

NetBSD LVM implementation

Adam Hamsik
haad@netbsd.org

EuroBSDcon 2009,
Cambridge, UK
September 2009.

Talk outline

- Introduce Logical Volume Manager in general
- Introduce NetBSD LVM implementation
- explain how, why was it done
- give usage examples
- Q&A

Introduction

- Missing major feature of a server OS
 - Linux, AIX, HP-UX, Solaris, FreeBSD has it
- Available disk space is growing
 - 3-4 terabytes in a common server
- Number of disks in a server is much higher than it used to be

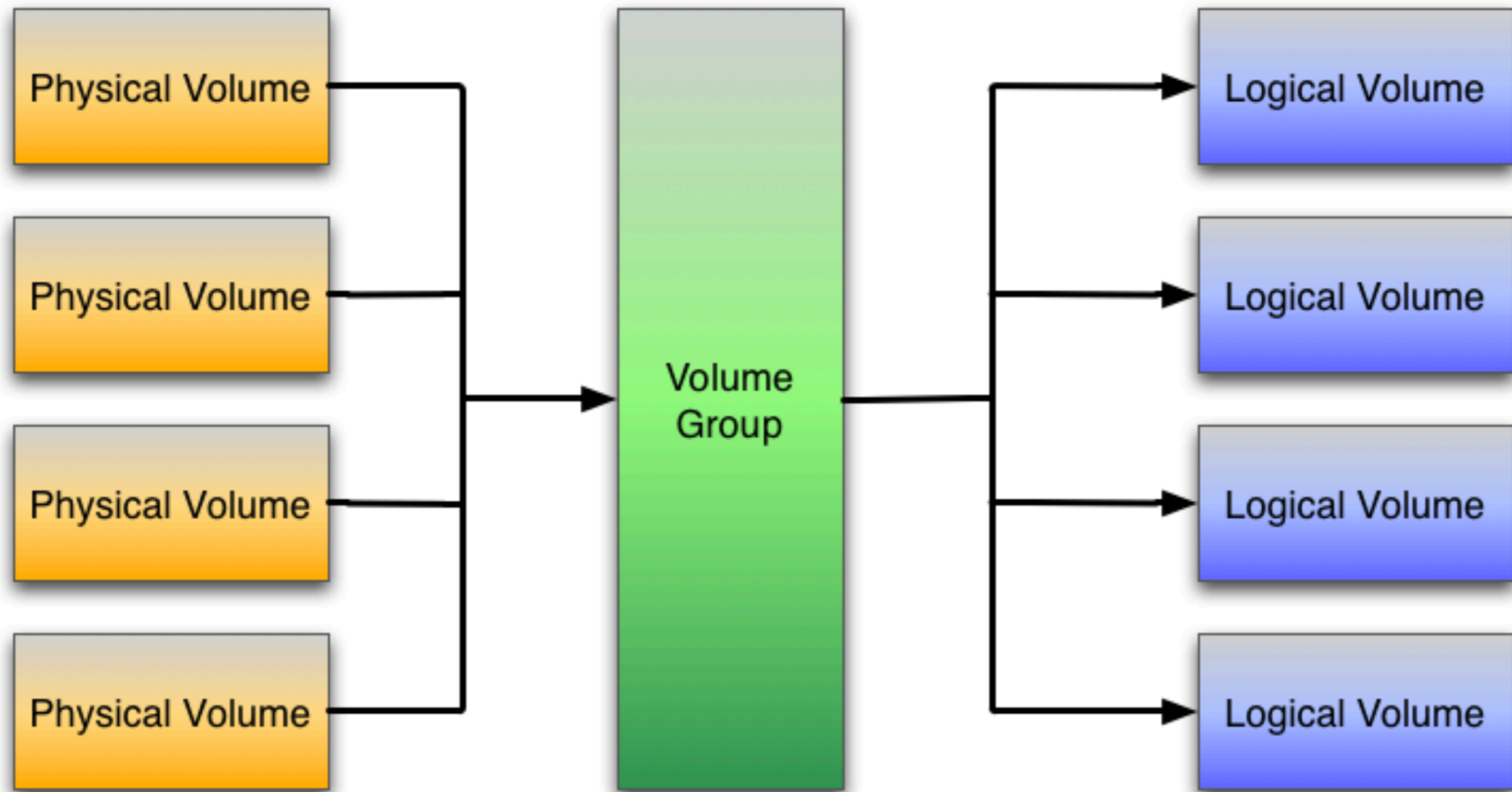
Motivation

- Hard and Interesting project
- Make administration of NetBSD servers much easier than it is now
- Learn something new about NetBSD internals

