

Inhalt:

LINUX Filesysteme
im Vergleich

- ▷ REISERFS
- ▷ XFS
- ▷ JFS
- ▷ EXT3

Aufgabe:

Für den Einsatz eines IDE RAID Systems sollte ein aktuelles Filesystem evaluiert werden, das hinsichtlich eventuell notwendiger Filesystemchecks eine schnellstmögliche online Schaltung gewährleisten kann. Vorhandene Hardware sollte dabei wieder verwendet werden, was sich auf die Auswahl der unten genannten Komponenten auswirkte. Trotzdem sollten sich allgemeine Tendenzen ablesen lassen.

Jedes der links genannten Filesysteme bietet spezifische Vor- und Nachteile auf die hier nicht näher eingegangen wird. Massgeblich für den Kunden war ein schneller Überblick über die offensichtlichsten Performance Parameter. Die Präferenz lag dabei auf dem Einsatz von XFS, weil es bereits in anderen Bereichen eingesetzt wird. Die Frage der auf den Grund gegangen werden sollte, ob nicht eine der Alternativen ein besserer Baustein wäre.

Ziel: Einsatz des RAID als rsync Backupdevice

Testumgebung:

IDE RAID Subsystem (ARENA 6 Slots – 6x 80GB IDE HDDs als RAID5 ohne Hotspare) 40MB RAM – Firmware Version 2.1U
Nettokapazität: 370GB partitioniert in 4 Test Partitionen (sda[1234])

**Serversystem:**

Linux: Debian Woody 3.0 + Vanilla Kernel 2.4.25 von kernel.org
(xfs, jfs, ext3 fest im Kernel, reiserfs als loadable Modul)

PCI Adaptec 2940UW SCSI Kontroller

128MB RAM / P200MMX CPU

Testtool:

bonnie++ (apt-get install bonnie)

Aufruf:

bonnie -d /raid[1234] -u user -s 250M

Partitionen sind mit den defaults im jeweiligen Filesystem formatiert worden

/proc/pci

```
Bus 0, device 10, function 0:
  SCSI storage controller: Adaptec AHA-2940U/UW/D / AIC-7881U (rev 1).
  IRQ 12.
  Master Capable. Latency=64. Min Gnt=8.Max Lat=8.
  I/O at 0xe000 [0xe0ff].
  Non-prefetchable 32 bit memory at 0xe5800000 [0xe5800fff].
```

/proc/scsi/scsi

```
Attached devices:
Host: scsi0 Channel: 00 Id: 02 Lun: 00
  Vendor: Arena   Model: RAID           Rev:
  Type:   Direct-Access          ANSI SCSI revision: 02
```

/proc/scsi/aic7xxx/0

```
Target 2 Negotiation Settings
  User: 20.000MB/s transfers (10.000MHz, offset 127, 16bit)
  Goal: 20.000MB/s transfers (10.000MHz, offset 8, 16bit)
  Curr: 20.000MB/s transfers (10.000MHz, offset 8, 16bit)
Channel A Target 2 Lun 0 Settings
  Commands Queued 124041
  Commands Active 1
  Command Openings 31
  Max Tagged Openings 32
  Device Queue Frozen Count 0
```



Output

bonnie++ Ergebnisse (2 Durchläufe um fehlerhafte Werte zu erkennen)

ReiserFS

```
Version 1.02b      -----Sequential Output----- --Sequential Input- --Random-
-Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
Machine   Size K/sec %CP K/sec %CP K/sec %CP K/sec %CP K/sec %CP /sec %CP
black    250M  1460  86 10171  64 4242  18  1563  83  8216  27 316.6  12
-----Sequential Create----- -----Random Create-----
-Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
files   /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
      16  1706  96 +++++ +++  1651  91  1581  90 25493  31  1562  92
black,250M,1460,86,10171,64,4242,18,1563,83,8216,27,316.6,12,16,1706,96,++++,+++1651,91,1581,90,25493,31,1562,92
```

XFS

```
Version 1.02b      -----Sequential Output----- --Sequential Input- --Random-
-Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
Machine   Size K/sec %CP K/sec %CP K/sec %CP K/sec %CP K/sec %CP /sec %CP
black    250M  1735  95 9819  26 4386  18  1491  79  8697  27 281.6  11
-----Sequential Create----- -----Random Create-----
-Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
files   /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
      16   783  85 +++++ +++   761  73   793  86 +++++ +++   444  41
black,250M,1735,95,9819,26,4386,18,1491,79,8697,27,281.6,11,16,783,85,++++,+++761,73,793,86,++++,+++444,41
```

EXT3

```
Version 1.02b      -----Sequential Output----- --Sequential Input- --Random-
-Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
Machine   Size K/sec %CP K/sec %CP K/sec %CP K/sec %CP K/sec %CP /sec %CP
black    250M  1556  88 9568  42 4333  21  1481  78  9125  26 273.7  9
-----Sequential Create----- -----Random Create-----
-Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
files   /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
      16   159  94 +++++ +++   9780  98   146  82 +++++ +++   828  96
black,250M,1556,88,9568,42,4333,21,1481,78,9125,26,273.7,9,16,159,94,++++,+++9780,98,146,82,++++,+++828,96
```

JFS

```
Version 1.02b      -----Sequential Output----- --Sequential Input- --Random-
-Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
Machine   Size K/sec %CP K/sec %CP K/sec %CP K/sec %CP K/sec %CP /sec %CP
black    250M  1560  85 9474  30 4606  17  1587  84  9037  26 327.1  12
-----Sequential Create----- -----Random Create-----
-Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
files   /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
      16  3072  83 +++++ +++  2411  73   980  53 +++++ +++   827  36
black,250M,1560,85,9474,30,4606,17,1587,84,9037,26,327.1,12,16,3072,83,++++,+++2411,73,980,53,++++,+++827,36
```

Ergebnis:

Erkennbar besser schlägt sich XFS und ReiserFS beim Write-Durchsatz als die Konkurrenz. Beim Lese-Durchsatz sieht das Bild schon wieder anders aus. Wichtig dabei, dass Benchmarks auf keinen Fall die Komplexität eines echten Systems mit IO Belastungen darstellen können. Die Abstände sind auf jeden Fall so gering, dass die Wahl eines der untersuchten Filesystems keine grundlegenden Probleme aufwerfen sollte. Im konkret geplanten Einsatz von XFS, ist kein Bottle-Neck zu erwarten.

Weitergehende Interpretationen auf den eigenen Einsatzzweck bleiben dem geeigneten Leser selbst überlassen, die Werte können Anhaltspunkte für die Auswahl eines der Journaling Filesystems bieten. Probleme bei der Aktivierung im Vanilla Kernel sind bei keinem der genannten Filesysteme aufgetreten. In Debian Woody 3.0 sind für alle Filesysteme die notwendigen Userspace Tools per apt-get install verfügbar, so dass auch hier keine Begrenzung im Einsatz vorliegt. Wie das bei anderen Distributionen aussieht müsste konkret noch untersucht werden.

Hinsichtlich erzeugen und löschen von Dateien legt JFS + ReiserFS die Marke höher als Sie von EXT3 + XFS erreicht werden kann. Da dies für den konkreten Einsatz keine Auswirkungen haben sollte wird auf eine Untersuchung dieser Eigenschaft verzichtet, je nach Einsatzzweck, z.B. Indizierung von vielen kleinen Dateien mit anschließendem Löschen dieser (z.B. News Import) mag JFS oder ReiserFS hier Vorteile ausspielen können. Sollte vor dem Einsatz aber genau geprüft werden.

