

FreeBSD - ZFS - und 4k-Festplattenproblematik

Beitrag von „guckux“ vom 6. April 2021, 15:13

Guckux

Dieser Tage bin ich mal wieder über das Festplatten-4k Problem gestolpert und dachte mir, guggst Dir mal Deine Config an - mal einfach so...

Dabei stellte ich fest - ich habe die Festplatten "raw" einfach per zpool create erstellt, so daß die genutzte Kapazität bei sec63 anfängt - da scheinen die zfs-tools wohl noch etwas "Modernisierung" vor sich zu haben...

Ergo: Spiegel neu aufbauen damit die Blöcke richtig gesetzt werden - dabei habe ich mal durch die Reihe performance-Tests gemacht - war vorher im praktischen Betrieb nicht "unzufrieden" - aber die Werte enttäuschen mich etwas...

Hardware: i5-6500 (4x 3.2GHz), 32GB DDR4 Ram, SATA-6G Anbindung,

2x ST4000DM004-2CV104 0001 (4TB), System auf nvme SSD Samsung 950 Pro 512GB

FreeBSD 12-STABLE von Anfang März, bonnie++ (1.98) zum stressen und messen... (ich weiß, daß Teil ist und kann böse sein, bin aber der Meinung, es ist mit am ehrlichsten)

Aufgefallen ist mir beim testen:

Beim rewriting beobachtet mit systat :zarc, daß die einzelne disk etwa 50% mehr Durchsatz anzeigt als zpool iostat auf dem pool...

Beim lesen passt es - rund 2x Durchsatz vom systat:zarc zum zpool iostat

Etwas enttäuschen mich die Lesewerte beim bonnie++ beim "final-raid"

bonnie++ -r 32768 -m server -u root

gpart show ada3

=> 40 7814037088 ada3 GPT (3.6T)

40 2008 - free - (1.0M)

2048 7814033408 1 freebsd-zfs (3.6T)

7814035456 1672 - free - (836K)

gpart show ada5

=> 40 7814037088 ada5 GPT (3.6T)

40 2008 - free - (1.0M)

2048 7814033408 1 freebsd-zfs (3.6T)

7814035456 1672 - free - (836K)

"Vernünftige" Darstellung bedingt 80Zeichen Breite! 😊

Code

1. single disk, mit zfs
2. Version 1.98 -----Sequential Output----- --Sequential Input- --Random-
3. -Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
4. Name:Size etc /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
5. alpinaR1-single 64G 306k 98 105m 7 30.1m 2 727k 98 179m 6 145.7 1
6. Latency 28417us 38531us 1541ms 45365us 173ms 478ms
7. Version 1.98 -----Sequential Create----- -----Random Create-----
8. alpinaR1-single -Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
9. files /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
10. 16 ++++++ +++ ++++++ +++ 4154.577626 8 ++++++ +++ ++++++ +++
3948.968091 7
11. Latency 1598us 36us 967ms 1535us 24us 468ms
- 12.
- 13.
14. raid1 - "naggisch"
15. Version 1.98 -----Sequential Output----- --Sequential Input- --Random-
16. -Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
17. Name:Size etc /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
18. alpinaR1 64G 299k 97 76.8m 6 40.8m 3 702k 96 232m 8 220.6 2

```

19. Latency 28550us 38345us 861ms 68444us 175ms 325ms
20. Version 1.98 -----Sequential Create----- -----Random Create-----
21. alpinaR1 -Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
22. files /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
23. 16 ++++++ +++ ++++++ +++ 4145.272690 8 ++++++ +++ ++++++ +++
    3683.089828 7
24. Latency 1502us 30us 899ms 1463us 83us 441ms
25.
26.
27. raid1 mit cache (L2ARC - auf 29GB nvme)
28. [root@alpina /raid1]# bonnie++ -r 32768 -m alpinaR1_cache -u root
29. Version 1.98 -----Sequential Output----- --Sequential Input- --Random-
30. -Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
31. Name:Size etc /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
32. alpinaR1_cache 64G 230k 89 68.9m 5 39.8m 3 720k 99 235m 8 241.0 3
33. Latency 33220us 7303ms 623ms 29652us 269ms 341ms
34. Version 1.98 -----Sequential Create----- -----Random Create-----
35. alpinaR1_cache -Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
36. files /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
37. 16 ++++++ +++ ++++++ +++ 2690.164896 5 ++++++ +++ ++++++ +++
    4142.792682 8
38. Latency 1511us 37us 2925ms 1473us 25us 558ms
39.
40.
41. raid1 mit cache und log (L2ARC - 29GB, logs 1GB beides auf nvme)
42. [root@alpina /raid1]# bonnie++ -r 32768 -m alpinaR1_cache_log -u root
43. Version 1.98 -----Sequential Output----- --Sequential Input- --Random-
44. -Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
45. Name:Size etc /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
46. alpinaR1_cache_ 64G 258k 99 78.9m 6 46.4m 4 700k 96 229m 8 229.2 2
47. Latency 33061us 7278ms 567ms 103ms 196ms 391ms
48. Version 1.98 -----Sequential Create----- -----Random Create-----
49. alpinaR1_cache_log -Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
50. files /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
51. 16 ++++++ +++ ++++++ +++ 3669.165249 7 ++++++ +++ ++++++ +++
    3656.157708 7
52. Latency 1560us 36us 1601ms 1470us 26us 580ms
53.
54.
55. alte config
56. 2x4GB false sec63
57. raid1 mit cache und log (L2ARC - 29GB, logs 1GB beides auf nvme)
58. Version 1.98 -----Sequential Output----- --Sequential Input- --Random-

```

```

59. -Per Chr- --Block-- -Rewrite- -Per Chr- --Block-- --Seeks--
60. Name:Size etc /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
61. alpinaR1 64G 287k 99 10.4m 1 8373k 1 720k 99 2.7g 99 200.0 4
62. Latency 29770us 437s 430s 14303us 391us 859ms
63. Version 1.98 -----Sequential Create----- -----Random Create-----
64. alpinaR1 -Create-- --Read--- -Delete-- -Create-- --Read--- -Delete--
65. files /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP /sec %CP
66. 16 ++++++ +++ ++++++ +++ 165.746758 0 ++++++ +++ ++++++ +++ 566.048702 1
67. Latency 28743us 41us 54405ms 1516us 20us 5952ms

```

Alles anzeigen

Hier sieht man deutlich schlechtere Werte als im vorstehenden Block - seq. output Block 10MB/s zu 78MB/s, rewrite 8k3 zu 46m, seq Input Block 2.7g zu 229m????? dafür mit einem Bruchteil der CPU-Last...