Bionano VIA™ AWS System Requirement

DOCUMENT NUMBER:
CG-30580

DOCUMENT REVISION:
C

Effective Date:
07/20/2023

Table of Contents

Legal Notice

For Research Use Only. Not for use in diagnostic procedures.

This material is protected by United States Copyright Law and International Treaties. Unauthorized use of this material is prohibited. No part of the publication may be copied, reproduced, distributed, translated, reverse-engineered or transmitted in any form or by any media, or by any means, whether now known or unknown, without the express prior permission in writing from Bionano Genomics, Inc. Copying, under the law, includes translating into another language or format. The technical data contained herein is intended for ultimate destinations permitted by U.S. law. Diversion contrary to U. S. law prohibited. This publication represents the latest information available at the time of release. Due to continuous efforts to improve the product, technical changes may occur that are not reflected in this document. Bionano Genomics, Inc. reserves the right to make changes in specifications and other information contained in this publication at any time and without prior notice. Please contact Bionano Genomics, Inc. Customer Support for the latest information.

BIONANO GENOMICS, INC. DISCLAIMS ALL WARRANTIES WITH RESPECT TO THIS DOCUMENT, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO THOSE OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. TO THE FULLEST EXTENT ALLOWED BY LAW, IN NO EVENT SHALL BIONANO GENOMICS, INC. BE LIABLE, WHETHER IN CONTRACT, TORT, WARRANTY, OR UNDER ANY STATUTE OR ON ANY OTHER BASIS FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING BUT NOT LIMITED TO THE USE THEREOF, WHETHER OR NOT FORESEEABLE AND WHETHER OR NOT BIONANO GENOMICS, INC. IS ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Patents

Products of Bionano Genomics® may be covered by one or more U.S. or foreign patents.

Trademarks

The Bionano logo and names of Bionano products or services are registered trademarks or trademarks owned by Bionano Genomics, Inc. (“Bionano”) in the United States and certain other countries.

Bionano™, Bionano Genomics®, Saphyr®, Saphyr Chip®, Bionano Access™, VIA™ software, and Bionano EnFocus™ are trademarks of Bionano Genomics, Inc. All other trademarks are the sole property of their respective owners.

No license to use any trademarks of Bionano is given or implied. Users are not permitted to use these trademarks without the prior written consent of Bionano. The use of these trademarks or any other materials, except as permitted herein, is expressly prohibited and may be in violation of federal or other applicable laws.

© Copyright 2023 Bionano Genomics, Inc. All rights reserved.

Revision History

REVISION	NOTES
A	Initial Release
B	Branding update
C	Product name change to VIA and updates to min.hardware recommendations.

VIA Architecture - AWS

The VIA system is designed as a truly scalable system that is easy to deploy and maintain by labs of any size, from a single user to a large enterprise. The key to this scalability is the modular design of the system. The core system is composed of three components: Clients, Server, and Processing Unit. Depending on the types of samples processed in VIA, additional utility software may need to be installed (see **Figure 1**).

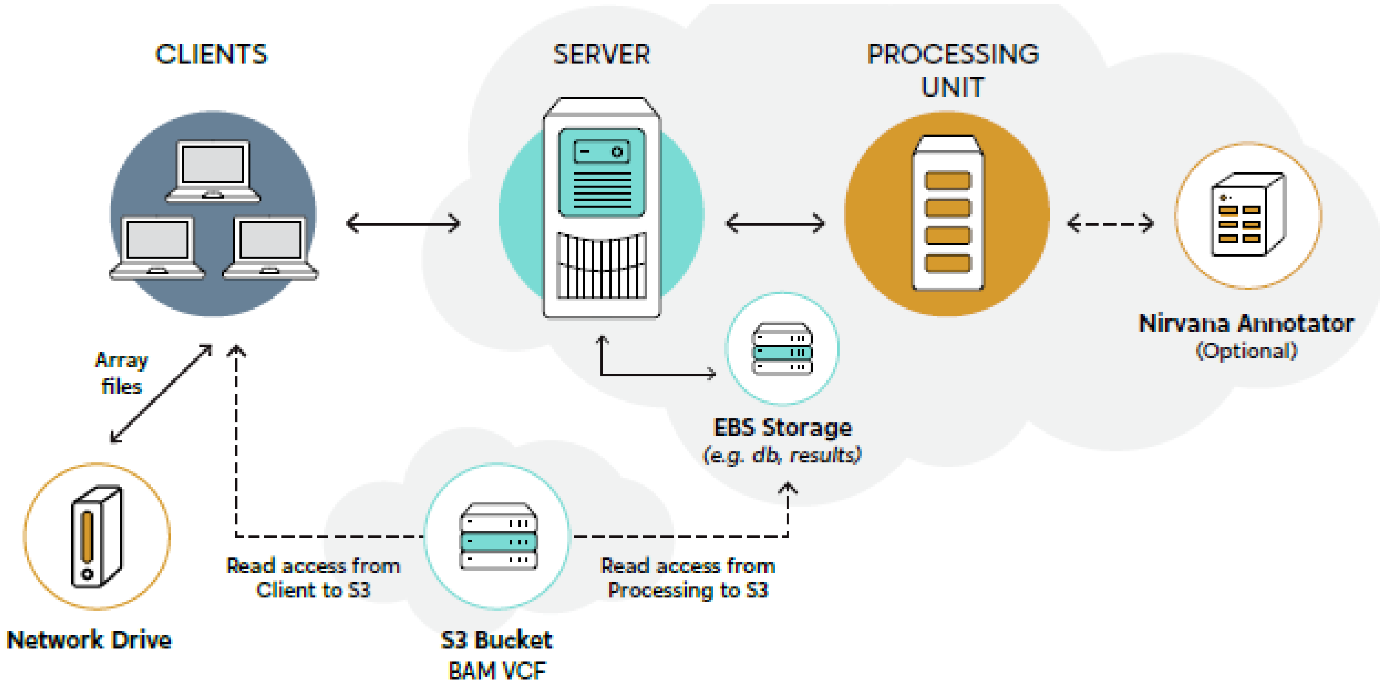

Figure 1. VIA on AWS EC2 with storage on EBS

Clients

The Client software is the Graphical User Interface (GUI) utilized by all users to access VIA. The Client software is installed on multiple computers, allowing multiple users to access VIA at the same time from within an institution’s network as well as remotely. After the initial installation, the system will be able to automatically update the client when the system administrator updates the software. This avoids manually updating each client.