Traditional disk partitioning methods

- DOS partitions
- BSD disklabel
- GPT partitions
 - new partitioning scheme developed by Intel as a part of EFI

Logical Volume Manager design



Details

- Every LVM implementation has some sort of Physical Volume which is placed on a hard disk
- Volume Group is a pool of available disk space from which virtual partitions can be created
- Logical volumes are virtual partitions which can be used in a system like normal disks

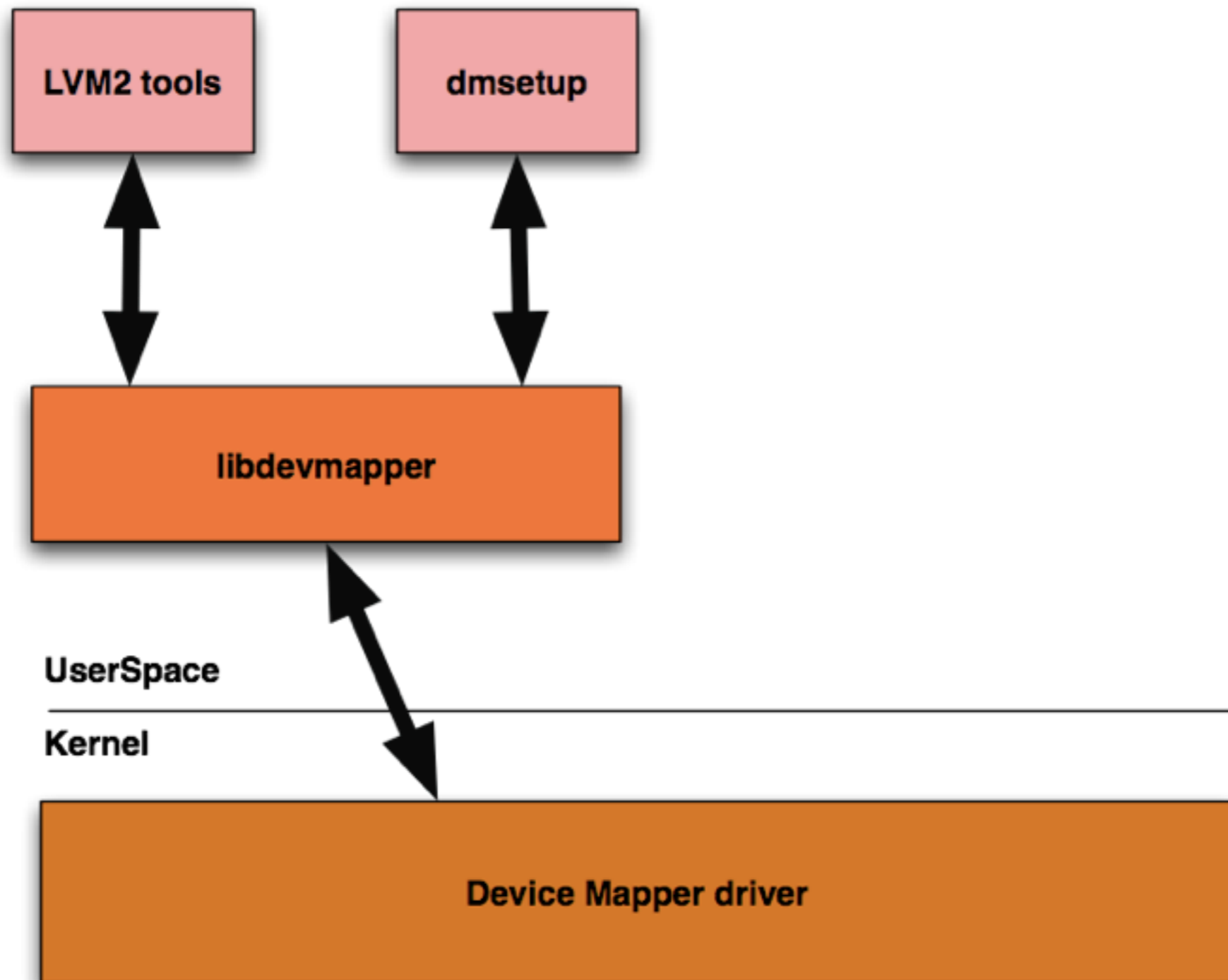
Elements in LVM

- PE - Physical Extends
- PV - Physical Volume
- VG - Volume Group
- LE - Logical Extends
- LV - Logical Volume

