HPC as a service

CI: X

ARCHITECTURE DEFINITION DOCUMENT

DOCUMENT PURPOSE

The Architecture Definition Document (ADD) is the deliverable container for the core architectural artifacts created or updated during a project. The Architecture Definition Document spans all architecture domains (business, data, application, and technology) and captures always the target architecture. The current and previous states are managed via the change management and version control of the document.

It is suggested that this document reference the various deliverables in the container. For instance, the Architecture Principles, Standards and Guidelines are documented separately and those documents are referenced here.

In order to achieve a common understanding of the architecture, the contents of this document are intended to be read, updated, reviewed and approved by all members of the architecture task force dedicated to project, the product managers, as well as the solution manager responsible for the support activities.

**ADD ownership**

Regarding any new asset, the implementation product manager initiates the ADD and ensures it’s duly updated during the release lifecycle by the responsible contributors. Hence the overall accountability for the document lies with the lead TPM (Technical Product Manager). However, responsibility for contributing to the specific sections are listed as follows

|  |  |
| --- | --- |
| **Document Section** | **Accountable** |
| Business Architecture (1-5) | Lead TPM & Lead Business Solution Architect |
| Application Architecture (6) | Lead SW Architect |
| Technology Architecture (7) | Experts from relevant Infrastructure/Platform/Technology/Support architecture domains |

The duly filled ADD is mandatory for the transfer of the asset from release to support organization and from then on it is responsibility of the solution manager to keep the document up to date to reflect any architecture-relevant changes.

When a new release related to the application is launched, the implementation product manager takes ownership of the concerned ADDs and ensures that it’s duly updated throughout the release lifecycle and hands it back over to support at the end of the project.

**Document guide**

The text in bold blue italic is meant to serve as a guide for the users of this document. It can be later removed for the sake of readability.

Please don’t delete sections which are not required by the Project! Just state them as “not applicable” or “n.a.” in the section text.

Where required, sections can be added.

Sections with newly introduced capabilities should be highlighted in the document with a yellow color to better identify them.

This aids the reader in quickly discovering the relevant sections where changes to the architecture have been introduced.

* Note that this rule can be ignored if this is the very first release or the baseline architecture documentation isn’t available for the baseline reference.

**TABLE OF CONTENTS**

[1 Administrative Information 5](#_Toc487813384)

[1.1 Document Classification 5](#_Toc487813385)

[1.2 Document History 5](#_Toc487813386)

[1.3 Product Team Structure 5](#_Toc487813387)

[1.4 Architecture and Design Team Structure 5](#_Toc487813388)

[2 Objectives 6](#_Toc487813389)

[2.1 General (Vision/Program level) 6](#_Toc487813390)

[2.2 Release specific 6](#_Toc487813391)

[3 Compliance 7](#_Toc487813392)

[3.1 Architecture Principles, Policies and Standards 7](#_Toc487813393)

[4 Architectural Risks & Issues 10](#_Toc487813394)

[4.1 Assumptions 10](#_Toc487813395)

[4.2 Risks 10](#_Toc487813396)

[4.3 Technical Design Considerations & Issues 10](#_Toc487813397)

[4.4 Criticality, EC, Data Classification, User Access & SOX 10](#_Toc487813398)

[4.5 Design Exception 11](#_Toc487813399)

[5 Business Architecture 12](#_Toc487813400)

[5.1 Business Processes 12](#_Toc487813401)

[5.2 Organization Structure and Units 12](#_Toc487813402)

[5.3 Process Allocation 12](#_Toc487813403)

[5.4 Data Architecture 12](#_Toc487813404)

[5.5 Supported Business Process Capabilities 13](#_Toc487813405)

[6 Application Architecture 14](#_Toc487813406)

[6.1 Key Design Decisions (KDD) 14](#_Toc487813407)

[6.2 Business Applications 14](#_Toc487813408)

[6.3 Application Component Architecture 15](#_Toc487813409)

[6.4 Application Interfaces 18](#_Toc487813410)

[6.5 Software Versions 20](#_Toc487813411)

[6.6 Software Licensing 21](#_Toc487813412)

[7 Technology Architecture 22](#_Toc487813413)

[7.1 Datacenter Technology Architecture 22](#_Toc487813414)

[7.2 Workplace 26](#_Toc487813415)

[7.3 Security Architecture 27](#_Toc487813416)

[7.4 Network Architecture 28](#_Toc487813417)

[Appendix I. Release References 29](#_Toc487813418)

[Appendix II. External References 30](#_Toc487813419)

[Appendix III. Standards & Guidelines 31](#_Toc487813420)

[Appendix IV. Glossary 32](#_Toc487813421)

# Administrative Information

## Document Classification

|  |  |
| --- | --- |
| **Does this Document contain EC or UK MoD sensitive data?**  **If yes then it cannot be uploaded to Box and non-EC personnel should not access it** | **NO** |

|  |  |
| --- | --- |
| If **YES**, select Data Type (US Vault / UK Bluenet / US P&W EC) |  |
| Country hosting the sensitive data: |  |
| **Impact** (What are the restrictions for viewing, attaching or editing the TAD) |  |
| Data Classification Level:  ***Please note:*** (Externally facing applications or applications that are critical for business continuity require SCABBA and SSL (https) ) | Choose an item. |

## Document History

| Revision | Author | Date | Brief Description of Change |
| --- | --- | --- | --- |
| 1.0 | Krutika Thorat | 08/24/2024 | Initial draft |
| 2.0 | Krutika Thorat | 09/18/2024 | Updated Risks, architecture diagram, network design |
|  |  |  |  |

## Product Team Structure

List the relevant team members (Architects see below) and release identification details in this sub-section. These members will be granted read-access to this Architecture document stored in the Central Architecture GE Box folder.

| Roles | Name |
| --- | --- |
| GE Business | GE Gas Power |
| Business PM |  |
| IM PM |  |
| ServiceNow Project Number | **GERITM46855849** |
| dPMM Project ID |  |

## Architecture and Design Team Structure

List the Architecture members of the project, who contribute to this Architecture document (see introduction page above for details). These members will be granted write-access to this Architecture document stored in the Central Architecture GE Box folder.

| Roles | Name |
| --- | --- |
| Enterprise Technology Leader |  |
| Solution Architect |  |
| Application Architect |  |
| Data Architect |  |
| Application Designer |  |
| Technical Lead |  |
| Module Lead |  |

# Objectives

## General (Vision/Program level)

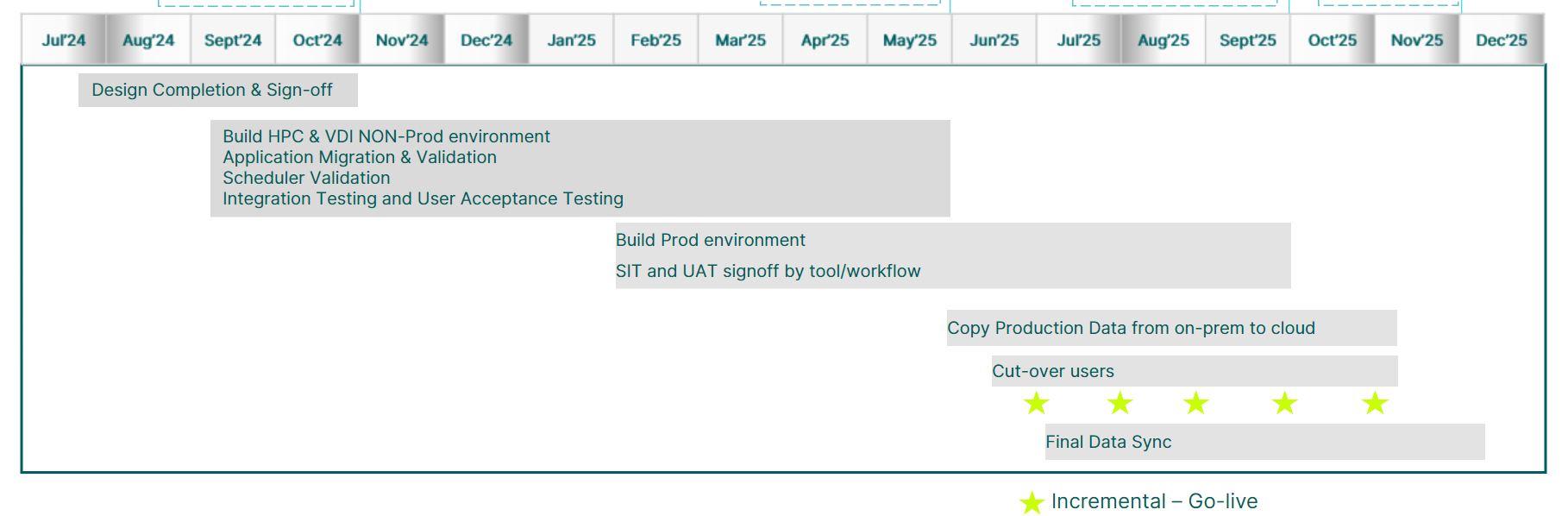
Summarize the general context of which this release is covered (initiative, program etc.). Document the overall scope and schedules of the general context. Also, indicate where this release fits in or delivers capabilities to the greater picture (strategic or enterprise level).

High Level Purpose:  The program aims to establish a cloud-based high-performance computing (HPC) environment. This initiative includes the development of a dedicated cloning/building of an Engineering Compute platform, enabling Vernova to run engineering simulations. This cloud HPC will facilitate complex engineering simulations, such as thermal dynamics, product definition, and mechanical component analysis, empowering engineering teams to enhance their capabilities and efficiencies. By leveraging advanced computational power, we can accelerate innovation cycles and improve the accuracy of our simulations, leading to better product designs and faster time-to-market.

### *Product Goals*

* Deliver a high-performance computing platform for GE Vernova engineering teams to run demanding applications.
* Ensure business continuity by migrating to a functional HPC environment.
* Establish an AS-IS HPC setup in the cloud.
* Enable job submission via a workflow scheduler and command-line interface.

### Capability Timeline



## Release specific

Describe the scope, deliverables, and the timeline the release. Outline the benefits and capabilities (business and technology like) delivered with this release to the customer.

The release is scheduled for October 2025 and will be incremental. Users will be migrated from the on-premises system to the cloud in batches, with access to the on-premises system restricted. The production build will also follow an incremental approach; applications and integrations tested in the build environment will be transitioned to production as they become ready.

* Initiate non-build – sept’24
* Initiate build- Feb’25
* Initiate Cut-over process – Jun’25
* Cut-over Complete – oct’25

# Compliance

## Architecture Principles, Policies and Standards

Refer to the GE architecture principles documentation (principles, policies and standards) which influence the architecture decisions. Outline deviations and exceptions from architectural standards and company guidelines in section 4, if any.

### General

Each project managing a change of the architecture design must follow GE standards. They are briefly described and available in SharePoint. The related link can be opened [here](https://teamspace.alstom.com/sites/ISTProjects_Team/Project%20and%20Portfolio%20Management/SitePages/PACT%20and%20Common%20Framework.aspx).

All given GE standards and policies must be met during the definition and design of the solution (see GE regulations in the appendixes for further details).

Special focus should be laid on the GE security and legal requirements to avoid data infringement or leakage by unauthorized parties by design.

Beside that the following architecture principles have to be reflected in the definition and design of the solution:

* Implement concepts of an Centralized architecture (includes utilization of strategic data centers)
  + Lowering running costs due to reduced server landscape complexity
  + Decreasing deployment efforts and the time to deliver
  + Stackable server base to enable efficient capacity management and to respond on changing resource requirements
* Virtualization over Physical servers
  + Where possible use virtual solutions (VM’s etc.) to implement applications and capabilities
* Cloud
  + Where feasible consider cloud technology to efficiently drive the solution
* OOTB Functionality
  + Realize as close as possible to the standard functionality to minimize customization efforts and to keep the impact of future PLM tool upgrades on the implemented code base manageable
  + Re-use of existing code portions where it makes sense
* Re-usability of existing components and GE technology stack first
  + Utilize existing GE technology stack and components in the solution
  + When creating new components with a release, assure that it’s ready to be re-used
  + Always favor standard technology, framework and protocols over non-standard implementations or solutions

### Security

Information Systems have a pre-eminent role in our daily operations and significantly contribute to the achievement of our strategic priorities. The main goal of the GE Security Policy is to maintain the Confidentiality, the Integrity, and the Availability of information while ensuring compliance with legislative, regulatory and contractual requirements.

GE is committed to deploy and follow its Security Policy which encompasses core security principles, security organization and based on a set of global security requirements.

Please refer to the following links for the [GE Information Security and IT Risk Management Policy](http://libraries.ge.com/download?fileid=660021302101&entity_id=53421501101&sid=101) and [GE Access Management Policy](http://libraries.ge.com/download?fileid=660173948101&entity_id=53421511101&sid=101).

### Export Control and IP

Excerpt taken from the “GE Export Enterprise Standard” document available on the [GE Power International Trade Controls](http://supportcentral.ge.com/products/sup_products.asp?prod_id=25521) website:

“The global movement of goods, technology and software is an important component of global business and national security interests. Export control laws regulate this activity, which may include domestic transactions, affecting many of GE’s operations from shipping products to collaborating across national boundaries. As a business with operations across the globe, GE is committed to complying with all applicable export control requirements including economic sanctions. It is the responsibility of all GE businesses and employees to comply with applicable export laws, regulations and GE policies.

…

The GE Export Controls Enterprise Standard (“Export Controls Standard”) contains minimum requirements for the execution of essential tasks required to request and use export authorizations and to ensure “regulatory excellence” in dealing with applicable export controls regimes…”

Please also refer to the [International Trade Compliance Center if Excellence (ITC CoE)](http://libraries.ge.com/FoldersData/51643647101/ITC%20CoE%20Website%20Library/ITC%20CoE%20Website%20Code%20(LIVE)/?fileid=650087307101&entity_id=51643427101&sid=101&version=0&encURL=Yes&fileid=650087307101&entity_id=51643427101&sid=101&version=0&encURL=Yes&fileid=650087307101&entity_id=51643427101&sid=101&nocache=Wed%20Sep%2016%2016:16:45%20EDT%202015&nocache=Wed%20Dec%2002%2017:38:30%20EST%202015&nocache=Fri%20Apr%2008%2006:32:27%20EDT%202016) for all related details regarding the Export Control (EC) and Intellectual Property (IP) regulations and resources.

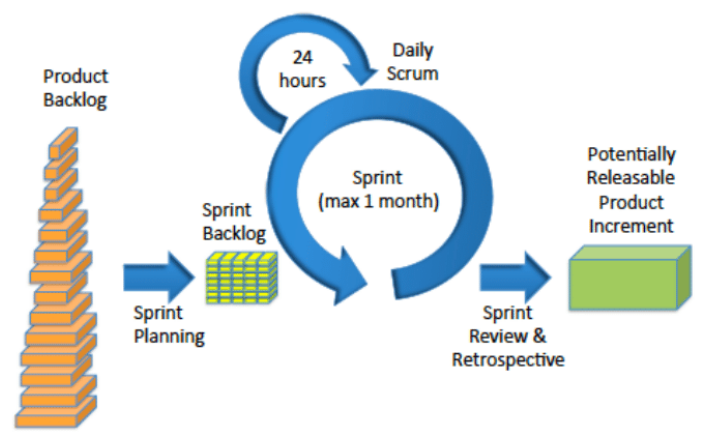
### Data Classification and Protection

Document classification procedures are an important element of an effective document management and information security program. With reference to the GE IT policy D13.02.013 for “GE Data Protection Standard Effective” (follow this [link](http://libraries.ge.com/download?fileid=664220557101&entity_id=53714827101&sid=101)), “*GE information must be classified in accordance with the GE Data Classification Policy as GE Restricted, GE Confidential, GE Internal, and Public – GE Primary Classification.*”

Also review the GE Data Classification guide (click this [link](http://libraries.ge.com/download?fileid=241135209101&entity_id=20820418101&sid=101)) for further guidance on classifying information generated or used by the Company. See Appendix III “Standards & Guidelines” for more references.

