

Prestazioni degli hard disk SCSI e Serial ATA in configurazioni
differenti: disco singolo, raid hardware e raid software.

Carlo Podda, Area HPC

CRS4

Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna

Polaris, Edificio 1

Loc. Pixina Manna

09010 Pula (Cagliari), Italia

E-mail: carlo@crs4.it

Prestazioni degli hard disk SCSI e Serial ATA in configurazioni differenti: disco singolo, raid hardware e raid software.

Con la serie di test eseguiti ci siamo proposti di verificare quali sono i limiti dei dischi rigidi moderni installati nelle macchine di calcolo del CRS4. In particolare abbiamo cercato di scoprire quali sono le prestazioni dei dischi rigidi in lettura o scrittura di file di dimensioni molto grandi , in questi test sino a 8 GigaByte.

Durante i test é emerso che:

la quantità di RAM installata nella macchina influenza parecchio le prestazioni in lettura/scrittura per i file di dimensione minore/uguale a questa. Per avere dei risultati che si potessero considerare non influenzati dalla RAM (caching) abbiamo fatto i test creando e scrivendo file di dimensioni sino a 4 volte più grandi della RAM.

La differenza di prestazioni, sia in lettura che in scrittura, tra raid0 (stripping) e raid1 (mirroring) si manifesta specialmente quando i file del test hanno una grandezza uguale o maggiore a quella della RAM installata nella macchina dove eseguiamo il test.

Da rilevare inoltre come il raid implementato con controller hardware (3Ware per SATA e Zero Channel per SCSI) non abbia dato risultati migliori rispetto al raid implementato via software (raidtools-1.00.3-7).

Per i test abbiamo utilizzato il software free Bonnie (<http://www.textuality.com/bonnie>) scritto in C e scaricabile con i sorgenti dal sito.

Con Bonnie abbiamo analizzato le prestazioni di dischi di diverso tipo, SCSI e SATA, in diverse configurazioni (raid hardware e raid software in configurazione mirror o stripping e a singolo disco SATA).

Nella scrittura di un file da 8 Gigabyte le prestazioni massime sono state registrate nella configurazione con due dischi SCSI in raid0, quasi 120MB/sec. Per lo stesso file, la scrittura su due dischi SATA in raid0 sfiora i 63MB/sec. Infine su un singolo disco SATA ha superato di poco i 44MB/sec.

Nella lettura di un file da 8 GigaByte con due dischi SCSI in raid0 sono stati superati i 121MB/sec, con due hard disk SATA in raid0 abbiamo raggiunto i 72MB/sec e con il singolo disco SATA sfiorato i 46MB/sec.

Le macchine utilizzate per i test:

aiodda2.crs4.it

dual CPU Xeon 3,06GHz,

Motherboard Supermicro X5DP8-G2, cache 512KB

1 GB di RAM 266MHz DDR Registered ECC

Linux CentOS-3, kernel 2.4.21-9.0.1.EL.c0smp

HD maxtor SATA MaxLine PLUS II, 8MB cache da 250 GB, su 3WARE hotswap
SATA cage, RDC-400-SATA

HD Seagate Ultra320 SCSI SCA 80 pin, 73 GB 10K rpm, su box SCA hotswap
controller 3WARE 8506-4LP

controller SCSI AIC 7902 DUAL channel U320, con controller raid zero channel 2010S

agumu02.crs4.it

dual CPU opteron 246 2.0GHz, cache L2 1MB

Motherboard TYAN Thunder K8s S2882UG3NR

2 GB RAM DDR, PC400 ECC Registered

Linux Rocks Clusters (modified), kernel 2.6.7

HD Seagate Ultra 320 SCSI SCA80 pin 73 GB 10K rpm

Controller SCSI AIC 7902 DUAL channel U320 (il raid e' implementato via software con raidtools 1.00.3-7)

nora13

dual CPU opteron 246 2.0GHz, cache L2 1MB

Motherboard TYAN Thunder K8s S2882

2 GB RAM DDR, PC400 ECC Registered

Linux Rocks Clusters (modified), kernel 2.6.7

HD Maxtor 6Y160M0 SATA

Il test consisteva nella creazione e successiva lettura un file di dimensioni crescenti: si partiva da 100MB, 500MB, 1GB, 2GB, 4GB e 8GB. Nei test fatti su aiiodda2, macchina con architettura e sistema operativo CentOS a 32 bit, ci siamo fermati al file da 2GB. Con file di queste dimensioni siamo comunque fiduciosi che le prestazioni della macchina rimarranno simili anche per file di dimensioni maggiori.