NetBSD LVM

- Done during Google Summer of Code 2008
- Simple BSD licensed kernel driver which maps real disk blocks to virtual ones
 - driver is called device-mapper
- Linux lvm tools which manage LVM metadata.

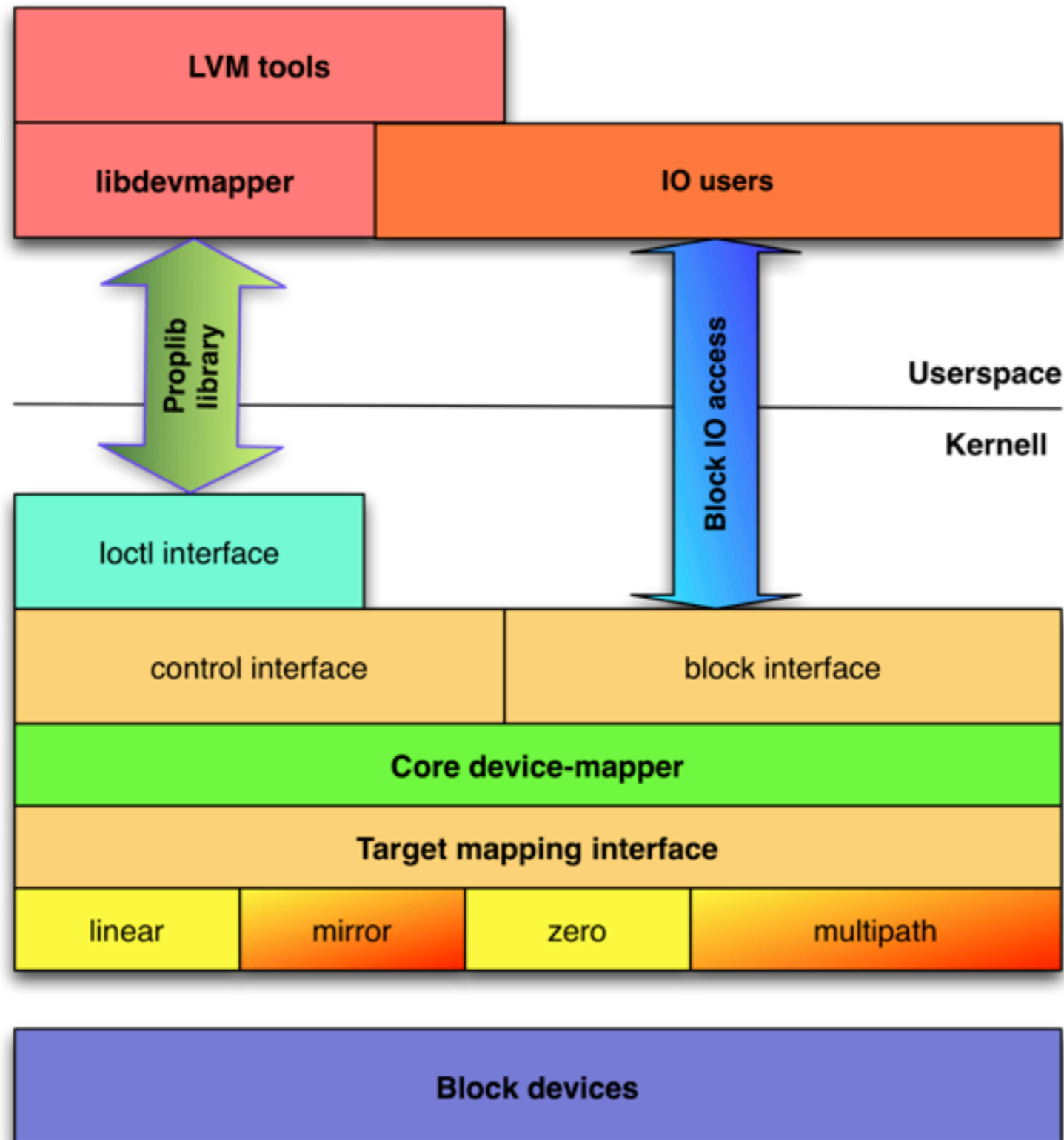
NetBSD LVM architecture



Device-mapper

- Clean and small reimplementations of Linux device-mapper driver
- BSD licensed
- Implements new kernel/userspace protocol based on proplib library
