ESP Technical Specifications

Server sizing, software dependencies, and network considerations

Prepared by

Bob Zambarano

Revision History

Revision	Date	Author	Notes
1	April 7, 2017	B Zambarano	Initial version
2	February 2, 2023	J Miller/B Zambarano	Added minimum system requirements, info about VPN, other minor updates

Table of Contents

ESP Overview

ESP stands for **E**lectronic medical record **S**urveillance for **P**ublic Health. ESP is an open-source, free-license software originally developed under a CDC Centers of Excellence grant to the Harvard Medical School's Department of Population Medicine.

ESP performs notifiable and chronic disease case detection and reporting against a DataMart of patient clinical data. ESP deals only with clinical data including patient demographics, patient visit and visit diagnoses and details, medications, lab test results, immunizations, and social behaviors (such as smoking and alcohol use).

More Information is available at: http://www.esphealth.org/.
ESP Software source code is available for download at: https://gitlab.com/ESP-Project/ESP.

ESP is typically installed and maintained within the health care data partner's data center.

The standard model of operation is to extract all new or updated patient clinical data from the site's EHR (Electronic Health Record System) on a nightly basis and load these files into ESP for processing. ESP then uses a set of processes and disease- or condition-specific plug-ins to identify reportable disease cases, identify all relevant available data, generate an HL7 or ECR message, then encrypt and transfer this message to the Massachusetts Department of Public Health (MDPH). ESP also tracks what case data has been sent, so when additional information for a given disease case is collected for a patient, it will trigger the creation and transfer of a case update message to MDPH.

Once disease detection plugins are configured and validated, ESP requires very little maintenance beyond mapping new lab tests when these are added to the EHR system that feeds ESP.

Sizing Hardware or Virtual System

Linux System or VM (Virtual Machine):
  Ubuntu 18.04 or later   or
  RHEL (RedHat Enterprise Linux) version 8.6 or later
        For other Operating Systems, please check with Commonwealth Informatics (CII) for support
        and version requirements.
 Minimum Linux System Requirements:
  4 CPU
  16 GB Memory
  100GB Disk Space

Storage

At its core, ESP is a DataMart of patient EHR data. Storage requirements can be significant. These requirements primarily depend on the number of patients to be included in the ESP system, and the number of years of data the ESP system will support (going back historically and going forward for the planned life of the server; however, space can be added later to accommodate events such as Medical Center/Hospital mergers, etc.).

Solid State Drives (SSD) are preferred over standard Hard Disk Drives (HDD).

ESP Standard System sizing rule is:

50GB base + (1GB * Number of active patients in your system/5000) *Number of years of patient data kept in the ESP system

> For example: An organization with approximately 200,000 active patients intends to maintain up to 12 years of patient data in the ESP DataMart. This DataMart would require:
>
> 50Gb + (1GB * 200,000/5000) * 12 = 530Gb of storage.

A third factor to consider when sizing storage is the density of patient data- meaning, the number of distinct observations, orders and results collected per patient. If your organization deals with long-term care of very sick patients, patient data density will be much higher and will require more storage capacity. Alternatively, a health care organization that primarily deals with transient short-duration patients with limited care provision will have much lower data density and will require less storage.

ESP does not support any significant transactional processing in the database, so disk I/O performance need only be moderate.

Memory

A good rule of thumb for adequate performance is 2Gb of memory for each 50Gb of storage. A system with 500Gb of storage, for example, would require 20Gb of memory.

CPU

ESP can run disease detection algorithms in parallel threads, so more CPUs will increase performance up to the point where disk I/O or memory reach capacity. Two CPU cores are a minimal requirement. For each 100Gb of disk storage, an additional CPU core will be useful.

OS and Software Stack

ESP runs on Linux OS. It is developed on Ubuntu Sever LTS systems, and has been run on other Linux distributions including RedHat, SuSe, and CentOS.

A basic Linux server would need the following additional software:

- Administrative logins from anywhere, but the system console requires Open-SSH service.
- Network firewall rules should be in place to restrict access to the system to only required users and systems.
- Git software pkg is used for ESP distribution.
- PostgreSQL version 12 or later as the Relational Database Management System (RDBMS).
- ESP software is currently running on Python 3.6 and the Django 2.1 software stack. An ESP installation uses the Python virtual environment infrastructure.
- The administrative web interface uses an Apache/Http2 web server.
- The MDPH-provided encryption software requires Java.

Network Requirements and Data Security

ESP will store and provide access to HIPAA regulated PHI (Personal Health Information) data, so network design for data security issues is of critical importance. The ESP server must exist within a firewalled and access controlled network environment.

NOTE: VPN access is required for CII staff to the site-hosted ESP Server for initial Setup and Configuration. If VPN access is granted for ongoing maintenance and support, it removes that responsibility from the site, and allows CII to perform monitoring and troubleshooting.

ESP software and updates are distributed via bitbucket.org. A firewall rule enabling outgoing requests to GitLab and the Python Project are required to install and maintain the ESP and Python software. Please ensure these two entities are accessible:
https://gitlab.com/
https://pypi.org/

Other Linux package management sites may be required for access, depending on your system.

Data provisioning for ESP occurs via a set of delimited text files, generated nightly or weekly, so there must be a network connection between the ESP server and the machine that will generate the EHR extract files. These files can be pushed to the server or pulled down from a secure location.
Data flows to the ESP system should use some encrypted protocol (Open-SSH supported transfer encryption such as SFTP or SCP, or other custom built encryption based on asymmetric/publickey cryptography).

Data reporting from ESP to MDPH utilizes an MDPH-built encryption and transfer software. This will require an outgoing rule in the data center firewall to allow traffic from ESP to the MDPH reporting servers.