

OOD/SAA Google Shared Drive Migration Project

Kick off February 2019

2019 OOD/SAA Servers and file management

- Multiple 2008 File servers
- Windows 2008 Clustered architecture
- Multiple ESX servers
- Multiple Dell EqualLogic SANs
- Complex networking setup
- Workgroup manager integration for user level permissions
- Offsite backups in Livermore DC with manual failover

Proposed migration to Google file management

- Migration to Google Shared Drives(formerly Google Team Drives)
- File access via Browsers or Google Drive File stream as well as mobile
- File access control via Workgroup manager(or individual access)
- Cloud backups to AWS S3

Challenges

- SAA - approximately 20 departments (200 people)
- OOD - approximately 35 departments (400 people)
- Approximately 20Tb, 6+ million files, 90+ shares
- Complex directory structures and naming conventions
- Every document is important
- “Our files are not compatible with Google docs”
- “We've Always Done it This Way” so why move
- Lift and shift issues (Google restrictions)
- No existing Google official migration tools

Plan

- Executive buy in and sponsorship (Barbara Pugliese and Katherine Kassaras)
- Appointment of project manager (Lanelle Neumann)
- Google Drive FAQs (separate document available)
- High level plan creation
- Executive presentation
- Communication plan
- Timeline
- Test Environments (Minimal set of files/directories)
- Test Migrations (Continually)
- Final Migrations (Thursdays)
- Post Migration support(TST, SNOW, Email, Slack)

Executive overview agenda

- Why Google Shared Drive
- Google Shared Drive vs Current Shared Drive
- Google Shared Drive Migration Workflow
- Timeline
- Coverage for all of SAA/OOD
 - Team Lead roles and responsibilities
 - Team Lead requirements
 - Team Lead checklist
- Migration Assumptions
- Next Steps
 - Deep dive with current shared drive
 - Test environment
 - Test migration
 - Clean-Up
- Google Shared drive demo

Why Google Shared Drive

- OOD Shared Drive will no longer exist by the end of 2019
- Google provides a secure infrastructure for file storage and aligns with the University's recommended architecture as well as campus IT Plan
- Minimizes ongoing maintenance expenses/overhead incurred with shared drives
- Files remain with the team even if an individual leaves
- Files are searchable via Web Browser and Explorer/Finder
- Non-Google files such as Excel spreadsheets, can be used in Google Shared Drive using File Stream extension(or "opened directly")

Google Shared Drive vs Current Shared Drive

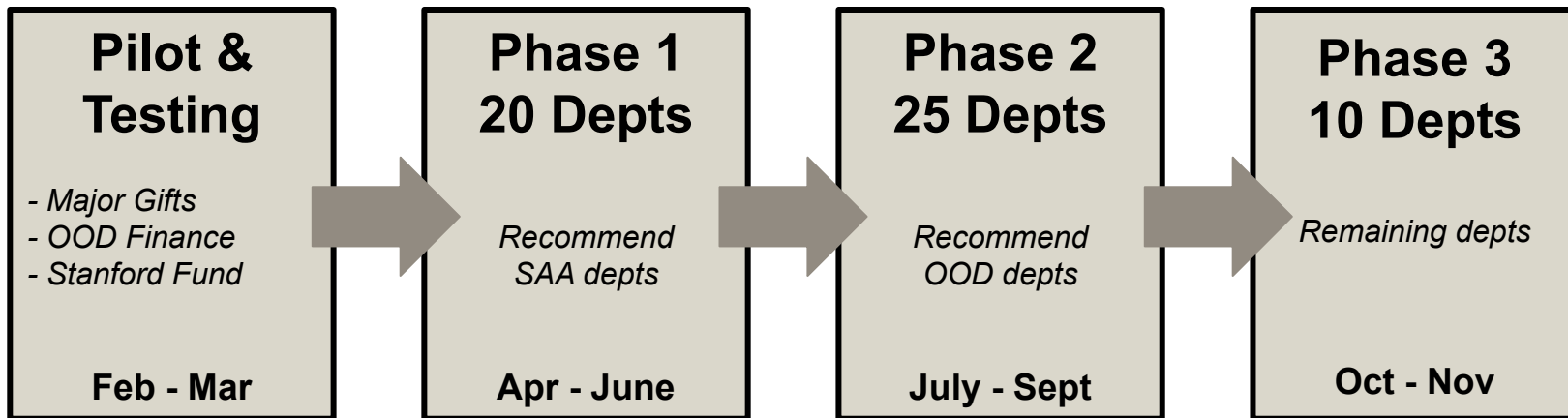
- Access is at the Google Shared Drive level or at the file level
 - Not the same as current shared drive which allows different access levels for folders
- Google Shared Drive has a 400k file limit and has a 20 folder deep limit
 - No limit for current shared drive
- Google Shared Drive will be integrated with workgroup manager to administer access
 - Same workflow with managing access to current shared drive
- Integration with workgroup manager means all members of Shared Drive will have content manager access (add, edit, delete)
 - Same access level as Shared Drive
- Google Shared Drive from the migration project will include the extended back-up by UIT
 - Same back-up as current shared drive
- Any new Google Shared Drives post Shared Drive migration will be determined by the department whether or not it will be integrated with workgroup manager
 - Sharing high risk data with external Stanford people cannot have workgroup manager integration

