

DE-DUPLIKACJA UWOLNIONA!

BoostFS

De-duplikacja na źródle dla każdego

Film omawiający BoostFS:

https://youtu.be/dSU_81xDoZ0

D~~E~~LL EMC

Agenda

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

A nighttime city skyline, likely New York City, with numerous skyscrapers illuminated. The image is overlaid with a semi-transparent grid of white text, resembling computer code or data, which is most visible in the lower half of the frame. The overall background is dark, with the city lights providing a high-contrast scene.

BoostFS

What is it?

BoostFS

What is it?

- BoostFS is source de-duplication that any
 - Backup software
 - Application
 - Database
 - Script
 - Commandcan 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 boxor
 - 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?



Tutorial jak zainstalować i skonfigurować Data Domain Virtual Edition

[Tutorial, krok po kroku, jak zainstalować i skonfigurować Data Domain Virtual Edition](#)

W powyższym pakiecie znajdziesz prezentacje które krok po kroku przeprowadzą Cię przez:

- a. Pobranie pakietu Data Domain Virtual Edition
- b. Zainstalowanie maszyny wirtualnej Data Domain Virtual Edition
- c. Backup SQL z de-duplikacją na źródle
- d. Konfiguracja sieci
- e. ...

9.

Wielu udanych odtworzeń!

BoostFS

I want to try – how?

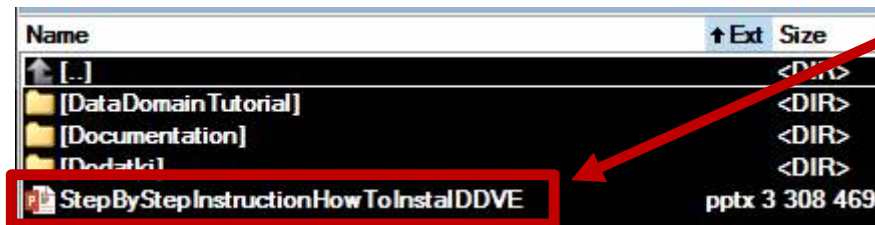
<http://gurago.pl/>



BoostFS

Chcę spróbować – jak?

<http://gurago.pl/>



Tutorial how to

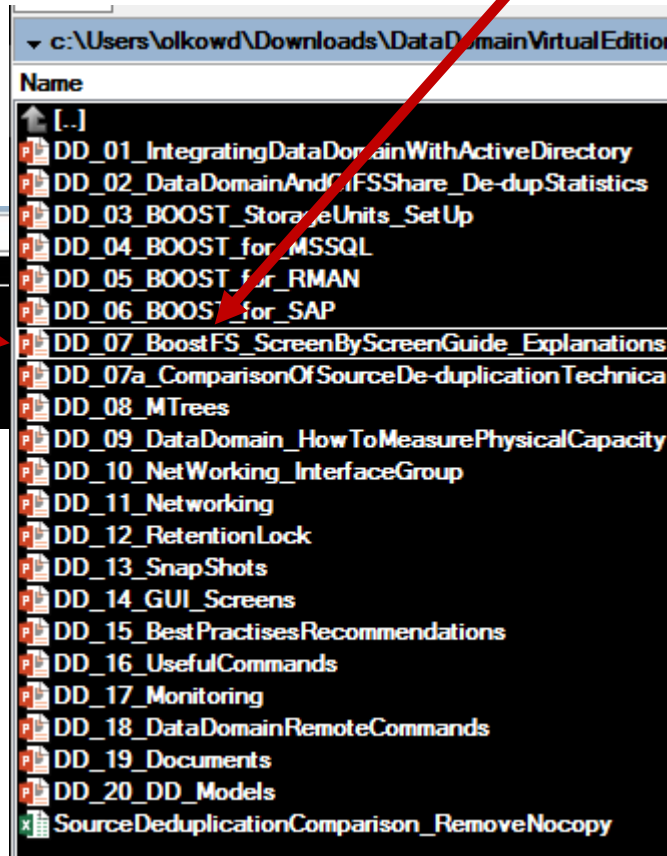
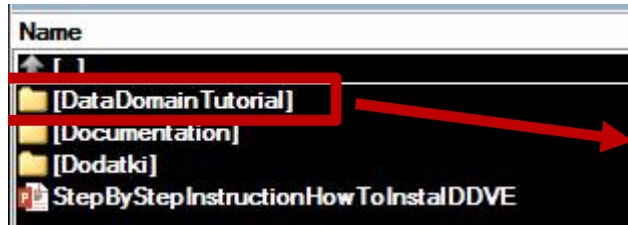
- Download
- Install
- Configure

