
DE-DUPLIKACJA UWOLNIONA!
BoostFS
De-duplikacja na źródle dla każdego
Film omawiający BoostFS:
https://youtu.be/dSU_81xDoZ0
DELL EMC

Agenda

- What is BoostFS?
- Let’s touch it
- Why source de-duplication
- Comparison of source de-duplication technics
- Announcements


BoostFS
What is it?
DELL EMC

BoostFS
What is it?

- BoostFS is source de-duplication that any
  - Backup software
  - Application
  - Database
  - Script
  - Command

  can use

BoostFS
What is it?

- Technically BoostFS is directory (folder) on server
- Any backup software / application / database / script / command writes to this directory (folder)
- Data are de-duplicated on the server and only unique chunks are sent to Data Domain

BoostFS
What is it?

- BoostFS uses the same BOOST technology that is used in any other application (backup software, databases, etc.)
- We will have
  - the same efficiency
  - no load of source machine (Intel)

  as we have in backup application / databases using BOOST

BoostFS
Licensing

- BoostFS is 100% free feature of Data Domain 6.0+
- Everyone who has Data Domain 6.0 or higher and
  - BOOST license on Data Domain box

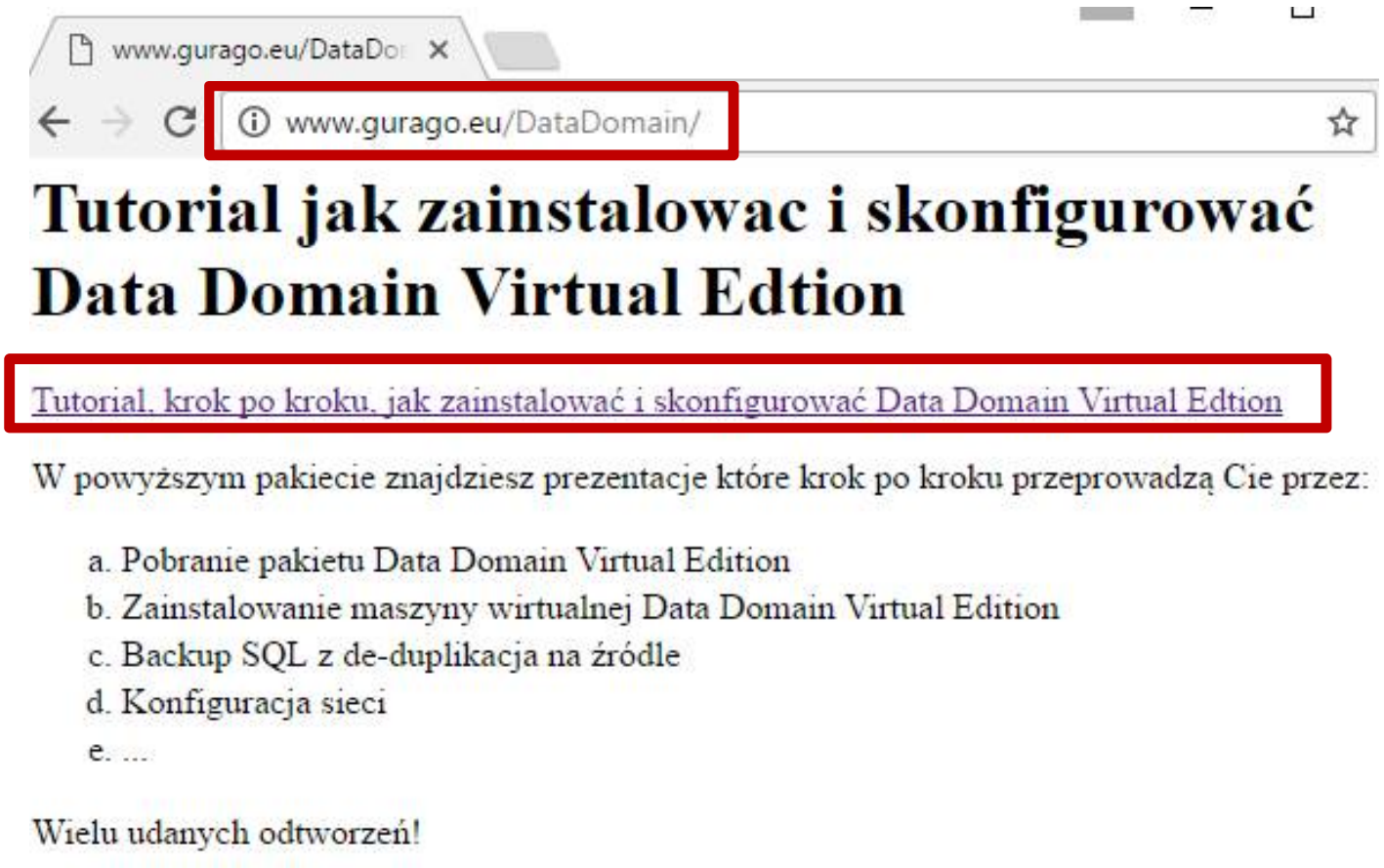
  or
  - Data Domain Virtual Edition (DDVE has BOOST license included within cost of DDVE)

  can use BoostFS without any additional fee

BoostFS
I want to try – how?

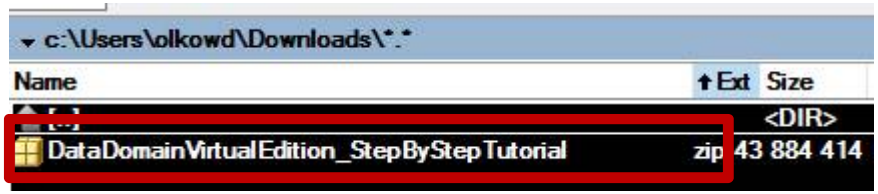
http://gurago.pl/

BoostFS
I want to try – how?



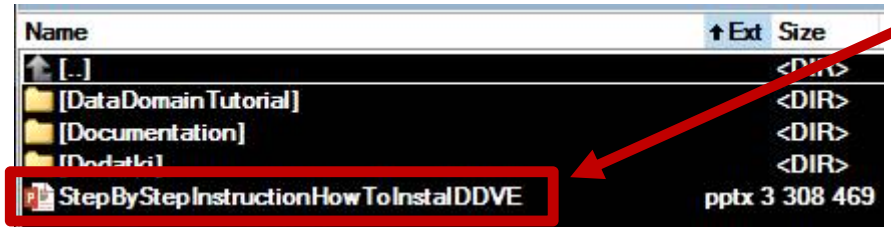
BoostFS
I want to try – how?

http://gurago.pl/



BoostFS
Chcę spróbować – jak?

http://gurago.pl/

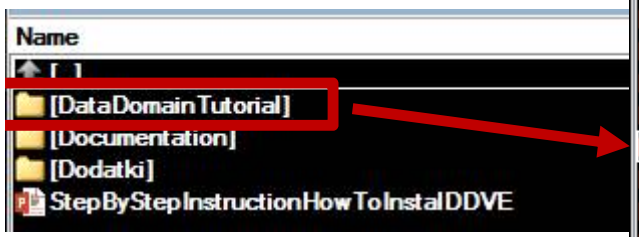


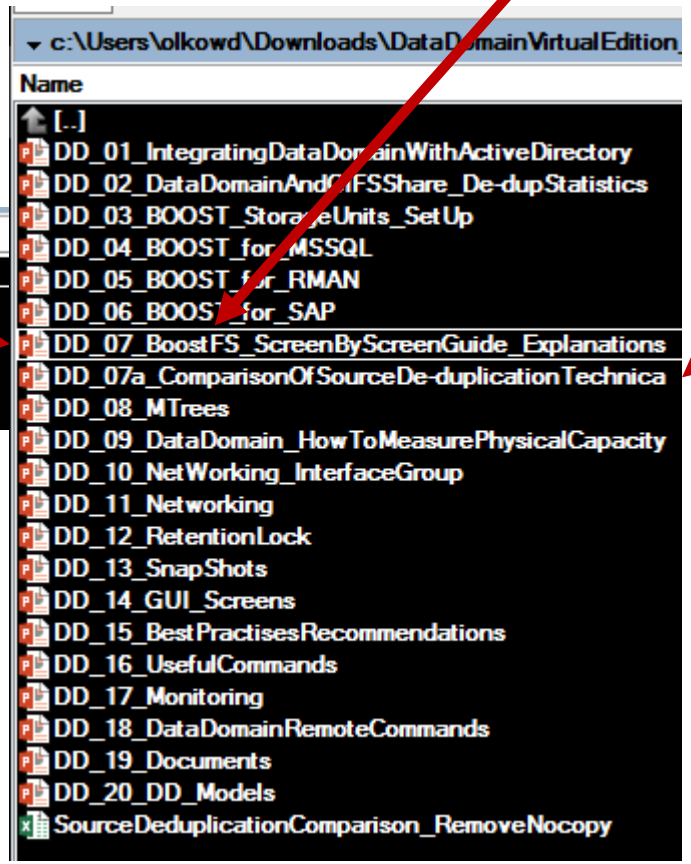
Tutorial how to
- Download
- Install
- Configure

Data Domain Virtual Edition (DDVE)

BoostFS
I want to try – how?

http://gurago.pl/





Step by step guide for BoostFS

BoostFS value presentation

Go for other ppt tutorials that show other super use cases of Data Domain!


