

Zum Selbststudium: So arbeitet die Datenkompression bei einer Data Domain

Nur keine Angst - ich möchte Sie nicht unbedingt die Datenkompressionsverfahren erläutern, die sich innerhalb einer Data Domain (genauer: im Betriebssystem DDOS) abspielen. Ich bin vielmehr davon überzeugt, daß sich ein Kunde eher von dem Ergebnis der Dateneduplizierung beeindrucken läßt, als daß er den Algorithmus detailliert nachvollziehen möchte.

Deshalb möchte ich Sie hier nur auf zwei Dokumente verweisen, die Sie in der Dell/EMC Knowledgebase finden und die sich detailliert darüber auslassen. Es handelt sich um folgende Dokumente:

KB 000023290 How to read Data Domain statistics for a NetWorker save set

KB 000003886 Data Domain Understanding DataDomain Compression

Auf den folgenden Seiten habe ich Ihnen dann noch ein paar Beispiele aus meiner kleinen NetWorker Installation aufgeführt, damit Sie einen Eindruck über die Zahlen für Ihr eventuelles Selbststudium bekommen - weitere Details hierzu werde ich Ihnen jedoch nicht vermitteln.

Ausgangspunkt sind diese 4 NetWorker Save Sets auf ein DDBoost Volume:

```
D:\>mminfo -avot
volume      type      client      date      time      size      ssid      fl      lvl      name
DDBoost.001 Data Domain 19-nwtest.eval.local 4/15/2021 12:28:14 4882 MB 4286059840 cb full D:\5GB
DDBoost.001 Data Domain 19-nwtest.eval.local 4/15/2021 14:29:27 4931 MB 4269289897 cb full D:\5GB
DDBoost.001 Data Domain 19-nwtest.eval.local 4/15/2021 14:43:59 4882 MB 4252513553 cb full D:\5GB
DDBoost.001 Data Domain 19-nwtest.eval.local 4/15/2021 14:57:29 4931 MB 4235737147 cb full D:\5GB

D:\>
```

1. 5GB Textfile - full

```
D:\>mminfo -q ssid=4286059840 -S
ssid=4286059840 savetime=4/15/2021 12:28:14 (1618482494) 19-nwtest.eval.local:D:\5GB
level=full sflags=vF size=5000154140 files=4 insert=4/15/2021
create=4/15/2021 complete=4/15/2021 browse=5/15/2021 23:59:59 retent=5/15/2021 23:59:59
clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
*backup start time: 1618482458;
*NSR_VSS_SHADOWCOPY_SET: {958F3D5B-A9B0-41C9-862E-87F7B2FBB9FE};
*policy action jobid: 32002;
*policy action name: "backup: 1618482494";
*policy name: "TEST: 1618482494";
*policy workflow name: "Backup: 1618482494";
*ss data domain backup cloneid: 1618482494;
*ss data domain dedup statistics: "v1:1618482494:5014205776:5014205776:272721499";
group: TEST;
saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618482494 time=4/15/2021 12:28:14 retent=5/15/2021 flags=
frag@ 0 valid= 7869421 file/rec= 0/0 rn=0 last=4/15/2021

D:\>
```

2. 5GB Textfile - full + 5x10 10 Mio Bytes Text

```
D:\>mminfo -q ssid=4269289897 -S
ssid=4269289897 savetime=4/15/2021 14:29:27 (1618489767) 19-nwtest.eval.local:D:\5GB
level=full sflags=vF size=5050158248 files=10 insert=4/15/2021
create=4/15/2021 complete=4/15/2021 browse=5/15/2021 23:59:59 retent=5/15/2021 23:59:59
clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
*backup start time: 1618489747;
*NSR_VSS_SHADOWCOPY_SET: {8985380F-B913-41CD-8F9A-B9DF03438471};
*policy action jobid: 32007;
*policy action name: "backup: 1618489769";
*policy name: "TEST: 1618489769";
*policy workflow name: "Backup: 1618489769";
*ss data domain backup cloneid: 1618489769;
*ss data domain dedup statistics: "v1:1618489769:5064350720:14896648:14232549";
group: TEST;
saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618489769 time=4/15/2021 14:29:29 retent=5/15/2021 flags=
frag@ 0 valid= 7869421 file/rec= 0/0 rn=0 last=4/15/2021
```