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!

A nighttime city skyline, likely New York City, with numerous skyscrapers illuminated. The image is dark, with the city lights providing a textured background. Overlaid on the image is a grid of faint, light-colored text that appears to be computer code or data, creating a digital or technical atmosphere.

BoostFS

Requirements

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

Configuration

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...
 1:ddbostfs
root@szczeliniec:/daniel/#:
```

Creating Storage Unit (source de-dup device)

The screenshot shows the EMC Data Domain System Manager interface. The left sidebar has 'PROTOCOLS' and 'DD Boost' highlighted with red boxes. The main content area shows 'DD Boost Status: Enabled' (highlighted with a red box) and a 'Storage Units' tab (also highlighted with a red box). Below the tab is a table of storage units with a '+' icon (highlighted with a red box) for adding new units.

<input checked="" type="checkbox"/>	Storage Unit ▲	User ◆	Quota Hard Limit ◆	Last 24hr Pre-Comp ◆	Last 24hr Post-Comp ◆	Last 24hr Comp Ratio ◆	Weekly Avg Post Comp ◆	Last Week Post-Comp ◆	Weekly Avg Comp Ratio ◆	Last Week Comp Ratio ◆
<input checked="" type="checkbox"/>	networker	ddbost	None	327.8 GiB	3.7 GiB	89.7x	0.7 GiB	3.7 GiB	89.7x	89.7x

Creating Storage Unit (source de-dup device)

Create Storage Unit

Name: BOOSTMySQL

User: Select a Local User...
Select a Local User...
Create a new Local User...
ddboost (none)
sysadmin (admin)

Quota Settings

Pre-Comp Soft Limit: sysadmin (admin)

Pre-Comp Hard Limit:
 Set to specific value: GIB
 None
 Set to specific value: GIB

Creating Storage Unit (source de-dup device)

Create Storage Unit ✕

Name:

User:

User:

Password:

Verify Password:

Management Role: none

i The user will be added to the DD Boost access list.

Quota Settings

Pre-Comp Soft Limit None
 Set to specific value: GiB ▼

Pre-Comp Hard Limit None
 Set to specific value: GiB ▼

? Create Cancel

Creating Storage Unit (source de-dup device)

The screenshot shows the EMC Data Domain System Manager interface. The left sidebar contains navigation options: HOME, HEALTH, DATA MANAGEMENT, REPLICATION, PROTOCOLS, DD Boost (selected), CIFS, NFS, HARDWARE, and ADMINISTRATION. The main content area is titled 'Protocols DD Boost' and shows 'DD Boost Status: Enabled' with a 'Disable' button. Below this, 'Kerberos Mode: Windows / Active Directory' is shown with a 'Configure' link. A tabbed interface below has 'Storage Units' selected. A table of Storage Units is displayed, with the 'BOOSTMySQL' row highlighted in blue and enclosed in a red box. The table columns include checkboxes, Storage Unit names, Users, Quota Hard Limits, and various performance metrics.

<input type="checkbox"/>	Storage Unit	User	Quota Hard Limit	Last 24hr Pre-Comp	Last 24hr Post-Comp	Last 24hr Comp Ratio	Weekly Avg Post-Comp	Last Week Post-Comp	Weekly Avg Comp Ratio	Last Week Comp Ratio
<input checked="" type="checkbox"/>	BOOSTMySQL	ddmysql	None	0.0 GiB	0.0 GiB	0.0x	0.0 GiB	0.0 GiB	0.0x	0.0x
<input type="checkbox"/>	Networker	ddboost	None	327.8 GiB	3.7 GiB	89.7x	0.7 GiB	3.7 GiB	89.7x	89.7x

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`

The image displays two screenshots of the EMC Data Domain System Manager web interface, illustrating the configuration of BoostFS. Red and blue arrows connect the command line parameters to their corresponding UI elements.

Left Screenshot: Storage Units

Storage Unit	US	Quota Hard Limit	Last 24hr Pre-Comp	Last 24hr Post-Comp	
<input type="checkbox"/>	BOOSTMySQL	ddmysql	None	0.0 GiB	0.0 GiB
<input type="checkbox"/>	Networker	ddboost	None	327.8 GiB	3.7 GiB

Items Selected: 0

Right Screenshot: DD Boost Settings

DD Boost

DD Boost Status: **Enabled** [Disable](#)

Kerberos Mode: Windows / Active Directory [Configure](#)

Settings | Active Connections | IP Network

Allowed Clients

Client	Authentication
<input type="checkbox"/> *	None

Users with DD Boost Access

User	Status
<input type="checkbox"/> a1	enabled
<input type="checkbox"/> ddboost	enabled
<input type="checkbox"/> ddmysql	enabled

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/#: █
```

