MinION Mk1B IT requirements

Checklist

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

Item/setup required	Reason	Provided?
Windows – 7, 8, 10 OSX – Sierra, High Sierra, Mojave, Catalina Linux – Ubuntu 16.04 or 18.04 Memory: 16 GB RAM CPU: i7 or Xeon with 4+ cores GPU: MinKNOW is only supported on GPUs built into Oxford Nanopore sequencing devices. Storage: 1 TB internal SSD Ports: USB3	Ensure computer has enough resource to run a MinION without deterioration of performance	
Account privilege: local admin	Ensure MinKNOW has correct CPU priority for data acquisition	
Internet connection	Required for updates and telemetry	
OS update settings: manual	Ensure updates do not impact performance or halt runs	
Ethernet: 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	Telemetry feedback, EPI2ME analysis	
Ethernet: HTTPS/port 443 and HTTP/port 80 TCP access to Linode-hosted IP addresses 178.79.175.200 and 96.126.99.215 for software distribution OR DNS rules ping.oxfordnanoportal.com, mirror.oxfordnanoportal.com, and *.nanoporetech.com	Software updates	

MinION nanopore sequencing device

Overview

The MinION™ is a small, self-contained device for nanopore sequencing. It plugs directly into a standard USB 3.0 port on a desktop or laptop computer.

The MinION device is controlled by MinKNOW™ software.

MinKNOW carries out several core tasks: data acquisition, real-time analysis and feedback, basecalling, data streaming, device control including selecting the run parameters, sample identification and tracking and ensuring that the platform chemistry is performing correctly to run the samples.

Despite its size, the MinION is capable of sequencing many gigabases (Gbases) in a single sequencing experiment. As such, it is recommended that users purchase a powerful PC to ensure all the technology features are accessible.

The default data analysis workflow when starting to use the MinION is below:

Minimum host computer specification

This example will allow a user to run a MinION with real-time local basecalling. Given the amount of data potentially generated, the basecalling may fall behind and need to be completed at the end of an experiment. An i7 CPU will basecall 36k bases/s with the fast basecalling model, and 4.4k bases/s with the high-accuracy model. In contrast, the GPU-containing MinIT will basecall 167k bases/s with the fast model and 22k bases/s with the high-accuracy model.

Component	Configuration
Operating system	Windows – 7, 8, 10 OSX – Sierra, High Sierra (64 bit), Mojave, Catalina Linux† – Ubuntu 16.04 or 18.04
Memory/RAM	16 GB RAM
CPU	i7 or Xeon††
GPU	MinKNOW is only supported on GPUs built into Oxford Nanopore sequencing devices. It is possible to use an external GPU to basecall with the standalone Guppy basecaller, and this will dramatically increase basecalling speed. Guppy works with only NVIDIA GPUs, with the following specific models supported: - NVIDIA Tesla V100 - NVIDIA Quadro GV100 - NVIDIA Jetson TX2 - NVIDIA Jetson Xavier It is possible to use other NVIDIA GPUs for basecalling, but support for those is limited. GPU basecalling with Guppy is supported on: - Ubuntu 16.04 (via deb) - Red Hat / Centos 7 (via RPM) - Generic Linux platforms
Storage	1 TB internal SSD
Ports	USB3‡

† Linux products are offered under limited support and may take the team longer to respond to queries

†† Users need to verify that their i7 is a four core model or better

‡ The MinION device is CE marked using USB3. If a user wished to use USBC, they may, but this invalidates the CE marking

We recommend that MinKNOW is installed to the internal SSD, rather than the main operating system drive. This avoids intermediate files being written to a non-SSD location during the experiment, potentially leading to a build-up of intermediate files that can cause the system to

run out of space and stop.

The below requirements are to ensure correct operation of the MinION during a sequencing experiment:

Component	Minimum requirement
User account privilege level	Local Administrator This is in order to give MinKNOW correct CPU priority for data acquisition.
Internet connection	An internet connection is required at all times for software updates and telemetry Offline configurations can be made available for field use and expeditions. Please contact Support.
Antivirus settings	Antivirus software scanning the entire hard drive uses significant amounts of resources and may cause performance issues. For this reason, it is recommended that the scans are switched to manual setting, rather than to automatic, so that they can be performed when the MinION is not in use.
OS update settings	OS system updates should be switched to manual mode, as downloading them during a sequencing experiment may impact performance. Updates that require restart will fatally halt a run.

