INL/RPT-22-70543
Revision 0

Light Water Reactor Sustainability Program

Development of a Cloud-based Application to Enable a Scalable Risk-informed Predictive Maintenance Strategy at Nuclear Power Plants

December 2022

U.S. Department of Energy

Office of Nuclear Energy

INL/RPT-22-70543
Revision 0

Development of a Cloud-based Application to Enable a Scalable Risk-informed Predictive Maintenance Strategy at Nuclear Power Plants

Rita Appiah, Cody Walker, Vivek Agarwal, Jonathan Nistor, Tom Gruenwald, Michael Muhlheim, Pradeep Ramuhalli

December 2022

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy

ABSTRACT

Light-water reactor operations and maintenance (O&M) costs are prohibitively high, thus contributing to the premature decommissioning of nuclear power plants (NPPs). This is partly due to how the equipment is monitored. In recent years, cloud computing has emerged as a dominant technology by virtue of its low costs, computing and storage adaptability, and ability to host applications over numerous types of virtual infrastructures. Cloud computing can be a cost-effective alternative to onsite storage and diagnostics. This paper conducts a techno-economic assessment of a provisional cloud deployment architecture for a NPP predictive monitoring (PdM) system. The cloud-based monitoring system would enable maintenance and diagnostics (M&D) analysts and other authorized plant users to remotely monitor equipment functionality so as to enable PdM practices and early detection of faults. The Microsoft Azure cloud platform is included in the proposed cloud architecture to provide data processing and storage, sensor device networking, and database management; however, this analysis could be extended to other cloud computing service providers as well. For the techno-economic assessment, technical feasibility is measured in terms of network performance metrics such as response time, latency, and throughput, whereas economic feasibility is measured in terms of operational costs and capital expenditures. Finally, this report covers certain regulatory and security aspects that may concern licensees looking to implement cloud computing. The report focuses on the integration of sensor database storage, the application of cloud resources to PdM, and the identification of technological and economic hurdles associated with moving to a cloud-computing-based architecture.

ACKNOWLEDGEMENTS

This report was made possible through funding from the U.S. Department of Energy (DOE)'s Light Water Reactor Sustainability program. We are grateful to Jason Tokey of DOE and Bruce P. Hallbert and Craig A. Primer at Idaho National Laboratory (INL) for championing this effort. We thank John M. Shaver and Judy Fairchild at INL for the technical editing and formatting of this report, respectively. We thank Barry Pike III and Lauren M. Perttula of RED, Inc. for some of the graphics contained in this report.

CONTENTS

ABSTRACT iii

ACKNOWLEDGEMENTS v

ACRONYMS x

1. INTRODUCTION 1
   1.1 Report Scope 1
   1.2 Organization of This Report 1

2. CLOUD COMPUTING 2
   2.1 NPP Cloud Transition Strategies 4
      2.1.1 Path Forward for the NPP Cloud Integration/Transfer Approach 5
   2.2 Framework of the Proposed Architecture 6
   2.3 Communication with Cloud Services 7
      2.3.1 Wireless Sensor to Cloud Services 7
      2.3.2 NPP to Cloud Services 9
   2.4 Cloud Architecture and Storage 11
      2.4.1 Data Ingest and Pipelines 12
      2.4.2 NPP Cloud Data Processing Architectures 13
      2.4.3 Data Storage 14
      2.4.4 Other Features 17
   2.5 Processing and Predictive Maintenance 19

3. REGULATIONS 29
   3.1 The Federal Government's Move to Cloud Computing 29
   3.2 NRC Adoption of Cloud Smart 30
   3.3 Regulatory Requirements for Cloud-based Servers 31
      3.3.1 Protection of Digital Computer and Communication Systems and Networks 31
      3.3.2 Export Control 33

4. ECONOMICS 33
   4.1 Sensor Costs 34
      4.1.1 Sensor Configuration 34
      4.1.2 Costs 36
   4.2 Distributed Antenna System Costs 37
      4.2.1 DAS Considerations 37
      4.2.2 Model Facility 38
      4.2.3 Costs 40
   4.3 Onsite Edge Computing Costs 41
   4.4 Cloud Costs 42
      4.4.1 Cloud Infrastructure 42
      4.4.2 Architecture Overview 42
      4.4.3 Cost Estimates for the NPP Feedwater and Condensate System 43
   4.5 RFID Network Costs 45
      4.5.1 System Components 45

4.5.2 RFID System Cost 47
4.6 Cost-Benefit Analysis 47
4.6.1 Present Costs (Base Case) 47
4.6.2 Cloud-Based Cost Summary 49
4.6.3 Cost-Benefit Analysis 50
5. SUMMARY AND PATH FORWARD 50
6. REFERENCES 50

FIGURES

Figure 1. Available cloud provider services and managed responsibilities. 3
Figure 2. Proposed high-level architecture of the hybrid cloud. 7
Figure 3. Wireless-sensor-to-cloud-database layout. 8
Figure 4. NPP-to-cloud-network connectivity. 10
Figure 5. NPP site edge server interface. 11
Figure 6. Data ingestion and egestion pipeline layout for the NPP cloud. 12
Figure 7. Lambda cloud data flow architecture. 14
Figure 8. Kappa cloud data flow architecture. 14
Figure 9. NPP cloud storage ingress and egress charge model. 16
Figure 10. NPP cloud storage layout. 17
Figure 11. Latest version of the AVEVA Predictive Analytics early detection model output [20]. 20
Figure 12. Layout of the AVEVA Predictive Analytics data hub [21]. 20
Figure 13. NPP cloud predictive maintenance layout. 22
Figure 14. Flow chart of the NPP cloud PdM infrastructure. 23
Figure 15. Near-real-time data analytics framework for NPP predictive maintenance. 24
Figure 16. Proposed modified cloud signal processing pipeline for NPP sensors [28]. 25
Figure 17. Proposed real-time FLSTM topology for NPP PdM. 29
Figure 18. Abstracted sensor data pipeline, with a focus on wireless sensors (in red, on the left-hand side of the diagram). 35
Figure 19. A sensor (thermocouple) married to a combined A/D converter and wireless transmitter. 35
Figure 20. An integrated vibration sensor with A/D conversion, down-sampling logic, and a wireless transmitter (also includes a magnetic mount) [55]. 36
Figure 21. Top: Basic DAS structure. Bottom: Components of a passive DAS system. 38
Figure 22. The DC Cook facility in Berrien County, Michigan, features two functioning power plants [58]. 39
Figure 23. DC Cook facility in Berrien County, Michigan, highlighting the six warehouses (W1–W6) [58]. 40

Figure 24. Abstracted sensor data pipeline, with a focus (in red) on onsite edge computing. 41
Figure 25. Breakdown of data generated during a full 2-year cycle at a commercial boiling-water reactor. 43
Figure 26. Microsoft Azure real-time analytics scenario used for cost estimation [60]. 44
Figure 27. General layout for an RFID System. 46
Figure 28. Impinj Speedway Revolution R420 ultra-high frequency RFID Reading (left), RFMAX Indoor RFID Antenna (middle), and Vulcan RFID printable labels (right) [61-63]. 46

TABLES

Table 1. Microsoft Azure cloud resources and charges (adapted from [17]). 18
Table 2. Range of sensor types required for diagnostic monitoring of NPP equipment. 35
Table 3. Estimate of sensor and installation costs associated with instrumenting the NPP feedwater and condensate system for online M&D [55-56]. 36
Table 4. Cost estimates based on coverage requirements and square footage for DAS equipment and the installation and commissioning thereof [58]. 41
Table 5. Costs pertaining to the onsite edge computing infrastructure for handling approximately 600 data streams associated with the feedwater and condensate system. 42
Table 6. Cloud cost estimates obtained utilizing Azure [60]. 44
Table 8. Total RFID system cost for model facility with 3 large (20,000 square-foot) and 3 small (10,000 square-foot) warehouses. 47
Table 10. Summary of cost categories and amortized costs associated with cloud-based implementation of a risk-informed predictive maintenance program 49

ACRONYMS

A/D	analog/digital
AI	artificial intelligence
AMQP	Advanced Message Queuing Protocol
API	application programming interface
AWS	Amazon Web Services
BTS	Base Transceiver Station
C2D	cloud to device
D2C	device to cloud
DL	deep learning
DOC	Department of Commerce
FEDRAMP	Federal Risk and Authorization Management Program
FLSTM	federated long short-term memory
HTTP	hypertext transfer protocol
IaaS	Infrastructure as a Service
IoT	Internet of Things
IP	Internet Protocol
IT	information technology
JSON	JavaScript Object Notation
LSTM	long short-term memory
M&D	maintenance and diagnostics
ML	machine learning
MQTT	Message Queuing Telemetry Transport
NIST	National Institute of Standards and Technology
NPP	nuclear power plant
NRC	Nuclear Regulatory Commission
O&M	operations and maintenance
OMB	Office of Management and Budget
OPC	Open Platform Communications
PaaS	Platform as a Service
PdM	predictive maintenance
RFID	radio frequency identification
RUL	remaining useful life
SCADA	supervisory control and data acquisition

SDK	software development kit
SQL	structured query language
SVM	support vector machine
VM	virtual machine
VPN	virtual private network
WLAN	wireless local area network

Development of a Cloud-based Application to Enable a Scalable Risk-informed Predictive Maintenance Strategy at Nuclear Power Plants

1. INTRODUCTION

Light-water reactor operations and maintenance (O&M) costs are prohibitively high, thus contributing to the premature decommissioning of nuclear power plants (NPPs). This is partly due to how the equipment is monitored. In recent years, cloud computing has emerged as a dominant technology by virtue of its low cost, its ability to host applications on various types of virtual infrastructures, and its computing and storage adaptability. Cloud computing can be a cost-effective alternative to onsite storage and diagnostics. One practical example of this is seen in conducting anomaly detection on mechanical components. Sensor drift, pump failures, and damaged motor bearings require multi-modal equipment condition monitoring. To efficiently utilize signal processing algorithms for fault diagnosis, enhanced real-time data-driven machine learning (ML) and artificial intelligence (AI) techniques should be employed. Implementing a real-time predictive maintenance approach using advanced AI models can alleviate the costs associated with periodic equipment repairs, labor, and unnecessary outages [1]. The NPP industry seeks to leverage cloud computing's elastic processing power and high-availability computing resources by advantageously shifting data storage locations from computers located at plant sites to cloud servers with integrated security management. Additional savings can be realized via the cloud service provider's pay-as-you-go method, which enables users to pay only for the resources required for the task at hand.

1.1 Report Scope

The goal of this report is to propose a NPP cloud architecture for database storage and analytics. This report is an extension of prior documents regarding the implementation of digital monitoring at NPPs. These preceding documents evaluated prospective wireless distributed antenna systems (DASs) and Long Term Evolution (LTE) communication networks that would be feasible for NPP digital sensors with wireless technologies [2]. ML capabilities, predictive maintenance (PdM) optimization, and data visualization techniques were also reviewed [3].

This report covers a techno-economic assessment of a provisional cloud deployment architecture for a NPP predictive monitoring system. This cloud-based monitoring system would enable maintenance and diagnostics (M&D) analysts and other authorized plant users to remotely monitor equipment functionality so as to allow for early detection of faults. The Microsoft Azure cloud platform is included in the proposed cloud architecture to provide data processing and storage, sensor device networking, and database management, though this analysis could just as easily be extended to other cloud computing service providers. Technical feasibility is measured in terms of network performance metrics such as response time, latency, and throughput, whereas economic feasibility is measured in terms of operational costs and capital expenditures. Finally, this report covers regulatory aspects that may concern licensees looking to implement cloud computing. Such aspects include data storage and export control concerns. This report focuses on integrating sensor database storage, applying cloud computing resources to PdM, and identifying any associated technological, economic, or regulatory hurdles.

1.2 Organization of This Report

Section 2 gives an overview of cloud computing and deployment methods. Section 3 covers the regulatory aspects that licensees should be aware of when considering the use of cloud computing resources. Section 4 compares the current costs of operating an onsite M&D center with the costs of operating a similar system using cloud resources.

2. CLOUD COMPUTING

Cloud computing refers to utility-based computing resources accessed over the internet. On the cloud platform, virtualization techniques permit the creation of multiple simulated environments and resources to be generated by a single physical hardware system through a type of software referred to as a hypervisor. The hypervisor provides an interface to various resources hosted on physical hardware systems and distributes them appropriately into secure environments known as virtual machines (VMs). Services, applications, and infrastructure resources on the cloud are available to users on-demand through network access, with resource pooling and rapid elasticity (i.e., automatic scaling of dynamic resources). Subscribing to cloud platforms require licensing agreements/management, security compliance, certificates, and the meeting of any regulatory measures required by your cloud provider. The advantages and disadvantages of various cloud platforms are discussed in [4].

National Institute of Standards and Technology (NIST) Special Publication 800-145 describes cloud computing as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [5]. The cloud model in that document consists of five essential characteristics, three service models, and four deployment models.