### SCRUM Release Process

The below depicted high-level SCRUM Release Process represents the actual product driven release process used in GE. Note that it replaced the obsolete [simplyPM process](https://devcloud.swcoe.ge.com/devspace/display/YFBSU/simplyPM+How-To+Documents)

. 

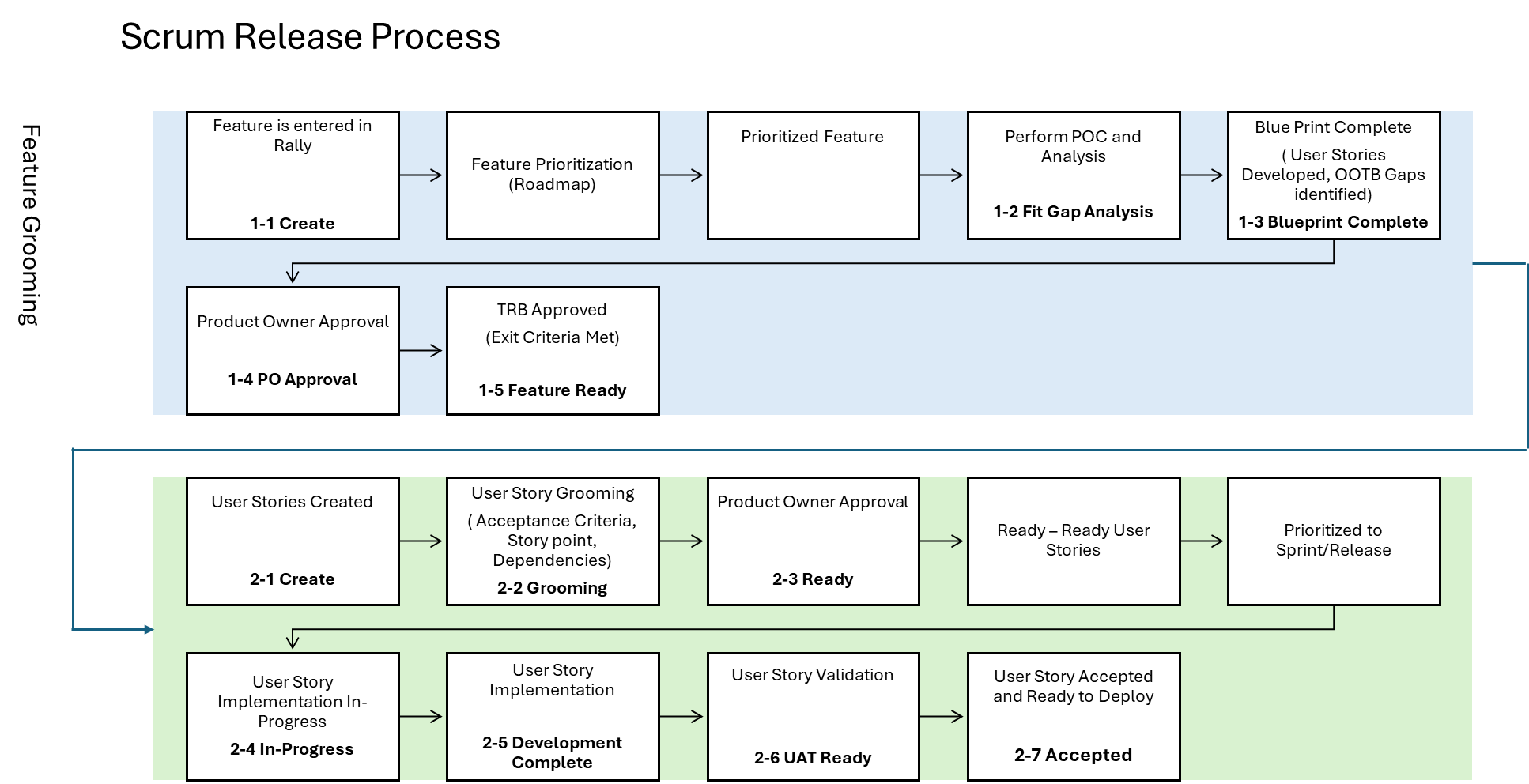


Figure 1: SCRUM Release Process

The blue colored process area is where the Blueprint team determines and defines the vision and its prioritized features, listed in the backlog. The green colored area shows the process steps owned by the SCRUM team, which define and implements the user stories per release, based on the prioritized features taken from the backlog.

The following table lists the deliverables per process steps for the SCRUM team (per the process numbers from Figure 1).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Type | Status | Deliverable(s) | Contributor | Responsible  (Performs/Completes the activity) | Accountable  (Ownership of Activity) | Approver |
| SW Factory | **User Story** | 2-1 Create | High Level User Stories being detailed in Rally |  | Process Area - Process Lead | Lead TPM | Lead TPM |
| 2-2 Grooming | All Details of the user story are captured in Rally | Process Lead | Process Area - TPM | Process Area - Process Lead | Lead TPM |
| User Story Splitting | Scrum Team | Process Area - TPM and TA | Process Area - TPM | Process Area - TPM |
| High Level Design |  | Process Area - TA | Lead Solution Architect | Lead Solution Architect |
| Proof Of Concept (Spike) are required | Scrum Team | Process Area - TPM | Lead TPM  Lead Solution Architect | Lead TPM |
| Data Migration Strategy (if applicable) |  | Data Architect | Lead Solution Architect Process Area - Process Lead | Lead Solution Architect |
| Acceptance Criteria | Scrum Team | Process Area - TPM | Process Area - TPM Quality Lead | Process Area - Process Lead Quality Lead |
| User Story Point Estimation | Scrum Team | Scrum Team Process Area TPM and TA | Scrum Team | Scrum Team |
| 2-3 Ready-Ready | All Details of the user story are captured in Rally | Process Lead | Process Area - TPM | Process Area - Process Lead | Lead TPM |
| Common understanding about Design, Development, Testing of the user story | Scrum Team | Process Area - TPM and TA | Process Area - TPM | Lead TPM |
| User Story Point Estimation | Scrum Team | Scrum Team Process Area TPM and TA | Scrum Team | Scrum Team |
| 2-4 InProgress | NO DELIVERABLE, only STATUS Representation. Indicates Start of the Implementation | Scrum Team | Scrum Team | Scrum Team | Scrum Team |
| 2-5 Dev. Complete | Code Complete and Checkin into Branch | Scrum Team | POD Team Developer | Scrum Team | Developer Lead |
| Automated Code Review performed, and all violations fixed | Scrum Team | POD Team Developer | Scrum Team | Developer Lead |
| Document Low Level Design | Scrum Team | POD Team Developer | Scrum Team | Developer Lead |
| Performance Load scripts development | Scrum Team | POD Team Developer | Scrum Team | Developer Lead |

Table 1: Deliverables from the SCRUM Team

# Architectural Risks & Issues

This section covers all the assumptions, risks and issues from the architecture perspective. Relevant topics from the release viewpoint will be covered within its risk register.

## Assumptions

Include all architectural relevant assumptions for the release.

|  |  |  |
| --- | --- | --- |
| **ID** | **Assumption Item** | **Description** |
| 1 | Active Directory Integration is possible | We assumed that by implementing the AD integration we can allow users to access files of Linux & windows with same username |
| 2 | Assumptions associated with Risks | We assume that all the risks will be mitigated |
| 3 | Development of frontend of scheduler | Job needs to be submitted via CLI (using scripts) & workflows |
| 4 | 4 months for Data Migration (Jun’25-Sept’25) | Time to transfer data between on-prem netApp to FSx. Snowball\* audit by GRC/Aerospace. |
| 5 | SnapMirror technology can transfer huge data volume | We assumed that SnapMirror will be a suitable technology to transfer the data but we aren’t sure about performance & approval on SnapMirror from GRC |
| 6 | There is no need to exchange data between IND & US users | We have received the statement from Brandon, but we are skeptic about this scenario |
| 7 | Only US citizens will part of SUDO usergroup | The user with SUDO access but not US citizen can access the EC data, which will have org level impact |
| 8 | Non-EC data will be available for non-PROD test | A subset of real case data is required for testing non-prod env. We can’t test with dummy data only. |
| 9 | Malaysia Latency will be >50ms | Due to availability of POP (in Singapore) & instances (in Malaysia) we are going ahead with Malaysia region , and we assume to get request route time from Bangalore to Malaysia below 50 ms |

## Risks

List all architectural risks detected and document impact, measures and mitigation plan. Don’t mix it up with release risks, as architectural risks contribute to the release risk list. Regularly review and update this architectural risk list.

To get latest updates in Risk please visit Rally - <https://rally1.rallydev.com/#/733053585421d/custom/601725600840>

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Date Created** | **Risk Category** | **Description** | **Probability** | **Impact** | **Criticality** | **Mitigation plan** | **GEV Owner** | **AWS Owner** |
| 28-08-2024 | Windows VDI | Non-Compliance with MS Office on Windows VDI | 9 | 9 | 81 | 1. Determine GEV license agreement with Microsoft 2. If risks exists even after acting on pt 1 -Evaluate other MS office options on Amazon 3. Create a service which will allow to use macro 4. Economize compliant office licenses on dedicated instances | Sridhar R | Craig B |
| 28-08-2024 | Cyber - Compliance | Unauthorized access to data | 9 | 9 | 81 | 1. Utilizing export policies and permissions to authorize data 2. Utilize network groups 3. Potentially have 2 file systems for EC & non-EC 4. Use netApp features to hide files 5. Resue what we have currently in place -e.g Export control flag in network security groups | Sridhar R | Henry |
| 28-08-2024 | Windows VDI | Windows VDI solution not robust | 9 | 9 | 81 | 1. Update domain join process - check if API returns domain controller 2. Optimize the domain join script 3. For enginFrame - implement optimized domain join script | Mahesh D | Craig B |
| 28-08-2024 | Windows VDI | Unavailability of AMI for windows 10 desktop | 3 | 9 | 27 | Requirement - Dedicated host/instance if VDA License 1. Prioritization on AMI Creation with Laszlo 2. Leverage existing VMDK to create AMI | Sridhar R | Craig B |
| 28-08-2024 | Cyber - Identity | Not using right protocol due to domain structure (using old NTLM protocols) | 3 | 9 | 27 | 1. Administration team should provide evidence on audit 2. Protocols & versions must be audited on regular basis | Yasen H | Henry |
| 28-08-2024 | Execution | long-lead time in build | 3 | 9 | 27 | 1. Clear communications on deadlines & risks 2. Share project plans including sprint schedules 3. Prioritization on resource allocation | Venu T | Seamus O |
| 28-08-2024 | Regions | Unacceptable latency performance for IND HPC Users | 1 | 9 | 9 | 1. Early feedback from users on performance 2. Build test env for malaysia | Yasen H | Seamus O |
| 28-08-2024 | Scheduler | Not having strategic support agreement with vendor NI-SP | 1 | 9 | 9 | 1. Contract between GEV & NI-SP to support enginFrame issues 2. Use AWS support for enginFrame issues till Sept'25\*(TBA) 3.Bring GEV's own orchestration - myHPC 4. Option 2 (backup) - bringing RES + skybridge in picture | Gabriele G | Sandeep S |
| 28-08-2024 | Cyber | Non-adherence to Cyber Standards - architect , code etc to meet standards | 1 | 9 | 9 | 1. understand requirements & work closely with cyber 2. act beforehand…let's not wait for last minute | Yasen H | Blake |

## Technical Design Considerations & Issues

Summarize any specific technical design decisions made in this release that are considered as deviations from design standards and the rationale for using them. Refer to the attached checklist to assist in determining whether the design decisions are considered exceptions. Each issue should be clearly documented in the table below.

Link to technical specifications document - <https://gevernova.ent.box.com/file/1667318738170>

## Criticality, EC, Data Classification, User Access & SOX

* Business Criticality Tier: Mission Critical
* Disaster Recovery Tier: Tier 1

**Note***: If the* ***Business Criticality Tier*** *differs from the* ***DR Recovery Tier****, an exception may be required. If this exception is appropriate, obtain business* *owner signoff and attach the approval.*

Application Export Compliance: non-EC

GE Data Classification: GE Confidential

User Access (External Facing?): Not - External (Internal)

**Note***: Externally facing applications or applications that are critical systems for business continuity require SCABBA and SSL (https URL)*

SOX Application:  **No**

**Project URL in service-now** **- GERITM46934737**

Design Exception - So far No design exception has been identified (There is possibility that we discover a few in future. example – AD integration and those will be explained in RISKs)

Is one or more PT Design Exception required? No

If YES, provide the links to any PT Design Exceptions for the project: < *insert* ***link*** *here*>

# Business Architecture

This is an abstract summary taken from the BRD (Business Requirement Documentation). Provide the link to the BRD as reference.

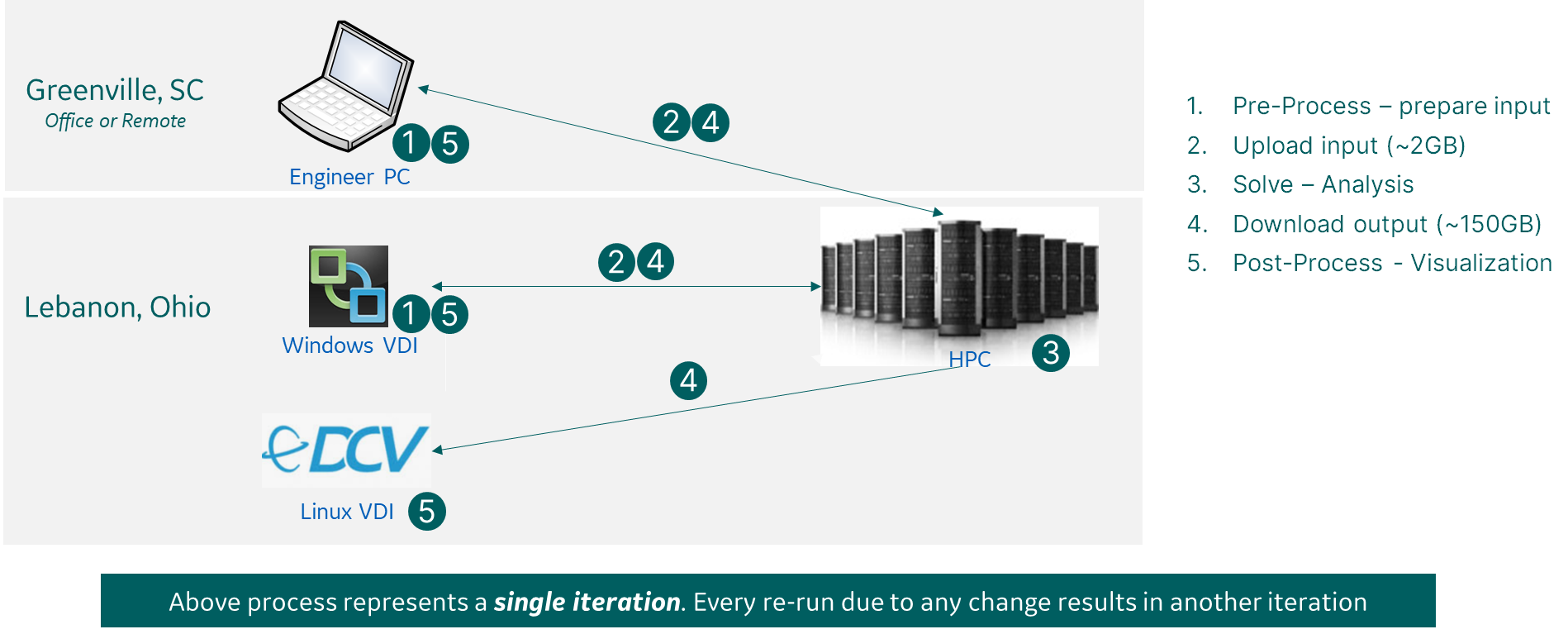
Document the Business Architecture to outline the transition from As-Is (baseline) to the To-Be (Target) architecture.