La scrittura del file veniva fatta in due modi differenti per avere due differenti tipi di test:

a carattere (funzione putc())

a blocchi (funzione write())

Anche la lettura del file veniva fatta in due modi diversi:

a carattere (funzione getc())

a blocchi (funzione read())

Le macchine del test non svolgevano in quel momento alcun compito particolare, erano chiuse per tutti gli utenti e le loro code disabilitate. Il loro carico di CPU era praticamente zero.

Per ogni test sono state avviate dieci istanze. Dai dati ottenuti abbiamo poi estratto il valore medio. I test sono stati effettuati cercando di eliminare eventuali le situazioni “contingenti” che potevano influenzarne il risultato di un tipo di configurazione raid o di un tipo di hard disk in un particolare momento. Per esempio, veniva avviato il test su un file di dimensione 100 MB sul singolo disco SATA, una volta terminato si passava con lo stesso test sul raid0 SATA e infine sul raid1 SATA. Si procedeva allora su un file da 500 MB e così via sino a 8000MB per tornare quindi a 100MB e ripartire.

I risultati completi dei test in formato ascii oltre al file pdf di questo stesso documento sono disponibili all' indirizzo: http://hpc.crs4.it/cluster/Test_and_Benchmark/sata_vs_scsi/

In sintesi le prestazioni migliori nella creazione di un file da 8GB si sono avute con i dischi SCSI in

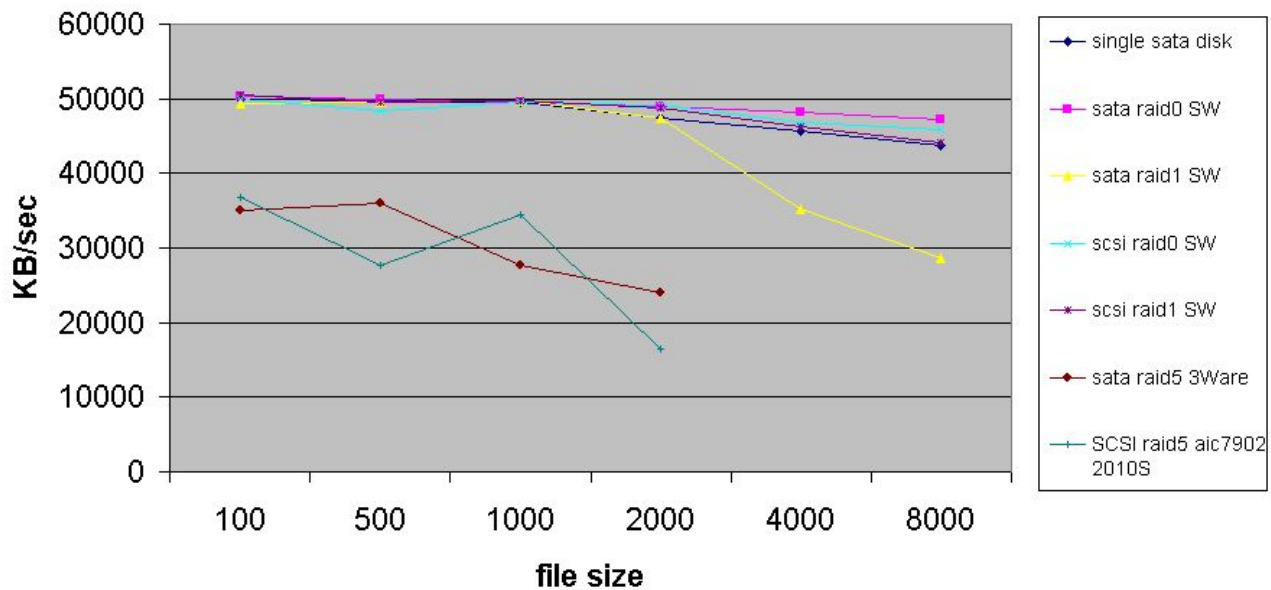
raid0 software. Anche nella lettura lo SCSI software è andato molto bene. Il guadagno di prestazioni rispetto a SATA è emerso specialmente nella creazione “a blocchi”. Non ci sono particolari differenze nella creazione del file carattere per carattere. In questa modalità di scrittura viene invece pesantemente sfruttata la CPU (come si può vedere nei file ASCII dei risultati).