The five essential characteristics are defined as follows:

1. ***On-demand self-service****:* A consumer can unilaterally provision computing capabilities (e.g., server time and network storage) as needed, without the need for human interaction with each service provider.
2. ***Broad network access****:* Capabilities are available over the network and are accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
3. ***Resource pooling****:* The provider's computing resources are pooled to serve multiple consumers via a multi-tenant model, with different physical and virtual resources being dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, and network bandwidth.
4. ***Rapid elasticity****:* Capabilities can be elastically provisioned and released to scale rapidly outward and inward as commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.
5. ***Measured service****:* Cloud systems automatically control and optimize resource usage by leveraging a metering capability at a level of abstraction appropriate to the type of service employed (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported on, providing transparency for both the provider and consumer of the utilized service.

The three main types of cloud services, as visually summarized in Figure 1, are described as follows:

1. ***Infrastructure as a Service (IaaS)***: This capability allows the consumer to provision processing, storage, networks, and other fundamental computing resources that enable the consumer to deploy and run arbitrary software, potentially including operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over the operating systems, storage, and deployed applications, as well as possibly limited control over select networking components (e.g., host firewalls). As shown in Figure 1, the customer is usually responsible for application maintenance, administration (e.g., middleware and run time), and data. These resources can be dynamically split and resized based on demand. Some examples of IaaS

services include DigitalOcean Amazon Web Services (AWS) EC2, Google Compute Engine, Azure VM and Containers, and Rackspace [6]. Industries may consider this option as solutions for high-performance computing, application development and production, as well as private database management for security purposes. This service is considered but should be used only by exception.

2. ***Platform as a Service (PaaS)***: This capability allows the consumer to deploy onto the cloud infrastructure consumer-created or acquired applications that were developed using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure, which includes the network, servers, operating systems, and storage, but can control the deployed applications and possibly the configuration settings for the application hosting environment. PaaS is usually built on top of IaaS, with the cloud provider managing the run time and middleware, as seen in Figure 1. It also features built-in scaling, load balancing, and a variety of services for messaging authentication and caching. For example, Azure Logic Apps helps speed up the application deployment and configuration processes. Examples of PaaS services include OpenShift, AWS Elastic Beanstalk, Microsoft Azure Content Delivery Network, Microsoft Azure App Service, and Apache Stratos [6]. These may also be considered for use in application development and data management. PaaS is considered as a potential candidate for the NPP cloud, since it offers the flexibility to customize integrated application programs and drives technology standardization for modernized systems and applications that require customization.

3. ***Software as a Service (SaaS)***: This service affords the customer the advantage of using the cloud provider's infrastructure and running applications without having to manage or maintain them. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email) or through a program interface. The consumer does not manage or control the underlying cloud infrastructure (e.g., the network, servers, operating systems, and storage) or even individual application capabilities, apart from limited user-specific application configuration settings. Examples of such cloud services include Google Apps, HubSpot, Dropbox, DocuSign, Office 365, and Gmail [6]. Industries may consider SaaS for remote management of specific types of software. SaaS can be used to support a low-code deployment approach and to optimize functional requirements.

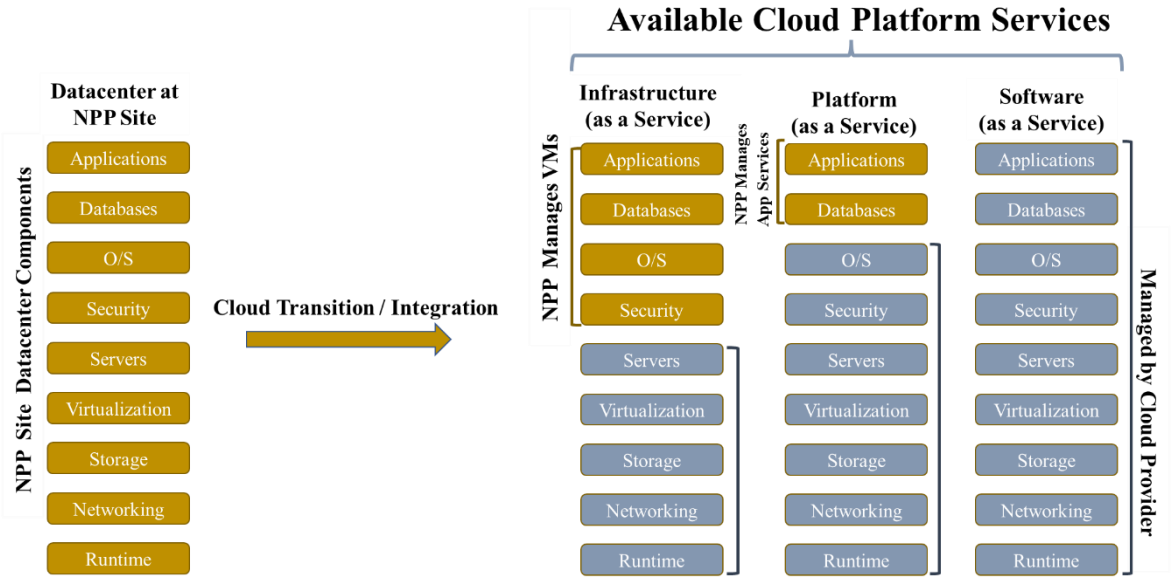

Figure 1. Available cloud provider services and managed responsibilities.

Once a cloud service model has been chosen, it must then be deployed. Deployment models should be selected in consideration of data privacy, encryption, compliance, retention, visibility, and security concerns. The four types of deployment models are summarized as follows:

1. ***Private cloud***: This cloud infrastructure is provisioned for exclusive use by a single organization consisting of multiple consumers (e.g., business units). It may be owned, managed, and operated by that organization, a third party, or a combination of the two, and may exist either on or off premises.
2. ***Public cloud***: This is a conventional platform by which resources and infrastructure can be made available over the internet by third-party offsite providers who are accessible to the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination thereof. The cloud customers are charged depending on the amount of data used. The three major cloud providers are Microsoft Azure, Google Cloud Provider, and AWS. Currently, the technological infrastructures of most NPP sites run on Windows platform, meaning that integrating existing data centers with Microsoft Azure cloud platform would be feasible from a compatibility perspective.
3. ***Community cloud***: This cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations with shared concerns (e.g., mission, security requirements, policy, and/or compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, by a third party, or by some combination thereof, and it may exist either on or off premises. The NPP industry could operate a community cloud in order to meet data security and regulatory requirements.
4. ***Hybrid cloud***: This cloud infrastructure is comprised of two or more of the aforementioned cloud infrastructure types (i.e., private, community, and public), each of which will remain distinct but be bound together by standardized or proprietary technologies that enable data and application portability (e.g., cloud bursting for load balancing between clouds). This hybrid cloud option is recommended for the NPP industry. The NPP industry seeks to expand existing data center capabilities so as to leverage Microsoft Azure public cloud storage and analytics services in order to reduce network latencies and achieve efficient scalability to enhance cost savings in real-time monitoring. Running workloads on the cloud can also be securely managed and optimized privately.

2.1 NPP Cloud Transition Strategies

To maximize the benefits of cloud services, information from the data historian, along with any currently employed procedures and applications, must be transitioned to the cloud. This process can be accomplished through rehosting, refactoring, re-architecting, or rebuilding (these are listed in order of increasing workload to complete). The details and applicable scenarios regarding each process are listed below.

Rehost: This process is often referred to as a "lift and shift," as it entails direct transfer of VMs from an onsite data center to the cloud. In this case, the type of application services is mainly the IaaS option. Databases and compactible file storage can be deployed on both managed VMs and file storage services. Onsite administration security as well as operational management would have to adjust to cloud services. Selecting the right infrastructure service and setting up auto-scaling, automatic backup services, and full infrastructure monitoring would mitigate any transition-related challenges. This approach can be used when the cloud transition process is hindered by a lack of human resources, insufficient time, and financial constraints. It is also feasible for when special cloud platform libraries and utilities are needed to support software application compatibility. If the applications are up to date, code and deployment changes are seldom needed. The two options for transitioning the VMs are Azure VM (IaaS) and Azure VMware Solution.

Refactor: This process refers to making small changes to the onsite codes prior to transitioning to the cloud environment (usually the PaaS). This process offers advantages in terms of availability, scalability,

and reliability, and does not require intense implementation. It can be used when deploying applications using Azure Docker containers in Kubernetes. Again, databases can be transitioned to managed database services on Azure, including Azure Structured Query Language (SQL), Azure SQL Managed Instance, and Azure Database for MySQL/PostgresSQL.

***Re-architect*:** This requires splitting application codes into smaller batches (microservices) for independent deployment on the cloud. The Azure Kubernetes container or serverless services could be used for the implementation. This option is feasible for independent private relational database data that must be transferred to non-relational managed NoSQL databases such as Azure Cosmos DB. Additional requirements may be needed to develop a separate continuous integration and continuous deployment (CI/CD) pipeline for every appropriate programming language in each microservice framework.

***Rebuild*:** This involves creating the system in the cloud from scratch by using required cloud services and practices. The implementation is extensive because most aspects of transitioned codes must be redesigned on the cloud platform. This process can be adopted if the microservices required for the small code partitions must be totally changed to support the relevant cloud infrastructure and development.

NPP cloud transition recommendations: After further analysis of the top three cloud provider services, the Azure cloud provider was explored for the NPP cloud transitions, based on the rationale found in the detailed services review [7], along with the fact that it is highly compactible with Microsoft applications and entails less configuration effort and faster deployment models by virtue of its memory-optimized central processing unit instances. A hybrid cloud platform is preferred, as the existing private cloud employs additional AVEVA cloud services and is envisioned to be integrated with Azure cloud services as proposed in this report. The rehost (i.e., lift and shift) combined with the refactor transitioning process (e.g., both IaaS and PaaS services) represents the best approach for conducting the NPP cloud transition, since the nuclear industry intends to scale and optimize maintenance prediction activities by leveraging cloud resources to reduce costs. The required deployment procedure is described in the subsequent sections.

2.1.1 Path Forward for the NPP Cloud Integration/Transfer Approach

The following is an outline of the technical considerations involved in supporting the recommended IaaS and rehost/refactor Azure cloud services model.

- *VMs and servers*: There is a need to access NPP site servers and to transfer workloads to Azure VM or Azure VMware Solution. The most efficient approach is to create a minimal amount of data for a version of the workload to get the application initially working on the cloud. NPP server applications already running on VMs can be easily transitioned to the cloud platform upon completion of the required application and software version updates.
- *Databases*: Existing NPP site databases can be moved to relational formats in Azure SQL Database, PostgreSQL, or Microsoft SQL Server in Azure VM. On the other hand, the non-relational databases can also be moved to blobs, tables, queues, Azure NoSQL Database, or Azure CosmosDB on the cloud platform. For enhanced performance, scalability, and security, it will be essential to rebuild the data model as a new Azure SQL Database. In addition, a data recovery plan should be considered for rectifying any errors and disasters. The option also exists to synchronize the plant site SQL database with the cloud SQL database servers by using the SQL Data Sync service in the Azure cloud. Large amounts of data can be moved either virtually or manually via network transfer or Azure Data Box, respectively. The NetApp Cloud Volumes ONTAP software package is recommended for cost-effective data storage and management [8].
- *Web applications*: NPP sites' web applications can be transferred to Azure Logic Apps and Azure Kubernetes Service.

- *Virtual desktop*: NPP sites' virtual desktops could also be transferred to virtual desktops in the Azure cloud.
- *Network*: A virtual private network (VPN) with an Azure ExpressRoute is recommended for speeding up data transmission and improving bandwidth performance.
- *Testing and optimization*: Functional and performance tests will be required to confirm application and workload compatibility on the cloud. The performance results could be compared with NPP site performance to resolve scalability and feature functionality issues in the cloud application.
- *Service management*: Workloads transferred to the cloud can be monitored by using Azure Application Insights to collect, analyze, debug, measure performance, allocate resources, troubleshoot, etc., with regards to the service-level agreements on the Azure cloud platform. Microsoft Operations Management Suite could also be employed to remotely manage both NPP site server applications as well the cloud.

2.2 Framework of the Proposed Architecture

In previous years, the nuclear fleet relied on time-consuming labor-intensive preventive maintenance programs to maintain and sustain plant assets leading to high O&M costs. Subsequently, with the increasing number of data-generating sensors, existing equipment reliability and maintenance processes can become complicated and inefficient to detect asset failures at an early stage. As a result, the nuclear industry seeks to leverage scalable cloud resources to detect plant asset disturbances that may elude other maintenance systems in real-time for better maintenance planning with minimal downtimes for safe long term economical operation. The proposed NPP cloud-based high-level architecture depicted in Figure 2 described in this section would be employed to collect and process the multi-source sensor data from various NPP sites for automatic monitoring alerts and diagnosis using advanced AI methods.