3. 5GB Random File - full

```
D:\>mminfo -q ssid=4252513553 -S
ssid=4252513553 savetime=4/15/2021 14:43:59 (1618490639) 19-nwtest.eval.local:D:\5GB
level=full sflags=vF size=5000039912 files=4 insert=4/15/2021
create=4/15/2021 complete=4/15/2021 browse=5/15/2021 23:59:59 retent=5/15/2021 23:59:59
clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
*backup start time: 1618490621;
*NSR_VSS_SHADOWCOPY_SET: {ACE48AFC-3222-478D-B67D-11A7D3443034};
*policy action jobid: 32012;
*policy action name: "backup: 1618490641";
*policy name: "TEST: 1618490641";
*policy workflow name: "Backup: 1618490641";
*ss data domain backup cloneid: 1618490641;
*ss data domain dedup statistics: \
"v1:1618490641:5016782124:5016782124:5016782124";
group: TEST;
saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618490641 time=4/15/2021 14:44:01 retent=5/15/2021 flags=
frag@ 0 valid= 7869421 file/rec= 0/0 rn=0 last=4/15/2021
```

D:\>

4. 5GB Random File - full + 5x10 10 Mio Bytes Text

```
D:\>mminfo -q ssid=4235737147 -S
ssid=4235737147 savetime=4/15/2021 14:57:29 (1618491449) 19-nwtest.eval.local:D:\5GB
  level=full  sflags=vF      size=5050158176  files=10      insert=4/15/2021
  create=4/15/2021 complete=4/15/2021 browse=5/15/2021 23:59:59 retent=5/15/2021 23:59:59
  clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
    *backup start time: 1618491432;
    *NSR_VSS_SHADOWCOPY_SET: {6B31DA77-5EF7-4FA9-BA33-FE40765675D1};
    *policy action jobid: 32017;
    *policy action name: "backup: 1618491451";
    *policy name: "TEST: 1618491451";
    *policy workflow name: "Backup: 1618491451";
*ss data domain backup cloneid: 1618491451;
*ss data domain dedup statistics: "v1:1618491451:5067041980:232302506:232302506";
  group: TEST;
  saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618491451 time=4/15/2021 14:57:31 retent=5/15/2021 flags=
  frag@      0 volid= 7869421 file/rec=      0/0      rn=0 last=4/15/2021
```

D:\>

```
D:\>mminfo -q "savetime=>today" -r "nsavetime,ssid,ssid(64)" -ot
  save time ssid      ssid
1618482494 4286059840 73a59fbe-00000006-ff781540-60781540-00020c00-35df0829
1618489767 4269289897 516593c0-00000006-fe7831a9-607831a9-00030c00-35df0829
1618490639 4252513553 8b5b4956-00000006-fd783511-60783511-00040c00-35df0829
1618491449 4235737147 7d9b6b0b-00000006-fc78383b-6078383b-00050c00-35df0829
```

D:\>

Herausfinden der Verzeichnisse des MTrees

```
sysadmin@ddve74-1#
sysadmin@ddve74-1# ddboost storage-unit show /data/coll/ssu
Storage-unit /data/coll/ssu not found
```