Create Storage Unit

Name: BOOSTMySQL

User: Create a new Local User...

User: ddmysql

Password: [Redacted]

Verify Password: [Redacted]

Management Role: none

The user will be added to the DD Boost access list.

Quota Settings

Pre-Comp Soft Limit: None Set to specific value: [] GIB

Pre-Comp Hard Limit: None Set to specific value: [] GIB

Buttons: Create, Cancel

Provide password for Boost user

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

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# ddbboost show stats interval 1
11/24 23:58:0
```

Backup KB/s	Post-comp Written KB/s	Network In KB/s	Restore KB/s	Network Out KB/s	Backup Conn	Restore 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 KB/s	Post-comp Written KB/s	Network In KB/s	Restore KB/s	Network Out KB/s	Backup Conn	Restore 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

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: 121, Bytes/Storage_Used: 238.8
  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 backed up 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 observe which machines are performing backups

ddboost clients show active all

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

```
sysadmin@ddt# ddbboost clients show active all
```

Client Hostname	Client Interface	Server Interface	Operation	Mode	Storage-unit
szczeliniec.labd.local	10.64.123.91	10.64.123.85	write	dsp	BOOSTMySQL

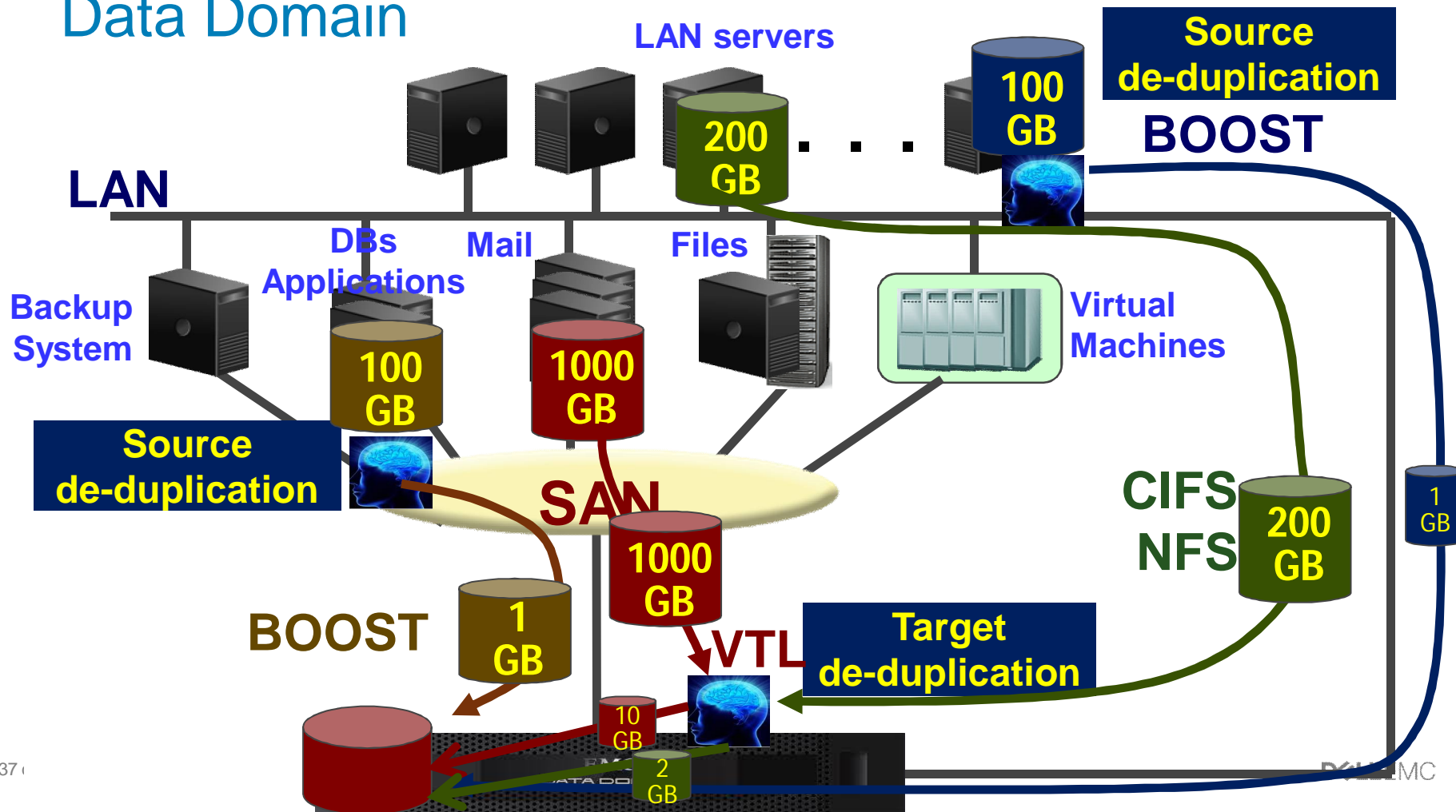
```
sysadmin@ddt# █
```



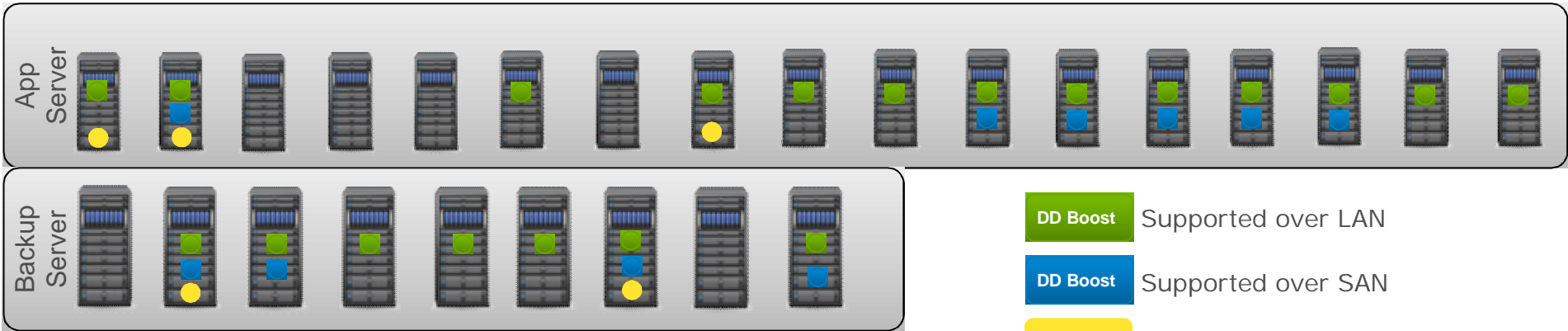
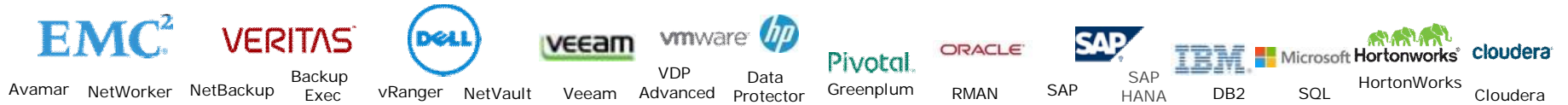
Data Domain

Why source de-duplication?

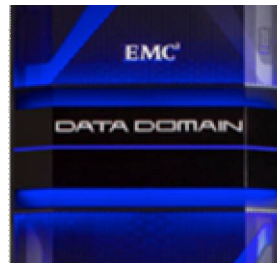
Data Domain



DATA DOMAIN BOOST ECOSYSTEM



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



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



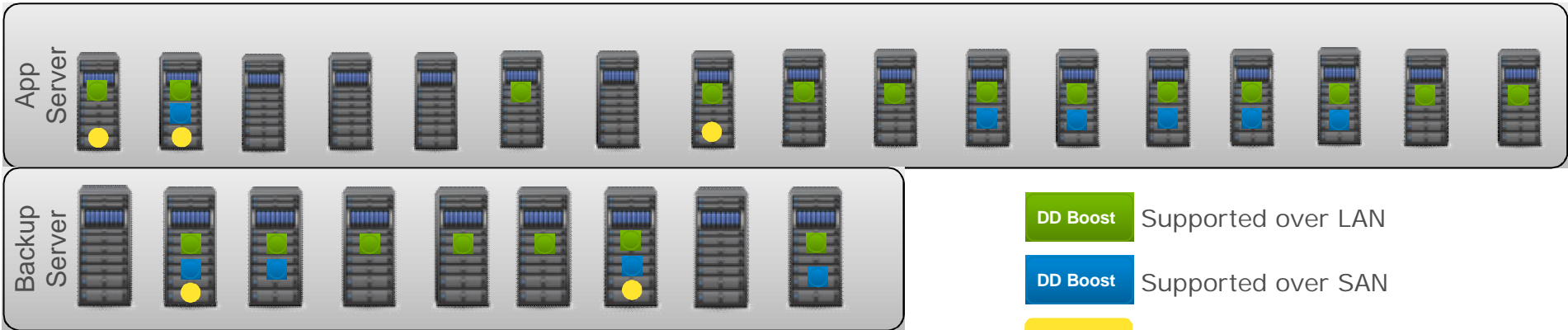
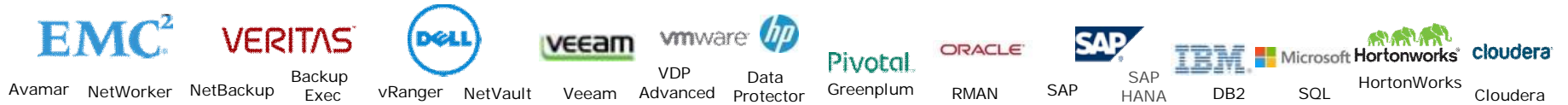
Source de-duplication

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

A nighttime photograph of a city skyline, likely New York City, with numerous skyscrapers illuminated against a dark sky. The lights from the buildings create a bokeh effect in the foreground.

