

PromethION 24/48 A-Series IT requirements

Checklist

This checklist represents the minimal requirements for installing the PromethION in your institution. For full explanation of requirements, please continue to the document. Please note either Fibre **OR** Copper connections can be used.

Item/setup required	Reason
2x SFP+ modules supporting 10GBASE-SR or 10GBASE-LR (not required if copper used)	Connection to IT infrastructure and internet
Appropriate 10 Gbps cabling for chosen SFP+ modules, OR Appropriate cabling for 10 Gbps Ethernet	Connection to IT infrastructure and internet
2x 10 Gbps fibre ports with DHCP service running, OR 2x 10 Gbps Ethernet ports with DHCP service running	Connection to IT infrastructure and internet
HTTPS/port 443 to 52.17.110.146, 52.31.111.95, 79.125.100.3 (outbound-only access) or DNS rule for ping.oxfordnanoportal.com	Telemetry feedback
Ethernet: HTTPS/port: 443 TCP access to AWS eu-west-1 IP ranges: http://docs.aws.amazon.com/general/latest/gr/aws-ip-ranges.html	EPI2ME analysis
HTTPS/port 443 to 178.79.175.200 and 96.126.99.215 (outbound-only access) or DNS rule for cdn.oxfordnanoportal.com	Software updates
Storage: *Sufficient infrastructure for required storage option	Streaming and long-term data storage
USB mouse	Device control
USB keyboard	Device control
VGA or DisplayPort-compatible monitor	Device control
3x C13 Power supply cables (2x for Data Acquisition Unit, 1x for Sequencing Unit.)	Device power

PromethION 24/48 A-Series

Overview

The PromethION is a benchtop device for nanopore sequencing designed to run and analyse up to 24 or 48 flow cells. It is ideal for labs with multiple projects that need the advantages of nanopore sequencing:

- Simple library preparation
- Real-time analysis
- Biological insights from long reads

In addition, the PromethION 24/48 also enables users to offer nanopore sequencing as a service when [certified](#).

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

All device control, data acquisition and basecalling on the PromethION is carried out by pre-installed custom software created by Oxford Nanopore Technologies. The default data analysis workflow when using the PromethION is as below:

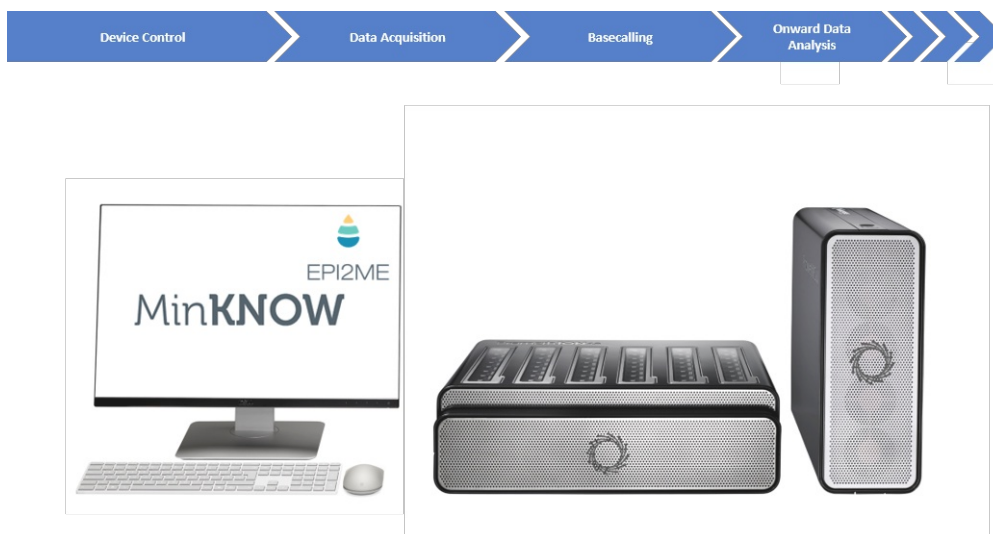


Figure 1: Default data analysis workflow of the PromethION 24/48 device

Specifications

The PromethION is designed around a simple user interface on top of cutting-edge custom electronics providing real-time analysis solutions. The specification of your device will depend on the P24 or P48 model:

