SciLifeLab
UPPSALA

Practical Knowledge

Transfering, moving and exporting files

160318
Martin Dahlö
martin.dahlo@scilifelab.uu.se

Enabler for Life Sciences

SciLifeLab UPPSALA

UPPMAX

- Moving files to and from UPPMAX
- Moving files within UPPMAX
- Compressing files
- Web export

SciLifeLab UPPSALA

To UPPMAX

- SFTP
- scp
- rsync

SciLifeLab UPPSALA

SFTP

- SFTP – SSH File Transfer Protocol
- Graphical clients
  - Filezilla
  - WinSCP
  - Transmit
  - ...

SciLifeLab UPPSALA

scp

- Secure copy
  - Widespread, not resumable
- scp [options] <source> <destination>

scp

- Secure copy
  - Widespread, not resumable
- scp [options] <source> <destination>

scp /local/path/file username@remoteurl:/remote/path/

scp /home/user/file.txt user01@milou.uppmax.uu.se:/home/user01/

SciLifeLab UPPSALA

scp

- Secure copy
  - Widespread, not resumable
- scp [options] <source> <destination>

  scp /local/path/file username@remoteurl:/remote/path/

  scp /home/user/file.txt user01@milou.uppmax.uu.se:/home/user01/
- Or the other direction..

  scp -r user01@milou.uppmax.uu.se:/home/user01/results/ /home/user/

SciLifeLab UPPSALA

rsync

- rsync [options] <souce> <destination>
  - Needs a server, resumable, checksums

SciLifeLab UPPSALA

rsync

- rsync [options] <souce> <destination>
  - Needs a server, resumable, checksums
- -tP
  - Preserve timestamps, show progress

SciLifeLab UPPSALA

rsync

- rsync [options] <souce> <destination>
  - Needs a server, resumable, checksums
- -tP
  - Preserve timestamps, show progress
- -rtP
  - Recursive, preserve timestamps, show progress

SciLifeLab UPPSALA

rsync

- rsync [options] <souce> <destination>
  - Needs a server, resumable, checksums
- -tP
  - Preserve timestamps, show progress
- -rtP
  - Recursive, preserve timestamps, show progress
- -nrtP
  - Dry run of -rtP

rsync

- rsync [options] <souce> <destination>
- -tP
  - Preserve timestamps, show progress
- -rtP
  - Recursive, preserve timestamps, show progress
- -nrtP
  - Dry run of -rtP