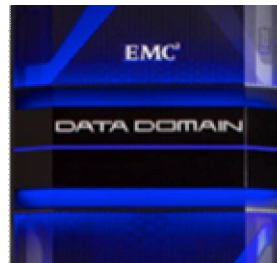
What type of
source
de-duplication
do we have?

DATA DOMAIN BOOST ECOSYSTEM

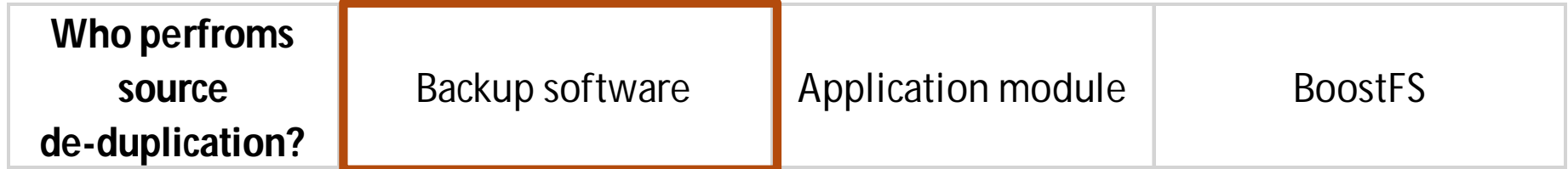


- DD Boost Supported over LAN
- DD Boost Supported over SAN
- DD Boost Supported over WAN

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



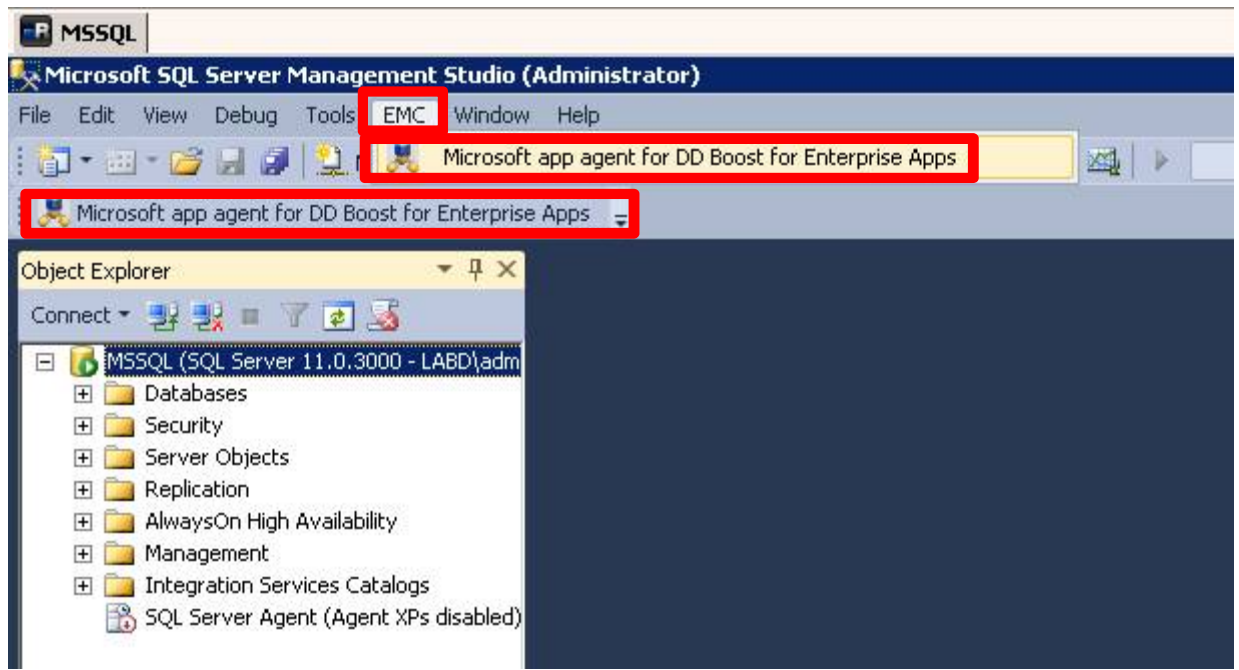
Who performs source de-duplication?	Backup software	Application module	BoostFS
--	-----------------	--------------------	---------



avamar71t1.labd.local Avamar Administrator - Activity (/)

Status	...	E... (... Elapsed	E... (... T...	Server	Progress Bytes	New Bytes	Client	Domain	OS	Client Release	Proxy
✓ Completed	...	00h:09m:40s	...	DD - d...	11.5 GB	34.7%	w2012r2....	Guest level backup	Windows 2012 R2	v7.1.100-302	N/A
✓ Completed	...	00h:03m:07s	...	DD - d...	11.4 GB	<0.05%	w2012r2....	Image level backup		v7.1.100-302	N/A
✓ Completed	...	00h:02m:29s	...	DD - d...	11.4 GB	<0.05%	w2012r2....	Image level backup		v7.1.100-302	aproxyt1.