Summarize any transition steps required. Ensure the described To-Be (Target) architecture covers the desired delivery scope.

## Business Processes

### Baseline Business Processes

Describe the As-Is Business Architecture with its business process, data model etc. in scope.



|  |  |
| --- | --- |
|  |  |

### Target Business Processes

Describe the To-Be Business Architecture with new, updated or obsolete business process, data model etc. in scope of the release. Highlight the deliverables from the business scope point of view.

## Organization Structure and Units

Name the affected organization structure/units, the locations and user number. Describe any impact to these items. Try to determine the potential increase of the user number for the next year or two. If there are different roles or user types (designer, manufacturer etc.) involved, try to difference them here.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Location Name** | **Active Directory Domain**  **(LOGON & RD for EC are standards)** | **Other Domain (describe)** | **Number of Users** | | **2 Year User Projection** | |
| **Total** | **Concurrent** | **Total** | **Concurrent** |
| Greenville USA | LOGON (logon.ds.ge.com) | None | 500 (HPC + VDI) |  |  |  |
| Bangalore IND | LOGON (logon.ds.ge.com) | None | 572 (HPC + VDI) |  |  |  |

## ~~Process Allocation~~

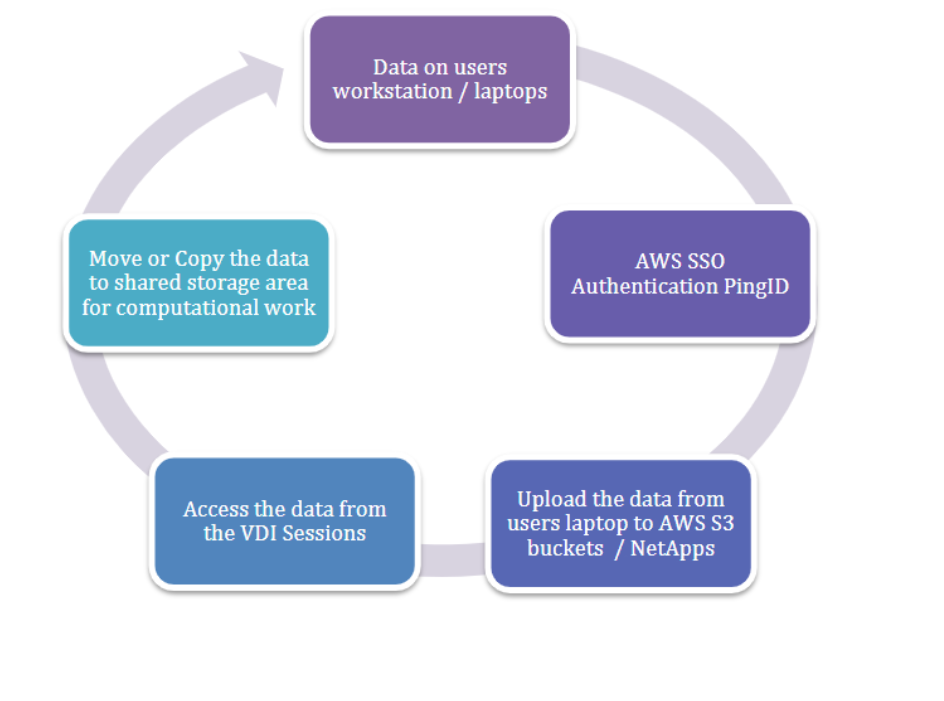
Map business processes in scope to the locations under deployment. Key locations can be represented geographically, functionally, or structurally.

| **Location Business Process** | **Location 1**  **Bangalore** | **Location 2**  **Greenville** | **Location 3** | **Location 4** | **Location 5** | **Location 6** |
| --- | --- | --- | --- | --- | --- | --- |
| Process 1 | X | X | X | X |  |  |
| Process 2 | X | X |  |  | X | X |
| Process 3 | X |  |  | X | X | X |
| Process 4 | X | X | X | X | X | X |

## Data Architecture

### High-Level Business/Logical Data Model

Describe on a high-level the underlying Business Data Model provided with the release (use diagrams for the overview).



### Business Data Objects

Provide the definition for each Business Data Object (BDO) in the scope. Where known, state the master repository (single point of truth) for each BDO listed.

| **Business Data Object ID** | **Business Data Object** | **Business Data Object Description** | **Master repository when available** |
| --- | --- | --- | --- |
|  |  |  |  |

Please refer to the Business Data Objects Classification (click this [link](https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/Taxonomies/SitePages/Home.aspx)) for a complete classification list. See Appendix III “Standards & Guidelines” for more references.

## Supported Business Process Capabilities

Describe the business capabilities delivered/supported with this release (business process capability mapping).

* A high-performance computing platform for GE Vernova engineering teams to run demanding applications.
* Business continuity by migrating to a functional HPC environment due to TSA exit
* Job submission via a workflow scheduler and command-line interface.

# Application Architecture

This section represents the content from the TDD (Technical Definition Document). Note that the focus laid here onto the logical architecture model level, not the physical implementation one (this will be covered with the Technology Architecture section below).

Describe the Application Architecture to outline the transition from As-Is (baseline) to the To-Be (Target) architecture. Summarize any transition steps required. Ensure the described To-Be (Target) architecture covers the desired delivery scope. Highlight which application components support which business capabilities. Use diagrams to outline the solution.

## Key Design Decisions (KDD)

Summarize the Business and Application Key Design Decisions (KDD) taken for this release in the table below (please add references to RALLY or ALM for further details if required). Don’t forget to review the Risk Register after KDD’s added.

**This section is covered in KDD 7.1.1**

## Business Applications

A Business Application usually is the overall service delivered to the business/customer. Some releases affecting or including multiple Business Applications. In such case it is important to describe the dependency/integration.

List of Business Applications with CI Details : <https://gevernova.box.com/s/997u6tw83ktntrz9jvxcrzqqc21q2gzl>

### Business Applications in scope

Provide definitions for the Business Applications in scope. Map to the AIP/planET repository where applicable.

List of Business Applications with CI Details : <https://gevernova.box.com/s/997u6tw83ktntrz9jvxcrzqqc21q2gzl>

### Mapping of business applications to supported business processes

List the Business Processes supported by the Business Applications in scope.

List of Business Applications with CI Details : <https://gevernova.box.com/s/997u6tw83ktntrz9jvxcrzqqc21q2gzl>

### Business Services

Provide definitions for the Business Services in scope. Map to the ServiceNow repository where applicable.

| **Business Service ID** | **Business Service Category** | **Business Service** | **Business Service Description or the link to the existing document** |
| --- | --- | --- | --- |
|  |  |  |  |

### Business Service Contracts

Provide definitions for the Business Services Contracts in scope. Map to the ServiceNow repository where applicable.

| **BS Contract ID** | **Business Service Contract Name** | **Business Service Contract Description or the link to the existing document** |
| --- | --- | --- |
|  |  |  |
|  |  |  |

## Application Component Architecture

Describe the breakdown of the Application Architecture into its Application Components on a logical level. Show the transition from As-Is to the To-Be Architecture (where applicable). List all software components, the capabilities delivered and required software licenses.

### Baseline Architecture (Logical Model)

Summarize the As-Is Application Components Architecture on a logical level. Let understand in short words & diagrams where this release is starting from.

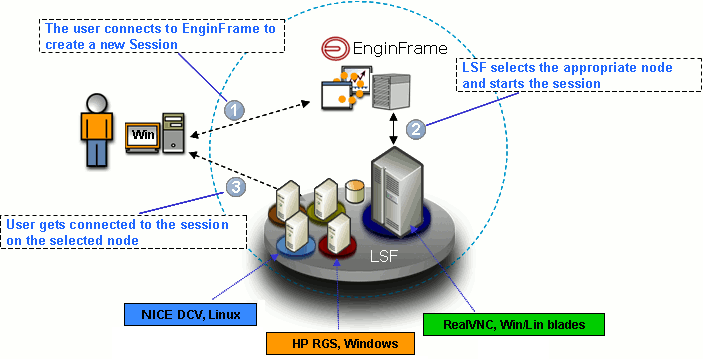


Figure 2: Example Baseline Application Architecture (Logical model)

Below table summarizes the major application components:

| **Application Component ID** | **Application Component Name** | **Application Component Details** |
| --- | --- | --- |
|  | EngineFrame |  |
|  | NICE DCV |  |

### Target Architecture (Logical Model)

Summarize the To-Be Application Components Architecture on a logical level. Let understand in short words & diagrams what this release wants to deploy.

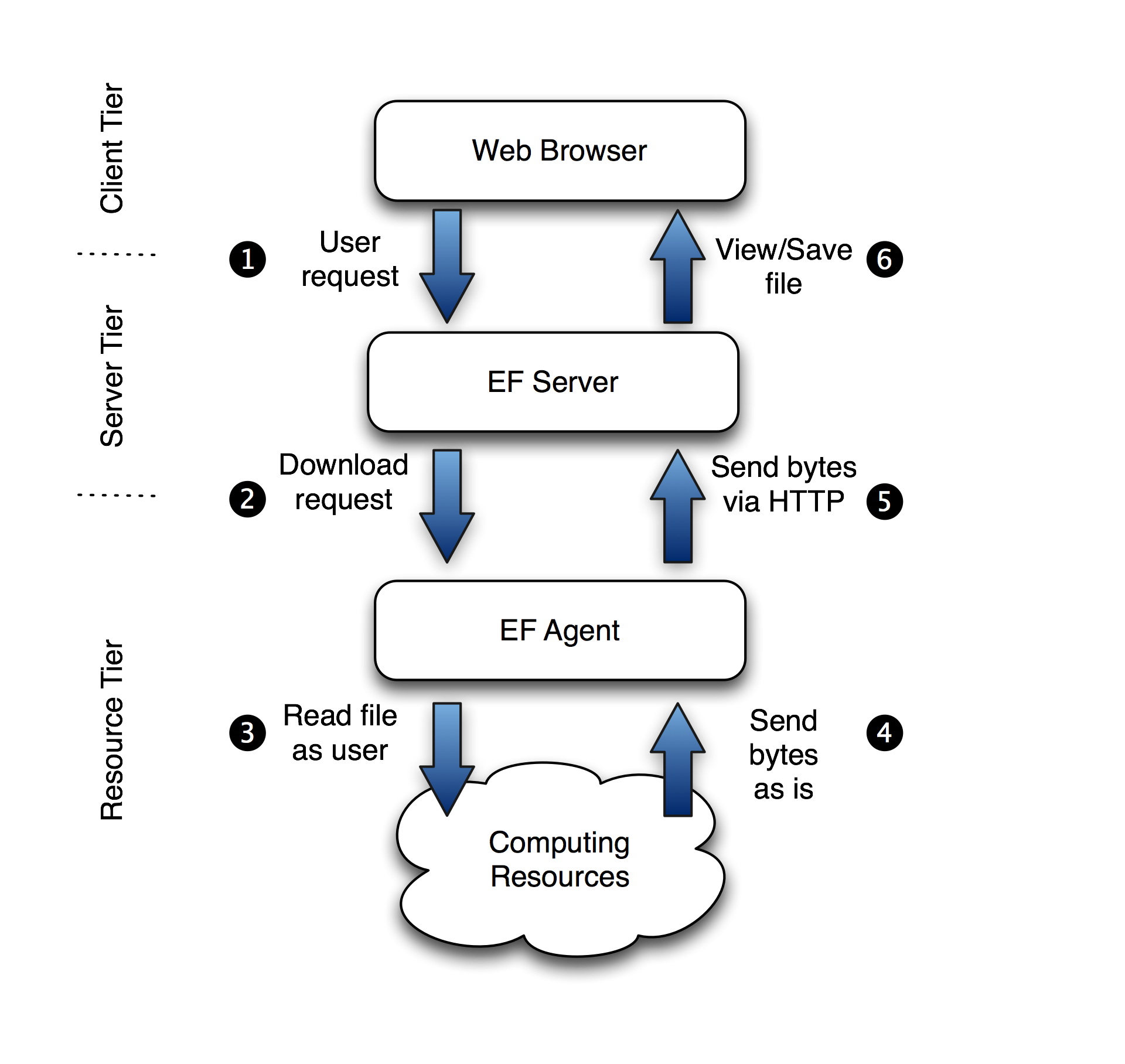


Figure 3: Example Target Application Architecture (Logical model)

| **Application Component ID** | **Application Component Name** | **Application Component Description** |
| --- | --- | --- |
|  |  |  |

### Reusable Components

Identify business specific common components used in the Application design from the list below. Please make sure you use the Application Development Toolkit available.

### Persistence Framework

Describe the persistence framework (custom built or third party) used in this release.

For example, does the design use an Object Relational Mapping (ORM) framework, Java Data Object framework (JDO), Data Access Object framework (DAO), LDAP to Object framework (e.g. Castor JAX), etc.?



Figure 4: Example Persistence Framework

#### Caching and Performance

Describe the caching mechanism of the application and how it delivers to the overall service performance.

#### Pooling

State the pooling mechanisms used in the design. For example, LDAP, EAI, Messaging, Object or Database connection pooling?

#### Thread Management

Describe the memory management mechanism for any component.

### User Interface Design

TG 3: Design: BI Doc: UI Doc- Reports

Describe the design of any user interface screens used with the project. Follow the standards described in the linked pages. Fill out a copy of the design template and attach it below.

UI Standards: <http://genet.ae.ge.com/ORG/ebats/cto/designcenter/ui/guide/index.html>

UI Design Template: <http://data.supportcentral.ge.com/books/17823/b15013/doc_239904.doc>

#### Presentation Framework

Document any 3rd party presentation frameworks used in the design (e.g. Apache Velocity, Tiles, etc.). Specify any non-standard User Interface components such as Applets, Active X Controls, Framesets, etc. Also, list any tag libraries used in the design.

### Reporting Solution Design

TG 3: Design: Appd: Reporting Design

Describe the design used for any external reporting solution used with the release (e.g. Cognos, Business Objects, etc.). Explain the reason for choosing a particular reporting solution. Attach any relevant templates for these designs.

If application decides to use high end reporting solutions like Cognos or Business Objects make sure design confirms to standards set for that tool.

**CloudWatch reporting details will be updated here.**

#### Business Intelligence

Enter the required information to outline the applied BI method. Make sure to have it aligned with the listed solution in this section.