File

rsync -tP /home/user/file.txt user01@milou.uppmax.uu.se:/home/user01/

Folder

rsync -rtP /home/user/datadir/ user01@milou.uppmax.uu.se:/home/user01/

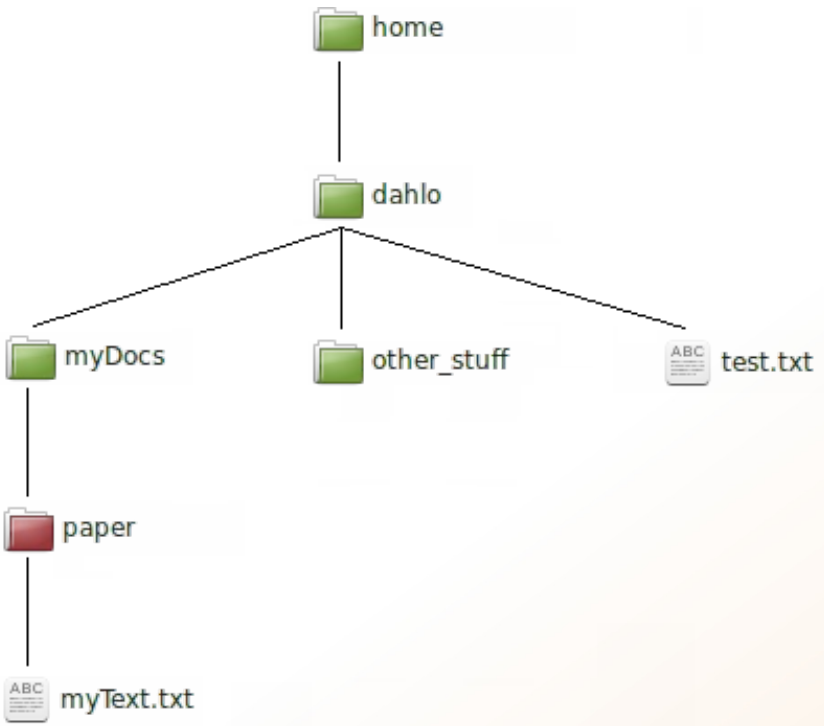
Within UPPMAX

- cp
- mv
- rsync

cp

- Copy a file

  cp <name of original> <name of copy>

