

# StaaS: almacenamiento como servicio

info@flossystems.com



Murcia, 1-2 de junio de 2012

StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

© 2012 FLOSSystems S.L.

This work is licensed under  
a Creative Commons Attribution 3.0 License



<http://creativecommons.org/licenses/by/3.0/es>

# Quiénes somos: FLOSSYSTEMS



- ▶ Larga experiencia como administradores de sistemas.
- ▶ Fundadores de **FLOSSYSTEMS**, compañía especializada en servicios avanzados basados en plataformas y sistemas con software libre.
- ▶ Diseño de plataformas: Virtualización, Clusters de Alta Escalabilidad, Clusters HA, Almacenamiento masivo, Seguridad...
- ▶ También impartimos formación especializada.
- ▶ Más info: <http://flossystems.com>



StaaS:

almacenamiento como servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes lógicos (LVM)

Ejemplo de LVM en Linux

ZFS



FLOSSystems S.L.,  
Madrid, Spain



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## RAID

Niveles estándar de RAID

## LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

## ZFS

## RAID

Niveles estándar de RAID

## LVM

Gestión de volúmenes lógicos (LVM)

Ejemplo de LVM en Linux

## ZFS



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## RAID

Niveles estándar de RAID

## LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

## ZFS

Por qué estudiar las tecnologías de almacenamiento:

- ▶ Es una de las tecnologías básicas que se combina con clusters, virtualización y cloud computing.
- ▶ Rendimiento y escalabilidad, replicación, migración no disruptiva de datos...
- ▶ Virtualización del almacenamiento.
- ▶ Reducción de los requerimientos de espacio y ahorro en costes de energía.



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## 6 RAID

Niveles estándar de RAID

### LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

### ZFS

# RAID



# RAID: Redundant Array of Independent Disks



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## 7 RAID

Niveles estándar de RAID

### LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

### ZFS

- ▶ Es un sistema que utiliza varios discos duros para distribuir o replicar datos a través de los discos.
- ▶ Evita pérdida de datos.
- ▶ Minimiza los tiempos de caída asociados a fallos de hardware (a menudo los reduce a cero).
- ▶ También puede incrementar el rendimiento.
- ▶ Se puede implementar en el hardware o vía software.



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## 8 RAID

Niveles estándar de RAID

### LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

### ZFS

RAID puede hacer dos cosas básicas:

1. Puede mejorar el rendimiento dividiendo (**stripping**) los datos a través de varios discos, que trabajan simultáneamente con un flujo único de datos.
2. Puede duplicar datos (**mirror**) a través de varios discos, reduciendo el riesgo asociado al fallo de un disco.





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

## RAID

9 Niveles estándar de RAID

## LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

## ZFS

- ▶ **RAID 0** (**stripping**): discos divididos sin paridad ni espejo
- ▶ **RAID 1** (**mirroring** o duplicación): es el primer nivel que ofrece redundancia.
- ▶ **RAID 4**: divide el volumen con paridad dedicada. Compite (y pierde en consistencia) con RAID 5.
- ▶ **RAID 5**: Volumen dividido (*stripped*) con paridad distribuida. RAID 5 requiere al menos 3 discos.
- ▶ **RAID 10** o **RAID 1+0**: es un volumen de datos espejado (RAID 1) que a su vez es dividido (RAID 0). RAID 10 requiere al menos 4 discos.

# RAID 0



StaaS:  
almacenamiento como  
servicio  
info@flossystems.com

- ▶ **RAID 0** (discos divididos sin paridad ni **mirroring**): usa dos o más discos de igual tamaño para reducir los tiempos de acceso y escritura. Se emplea exclusivamente para mejorar rendimiento.
- ▶ Tolerancia a fallos: 0 discos