Business Intelligence platform: Choose an item.

If BI platform selection is *other*, please describe: < *insert* ***BI platform*** *here*>

**Note:** *Tableau is the strategic choice (leverage existing GE Power solutions & shared platform capacity)*

*Cognos and OBIEE are in sunset mode and will require a Design Exception.*

*Note: Costing is liquidated by the BI Team and is not required to be submitted by the SA*

Has the BI Team Lead for your P&L been engaged? Choose an item.

Has the PT BI Team been engaged (by the BI Team Leader)? Choose an item.

***REFERENCE:***

* *Power Technology Tableau Portal:* [*http://supportcentral.ge.com/products/sup\_products.asp?prod\_id=320692*](http://supportcentral.ge.com/products/sup_products.asp?prod_id=320692)
* *Tableau Reference Architecture: https://ge.box.com/s/8unjrt5qsist98h8qk2hzdzxnrb4jsmj*
* *Cognos:* [*http://supportcentral.ge.com/products/sup\_products.asp?prod\_id=14983*](http://supportcentral.ge.com/products/sup_products.asp?prod_id=14983)
* *Informatica:* [*http://supportcentral.ge.com/products/sup\_products.asp?prod\_id=271414*](http://supportcentral.ge.com/products/sup_products.asp?prod_id=271414)
* *Informatica Reference Architecture:* [*https://ge.box.com/s/gjkfhyag9ykgrhaw39uarjgglabxdttc*](https://ge.box.com/s/gjkfhyag9ykgrhaw39uarjgglabxdttc)

### Database Design

The Database Technical Design covers the complete data model and will be reviewed by the TSG Data Architecture team. Please link to or embed the ERWIN diagram of the DB design used by the application.

For eMatrix applications, explain in detail the eMatrix Objects that will be used.

**These details are covered under 7.1.3 Database Architecture**

#### Data Model

TG 3: Design: Doc: Data Model

If a Data Model is being altered, provide the Physical Diagrams of the facts, dimensions, attributes, plus entities that satisfy reporting requirements.

Create and name as many modules as needed.

### ETL-BIO Design

Explain any integration design that is part of the system. Document the details of the Input and Output data structures. Document the details of the data processing including field mapping and transformations. The easiest way is to include references to external reference documents and spread sheets.

TG 3: Design: DM Update: ETL Design

Describe the ETL (Extract, Transform and Load) approach, "Source" to "Target" mapping, processes, data population strategies, load scheduling strategies, warehouse admin strategies.

Copy the <Integration Name> section for each Integration Function. Provide full detailed documentation of the information flow and transformation.

**Each one of the ETC/BIO application have very specific requirement on Technical design and have a technology specific template with all the required information. Provide link to ETL/BIO technology specific document(s) used (this technology document should not be submitted for approval until the required ETL/BIO specific document has not been completed):**

|  |  |
| --- | --- |
| **Webmethods** |  |
| **ETL/Informatica** |  |
| **Cognos** |  |

## Application Interfaces

Shortly summarize the today existing As-Is and to-Be Application Interface Architecture design (where applicable). List the interfaces, its coverage and technology. Describe strategies and design used for messaging integrations with other applications and database.

### Baseline Architecture (Logical Model)

Shortly summarize the existing As-Is Application Interface Architecture design.

### Target Architecture (Logical Model)

Shortly summarize the To-Be Application Interface Architecture design. List all interfaces which will be changed or being part of the transition.

Find below the summary of applied interface changes during the target architecture transition:

| **Interface ID** | **Interface Name** | **Sending Application Component 1** | **Receiving Application Component 2** | **Technology (Baseline)** | **Technology (Target)** | **Remark** |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |

In the following table all interfaces (including the internal ones) have been summarized:

| **Interface ID** | **Interface Name** | **Sending Application Component 1** | **Receiving Application Component 2** | **Frequency** | **Automation Level** | **Input and/or Output** | **Business data objects** |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |

The following table lists all interfaces with provided services consumed:

| **Operational Contract ID** | **Application Component 1** | **Application Component 2** | **Contract Description or link to the dedicated document** |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

The following below table summarizes the Application interfaces which are connected to external business applications.

| **Business Application ID** | **Business Application Name** | **Interface ID** | **Interface Name** |
| --- | --- | --- | --- |
|  |  |  |  |

### Data Integrations – High Level Requirements (Informatica & WebMethods)

Enter the required information to outline the applied integration method. Make sure to have it aligned with the listed integrations in this section.

1. **webMethods**

Provide the following details for each webMethods integration

Data flow #*n*

* + - Source: < *insert* ***Source*** *here*>
    - Target: < *insert* ***Target*** *here*>
    - Data volume: < *insert* ***Data volume*** *here*>
    - Frequency: < *insert* ***Frequency*** *here*>
    - Integration Size: Choose an item.

1. **Informatica**

New or existing folder required?Choose an item.

Application folder name:< *insert* ***folder name*** *here*>

*(Naming standard for application folder name: Business-INF-APPCI-Instance)*

*Example*: PGS-INF-MAXIMO1-PROD

***Note:*** *For new application folders, the CI needs to be created*

Provide the following details for each Informatica integration

Data flow #*n*

* + - Source: < *insert* ***Source*** *here*>
    - Target: < *insert* ***Target*** *here*>
    - Est. data flow size (MB): < *insert* ***data flow size*** *here*>
    - Frequency: < *insert* ***Frequency*** *here*>

Are there file transfer requirements (SFTP) with Informatica? Choose an item.

**Note:** *Flat file transfer require server / folder name location for source and target*

**Note:** *If the Informatica integration is with Oracle, Greenplum,Teradata or SAP ERP systems, the connector driver license is already in place. If not one of these, a discussion may be needed to determine if the generic ODBC driver is sufficient or if another driver is required, which may require an additional license agreement. In any event, please reach out to PT BI and Data Management Team for design considerations.*

**Note:** *If the business requires replication of ERP tables to the ODS in a real-time or near real-time basis, then Informatica CDC utilities should be employed. An additional license will be required. Please reach out to PT BI and Data Management Team for design considerations.*

1. **Other Integrations**

Choose an item.

## Software Versions

Provide the software versions used in the Application, including Databases and third-party software.

Find details here : <https://gevernova.box.com/s/997u6tw83ktntrz9jvxcrzqqc21q2gzl>

## Software Licensing

Provide the required software licenses to operate the Application, including Databases, OS and third-party software.

### Licensing

***REFERENCE:***

* [*SW Procurement Approval Process Link*](http://internal.energy.ge.com/srl/index.jsp)

# Technology Architecture

This section represents the content from the TAD (Technical Architecture Document). Note that the focus laid here onto the physical architecture model level, which materialize the logical model (described in the Application Architecture section before).

Describe the Technology Architecture to outline the transition from As-Is (baseline) to the To-Be (Target) architecture. Summarize any transition steps required. Ensure the described To-Be (Target) architecture covers the desired delivery scope. Highlight which logical application components will be realized through a physical one. Use diagrams to outline the solution.

The purpose of this section is to provide a high-level view of the technology architecture and provide references to the relevant technology architecture documentation (if exists) that complements this document.

## Datacenter Technology Architecture

Describe the physical representation of the Application Architecture and its Application Components on a infrastructure level. Show the transition from As-Is to the To-Be Architecture (where applicable). List all infrastructure and network components, the service and security requirements and the client delivery.

Purpose: Provide a scalable VDI and HPC architecture for users across the business to access HPC and VDI applications. It will allow access for both pre and post processing as well as simulation data on a shared file system to simplify file movement and improve end user experience. User jobs will be tied back to business units via Slurm database to track budget to projects. This platform will be deployed independently for US and India users: US will leverage AWS Virginia region and India will leverage AWS Malaysia region.

### Key Design Decisions (KDD)

Summarize the Technology Architecture Key Design Decisions (KDD) taken for this release in the table below (please add references to RALLY or ALM for further details if required). Don’t forget to review the Risk Register after KDD’s added.

| **KDD ID** | **ASRB Required** | **Title** | **Rationale** | **Approver/Date** | **Reason for Decision** | **Implementation plan** | **Required Date** |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Y | No Storage Segmentation for EC / non-EC Amazon FSx ONTAP | Same Storage | Gabriele Giovani & Yasen Hristov | RESOLVED (September 11, 2024) | All HPC nodes mounted to both file systems; VDI node access is managed by AD group permissions. | - |
|  | Y | Subnet segmentation for App, DB, HPC, and VDI | Simplifies network security group control and access back to GE Vernova network. HPC and VDI nodes do not require direct routing from GE Vernova network but will require outbound network access. | Gabriele Giovani & Yasen Hristov | RESOLVED (September 11, 2024) | Deploy 4 subnets for each availability zone (2 routable, 2 non-routable)  App & DB are using 2 AZs; HPC & VDI are in 2+ | - |
|  | N | NI-SP Engineframe | EnginFrame supports a web-based job submission portal for HPC users to submit jobs without creating CLI scripts. Currently, AWS RES does not support this feature. |  | UNRESOLVED; NISP support model unclear. Next steps on AWS team | RES did not support web-based job submission portal; EnginFrame supports and is in use today |  |
|  | N | ParallelCluster vs MyHPC |  |  | UNRESOLVED; conversion of MyHAL. Need to evaluation all scheduler options. Conversion to PBS or SLURM directly. |  |  |
|  | Y | AWS Region in Malaysia for India users | Singapore constraints on land and power limit expansion, instance availability, and future conversion to better hardware. Malaysia is the newest AWS region, announced August 21, 2024 |  | UNRESOLVED:  in test planning. Plan for ASRB is to go with Malaysia region. Can be resolved immediately upon completion of Malaysia latency test. Otherwise repeat ASRB. [update 9/27 – AWS Guardduty not available in Malaysia; scheduled for mid-October. This is blocking testing] | 1/ deploy in Mumbai and test latency (traffic coming through Singapore)  2/ short term Malaysia region test (with AWS Gaurduty exception).  3/ updated in AWS-GEV PPA to extend negotiated rates to this region | September 30th |
|  | Y | Asia Landing Zone | What if Malysia doesn’t give us expected latency , what is going to be an alternate landing zone for India requests |  | RESOLVED   |  |  | | --- | --- | | Region | ~ latency [ms] | | Singapore | 45 | | Malaysia | 55 | | Jakarta | 70 | | Mumbai | 110 | |  |  |
|  | N | Non-Compliance with MS Office on Windows VDI | 1. Determine GEV license agreement with Microsoft 2. If risks exists even after acting on pt 1 -Evaluate other MS office options on Amazon 3. Create a service which will allow to use macro 4. Economize compliant office licenses on dedicated instances |  | UNRESOLVED; details on MSFT agreement forthcoming. |  |  |
|  | Y | Unauthorized access to data | 1. Utilizing export policies and permissions to authorize data 2. Utilize network groups 3. Potentially have 2 file systems for EC & non-EC 4. Use netApp features to hide files 5. Resue what we have currently in place -e.g Export control flag in network security groups |  | RESOLVED; dependency matrix created. [@Krutika will update] | Consult AWS Soc2 report to validate no access to data (EC or non-EC). Ensure network security groups extend to the EC FSx OnTAP file system when joined to CloudAD. | September 20th |
|  | N | Windows VDI solution not robust | 1. Update domain join process - check if API returns domain controller 2. Optimize the domain join script 3. For enginFrame - implement optimized domain join script |  | UNRESOLVED; discussion with AD team scheduled for September 12; review with Mahesh D. Update 9/27 – waiting on ticket from Yasen to Mark S / Identity. Potential resolution – cycle through AD servers individually at the app layer. |  |  |
|  | Y | Unavailability of AMI for windows 10 desktop | Requirement - Dedicated host/instance if VDA License 1. Prioritization on AMI Creation with Laszlo 2. Leverage existing VMDK to create AMI |  | UNRESOLVED; no direction. Budget required for image creation. 9/27 update – image is encrypted and unable to be shared across accounts. |  | September 30th |
|  | Y | Not using right protocol due to domain structure (using old NTLM protocols) | 1. Administration team should provide evidence on audit 2. Protocols & versions must be audited on regular basis |  | UNRESOLVED; testing Kerberos encryption in non-prod and if NTLM is the only option, work with security on exception. |  | ? |
|  | N | long-lead time in build | 1. Clear communications on deadlines & risks 2. Share project plans including sprint schedules 3. Priortization on resource allocation |  | UNRESOLVED; confer with Venu |  |  |
|  | N | Not having strategic support agreement with vendor NI-SP | 1. Contract between GEV & NI-SP to support enginFrame issues 2. Leverage NI-SP in short term until AWS PCS can support a job submission portal natively without requiring EnginFrame. |  | UNRESOLVED; Need to resolved with AWS and NISP |  |  |
|  | Y | Non-adherence to Cyber Standards - architect , code etc to meet standards | 1. understand requirements & work closely with cyber 2. act beforehand…let's not wait for last minute |  | UNRESOLVED; Yasen meet with Cyber to review non-prod to document any non-adherent items and Identify gaps. AWS Well Architected Framework Review to use as a standard for cyber standards AFTER non-prod and AFTER Cyber review. | Schedule a WAFR Security Pillar review –review the questions and decision impacts. | - |

### Baseline Technology Architecture (Physical Model)

Summarize the As-Is Technology Architecture. Let understand in short words & diagrams where this release is starting from. Minimum here is to document production & QA environment.

The existing HPC environment is maintained in Lebanon (US) and Bangalore (IN) data centers using GRC-maintained clusters of Dell (?) compute hardware and Isilon and NetApp storage connected over Infiniband. The existing VDI environment runs on VMWare Horizon View out of Lebanon and Bangalore and uses V: Drive for application binaries.

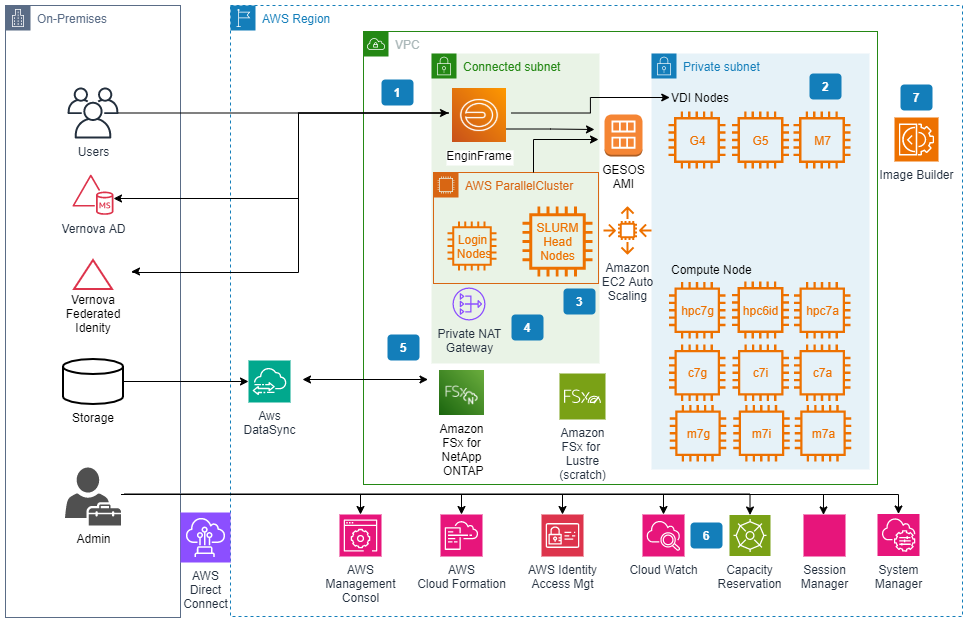
HPC and VDI high level reference architecture:

[EXISTING ARCHITECTURE DIAGRAM]

### Target Technology Architecture (Physical Model)

Summarize the To-Be Technology Architecture. Let understand in short words & diagrams where this release will deploy. Minimum here is to document production & QA environment.