✓ Completed	...	00h:02m:40s	...	DD - d...	15.0 GB	4.5%	w2012r2	Image level backup		v7.1.100-302	aproxyt1.
✓ Completed	...	00h:00m:51s	...	DD - d...	15.0 GB	0.1%	w2012r2	Image level backup	Windows 8.1	v7.1.100-302	aproxyt1.
✓ Completed	...	00h:00m:50s	...	DD - d...	15.0 GB	<0.05%	w2012r2	Image level backup		v7.1.100-302	aproxyt1.
✓ Completed	...	00h:12m:24s	...	DD - d...	12.7 GB	9.3%	w81.labd...	Guest level backup		v7.1.100-302	N/A
✓ Completed	...	00h:04m:04s	...	DD - d...	12.7 GB	<0.05%	w81.labd...	Image level backup		v7.1.100-302	N/A
✓ Completed	...	00h:03m:21s	...	DD - d...	12.7 GB	<0.05%	w81.labd...	Image level backup	Windows 8.1	v7.1.100-302	N/A
✓ Completed	...	00h:02m:43s	...	DD - d...	15.0 GB	3.1%	w81	Image level backup		v7.1.100-302	aproxyt1.
✓ Completed	...	00h:00m:52s	...	DD - d...	15.0 GB	0.2%	w81	Image level backup		v7.1.100-302	aproxyt1.
✓ Completed	...	00h:00m:49s	...	DD - d...	15.0 GB	<0.05%	w81	Image level backup		v7.1.100-302	aproxyt1.

<p>Who performs source de-duplication?</p>	<p>Backup software</p>	<p>Application module</p>	<p>BoostFS</p>