NPP site wireless sensor network: The left-hand side of

Figure 2 (i.e., “local plant site”) illustrates the local NPP site connectivity that supports various network types among sensors and actuators, the Internet of Things (IoT) Hub gateway, and VM servers. For local wireless connectivity, a DAS-LTE combines wireless amplifiers and fiber optic cables to distribute wireless signals to antennas over a wide frequency range (kHz to GHz). Large volumes of sensor data are continuously collected at the plant site and transmitted to the cloud via a Wi-Fi router and the IoT Hub gateway point that ensures a high data transmission rate with bidirectional communication. An edge device can be employed to enhance the data processing activities (e.g., data cleaning and feature extraction) prior to sending the data to the cloud resources. The VMs and data servers at the plant site are accessed from the cloud, using a VPN connection to ensure data security. Currently, some data sources are taken periodically by hand, but by deploying new wireless sensors, these infrequent, route-based measurements can be replaced with frequent and reliable sensor measurements. Continuous collection and monitoring of sensor data, as enabled by wireless sensor monitoring, will ensure that equipment remains healthy and is operating within the acceptable limits in a cost-effective manner. Thus, there is a need to ensure that the wireless network can support these new sensors and their corresponding frequency ranges.

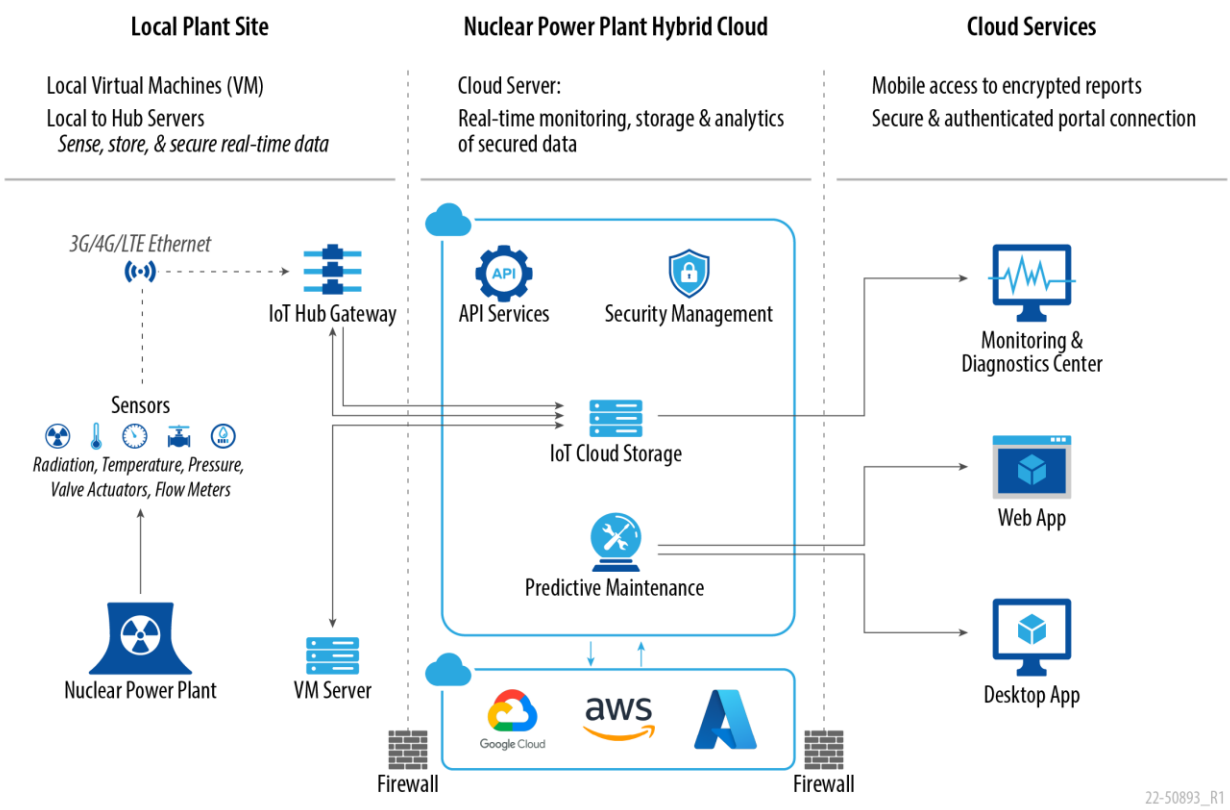

Figure 2. Proposed high-level architecture of the hybrid cloud.

NPP cloud network: The second aspect of the proposed architecture shown in

Figure 2 encompasses further real-time data processing and analytics, storage, and security management aspects. On the cloud platform, a series of applications can be integrated, as needed, into the central IoT Hub infrastructure adaptors. Data routing and authentication of all incoming messages received from various sensor devices and locations is authenticated by the Event or IoT Hub gateway prior to further data processing and AI data analytics. Web applications are also developed and hosted by container and Kubernetes services for real-time visualization of archived historical data.

NPP cloud services: The third aspect of the proposed architecture shown in

Figure 2 encompasses the sharing of analysis reports and data visualizations through a web browser or portal. M&D reports are also made available to authorized plant and technical staff. In addition, administrative management functions are made accessible to front-end users. Using services such as Azure, GE Predix, and AVEVA, cloud service providers have shown that data can be shared and demanded services provided through secure database servers.

2.3 Communication with Cloud Services

2.3.1 Wireless Sensor to Cloud Services

To realize the maximum benefits of cloud computing, sensors must be in regular communication with the cloud computing site. This will most likely entail installation of new wireless sensors to replace the manual, route-based measurements and thereby improve the reliability and consistency of these measurements. With frequent measurements being sent to an onsite database, the data will be refined

prior to being sent, in either a streaming or batch-like fashion, to cloud computing services for further analysis and eventual storage.

The existing plant site DAS-LTE/wireless local area network (WLAN)/long-range wide area network (LoRaWAN) sensor network uses a ASP.NET (Windows) application on a Microsoft SQL server database. High bandwidths (700–950 MHz) are supported by the WLAN/Wi-Fi for low-latency data transmission rates. Signals from low-power, low-bit-rate sensors within a range of 10 km are wirelessly connected via the LoRaWAN. The radio frequency identification (RFID) network also operates in a low signal frequency range (Hz–GHz), with a sensing distance of 1–100 m. This RFID network server will be used to detect whenever equipment is removed from inventory. Other sensor devices must be accommodated as well, including valve position monitoring gauge readers in the 433 MHz frequency range, hundreds of vibration sensors, dozens of dosimetry sensors in the 915 MHz frequency range, and other devices such as smart phones, tablets, and ultrasonic/infrared cameras with a frequency range of 700–850 MHz [9].

The wireless-sensor-to-cloud-database layout shown in Figure 3 is basically applicable to all registered sensor devices at the NPP site. When the sensor device is activated or provisioned, the device-specific schema information (e.g., equipment ID and authentication) is stored in the cloud device registry, irrespective of hardware differences or sensor device model. To ensure a unified telemetry data format for all connected sensor devices, the plant-side device management takes care of all device-to-cloud bi-directional messages, property updates from the NPP site to the cloud, and direct data path communications among individual connected devices. The Azure IOT Hub maintains a device twin for DAS-LTE/WLAN/LoRaWAN sensors, and this twin contains the following:

- *Device tags*: Read/write JavaScript Object Notation (JSON) documents to identify sensor devices and subsequently aid in organizing them.
- *Desired property*: Synchronizes device configuration together with the reported properties. The notification back end (e.g., Azure storage containers) can establish desired properties—as well as changes to desired properties—to be read by the NPP device application on the edge server.
- *Reported property*: This is also used in tandem with the desired property to synchronize sensor device conditions. The NPP site device application can set the reported properties to be read and queried by the cloud back-end service queue.

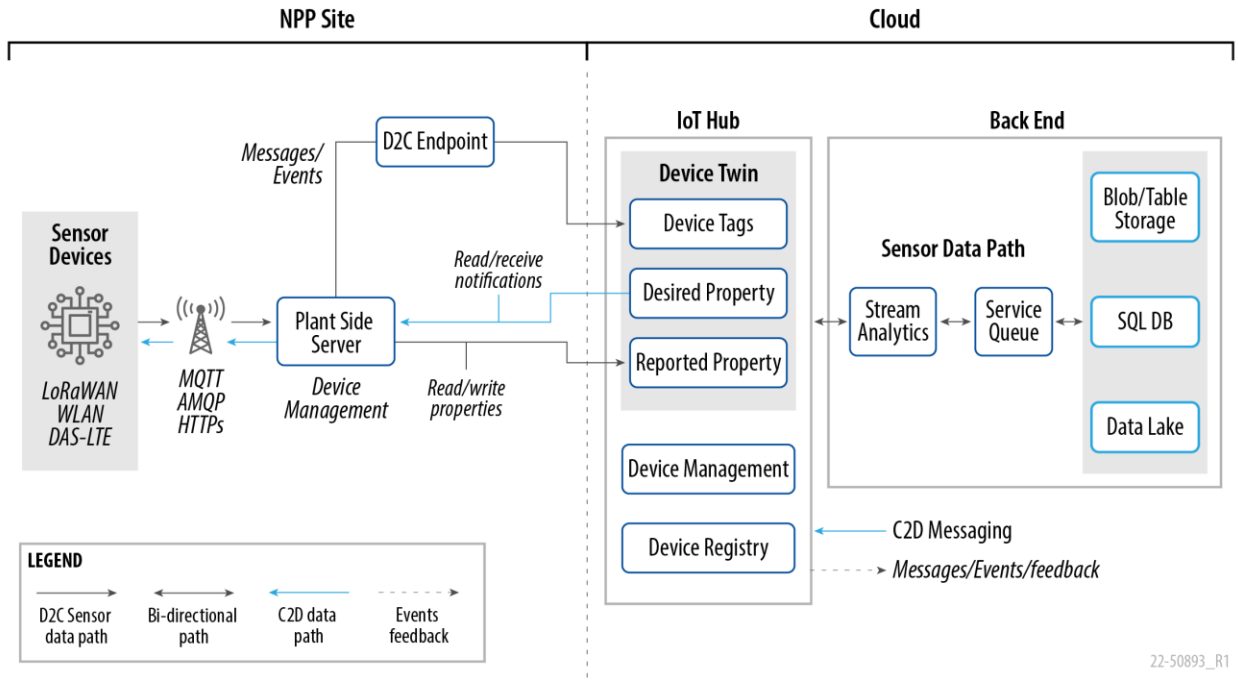

Figure 3. Wireless-sensor-to-cloud-database layout.

To ensure consistency in sensor data throughout the sensor-to-database layout given in Figure 3, the JSON encoding format was adopted. Incoming JSON-formatted messages are automatically routed to different devices, as per the IoT Hub's message structure. Wi-Fi network sensor devices publish sensor data to the Message Queuing Telemetry Transport (MQTT) protocol to aid in constructing a device-to-cloud (D2C) data path. Sensor data from the LoRaWAN low-powered devices are first reported to the plant server. Aggregation of these data is later transferred to the cloud. Similarly, on the cloud side, updates to the desired properties are transmitted to the Wi-Fi devices via the Advanced Message Queuing Protocol (AMQP) or the Hypertext Transfer Protocol (HTTP) notification protocols in order to support the file upload data flow, which is enabled by standalone applications.

2.3.2 NPP to Cloud Services

After wireless sensors transmit their data to onsite plant databases or historians, the plant database must communicate with the cloud. The cloud IoT Hub accomplishes this task and is essential for connecting the plant database to cloud resources. The cloud IoT Hub is at the center of the sensor-to-database layout in Figure 3 above. The cloud IoT Hub management encompasses three types of inputs: property update, bi-directional, and direct message inputs. Each device corresponds to two identical objects: the local device and the digital twin representation on the IoT Hub. The device twin contains device tags, desired properties, and reported properties. Property updating consists of synchronizing the locally desired properties from the plant server to the cloud desired properties, then updating the local reported properties based on the cloud reported properties. Bi-directional (i.e., D2C and cloud to device [C2D]) messages are natively supported by the IoT Hub in order to enable telemetry data transmission and notification delivery. Direct messaging inputs are used to instantly invoke local commands in the end device for execution. The proposed layout is capable of transmitting information from plant site devices to the cloud database for storage. In this scenario, the Azure IoT device software development kit (SDK) is recommended as an open-source software framework for enabling simple and secure cloud gateway management and connectivity. Some supported platforms within the Azure SDK are Windows 7/8/10, Linux (Ubuntu, Fedora, Raspbian), and Android. Languages such as .NET, C#, JavaScript, and C are compatible, as well.

Figure 4 shows the DAS-LTE/LoRaWAN/WLAN sensor network as being the backbone of the layout that enhances the connectivity between the plant—or onsite local infrastructure—and the cloud. Thanks to native Internet Protocol (IP) support devices with Wi-Fi interfaces, some devices can directly communicate with the cloud gateway by using any of the supported messaging protocols (e.g., MQTT and the Constrained Application Protocol) through the JUNIPAR PJAR router. In parallel, the message exchange between resource-limited devices and the cloud is bridged by the edge server. Back-end C2D messages can also be transmitted through the VPN and IPSec/IKE tunnel thanks to the switch to the edge device, then further relayed to the target devices.