The HPaaS Platform is designed using EnginFrame and AWS Parallel Cluster – both on the AWS cloud platform. VDI application binaries will be stored on V: Drive and mounted to Windows VDI. For HPC only 2 AMIs: 1/ start with GESOS, 2/ add application layer, 3/ create new AMI. For VDI: 1/ start with GESOS, 2/ add application layers (Nvidia, etc), 3/ create one AMI for RHEL and one for Windows 10. ~~VDI and HPC applications will be built using Amazon Machine Images (AMI) customized by GESOS team. We will need to add the 3~~~~rd~~ ~~party drivers (ie Nvidia) and then install specific applications (eg Ansys) from V:/ Drive. The following image illustrates various AWS services used to implement the platform. The following sections describe how each service is used.~~



\*\* License server (Ansys, Hyperworks, Teamcenter, Unigraphics). Add license server to on-premises box

|  |  |
| --- | --- |
|  | Engine Frame is the web portal to access the system (see more detailed below). Engine frame uses Vernova Federated Identity (Ping) for portal authentication. EngineFrame control the life cycle of the VDI nodes and allows the submission of HPC jobs to the ParallelCluster Slurm head node. The windows VDI nodes are domain joined to Vernova AD. The Linux VDI nodes authenticate via the Vernova LDAP server. |
|  | VDI nodes are created in a private subnets. To minimize the chance of insufficient capacity error for the onDeman VDI nodes, multiple subnets across multiple AZ are used. See detailed below |
|  | ParallelCluster provides Slurm Head node and |
|  | A private NAT Gateway. |
|  | Amazon FSX for NetApp ONTAP is a managed storage service that can meet the needs of high IOPS and Bandwidth requirements for HPC. It also supports SMB and NFS protocols, allowing nodes running Linux or Windows to join and mount the same namespace. |
|  | EC2 instances will use reserved capacity on AWS, meaning that the capacity is always available and committed for use by GE Vernova only. Pricing will not fluctuate for usage of these EC2 instances, |
|  | Custom code will be required to captured the latest GESOS AMI and populate it into the EnginFrame configuration for use for VDI nodes (Windows and Linux). Further, AWS ParallelCluster will need this AMI for creating compute nodes (Linux) |

AWS cloud is the landing zone to setup the platform.

* GitHub Actions would deploy the Java web app for EnginFrame Frontend
* Reserved Capacity EC2 instances in a single availability zone will be orchestrated by EnginFrame
  + Instances reserved include HPC7a, G4nd, and P4
* On Demand EC2 instances will be used in the event of disaster recovery or a burst capacity queue
* VDI instances will join GE CCMGMT Domain; HPC instances will join LDAP using nix account (all on EC2)
* Amazon FSx for NetApp ONTAP will be the primary storage technology mounted to VDI and HPC for simulation storage and for V:/Drive (used for application binaries)
  + FSx will use SnapMirror for data migration directly into FSx ONTAP
  + Isilon will use AWS Datasync for moving data into FSx ONTAP
* AWS Autoscale Groups will be used for managing AWS instance scaling
* Desktop-to-Cloud Visualization (DCV) technology will be used for remote access to desktops for performance
* Application Load Balancer (ALB) is used for load balancing, application components with SSO + MFA authentication for EnginFrame
* Network Load Balancer (NLB) is used for load balancing DCV sessions to VDI nodes

#### Technology Summary (Web/App)

Summarize the deployed technology, platform and its provision.

##### CloudFormation & Infrastructure:

Platform Area : Public Cloud (AWS) - IaaS/PaaS

Product Type : Web/App - Dedicated Build (Custom)

Product : Dedicated WS - Internal App

App Server Tech : AWS Cloudformation

Containerization : none

Build tool : GitHub Actions

##### Applications

Application Components : EnginFrame, ParallelCluster, SLURM, HPC & VDI applications (Ansys, Cadence, etc)

Platform Area : Public Cloud (AWS) - IaaS/PaaS

Product Type : Web/App - Dedicated Build (Custom)

Product : Dedicated AS w/ Dedicated WS - Internal App

App Server Tech : java

Containerization : none

Java : Open JDK 11

Build tool : GitHub Actions

#### Solution Building Blocks (SBB)

Summarize the given solution blocks and how they’re providing their capabilities to the Application. Let understand the reader the motivation to realize/configure it that way.

##### EnginFrame:

#### EnginFrame is a grid-enabled application portal for user-friendly HPC job submission, control, and monitoring. It includes sophisticated data management for all stages of job lifetime. It provides a modular system where you can easily add new functionality, such as application integrations, authentication sources, and license monitoring. It also features a sophisticated web services interface that you can use to enhance existing applications and develop custom solutions for your own environment. EnginFrame is a computing portal that uses existing scripting solutions when available. This means that, while using EnginFrame, you can avoid interacting with a command line interface in situations where you might have not had that option before.

#### These are key considerations for using EnginFrame for HPCaaS: 1/ it is customizable using a process already employed by Anemoi HPC team, and 2/ it provides a single web-based portal for launching VDI nodes and HPC jobs, including job submission (avoiding CLI only)

##### EC2 (p4d.24xlarge, hpc7a.96xlarge, hp6id.32xlarge, g4dn):

#### Amazon Elastic Compute Cloud (Amazon EC2) offers the broadest and deepest compute platform, with over 750 instances and choice of the latest processor, storage, networking, operating system, and purchase model to help you best match the needs of your workload.

#### [Amazon EC2 Hpc7a instances](https://aws.amazon.com/ec2/instance-types/hpc7a/) feature 4th Gen AMD EPYC processors and are designed for tightly coupled, compute-intensive high performance computing (HPC) workloads such as computational fluid dynamics (CFD), weather forecasting, and multiphysics simulations. Includes up to 192 cores of 4th Gen AMD EPYC processors with 768 GiB of memory (AMD EPYC 9R14).

#### [Amazon EC2 G4dn instances](https://aws.amazon.com/ec2/instance-types/g4/) are designed to help accelerate machine learning inference and graphics-intensive workloads. Includes 2nd Generation Intel Xeon Scalable Processors (Cascade Lake P-8259CL) and up to 8 NVIDIA T4 Tensor Core GPUs.

#### [Amazon EC2 P4 instances](https://aws.amazon.com/ec2/instance-types/p4/) provide high performance for machine learning training and high performance computing in the cloud. Includes 3.0 GHz 2nd Generation Intel Xeon Scalable processors (Cascade Lake P-8275CL) and up to 8 NVIDIA A100 Tensor Core GPUs

#### These instance families and types are chosen for HPCaaS because they are sized to meet existing on-premises core counts but using the latest processors available. They cover the existing HPC clusters in Lebanon and Bangalore data centers as well as VMware Horizon View hardware requirements. To maximize cost optimization, Windows Bring Your Own Licensing is required.

##### Capacity Reservation:

#### On-Demand Capacity Reservations allow you to reserve compute capacity for your Amazon EC2 instances in a specific Availability Zone for any duration. If you have strict capacity requirements, and are running business-critical workloads that require a certain level of long or short-term capacity assurance, Capacity Reservation ensures that you always have access to Amazon EC2 capacity. The capacity becomes available and billing starts as soon as the Capacity Reservation is provisioned in your account. When you no longer need the capacity assurance, cancel the Capacity Reservation to release the capacity and to stop incurring charges. You can also use the billing discounts offered by Savings Plans and Regional Reserved Instances to reduce the cost of a Capacity Reservation.

#### Capacity Reservation is chosen due to the predictable pricing model (prenegotiated flat rate per month for instance usage). This also means that unless an on-demand HPC queue is created, the pricing will be 100% predictable outside network and storage fluctuations.

##### On-Demand Instances

#### On-Demand Instances let you pay for compute capacity by the hour or second (minimum of 60 seconds) with no long-term commitments. This frees you from the costs and complexities of planning, purchasing, and maintaining hardware and transforms what are commonly large fixed costs into much smaller variable costs.

#### On-Demand Instances is the default option for any bursting requirements for this workload that would require special approval to consume. Jobs that are time sensitive or otherwise require access to a specific instance can use ondemand instance and pay the higher rate. On-Demand is also the choice for any Disaster Recovery deployment as the capacity reservations do not exist in the DR region.

##### Dedicated Instances:

#### Dedicated Instances are EC2 instances that run on hardware that's dedicated to a single AWS account. This means that Dedicated Instances are physically isolated at the host hardware level from instances that belong to other AWS accounts, even if those accounts are linked to a single payer account. However, Dedicated Instances might share hardware with other instances from the same AWS account that are not Dedicated Instances. Dedicated Instances provide no visibility or control over instance placement, and they do not support host affinity. If you stop and start a Dedicated Instance, it might not run on the same host. Similarly, you cannot target a specific host on which to launch or run an instance. Additionally, Dedicated Instances provide limited support for Bring Your Own License (BYOL).

#### To maximize cost efficiency, Windows BYOL is required for this workload. To meet Microsoft licensing standards, dedicated instances are needed and GE Vernona must acquire the licenses separately. Therefore all instances as part of this workload will use AWS dedicated instances.

##### Amazon FSx for NetApp ONTAP:

#### Amazon FSx for NetApp ONTAP is a storage service that allows customers to launch and run fully managed ONTAP file systems in the cloud. ONTAP is NetApp’s file system technology that provides a widely adopted set of data access and data management capabilities. Amazon FSx for NetApp ONTAP provides access to shared file storage over all versions of the Network File System (NFS) and Server Message Block (SMB) protocols, and also supports multi-protocol access (i.e. concurrent NFS and SMB access) to the same data.

#### This is the primary storage solution due to the multi-protocol support, minimizing file movement between Linux and Windows Operation Systems. Today, this leads to a burdensome process for file transfer. Amazon FSx for NetApp ONTAP will also accelerate migration, which can leverage NetApp SnapMirror technology to migrate data live over network connections. This storage also support active “hot” storage that can be segmented for backup & restore.

##### NICE DCV:

#### NICE DCV is a high-performance remote display protocol that provides customers with a secure way to deliver remote desktops and application streaming from any cloud or data center to any device, over varying network conditions. With NICE DCV and Amazon EC2, customers can run graphics-intensive applications remotely on EC2 instances, and stream their user interface to simpler client machines, eliminating the need for expensive dedicated workstations.

#### NICE DCV is the chosen streaming protocol as it comes natively with EnginFrame, but also supported by AWS and comes natively with a number of their services. It’s in use today with the Anemoi HPC workload and performance is acceptable.

##### Amazon S3:

#### Amazon Simple Storage Service (Amazon S3) is an object storage service offering industry-leading scalability, data availability, security, and performance. Millions of customers of all sizes and industries store, manage, analyze, and protect any amount of data for virtually any use case, such as data lakes, cloud-native applications, and mobile apps. With cost-effective storage classes and easy-to-use management features, you can optimize costs, organize and analyze data, and configure fine-tuned access controls to meet specific business and compliance requirements.

#### Amazon S3 provides long term and cost effective storage and is required into a number of the services part of the HPCaaS design (EnginFrame, Parallel Cluster).

##### Application Load Balancer

#### Application Load Balancer operates at the request level (layer 7), routing traffic to targets (EC2 instances, containers, IP addresses, and Lambda functions) based on the content of the request. Ideal for advanced load balancing of HTTP and HTTPS traffic, Application Load Balancer provides advanced request routing targeted at delivery of modern application architectures, including microservices and container-based applications. Application Load Balancer simplifies and improves the security of your application, by ensuring that the latest SSL/TLS ciphers and protocols are used at all times.

#### Application Load Balancers will support HPCaaS high availability across AWS availability zones as well as SSO redirection to Ping Federate. In the event of an AWS availiabilty zone outage, it will redirect traffic to each additional availability (and subnet) into which it is configured. It is an AWS supported service and is a fit for EnginFrame highly-available design as well as Parallel Cluster login nodes.

##### Network Load Balancer

#### Network Load Balancer operates at the connection level (Layer 4), routing connections to targets (Amazon EC2 instances, microservices, and containers) within Amazon VPC, based on IP protocol data. Ideal for load balancing of both TCP and UDP traffic, Network Load Balancer is capable of handling millions of requests per second while maintaining ultra-low latencies. Network Load Balancer is optimized to handle sudden and volatile traffic patterns while using a single static IP address per Availability Zone. It is integrated with other popular AWS services such as Auto Scaling, Amazon EC2 Container Service (ECS), Amazon CloudFormation, and AWS Certificate Manager (ACM).

#### Network Load Balancers will support HPCaaS high availability for DCV sessions across VDI nodes. As this is not web traffic, we must distribute traffic at a lower level in the OSI model to distribute sessions. In the event of an AWS availiabilty zone outage, it will redirect traffic to each additional availability (and subnet) into which it is configured. It is an AWS supported service and is a fit for EnginFrame DCV streaming for VDI nodes.

##### Amazon RDS (EnginFrame supports MySQL® Database 8.0.x and later with the InnoDB storage engine)

#### [MySQL](https://aws.amazon.com/rds/mysql/what-is-mysql/) is the world's most popular [open source](https://aws.amazon.com/products/databases/open-source-databases/) relational database and Amazon RDS makes it easier to set up, operate, and scale MySQL deployments in [the cloud](https://aws.amazon.com/what-is-cloud-computing/). With Amazon RDS, you can deploy scalable MySQL servers in minutes with cost-efficient and resizable hardware capacity.

#### Amazon RDS for MySQL frees you up to focus on application development by managing time-consuming database administration tasks, including backups, upgrades, software patching, performance improvements, monitoring, scaling, and replication. Amazon RDS supports MySQL Community Edition versions 5.7 and 8.0 which means that the code, applications, and tools you already use today can be used with Amazon RDS.

#### In order to make EnginFrame and SLURM highly available, we must distribute the database across availability zones. Amazon RDS will provide the managed mySQL database to minimize administrative tasks and highly-available configuration. An empty database instance that's named EnginFrameDB must be created before EnginFrame first startup.

##### AWS KMS

#### AWS KMS is a managed service that helps you more easily create and control the keys used for cryptographic operations. The service provides a highly available key generation, storage, management, and auditing solution for you to encrypt or digitally sign data within your own applications or control the encryption of data across AWS services.

#### AWS allows using GE Vernova to provide their own keys for encryption at rest as it allows GE Vernova to create and manage their own encryption keys

#### SLURM

#### Slurm is an open source, fault-tolerant, and highly scalable cluster management and job scheduling system for large and small Linux clusters. Slurm requires no kernel modifications for its operation and is relatively self-contained. As a cluster workload manager, Slurm has three key functions. First, it allocates exclusive and/or non-exclusive access to resources (compute nodes) to users for some duration of time so they can perform work. Second, it provides a framework for starting, executing, and monitoring work (normally a parallel job) on the set of allocated nodes. Finally, it arbitrates contention for resources by managing a queue of pending work.

