Beyond Disk Imaging for Preserving User State in Network Testbeds

Jelena Mirkovic, Abdulla Alwabel and Ted Faber
USC Information Sciences Institute
Disk Images In Testbeds

- Testbed experimentation takes a long time
  - Setup, running, analysis
  - Weeks to months
- Testbeds encourage release of resources if they are not actively used (idle swap, emails)
- Users need a way to preserve state
  - OS configuration
  - Applications installed and data
⇒ Solution: disk images
  - Base and custom images
What’s the Downside?

- Images use a lot of space
  - 340 GB on Deterlab = 1/3 of shared storage
- Older images are not updated
  - Vulnerabilities
  - Swap-in problems
- Older images reduce experiment’s chance of swap in
  - Not supported on new hardware – testbeds and their user base grow
  - Users do not understand this limitation
Image Usage on DeterLab

- Mined from DB/disk and usage logs
- 1,128 images
  - 645 in DB/disk (22% never used)
  - 483 used and deleted
- 47% projects and 14.4% allocations use custom images
  - 48% of research projects = 735 images/77 projects
  - 44% of class projects = 124 images/19 projects
  - 47% allocations use images older than 1 year
⇒ Disk images are popular and numerous
Custom Image Use/Support

Year

Images used

< 50%
50-80%
80-90%
> 90%
 Alternatives to Disk Imaging

- Goals:
  - reduce storage requirements
  - facilitate easier update/upgrade
- Catalogue and reapply image customizations at file level
  - DiffBase
  - DiffCustom
  - AppStore
- Help users script/reapply installations + store large data locally (dedicated disk/network)
OS Loading Process

- MFS load
- Frisbee OS image
  - $f(\text{image\_size})$
- CURRENT
  - Boot custom OS
    - $f(\text{image\_complexity})$
- PROPOSED
  - Boot base OS
    - $f(\text{image\_complexity})$
    - Restore files / packages
      - $f(\text{state\_size})$
      - Reboot
        - $f(\text{image\_complexity})$

T1

T2
DiffBase

- Compare custom image with base during creation, file by file
  - Store only differences
  - Added/modified files, list of deleted files, catalogue of sparse files and their contents
- Saves 75% disk space
- Larger images load faster!
- Helps a little with OS update/upgrade
DiffBase – Storage

![Bar chart showing the percent of an image that must be saved. The x-axis represents the percent range (0-90), and the y-axis represents the image count. The bars indicate the frequency of images in each percent range.]
DiffBase – OS load
- Compare custom images within one project with each other
  - Store one using DiffBase
  - For others store file differences between each and that first custom image
- Saves 78% disk space
- Helps as much as DiffBase with OS update/upgrade
DiffCustom – Storage

![Histogram showing the percent of an image that must be saved](image)
Run DiffBase approach
Automatically compare files identified as added/modified with application packages
For matches:
- Remove them from DiffBase state
- Store configuration changes and package name
Saves 81% disk space
Facilitates easy OS update/upgrade
App Store – Storage

Image count

Percent of a DiffBase state that must be saved

Ubuntu 8.04
Ubuntu 10.04
Testbed users need a way to save state
  - Disk images are not a good option
Our proposed alternatives save significant amount of storage
  - But do little to help users update/upgrade images
Future research
  - Help users catalogue changes they made to a base image and their purpose
  - Automate migration to other images / upgrade
Contact

- Jelena Mirkovic
  - sunshine@isi.edu
- Abdulla Alwabel
  - alwabel@usc.edu
- Ted Faber
  - faber@isi.edu