BoostFS
Requirements
DELL EMC

BoostFS
Tested Applications

Backup Applications

- Commvault Simpana V11, V10, V9

NoSQL Databases

Mongo DB v2.6, 3.0., 3.2

- Backup Tools
- Mongodump
- OpsMgr v2.4

MySQL v5.7, v5.6

- Backup Tools
- MySQL Enterprise Backup v5.6, v57
- Mydumper
- Percona XtraBackup

BoostFS
Tested Applications

Backup Applications

For any other application contact EMC for formal approval (RPQ) in case of production use.

- MySQL Enterprise Backup v5.6, v57
- Mydumper
- Percona XtraBackup

BoostFS
Supported Operating Systems

The following Linux distributions are supported:

- Red Hat Enterprise Linux versions 6 and 7
- SUSe Linux Enterprise Server versions 11 and 12
- Ubuntu 14.04 and 15

BoostFS
Supported Operating Systems

The following Linux distributions are supported:

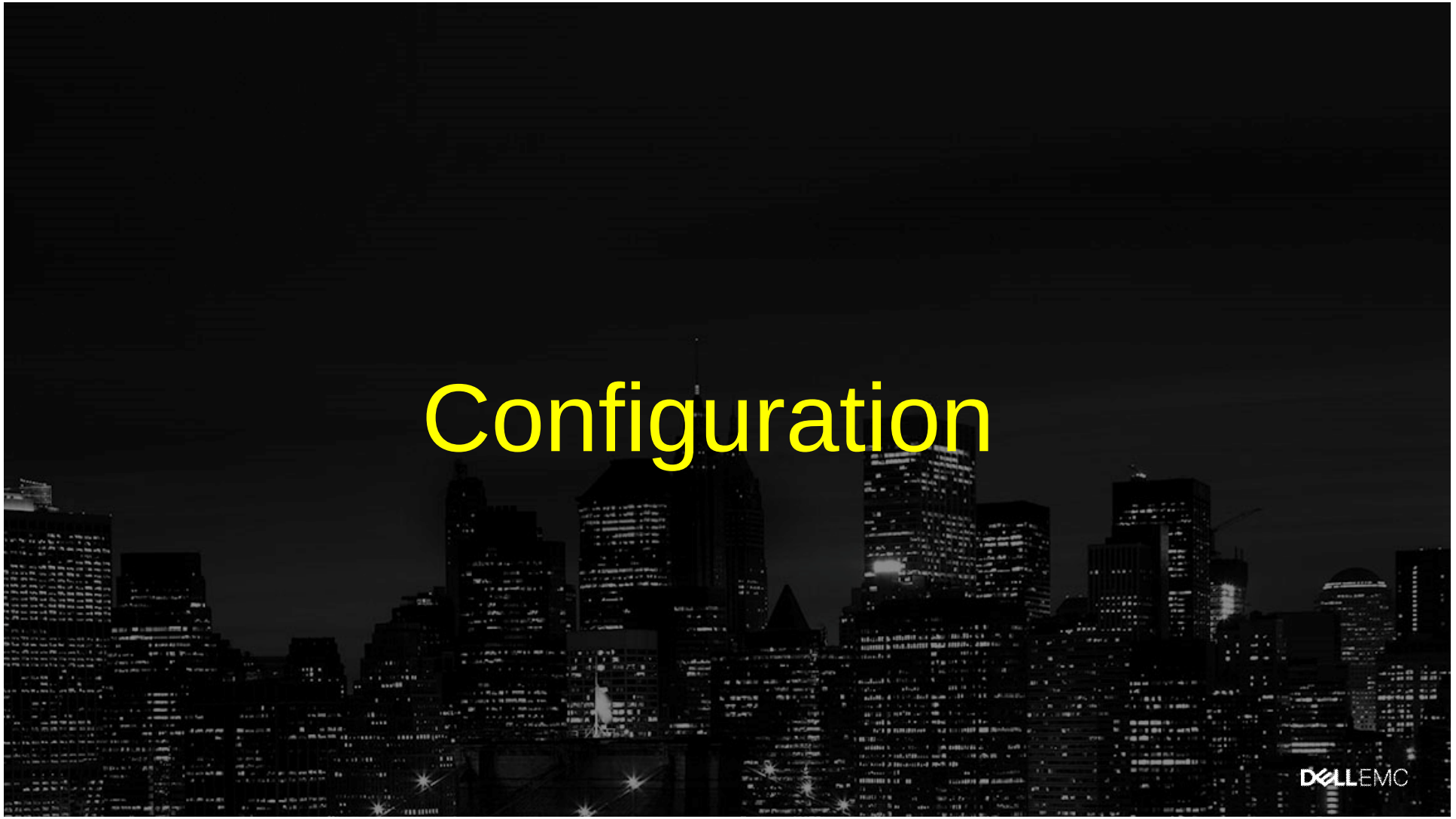
- Red Hat Enterprise Linux versions 6 and 7
- SUSe Linux Enterprise Server versions 11 and 12
- Ubuntu 14.04 and 15

This guide shows the of installation of BoostFS on SuSE 11

BoostFS
Data Domain requirements

- BoostFS requires Data Domain Operating System (DDOS) 6.0 or higher
- If you do not have Data Domain, install Data Domain Virtual Edition (DDVE).
- Data Domain Virtual Edition (DDVE) **is free** for tests/trials with perpetual **0.5TB license**


BoostFS
Let’s touch
DELL EMC


Configuration
DELL EMC

BoostFS Installation
Checking if we see binaries from Linux server

pwd

ls -al

```
root@szczeliniec:/daniel/#: pwd
/daniel
root@szczeliniec:/daniel/#: ls -al
total 2400
drwxr-xr-x  2 root root    4096 Nov 24 13:15 .
drwxr-xr-x 28 root root    4096 Nov 24 13:08 ..
-rw-r--r--  1 root root 2441623 Nov 24 12:47 DDBoostFS-1.0.0.4-546961.sles.x86_64.rpm
```