#### Slurm is the planned scheduler for this deployment as it is open source, can be deployed in a multi-AZ design for fault tolerance, and supports SLURM accounting for charging back compute resources to consuming business units. It is also supported by AWS ParallelCluster.

### System Architect Diagrams

Provide the system architecture diagrams for the major building blocks and its relationships (e.g. interfaces, middleware, hosting etc.). Describe the system architecture for all environments of the Application.

Following sections show architecture diagrams for the major components and its relationships for the provided service.

**Single Points of Failure:**  
- Amazon FSx for NetApp ONTAP HA Storage ( + required for EnginFrame HA)

- Amazon FSx for Lustre (single-AZ by design)

- Instance limitation if Primary AZ goes down (VDI and compute instances will be on-demand)

- Put License server in AWS; Ansys confirmed we can port into AWS. – but TSA limits movement

- EngineFrame Scheduler – should be highly available through ALB and HA design

- ParallelCluster Scheduler – should be configured Active / Passive design so scheduler can resume after manual cutover.

~~a automated cutover (no manual intervention required).(@seamusowens)~~

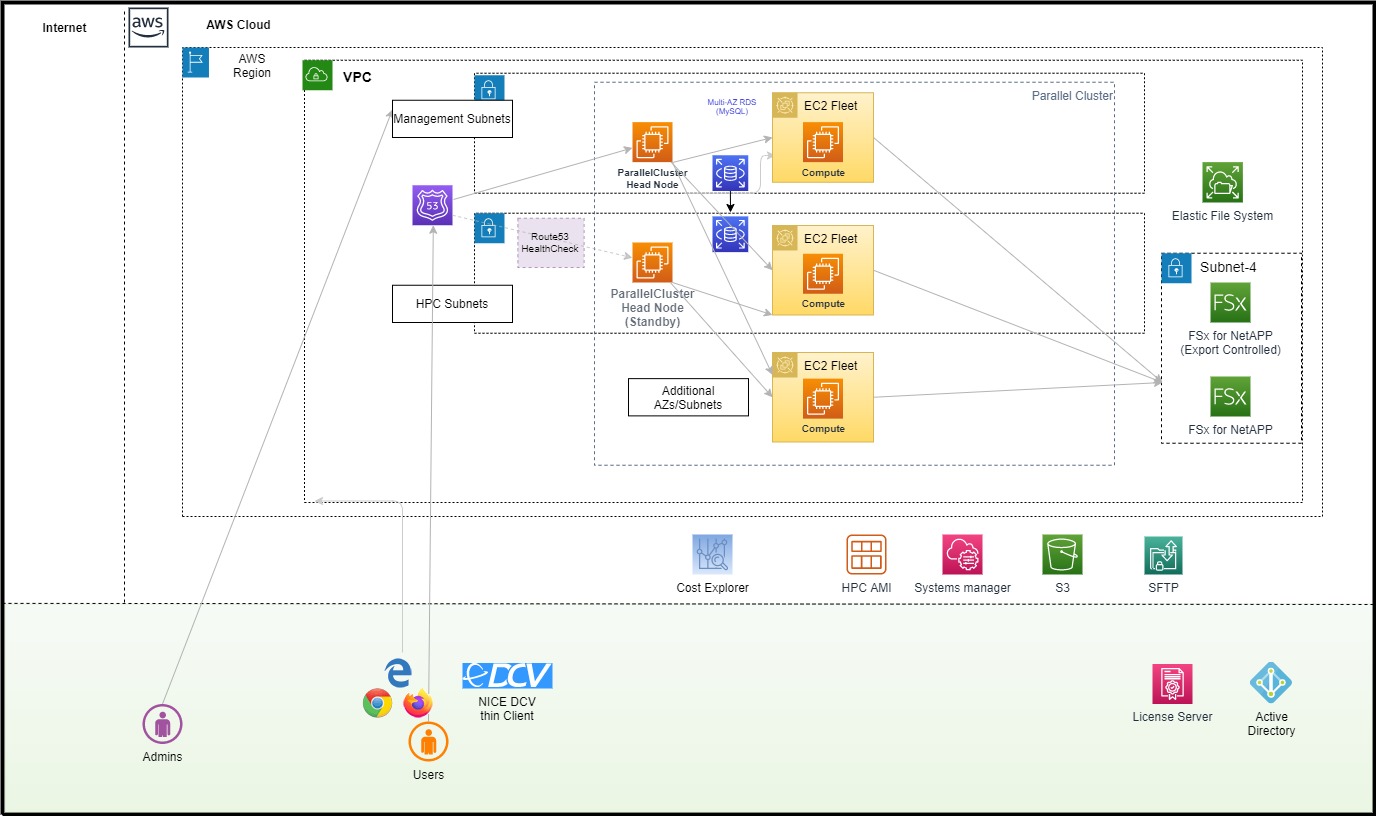
diagram will be updated with separate head nodes for AZ 1 & 2 (@seamusownes)

approach – load balancer

#### EngineFrame

EnginFrame application module is developed to create and connect to VDI instances, maintain a SLURM-based job scheduler – including web-based job submission portal, and create and connect to HPC instances user interface.

Below diagram depicts the architecture of solution building block and their interactions with other components.



### DC Sites and Infrastructure Services (Based on CMDB)

Name the strategic datacenter placement for all environments of the Application.

|  |  |  |  |
| --- | --- | --- | --- |
| **Environment** | **Disposition** | **Data Center** | |
| **Location** | **Preferred / Required /Remote**  **Other Location not listed** |
| DEV | New | AWS | Us-east-1 (Virginia) |
| PROD | New | AWS | ap-southeast-5 (Malaysia) & us-east-1 (Virginia) |

Note: ***Compliance requires the use of*** *PT Design Exception* ***if infrastructure is placed*** *outside of a strategic data center*

***REFERENCE***

* *PT Design Exception request:* <http://supportcentral.ge.com/ProcessMaps/form_new_request.asp?prod_id=276282&form_id=484388&node_id=1444120&map_id=384811&reference_id=6712390007&reference_type=4>

The building blocks are ideally based on the service catalogue, except in the case of a deviation from the standards.

| **Application Component ID** | **Application Component Name** | **Environment** | **Infrastructure Service Category** | **CMDB Catalogue Item\*** | **Server Name\*** | **Datacenter Site & Country** |
| --- | --- | --- | --- | --- | --- | --- |
|  | EnginFrame Web Application | Production |  |  |  | AWS Virginia (us-east-1) |
|  | DCV Gateway | Production |  |  |  | AWS Virginia (us-east-1) |
|  | DCV Broker | Production |  |  |  | AWS Virginia (us-east-1) |
|  | ParallelCluster Login Node | Production |  |  |  | AWS Virginia (us-east-1) |
|  | ParallelCluster Head Node | Production |  |  |  | AWS Virginia (us-east-1) |
|  | VDI Nodes | Production |  |  |  | AWS Virginia (us-east-1) |
|  | HPC Compute Nodes | Production |  |  |  | AWS Virginia (us-east-1) |
|  | AWS DataSync | Production |  |  |  | AWS Virginia (us-east-1) |
|  | Amazon FSx for NetApp ONTAP | Production |  |  |  | AWS Virginia (us-east-1) |
|  | Amazon FSx for Lustre | Production |  |  |  | AWS Virginia (us-east-1) |
|  | EnginFrame Web Application | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | DCV Gateway | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | DCV Broker | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | ParallelCluster Login Node | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | ParallelCluster Head Node | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | VDI Nodes | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | HPC Compute Nodes | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | AWS DataSync | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | Amazon FSx for NetApp ONTAP | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | Amazon FSx for Lustre | Production |  |  |  | AWS Malaysia (ap-southeast-5 |
|  | EnginFrame Web Application | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | DCV Gateway | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | DCV Broker | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | ParallelCluster Login Node | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | ParallelCluster Head Node | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | VDI Nodes | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | HPC Compute Nodes | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | AWS DataSync | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | Amazon FSx for NetApp ONTAP | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | Amazon FSx for Lustre | Non-Production |  |  |  | AWS Virginia (us-east-1) |
|  | ADD DR PIECES??  All DR components will be exact replica of PROD in small scale and switched off (Virginia and India users)…. Including storage with minimal amount should be replicated from Virginia to Ohio  Stopped instance patching |  |  |  |  |  |

*\*The server configuration could be looked up in the CMDB (ServiceNow).*

***REFERENCE***

* *CMDB Data Model Documentation:* [*https://ge.service-now.com/kb\_view.do?sysparm\_article=KB0675960#Data%20Model*](https://ge.service-now.com/kb_view.do?sysparm_article=KB0675960#Data%20Model)

### Infrastructure Support Matrix

Summarize the infrastructure support contacts (especially local contacts) to reach out when Technology & Hardware issues arise, upgrades are required, planned outages etc. Try to identify these contacts to cover all components run within the Application solution provided (globally).

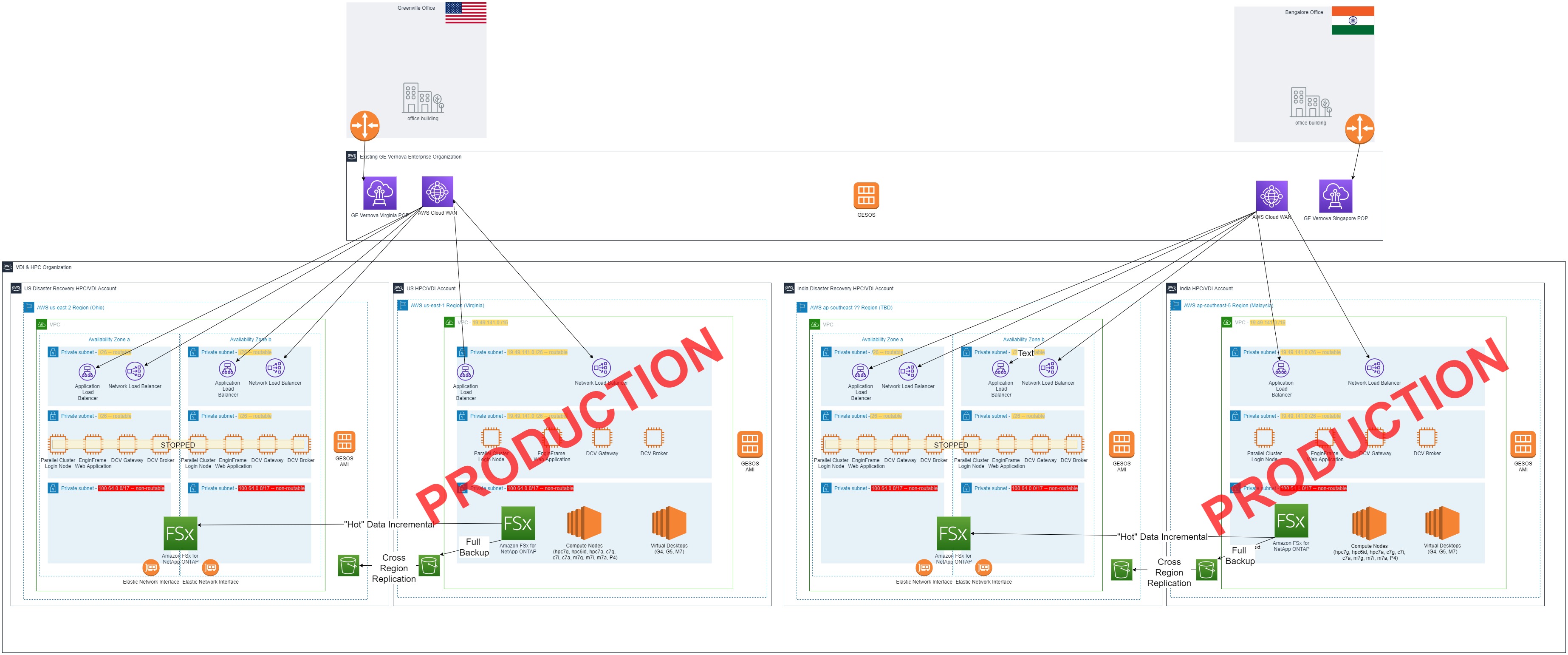
| **Application Component ID** | **Application Component Name** | **Local contact details** | **Datacenter Site & Country** | **Remarks** |
| --- | --- | --- | --- | --- |
|  | EnginFrame | [AWS/NI-SP Support for first year]\* |  |  |
|  | AWS ParallelCluster | AWS Enterprise Support [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) | US Lead: Aaron Bluestone (ablue@amazon.com) India: Puneetha Kumara ([ckpuneet@amazon.com)](mailto:ckpuneet@amazon.com)) US West: Nik Hutson (hutsnico@amazon.com) | [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) DL is preferred as it will always capture TAM |
|  | AWS FSx for NetApp ONTAP | AWS Enterprise Support  [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) (can also include NetApp Team) | US Lead: Aaron Bluestone (ablue@amazon.com) India: Puneetha Kumara ([ckpuneet@amazon.com)](mailto:ckpuneet@amazon.com)) US West: Nik Hutson (hutsnico@amazon.com) | [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) DL is preferred as it will always capture TAM |
|  | DCV | AWS Enterprise Support  [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) | US Lead: Aaron Bluestone (ablue@amazon.com) India: Puneetha Kumara ([ckpuneet@amazon.com)](mailto:ckpuneet@amazon.com)) US West: Nik Hutson (hutsnico@amazon.com) | [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) DL is preferred as it will always capture TAM |
|  | Standard AWS Infrastructure [EC2, RDS, ALB, EFS,] | CloudOps DL & AWS Enterprise Support | US Lead: Aaron Bluestone (ablue@amazon.com) India: Puneetha Kumara ([ckpuneet@amazon.com)](mailto:ckpuneet@amazon.com)) US West: Nik Hutson (hutsnico@amazon.com) | [aws-tam-ge-vernova@amazon.com](mailto:aws-tam-ge-vernova@amazon.com) DL is preferred as it will always capture TAM |
|  |  |  |  |  |

### Backup & Recovery

Summarize the backup and disaster recovery details (reference to the DR plan).