- Implements linear and stripe targets
- SMP ready

Device-mapper architecture



DM implementation

- For every device in device-mapper there is a table which describes how are physical blocks mapped to virtual one

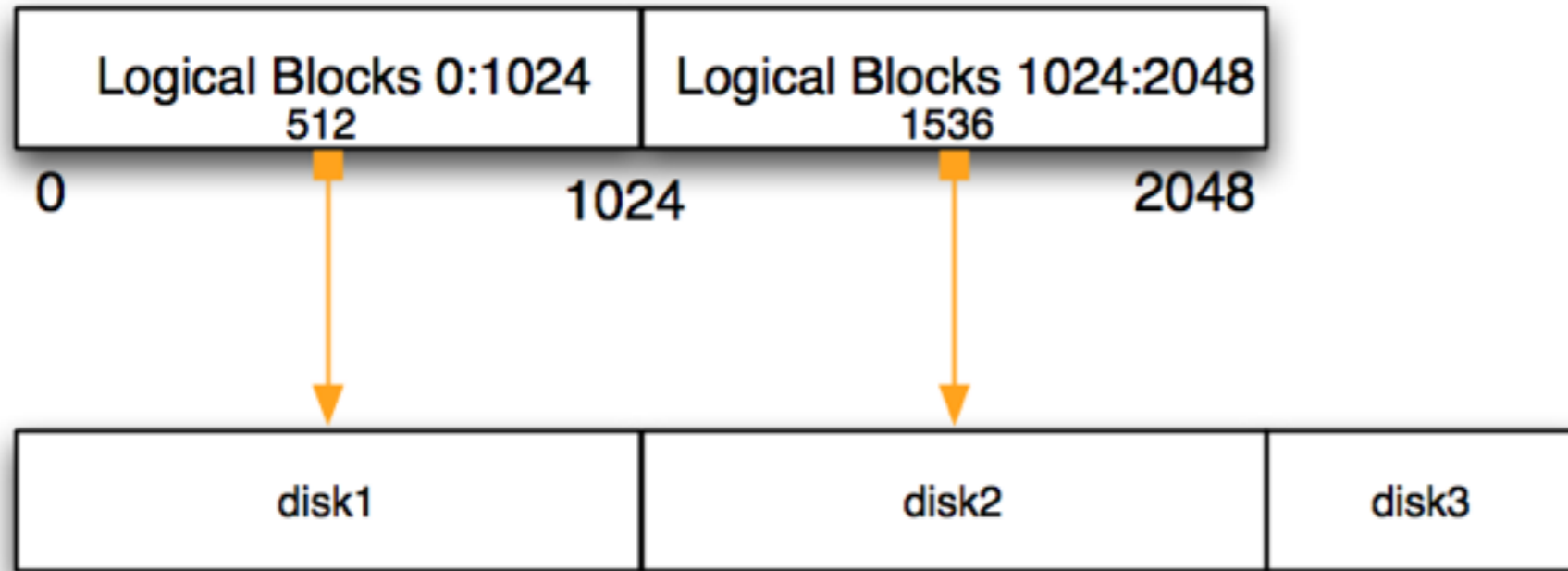
```
vg01-home: 0 67108864 linear /dev/wd0a  
1024
```

- Device can't have more than one table loaded at time but one table can have more lines