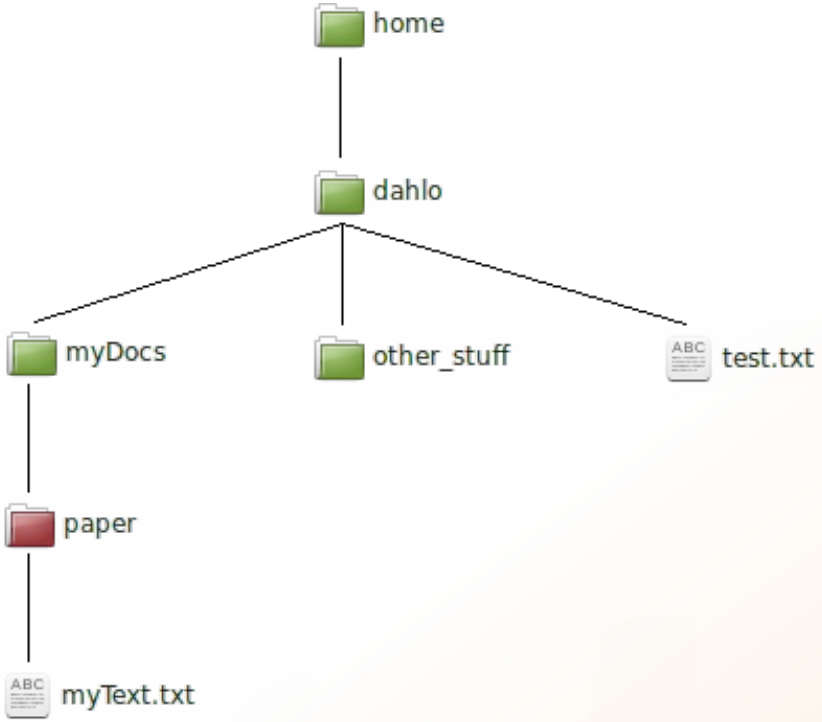


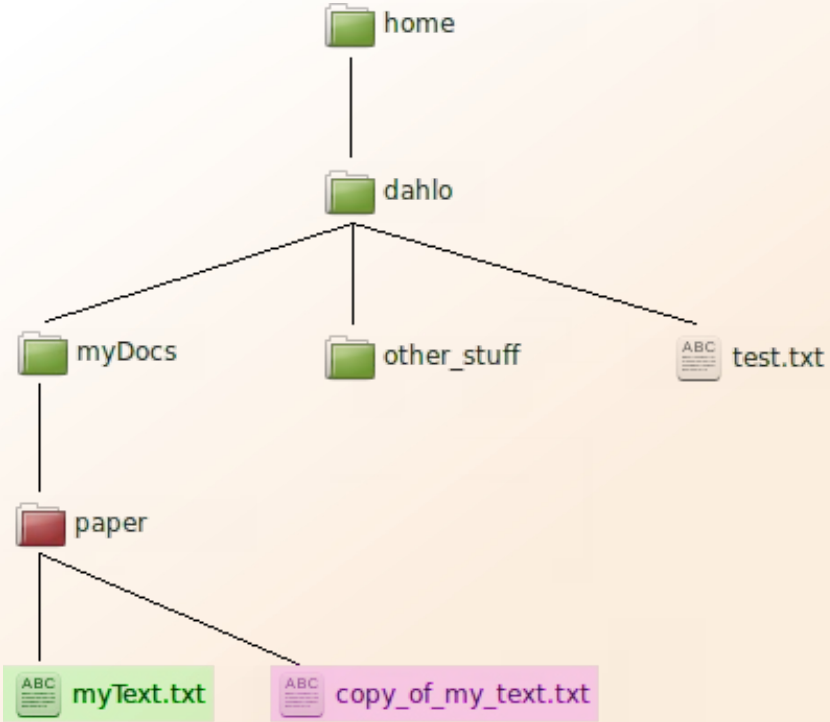
cp

- Copy a file

  cp <name of original> <name of copy>

  cp myText.txt copy_of_my_text.txt



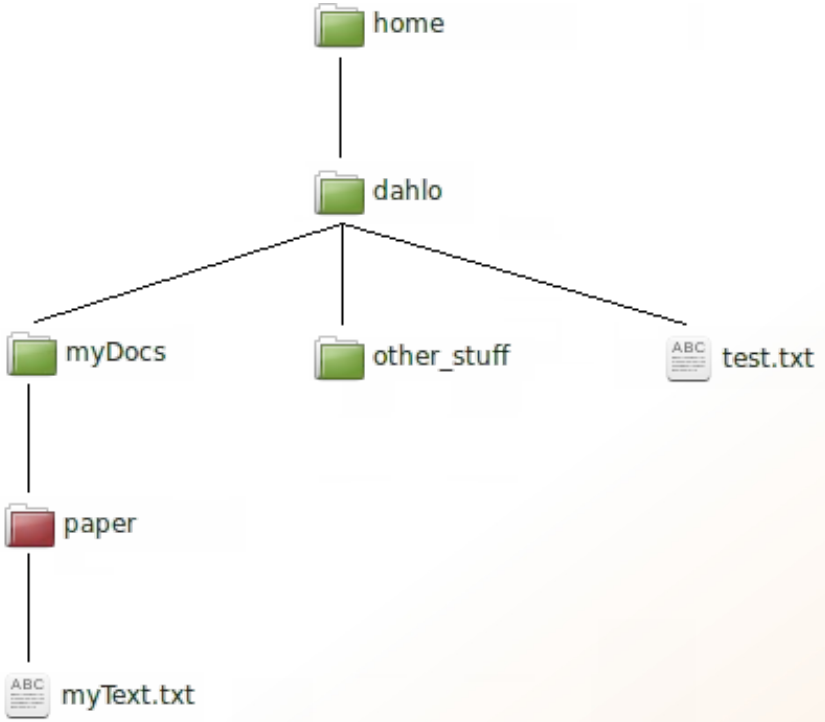


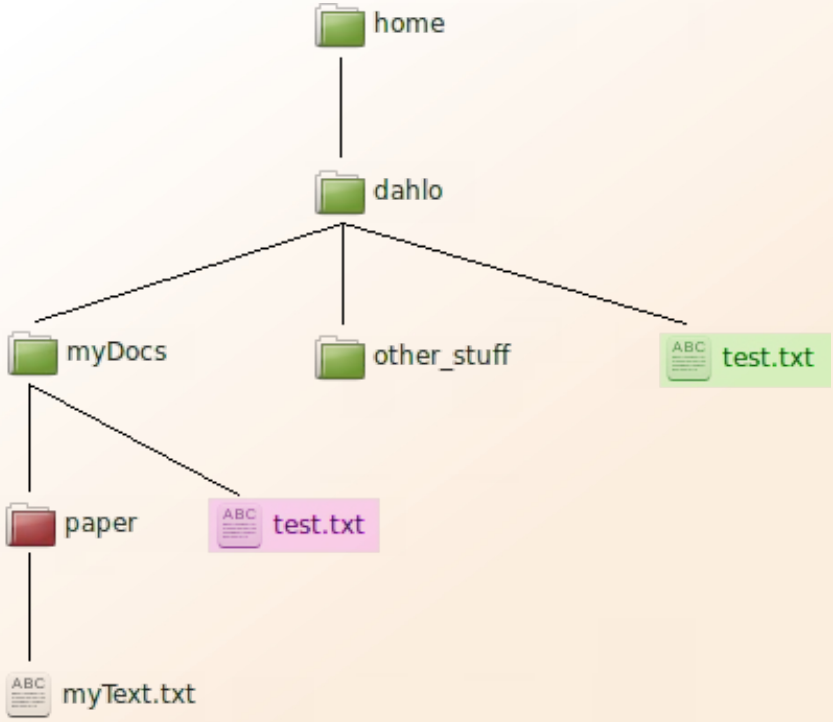
cp

- Copy a file

  cp <name of original> <name of copy>