Component	Specification
Operating system	Ubuntu Focal (20.04) running off Intel CPU Customers are advised to keep up to date with all software and security patches
Storage	60 TB internal SSD
GPU cards	Data Acquisition Unit: 4x NVIDIA Ampere-series GPUs connected with NVLink
Memory	512 GB RAM
Size and weight	Sequencing Unit: - H 190 x W 590 x D 430 mm - Weight 28 kg Data Acquisition Unit: - H 440 x W 178 x D 470 mm - Weight 25 kg
Environmental ranges	Designed to sequence at +18° C to +22° C*
Maximum heat output	Sequencing Unit: 4,094 BTU/hr Data Acquisition Unit: 6,824 BTU/hr
Input voltage range: PromethION Sequencer	100-240 V (50/60 Hz) Maximum power: 1.2 kW
Input voltage range: PromethION A100 Data Acquisition Unit**	Input voltage range: 200–240 V (50/60Hz) Maximum power (240 V input): 2.2 kW Operating current (240 V input): 9.1 A

*Functional range of electronics +5° C to +40° C

**PromethION A-Series Data Acquisition Unit contains an additional redundant power supply for reliability.

The PromethION A100/A-series requires electricity from a mains source at 200-240 V. If you are in a region where this is not standard (e.g., the U.S.), you will need access to two main sockets with mains electricity at 200-240 V.

Telemetry

MinKNOW collects telemetry information during sequencing runs as per the Terms and Conditions to allow monitoring of device performance and enable remote troubleshooting. Some of this information comes from free-form text entry fields, therefore no personally-identifiable information should be included. We do not collect any sequence data.

The EPI2ME platform is hosted within AWS and provides cloud-based analysis solutions for multiple applications. Users upload sequence data in FASTQ format via the EPI2ME Agent, which processes the data through defined pipelines within the EPI2ME Portal. Downloads from EPI2ME are either in Data+Telemetry or Telemetry form. The EPI2ME portal uses telemetry information to populate reports.

Software updates

The IP address from which you receive software updates will depend on your geographical location. You can update through the software UI or through **apt** on the terminal, so you require outbound-only access. We notify users about software updates through the Nanopore Community and provide full instructions for updating in each release note.

Storage

File types

Nanopore sequencing data is stored in three file types: POD5, FASTQ and BAM. Basecalling summary information is stored in a sequencing_summary.txt file:

- POD5 is an Oxford Nanopore-developed file format which stores nanopore data in an accessible way and replaces the legacy .fast5 format. This output also reads and writes data faster, uses less compute and has smaller raw data file size than .fast5.
- .fast5 is a legacy file format based upon the .hdf5 file type, which contains 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.
- FASTQ is a text-based sequence storage format, containing both the sequence of DNA/RNA and its quality scores. By default, nanopore sequencing experiments save up to 4000 DNA sequences in one FASTQ file. File size can vary from <1 Mb to tens of Mb depending on the number and length of sequences.
- BAM files are output if you perform alignment or modified base calling on the basecalled dataset.
- 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.

Example file sizes below are based on different throughputs from an individual flow cell, with a run saving POD5, FASTQ, and BAM files with a read N50 of 23 kb.

Example file sizes below are based on different throughputs from an individual flow cell, with a run saving POD5, FASTQ, and BAM files with a read N50 of 23 kb. TMO = theoretical maximum output.

Flow cell output (Gbases)	POD5 storage (Gbytes)	FASTQ.gz storage (Gbytes)	Unaligned BAM with modifications (Gbytes)
100	700	65	60
200	1,400	130	120
290 (TMO)	2,030	188.5	174

As an experiment progresses, POD5 files are produced for all reads. If you choose to basecall your data, these reads are used by the MinKNOW software to generate sequence data which is then stored in FASTQ files and/or BAM files.