Regarding the NPP connectivity network depicted in Figure 4, the various data link and communication protocols have already been mentioned in the previous section. These protocols permit interoperability among devices from different networks. The network components can be described as follows:

- *RFID front-end server*: Radio frequency identification for automatic data acquisition, scalability, and data inventory tracking. The RFID front-end server would act as an inventory management system consisting of RFID readers and RFID tags, as well as a RFID software application. The RFID inventory management system would track and manage the sensor data inventory (e.g., status and location), with real-time updates to cost-effectively automate and improve the accuracy of inventory counts. Again, the RFID batteryless technology provides numerous benefits, including large-scale integration, reduced complexity, and flexible mobility [2]. The digitally processed signals stored on the sever would be transmitted over a transmission control protocol / internet protocol (TCP/IP) network. In addition, the *RFID backhaul server* receives the filtered signals for various frequency bands by employing a modulation scheme (e.g., the quadrature amplitude or the spread spectrum for the WLAN/LTE and LoRaWAN communication protocols, respectively) [2]. The modulated signal is then transmitted over to the cloud for storage, analytics, and monitoring.
- *Edge server*: Acts as a local hub where all sensor data are aggregated and forwarded to the cloud.
- *IoT Hub gateway*: The centralized device that transmits aggregated data to the cloud storage.
- *Express route circuit*: Joins the onsite network with the cloud via routers.
- *Express load balancer*: Express route gateway that reduces latency and connects the private virtual network (VNET) to the cloud subnet.
- *VPN gateway*: Provides external connectivity to the on-premises network.
- *IPsec/IKE tunnel*: A site-to-site VPN gateway connection linking the NPP site to the cloud's virtual network.
- *VPN load balancer*: Routes network traffic from the VPN gateway to the cloud.
- *Virtual machine*: Enables hosting applications such as Windows, Linux, and iPhone Operating System (iOS) to achieve virtual scalability.

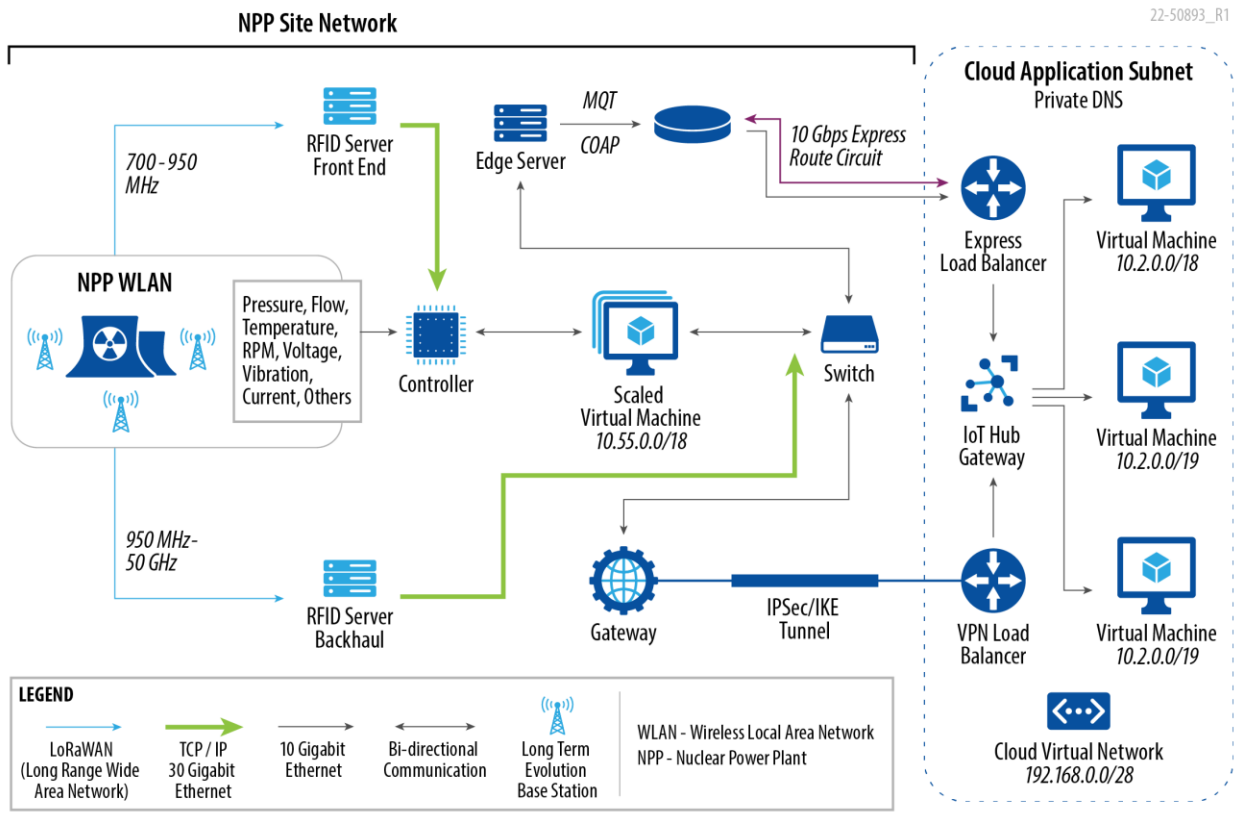

Figure 4. NPP-to-cloud-network connectivity.

The edge server is where all the sensor data are aggregated on the plant side before being forwarded to the cloud. This component is essential because, when used in tandem with methods such as feature extraction or dimensionality reduction, it can reduce the total amount of data being transmitted to the cloud, thus leading to cost savings in terms of the subsequent data storage and processing. Figure 5 gives the layout of the local edge server, showing it to be mainly comprised of an edge manager, edge broker, user account and authorization, edge VM, edge operating system, and edge agent. The edge VM runs in an isolated process in parallel with other components to interact with sensors and actuators that are directly connected to the edge server.

The data flow from the plant site's edge server to the cloud is as follows:

1. Sensor data are ingested to and transmitted from the edge server. The hypervisor creates and oversees the VMs as they run.
2. The data are transformed and sent to a data bus, depending on the communication protocol from the protocol adapter.
3. Following subscription to the protocol adapter, data are sent from the data bus to the Event Hub and time series gateways.
4. Data pre-/post-processing is conducted by the custom application, which acts as a computing analytics unit on the edge server and publishes the data back to the data bus or the cloud.
5. The historian is used to archive the processed data.
6. Finally, the processed data are published to the respective gateways on the cloud.

In addition, the edge analytics services that are also available through the custom application can perform data cleaning and compression, as well as some feature extraction prior to the data being transmitted to the cloud.

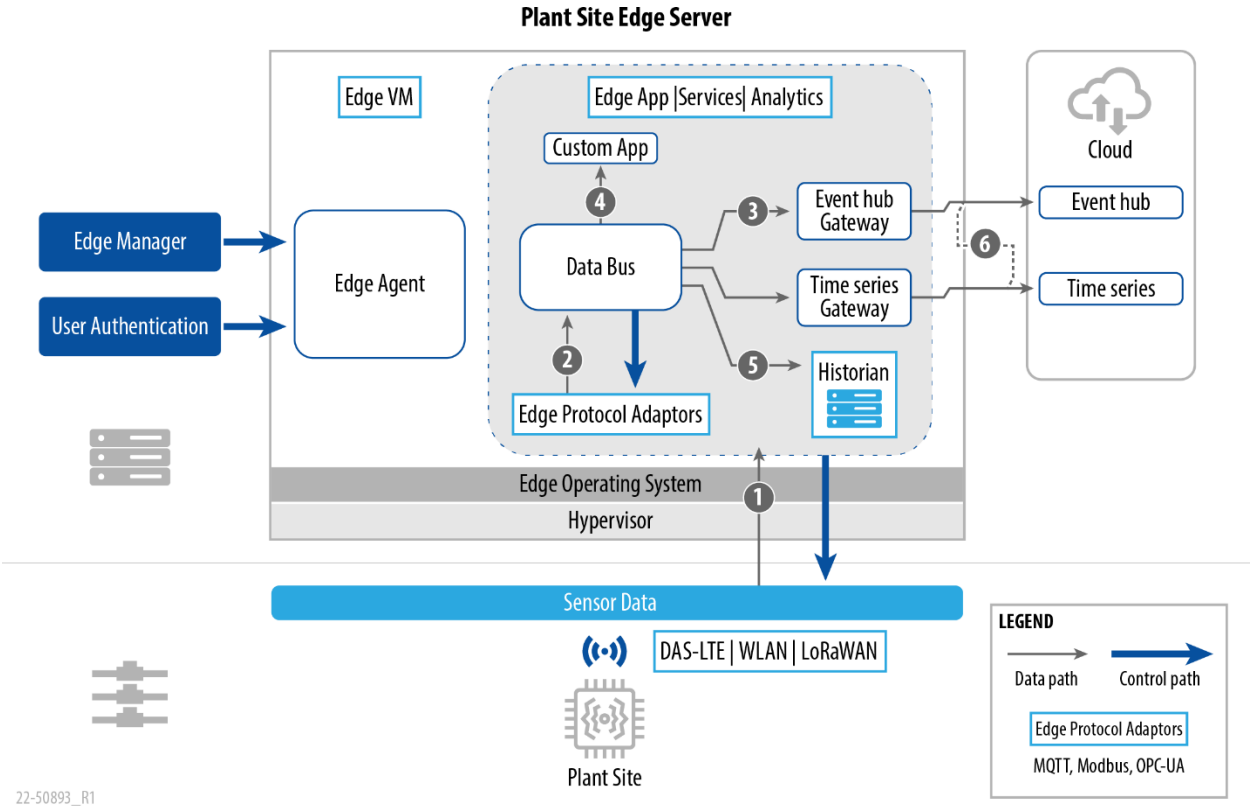

Figure 5. NPP site edge server interface.

2.4 Cloud Architecture and Storage

Architecture and storage pertain to the structure of the data pipeline once the data have been transmitted inside the cloud.

2.4.1 Data Ingest and Pipelines

This section presents the data ingestion and egestion pipelines required by the architectures proposed in the previous sections in terms of performing remote analytics and maintenance. The architecture of the data pipeline illustrated in Figure 6 takes into consideration the higher cost of streaming the data as opposed to processing them in batches. In this regard, the frequencies of the input data for near-real-time and batch processing were considered to mitigate data loss and errors in the data transmission process. The two main sources of data considered were vibration and NPP equipment process data, which pertain to flow, pressure, and time series information obtained from the Open Platform Communications (OPC) Unified Architecture as well as the supervisory control and data acquisition (SCADA) systems or historian server at the NPP site. An example of cloud deployment using data from a NPP site is given in [10]. For near-real-time analysis, the sampling frequency for the vibration data could be as high as 8,192 samples/sec, transferred on an hourly basis. The data ingestion pipeline in Figure 6 describes the various aspects of feasibly transmitting vibration and plant process data—namely, ingestion, transformation and analysis, storage, and visualization.

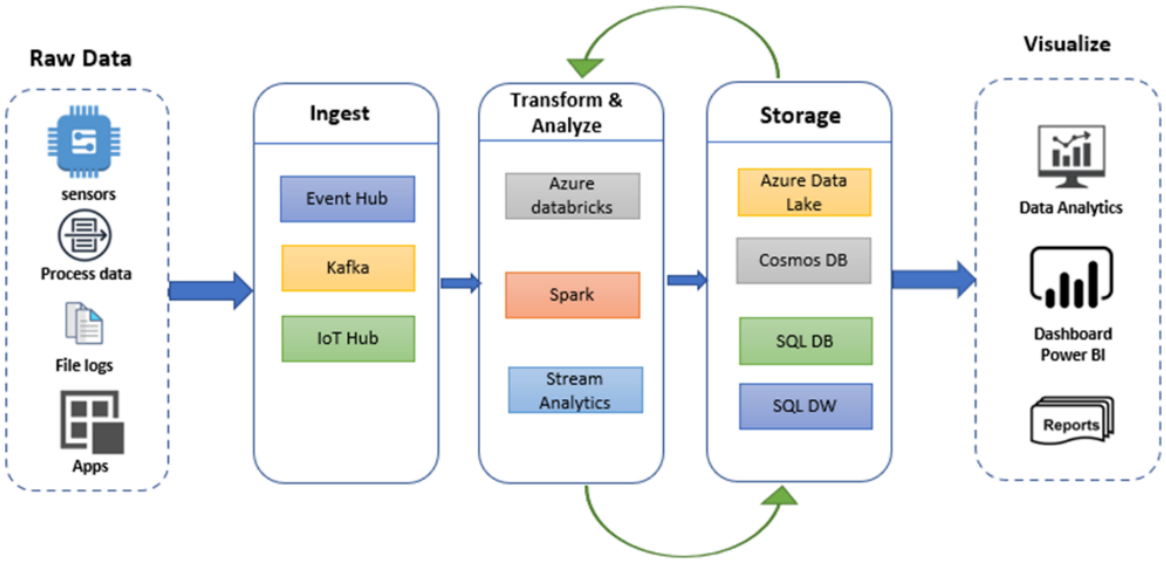

Figure 6. Data ingestion and egestion pipeline layout for the NPP cloud.

