MinION Mk1C IT Requirements

Checklist

This checklist represents the minimal requirements for installing the MinION Mk1C in your institution. For full explanation of requirements, please continue to the document

Item/setup required	Reason	Provided?
1x 1 Gbp/s RJ45 port with DHCP service running (initial configuration only) OR Wi-Fi hotspot with internet connectivity (captive portals unsupported. Initial configuration only)	Connection to IT infrastructure and internet	
If using wired connection: 1x 1 Gbp/s Ethernet cable	Connection to IT infrastructure and internet	
HTTPS/port 443 TCP access to AWS eu-west-1 IP ranges listed at http://docs.aws.amazon.com/general/latest/gr/aws-ip-ranges.html OR DNS rules ping.oxfordnanoportal.com and *.nanoporetech.com	Telemetry feedback, EPI2ME analysis	
HTTPS/port 443 and HTTP/port 80 TCP access to Linode-hosted IP addresses 178.79.175.200 and 96.126.99.215 OR DNS rules cdn.oxfordnanoportal.com, and *.nanoporetech.com	Software updates	
Standard wall power socket	Device power	

MinION Mk1C device

Overview

The Oxford Nanopore Technologies® MinION™ Mk1C is a compact, portable device, combining the hardware for running nanopore sequencing experiments with fully-integrated compute used for basecalling and onward analysis. The device is entirely self-contained and requires no other computing resource.

The MinION Mk1C can be used with MinION flow cells and the Flongle adapter and flow cells. Users can operate the system in an offline or online mode, enabling full sequencing and analysis to be performed in the lab and field. A high-definition touch screen means users receive constant feedback regarding experiment progression and can easily interact with the device without additional hardware.

The MinION Mk1C benefits from the inclusion of on-board compute which permits data acquisition, analysis and feedback, basecalling, data streaming and device control, all without placing any additional burden on existing IT infrastructure.

All device control, basecalling, analysis and orchestration on the MinION Mk1C is carried out by pre-installed custom software created by Oxford Nanopore Technologies.

Specifications

The MinION Mk1C is designed around a simple user interface on top of cutting-edge custom electronics providing real-time analysis solutions:

Component	Specification
Size and weight	H 33 mm x L 142 mm x W 118 mm; 455 g
Power	Supplied with a 6.3-19.6 VDC power supply Max rated current 10 A Max rated power 60 W
Compute spec	1 TB SSD Storage, 8 GB RAM, GPU embedded analysis accelerator
Connectivity	WiFi, USB 2.0 with eSATA (the eSATA is currently non-functional), microSD, Ethernet
Pre-loaded software	Linux OS, MinKNOW and Guppy
Security	Kensington Nano Security Slot
Environmental conditions	Tested to function between 0°C to +40°C. Do not cover vents on the top or sides of the device

Networking explanations

The two requirements for firewall permissions are:
1. Access to the AWS eu-west-1 region for telemetry feedback and use of the EPI2ME platform through port 443, OR DNS rules ping.oxfordnanoportal.com and *.nanoporetech.com
2. Access to the Linode-hosted IP addresses 178.79.175.200 and 96.126.99.215 for software distribution, OR DNS rules cdn.oxfordnanoportal.com and *.nanoporetech.com

Telemetry

Telemetry information is collected by MinKNOW during sequencing runs as per the Terms and Conditions to allow monitoring of device performance and enable remote troubleshooting. Some of this information is obtained from free-form text entry fields, therefore no personally-identifiable information should be included. No sequence data is collected.

The EPI2ME platform is hosted within AWS and provides cloud-based analysis solutions for multiple applications. Use of this platform entails users uploading sequence data in FASTQ format from the EPI2ME Agent, which is then processed through defined pipelines within the EPI2ME Portal. Downloads from EPI2ME are either in Data+Telemetry or Telemetry form. Telemetry information is used to populate reports within the EPI2ME portal.

Software updates

The IP address from which you receive software updates will depend upon your geographical location. Updates are performed through **apt** on the terminal, therefore outbound-only access is required. Users are notified of the availability of software updates through the Nanopore Community, and full instructions for updating are provided in each release note.

Storage

File types

Nanopore sequencing data is stored in two file types, .fast5 and FASTQ. Basecalling summary information is stored in a *sequencing_summary.txt* file:

- .fast5 is a customized file format based upon the .hdf5 file type, which is designed to contain all information needed for

analysing nanopore sequencing data, including raw signal data, and tracking it back to its source. As default, each .fast5 file will contain 4000 reads although this can be configured when starting a run. The size of a .fast5 file size will depend on the number of reads contained and the length of DNA/RNA sequenced. See the table below for estimates of storage space required for a typical sequencing run.

- FASTQ is a universal text-based sequence storage format, generated when the nanopore signal data is basecalled and containing both the sequence of DNA/RNA and its quality scores. By default, the device saves up to 4000 sequences in one FASTQ file. The size of a FASTQ file will depend on the number of reads contained and the length of DNA/RNA sequenced. See the table below for estimates of storage space required for a typical sequencing run.

