Managing Data from UASs

Presenter: Connar Stone
A Little Bit About Me

- Graduate of EOU (2016) in the field of Computer Science
- Current IT Manager/Drone Program Manager for Grande Ronde Model Watershed (GRMW)
- Other Hats: Website, Database, Network, & Hardware Management
- Father and Husband
- Novice Magician (No, I don’t perform at birthday parties)
Checklist
Ready to Fly
On Site

5 HQ photos
4000 videos
30 Gigabytes
Back at the Office

Disk Management

There is not enough space available on the disk(s) to complete this operation.

OK
Have No Fear!
Solutions

[LOCAL] You could try to delete files from your machine to make room (but what happens next time) YOU CAN’T KEEP DOING THIS

[LOCAL] You could get a large external drive (Like other local storage solutions this is not protected against flood, fire, or natural disaster) EVEN THESE RUN OUT OF SPACE

[NETWORK] You could create some kind of Network Attached Storage (NAS), or perhaps a large server (Again, perhaps not protected against flood, fire, natural disaster, or SECURITY BREACH, but maybe solves the space issue).

[CLOUD] You could use something like Dropbox or Google Drive (UL/DL speeds could prove problematic)

[HYBRID] You could build your own server and use something like Dropbox or Google Drive as a backup service. THIS IS EFFECTIVELY WHAT WE DO AT GRMW.

There may be a number of other solutions out there. These are just a few.
2020-2021 YEAR IN REVIEW

UAV METRICS:

• 42 Missions Flown
  (46 missions in 2019-2020)

• 409GB Raw Imagery Collected
  (178GB in 2019-2020)

• Added 29.5 Flight Hours (401 miles)
  (28 flight hours & 361 miles in 2019-2020)
GRMW’s Network Server Solution

Office

Internal Sharing

External Sharing

Samba (FREE)

Continuous Backups

Owncloud (FTP Server) (FREE)

Sync

Dropbox

20TB Raid0

(Running Ubuntu OS)

20TB Remaining Free Space

(~200,000 Photos)

409 GB Collected 2020

35+ Years of Collection Left
ThinkSystem SR550 Rack Server

Affordable, all-purpose rack server for local/remote sites

- Versatile 2U rack design
- Flexible storage configurations
- SW and HW RAID options
- Enterprise-class RAS features
- XClarity HW/SW/FW management suite
- Centralized, automated management

Starting at

$1,832.02
A Simpler Solution

*Please note that some cloud services do NOT support external drives for syncing*

16TB Internal HDD $290
Newegg.com!
A Word on Cloud Solutions

- Be careful which cloud backup solution you go with
- Some cloud backups charge network download fees (by the GB)
- Some limit your total storage
- Some only work for specific Operating Systems
- Some have TERRIBLE download speeds (weeks to download your data)
- All typically have some kind of subscription cost (monthly or yearly), which varies dramatically from service to service
- Some even have limits on the size of singular files!
- BE PICKY! There are many cloud services out there for backing up your data, make sure you pick the right one for you.
What to Store?

- Types of Media
  - Textual — TXT, CSV, DB (kilobytes) [Raw]
  - Photos — JPG, PNG, TIFF, GIF (megabytes) [Raw]
  - Videos — MP4, AVI, WAV, MOV, WMV (megabytes to gigabytes) [Raw]
  - Ortho — TIFF, JPEG, Meta Raster (gigabytes) [Post-Processed] (What is an Ortho?)

- Storing Raw Data is usually the best because you can always regenerate post-processed data if needed—however, the time it takes to process something must also be considered.

- You might ask when you might need a specific data set again, if the answer is “probably not for a while” and it is composed of large files that can easily/quickly be regenerated using your raw data sets, then consider just storing the raw data and regenerating the larger sets later if needed.

- How much space do you actually need? The more the better, but cost may limit your capacity.
Working With UAS Data Sets

- Picking the right software for the job (Global Mapper [Blue Marble], ArcGIS, Pix4D, Drone Deploy, DJI GS Pro, Photoshop...)

