

Storage: What is offered, what do we archives really need, and how do we get there?

Peter Bubestinger

Introduction

The default

The alternatives

Example

End

Storage: What is offered, what do we archives really need, and how do we get there?

Peter Bubestinger

26. Jun 2014

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Who's the target audience?

Who's the main target audience?

Hint: It's not long-term preservation. . . :)

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Designed for:

- High performance computing
- Webservers
- Banks, Hospitals, etc.
- Data-centers
- . . .

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Regular professional storage?

Regular professional storage

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Main concerns:

- High speed
- Minimal downtime
- Real-time replication (failover)

Regular professional storage

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Main concerns:

- High speed
- Minimal downtime
- Real-time replication (failover)

What is offered:

- Hardware RAID
- Block-based clustering
- Quick release connectors for:
 - Disks
 - Fans
 - Power supply

Why is this offered?

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Quick tech crash-course

- **Hardware RAID:**
 - Array of disks, acting like one big disk
 - Saves you if individual disks are failing
- **Block-based clustering:**
 - Files are scattered in byte-blocks
 - Needed to scale beyond one physical machine
- **Real-time replication:**
 - Synchronous, transparent replication of data
 - One failover node = requires 2x capacity
- **Quick release connectors:**
 - Good for being fast at changing components

Everything has its price

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

The devil is in the details:

- **Hardware RAID:**
 - Proprietary, non-interoperable RAID protocols.
 - Vendor-specific interface/tools/features.
 - Dependence on special hardware
- **Block-based clustering:**
 - Failure = “byte pulp”
 - Dependence on *a specific* cluster technology
- **Real-time replication:**
 - This is not a backup
 - Wait for sync: Problem over low bandwidth (tape, internet)
- **Quick release connectors:**
 - Incompatible, proprietary form factors

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

What do archives need?

What do archives need?

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Main concerns:

- Data robustness
- Infinite migration to “future technology x”
- Scalability (e.g. size)
- Asynchronous backup copies

What do archives need?

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Main concerns:

- Data robustness
- Infinite migration to “future technology x”
- Scalability (e.g. size)
- Asynchronous backup copies

What we'd like:

- Affordable, maintainable hardware
- Spare parts
- Technology- & vendor-neutrality
- File-based clustering
- Well-known, high quality, open solutions

What's the option?

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Possible alternatives:

- Software RAID (Linux Kernel)
- GlusterFS (=file-based cluster filesystem)
- Asynchronous backups
- (More) regular hardware - where possible

Again: Everything has its price

Storage: What is offered, what do we archive really need, and how do we get there?

Peter Bubestinger

Introduction

The default

The alternatives

Example

End

Different “devils”:

- **Software RAID:**
 - Maybe less performance
- **GlusterFS:**
 - Maybe less performance
- **Asynchronous backups:**
 - Manual failover/recovery necessary
- **Regular hardware:**
 - Maybe less performance
 - More time to change disks/fans/etc

Again: Everything has its price

... but it might be worth paying it

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Different devils, but better for archiving:

- **Linux Software RAID:**
 - No hardware dependency
 - Documented, interoperable RAID protocols
 - Common tools/features
- **GlusterFS:**
 - No “byte-pulp” in case of failure: Always whole files
 - No specific cluster-system dependency

Again: Everything has its price

...but it might be worth paying it

Storage: What is offered, what do we archive really need, and how do we get there?

Peter Bubestinger

Introduction

The default

The alternatives

Example

End

Different devils, but better for archiving:

- **Asynchronous backups:**
 - Individual recovery possible
 - Backups to low-bandwidth storage (tape, internet)
- **Regular hardware:**
 - More choice/options
 - Broader userbase
 - (Better) documentation
 - Spare parts
 - Less vendor-dependence

The Austrian Mediathek

A real world example

Storage: What is offered, what do we archive really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Our new storage: “DLP-Storage”

- 2 x 300 TiB disk-based
- Open Hardware: “Backblaze pods”
- Each pod = a normal PC server mainboard + 45 disks
- Linux SoftRAID-6 + XFS + GlusterFS
- Hashcode-aware backups

The Austrian Mediathek

A real world example

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Our new storage: “DLP-Storage”

- 2 x 300 TiB disk-based
- Open Hardware: “Backblaze pods”
- Each pod = a normal PC server mainboard + 45 disks
- Linux SoftRAID-6 + XFS + GlusterFS
- Hashcode-aware backups

Note: You can use *any* hardware you want.
It just needs to run GNU/Linux.

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Oh!
Did I mention...

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

... it's completely
100% Free-and-Open-Source
software? :)

Sounds difficult?

...or just too good to be true?

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Your options to have one:

- Build your own
- Ask “Linbit” (www.linbit.com)
- ...or other Linux-storage administrators of *your* choice.

Storage: What is offered, what do we archives really need, and how do we get there?

Peter
Bubestinger

Introduction

The default

The alternatives

Example

End

Questions?

Comments?

Thank you very much for your attention!

Some rights reserved...

This presentation is available under a Free License:
Creative Commons Attribution Share-Alike
(CC-BY-SA)

Contact:

- Österreichische Mediathek: www.mediathek.at
- Peter Bubestinger: peter.bubestinger@mediathek.at