1. **Ingestion**: This, the first step in the layout, involves extracting raw data and/or features from the sensor devices or edge server, server logs and files, process data file uploads, and SCADA systems or data historians at various frequencies. Because data at this stage originate from various sources at variable speeds/frequencies and in different formats (i.e., structured, unstructured, or semi-structured), it is imperative that the data ingestion be completed effectively. The following data ingestion tools were employed for our specific data:
   - *Event Hub*: This is a low-latency seamless integration streaming platform for receiving and processing a large volume of data events per second. It also has built-in adapters for real-time and batch-event transformation.
   - *Apache Kafka*: This is a high-throughput distributed messaging system that features enhanced integrated portioning and fault tolerance for large-scale data streams.
   - *IoT Hub*: This cloud platform service, primarily managed in Microsoft Azure, acts as the central communication layer between the connected devices. It also supports monitoring to assist in workload scaling and the tracing of device connections.
2. **Transformation and Analysis**: This step tackles the question of whether to stream data or process them in batches. The collated data are converted, normalized, cleaned, merged, and sorted for further analytics. Recommended tools include:
   - *Apache Spark*: A parallel processing cloud application in Azure HDInsight, it supports in-memory huge data analytics. It can load and cache data repeatedly during in-memory computations. Spark scales up in performance and data optimization, providing interfaces for Python, R, Scala, and SQL. This tool is useful for applying large-scale data analytics to the vast historical data currently stored at each plant.
   - *Azure Databricks*: A cloud-based engineering data processing tool that is useful for transforming, processing, and exploring huge amounts of data through AI models. To increase performance, the Databricks runtime augments the capacities of Apache Spark workloads. In addition, cost minimization is achieved through auto-scaling and auto-termination modes for Spark clusters. This tool allows for integration of open-source libraries, thus enabling the latest AI and ML techniques to be used.

- *Stream Analytics*: A managed cloud application service designed for conducting large-scale data analytics on real-time data by employing SQL to generate meaningful insights with low latency. It supports real-time dashboarding, anomaly detection, remote monitoring, and maintenance—each of which help enable PdM.

3. **Storage**: This step refers to storing the data at a specific cloud location for convenient read/write access. Again, it unifies the various data sources by activating the standard data access protocol for various data formats. The following are some storage locations identified for real-time streaming:
   - *Azure Data Lake* (preferably the Azure Data Lake Gen 2 storage) is well suited for unlimited data storage at different ingestion speeds at a specific location for operational and exploratory analytics. Additionally, it permits the storage of both relational (i.e., data contained in one or more tables, with a predefined relationship) and non-relational data, in formats such as JSON, XML, Binary, Avro, Excel, and Parquet.
   - *Cosmos Database (DB)* provides fast, flexible storage for replicated database and operational logging applications. Changes applied to the data can be used to trigger additional actions or patterns by integrating streaming analytics, performing cache synchronization, and archiving data to the cold storage (i.e., rarely accessed data). Cosmos DB currently supports table storage, and muti-model operational data can be stored in either the transactional store format (rows) or the analytical store format (columns).
   - *SQL DB*, a SQL database storage type that supports storage of up to 100 terabytes (TB), is responsible for backup, scalability, and data restoration services, with added data protection. Row and column store formats are also supported by this option.
   - *SQL Data Warehouse (DW)*, a SQL data warehouse for relational database storage, integrates data from various sources. More importantly, it reduces costs by leveraging a scale-out for elastic computing that separates compute and storage processing independently.
4. **Visualization**: This step uses dashboards and Power BI to generate scientific and engineering insights into analyzed data for both real-time and near-real-time reporting. Additional analytical tools such as Python, R, and HDInsight can also be used to visualize data on the cloud platform.

2.4.2 NPP Cloud Data Processing Architectures

For streaming both NPP equipment process data and real-time sensor data inputs to the cloud, two data processing architectures would suffice for both real-time and near-real-time applications. The Lambda architecture is better suited to real-time streaming, whereas the Kappa architecture is better suited to near-real-time streaming so as to take care of any lags in the data ingestion process.

Lambda Data Flow Architecture: Figure 7 is a schematic representation of the Lambda data flow for cloud platforms that process large quantities of data by leveraging the batch and stream processing approaches. This architecture balances latency, throughput, and fault tolerance by merging batch and real-time stream processing data in order to provide comprehensive insights into online data. The batch layer (cold path) precomputes results by using a distributed system in archiving the voluminous amount of processed data [11]. On the other hand, the speed layer fills in the gap caused by the batch layer's lag, thus providing up-to-date views during real-time streaming. Outputs from the batch and speed layers are merged and queried by returning precomputed views from the processed data. The master data then stores the processed data within the stream path.

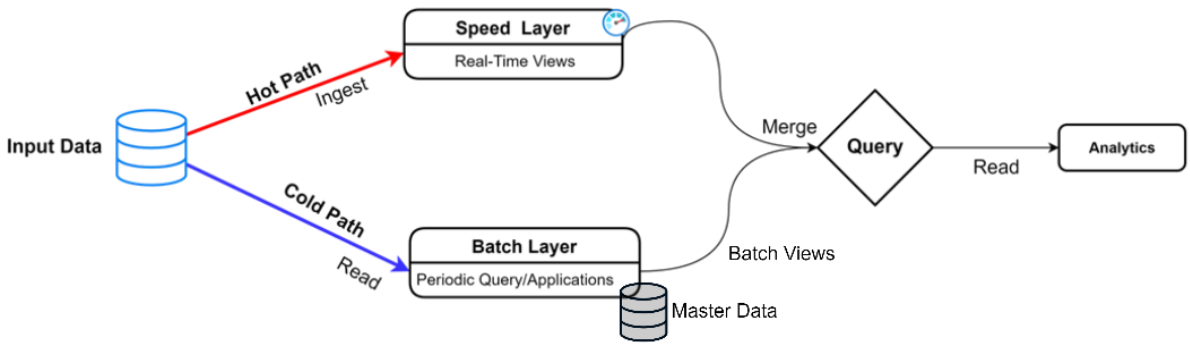

Figure 7. Lambda cloud data flow architecture.

Kappa Data Flow Architecture: The layout of this architecture, given in Figure 8, represents a simplification of the Lambda architecture by eliminating the cold path in the near-real-time streaming mode [11]. It seeks to reduce the Lambda architecture's complexity in using two separate layers in tandem for the same functional requirement. The architecture centers on a highly scalable queue log that ingests all data. Computations on the streaming data occur in the speed layer. Mirrored events are stored in the long-term storage to enable re-computation on the archived data when necessary. Near-real-time visualizations of the processed data are performed. The long-term storage capabilities afford scalable and fault-tolerant storage that recomputes historical data in unified log events.

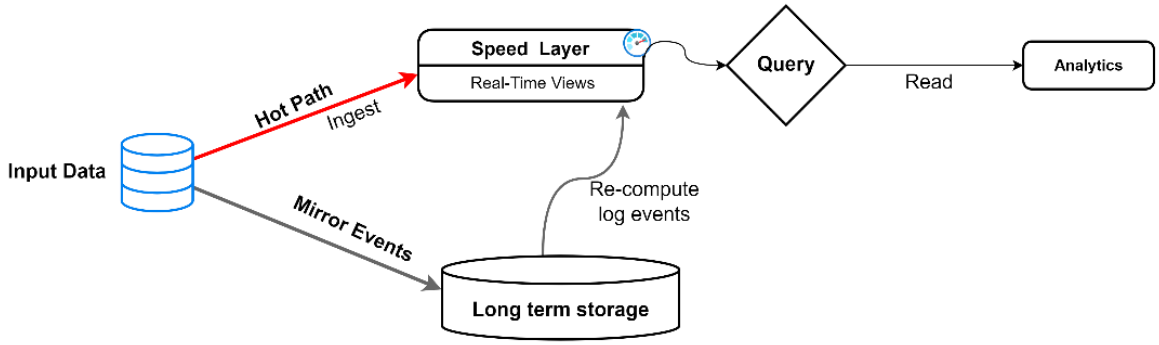

Figure 8. Kappa cloud data flow architecture.

2.4.3 Data Storage

The cloud data storage architecture would afford a single location for managing and scaling complexities and costs, with enforced data security. Cloud data integration storage of NPP historian data would require the aggregation of data from various plant locations and data sources (e.g., routine operational records, sensor equipment process instrumentation and controls, and data management systems [12]). Most importantly, leveraging the cloud's data storage capabilities would enable the aggregation of databases from several NPP sites. Centralized data storage would enhance data transformation and cleansing with improved real-time streaming analytics and visualization.

And though the cloud data storage architecture would serve as a single-point data access, cloud data storage can be replicated multiple times across global locations or regions to ensure redundancy [13]. Typically, there are four ways to ensure that data are stored redundantly: locally redundant storage, zone-redundant storage, geo-redundant storage, and read-access geo-redundant storage. These replications provide scalable features and load balancing options to meet peak demands, thereby ensuring a high degree of automated data availability. Stored cloud data can be accessed from any location worldwide over HTTPS/HTTP, while the supporting languages include NET, Java, Python, Ruby, Go, Node.js., and PHP.

For Azure storage accounts, the available storage options are Azure blobs, files, queues, tables, and disks. Azure storage accounts act as containers that merge a set of Azure storage services. The various types of storage accounts include standard general-purpose v2, premium block blobs, premium file shares, and premium page blobs [14].

From a NPP perspective, these various types of cloud storage can be categorized into three distinct groups: short-term, long-term, and optimized.

Short-term storage: This refers to short-duration, expiry-based storage such as Azure IoT Hub, Event Hubs, and queues. Backlog or exchange messages between NPP site components and the cloud storage are accomplished through Azure Queue Storage services via HTTP or HTTPS/AMQP/MQTTP.

Long-term storage: This type of storage permits the storing of large volumes of data for long durations in Azure blobs (i.e., data lakes), files, and disks. Blob storage can be used to store unstructured data such as images, media files, documents, logs, or raw data, along with the corresponding metadata information. Blobs can store up to 500 TB of unstructured data via a directory-like method known as containers [15]. And because data stored on a cloud can increase exponentially, Azure provides three access storage tiers (i.e., hot, cool, and archive) to provide cost-effective long-term storage, based on how the data are used [16]. Files services such as representational state transfer (REST) application programming interface (API) and VM file shares are used for accessing data storage from the NPP site and the cloud, respectively. Disk storage is accessed via virtual hard disks (e.g., ultra-disks, solid-state drives, and standard hard disk drives). The managed Azure disks are also stored as page blobs, which are random input/output (I/O) storage objects.

Optimized storage: In general, this supports flexible, low-latency querying scenarios and serves as a repository for processed real-time data. Azure Table storage, an optimized storage method, is often preferred for tabular, structured, relational schema-less data accessed via REST API. The Azure Cosmos DB and NoSQL data stores include table storage options. Other storage components in this category include Azure SQL Database, Document DB, SQL Data Warehouse, Redis Cache, and HDInsight. The proposed NPP cloud storage charge model for the above storage types is shown in Figure 9 below.

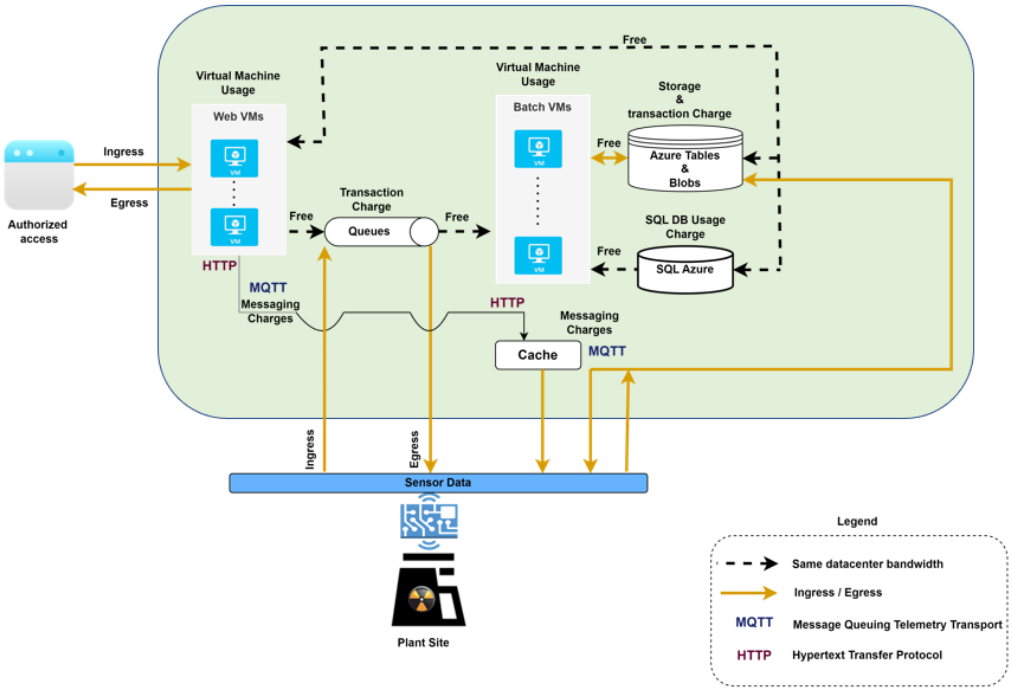

Figure 9. NPP cloud storage ingress and egress charge model.