Team Drive Leads Roles and Responsibilities

- Represent their department and identify any specific use cases and areas of concerns
- Provide “informal” support and be the initial point person for any questions from their department
- Review As-Is Shared Drive
 - Define To-Be Team Drive settings and structure
- Attend meetings and distribute communication to their organization
- Obtain milestone approvals for their departments (i.e. Team Drive structure, timeline, migration date, etc.)

Timeline

- SAA - approximately 20 departments
- OOD - approximately 35 departments



SAA Timeline

Plan & Define

03/04/2019 - 05/31/2019

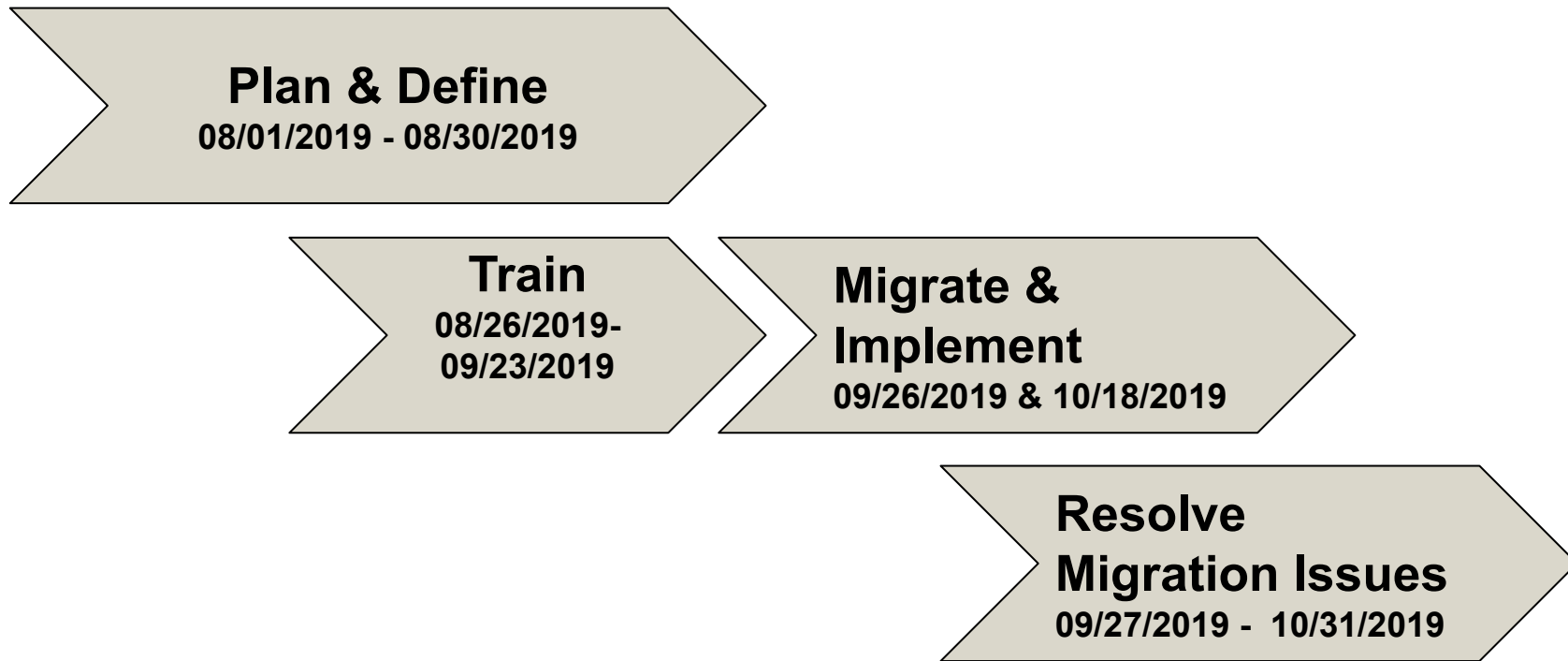
Train

05/06/2019-
05/31/2019

**Migrate &
Implement**

06/10/2019 - 06/28/2019

OOD Timeline - Phase 2



Team Drive Migration Workflow



1. Document As-Is Shared Drives

UIT:

- Identify current shared drive structure and nested folders
- Identify number of files in each shared drive (>400)
- Identify current admins and members of shared drive

PM:

- Review current use cases with depts

2. Review As-Is Shared Drives

PM:

- Review As-Is w/ OOD/SAA
- Propose Team Drive structure based on number of files, nested folders, team members and others

OOD & SAA Depts:

- Determine if structure is the same as As-Is or need to change; complete any data-clean-up