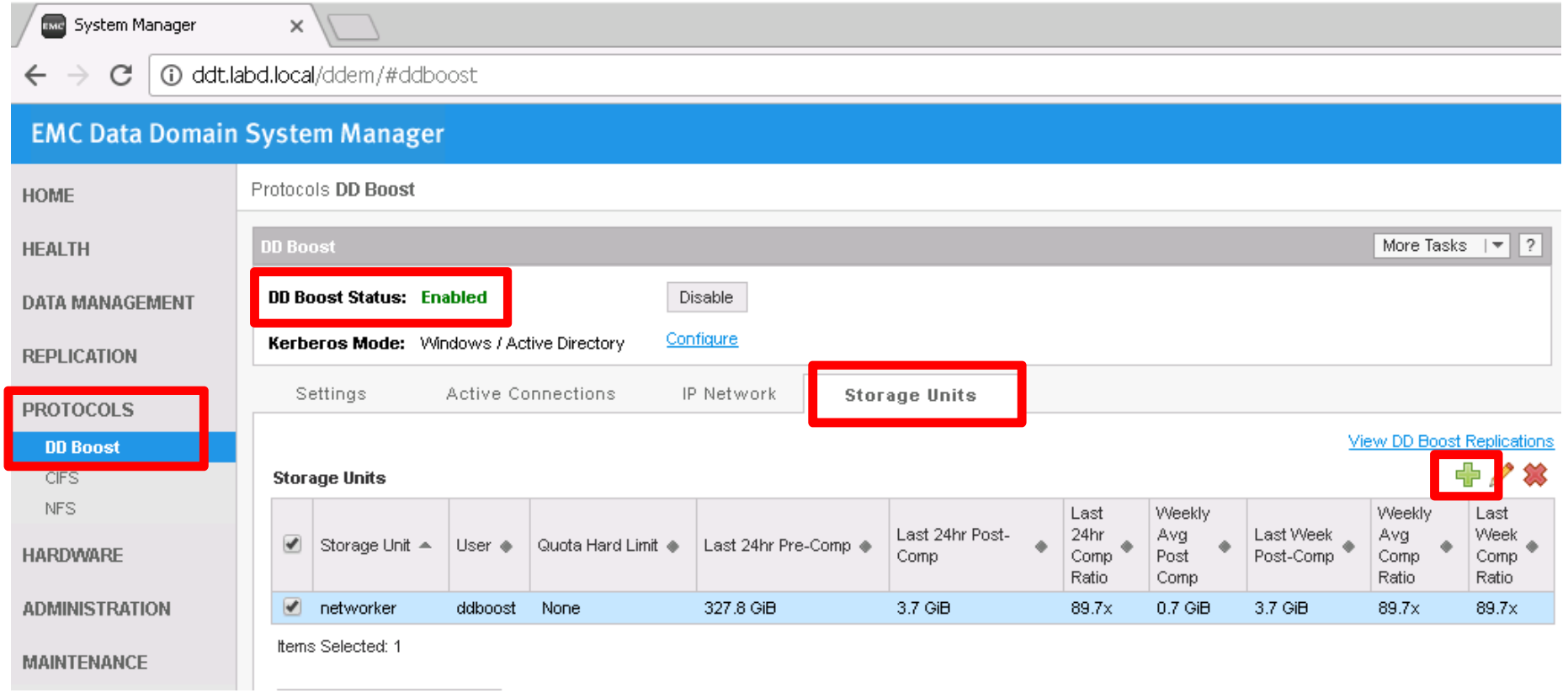
BoostFS Installation

Installing BoostFS on our Linux server

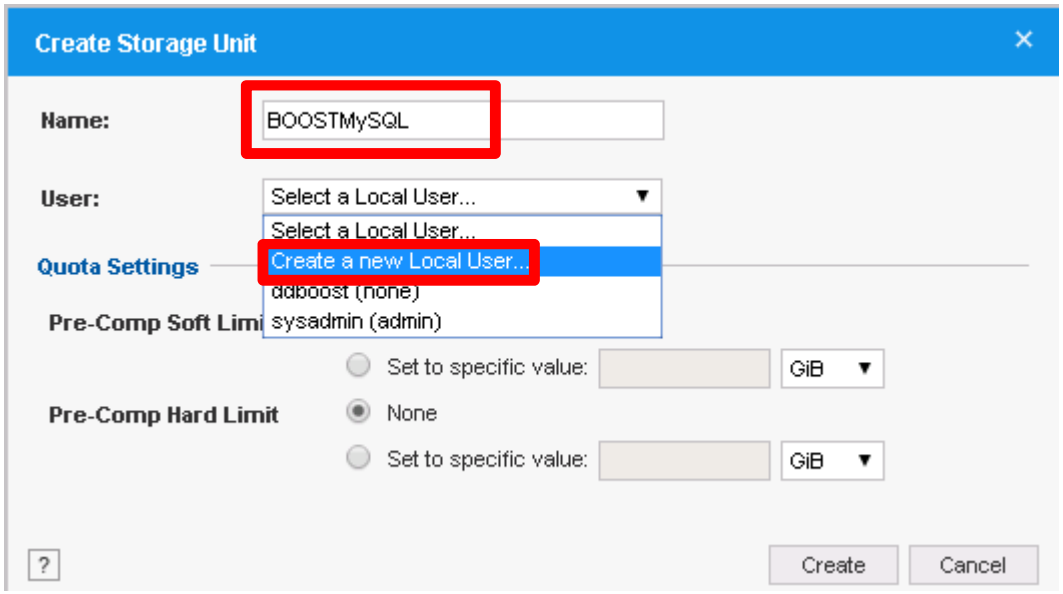
rpm -Uvh DDBoostFS-1.0.0.4-546961.sles.x86_64.rpm

```
-rw-r--r--  1 root root 2441623 Nov 24 12:47 DDBoostFS-1.0.0.4-546961.sles.x86_64.rpm
root@szczeliniec:/daniel/#:
root@szczeliniec:/daniel/#: rpm -Uvh DDBoostFS-1.0.0.4-546961.sles.x86_64.rpm
Preparing...                ########################################### [100%]
   1:ddboostfs              ########################################### [100%]
root@szczeliniec:/daniel/#:
```

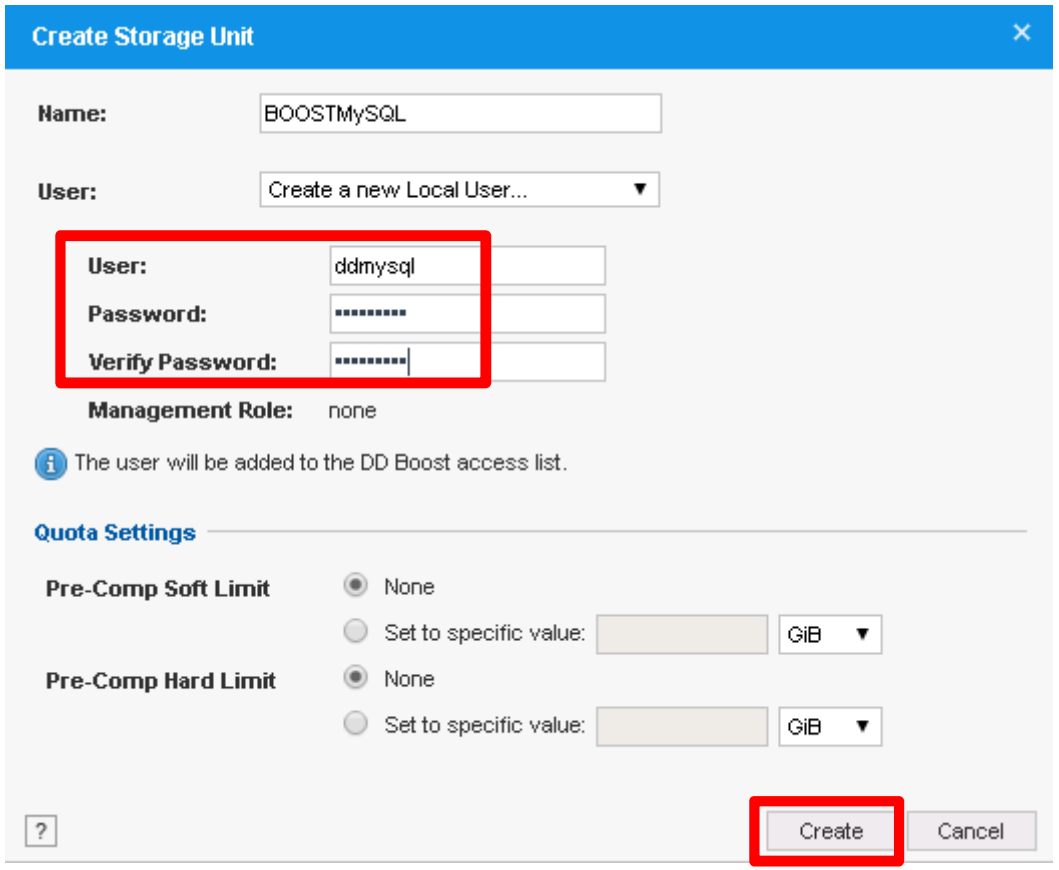
Creating Storage Unit (source de-dup device)



Creating Storage Unit (source de-dup device)