Figure 10 describes the scalable cloud storage alternatives for the NPP virtual data center. This storage layout of collaborative models would promote and facilitate the sharing of remote monitoring alerts and reports among different NPP site locations.

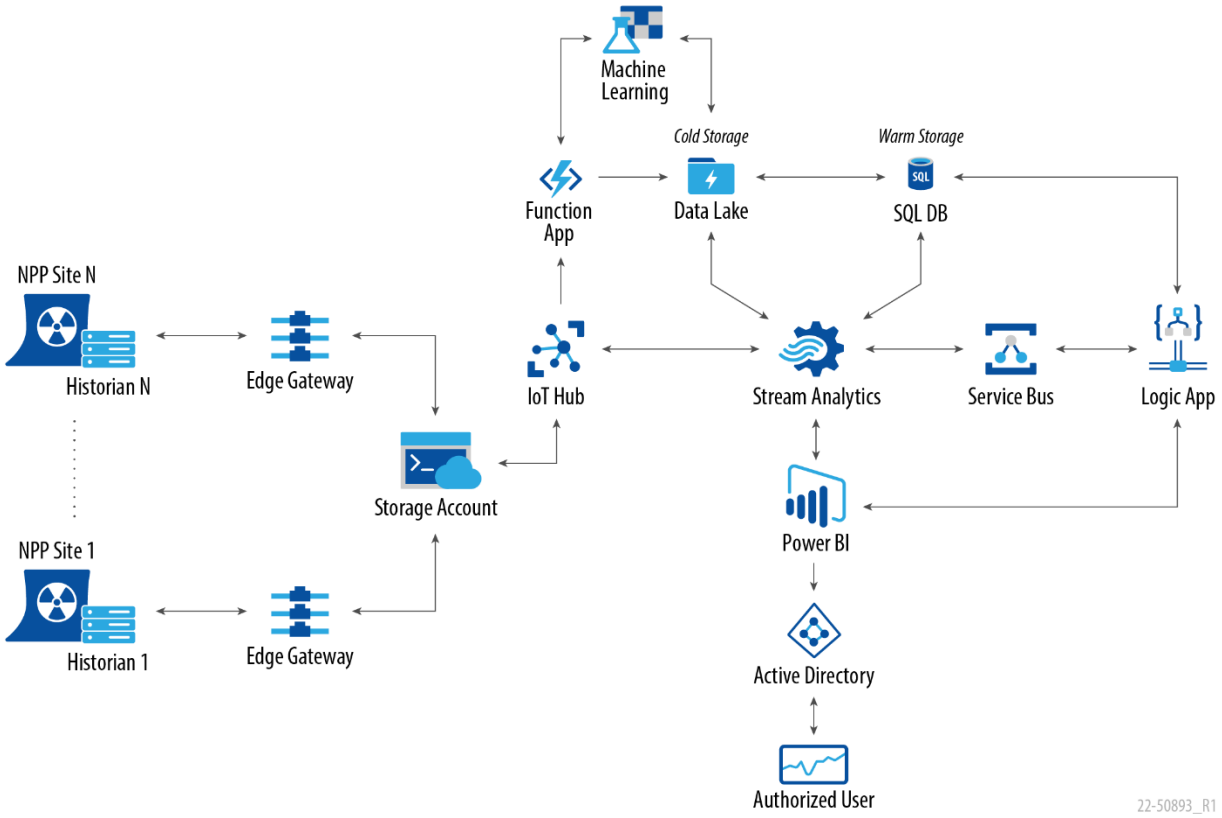

Figure 10. NPP cloud storage layout.

Azure storage accounts can be accessed through the portal, command line interface, or power shell. As illustrated in Figure 10, sensor data from individual NPP site historians are sent to an edge device. Technical staff can then monitor the archived data on the edge device by using a local dashboard. Data are then transmitted to the cloud IoT Hub gateway for cold storage in a data lake. Time series data are then analyzed in real-time via streaming analytics, with access from the hot data storage. The bi-directional message communication and data flow paths illustrate how recommendations can be sent from the cloud storage to the NPP site, or vice versa, to support prompt actions for ensuring that operations are optimized. Final processed data can be stored either on the cloud, sent to the NPP site historian for archival, or visualized using Power BI. Cloud data can also be integrated with web applications for authorized users.

2.4.4 Other Features

Virtual Machines: A virtual IaaS function provided by Microsoft Azure, these VMs support both Windows and Linux platforms. In the NPP cloud IaaS case, total control is gained over the VM configuration in terms of software installation and maintenance. Accessing a VM through the internet requires that an associated network interface or load balancer be configured with a public IP address. Accelerated VM networks enable acceleration of low latency and high throughput. Backup and recovery services can be established using Azure Site Recovery. The size of a VM determines the resource allocation (e.g., processing, memory, and storage capacity), as well as the cost involved [17]. VMs can also be task-specific (e.g., general purpose, compute optimized, memory optimized, storage optimized, high performance, or graphics processing unit (GPU) optimized). Table 1 gives information on Microsoft Azure cloud services VMs and other cloud resource charges relevant to NPP cloud integration.

Table 1. Microsoft Azure cloud resources and charges (adapted from [17]).

Service	Minimum Monthly Charges	Details
Virtual Machines† (AI/ML)	A-series: $31.39 /month F-series: $36.28/month D-series: $53.29/month N-series: $170.09/month	Workloads and memory capacities determine the cost A series: (2.3–3.5 Ghz), 2–64 GB (general purpose) F series: (2.5–4.0 Ghz), 2–1000 GB (compute optimized) D series: (2.3–3.5 Ghz), 8–384 GB (memory optimized) N series: (2.5–3.5 Ghz), 14–6,500 GB (GPU optimized)
Azure Blobs and Tables†	Premium: $0.15/GB Hot Tier: $0.018/GB Cold Tier: $0.01/GB Archive: $0.00099/GB	Total cost for blob storage depends on: • Volume of data stored per month • Quantity and types of operations performed, along with any data transfer costs • The data redundancy option selected (e.g., LRS, ZRS, GRS, and RA-GRS).
Private Network Bandwidth†	ExpressRoute Circuits: $55/month ExpressRoute Gateways: $136.8/month	• ExpressRoute Circuits bandwidth: 50 Mbps – 10-Gbps • *Indicated prices are for ExpressRoute Circuits Zone 1. • *The U.S. GOV Zone 1 costs are a bit higher. • ExpressRoute Gateway bandwidth: 1 Gbps – 10 Gbps
Transactions†	$0.00036/10K transactions	Create, Read, Update, and Delete into Azure Queues, Blobs, and Tables is considered a transaction.
Azure SQL Database†	$368.19/month	Locally redundant database: standard series (Gen 5) 10–396 GB memory
Azure Service Bus†	$0.05/1000K message operations	A message operation may be a service bus message, access control token request, or service management API call.
Ingress GB†	$0.00/GB (Free in the U.S.)	Data transferal into the data center is free in the U.S. (except for backup recovery services).
Egress GB†	$0.01/GB	Only the data transferred out of the data center are billed.

† From Microsoft 2022 [17] (information current as of the time of this writing).

Power BI: Together with Azure analytics and other services such as Event Hub, Stream Analytics, Azure ML, Azure HDInsight, and Azure storage, this provides visualized insights into complex processed data. Both NPP site historian data and cloud-based data can be integrated to build meaningful reports using the built-in connectivity of Power BI. Data can also be transformed or shaped via embedded dashboards to enable interactive visualizations in web applications, other applications, and portals. The Power BI standalone desktop application can be used to develop data modeling and visualization tools on the NPP edge server prior to publishing the solution on the portal.

2.5 Processing and Predictive Maintenance

After the sensor or feature data have been uploaded to the cloud, processing and PdM enable component health predictions, as well as estimations of their remaining useful life (RUL). Sensor data processing focuses on computational requirements and the optimization thereof, whereas PdM enables the use of AI and ML to predict component failures and anomalies. This section begins by describing the current PdM approach involving the AVEVA Connect cloud architecture, which the collaborating NPP is familiar with. It concludes by describing a similar technique for accomplishing PdM by using the Microsoft Azure Cloud platform.

Industry 4.0 is primarily focused on enhancing intelligent processes by bringing enhanced digital technologies such as IoT, big data, and cloud computing to enable continuous real-time monitoring of sensor data flows from manufacturing to organizational power systems [18]. The AVEVA PdM cloud platform provides solutions that aid organizations in autonomously responding to failure alerts and diagnoses, allowing them to efficiently cope with the increasing volume of historical data. AVEVA's Insight Advanced Analytics module is a single-tenant software offered as a SaaS on the AVEVA cloud platform (AVEVA Connect) for monitoring the real-time heath of critical assets [19]. Because this software does not specify which equipment type to monitor, it is thus customizable and can be adapted to monitor the performance of a variety of industrial processes. Furthermore, AVEVA Predictive Analytics integrates with a variety of data historians and monitoring systems, is highly scalable, and can be deployed either on-premises or in the cloud.

Again, the updated version of the AVEVA Predictive Analytics application utilizes advanced AI and ML techniques to provide near-real-time asset information for notifying operators whenever a piece of equipment malfunctions. Organizations can use this application to diagnose equipment issues prior to failure, thus helping meet global market demands by avoiding unplanned downtimes, increasing asset utilization, extending equipment lifetimes, and improving safety.

The salient aspects of the software can be summarized as follows:

- *Sensor fault detection*: To ascertain and be more confident of results, AVEVA'S new configurable sensor analysis, which is based on an advanced pattern recognition technique, identifies bad sensor signals so that they can be excluded from model calculations, and forecasts assets' RULs to better inform decisions on whether the assets should be replaced or overhauled. When adjusting operating conditions, the time/distance to a certain failure threshold (e.g., bearing temperature, filter differential pressure, and vibrational data) is evaluated.
- *Asset framework adapter for the AVEVA PI System™*: The AVEVA PI System Asset Framework can identify asset tags and thus better leverage enterprise deployment [20]. The new adapter reduces model configuration errors by ensuring consistency across multiple assets of the same type.

AVEVA Predictive Analytics Features

AVEVA Predictive Analytics can identify specific sensors that deviate from normal conditions, as seen in Figure 11. By accurately forecasting sensor data, any discrepancy between the actual and the predicted values can serve as a dynamic alarm. Traditional alarm thresholds require the signal to cross a predefined static value, whereas dynamic thresholds can provide earlier indication of equipment

degradation or a change in the process. Dynamic thresholds can be used in conjunction with traditional static thresholds. Less subject matter expertise and domain knowledge is required for model development, as most model information templates can be easily accessed from the pre-built model builder library.

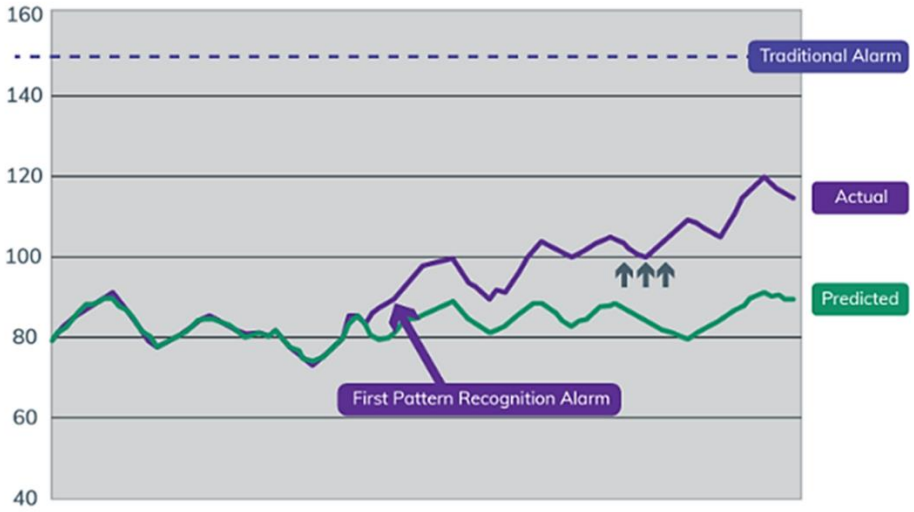

Figure 11. Latest version of the AVEVA Predictive Analytics early detection model output [20].