**DR for Virginia is Ohio Region**

**DR for Malaysia is TBD Region**



**20TB / last week’s data/hot data?**

**Can FSx replicate data from primary to standby ? – Seamus to check with niki**

* **Patching process needs to be validated by OPS - Yasen**
* **Application validation process needs to be validated**
* Please include link to Disaster Recovery Plan link if relevant: < *insert* ***link*** *here*>

|  | **Non-Production** | **Production** | **Notes** |
| --- | --- | --- | --- |
| DR tier | Choose an item. | Choose an item. | Storage and HPC Compute = Single AZ  VDI and EnginFrame = Multi AZ |
| DR method | Choose an item. | Choose an item. | Stand up new FSX OnTAP, Restore from S3 for Storage; OnDemand capacity in another Region. |
| Current sizing requirements for file system | ? | US: 1 PB India: 500TB | Archive: S3 archival |

* Below details need to be validated with business

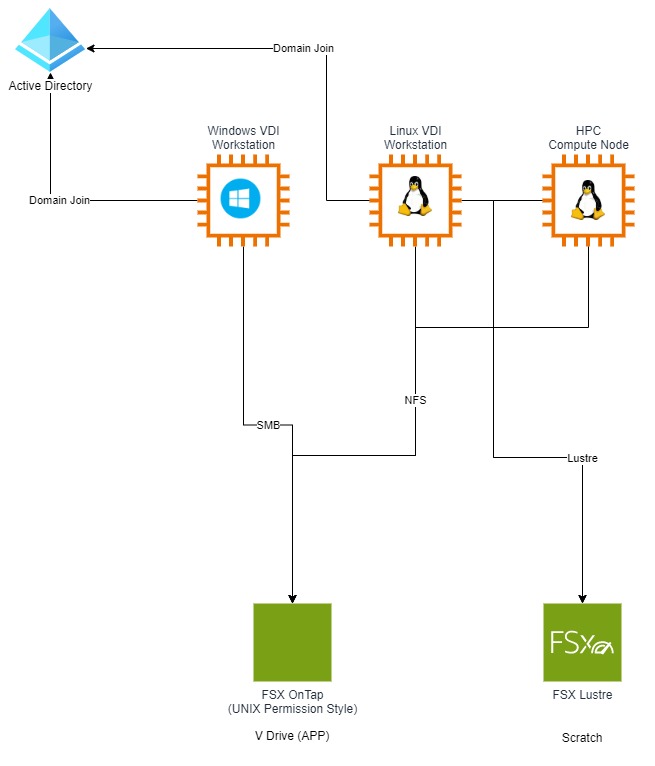
| **Backup interval** | **Backup type** | **Retention time** |
| --- | --- | --- |
| Daily | Incremental | 2 weeks |
| Weekly | Incremental | 4 weeks |
| Monthly | Incremental | 6 months |

### Storage Architecture

Summarize the Storage architecture, its capacity and storage devices & configuration used to supply the Application. This includes the central application site and also any remote site included as part of the deployed Solution.

Amazon FSX for NetApp OnTAP is being used as a solution for file collaboration requirements for both VDI and HPC applications. We will use a single file system and use Active Directory group permission to manage file access between EC and non-EC users and data. This service provides access to shared file storage over all versions of the Network File System (NFS) and Server Message Block (SMB) protocols, and also supports multi-protocol access (i.e. concurrent NFS and SMB access) to the same data. As a result, you can access Amazon FSx for NetApp ONTAP from virtually any Linux, Windows, or macOS client.

Amazon FSx for Lustre will be included in the design to due to potential scaling issue with FSx OnTAP at jobs at or above 500 cores. Lustre can be adjusted or removed later if no necessary.



Each Amazon FSx for NetApp ONTAP file system has two storage tiers: primary storage and capacity pool storage. Primary storage is provisioned, scalable, high-performance SSD storage that’s purpose-built for the active portion of your data set. Capacity pool storage is a fully elastic storage tier that can scale to petabytes in size and is cost-optimized for infrequently-accessed data.

US:

Primary Storage Pool: 20 TB (each VDI will have a maximum storage quote for 200GB)   
Capacity Storage Pool: Remainder (980 TB)

India:

Primary Storage Pool: 20 TB (each VDI will have a maximum storage quote for 200GB)

Capacity Storage Pool: Remainder (480 TB)

Access to shared storage for VDI nodes will be governed by Active Directory group permissions and will extend down to user authorization at the file and folder level. This

Usage includes three patterns–

1. Pre and Post Processing storage on Windows desktops (VDI) over SMB
2. HPC cluster mounted storage on scaling compute notes for high performance scratch and persistent storage (also accessed over shared space with VDI nodes) over NFS.
3. V: Drive access to engineering application binaries

|  |  |
| --- | --- |
| **General details** | |
| Hostname | Details can be filled in once infrastructure is built |
| Function |  |
| Domain |  |
| Site |  |
| Technical Administrator 1 |  |
| Physical/Virtual |  |
|  | |
| Model |  |
| RAM |  |
| CPU |  |
| Disks |  |
| Storage Devices |  |
| Further storage components |  |
|  | |
| Operating system |  |
| Installed software |  |
|  | |
| Login name |  |
| Login type |  |
| Ports for firewall |  |

### Database Architecture

Summarize the physical Database architecture, its capabilities and capacity to supply the application.

Yet to fill some details post infrastructure is setup

| **Attribute** | **Value** | **Description** |
| --- | --- | --- |
| Database vendor | AWS | **Amazon RDS service** |
| Database disposition | Create new database (instance) | Yes |
| Type of database | MySQL Community Edition |  |
| Database hosting | RDS (Managed Service) |  |
| Shared or dedicated database instance | Dedicated |  |
| Number of databases or instances per environment | 2 | SLURM and EnginFrame |
| Database size (in GB) | 50 GB | Good to start, may change based on slurm storage utilization (check with Seamus) |
| Database use | OLTP | Not very db intensive , just to track activities |
| Database version | MySQL 8.0 |  |
| RDS Instance size | db.m6g.xlarge | 4 vcpu, 16GB RAM |
| ~~SQL Server Services~~ |  | ~~Database Engine Service, Reporting Service, Analysis Service, Integration Service~~ |
| Instance CI – PROD (if known) |  |  |
| Instance CI – QA (if known) |  |  |
| Instance CI – STG (if known) |  |  |
| Instance CI – DEV (if known) |  |  |
| Instance CI – DR (if known) |  |  |
| Instance CI – Other (if known) |  |  |
| Number of schemas per instance | Choose an item. | Note: Oracle request to create instance & schema imposes a limit of 3 schemas |
| Schema names |  |  |
| Total number of users |  |  |
| Number of active sessions |  | Total number of concurrent users and their sessions |
| Character Set | N/A (Only Oracle) |  |
| Custom backup / archival requirements | Automated RDS Snapshot (daily) | Specify any custom backup or archival requirements |
| Type of replication required | Snapshot |  |
| Database Functional Account ID |  |  |
| Non-Standard database User Account ID |  |  |

## Workplace

Summarize the client-side view, including the deployment strategy, client types and pre-requisites to be considered.

### Client Devices and Operating Systems

The following client platforms maybe considered as a reference at the time of writing. However, this list is expected to evolve and for the latest please refer to the Workplace Standards.

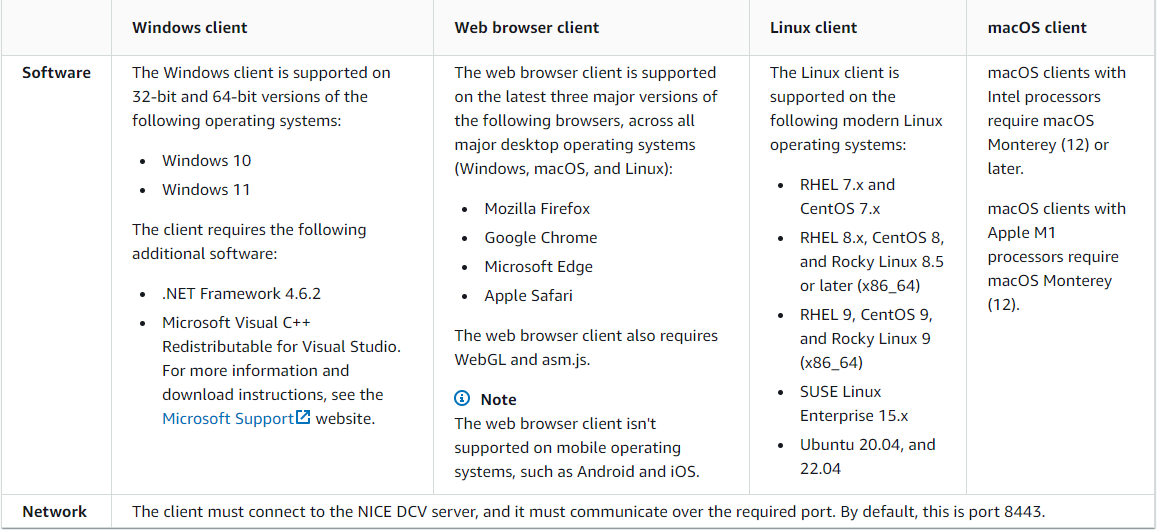
|  |  |  |
| --- | --- | --- |
| **Computers** | **Mobile Tablets** | **Phones** |
| User laptop | N/A | N/A |

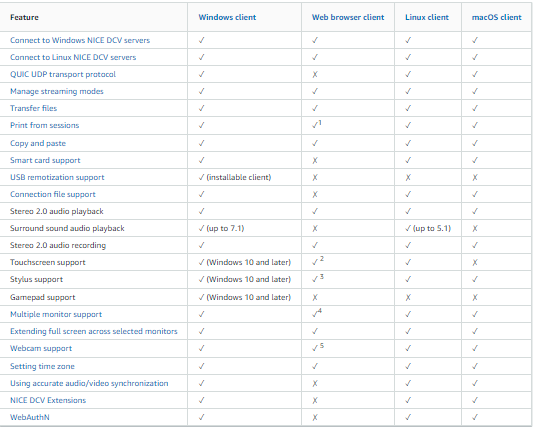
### Client deployment Process

Summarize the client deployment process and the deployment strategy.

Windows desktops created using EnginFrame to instantiate Virtual Desktops on EC2 instances . These can be accessed via any HTML5 web browser. However, native clients support additional features such as multi-monitor support, with the Windows native client also supporting USB support for 3D mice, storage devices and smart cards.

The DCV Client will be distributed via V: Drive shares to end users via local laptop/workstation mount (alternatives is to use myTech?

Native Client Requirements:  
  




### User Client types

List the existing client types to be considered when deploying the client application package (for e.g. OS, Windows, Linux, CITRIX client access etc.).

Below are client types supported -

OS – Windows

### Prerequisites & Compatibility Matrix

Describe the client side pre-requisites to be considered for running the client application.

This section is not applicable.

| **Application Component ID** | **Application Component Name** | **Prerequisites** | **Compatibility** |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

## Security Architecture

Summarize the security architecture topics and how it is handled by the provided solution. Typical topics are Data classification, Password policies, User authentication, EC etc.

Ensure that security review results with the security team were reflected here in this section.

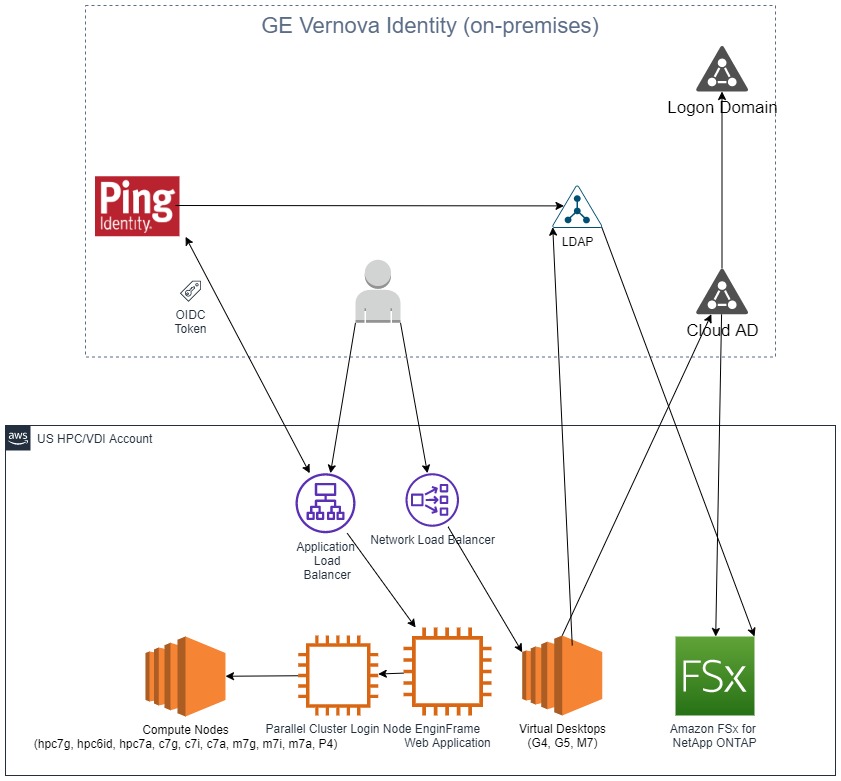
[Create security architecture diagram – Seamus]

### Security and Single Sign-On integration

Describe the security design for the application and how the application integrates with SSO. State what specific groups and roles are created and how they are enforced. Also mention any additional security that the application enforces above that supplied by Siteminder (SSO) and LDAP.

On AWS, at Application Load Balancer (ALB), SSO openid configurations are enabled to enforce SSO authentication for each user request.

**SSO Authentication Flow:**



### Data Classification and Export Control

Describe the solution provided to handle the Data Classification and/or Export Control within this Application release. Add in Diagrams tables to summarize the concept and mechanism. Provide the links/references to the detailed design specifications.

A subset of the data and tools are designated export controlled and can only be accessed by individuals who meet the export controlled requirements. These individuals are segmented within Active Directory and are partitioned from this data and tools based on AD permissions and network security controls. This will be the controlling mechanism to managed access to export-controlled data.

One identified risk is sudo access potentially elevating a non-EC cleared employee into permissions not granted through active directory. This will be mitigated by limiting sudo access to US citizens only. This will exclude anyone in the US from elevating privileges who is not a US citizen. There is no export-controlled data in the India deployment so there is no need to limit sudo access in this case.

Managing access to sudo will be managed by creating or using existing Active Directory user group for US citizens. You would designated this user group in /etc/sudoers file with “%ActiveDirectoryUserGroup ALL=(ALL:ALL) ALL” . This will grant sudo access to “ActiveDirectoryUserGroup”. Then, we will remove sudo access from any remaining groups in /etc/sudoers in the US. We will not adjust this file in the India deployment.

Further, per AWS agreement, GE Vernova controls the data. Customers determine who can access your data. Using AWS Regions, you control where your data is stored, based on your specific needs. We make it easier for you to encrypt your data in transit and at rest. We do not access, use, or share customer data without your agreement, except as required to prevent fraud and abuse, or to comply with law, as described in our Customer Agreement.

* AWS SOC 2 Security, Availability, Confidentiality & Privacy Report available to AWS customers from [AWS Artifact](https://aws.amazon.com/artifact/getting-started/).

**SOC 2: Security, Availability, Confidentiality & Privacy:** A description of the AWS controls environment and external audit of AWS controls that meet the AICPA Trust Services Security, Availability, Confidentiality, and Privacy Criteria

### Error Handling and Logging

Describe how exceptions and errors are handled by completing the following table.

Also, discuss below how application design accounts for logging application information. Is the logging persisted? If so, what is it saved to (file, database, SNMP messages, publish messages on a JMS Queue)?

Design yet to be finalized

| **Mechanism** | **Internal Log Message** | **External user constraints** |
| --- | --- | --- |
| CloudWatch |  |  |
| CloudWatch Logs |  |  |

@Yasen – what is expected to be captured? Application & Infrastructure logs?

* Application level logs

### Password Management

| **Component** | **Password Storage** | **Comments** |
| --- | --- | --- |
| RDS Passwords | AWS Secrete Manager |  |

### Security assessment

Work in progress. Need update few items once security assessment is done

AWS Well Architected Framework Review—Security Pillar Outcomes Matrix (9/20-23)

|  |  |
| --- | --- |
| Security Principle | Question-based Risk |
| Secure Operations | High |
| Identity Management | High |
| Permissions Management | Medium |
| Incident Detection | None |
| Network Protection | High |
| Compute Resource Protection | Medium |
| Data Classification | Medium |
| Data at Rest | None |
| Data in Transit | High |
| Incident Response | Medium |
| DevSecOps | High |

#### Secure Operation:

Need a completed thread model assessment. A threat model is a living document and should continue to evolve as your workloads change. Revisit your threat models over time, including when there is a major change, a change in the threat landscape, or when you adopt a new feature or service. [Pending]

#### Identities and Permission:

Clarify Wiz responsibilities – 1/ Audit and rotate credentials regularly, 2/ account access (cross account & public), 3/

#### Network Protection:

Network layer inspection for east-west traffic. Currently not in place, CloudWAN segmentation is open to all internal access.