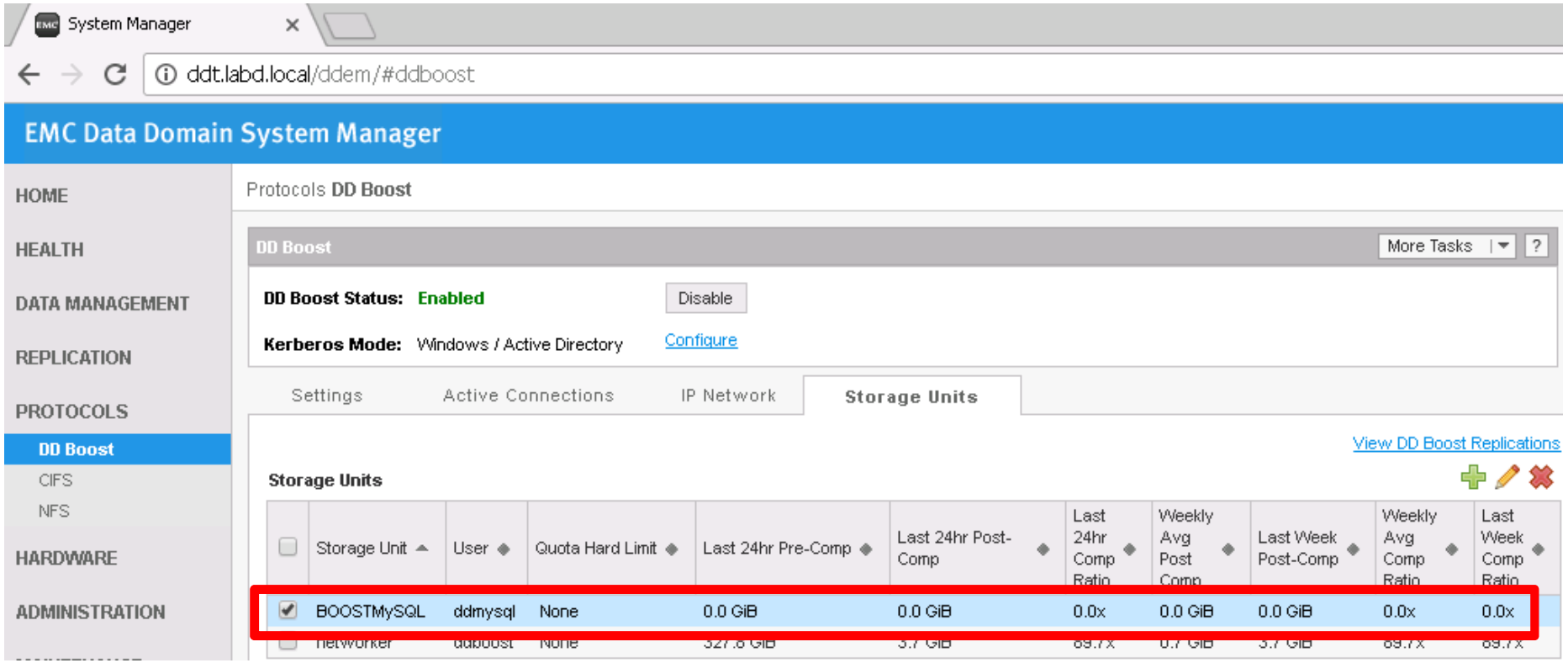


Creating Storage Unit (source de-dup device)



DELL EMC

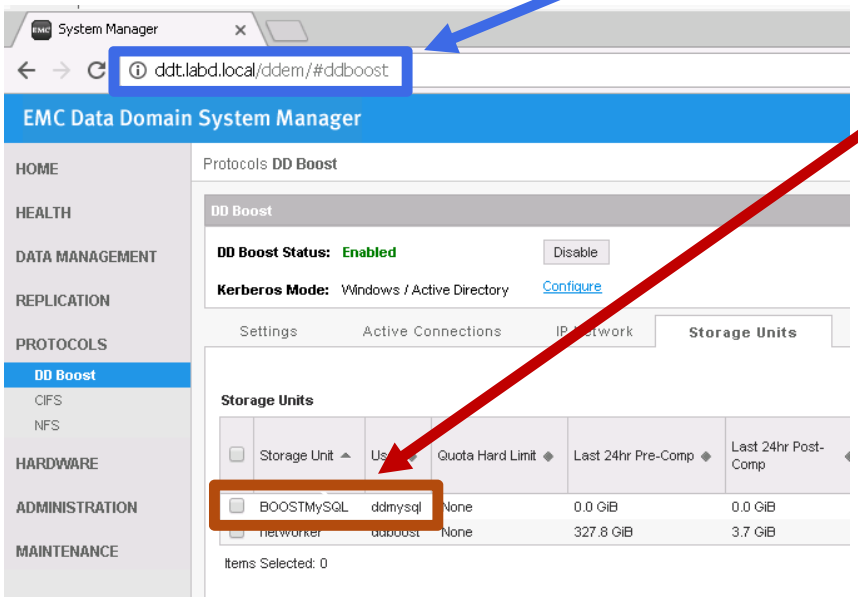
Creating Storage Unit (source de-dup device)

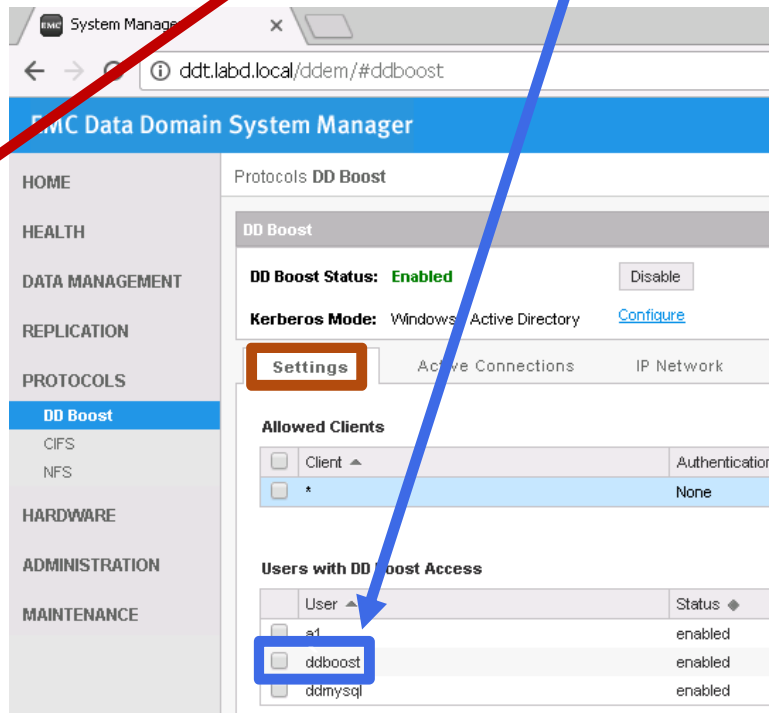


BoostFS: Defining Data Domain target
boostfs lockbox set (command)

FQDN of Data Domain *Storage Unit* *Boost user*

/opt/emc/boostfs/bin/boostfs lockbox set -d ddt.labd.local -s BOOSTMySQL -u ddmysql



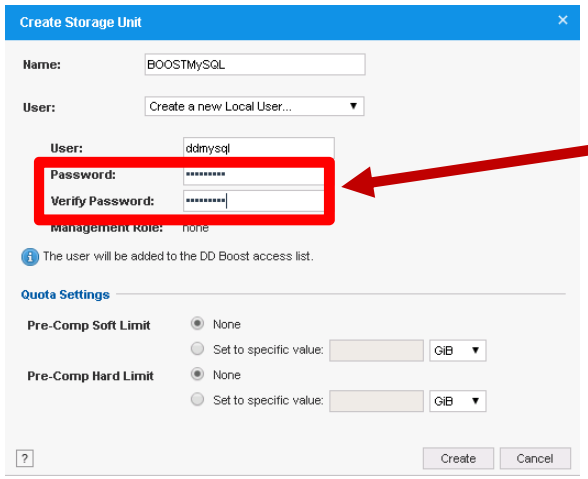


BoostFS: Defining Data Domain target
boostfs lockbox set (command)

FQDN of Data Domain *Storage Unit* *Boost user*

/opt/emc/boostfs/bin/boostfs lockbox set -d ddt.labd.local -s BOOSTMySQL -u ddmysql

```
root@szczeliniec:/daniel/#: /opt/emc/boostfs/bin/boostfs lockbox set -d ddt.labd.local -s BOOSTMySQL -u ddmysql
Enter storage unit user password:
Enter storage unit user password again to confirm:
Lockbox entry set
root@szczeliniec:/daniel/#:
```



Provide password for Boost user

DELL EMC

Directory with source de-dup to Data Domain

Creating directory to which we write with source de-dup on DD

mkdir /backup

ls -al

```
root@szczeliniec://#: mkdir /backup
root@szczeliniec://#: ls -al /backup
total 8
drwxr-xr-x  2 root root 4096 Nov 24 21:43 .
drwxr-xr-x 29 root root 4096 Nov 24 21:43 ..
root@szczeliniec://#:
```