## RAID

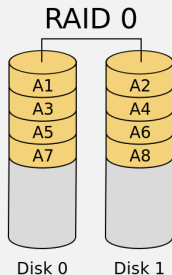
10 Niveles estándar de RAID

## LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

## ZFS



# RAID 1



StaaS:  
almacenamiento como  
servicio

info@flossystems.com

RAID

11 Niveles estándar de RAID

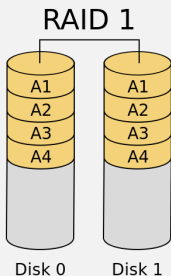
LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

- ▶ Volumen duplicado (“espejado”) sin paridad ni stripping: ofrece redundancia. Los datos son duplicados en dos o más discos de forma simultánea.
- ▶ Tolerancia a fallos:  $n-1$  discos



# RAID 4



StaaS:  
almacenamiento como  
servicio

info@flossystems.com

RAID

12 Niveles estándar de RAID

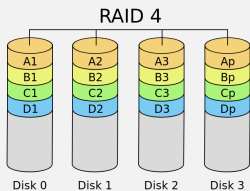
LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

- ▶ Discos divididos con un disco dedicado a información de paridad.
- ▶ Incurre en tiempos de espera cuando escribe la paridad.
- ▶ Pierde en comparación con RAID 5, su competidor.
- ▶ Tolerancia a fallos: 1 discos
- ▶ Requiere al menos 3 discos.



# RAID 5



StaaS:  
almacenamiento como  
servicio

info@flossystems.com

RAID

13 Niveles estándar de RAID

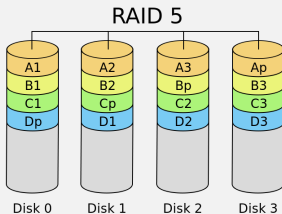
LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

- ▶ Volumen dividido con paridad distribuida: es el nivel estándar más completo de RAID. Dividiendo datos e información de paridad, crea una arquitectura redundante que al mismo tiempo mejora los tiempos de lectura/escritura.
- ▶ Tolerancia a fallos: 1 disco.
- ▶ RAID 5 requiere al menos 3 discos.





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

14 LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

# LVM

# Antes de los volúmenes lógicos



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

15 LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

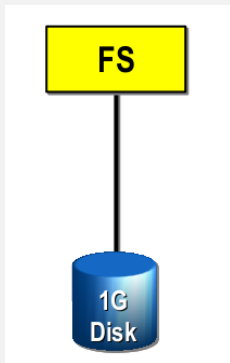


Figura: No era muy grande.

# Por qué existen volúmenes lógicos



StaaS:  
almacenamiento como  
servicio

info@flossystems.com

RAID

Niveles estándar de RAID

16 LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

Los usuarios precisaban más espacio, ancho de banda, fiabilidad y flexibilidad:

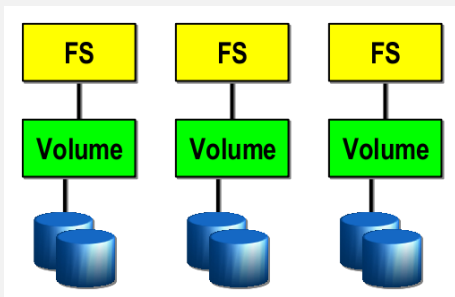


Figura: Fácil: inserta un “volumen” para juntar discos.





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

17

- ▶ LVM es un método genérico de asignar el espacio de almacenamiento.
- ▶ Más flexible que los esquemas de particionado convencionales.
- ▶ Existen distintas implementaciones:
  - ▶ Vinum (FreeBSD)
  - ▶ LVM (NetBSD)
  - ▶ LVM (Linux)
  - ▶ ZFS
  - ▶ LVM (AIX –1989–, HP-UX)
- ▶ LVM ayuda a los sysadmins a asignar eficientemente el espacio disponible en disco.
- ▶ LVM es una de las muchas formas de **virtualización del almacenamiento**.