#### Compute Protection:

"Validate software integrity" - we are not doing much here at present"Reduce manual management and interactive access" - goal is to have very little to no manual intervention"Automate compute protection" - make sure we have cron log trimming

#### Data Classification:

Automate identification and classification --- look for ways to identify EC data and flag for compliance violations or risks

"Define scalable data lifecycle management" – clarify how data on VDI instances will be lifecyled off or restricted from storing EC data.

#### Encryption of Data in Transit:

"Implement secure key and certificate management" - URL specific SSL certs are stored in ACM"Enforce encryption in transit": Traffic within the VPC is marked with (\*), crossing GE network with ($), across the Internet (I) - ($) SMB v3.1.1 - (\*) NFS v4 - ($) SSH/SFTP - ($) HTTPS - ($) License servers (custom encryption) - (I) from VPC to service endpoints - always use HTTPS; enforce HTTPS access in bucket policy - (I) Box, etc. - (I) yum repos for OS updates

In-transit encryption, Lack of Kerberos for SMB – Inherent flaw because of lack of trust between CloudAD and LOGON

#### Incident response:

"Run simulations" - Dave Strum: we are in the middle of a Red Team/Blue Team Cloud audit so you can eventually check that off the list."Prepare forensic capabilities" - Platform logs and network logs analytics + EC2 ability to contain an instance and take snapshots and do OS forensic analysis"Develop and test security incident response playbooks" - Matt confirmed playbooks exist for GE Vernova

"Pre-deploy tools" - Endpoint security agents (ex. Crowdstrike) and forensic collection tools are available on all EC2

#### DevSecOps Process:

"Perform regular penetration testing" - plan to ask for Pen test for web app and/or other aspects"Deploy software programmatically" - striving to have all apps deployed via code"Regularly assess security properties of the pipelines" - Git PR approval processes; who can access/change the pipeline?

"Train for application security" - consider AWS certifications for team members"Automate testing throughout the development and release lifecycle" - we use Aqua and Wiz scanning to do IaC misconfiguration assesment; manual load testing"Manual code reviews" - done for custom code"Centralize services for packages and dependencies" - typically we pull from source repos directly (BAD?)"Build a program that embeds security ownership in workload teams" - EVM reports are avail to app teams(s)

|  | **Description** | **Notes** |
| --- | --- | --- |
| Did the team submit the Project Complexity & Governance Assessment workflow? Provide link to workflow in Notes column. (see link to workflow below) | Choose an item. |  |
| Did the team complete the GE Power Application Security Checklist? (see link below) | Choose an item. |  |
| Does the release require a Security Exception? | Choose an item. |  |
| [Data Classification](http://supportcentral.ge.com/products/sup_community_portal_home.asp?prod_id=262056) | GE Confidential |  |
| SAST/DAST  (Externally facing applications or applications that are critical systems for business continuity require SAST/DAST and SSL (https URL) | Yes (describe) |  |
| How will users log into the system? (Authentication) | Other (describe in Notes) | MFA PingID |
| How will system control access? (Authorization) | Other (describe) | Role based access |
| Describe the creation, deletion and changes to authorization on a system (Access Provisioning) | IDM | Connected with AD authorization |
| Encryption protocols employed | Other or multiple (describe) | SSL & SSH |
| Functional Accounts (non-DB related) - Does your application use functional accounts (non-personal SSO)as part of the architecture? \* | Yes (describe) | Service /Master Payer Account |
| Does your application require that multiple users log into the application with the same user name and/or password? | No | No |

## Network Architecture

Summarize the network architecture topics and how it is provided with the solution. Typical topics are Network information and communication, protocols.

Ensure that security review results with the security team were reflected here in this section.

A dedicated VPC would be provided to CPL application deployment. The details to be updated once infrastructure is provided.

### Network Design

Summarize the relevant sites which are part of the Application network communication. This includes the latency and bandwidth between the sites as minimum (e.g. Central DC site to its regional DC’s, the regional DC’s with the client sites).

This section covers the network capabilities and characteristics within the overall LAN/WAN design.

<https://confluence.apps.gevernova.net/devspace/pages/viewpage.action?pageId=108076483#GEPRJ5298587ARCENG|HPCProjectforARC[DEV]-(DESIGN)ArchitectureDiagram>

### Network Communication

This section summarizes the applied communication protocols and ports between the major application components on server/client side and over the LAN/WAN.

The details to be updated once infrastructure is provided.

1. CloudWAN network segmentation only north-south for now; east-west is open for internal traffic
2. Release References

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Document Name** | **Description** | **Link** |
| 1 | Project Brief & Business Case | Gate presentation of Scope and Business Case |  |
| 2 | Business Requirements | GE Box Folder with all Business Requirements |  |
| 3 | Functional Design Specifications | GE Box Folder with all Functional Specs |  |
| 4 | Technical Design Specifications | GE Box Folder with all Technical Specs |  |
| 5 |  |  |  |
| 6. |  |  |  |
| 7. |  |  |  |
| 8. |  |  |  |

1. External References

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Subject** | **Description** | **Link** |
| 1 | ITAR | International Traffic in Arms Regulations:  Control the export and import of defense-related articles and services on the [United States Munitions List](https://en.wikipedia.org/wiki/United_States_Munitions_List) (USML) | <http://www.pmddtc.state.gov/regulations_laws/itar.html> |
| 2 | EAR | Export Administration Regulations:  US Code of Federal Regulations Title 17, chapter VII, subchapter C | <http://www.ecfr.gov/cgi-bin/text-idx?gp=&SID=8dfe4a8caaadb2617899e0228f49ec0f&mc=true&tpl=/ecfrbrowse/Title15/15CVIIsubchapC.tpl> |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6. |  |  |  |
| 7. |  |  |  |
| 8. |  |  |  |

1. Standards & Guidelines

| **No.** | **Document Name** | **Description** | **Link** |
| --- | --- | --- | --- |
| 1 | Solution Selection Guidelines |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/processes/Documents/Architecture%20Principles%20Applied%20to%20a%20New%20Solution%20Choice.pptx> |
| 2 | Alstom Architecture Standard Profile |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/standard/SitePages/Home.aspx> |
| 3 | ADD User Guide |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/processes/Shared%20Documents/ADD%202.0%20User%20Guide.pptx> |
| 4 | Alstom Architecture Principles |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/principles/Shared%20Documents/Alstom%20Enterprise%20Architecture%20Principles.pdf> |
| 5 | Alstom Process Classification Framework |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/Taxonomies/Shared%20Documents/Alstom%20Capabilities%20Classification%20Framework%20V3.xlsx> |
| 6 | Alstom Business Data Objects Classification |  | <https://teamspace.alstom.com/sites/ASCArchitecture_Team/framework/Taxonomies/Shared%20Documents/Alstom%20Business%20Data%20Objects%20Classificaton%20V2.xlsm> |
| 7 | Data Object Inventory | Data object definitions managed by CDO (Chief Data Office) | <https://teamspace.alstom.com/sites/TpSectorIS_Community/Enterprise-Architecture/CDO/Key%20Documents/CDO_DataObjectInventory.xlsx> |
| 8 | GE IT Policy Community | Consists GE IT Policies, GE IT Standards, GE IT Procedures, GE IT Bulletins and GE IT Leading Practices | <http://libraries.ge.com/foldersIndex.do?entity_id=2630983101&sid=101&SF=1#25459001101> |
| 9 | PT SA Wiki |  | <http://libraries.ge.com/LibrariesWiki/28925198101/Links> |
| 10 | PT - SA working library (Box) |  | <https://ge.box.com/s/i3nf5myzltld8v4e2m4f> |
| 11 | PIT Web Ops Portal |  | <https://webops.pw.ge.com> |
| 12 | planET |  | <http://planet.pw.ge.com/planet/Home.aspx?gpstate=none,it_0> |
| 13 | AWS Ref Architecture Templates |  | <http://libraries.ge.com/foldersIndex.do?entity_id=51846442101&sid=101&SF=1> |
| 14 | Power Security - IT Controls Playbook |  | <https://devcloud.swcoe.ge.com/devspace/display/YFBSU/Home> |

1. Glossary

Terminology, Definitions and Abbreviations:

| **Notation** | **Object/Description** |
| --- | --- |
| **AM** | **A**ccess **M**anager Teamcenter application that enables the system administrator to grant users access to Teamcenter objects. |
| **APC-RPDM** | **A**LSTOM **P**ower **C**omponents - **R**eal **P**roduct **D**ata **M**anagement |
| **APC-VPDM** | **A**LSTOM **P**ower **C**omponents - **V**irtual **P**roduct **D**ata **M**anagement |
| **COBIT** | **C**ontrol **Ob**jectives for **I**nformation and Related **T**echnology (IT-Governance framework) |
| **Corporate Server** | Host computer at the center of a Teamcenter network. This host contains the Teamcenter application root directory, Teamcenter data directory, licensing, file managers (Teamcenter File Services and File Management System), and volumes. For installations that include the Web tier (four-tier architecture), the corporate server also contains the Teamcenter server manager. Multiple application clients can map to or mount the corporate server. |
| **CRT** | **C**e**r**tifica**t**ion/Integration (beta) |
| **DEV** | **Dev**elopment |
| **EAR** | **E**nterprise **Ar**chive. Enterprise application that requires a J2EE application server |
| **EPI** | **E**nterprise **P**rogram **I**nterface |
| **ERP** | **E**nterprise **R**esource **P**lanning |
| **FCC** | **F**ile **C**lient **C**ache |
| **FMS** | **F**ile **M**anagement **S**ystem |
| **Four-Tier Architecture** | Teamcenter architecture that includes four tiers: resource tier, client tier, Web tier, and enterprise tier. |
| **Four-Tier Deployment** | Deployment of the Teamcenter four-tier architecture. The Web tier, enterprise tier, resource tier, and client tier can each be hosted on the same or separate computers. |
| **FSC** | **F**ile Management System **S**erver **C**ache |
| **FSI** | **F**ield **S**ervice **I**nstruction |
| **HAL** | **H**igh **A**vailability **L**icense server |
| **IAT** | Development/Deployment **I**nternal **A**cceptance **T**est |
| **IDSM** | **I**ntegrated **D**istributed **S**ervices **M**anager A network node that runs a daemon process to handle the transfer of data objects among sites (databases) within a Multisite environment. One IDSM server node must be designated for each Teamcenter database from which objects are published; each server node can act only for one database. |
| **ITIL** | **IT** **I**nfrastructure **L**ibrary (de-facto standard of best practices for IT Service Management) |
| **ITSM** | **IT** **S**ervice **M**anagement(discipline for managing IT systems with a customer centric view) |
| **LM** | **L**icense **M**anager |
| **LOV** | **L**ist **o**f **V**alues |
| **LPAR** | **L**ogical **Par**tition |
| **LUM** | **L**icense **U**se **M**anagement |
| **NS** | **N**aming **S**ervice Service that maintains a list of names and the objects associated with them. For the Teamcenter rich client, the naming service maintains a list of servers and the objects and methods used by those services. The naming service is an integral part of the server directory. |
| **ODS** | **O**bject **D**irectory **S**ervice |
| **OOTB** | **O**ut **O**f **T**he **B**ox |
| **POM** | **P**ersistent **O**bject **M**anager Interface between Teamcenter objects and the Relational Database Management System (RDBMS). The persistent object manager provides definition of classes by inheritance from other classes and definition of attributes, manipulation of in-memory objects and support for their saving and retrieval to and from the underlying RDBMS, support for applications accessing the same data concurrently, protection against the deletion of data used by more than one application, and support for the access control lists attributed to objects. |
| **PRD** | **Pr**o**d**uction |
| **PRO** | **Pr**oduct **O**rder |
| **QPL** | **Q**uick **P**art **L**ocator Component of Repeatable Digital Validation/Quick Part Locator (RDV/QPL) that creates and queries tables in an Oracle database. The quick part locator enables rapid lookup of components in an assembly by their location or properties. These tables cache pre-computed results of location, attributes, and spatial relationship of all components in a product assembly. |
| **Resource Tier** | Teamcenter architectural tier comprising the database Server, database file Servers, and volumes |
| **Rich Client** | Java-based user interface to Teamcenter installed on user workstations. The rich client accesses Teamcenter databases using a remote or local server |
| **RSA** | **R**ational **S**ystem **A**rchitect |
| **Server Manager** | Process that manages a pool of Teamcenter Server processes in a deployment of the four-tier architecture. The Server manager starts and times out a configurable number of Server processes to communicate with the Teamcenter database. A Server assigner process assigns available Server processes to user sessions. The Server manager communicates with the Web tier application using either TCP or multicast protocol. |
| **Server Pool** | Pool of Teamcenter Server processes running in the enterprise tier. A small deployment may have only one pool of Server processes. For larger deployments, the pool of Server processes is distributed as sub pools across multiple hosts, with a Server manager for each sub pool. Server pools are applicable for deployments of the Teamcenter four-tier architecture only. |
| **SID** | Oracle **S**ystem **Id**entifier Alphanumeric text used to identify a collection of processes and associated memory structures as belonging to a particular Oracle database instance. The ORACLE\_SID environment variable defines the Teamcenter-Oracle system identifier. |
| **TCE** | **TC** **E**ngineering |
| **TCFS** | **T**eam**c**enter **F**ile **S**ervice |
| **TDP** | **T**echnical **D**ata **P**ackage |
| **TEM** | **T**eamcenter **E**nvironment **M**anager Tool with a wizard-style interface that installs Teamcenter servers and two-tier and four-tier rich clients. TEM also performs maintenance operations, such as upgrading servers, applying maintenance packs, and installing patches. Teamcenter installers launch TEM using the tem.bat command (on Windows systems) or the tem.sh command (on UNIX or Linux systems). |
| **Thin Client** | Teamcenter user interface that provides a streamlined browser-based view of product information stored in a Teamcenter database. The thin client is configured in the Web tier, which creates and serves its Web pages to the client. |
| **TMI** | **T**C Engineering **M**etaphase **I**nterface |
| **Transient Volume** | Operating system directory controlled by Teamcenter and used to store temporary data for transport of reports, PLM XML, and other non-volume data between the Web tier and client tier in a deployment of the Teamcenter four-tier architecture. |
| **Two-Tier Architecture** | Teamcenter architecture that includes a resource tier and a client tier. The resource tier comprises the database Server and database. The client tier comprises the Teamcenter rich client, third-party applications that integrate with the rich client, and a local Server. This architecture supports only the Teamcenter rich client. |
| **Two-Tier Deployment** | Deployment of the Teamcenter two-tier architecture. In a typical deployment of the two-tier architecture, the rich client and its local Server are installed on a user's workstation as are third-party applications that integrate with the rich client. The database Server and the Teamcenter corporate Server are installed on one or more separate computers. |
| **VLD** | **V**a**l**i**d**ation (alpha) |
| **Web Tier** | Teamcenter architectural tier that comprises a Java application running in a Java 2 Enterprise Edition (J2EE) application Server. The Web tier is responsible for communication between the client tier and enterprise tier. The Web tier application also includes the Application Interface Web Service (AIWS), WebDAV service, and thin client. |