Server

The server is the heart of the system; it manages the data in the repository as well as all the resources. (e.g., tracks, users, configurations, etc.). The database storage component of the VIA server is placed in the EBS Storage with some exceptions for “raw” NGS data (see NGS section). The server component is installed on Amazon AWS EC2, and all clients communicate with the server through a secure http connection. The server should be up and running all the time to service incoming requests from the clients and include a back-up policy to back up the content of the repository on a regular basis.

Processing Server(s)

The processing server is used to convert raw sample datapoints into variants. It will be hosted on the same EC2 instance. For large scale deployments, the system can utilize multiple Processing Servers in cases where there is demand for processing many samples simultaneously. The Processing Server communicates with the server via a secure protocol. The VIA admin can specify different processing servers for different types of samples for efficient processing; smaller files (e.g., panels/arrays) can be run on a less powerful processing server while WGS samples can be assigned to a much more powerful processing server. This unit does not require backups.

VIA Data File Types

Microarrays

All data including raw array data is stored on the server. Hardware requirements vary based on the type and size of data being processed and managed.

NGS

VIA software can store and process different types of NGS data. BAM files are processed to derive copy number and BAF. They are also used to process, and display read depth and individual reads in the genome browser. VCF files are used for sequence variants.

Processed data from VCF and BAM files is stored in the repository, but the original files are not copied to the server and are kept in their original location. The server keeps track of where these files are located by storing the file path to the files and updating this if the files move. In the case of an AWS setup, these files should be stored in an S3 bucket which is accessible to the EC2 instance running the Processing unit and to the clients on-prem.

NGS data comes in many flavors (VCF files, BAM files from targeted panels, BAM from WES, BAM from WGS) and the size of these files varies. Hardware requirements vary based on the type and size of data being processed and managed. For example, if processing deep WGS data, more RAM and faster processing are necessary.

VIA Software Additional Utility Software

Nirvana Annotator

VIA software uses the **Illumina Nirvana** annotation tool to annotate VCF files in a single seamless workflow beginning with loading of files into VIA. The **Nirvana** tool is installed separately from the VIA system. It is recommended that Nirvana be installed on the same machine as the Processing Server. The installation files are rather large and require at least 60GB of space (including supplementary databases).

BAM Multiscale Reference Builder

The **BAM MultiScale Reference Builder** is a separately installed application that is used to generate reference files to be used with the VIA software when processing BAM files to derive copy number from NGS platforms.

Building a reference file with this utility may only need to be done once, if the same reference file will be used for all the experimental NGS samples. This utility should be installed on a user machine that has access to BAM files. Hardware requirements vary based on the size of BAM files that will be used for creating the reference file. For a small NGS panel, 8GB RAM is minimum, but more RAM (up to 32GB) will be needed for deep WGS. The required space for this utility is 6GB.

VIA Clinical Sample Importer

The **VIA Sample Importer** is an optional command line utility that streamlines and automates sample loading and processing for VIA software. It requires as input the sample descriptor file that specifies samples, location, and processing settings. The installation space required for this utility is ~500MB in size. This utility should be installed on a user machine.

NOTE: See hardware recommendation on the next page.

Hardware Recommendation

Below are minimum hardware recommendations. Adjust the RAM and CPU based on expected volume and type of data being processed and expand disk space as the database grows.

Table 1. Minimum Hardware Recommendations

	Client
Operating Systems	64bit Windows Mac OS X (Java 1.8)
RAM	16GB recommended (minimum 8GB) solely for Client application
Disk Space	500MB

Table 2. AWS Server and Processing Unit Nirvana (optional)

	AWS Server and Processing Unit Nirvana (optional)
Operating Systems	CentOS 7 or 8 (Java 1.8)
Platform	EC2 type m4.xlarge (Includes RAM 16GB, Quad core)
EBS storage	Primary: VIA database storage. Size varies, recommended 100GB minimum Secondary: 500GB depending on throughput & sample data size
Additional Storage	S3

NOTE: File sizes can vary widely but here are some examples to help with estimating the amount of disk space needed: VIA database storage — 10–170MB per sample, S3 Bucket — Low-pass BAM file: 100MB, Targeted Panel BAM file: 2GB, WES BAM file: 5GB, 30x WGS BAM file: 30GB

Technical Assistance

For technical assistance, contact Bionano Genomics Technical Support.

You can retrieve documentation on Bionano products, SDS's, certificates of analysis, frequently asked questions, and other related documents from the Support website or by request through e-mail and telephone.

TYPE	CONTACT
Email	Software-support@bionano.com
Phone	Hours of Operation: Monday through Friday, 9:00 a.m. to 5:00 p.m., PST US: +1 (858) 888-7663
Website	www.bionano.com/support
Address	Bionano Genomics, Inc. 9540 Towne Centre Drive, Suite 100 San Diego, CA 92121