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

18 Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

- ▶ LVM permite que el espacio sea dinámicamente asignado desde una partición grande a las particiones que van necesitándose.
- ▶ Permite concatenar, dividir, juntar o cualquier otra combinación entre particiones en una partición virtual mayor, que los sysadmins pueden cambiar el tamaño o mover.
- ▶ Idealmente sin interrupción del sistema.
- ▶ Desventajas:
  - ▶ puede complicar el arranque en un *disaster recovery*.
  - ▶ puede sufrir *fragmentación externa* a causa del FS subyacente, reduciendo el rendimiento de E/S.

# Ejemplo de LVM en Linux



StaaS:  
almacenamiento como  
servicio

info@flossystems.com

RAID

Niveles estándar de RAID

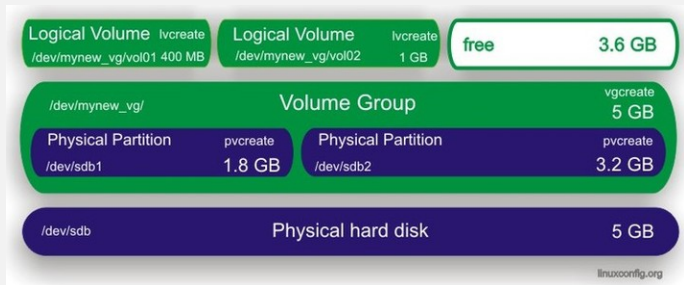
LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

ZFS

19





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

20

Ejemplo de LVM en Linux

ZFS

## Creación de volúmenes físicos

```
# pvcreate /dev/sdb1  
# pvcreate /dev/sdb2
```

## Creación del Virtual Group

```
# vgcreate mynew_vg /dev/sdb1 /dev/sdb2
```

## Añadir nuevos volúmenes físicos a un grupo virtual

```
# vgextend mynew_vg /dev/sdb3
```



## Creación de Volúmenes Lógicos

```
# lvcreate -L 400 -n vol01 mynew_vg  
# lvcreate -L 1000 -n vol02 mynew_vg
```

## Mostrar Grupos y Volúmenes Lógicos

```
# vgdisplay  
# lvdisplay
```

## Creación de un sistema de ficheros en volúmenes lógicos

```
# mkfs.ext3 /dev/mynew_vg/vol01  
# mount /dev/mynew_vg/vol01 /home/foobar
```



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

#### RAID

Niveles estándar de RAID

#### LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

# ZFS

22 ZFS





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

23 ZFS

- ▶ ZFS es un potente, escalable (128bit) y moderno sub-sistema de almacenamiento.
- ▶ Fiable, administración sencilla, integridad de datos y servicios integrados.
- ▶ ZFS combina los roles tradicionales de Volume Manager (RAID) y Sistema de Ficheros.
- ▶ La idea es que el disco debería ser algo similar a los módulos DIMM de RAM, conectar y usar.
- ▶ ZFS se lleva muy bien con SSD, y sabe cómo usarlo.

# ZFS: características únicas (1)



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

24

ZFS

- ▶ *Pool* de almacenamiento: elimina por completo el viejo concepto de volumen lógico como capa aparte.
- ▶ ZFS es Copy-on-Write transaccional: si múltiples procesos piden recursos iguales, se les devuelven punteros al mismo recurso.
- ▶ Siempre consistente (no necesidad de fsck)
- ▶ Integridad de datos: detecta y corrige silenciosamente corrupción de datos.
- ▶ Inmensa escalabilidad



## ZFS: características únicas (y 2)



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

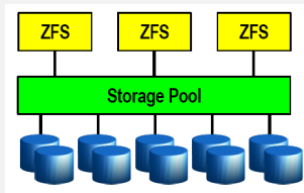
Ejemplo de LVM en Linux

25

ZFS

- ▶ Características avanzadas: snapshots, clones, rollbacks, deduplication, compresión, replicación, cifrado, compartición nativa vía nfs, cifs o iscsi...
- ▶ Administración simple: zfs y zpool.
- ▶ Estado del arte, marca el camino a los futuros FS (como btrfs)
- ▶ Limitaciones:
  - ▶ ZFS no es un FS de tipo cluster ni un sistema distribuido o paralelo.
  - ▶ Muy exigente en recursos.