Storage-unit /data/coll/ssu does not exist - Failed to get file information.

```
sysadmin@ddve74-1# system show serialno
Serial number: AUDVZF9H9RS9CY
System software-id: Not available
Instance software-id: Not available
sysadmin@ddve74-1#
sysadmin@ddve74-1# priv set se
Enter system password: serial_number
```

```
SE@ddve74-1## se sfs_dump -c /data/coll/ssu
name      mtime  fileid size  seg_bytes  seg_count  redun_seg_count pre_lc_size  post_lc_size
/data/coll/ssu/Test/.nsr_serial 3038552547000000000 18 328 1396 3 0 1396 1318
/data/coll/ssu/Test/nsr.dir 3038552558000000000 19 52 1120 3 0 1120 1121
/data/coll/ssu/Test/volhdr 3038552558000000000 20 332 1400 3 0 1400 1203
/data/coll/ssu/Test/11/07/notes/8b5b4956-00000006-fd783511-60783511-00040c00-35df0829 1618491176125150676 139 836
2928 4 0 29282611
/data/coll/ssu/Test/11/07/8b5b4956-00000006-fd783511-60783511-00040c00-35df0829 1618491175352942995 134 5000039912
5016782124 597182 0 5016782124 5095875295
/data/coll/ssu/Test/28/38/notes/516593c0-00000006-fe7831a9-607831a9-00030c00-35df0829 1618489840417802323 133 832
2924 4 0 29242628
/data/coll/ssu/Test/28/38/516593c0-00000006-fe7831a9-607831a9-00030c00-35df0829 1618489839834277754 128 5050158248
5064350720 506113 504825 14896648 14232549
/data/coll/ssu/Test/73/82/notes/7d9b6b0b-00000006-fc78383b-6078383b-00050c00-35df0829 1618491607884196557 145 832
2924 4 0 29242626
/data/coll/ssu/Test/73/82/7d9b6b0b-00000006-fc78383b-6078383b-00050c00-35df0829 1618491605947406220 140 5050158176
5067041980 602232 576623 232302506 235192565
/data/coll/ssu/Test/76/06/notes/73a59fbe-00000006-ff781540-60781540-00020c00-35df0829 1618482753020433005 127 836
2928 4 0 29282626
/data/coll/ssu/Test/76/06/73a59fbe-00000006-ff781540-60781540-00020c00-35df0829 1618482751854206930 122 5000154140
5014205776 501090 0 5014205776 272721499

files bytes seg_bytes segs redundant pre_lc post_lc bytes/pre_lc bytes/post_lc
11 20100514524 20162396220 2206642 1081448 10278202674 5618036041 2.0 3.6
SE@ddve74-1##
```

Der Abgleich mit dem NetWorker Medien Index

```
D:\>mminfo -q "savetime>=today" -r "nsavetime,ssid,ssid(64),sumsize(10)" -ot
  save time ssid      ssid
1618482494 4286059840 73a59fbe-00000006-ff781540-60781540-00020c00-35df0829 5000154140
1618489767 4269289897 516593c0-00000006-fe7831a9-607831a9-00030c00-35df0829 5050158248
1618490639 4252513553 8b5b4956-00000006-fd783511-60783511-00040c00-35df0829 5000039912
1618491449 4235737147 7d9b6b0b-00000006-fc78383b-6078383b-00050c00-35df0829 5050158176
```

D:\>

Jetzt können Sie die entsprechenden Größen vergleichen und analysieren

Und hier noch 2 weitere Save Sets für ein AFTD Volume:

```
D:\>mminfo -q "savetime>=-2days" -r "volume,nsavetime,ssid,ssid(64),sumsize(10)" -ot
volume          save time  ssid          ssid          size
DDBoost.001     1618482494 4286059840 73a59fbe-00000006-ff781540-60781540-00020c00-35df0829 5000154140
DDBoost.001     1618489767 4269289897 516593c0-00000006-fe7831a9-607831a9-00030c00-35df0829 5050158248
DDBoost.001     1618490639 4252513553 8b5b4956-00000006-fd783511-60783511-00040c00-35df0829 5000039912
DDBoost.001     1618491449 4235737147 7d9b6b0b-00000006-fc78383b-6078383b-00050c00-35df0829 5050158176
AFTD_1.001      1618608361 4202299626 90d266c7-00000006-fa7a00ea-607a00ea-00070c00-35df0829 5000154200
AFTD_1.001      1618608784 4185522833 13baf293-00000006-f97a0291-607a0291-00080c00-35df0829 5050158216
```