BoostFS: Mounting Data Domain to dir

Attaching Data Domain to our source de-dup directory

/opt/emc/boostfs/bin/boostfs mount -d ddt.labd.local -s BOOSTMySQL /backup

```
root@szczeliniec://#: /opt/emc/boostfs/bin/boostfs mount -d ddt.labd.local -s BOOSTMySQL /backup

mount: Mounting ddt.labd.local:BOOSTMySQL on /backup
root@szczeliniec://#: ls -al /backup
total 8
drwxrwxrwx  2  501 users  101 Nov 24 16:45 .
drwxr-xr-x 29 root root  4096 Nov 24 21:43 ..
-r--r--r--  1 root root    83 Nov 24 21:54 .boostfs_cache
-r--r--r--  1 root root    79 Nov 24 21:54 .boostfs_connections
-r--r--r--  1 root root   114 Nov 24 21:54 .boostfs_nodes
-r--r--r--  1 root root    70 Nov 24 21:54 .boostfs_streaminfo
-r--r--r--  1 root root    27 Nov 24 21:54 .boostfs_sulist
-r--r--r--  1 root root    86 Nov 24 21:54 .boostfs_sysinfo
root@szczeliniec://#:
```


Performing backup
DELL EMC

Performing backup to BoostFS directory

Manual backup with just cp command

cp –R /production /backup/production_copy1

```
root@szczeliniec://#: cp -R /production/ /backup/production_copy1
root@szczeliniec://#:
```

We are the backup application
We – using the "cp" command – are writing to Data Domain.
Normally the backup is performed by backup application that automatically creates directories and files

Data Domain shows backup performance

Statistics from Data Domain

ddboost show stats interval 1

Data written to our /backup directory

Data that DD has got (what a nice de-dup!)

We write to /backup directory with 1 stream

```
sysadmin@ddt# ddboost show stats interval 1
11/24 23:58:0
  Backup          Post-comp          Network          Restore          Network       Backup   Restore
   KB/s           Written KB/s       In KB/s           KB/s            Out KB/s      Conn     Conn
--------------- --------------- --------------- --------------- --------------- ------- -------
              0               0               0               0               0       0       0
              0               0               0               0               0       0       0
              0               0               0               0               0       0       0
              0               0               0               0               0       0       0
         15,920               1               1               0               0       1       0
         31,228              44              44               0               0       1       0
          7,995               0               0               0               0       1       0
11/25 00:05:38
  Backup          Post-comp          Network          Restore          Network       Backup   Restore
   KB/s           Written KB/s       In KB/s           KB/s            Out KB/s      Conn     Conn
--------------- --------------- --------------- --------------- --------------- ------- -------
         23,291               5               5               0               0       1       0
         39,217               6               6               0               0       1       0
         41,250               9               9               0               0       1       0
         27,708              17              17               0               0       1       0
         31,421               2               2               0               0       0       0
```

DELL EMC

We see the huge de-duplication

Command
filesys show compression /data/col1/BOOSTMySQL
shows de-duplication statistics for our logical DD
BOOSTMySQL
which is our BOOST device mounted on server
Szczeliniec.labd.local

filesys show compression / data/col1/BOOSTMySQL

```
sysadmin@ddt# filesys show compression /data/col1/BOOSTMySQL
Total files: 124;  bytes/storage_used: 230.6
       Original Bytes:          991,051,760
  Globally Compressed:            4,310,188
   Locally Compressed:            3,814,613
            Meta-data:               47,792
```

We see the size of all protected data (991 051 760)

Amount of unique block (4 310 760).
Data Domain divided all data (991GB) on chunks 4K to 12K. Amount of unique blocks was just about 4,3MB. Probably similar data were backuped to different logical Data Domain before. In DD we have global de-dup!!!

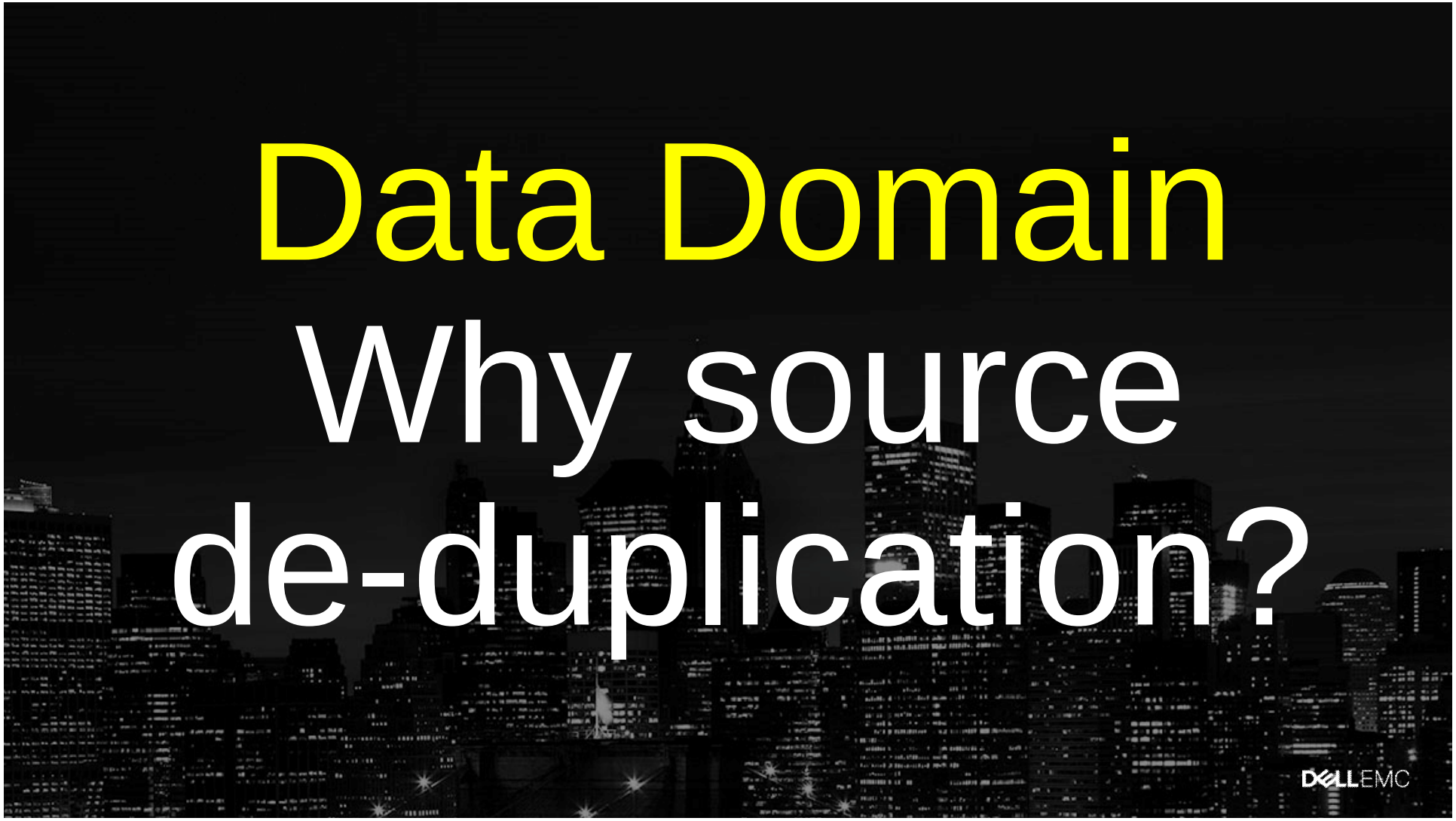
Unique blocks (4,3MB) were compressed before written to disk. After compression, unique blocks occupy on disk 3 814 613 bytes

