US Department of Health and Human Services

Privacy Impact Assessment

Date Signed:

09/24/2013

OPDIV:

FDA

Name:

CDRH High Performance Computing

PIA Unique Identifier:

P-9325113-332067

The subject of this PIA is which of the following?

General Support System (GSS)

Identify the Enterprise Performance Lifecycle Phase of the system.

Operations and Maintenance

Is this a FISMA-Reportable system?

No

Does the system include a Website or online application available to and for the use of the general public?

No

Identify the operator.

Agency

Is this a new or existing system?

Existing

Does the system have Security Authorization (SA)?

No

Indicate the following reason(s) for updating this PIA.

New system functions and revised security sensitivity.

Describe in further detail any changes to the system that have occurred since the last PIA.

Added scientific workstation virtualization

Describe the purpose of the system.

The FDA High Performance Computing (HPC) system is a wholly internal supercomputing environment that provides high-performance computational clusters specifically engineered to support large scale modeling and simulation projects needed by FDA scientists. The clusters enable analytic programs to run at a speed not obtainable with a scientific workstation. HPC also provides secure storage for the intermediate work products of research and development in FDA regulatory science. Additionally, it will store working data and intermediate work products of genomic analyses performed in support of certain regulatory cases.

Describe the type of information the system will collect, maintain (store), or share.

HPC holds scientific de-identified reference libraries, mainly genomic files, and provides secure storage for the intermediate work products of regulatory science analytical efforts (e.g., trial design, modeling and simulation, computational, statistical and scientific research data analysis). It will also store the working data and intermediate work products of genomic analyses performed in support of regulatory processes. HPC does not collect, store, use, maintain, or disseminate PII.

Provide an overview of the system and describe the information it will collect, maintain (store), or share,

The HPC system includes two semi-autonomous computational clusters with large storage capacity designed to scale (expand) to multiple Petabytes and beyond. Control and administrative functions for both clusters are hosted in the FDA's Center for Devices and Radiological Health (CDRH) Scientific Computing Lab. All users are registered on the FDA network and access the system from inside the FDA firewall. Scientists configure their applications to leverage the high-performance resources of the clusters.

The HPC enables CDRH to rapidly run analytic programs. The system’s immense storage and computational resources may hold de-identified genomic and other reference libraries and provide secure storage for the intermediate work products of both regulatory science research (e.g., trial design, modeling and simulation and research data analysis) and regulatory case work. Certain applications hosted by the HPC will download publicly available genomic reference data from reputable and secure genomic repositories such as the National Center for Biotechnology Information (NCBI) at the National Institutes for Health (NIH) and will subsequently contribute analytical results, known as annotations, to those repositories.

Does the system collect, maintain, use or share PII?

No