  cp /home/dahlo/test.txt ../../myDocs/



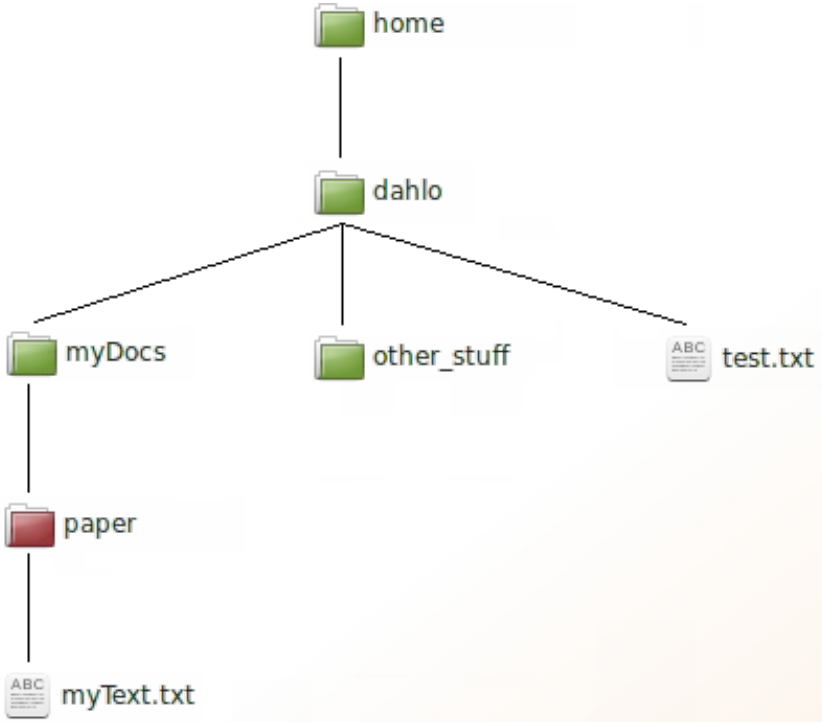


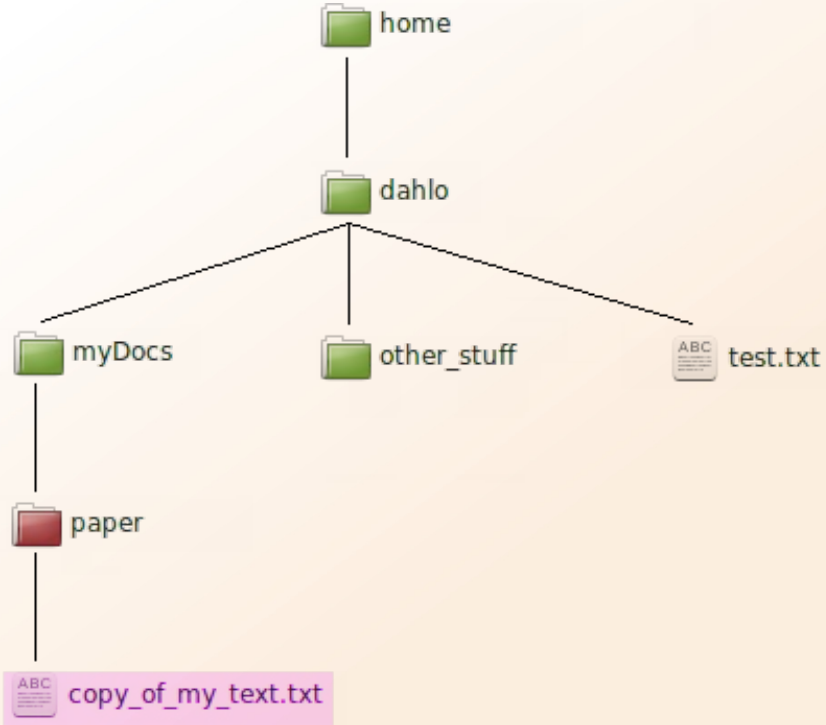
mv

- Move a file

  mv <name of original> <name of copy>

  mv myText.txt copy_of_my_text.txt



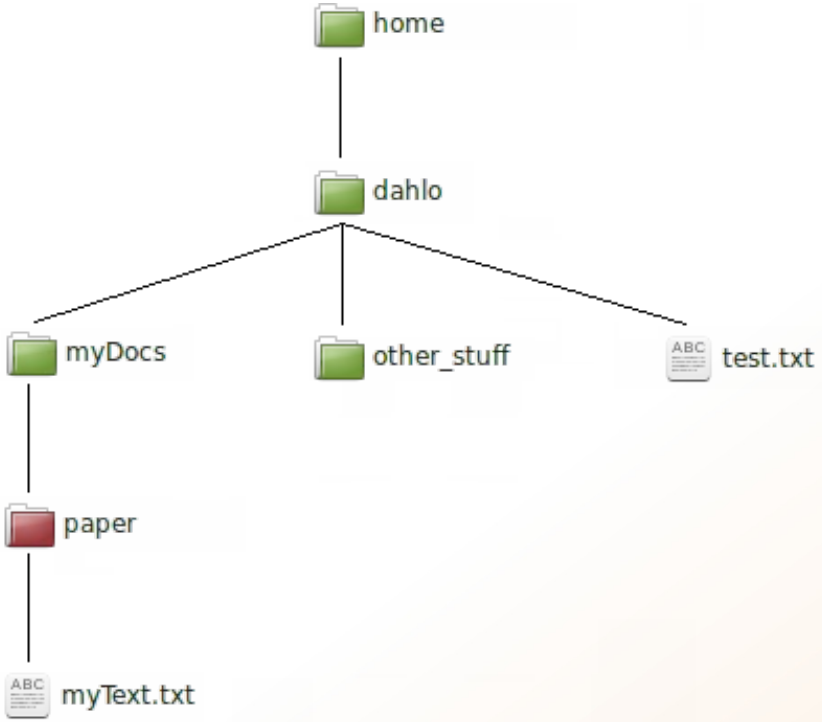


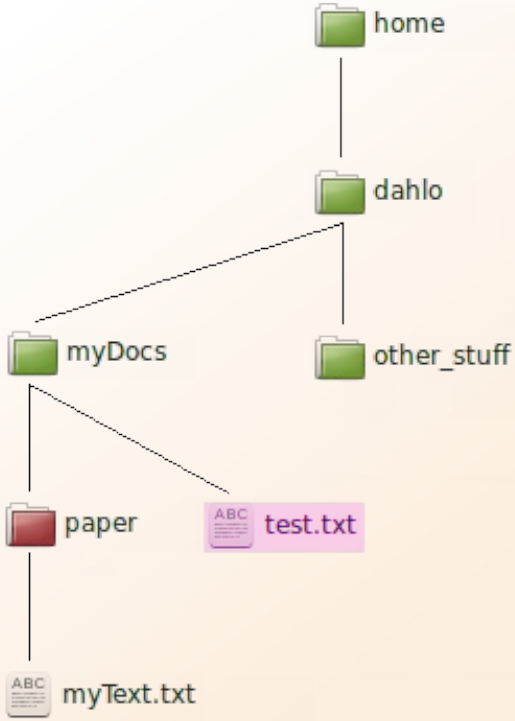
mv

- Move a file

  mv <name of original> <name of copy>

  mv /home/dahlo/test.txt ../../myDocs/





rsync

- rsync [options] <souce> <destination>
- -tP
  - Preserve timestamps, show progress
- -rtP
  - Recursive, preserve timestamps, show progress
- -nrtP
  - Dry run of -rtP