---	------------------------	---------------------------	----------------



Who performs 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 performs 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 performs 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 available
Price	Up to backup software vendor rules	Per amount of protected databases (source TB)	No additional charge
BOOST on Data Domain hardware is required			

A nighttime photograph of a city skyline, likely New York City, with numerous skyscrapers illuminated against a dark sky. The lights from the buildings create a bokeh effect in the foreground. The text "I want more info" is overlaid on the image.

I want
more info

DELL EMC

More info

Where to find?

<http://gurago.pl/>

Technology guides

<http://backuprecoveryman.blogspot.com/2016/02/przegląd-materiałów-po-polsku.html#DataDomain>

Films, values, use cases, ...

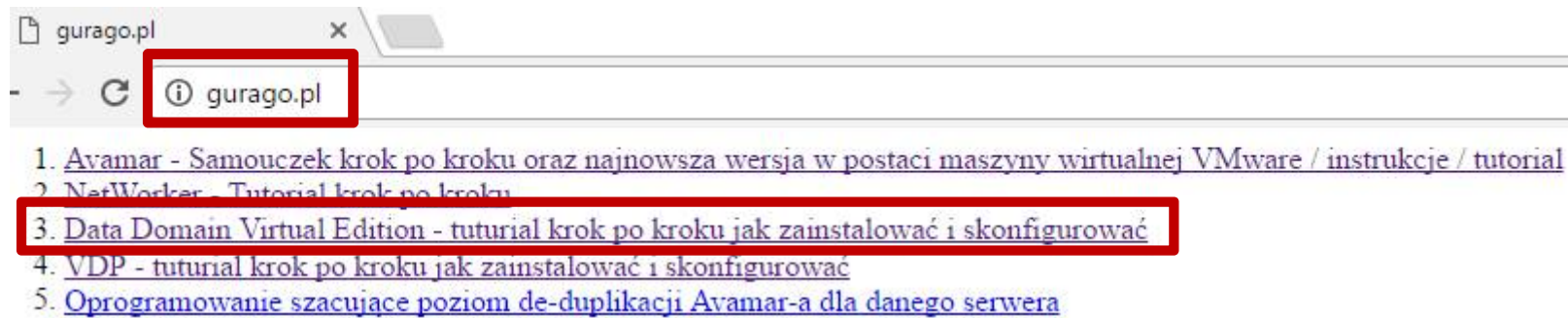
Just contact me....

Daniel.Olkowski@dell.com

gurago.pl