- *sequencing_summary.txt* contains metadata about all basecalled reads from an individual run. Information includes read ID, sequence length, per-read q-score, duration etc. The size of a sequence summary file will depend on the number of reads sequenced.

Typically, 1 Gbase of sequence data takes up approximately 10 Gbytes of storage. This typically comprises 90% .fast5 files, 9% FASTQ files and 1% sequence summary file. However, the exact number will depend on factors such as chemistry version, read length, etc.

Example file sizes below are based on different outputs from a flow cell, with a run saving both .fast5 and FASTQ files with a read N50 of 25 kb.

Output (Gbases)	.fast5 storage (Gbytes)	FASTQ storage (Gbytes)	.fast5 + FASTQ storage (Gbytes)
10	90	10	100
15	135	15	150
30	270	30	300

As an experiment progresses, .fast5 files are produced for all reads. If basecalling is chosen, these reads are utilised by the on-board software (more information below) to generate sequence data which is then stored in FASTQ files.

Long-term storage

The MinION Mk1C has sufficient SSD disk space for multiple runs to be carried out, storing both .fast5 and FASTQ data. However, it is imperative this data store is cleared regularly to prevent successive runs from terminating due to lack of storage space. For this, a user must provide storage to transfer data off the device.

The MinION Mk1C runs onboard admin software that is able to mount multiple filesystem types. If being installed within an institution, filesystem mounting can also be achieved through SSH and command line interactions. We recommend storage presented as NFS or CIFS.

The device can also directly mount and transfer data to USB and microSD storage peripherals. In addition, a password-protected samba share can be operated through a Wi-Fi hotspot for wireless data transmission.

The form (and volume) of data to be stored and transferred will depend on customer requirements:
- Storing .fast5 files with raw read data in will permit re-basecalling of data when new algorithms are released by Oxford Nanopore. In such cases, new releases of basecallers have enabled significant improvements in basecalling accuracy of existing datasets through re-basecalling. Further, selected Oxford Nanopore and third-party tools use the raw signal information contained within the .fast5 to extract additional information from the raw signal e.g calling modified bases, reference-guided SNP calling or polishing of data.
- Retaining just FASTQ files will allow use of standard downstream analysis tools using the DNA/RNA sequence, but no further sequence data can be generated when improvements in basecalling become available.

Included software

Oxford Nanopore Technologies build and provide numerous software types involved in acquisition, orchestration and analysis:

MinKNOW

MinKNOW carries out several core tasks:

- Device control, including run parameter selection
- Data acquisition
- Real-time analysis and feedback
- Data streaming
- Basecalling (through integrated Guppy)

The MinKNOW software carries out several core tasks: data acquisition, real-time analysis and feedback, basecalling, data streaming, controlling the device, and ensuring that the platform chemistry is performing correctly to run the samples. MinKNOW takes the raw data and converts it into reads by recognition of the distinctive change in current that occurs when a DNA strand enters and leaves the pore. MinKNOW then basecalls the reads, and writes out the data into .fast5 or FASTQ files.

Guppy

Guppy is a data processing toolkit that contains the Oxford Nanopore Technologies' basecalling algorithms, and several bioinformatic post-processing features. It is run from the command line, but has a low entry level for use. A selection of configuration files allow basecalling of DNA and RNA libraries, made with Oxford Nanopore Technologies current sequencing kits, in a varied range of flow cells.

If rebasecalling off the MinION Mk1C, it is possible to use an external GPU to basecall with Guppy, and this will dramatically increase basecalling speed over using CPU. Guppy currently works with only NVIDIA GPUs, with the following specific models supported:

- NVIDIA Volta GV100
- NVIDIA Tesla V100
- NVIDIA GTX1080Ti
- NVIDIA Jetson TX2

It is possible to use other NVIDIA GPUs for basecalling, but support for those is limited.

EPI2ME

The EPI2ME platform is a cloud-based data analysis service offering a number of workflows for end-to-end analysis of nanopore data. The basecalled read files from a sequencing run can be uploaded to the EPI2ME platform via another piece of software called the Desktop Agent. The user selects a workflow in the Agent such as 16S alignment or barcoding, and the Agent transfers the read files one by one into the cloud for real-time analysis. The analysis results are generated and updated in real-time in the form of a report in the EPI2ME platform.

Frequently asked questions

Does my device require an internet connection to function?
No, an internet connection is not a requirement for the device to start or to continue sequencing. An internet connection will however be required to provide telemetry and debug information in the case of troubleshooting support being required. Additionally, an internet connection will be required to make use of any cloud-based analysis through EPI2ME.

Does the device have a battery?
No, the MinION Mk1C does not contain an internal battery. It is however compatible with commercially-available power supplies for e.g. laptops.

Change log

Date	Version	Changes made
19th August 2022	V4	The .fast5/FASTQ file size table has been updated.

Date	Version	Changes made
21st April 20221	V3	Updated the portal link from mirror.oxfordnanoportal.com to cdn.oxfordnanoportal.com in the checklist and network explanations
28th April 2021	V2	- Added microSD under Connectivity and Kensington Nano Seciruty slot under Security in "Technical specifications"
December 2020	V1	Initial version