File

rsync -tP /home/user01/file.txt /proj/b2016999/

Folder

rsync -rtP /proj/b2010000/datadir/ /proj/b2016999/private/

SciLifeLab UPPSALA

File Compression

- Many different formats
  - tar?
  - gzip
  - bzip2

SciLifeLab UPPSALA

File Compression

- Many different formats
  - tar?
    - No compression (can be combined with gzip)
    - Many files → single file
  - gzip
  - bzip2

SciLifeLab UPPSALA

File Compression

- Many different formats
  - tar?
    - No compression (can be combined with gzip)
    - Many files → single file
  - gzip
    - Fast
    - Not ”blocked”
  - bzip2

SciLifeLab UPPSALA

File Compression

- Many different formats
  - tar?
    - No compression (can be combined with gzip)
    - Many files → single file
  - gzip
    - Fast
    - Not ”blocked”
  - bzip2
    - Slow (~6x), but can be parallellized with **pbzip2**
    - Blocked
    - Compresses slightly more than gzip

SciLifeLab UPPSALA

File Compression

- Many different formats
  - tar?

    tar -czvf archive.tar.gz /path/to/files/

    (**C**reate, **Z**ipped, **V**erbose, **F**ile)

  - gzip

  - bzip2 (pbzip2)

File Compression

- Many different formats
  - tar?

    tar -czvf archive.tar.gz /path/to/files/

    (**C**reate, **Z**ipped, **V**erbose, **F**ile)

  - gzip

    gzip sample.fastq

  - bzip2 (pbzip2)

File Compression

- Many different formats
  - tar?

    tar -czvf archive.tar.gz /path/to/files/

    (**C**reate, **Z**ipped, **V**erbose, **F**ile)

  - gzip

    gzip sample.fastq

  - bzip2 (pbzip2)

    bzip2 sample.fastq

    (pbzip2 sample.fastq)

File Decompression

- Many different formats
  - tar?

    tar -xzvf archive.tar.gz

    (e**X**tract, **Z**ipped, **V**erbose, **F**ile)

  - gzip

  - bzip2 (pbzip2)

File Decompression

- Many different formats
  - tar?

    tar -xzvf archive.tar.gz

    (e**X**tract, **Z**ipped, **V**erbose, **F**ile)

  - gzip

    gunzip sample.fastq.gz

  - bzip2 (pbzip2)

File Decompression

- Many different formats
  - tar?

    tar -xzvf archive.tar.gz

    (e**X**tract, **Z**ipped, **V**erbose, **F**ile)

  - gzip

    gunzip sample.fastq.gz

  - bzip2 (pbzip2)

    bunzip2 sample.fastq.bz2

    (pbunzip2 sample.fastq.bz2)

File Decompression

- FAQ articles at http://www.uppmax.uu.se/faq

- http://www.uppmax.uu.se/faq/should-i-use-a-specialty-compression-tool-for-fastq-format-files
- http://www.uppmax.uu.se/faq/which-compression-format-should-i-use-for-ngs-related-files

Web Export

/proj/b2016999/webexport/

available at

https://export.uppmax.uu.se/b2016999/

SciLifeLab UPPSALA

Web Export

- Served by Apache
  - Custom .htaccess and .htpasswd possible
- module load www-tools
- http://www.uppmax.uu.se/webexport-guide

www-tools

- www-add-dir <path>
  - Create/modify a directory with Apache settings

www-tools

- www-add-dir <path>
  - Create/modify a directory with Apache settings

- www-add-user <username>
  - Add user to password protected directory

www-tools

- www-add-dir <path>
  - Create/modify a directory with Apache settings
- www-add-user <username>
  - Add user to password protected directory
- www-rem-user <username>
  - Remove user from password protected directory

SciLifeLab
UPPSALA

Thanks!

Questions?