MinIT for running MinIONs and basecalling
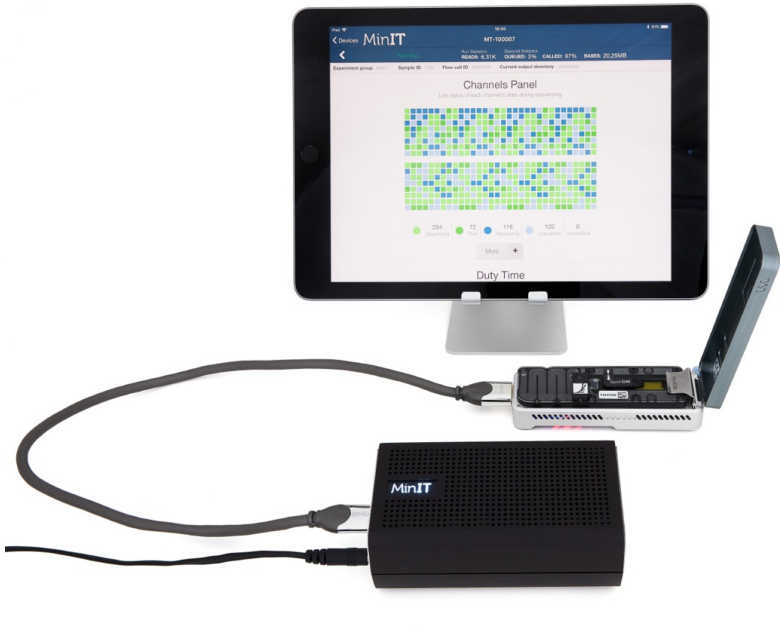
MinIT is a companion to the MinION personal DNA/RNA sequencer. It is a pre-configured compute module with the software that controls the MinION, carries out data acquisition, and performs basecalling.

The MinIT is:

- Simple to install

- Easy to connect to via Wi-Fi from a tablet or laptop

- Enabled to keep up with real-time basecalling and analysis utilising the GPU accelerators on board

MinIT specifications:

- Pre-installed software: Linux OS, MinKNOW, Guppy, EPI2ME

- Wi-Fi enabled; you can control your experiments using a laptop, tablet or smartphone

- FASTQ files are written to an onboard storage: 512 GB SSD
- Processing: GPU accelerators (ARM processor 6 cores, 256 Core GPU), 8 GB RAM
- Small footprint: 290 g
- 1 x USB3 (to connect your MinION), 1 x USB2 (for data transfer), 1 x 1 GB Ethernet (for networking)

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

2. a. Access to the Linode-hosted IP addresses 178.79.175.200 and 96.126.99.215 for software distribution

b. OR DNS rules ping.oxfordnanoportal.com, mirror.oxfordnanoportal.com, and *.nanoporetech.com

Telemetry

Telemetry information is collected by MinKNOW sequencing runs as per the Terms and Conditions to allow monitoring of experiment performance and provide support in troubleshooting cases. 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 from Metrichor is hosted within AWS and provides cloud-based analysis solutions for multiple applications. Use of this platform entails users uploading sequence data in .fast5 or 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

Depending on your geographical region, only one of 178.79.175.200 or 96.126.99.215 will be used for provision of updates to device software. The Updates are triggered as pull requests, therefore outbound-only access is required.

File types

The MinION stores biological data in two file types, .fast5 and FASTQ:

- .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 and tracking it back to its source. A .fast5 file contains data from multiple reads (4000 reads as default), and is several hundred Mb in size. Storing .fast5 files with raw read data in will permit re-basecalling of data when new improvements are released by Oxford Nanopore.
- FASTQ is a universal text-based sequence storage format, containing both the sequence of DNA/RNA and its quality scores. By default, the MinION saves up to 4000 DNA sequences in one FASTQ file. File size can vary from < 1Mb to tens of Mb depending on the number and length of sequences. Retaining only FASTQ files will allow use of downstream analysis tools, but no further sequence data can be generated when improvements in basecalling become available.

As an experiment progresses, .fast5 files are produced for all reads. If basecalling is chosen, these reads are utilized by the onboard software (more information below) to generate sequence data which is then stored in FASTQ files and as a FASTQ record within the .fast5 file.

