SunSoft
A Sun Microsystems Company

Large Files for Solaris: Something big

Rob Thurlow

thurlow@eng.sun.com

Large Files Program

- **Target: Large files in next major Solaris release**
  - **Had something in Solaris 2.5 timeframe, but it just wasn't ready for prime time**
  - **Large File Summit draft API specification almost complete, so compliance works**
  - **Want to present minimal surprises to users (converting many utilities)**
  - **No changes to UFS layout => 1 terabyte max filesize**
  - **VM & other support for larger file sizes when ready in the underlying filesystems**
- **64-bit process address space will come later**

Large File Support for NFS

- NFS Version 3 support, Version 2 safety
  - V3 will reflect server limitations only, to 2^63-1
  - V2 will fail politely when appropriate
  - NLM Version 4 will match NFS V3 limits
  - Most changes obvious:
    - offset, limit variables => 64 bits
    - awareness of 32-bit limits must be made more general
    - 64-bit variables must be mutex-protected
    - lots of data structures and interfaces got jostled
    - new checks for EOVERFLOW/EFBIG conditions
  - Some oddities: pathconf(fd, _PC_FILESIZEBITS) has to be derived from the FSINFO.maxfilesize value

Large File Summit API

- **Unix API changes for manipulating files larger than 2Gb on 32-bit operating systems; aimed at X/Open**
  - **never lie - no syscall will return an truncated value**
  - **open protection - naive programs will fail early**
  - **explicit *syscall*64() interfaces or compilation flags to access new types and interfaces with POSIX source**
  - **not all problems solvable, e.g. fd inheritance**
- **Vendor participants include Sun, IBM, HP, SGI, Convex, Sequent, Tandem, Pyramid, Novell, Amdahl, Unisys, HaL**
- **ISV participants include MSC, SAS, Oracle, Sybase**
- **See http://www.sas.com/standards/large.file/index.html**

Large File Status

- **Working at Connectathon! Please test against us!**
- **Server supports I/O and locking of sparse files to 2^40 - 1 (you filla our disk, I breaka you fingers :-)**
- **Machine name is ss-nfs03, billboard IDs are:**
  - **sun-lf-nfs-v2-udp**
  - **sun-lf-nfs-v2-tcp**
  - **sun-lf-nfs-v3-udp**
  - **sun-lf-nfs-v3-tcp**
- **Client should support up to 2^63 - 1, if we can find a server**
- **We'd like to test against all V3 servers; please let us know your offset limits**