

XFS

Zbigniew Jarosik
zibi@nora.pl
<http://zibi.nora.pl/>

Zimowisko 2010

Szczęśliwość

- Po co nam w ogóle eFeSy?
 - Limity rozmiarów
 - Utrata spójności
 - Wielogodzinne fsck
 - Denne transfery
- Czego oczekujemy?
 - Dużo czasu na granie w kłejka
 - Gęste włosy w kolorze innym niż siwy
 - Czas dla rodziny
- Co mamy w menu?

Menu

ISO9660 minix qnx4
ZFS ext UDF HFS UFS
NTFS ext4 UFS vfat
JFS BeFS minix HPFS
fat32 ext3 cramfs
ext2 btrfs Reiser4
ReiserFS fat 9p XFS fat12

Menu

ISO9660 minix qnx4
ZFS ext UDF HFS UFS
NTFS ext4 UFS vfat
JFS BeFS minix HPFS
fat32 ext3 cramfs
ext2 btrfs Reiser4
ReiserFS fat 9p **XFS** fat12

O co tyle szumu?

- SGI 1994 (IRIX v5.3)
- 2000 GPL
- Kernel 2.4
- 2001 pierwsze wdrożenia dystrybucyjne

Trybiki

- FS size – 16EB
- File size – 8EB-1B
- 1 EB = 1,000,000,000,000,000,000 B = 10^{18}
- 1 EB = 1024^6 (ZFS – 1024^7)

- Linux 32bit – 16TB (fs i file)
- Brak wsparcia dla volume managera (XLV)
- Warstwa translacji

Diabeł

- Księgowanie
- Grupy alokacji
- Extent based allocation
- Rozmiar bloku 512B – 64KB
- Opóźniona alokacja + defrag online
- RAID-0
- Quota, QoS
- Atrybuty i ACL
- Online growfs
- Wsparcie dla snapshotów
- Natywne dump/restore
- DMAPI dla HSM

Buchalteria

- Księgowanie metadanych
- Max 64k bloków i max 128MB
- Możliwe wyprowadzenie żurnala na osobne urządzenie
- Automatyczne recovery przy mount – bardzo szybkie
- W razie czego - `xfs_repair`

Mała komanda po milion żołnierzy

- Grupy alokacji (wewnętrznie 32bit)
 - Wielowątkowość
 - Skalowalność
- Extent based allocation (nawet 4GB)
- Opóźniona alokacja
 - defrag online
- Rozmiar bloku – 512B – 64KB
- RAID-0

Dwie literki Q

- QoS
- Quota
 - user
 - group
 - project

Second life

- Extended attributes - drugi strumień danych
 - Baza metadanych o pliku
 - Notatki
 - Indeksy
 - Uprawnienia
- Klucz – wartość
 - Klucz – 256B
 - Wartość – 64KB
- Można podłączyć do dowolnego inode
 - xfsdump, xfsrestore
- ACL

Digging to the Core

- Online growfs
 - Współpraca z LVM
- Wsparcie dla snapshotów
 - Współpraca z LVM
- Natywne dump/restore
 - Inkrementacyjny
 - Per folder
 - Resume
- DMAPI dla HSM
 - Jako łąta

Rewers waleta

- File truncating
- Tempo xfs_repair
- Wolna praca przy modyfikacji złożonych strukturach katalogów
- Problemy z GRUB-em
- Brak wsparcia dla szyfrowania/kompresji
- Brak block journalingu
- No shrink ;/
- Brak undelete

I sam walet

- Bardzo szybki przy przetwarzaniu dużych plików
- Rozsądnie szybki przy małych plikach
- Obsługa dużych dysków i macierzy
- Brak limitu ilości plików
- ACL-e
- Dobra współpraca z LVM-em
- Snapshoty (na LVM-ie)

¿Preguntas?

Zbigniew Jarosik
zibi@nora.pl
<http://zibi.nora.pl/>

Zimowisko 2010