Ground Station  Processing  Post-Processing

LITOH | Pix4D | Global Mapper

DroneDeploy | ArcGIS

GSP
Ways to Reduce the Size of Data Sets

- Compression (Global Mapper .GMG files [Global Mapper Grid], JPEG vs. PNG) — BE AWARE OF THE DIFFERENCE BETWEEN LOSSY AND LOSSLESS COMPRESSION. JPEG IS LOSSY.
- Discard Unneeded files (cleaning datasets before storage)
- Archival of Old Files to off-system storage
- Remove Redundancies (multiple backups)

Types of Backups

- Full vs. Differential vs. Incremental Backups
  - Full – Most basic; all data is sent to another location (daily, weekly, monthly)
  - Differential – Makes a full backup of each change (daily, weekly, monthly)
  - Incremental – Only backs up the most recently changed data (continuous—most common for cloud services)
Organizing Your Data

- Databases (Spreadsheets, MySQL, Access, Indexing)
- File Structures (UAS -> Flights -> 2019 -> 2019-5-23 Site 1)
- Naming Conventions (Date vs. Location)
- Searchability — If I was looking for a specific flight in my dataset, how long would it take me to find it?

<table>
<thead>
<tr>
<th>Date</th>
<th>Flight Details</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-5-26</td>
<td>Chicken Creek (RGB + Multispectral)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-16</td>
<td>HUGR (RGB + Multispectral, Full Rich)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-20</td>
<td>Longley Meadows (RGB + Multispectral)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-20</td>
<td>Bird Trap Tracks (RGB + Multispectral)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-11</td>
<td>Mine Tailing (RGB + Multispectral)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Whiskey Creek (RGB + Multispectral)</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Spring Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>CCYI Hoffler</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Bird Trap Spring</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wonder</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wolfe Wetland</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wallowa Wilson</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wallowa McCharen</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wallowa Lettrop</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Wallowa Baker</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Tamahke</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Lookie Town Project</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-17</td>
<td>Dry Creek Javitz</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-3</td>
<td>Southern Cross</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-11</td>
<td>Sheep Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-11</td>
<td>East Sheep Meadow</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>HUGR</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Hall Ranch</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Windrow</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Sheep Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-15</td>
<td>Mine Tailing</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-15</td>
<td>Limbo Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-15</td>
<td>Chicken Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-15</td>
<td>Indian Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>UHR Upper Fly Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>UHR Lower Fly Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>UHR Eager</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>McCoy Meadows</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Whiskey Creek</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Longley Meadows</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-12</td>
<td>Little Creek Owensome</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-19</td>
<td>Wallowa Wilson-Hamlet</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-19</td>
<td>Wallowa 39 Loop Rd</td>
<td>Filedorf</td>
</tr>
<tr>
<td>2019-5-19</td>
<td>Woodlce (Multispectral)</td>
<td>Filedorf</td>
</tr>
</tbody>
</table>
Speeding Up Transfer Rates

◊ Ensure high quality Ethernet connections to your local and wide area networks (Cat5e or better for modern connections, Cat6 is ideal) — if using wifi, check your router and adapter.

◊ USB 3.0 Connections for any external drives (Sata3 or better for Internal Drives)
◊ Using an SSD over an HDD where possible (space/price tradeoff)
◊ If using cloud services, consider upgrading ISP connection for faster upload/retrieval
◊ Try a speed tester such as speedtest.net to test your connection (<30mbps—Bad, <100mbps—Okay, >100mbps—Good, >500mbps—Excellent)
Plan Ahead

Know how long your solution is going to last—All storage solutions eventually have to be changed, adjusted, upgraded, or even removed in favor of something new.

Storage is finite—it is better to have too much than too little, but there is a line where too much space becomes redundant to your solution. Plan out how much you need by how much you use, how much more you might use, and how long you plan to use it for.

Accidents happen—Having a local backup as well as a cloud backup may save your data and at the very least, it may save you a lot of time. Consider a natural disaster scenario vs. a failed drive scenario. Which is more likely to occur? Which is more devastating?

Keep it flexible—All-in-one solutions like Google Drive can be convenient, but they can also be very rigid in that there is really only one way to access your data, through the internet. If you suffer an outage or must retrieve large files and you have a poor connection, you may be waiting a long time to access your data. There are also some organizations that cannot work with Google Drive due to network restrictions. Having options for getting data to someone who might need it is a big part of managing your data.
Recap/Conclusion

- Practical Solutions for dealing with large UAS datasets
- Specific types of software for working with UAS data
- Cloud Solutions and Media Types
- Methods for reducing total storage
- Organization
- Transfer Speeds
- Planning for Future Storage Needs
Questions?
Helpful Links & Resources

- https://ubuntu.com/download/server
- https://www.samba.org/samba/download/
- https://owncloud.com/download-server/
- https://www.newegg.com/
- https://www.speedtest.net/
- https://www.bluemarblegeo.com/global-mapper-pro/
- https://www.arcgis.com/index.html
- https://flylitchi.com/