Device-mapper targets

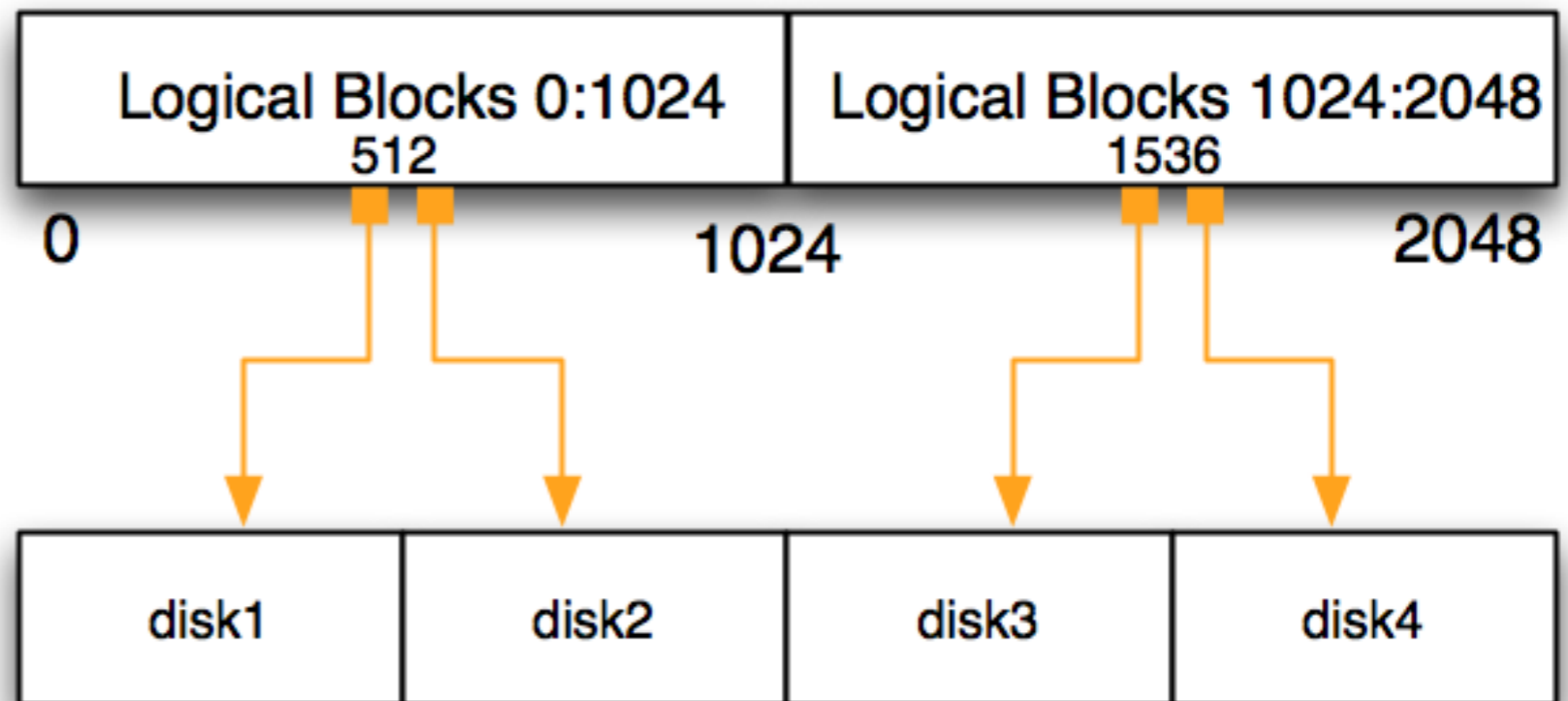
- Targets maps real block devices to virtual ones
- There are linear, stripe, mirror, multipath and snapshot targets in Linux
- In NetBSD linear and stripe targets were implemented which is enough to get LVM properly working

Available Targets



Linear Target

Stripe Target



Use cases

- Xen server with Virtual Machine disks on Logical Volumes
- NAS device with iSCSI disks placed on a Logical volumes
- Backup servers

Future Work

- Cluster LVM support
- Adding new targets mirror and snapshot
- Implementation of DRBD like device which can be used for creation of raid disks through the network

More Info

- <http://netbsd.org/docs/guide/en/index.html>
- `src/sys/dev/dm` and `src/external/gpl2/lvm2`
- NetBSD mailing lists
- EuroBSDcon 2009 paper

Questions & Answers

???