OpenSolaris Storage Platform

Storage Revolution in the Making

Mads Toftum

http://soulfood.dk
WHY?
3 versions

- Solaris 10

- Solaris Express Community Edition (SXCE)

- OpenSolaris 2008.5
  - the "artist" formerly known as Indiana
ZFS

"The Last Word in filesystems"
ZFS

- Data Integrity
- Easy Administration
- Easy Data Migration
- Snapshots
- Compression
- Encryption
# zpool create tank mirror c1t2d0 c2t2d0

# zfs create tank/home
# zfs set mountpoint=/export/home tank/home

# zfs set compression=on tank/home

# zfs set sharernfs=rw tank/home

# zfs create -o casesensitivity=mixed tank/home
# zfs set sharecifs=on tank/home
ZFS

- New developments
  - SSD

- Long term
  - de-duplication?
The last words about ZFS

almost...
Storage Area Network
Comstar

- SCSI target
- Multipathing across different transport protocols
- LUN Masking and mapping
- Qlogic 4G & 8G
- Emulex
Fibre Channel over Ethernet

- http://www.open-fcoe.org/
- Target
- Initiator
- 1, 10, ?
- Comstar
iSCSI

- Target
- Initiator

# zfs create -V 2g tank/iscsi/vol1
# zfs set shareiscsi=on tank/iscsi/vol1
# iscsitadm list target
Target: tank/iscsi/vol1
iSCSI Name:
iqn.1986-03.com.sun:02:7b4b02a6-3277-eb1b-e686-a24762c52a8c
Connections: 0
Networked File Systems
pNFS - parallel nfs

- NFS version 4.1
- Extension to NFS
- Data striped across multiple servers
- Separate data and metadata
QFS

- Quick File System
- A distributed file system that can be mounted on multiple host systems.
- striped individually or device groups
- tunable
- linux clients (redhat, suse)
- often used with SAM
CIFS

- Samba replacement
- fully integrated
  - kernel based protocol
  - credential (access token) support
  - sharemgr
  - case insensitive (optional)
- Security Descriptors (Super-set of ACLs)

```bash
zfs create -o casesensitivity=mixed -o nbmand=on -o sharesmb=on tank/cifstest
```
Archiving and replication
Network Data Management Protocol (NDMP)

- NDMPv4
- transport data between NAS devices, and backup devices
- zfs snapshot
AVS

- Availability Suite
- Remote Mirror Copy
- Point in Time Copy
- UFS, QFS, VxFS, and ZFS
- block level
- GEO Cluster
SAM

- Storage Archive Manager
- Hierarchical Storage Management
- Different speeds / prices for storage
- File Metadata stays
- Speeds recovery time
- Often used with QFS
ADM / DMM

- Automatic Data Migration
- Similar to SAM
- Hierachical Storage Management
- Lustre, QFS, ZFS

- Data Migration Manager
- move filesystems
- NFS
- CIFS
- immediately available on target system
Misc

- Honeycomb
- Project Celest
HoneyComb

- Fixed Content Storage
- Software
- Scalable, multi-cell based symmetric, cluster
- Hardware
  - Sun StorageTek 5800
Project Celeste

- High Availability Data Store
- Distributed
- Peer-to-peer
- Masterless
- Mutable
## Endorsed Projects & Communities

<table>
<thead>
<tr>
<th>Storage Networking</th>
<th>File Systems</th>
<th>Data Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Brocade FC HBA, Initiator &amp; Target</td>
<td>- CIFS client for Solaris</td>
<td>- Automatic Data Migration (ADM)</td>
</tr>
<tr>
<td>- COMSTAR (Common Multiprotocol SCSI Target)</td>
<td>- CIFS server</td>
<td>- Project Celeste</td>
</tr>
<tr>
<td>- Emulex Fibre Channel Device Driver</td>
<td>- DTrace Provider for NFSv4</td>
<td>- Common Array Manager</td>
</tr>
<tr>
<td>- FibreChannel over Ethernet</td>
<td>- Ext3 file system support</td>
<td>- Data Migration Manager</td>
</tr>
<tr>
<td>- Hitachi</td>
<td>- File System Driven Device Naming</td>
<td>- Device Mapper</td>
</tr>
<tr>
<td>- iSNS Server</td>
<td>- Fuse on Solaris</td>
<td>- Honeycomb fixed content data management</td>
</tr>
<tr>
<td>- Mega SAS drivers</td>
<td>- JFS: Journaled File System</td>
<td>- Media Management System</td>
</tr>
<tr>
<td>- Multi-pathing (MPxIO) software</td>
<td>- loki compression &amp; cryptography support</td>
<td>- Object Storage Device (OSD) support for Solaris</td>
</tr>
<tr>
<td>- NPORt ID Virtualization (NPIV)</td>
<td>- NFS Community</td>
<td>- OpenPegasus CIM Server</td>
</tr>
<tr>
<td>- Network Storage</td>
<td>- NFS RDMA transport update and performance analysis</td>
<td>- Solaris Volume Manager Community</td>
</tr>
<tr>
<td>- Object Storage Device (OSD) support for Solaris</td>
<td>- NFS Server in non-Global Zones</td>
<td>- Sun StorageTek Availability Suite</td>
</tr>
<tr>
<td>- Solaris iSCSI Target</td>
<td>- NFS version 4.1 dNFS</td>
<td>- Storage Archive Manager</td>
</tr>
<tr>
<td>- Solaris iSCSI Extensions for RDMA (iSER)</td>
<td>- NFSv4 namespace extensions</td>
<td>- VSCAN service</td>
</tr>
<tr>
<td></td>
<td>- Shared QFS file system</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Squashfs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- UFS Community</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- YANFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ZFS Community</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ZFS on disk encryption support</td>
<td></td>
</tr>
</tbody>
</table>
New Toys

www.sun.com/unifiedstorage
overview

- **7110**
  - 2u box with 2TB of storage
- **7210**
  - 4u box with 46TB of storage
- **7410**
  - 2u box with up to 288TB (576TB)
  - can be clustered
Features

- NFS v3 and v4
- CIFS
- iSCSI
- HTTP
- WebDAV
- FTP
Data Services

- RAID-Z (RAID-5 and RAID-6), Mirrored, and Striped disk configurations
- Unlimited Read-only and Read-write Snapshots, with Snapshot Schedules
- Built-in Data Compression
- Remote Replication of data for Disaster Recovery
- Active-Active Clustering (in the Sun Storage 7410) for High Availability
- Thin Provisioning of iSCSI LUNs
- Virus Scanning and Quarantine
- NDMP Backup and Restore
more info

OpenSolaris Storage Community
http://opensolaris.org/os/community/storage/

Danish OpenSolaris User Group
http://opensolaris.org/os/project/dosug/
Credits

Images and text from sun.com and opensolaris.org

http://flickr.com/photos/saschaaaa/

http://flickr.com/photos/dtype/

http://flickr.com/photos/zzathras777/