D:\>

```
SE@ddve74-1## se sfs_dump -c /data/coll/aftd
name mtime fileid size seg_bytes seg_count redun_seg_count pre_lc_size post_lc_size
/data/coll/aftd/aftd_1/.nsr_serial 1618591974719630000 18 340 1408 3 0 1408 1328
/data/coll/aftd/aftd_1/nsr.dir 1618592034011519000 19 52 1120 3 1 1068 1068
/data/coll/aftd/aftd_1/volhdr 1618592034109024000 20 332 1400 3 0 1400 1198
/data/coll/aftd/aftd_1/12/53/notes/90d266c7-00000006-fa7a00ea-607a00ea-00070c00-35df0829 1618608445834290000 127
672 1740 3 0 1740 1505
/data/coll/aftd/aftd_1/12/53/90d266c7-00000006-fa7a00ea-607a00ea-00070c00-35df0829 1618608445426571000 122
5000154200 5017392892 597303 595936 17802751 17810971
/data/coll/aftd/aftd_1/37/22/notes/13baf293-00000006-f97a0291-607a0291-00080c00-35df0829 1618608869255307000 133
672 1740 3 0 1740 1504
/data/coll/aftd/aftd_1/37/22/13baf293-00000006-f97a0291-607a0291-00080c00-35df0829 1618608868906188000 128
5050158216 5067633720 602338 600933 18218296 18205859

files bytes seg_bytes segs redundant pre_lc post_lc bytes/pre_lc bytes/post_lc
7 10050314484 10085034020 1199656 1196870 36028403 36023433 279.0 279.0
SE@ddve74-1##
SE@ddve74-1##
```



```
D:\>mminfo -q "ssid=4202299626" -S
ssid=4202299626 savetime=4/16/2021 23:26:01 (1618608361) 19-nwtest.eval.local:D:\5GB
level=full sflags=vF size=5000154200 files=4 insert=4/16/2021
create=4/16/2021 complete=4/16/2021 browse=5/16/2021 23:59:59 retent=5/16/2021 23:59:59
clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
    *backup start time: 1618608343;
*NSR_VSS_SHADOWCOPY_SET: {F5F4DC86-7F90-4CF2-B54B-5DC05935C07E};
    *policy action jobid: 32065;
    *policy action name: "backup: 1618608362";
    *policy name: "TEST: 1618608362";
    *policy workflow name: "Backup: 1618608362";
        group: TEST;
        saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618608362 time=4/16/2021 23:26:02 retent=5/16/2021 flags=
frag@ 0 valid=4219060511 file/rec= 0/0 rn=0 last=4/16/2021
```

```
D:\>
D:\>
D:\>mminfo -q "ssid=4185522833" -S
ssid=4185522833 savetime=4/16/2021 23:33:04 (1618608784) 19-nwtest.eval.local:D:\5GB
level=full sflags=vF size=5050158216 files=10 insert=4/16/2021
create=4/16/2021 complete=4/16/2021 browse=5/16/2021 23:59:59 retent=5/16/2021 23:59:59
clientid=d7d6416b-00000004-60610392-60610391-00010c00-35df0829
    *backup start time: 1618608767;
*NSR_VSS_SHADOWCOPY_SET: {82AC53DF-2DDC-4363-9E23-75BB0BA0F980};
    *policy action jobid: 32070;
    *policy action name: "backup: 1618608785";
    *policy name: "TEST: 1618608785";
    *policy workflow name: "Backup: 1618608785";
        group: TEST;
        saveset features: CLIENT_SAVETIME;
Clone #1: cloneid=1618608785 time=4/16/2021 23:33:05 retent=5/16/2021 flags=
frag@ 0 valid=4219060511 file/rec= 0/0 rn=0 last=4/16/2021
```

```
D:\>
```