Technology guides

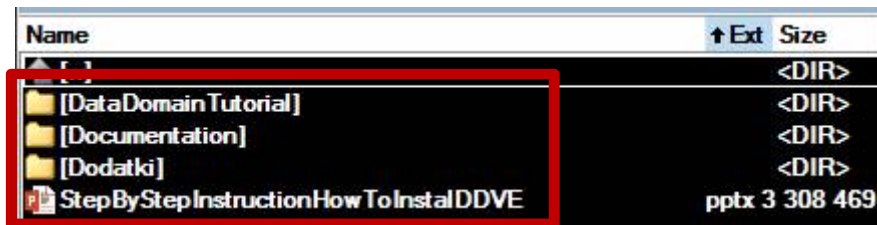
<http://gurago.pl/>



gurago.pl

Technology guides

<http://gurago.pl/>



Name	↑ Ext	Size
[DataDomain Tutorial]		<DIR>
[Documentation]		<DIR>
[Dodatki]		<DIR>
StepByStepInstructionHowToInstalDDVE		pptx 3 308 469

Ogłoszenia

Warsztaty technologiczne 2017

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

Data Domain – Architektura i podstawy administracji

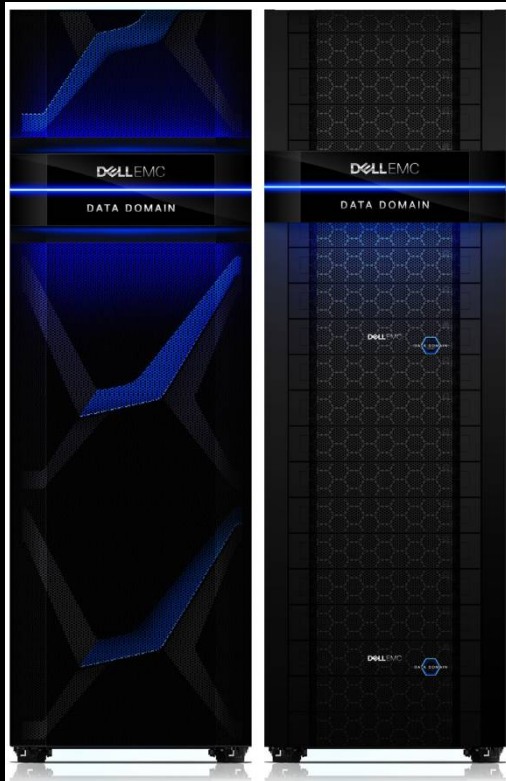
Do wyboru 2 terminy:

- 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...



OS: 5.5.1.4-464376 Model: DD2500

Status Data Management Replication Hardware System S

Summary Alerts Active Users Stats

Alerts

Count	Type	Most recent Alerts
3	Hardware	SCSI Target FibreChannel Port 1b is offline.
0	Replication	No Alerts
0	File System	No Alerts
1	Others	Error communicating with mail server.

File System	Services	Hardware
Status: Running 1988.7x Used: 252.6 / Total: 57422.0	Replication Errors: 0 Warning: 0 Normal: 0 VTL Stopped	Enclosures 2

Daniel.Olkowski@dell.com