	32	0,12302	0,06913	69,13	19,94
60	65	0,19215	0,09219	92,19	8,01
61	142	0,28434	0,10540	105,40	12,70
62	176	0,38974	0,11425	114,25	33,37
63	224	0,50399	0,11392	113,92	106,36
64	219	0,61791	0,10784	107,84	114,58
65	109	0,72575	0,08752	87,52	5,27

$$\Sigma = 362.58$$

## Reference

- [1] Příloha: Naměřená data.
- [2] Tabulky normálního rozložení: <http://www.fp.vslib.cz/kap/picek/tabks.pdf>
- [3] Kvantily  $\chi^2$ :  
[http://www.pef.zcu.cz/  
/pef/kmt/person/Kohout/info\\_soubory/letnisem/ruzne/tabchi.htm](http://www.pef.zcu.cz/pef/kmt/person/Kohout/info_soubory/letnisem/ruzne/tabchi.htm)
- [4] M. Bradáč: Domácí úloha z předmětu M6F:  
[http://cmp.felk.cvut.cz/navara/MVT/M6F\\_xbradac\\_domaci\\_prace\\_v2.pdf](http://cmp.felk.cvut.cz/navara/MVT/M6F_xbradac_domaci_prace_v2.pdf)