Included software

Oxford Nanopore Technologies builds and provides numerous software types involved in acquisition, orchestration and analysis:
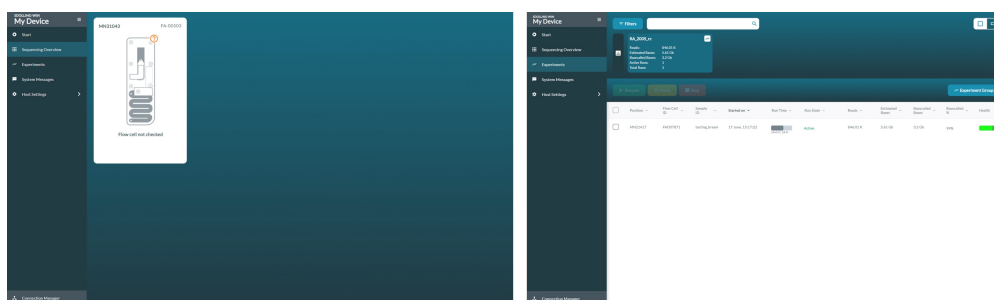
- MinKNOW
- Guppy
- EPI2ME

MinKNOW

MinKNOW carries out several core tasks:

- Data acquisition
- Real-time analysis and feedback

- Data streaming
- Device control, including run parameter selection
- Sample identification and tracking
- Ensuring chemistry is performing correctly

MinKNOW utilizes an intuitive graphical user interface (GUI) and receives updates on a regular basis. This is the core software provided by Oxford Nanopore, without which the sequencing devices cannot be run. Data from MinKNOW is packaged into individual read .fast5 files (over 1 million of which can be generated by a single flow cell), which are a customised file format based upon the .hdf5 file type. These .fast5 files are then used by other downstream software.

Figure 1: Example screen shots of the MinKNOW GUI. This permits selection and initiation of experiments, as well as providing real-time feedback on experiment progression

Guppy

Guppy is a production basecaller provided by Oxford Nanopore, and uses a command-line interface. It utilizes the latest in Recurrent Neural Network algorithms in order to interpret the signal data from the nanopore, and basecall the DNA or RNA passing through the pore. It is optimiszed for running with basecall accelerators e.g. GPUs. Guppy implements stable features into Oxford Nanopore Technologies' software products, and is fully supported. It receives .fast5 files as an input, and is capable of producing:
- .fast5 files appended with basecalled information
- .fast5 files that have been processed, but basecall information present in a separate FASTQ file

EPI2ME

EPI2ME™ is an onwards data-analysis platform created by Oxford Nanopore's subsidiary company, Metrichor. It provides users with real-time analysis such as species identification, alignment workflows and other bioinformatics solutions. It is currently provided as a cloud-based analysis platform, which is initiated through the local EPI2ME Agent.
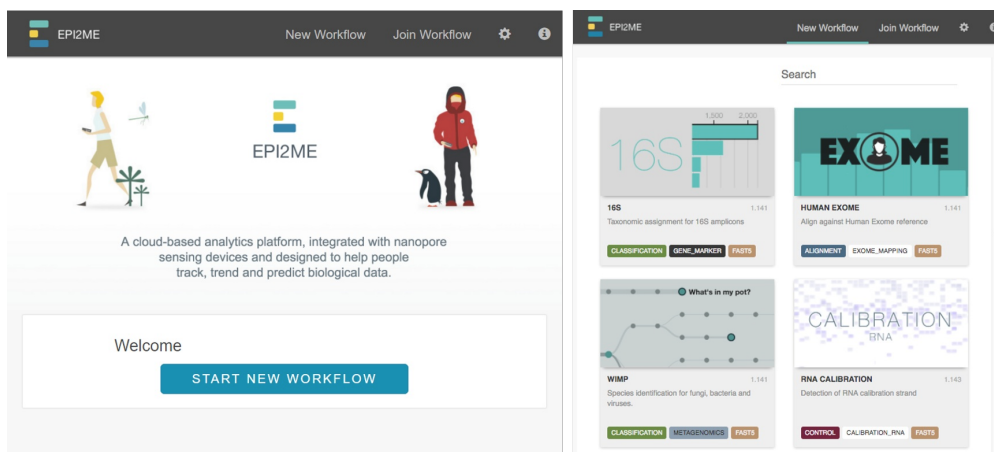
Figure 2: The EPI2ME Agent. Installed on the device, this allows selection and initiation of workflows within the EPI2ME portal.

Users will be able to review basic experimental statistics as well as use prepared workflows such as real-time taxonomic identification, and reference

alignment.

Figure 3: Example reports and plots from the EPI2ME portal.

Change log

Date	Version	Changes made
Feb 2020	V3	More information about basecalling on GPUs using Guppy
18th May 2020	V4	Removed Bluetooth compatibility for MinIT
3rd July 2020	V5	Updated Fig 1 to 4.0.8 UI
23rd Oct 2020	V6	Updated OSX compatibility
23rd Dec 2020	V7	- Removed requirement for accessing IP address 106.187.40.102 - Updated the supported external NVIDIA GPUs to Tesla V100, Quadro GV100, Jetson TX2 and Jetson Xavier