Data transfer and long-term storage

The storage capacity of the PromethION A-Series Data Acquisition Unit is designed as a temporary cache for data being written by MinKNOW. The storage on the device is not designed for and should not be used for long-term storage of sequence data. When running the device at full capacity, it is essential that this data is streamed from the device in real-time to prevent runs from terminating due to lack of storage space. For this, a customer site must ensure connectivity to the local infrastructure is of sufficient bandwidth to prevent data backing up. The PromethION offers two 10 Gbit/s ports for this purpose, with the customer able to choose between Ethernet or Fibre solutions. USB ports do not provide sufficient bandwidth for real-time data transfer so should not be used. Below is a worked example showing idealised data transfer speeds; real transfer speeds could be slower depending on network configuration.

	1 Gbit/s	10 Gbit/s
1x 200 Gbase flow cell	~7 hours	~1 hour
48x 200 Gbase flow cell	~320 hours	~32 hours

The PromethION runs on Ubuntu 20.04 and is able to mount multiple filesystem types. We recommend storage presented as NFS or CIFS. This storage streamed to in real-time should be SSD for its high write speed compared to HDD. After initial writing to networked SSD drives, data can be moved to storage with a slower write speed for long-term storage.

The form and volume of data to be stored will depend on customer requirements:

- Storing POD5 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 POD5 to extract additional information from the raw signal e.g modified bases calling, 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.

Oxford Nanopore is unable to provide exact recommendations for storage, as these will be site-specific. The above guidelines and requirements should be taken into consideration.

Frequently asked questions

Can I use an uninterruptible power supply (UPS) with the PromethION 24/48?

Yes you can use a UPS, however we are unable to recommend a specific product to use.

Which SFP+ modules are compatible with the fibre ports on the PromethION 24/48?

Any SFP+ modules will be compatible, please select the SFP+ module that fits with the infrastructure of your institution.

Do you have any recommendations for how to move data off the PromethION 24/48 in real-time during a sequencing run?

Scripts are provided within the PromethION software packages to enable the user to set up real-time data transfer to a remote location. For full details, contact support@nanoporetech.com

Do you have recommendations for storage I can connect to the PromethION 24/48?

The PromethION can communicate with any NAS device; we suggest you discuss your storage requirements with your local IT team based on the information provided above to get the right solution for your needs.

Change log

Date	Version	Changes made
30th October 2023	V11	The "Specification" table has been updated with a requirement for a 200-240 V input voltage range.
28th September 2023	V10	<ul style="list-style-type: none"> - A file size table has been added to "File types" - Minor corrections and clarifications throughout the document
August 2023	V9	Device name change from A100 to A-Series
28th June 2023	V8	<ul style="list-style-type: none"> - In "Checklist", the specs for telemetry feedback, EPI2ME analysis, and software updates have been changed. - "Networking explanations" has been removed. - In "File types", information has been added about POD5 files and BAM files. The file size comparison table has been temporarily removed. - The "Included software" section has been removed. - Minor corrections and clarifications throughout the document.
6th September 2022	V7	<ul style="list-style-type: none"> - In the Checklist, the device power requirements have been updated to "3x C13 Power supply cables". - The example file size table has been updated to reflect data outputs characteristic of PromethION flow cells.
19th August 2022	V6	The .fast5/FASTQ file size table has been updated.
May 2022	V5	<ul style="list-style-type: none"> - Device power requirements have been clarified to "3x 1.2 kW power supplies for C13 cable (2x for Data Acquisition Unit, 1x for Sequencing Unit.)" - Updated the monitor requirements to "VGA or DisplayPort-compatible monitor" - In the specification table, the storage has been updated to 60 TB internal SSD, and the RAM to 512 GB - The functional temperature range for electronics has been corrected to +5°C to +40°C
April 2022	V4	Updated the portal link from mirror.oxfordnanoportal.com to cdn.oxfordnanoportal.com in the checklist and networking explanations section
March 2022	V3	<ul style="list-style-type: none"> - Updated to include information on the A100 Data Acquisition Unit - "Site requirements" section has been deleted - The list of supported GPUs in the "Guppy" section has been updated - FAQs have been updated
April 2020	V2	Amended output file sizes after the introduction of file compression
July 2019	V1	Initial version