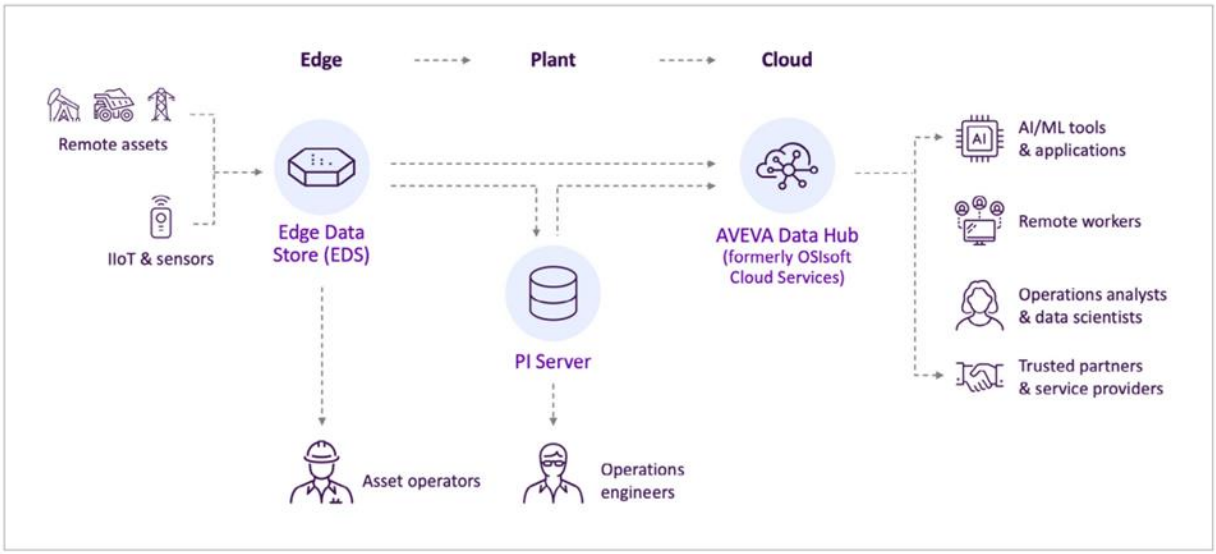

Figure 12. Layout of the AVEVA Predictive Analytics data hub [21].

Figure 12 shows the edge-to-cloud data processing layout for AVEVA Predictive Analytics. The model results can be integrated with other industrial systems via web services, servers, or desktop client with an available REST API [22]. The various aspects of the AVEVA Predictive Analytics module can be summarized as follows:

Data Ingestion and Processing: The AVEVA data hub found on AVEVA Connect is the cloud platform used for aggregating, storing, enriching, accessing, and analyzing real-time operations data from historians, PI servers, edge devices, etc. Normalization, filtering, and contextualization of sensor data are conducted on the human-machine interface/SCADA edge device. A brief list of monitored equipment would include heat exchangers, condensers, pumps, turbines, compressors, motors, and electrical generators.

Data Analytics via Multiple Algorithms: This includes data analysis techniques (e.g., advanced pattern recognition, ML technology, and artificial neural network technology) for monitoring and predicting the operational conditions of thousands of pieces of remote equipment at different sites. Supervised ML is used to understand the individual operating history of a given piece of equipment, and a series of normal operational profiles is developed for each specific device. Batch data processing is also supported in addition to near-real-time mode.

Early Detection: Known operational profiles are compared against real-time operational data in order to detect subtle system behavioral changes that serve as early warning signs of possible equipment failures, in advance of existing operational alarms. This equipment-agnostic method affords users the opportunity to quickly transform raw data into actionable insights so as to prevent equipment failures and make informed decisions to enhance operations.

BI Gateway Visualization: The AVEVA Insight BI Gateway automates the extraction, transformation, and storage of operational key performance indicators, metrics, and information across various data structures. Core capabilities include visualization of historical data tags, using dashboards sharing of model outputs with no-touch analytics on incoming data.

Notifications and Alerts: Thresholds are used to communicate alerts whenever discrepancies between actual and predicted values exceed the specified limits. These alerts can be managed in various ways, such as category level, criticality, frequency, and duration. Alert events are linked to a graphical trend that illustrates event data, threshold limits, and the number of alarm occurrences. Through email, authenticated users can receive real-time notifications of equipment status alerts.

NPP Cloud Predictive Maintenance Framework

PdM includes failure prediction, failure detection, failure diagnosis, failure type classification, and recommendations for maintenance actions or equipment failure mitigation [23]. Microsoft Azure IoT Suite provides resources that assist in deploying PdM solutions. Azure ML services are focused on techniques for predicting when equipment will fail, then facilitating maintenance planning in advance of that event. These services also include a collection of preconfigured ML modules, as well as user-defined scripts in programming languages such as Python, R-studio, Scala, and C# for enabling an end-to-end solution that encompasses everything from data processing to the final stage of deployment.

The high-level architecture created for efficient deployment of NPP PdM cloud components is detailed in Figure 13 below. As seen in the figure, the relevant steps involved are data recording, data ingestion, processing and storage, preparation and training, modeling and predictions, and visualization.

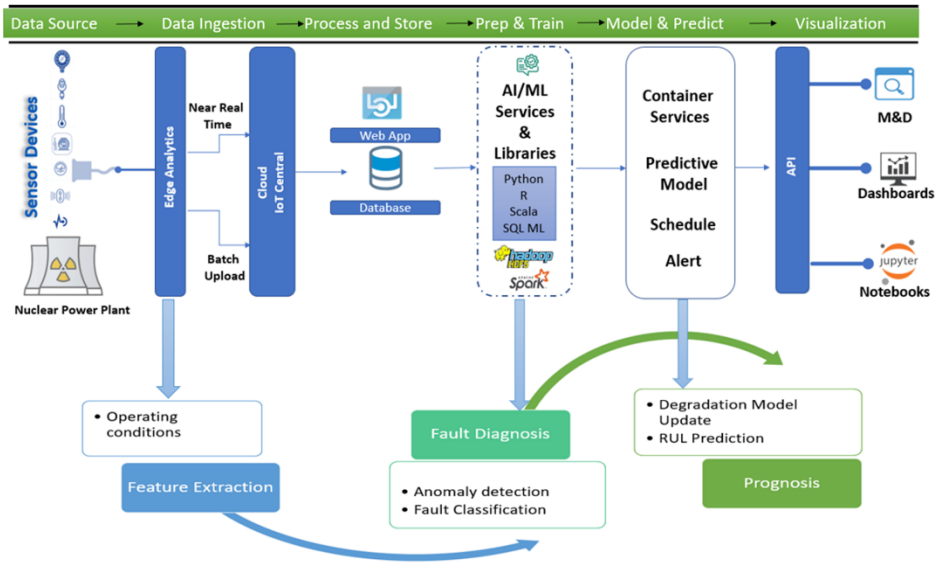

Figure 13. NPP cloud predictive maintenance layout.

Data Source and Data Ingestion: The initial stage of the PdM process is the data source, as well as the communication protocols needed to transmit the data to the cloud for processing and storage. IoT sensor devices translate physical actions from machines into digital signals such as vibrations, voltages, and temperatures. Data can also be streamed from other sources such as NPP process data, maintenance logs, and programmable logic controllers. Additional connectors can be built to incorporate other data schemas. A centralized data aggregation approach on the NPP edge server can be employed to merge data from over 600 sensors with numerous historical maintenance records. Communication protocols such as the OPC Unified Architecture, Modbus, and Transmission Control Protocol are supported. Azure IoT Central is a cloud gateway software application that performs data preprocessing, exports telemetry time-series data and device determines equipment efficiency trends, and sends the data to other services for storage processing and analysis, based on user-defined functions.

Process and Store: The link between the edge server and the cloud IoT Central reflects the idea behind utilizing edge computing to perform certain feature extraction tasks to provide near-real-time data while leveraging the cloud's batch and database storage for low-latency computations. At this stage, information such as labeling operating conditions, equipment-level features, health indicators, and sensor equipment RULs are also performed on the training data.

Prep & Train: The cleaned and processed sensor data in the cold storage were prepared by labeling and classifying fault signals using the Azure Databricks component, which supports the training of large datasets at scale. Model training is performed by the Azure ML services and deployed to container instances or an Azure Kubernetes web service cluster. User-defined scripts such as Python, R, Scala, and Java can also be employed in building the ML training model by using classification algorithms. The trained model is further integrated with historical data for retraining on new diagnostic and prognostic models. A large portion of the data can be used for training, whereas only a small fraction is needed to validate the performance of the trained model. Apache Spark running on a Hadoop Kubernetes cluster is able to leverage Spark SQL and DataFrame API for data cleaning, feature engineering, and development of the ML training model [24].

Model & Predict: The ground truth data obtained from the degraded sensor's historical data, along with the lifetime of each piece of equipment, are employed to develop diagnostic and prognostic models. The sensor measurements are used as inputs to the predictive model to calculate the RUL. Several neural-network-based anomaly detection methods can be adopted during model prediction. These include the artificial neural network, recurrent neural network, autoencoder, and long short-term memory (LSTM)

neural network. Regression-based methods can also be deployed to an Azure web application container and then scored to provide real-time predictions regarding the input sensor data. A multi-class model with well-tuned hyperparameters can predict the time range for—and likelihood of—failure. In the Azure ML Studio library, Apache Spark is used for model predictions, and it also supports regression analysis, feature engineering, and multi-/binary classification. Available classification algorithms in the Apache Spark ML library include random forest, logistic regression, gradient-boosted tree classifier, and naïve Bayes. Spark also offers regression algorithms such as generalized linear regression, gradient-boosted tree regression, and random forest regression [25].

Visualization: Once the most important parameters of a model have been identified, the model would have to be deployed as a web service application, using a REST API endpoint. In the case of the NPP cloud framework, the PdM procedure often involves collecting and processing real-time or batched sensor data by using the Azure IoT/Event Hubs. The processed data would then be transmitted to a predictive model web service, where the results would then be displayed on a Power BI dashboard for alerting the M&D staff of possible issues. Ingested data stored on the cloud can be merged with external data from the NPP site historians so that they can be fed back for re-modeling and training purposes. Models deployed on the NPP edge server can expedite anomaly detection. Again, understanding the time series data patterns is essential for identifying anomalies through advanced dynamic alarming techniques rather than through static thresholds. In Microsoft Azure, the real-time data stream enables users to monitor and analyze events as they occur. The control aspect of the real-time visualization is used to communicate with sensor devices and thereby alter the digital twin properties by invoking direct methods in synchrony with the back-end Azure SDK. During real-time visualization, line graphs are mostly used to track the status of the monitored device [26]. Historical events can also be visualized by using line graphs, stacked area graphs, and horizon graphs when analyzing and monitoring the long-term trends of archived sensor data.

Development of a real-time PdM method requires a complete end-to-end approach that includes sensor devices, communication protocols, data analytics tools, data visualizations, and collaborative insights. A flow chart of the required infrastructure for the NPP PdM framework is given in Figure 14.

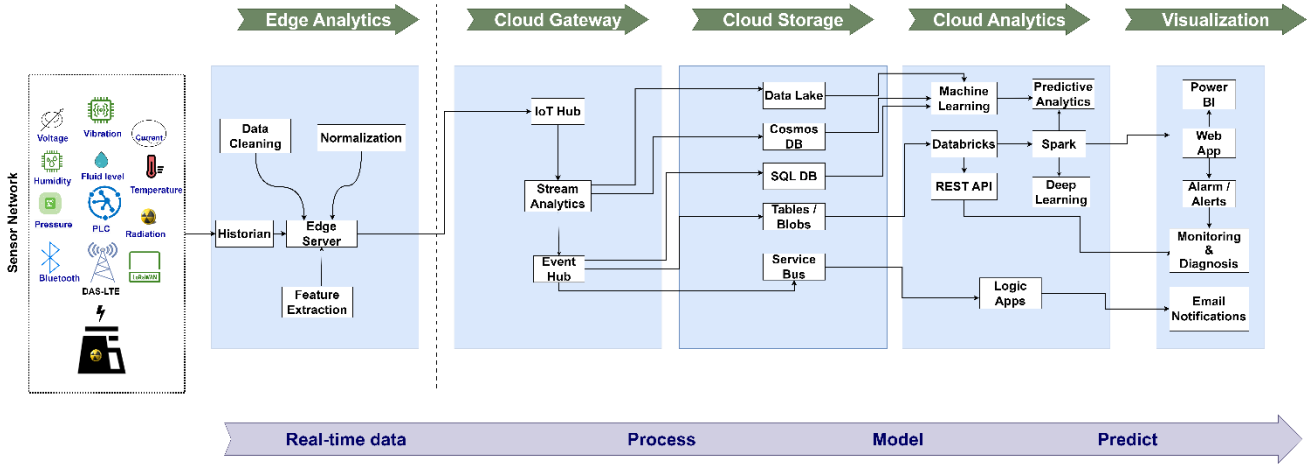

Figure 14. Flow chart of the NPP cloud PdM infrastructure.

NPP Predictive Maintenance Technical Implementation

This section expounds on the technical details and considerations needed to efficiently deploy the NPP cloud PdM infrastructure (summarized in Figure 15) by merging sensor data from various NPP sites with the available ML models in Microsoft Azure.