We can observer which machines are performing backups

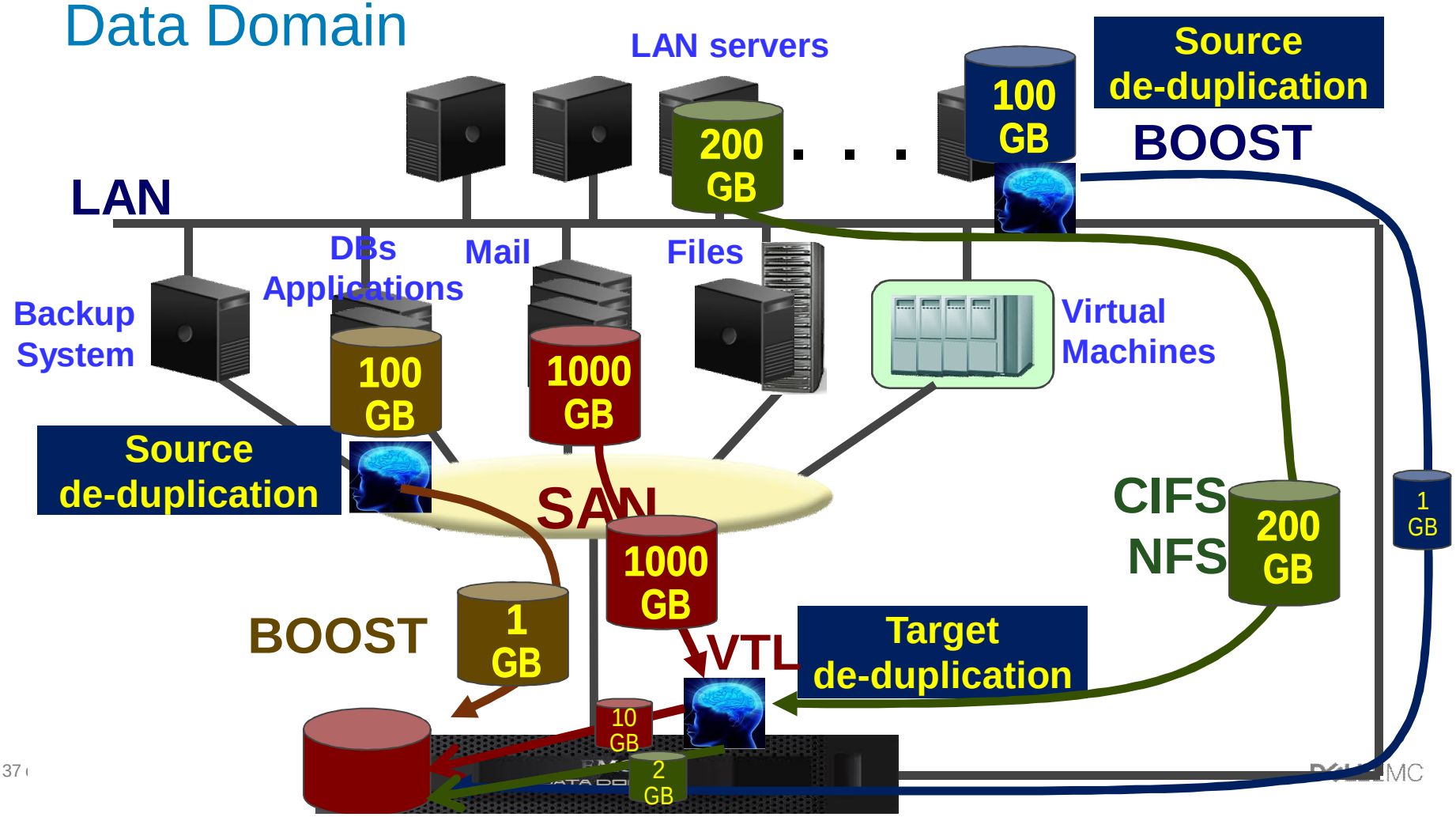
ddboost clients show active all

We can observe what clients are writing to our Data Domain!

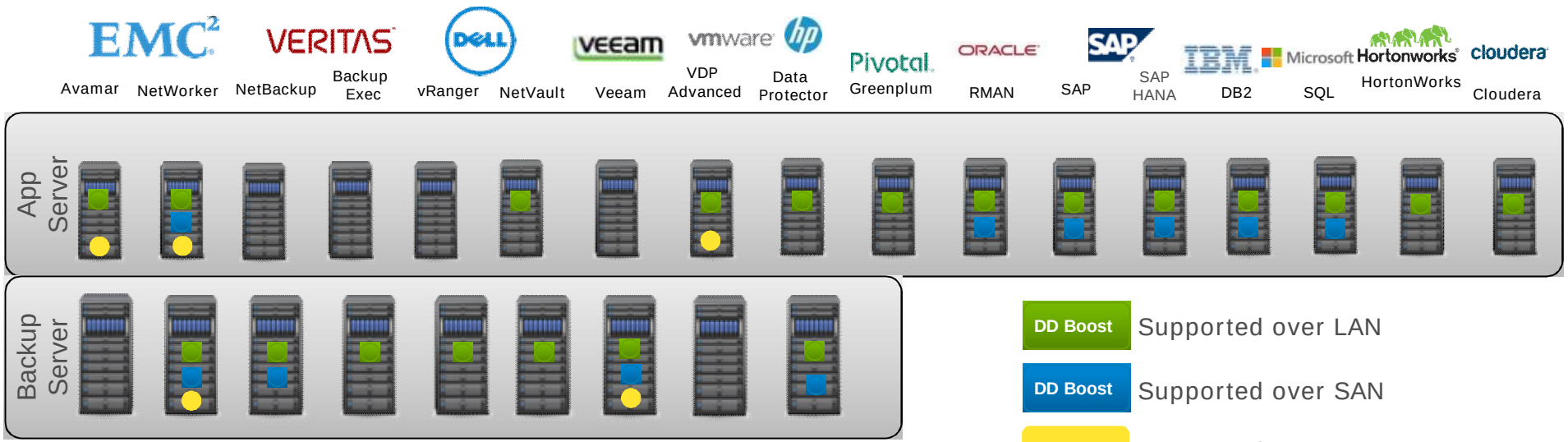
```
sysadmin@ddt# ddboost clients show active all
Client                  Client          Server          Operation   Mode   Storage-unit
Hostname                Interface       Interface
----------------------  ------------    ------------    ---------   ----   ------------
szczeliniec.labd.local  10.64.123.91    10.64.123.85    write       dsp    BOOSTMySQL
----------------------  ------------    ------------    ---------   ----   ------------
sysadmin@ddt#
```


Data Domain
Why source
de-duplication?
DELL EMC

Data Domain


LAN servers
200 GB
100 GB
Source de-duplication
BOOST
LAN
DBs Applications
Mail
Files
Backup System
100 GB
1000 GB
Virtual Machines
Source de-duplication
SAN
CIFS
NFS
200 GB
1 GB
1000 GB
BOOST
1 GB
VTL
Target de-duplication
10 GB
2 GB

DATA DOMAIN BOOST ECOSYSTEM



DD Boost Supported over LAN

DD Boost Supported over SAN

DD Boost Supported over WAN

For everything else, use the DD Boost file system plug-in

DELL EMC

Data Domain:
Why source de-duplication?

- Performance
  - The highest possible performance
  - Performance scalability
    - Across different machines
    - On the same machine

Data Domain:
Why source de-duplication?

- Performance
  - The same speed no matter if LAN or SAN
  - Over 60TB/h continues backup speed in real environments

Data Domain:
Why source de-duplication?

- No network bottleneck
- Source de-duplication for every backup software

Data Domain:
Why source de-duplication?

- No network load
- Client direct
- Simple management
- Network load balancing & **live** failover
  - Interface Group

Data Domain:
Why source de-duplication?

- Source de-duplication device (*Storage Unit)* is visible at the same time via LAN/SAN
  - You can backup via SAN and the same backup via LAN
  - You can backup via LAN and the same backup via SAN
  - Clients can backup / restore at the same time via LAN/SAN

Data Domain:
Why source de-duplication?

- Backup via WAN
  - Possible delays up to 20ms / 35ms / 200ms

Data Domain:
Why source de-duplication?