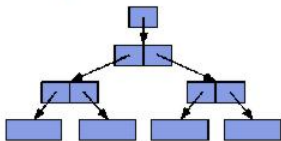
En el proceso de escritura (I/O), un bloque puede ser comprimido, cifrado, realizada la suma de comprobación y a continuación deduplicado, en ese orden.



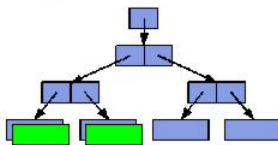
- ▶ Los sistemas de ficheros se crean sobre pools de almacenamiento virtual llamados *zpool*s.
- ▶ Un *zpool* se construye a partir de dispositivos virtuales (*vdevs*) desde dispositivos de bloques: ficheros, particiones de disco duro o discos enteros.

## Copy-On-Write Transactions

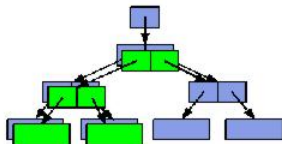
1. Initial block tree



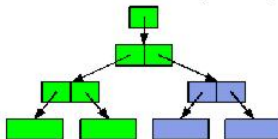
2. COW some blocks



3. COW indirect blocks



4. Rewrite uberblock (atomic)

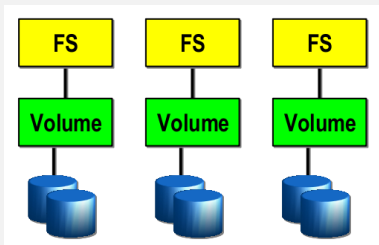


# Volumen/FS vs. pool de almacenamiento



## Volúmenes tradicionales:

- ▶ Abstracción: disco virtual
- ▶ Partición/volumen para cada FS
- ▶ Crece o se reduce manualmente
- ▶ Cada FS tiene ancho de banda limitado
- ▶ El almacenamiento se fragmenta



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

Gestión de volúmenes  
lógicos (LVM)

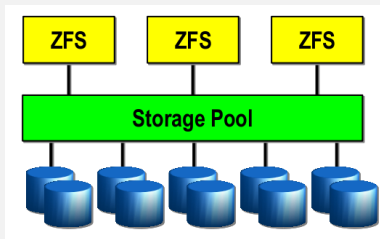
Ejemplo de LVM en Linux

28

ZFS

## Almacenamiento con ZFS pools:

- ▶ Abstracción: malloc/free
- ▶ No hay particiones que manejar
- ▶ Crece o se reduce automáticamente
- ▶ Todo el ancho de banda está siempre disponible
- ▶ Todo el almacenamiento en el pool es compartido



StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

RAID

Niveles estándar de RAID

LVM

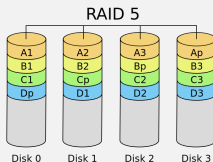
Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

29

ZFS

- ▶ RAID no estándar: específico de ZFS.
- ▶ Similar a RAID 5, pero evita el *write hole* de RAID 5 (si se produce un apagón durante una escritura, paridad o datos pueden quedar inconsistentes/corruptos).
- ▶ Existe también RAID Z2 que dobla (o triplica) la estructura de paridad alcanzando resultados similares a RAID 6.
- ▶ En Julio de 2009, se incorporó RAID Z de triple paridad a OpenSolaris.
- ▶ No precisa ningún hardware especial.





StaaS:  
almacenamiento como  
servicio

[info@flossystems.com](mailto:info@flossystems.com)

#### RAID

Niveles estándar de RAID

#### LVM

Gestión de volúmenes  
lógicos (LVM)

Ejemplo de LVM en Linux

## ► ZFS Demo

31 ZFS



FLOSSystems S.L.,  
Madrid, Spain

32

# StaaS: almacenamiento como servicio

info@flossystems.com



Murcia, 1-2 de junio de 2012