Il raid software nei dischi SCSI impegna in scrittura un po' più CPU rispetto allo stesso raid su dischi SATA; questo specialmente quando il file da generare ha dimensioni almeno doppie rispetto alla RAM installata e comunque in particolare nelle operazioni di scrittura per blocchi.

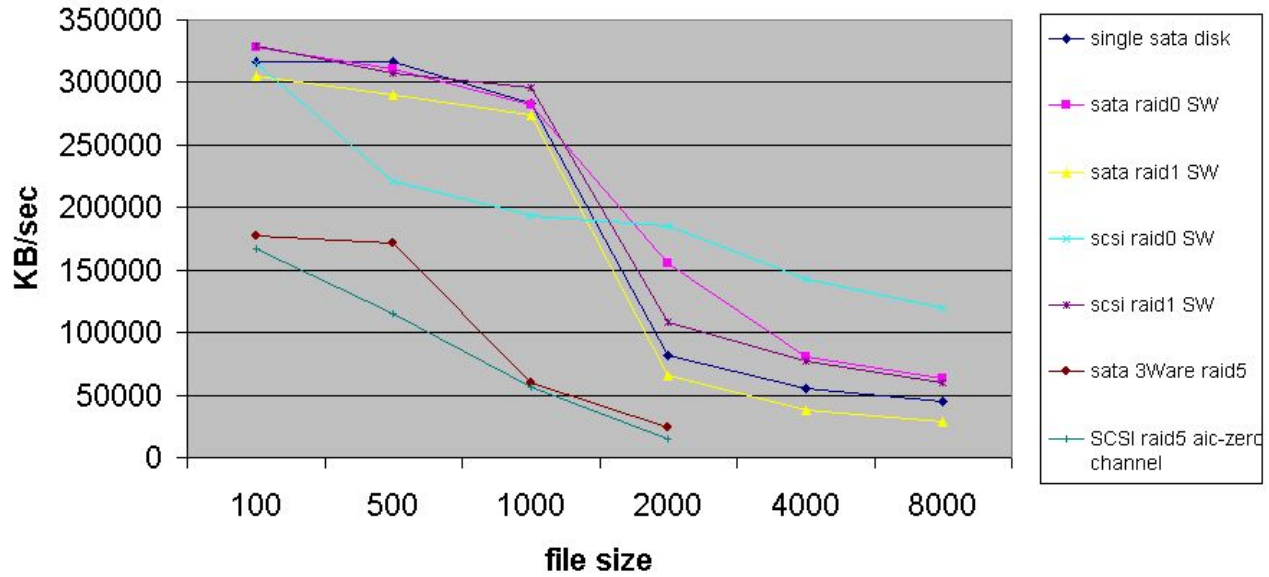
Abbastanza modesti i risultati della macchina con controller RAID hardware 3Ware e Zero Channel Adaptec.