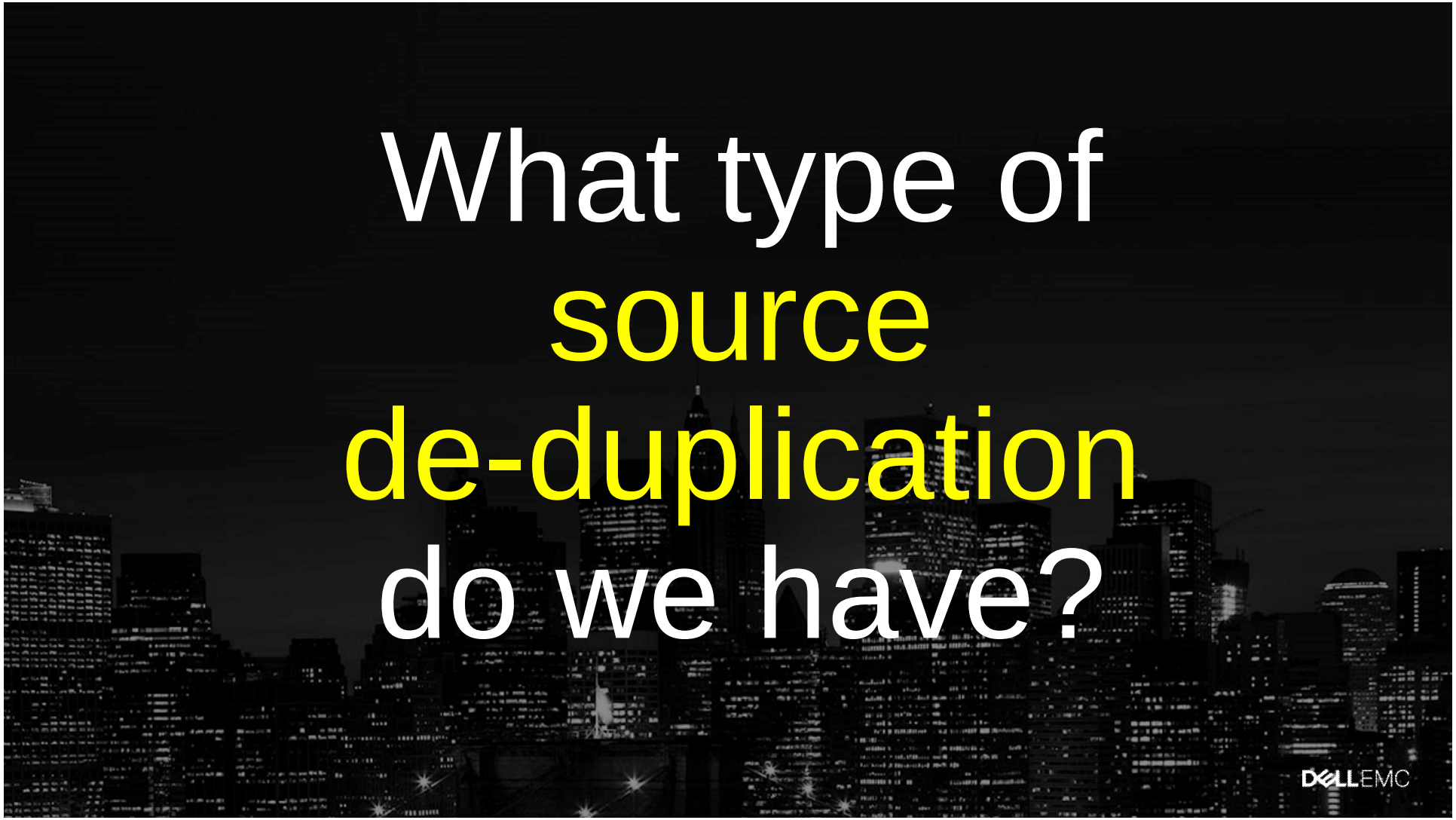
- Smaller costs
  - Fewer Media Servers
  - No need for dedicated network
  - Smaller amount of backup software licenses
  - Management

Data Domain:
Why source de-duplication?

- Disaster Recovery Managed by backup software
- Replication with granularity for single file
- Simple Disaster Recovery
  - Backup software which is aware of backup copies in both locations
-

Source de-duplication

- provide the highest backup performance
- technically best approach to backup/restore data from Data Domain
- provides the biggest value and uniuqe possibilities for backup/restore
- shall be used if possible
- is recommended way to backup Data Domain


What type of
source
de-duplication
do we have?
DELL EMC

DATA DOMAIN BOOST ECOSYSTEM

Avamar NetWorker NetBackup Backup Exec vRanger NetVault Veeam VDP Advanced Data Protector Greenplum RMAN SAP SAP HANA DB2 SQL HortonWorks Cloudera

App Server

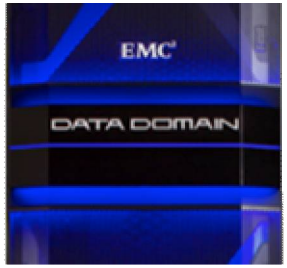
Backup Server

DD Boost Supported over LAN

DD Boost Supported over SAN

DD Boost Supported over WAN

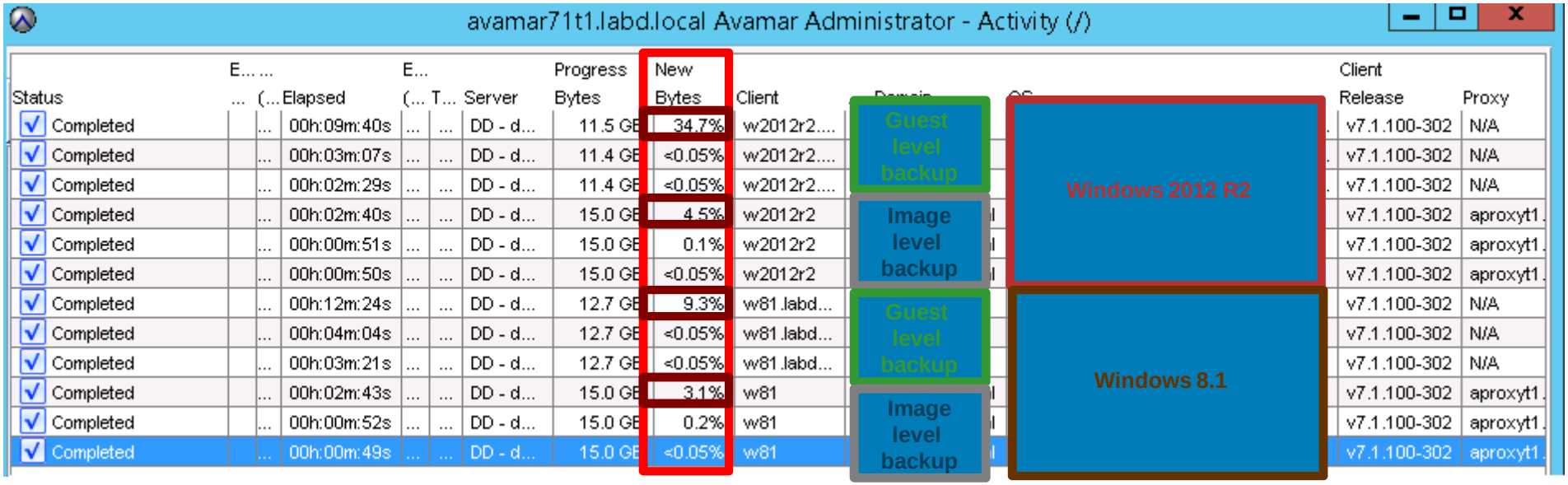
For everything else, use the DD Boost file system plug-in



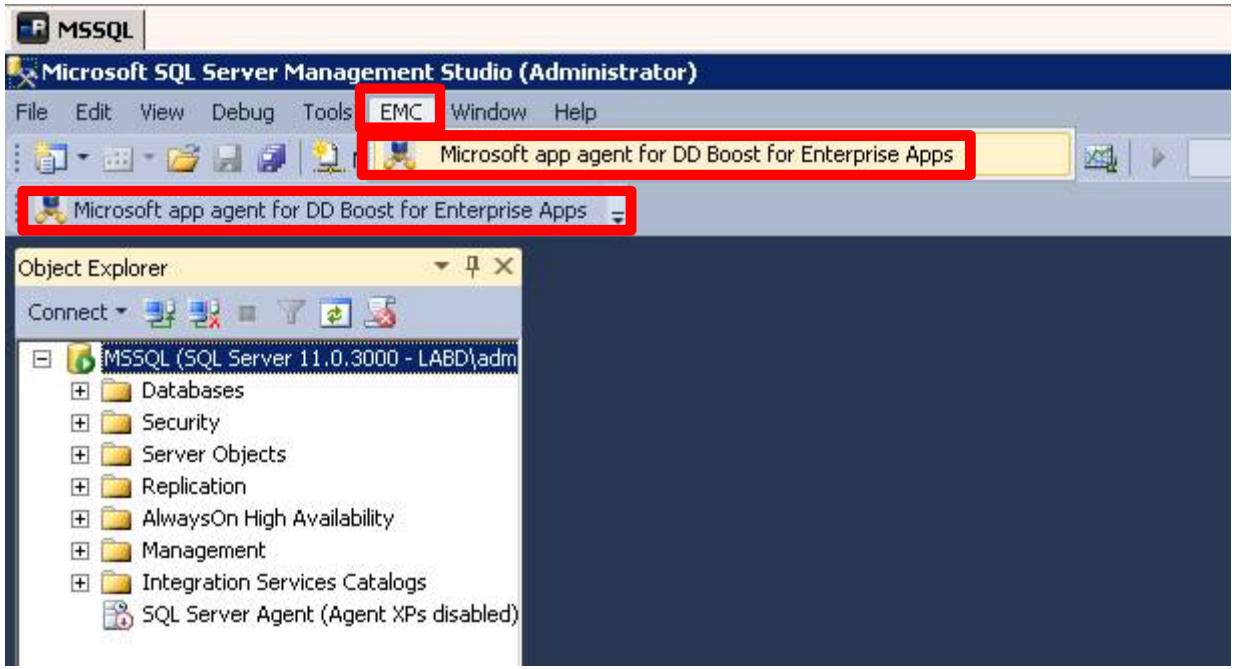
DELL EMC

Who perfroms source de-duplication?	Backup software	Application module	BoostFS

Who perfroms source de-duplication?	Backup software	Application module	BoostFS



Who perfroms source de-duplication?	Backup software	Application module	BoostFS



Who perfroms source de-duplication?	Backup software	Application module	BoostFS

cp –R /production /backup/production_copy1

```
root@szczeliniec://#: cp -R /production/ /backup/production_copy1
root@szczeliniec://#:
```

