Use the following procedures to capture data using a Linux workstation. See Data capture (/transfer-appliance/docs/2.0/data-capture) for more information on data capture options.

You need to know the following information in order to capture data from a Linux workstation:

- The URL for Transfer Appliance. This is the URL presented when you connect Transfer Appliance (/transfer-appliance/docs/2.0/connect-appliance).
- The capture user credentials. These credentials are initially supplied to you by Google, but you can change the capture password (/transfer-appliance/docs/2.0/resetting-password).

In addition, make sure that:

- You are using a supported version of Linux. See supported operating systems (/transfer-appliance/docs/2.0/specifications).
- You have prepared for data transfer (/transfer-appliance/docs/2.0/preparing-data-transfer).
- You are using the correct data transfer method (/transfer-appliance/docs/2.0/choosing-capture-method).
- Your network meets Transfer Appliance requirements (/transfer-appliance/docs/2.0/preparing-data-transfer#prepare-network).
- Transfer Appliance is unpacked and configured (/transfer-appliance/docs/2.0/preparing-to-unpack).

Use the following procedure to install the Capture Utility on a Linux workstation.

1. On a workstation, connect to the Transfer Appliance web interface.
2. Select **Linux** in the **Data Capture** menu to download the Capture Utility installer. The Capture Utility contains configuration information specific to the Transfer Appliance from which it was downloaded. Only download the Capture Utility from the appropriate Transfer Appliance.
3. Extract the downloaded archive.

This creates the **CaptureUtility** directory containing the installer script, `installer.sh`.

★ **Note:** Extract the Capture Utility to a local filesystem. The Capture Utility is not designed to be run from remote filesystems.

4. Run the `installer.sh` in the **CaptureUtility** directory:

   The Capture Utility `tacapture.sh` is installed in the same **CaptureUtility** directory .

Before running your first data capture job, confirm that the workstation can connect to the Transfer Appliance.

1. On the Linux workstation, open a terminal window.
2. Test connectivity with the Transfer Appliance by running `tacapture.sh` with the `-t` option.

3. You are prompted to enter the capture user password.

   The utility tests connectivity between the workstation and Transfer Appliance and returns results similar to the following:

If the connectivity test fails, refer to the error message returned to determine the cause. The most common reason for connectivity test failure is a firewall blocking the ports needed for data capture. See Preparing the Network (/transfer-appliance/docs/2.0/preparing-network) for details on which ports need to be open.

Before you start a Linux workstation capture, make sure you have the following permissions to your data source:

- Read and execute permissions for folders.
- Read permissions for files.

If you are only capturing data that you own, you already have the required file and folder permissions. If you want to capture data owned by other people, contact your IT administrator and ask for an account that grants you access to the data. If you are in IT administrator in charge of moving other people's data, use a backup operator or service account that has read access to all of the data.

After checking connectivity, use the Capture Utility to start a data capture job.

1. Run **`tacapture.sh`** and specify a job name and capture target:

   where `[JOB NAME]` is the name of the data capture job and `[CAPTURE DIRECTORY]` is the directory that contains the data to capture. The Capture Utility recursively captures all data in the directories under the one specified.

   Use a meaningful job name. A job name is used to identify a capture job and the files it contains for the rest of the data migration project. A job name can contain alphanumeric, underscore, and hyphen characters.

★ **Note:** It is recommended that you create a maximum of 1,000 data capture jobs. If the number of capture jobs exceeds 1,000, Transfer Appliance and the rehydration process can experience degraded performance.

2. When prompted, enter the capture user password provided to you by Google:

   The capture job runs and displays a completion message when it is finished. Leave the terminal window open while the capture job is running, or the job will be terminated. It takes up to ten minutes for the Transfer Appliance Web interface to display a "Failed" status.

   For example, the following command creates a job named `data-capture` to capture the data in the directory `/mnt/data` and its subdirectories.

By default, the Capture Utility uses up to 8 parallel tasks for each capture job. To use fewer parallel capture tasks, use the `-m` option with `tacapture.sh`. If you have poor network bandwidth, reduce the number of parallel tasks for each capture job.

For example, the following command creates a job named `dataFactory` that uses at most 6 parallel data capture tasks to capture the data in the directory `/mnt/data` and all subdirectories.

For more information about the options for the Capture Utility, see Capture utility reference (/transfer-appliance/docs/2.0/capture-utility-reference).

You can specify data paths for capture in one of two ways:

- If you want to capture a single directory or file, provide a single directory when running a capture job. This method captures data from the specified directory and its subdirectories.
- To capture multiple target files and directories, run `tacapture.sh` with the `-f` option. You must provide a line delimited text file containing absolute paths to the files you want to capture.

The following is an example of the contents of the text file you must provide when you use the `-f` option.

The following command creates a job `dataFactory` that captures data from the above targets specified in `/home/user/filespec.txt`.

For more information about Capture Utility options, see Capture utility reference (/transfer-appliance/docs/2.0/capture-utility-reference).

To direct the Capture Utility to skip symbolic links to files, use the `-e` option with `tacapture.sh`. For example, the following command creates a job named `data-capture` to capture the data in the directory `/mnt/data` (including its subdirectories), but skips symbolic links to files.

For more information about the options for the Capture Utility, see Capture utility reference (/transfer-appliance/docs/2.0/capture-utility-reference).

The Capture Utility (`tacapture.sh`) preserves the following file metadata by default:

- The user ID of the owner.

- The group ID of the owning group.
- The file permissions.
- The last modification time of the file.
- The last file access time.

For example, the following command creates a job named data-capture to capture the data and file metadata in the directory `/mnt/data` and all subdirectories.

For more information about the options for the Capture Utility, see Capture utility reference (/transfer-appliance/docs/2.0/capture-utility-reference).

To perform a parallel data capture task, follow these instructions on a separate workstation. You can also perform a Microsoft Windows workstation capture (/transfer-appliance/docs/2.0/capturing-data-windows) using the same Transfer Appliance.

If your data size exceeds the capacity of a single Transfer Appliance, capture your data using multiple appliances in succession (/transfer-appliance/docs/2.0/performing-serial-captures).

To retry a transfer job, see Retrying unsuccessful data capture jobs (/transfer-appliance/docs/2.0/retrying-failed-jobs).

To cancel a transfer job, see Canceling transfer jobs (/transfer-appliance/docs/2.0/canceling-jobs).

To monitor:

- data capture jobs, see Monitoring data capture jobs (/transfer-appliance/docs/2.0/monitoring-capture-jobs).
- Transfer Appliance storage usage, see Monitoring appliance storage usage (/transfer-appliance/docs/2.0/monitoring-appliance-storage).
- disk status, see Monitoring disk status (/transfer-appliance/docs/2.0/monitoring-disk-status).

If you are done capturing data, see Preparing and shipping an appliance (/transfer-appliance/docs/2.0/shipping-appliance).