Bisogna comunque sottolineare che si tratta di una configurazione RAID5 e che la macchina in questione è una macchina a 32 bit dual Xeon con un solo GigaByte di RAM. Possiamo ipotizzare che se nella macchina dual Xeon ci fossero stati 2 GigaByte di RAM invece di 1 solo GigaByte, i grafici risultanti sarebbero stati, perlomeno spostati di una unità verso destra. Il comportamento della macchina e il suo grafico delle prestazioni è, infatti, paragonabile a quello delle altre due macchine più performanti.

file create x character - putc()

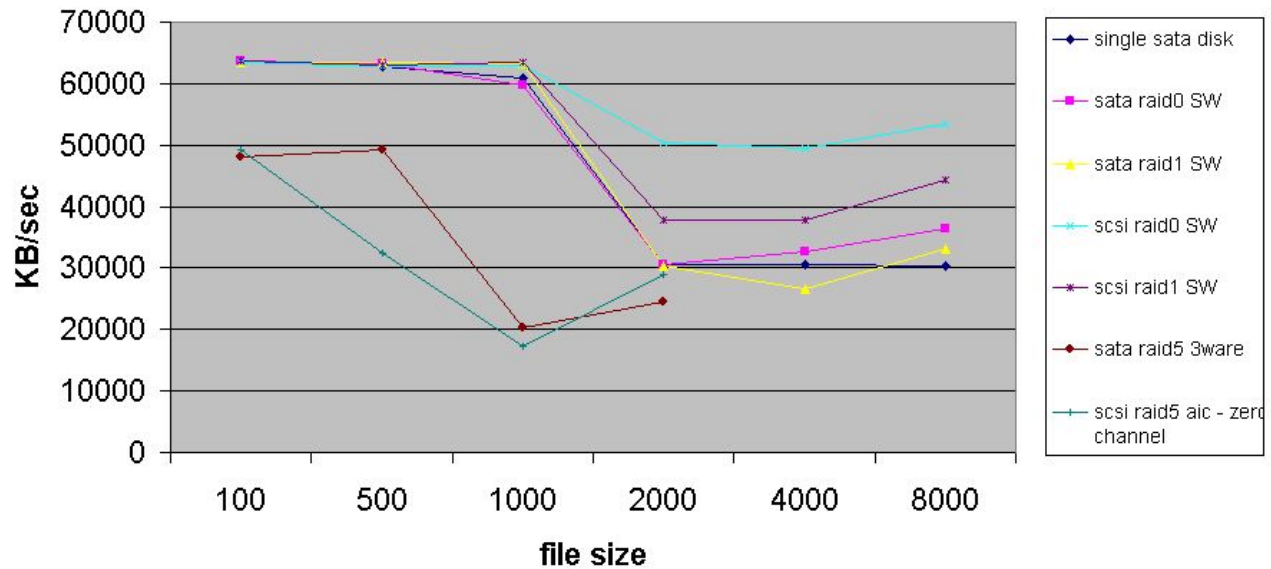


file create per block - write()