Who perfroms source de-duplication?	Backup software	Application module	BoostFS
Performance	Highest	Highest	Highest
Management	Backup software admin + required data (configs) from database admins	Database admins	Backup software admin or database admin
Monitoring	Backup software	Database admin / eCDM	According to backup type (backup Depends on backup method
Replication	Backup software is aware about replicated copy	Mtree (Data Domain level) replication	Mtree (Data Domain level) replication

Who perfroms source de-duplication?	Backup software	Application module	BoostFS
Backup /restore network	SAN/LAN (required support from backup application)	SAN/LAN	LAN
Number of streams	Any / Up to backup application	Any	64 per Storage Unit
Recommendation	Recommended if central backup software is responsible for backups	Recommended if databases admins own the backup	Only if source de-duplication at * backup software level * application module level is not avaialable
Price	Up to backup software vendor rules	Per amount of protected databases (source TB)	No additional charge
	BOOST on Data Domain hardware is required		


I want
more info
DELL EMC

More info
Where to find?

http://gurago.pl/

Technology guides

http://backuprecoveryman.blogspot.com/2016/02/przeglad-materiaow-po-polsku.html#DataDomain

Films, values, use cases, …

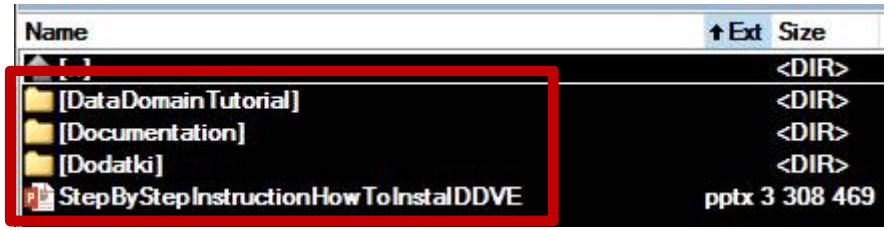
Just contact me….
Daniel.Olkowski@dell.com

gurago.pl
Technology guides

http://gurago.pl/

gurago.pl
Technology guides

http://gurago.pl/




Ogłoszenia
DELL EMC

Warsztaty technologiczne 2017

- Data Domain
- Avamar
- NetWorker
- VDP
- Zainteresowanych proszę o kontakt
  - *Daniel.Olkowski@dell.com*

Data Domain – Architektura i podstawy administracji

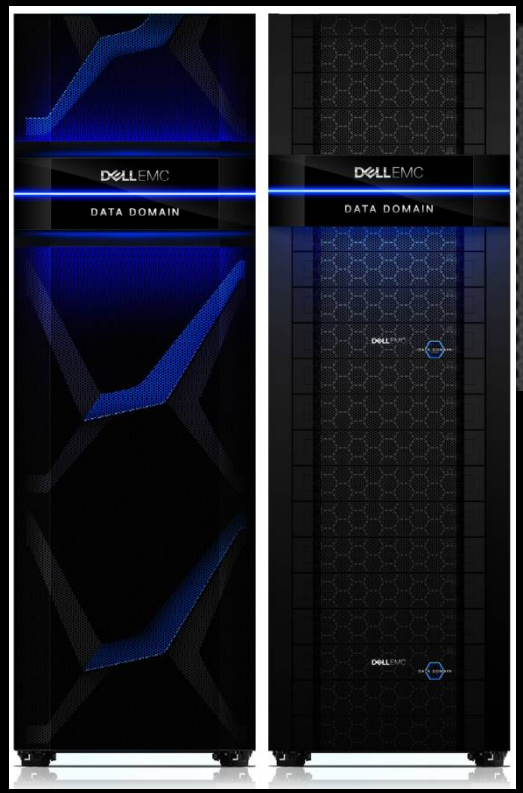
Do wyboru 2 terminy:

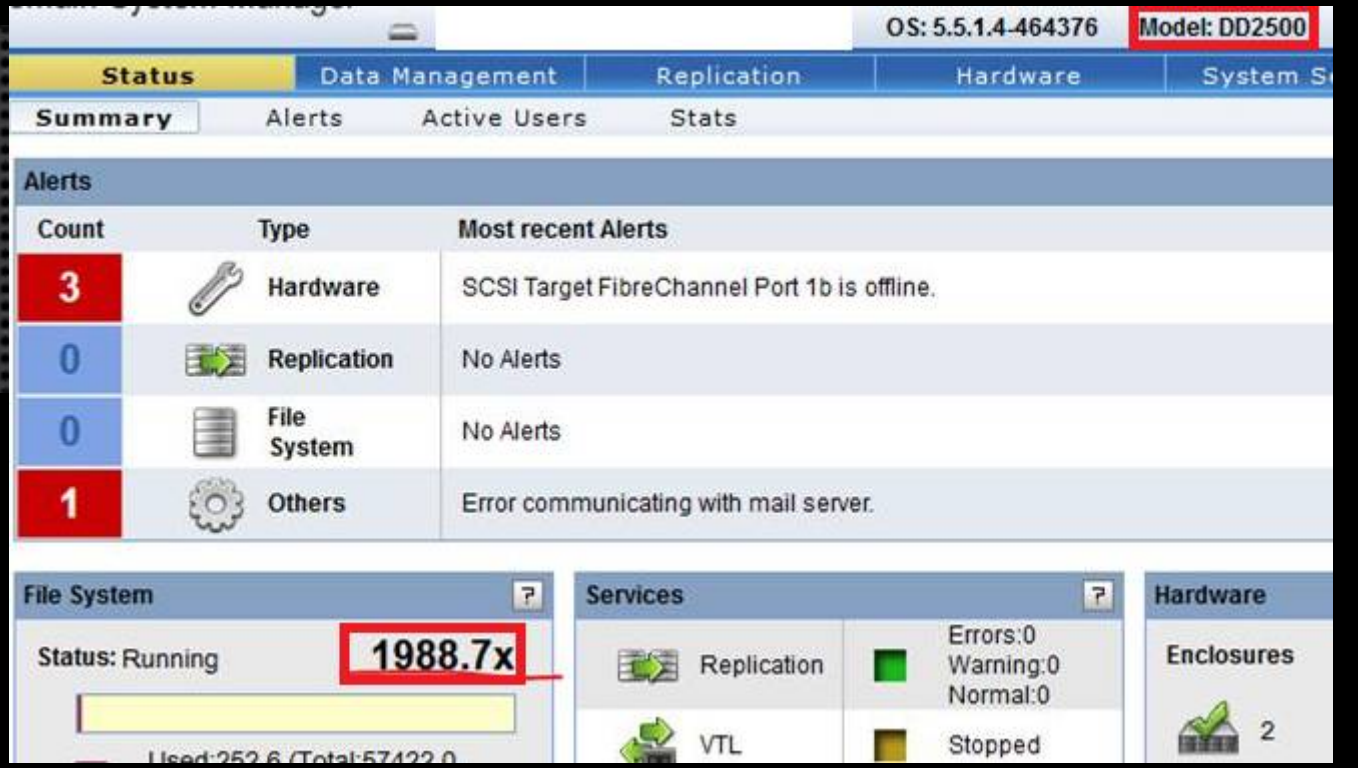
- 7 luty 2017 wtorek, 10:00 – 16:30. Rejestracja: https://emcinformation.com/472401/REG/.ashx
- 12 września 2017 wtorek, 10:00 – 16:30. Rejestracja: https://emcinformation.com/472501/REG/.ashx

Następny webcast

- 13 stycznia, piątek 10:00 – 10:30
- Nowości w Avamar 7.4

Questions…





Daniel.Olkowski@dell.com