3. Provide Initial Review of Team Drive

PM:

- Provide initial review/training of Team Drive functionality to help depts understand how to best structure their Team Drives

OOD & SAA Depts:

- Access Team Drive Test Environment

4. Define & Approve New Team Drives

OOD & SAA Depts:

- Based on As-Is and Team Drive functionality finalize and approve Team Drive structure
- Complete data clean-up

5. Train OOD & SAA Depts

OOD & SAA Depts:

- Train the trainer (Team Leads trained as super users)
- Team Drive training available on Lessonly

6. Set-Up Team Drive & Migrate Data

UIT:

- Set-up approved Team Drive structure
- Migrate shared drive files/data to Team Drive
- Back-up

7. Use Team Drives

OOD & SAA Depts:

- Use Team Drive to add, edit, move, delete files and folders
- Dept super users to provide initial support

UIT:

- Provide initial and ongoing support

Team Lead Checklist

TASK	Start Date	End Date	STATUS
Identify any concerns with using Team Drive	03/04/2019	04/12/2019	IP
Review current shared drive	03/11/2019	04/12/2019	Complete
Review and approve proposed team drive settings and structure	03/18/2019	04/26/2019	
Test Google team drive in test environment	04/01/2019	04/26/2019	
Obtain approval of migration date	04/01/2019	04/26/2019	
Data clean-up	04/01/2019	05/31/2019	
Training for all	05/06/2019	05/31/2019	
Spot check migrated files	06/13/2019	06/28/2019	
Provide initial “informal” support for your dept(s)	03/04/2019	06/28/2019	
Provide project update communication for your dept(s)	03/04/2019	06/28/2019	

Team Lead Requirements

Complexity	Estimated Hours per week	Team Leads
(H)igh - Complex structure with greater than 10 unique access in nested folders, >5 Shared Folders, >400K files	5-8 hrs/week	Ross and David
(M)edium - Medium complex structure with less than 10 unique access in nested folders, 3-5 Shared Folders	3-5 hrs/week	Anna-Alycia
(L)ow - Simple structure with no unique access in nested folders, < 3 Shared Folders	1-2 hrs/week	Michelle, Katy/Jeremy

Tools we used

- Rclone
 - Free on Github
 - Written in Go
 - Supports most cloud providers
 - Large number of command line switches
- Treecize (Free)
- Google Drive File Stream(deployed via BF)
- Workgroup Manager
- CloudHQ

The Good, the Bad and the Ugly

- Prep work
 - Perform cleanup and deletions prior to migrations
 - Establish firm freeze dates for changes to file/folder structure(sync issues)
 - Mapping of existing data to new Google Shared Drive(s)
 - Mapping workgroups/users and access levels
 - On migration day disconnect and remove users' mapped drives
 - Address Workgroup manager and Google integration issues

The Good, the Bad and the Ugly (part 2)

- Google drive
 - Naming restrictions (combined length of file and folder names)
 - 400K objects and 20 nested folders
 - Deleted items count towards 400K limit
 - Do not delete - move (clean up after)
 - Deleted files are only recoverable for 30 days
 - Content Manager vs Manager access(mv vs cp)
 - Google file stream and BF install signin missed
 - Files uploaded via G Drive browser upload tool change time stamps
 - In Google shared drive you can't share folders
 - Default GSD search is content(vs file/folder name and no wild card)
 - Sunet ID vs alias email sharing

The Good, the Bad and the Ugly (part 3)

- Microsoft Office considerations
 - Google drive limited Ms Office documents sharing support(notifications and errors for multiple people)
 - File locking issues and sharing(3rd party)
 - Temp files visible (~.doc)
 - Consider Google drive levels of access(Content Manager vs others)
 - Ms Access issues (corruptions due to file locks)
 - Excel links to other documents will be broken
 - Opening Excel documents and saving in Gsheet may remove macros/comments

The Good, the Bad and the Ugly (part 4)

- Rclone
 - Source/destination prep (in file)
 - Dedupe function(Google allows for identical file names)
 - Rclone pair scripting
 - Daily upload limit is 750Gb(hard set per user account)
 - Must setup GCP key/token pair(rclone and Google API limits)
 - Retains file creation/modification time stamps(not folders)
 - Use rclone to delete files post migration
 - We used 3rd party tool to compare objects(Tree Size and GFS)

The Good, the Bad and the Ugly (part 5)

- CloudHQ
 - Can be used as Cloud to Cloud migration tool
 - Inexpensive
 - Will create buckets/folders
 - Archives deleted files
 - Supports Glacier restores
 - Poor logging
 - Poor support

In conclusion

- Many ways to improve and automate migration process
- Rclone may not be the only choice
- Google limits will improve

Thank you

- OOD/SAA Executive cabinet
- Lanelle Neumann (PM)
- OOD/SAA Team Leads
- Jerry Nagano(Technical implementation)
- TST support team of OOD/SAA
- Marcello Golfieri

Q & A