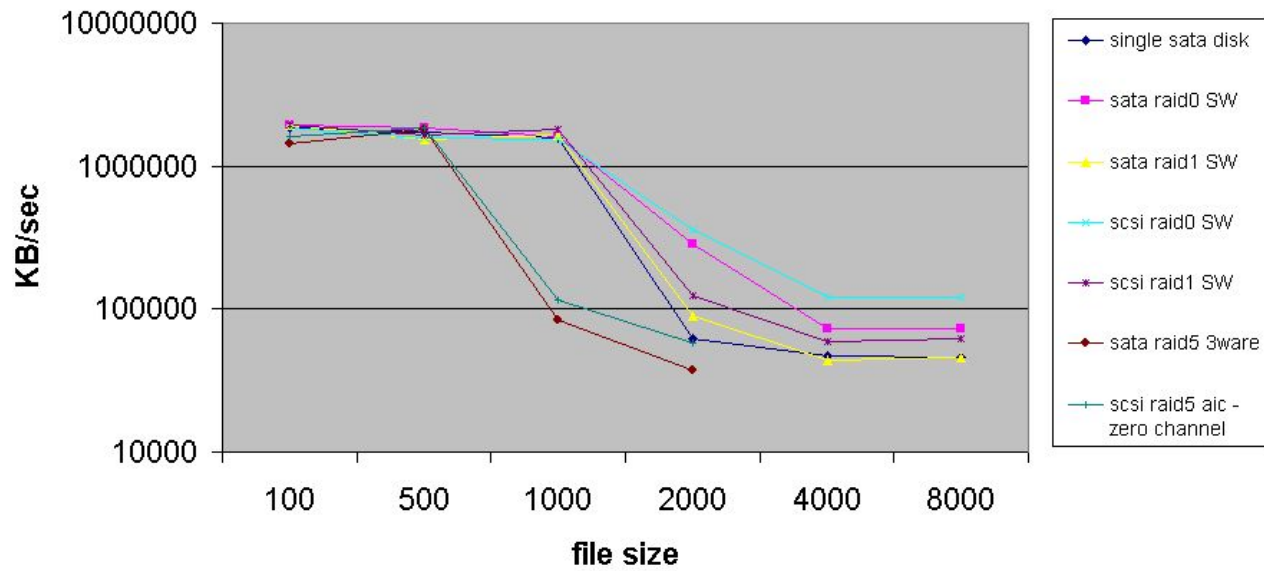
| file write per block- write() function | | | | | | | |
|--|---------------|-----------------------|----------------|------------------|------------------|------------------|------------------|
| file size MB | Kbyte/sec | | | | | | |
| | sata 3ware | scsi zero channel. | single sata | sata raid1 sw | sata raid0 sw | scsi raid1 sw | scsi raid0 sw |
| 100 | 177456,5 | 166733,9 | 316099,9 | 304573,5 | 328425,9 | 329074,4 | 315715,9 |
| 500 | 171698 | 114614,2 | 316864,6 | 289877,2 | 311237,8 | 307402 | 221147,7 |
| 1000 | 60332,9 | 56790,4 | 283112 | 273851,5 | 281923,1 | 296456,6 | 193743,1 |
| 2000 | 24060,4 | 15268,2 | 82055,6 | 65545,7 | 155954 | 107900,1 | 185235,7 |
| 4000 | | | 55143,1 | 38071,6 | 80524,5 | 76612,2 | 143186 |
| 8000 | | | 44391,4 | 29190,1 | 62912,7 | 59500 | 119652,4 |

**file read
character - getc()**



| file read per character- getc() function | | | | | | | |
|--|---------------|-----------------------|----------------|------------------|------------------|------------------|------------------|
| file size MB | Kbyte/sec | | | | | | |
| | sata 3ware | scsi zero channel. | single sata | sata raid1 sw | sata raid0 sw | scsi raid1 sw | scsi raid0 sw |
| 100 | 48014,8 | 49122,1 | 63736,2 | 63517,3 | 63584,7 | 63669,5 | 63465,4 |
| 500 | 49196,6 | 32397,6 | 62758,4 | 63398,9 | 63259,7 | 63086,8 | 62690,8 |
| 1000 | 20289,9 | 17182,5 | 60799 | 63280,4 | 59630,4 | 63548,6 | 62959,3 |
| 2000 | 24579,2 | 29018,4 | 30609,7 | 30229,9 | 30652,1 | 37744,4 | 50432,2 |
| 4000 | | | 30504,3 | 26705,3 | 32760,9 | 37774,8 | 49532,7 |
| 8000 | | | 30344 | 33125,3 | 36349,8 | 44329,8 | 53468,6 |

file read per block - read()



| file read per block - read() function | | | | | | | |
|---------------------------------------|---------------|-----------------------|----------------|------------------|------------------|------------------|------------------|
| file size MB | Kbyte/sec | | | | | | |
| | sata 3ware | scsi zero channel. | single sata | sata raid1 sw | sata raid0 sw | scsi raid1 sw | scsi raid0 sw |
| 100 | 1421903 | 1611204 | 1870445 | 1925690 | 1920189 | 1927977 | 1866481 |
| 500 | 1796381 | 1836299 | 1712440 | 1538000 | 1831582 | 1654025 | 1590681 |
| 1000 | 83475,2 | 116190,6 | 1577063 | 1661820 | 1599907 | 1829582 | 1543364 |
| 2000 | 37285,6 | 57423,1 | 61514,5 | 89353,4 | 282392,4 | 124120,5 | 356248,6 |
| 4000 | | | 47167,6 | 43764,7 | 72693,7 | 59524 | 121993,5 |
| 8000 | | | 45950,7 | 45842,1 | 72156,9 | 61608,6 | 121943,4 |