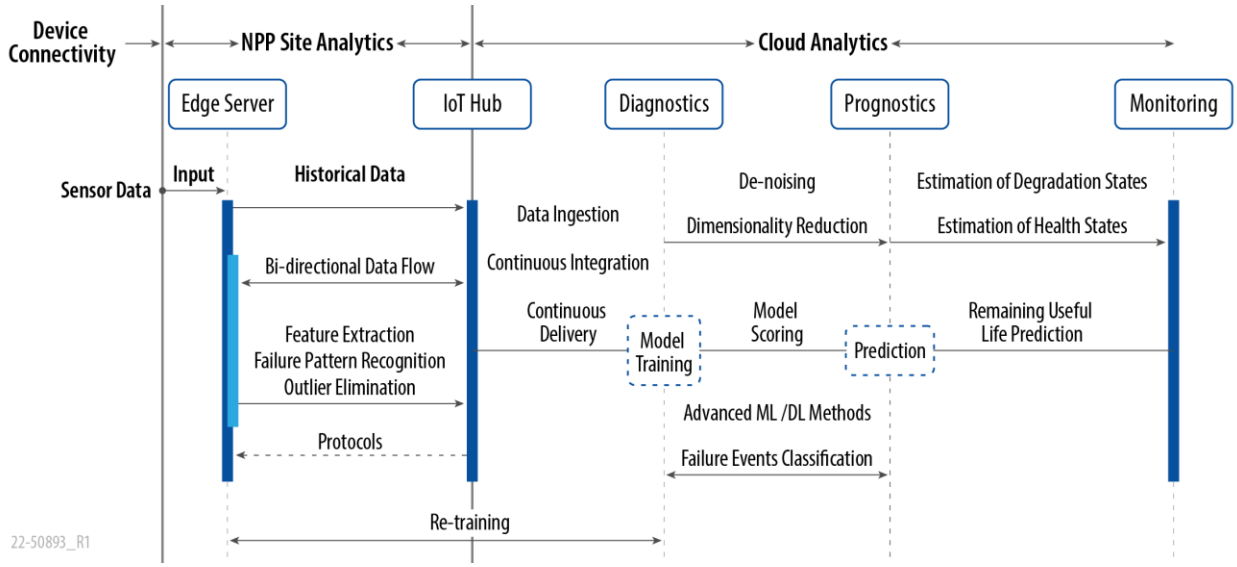

Figure 15. Near-real-time data analytics framework for NPP predictive maintenance.

Device Connectivity

The NPP industry has identified and assessed the need to mitigate O&M costs, and the proposed real-time PdM approach is a step in the right direction toward connecting sensor devices and identifying patterns that can aid in predicting component failures. Insights gained from the analyzed patterns could be used to address issues such as unscheduled downtimes and the optimization of equipment usage. At NPP sites, challenges such as the unavailability of sensors to generate data and the lack of computational resources to aggregate historical data can be alleviated by advanced cloud-based AI PdM methods. Again, PdM augments previous strategies such as reactive and preventive maintenance procedures by maximizing both uptime durations and device lifetimes [27].

PdM is mostly data-driven, and for model training and prediction, would thus require operational data stemming from both normal and abnormal operations. Normally, the data structure (i.e., relational, structured, or unstructured) depends on sensor equipment scenarios that occur before, during, or after failures identified via network sensors, maintenance logs, machine specifications, etc. (see Figure 6 and Figure 14). Records from various NPP sites can be aggregated on a historian VM/server prior to transmission to the edge server.

As the NPP industry seeks to monitor both time domain and frequency signals from various sensor equipment at different time intervals, it is imperative to have a reliable interface that supports the landscape of heterogeneous devices that send signals from the plant site to the cloud. In this regard, Microsoft Azure has a signal processing technology pipeline that identifies and captures signals, metadata, and other data formats from connected sensor devices by using an IoT gateway and application programming interface (API). Azure's signal processing pipeline has built-in traceability, versioning, and rollback functionality to ensure that the data from all the connected sensors follow a standard format, with specific device tag IDs to provide origin, location, and contextual information. With the connected signal processing pipeline, the NPP site can track the status of all connected sensors, conveniently transition, and integrate local PostgreSQL datacenters to the cloud-equivalent SQL databases then connect, process, and monitor equipment health to enhance productivity.

Figure 16 illustrates a modified signal processing pipeline from Microsoft Azure for NPP sensor connectivity. This signal pipeline configuration tracks signals from the NPP-connected sensor network by using the ASP.NET web API on the Azure Kubernetes service. Step 1 is the data transmission from the NPP site to either the Azure IoT edge server or a private NPP edge device by means of a KEPServer configuration with built-in libraries (i.e., single-source software for connecting devices, managing data, and controlling the automation of the data flow). In Step 2, the OPC Publisher and OPC Twin edge module is a Windows/Linux run-time application that conveys telemetry data to the cloud gateway in batch packets to ease the network load. The configuration controller communicates with OPC Publisher via transfer protocols such as HTTPS, AMQP, and MQTT in order to apply the requests in Step 3. Step 4 acts as a hybrid integration messaging service that sends sensor device information from the cloud gateway to other cloud resources and applications. The pipeline configuration in Step 5 is an API that creates, reads, updates, and modifies redundant operations. Signal sampling rates and intervals are also defined. Step 6 uses storage and streaming analytics services such as Event Hubs, Data Lake, and Data Explorer to enrich the ingested time series data. In Step 7, the asset registry performs verification of the signal metadata and operations on associated server machines. In Step 8, the pipeline publisher sends configuration file information to the cloud gateway to request updated pipeline versions from the edge configuration controller module.

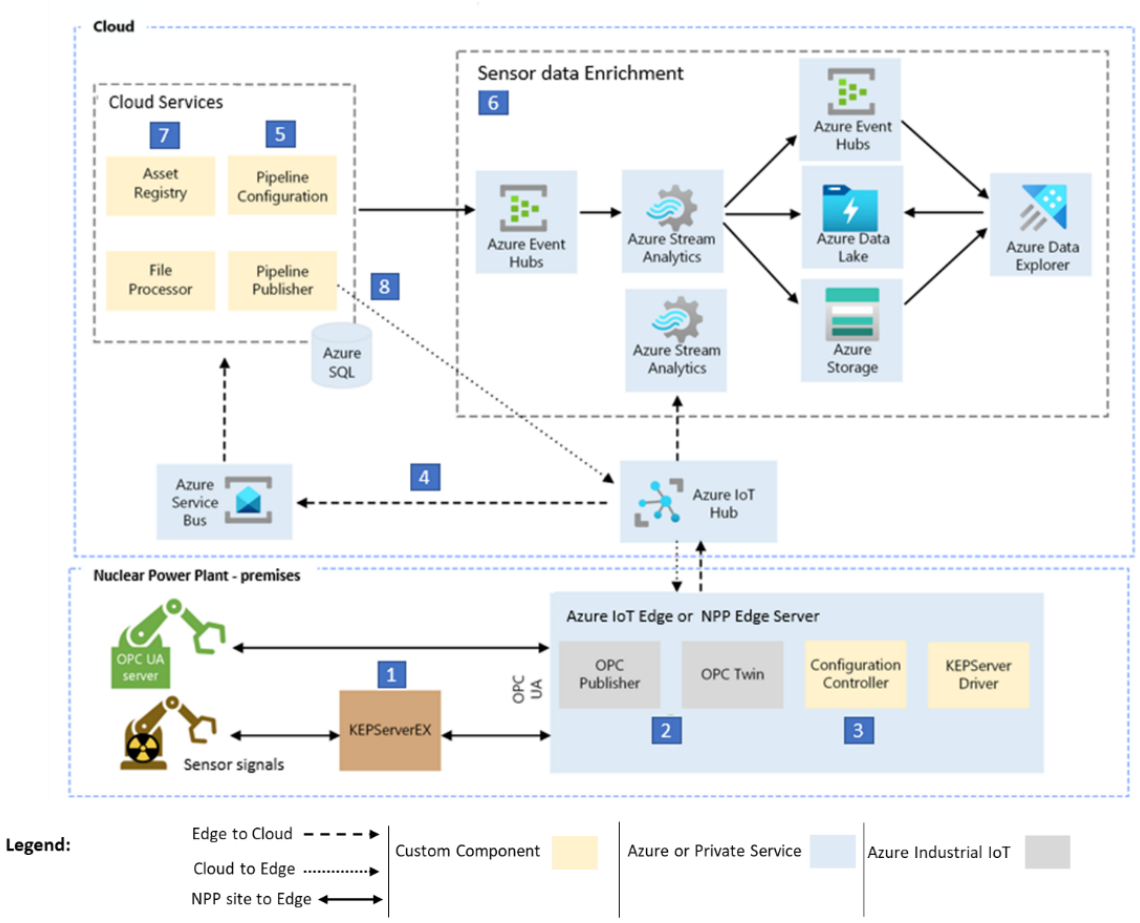

Figure 16. Proposed modified cloud signal processing pipeline for NPP sensors [28].

NPP Site Analytics for PdM

A possible challenge related to real-time monitoring of sensor data from NPP sites pertains to the transmission of large volumes of data to the cloud for modeling and prediction. However, an efficient cloud data management system would suffice to mitigate such challenges. Big data can be defined based

on characteristics such as volume (quantity of available data to be stored; data type, format, and source), velocity (i.e., the data speed and the sampling rate frequency), and veracity (i.e., quality of the streamed data for fostering accurate evaluations). Microsoft Azure's cloud platform offers a ML visual studio for continuous integration and delivery of new training models based on recently ingested real-time data.

Real-time PdM is a dynamic problem that requires associated ML models to be continuously refreshed and retrained to resolve changes in failure predictions. Installed sensor devices could experience several failures before the end of the equipment lifecycle. After the first failure, historical maintenance data can be collected, while the failure features can be captured once the faulty sensors are repaired. Real-time telemetry data such as velocity and vibration frequency can reveal health condition changes in sensor devices.

Edge-based analytics at NPP sites would be capable of performing short-term analytics on the aggregated data from various sites over short periods of time such as statistical and gradient feature extraction, failure pattern recognition, and outlier detection. This would ease high bandwidth demands and improve network latency throughout the entire modeling process. Data with long-term patterns would be transmitted to the cloud for advanced pattern recognition and deep learning (DL) analytics. VMs with run-time Python applications could be employed for the sensor data preprocessing on the edge server. To enhance the supervised feature selection and failure pattern recognition algorithms, a time windowing technique could be employed to aggregate the data within a specified window. A feature extraction ensemble algorithm using the scikit-learn recursive feature elimination, cross-validated (RFECV) module is recommended. This ensemble algorithm combines decision trees, random forests, support vector machines (SVMs), and a linear regression module to rank all features in accordance with their respective RFECV, with reduced bias [29]. Unsupervised feature extraction algorithms such as spectral embedded analysis and the Laplace score are also available using the scikit-learn feature utility and should be considered.

Cloud Analytics for PdM

PdM modeling on the cloud enhances the continuous monitoring of real-time NPP sensor equipment conditions in order to predict upcoming failures, thus enhancing proactive maintenance activities. Anomaly detection can trigger automatically, with alarm notifications being sent to authorized M&D staff to alert them as to the required action to be taken. In the diagnostics modeling and training stage, two categories of faults can be classified as point and pattern anomalies respectively, since the time series sensor data may contain considerable noise. Several AI models can be adopted for diagnostics, including the convolutional neural network, recurrent neural network, autoencoder, LSTM, and deep belief network. The prognostics models are used to predict the RUL, which is the length of time a machine can operate prior to requiring repair or replacement. NPP technical operators can schedule maintenance and optimize operating efficiencies to avoid unplanned downtimes. Some ML models for RUL prediction are readily available in Azure's ML Studio library. These include SVMs, decision trees, reinforcement learning, and gated recurrent unit [30]. Algorithms from the open-source scikit-learn library can be used to determine the most suitable models for anomaly detection and fault classification. Microsoft Azure embraces user-defined functions using open-source tools such as Python and R. Available improved infrastructure components such as large memory resource machines, central processing units, field-programmable gate arrays, and GPUs can be utilized to scale and provide improved capabilities for a robust AI predictive model. After the model is trained and the hyperparameters are tuned in real-time using hot tier storage on the Azure cloud platform, a web host can be utilized to deploy it for streaming analytics. Monitoring, reporting, and alerts can be realized through the Power BI dashboard and notification hub. The cloud-edge data integration proposed in this report can be used for model retraining, as processed models can be fed back to historical databases on the edge server, enabling failure detection events to also be monitored on the NPP edge server.

Proposed Scalable PdM DL/ML Algorithms

The main goal of the NPP PdM model is to precisely estimate sensor equipment RULs to provide ample time for maintenance to be planned accordingly. However, this can be cumbersome, since sensor degradation history data may be unavailable, and the available data may contain different failure modes. To efficiently model this degradation, advanced DL and ML models can be used to interpolate and map the relationships between training features.

Centralized cloud computing is an ideal data integration strategy for the NPP fleet, but the data privacy, commercial, and legal concerns that surround data sharing make data integration challenging for several NPP sites. In view of this, advanced federated DL and ML methods have been proposed for scaling the AI algorithms on an aggregated edge server with training data privately distributed from individual NPP servers at various locations.

In 2017, federated learning was first introduced by Google [31] as a variant of distributed learning. In federated learning, the cloud server acts as an aggregation point for the DL and ML models by coordinating multiple edge device analytics, in parallel, to perform a single learning task. One major difference between federated and distributed learning is that federated learning combines edge historians and analytics to perform a task, whereas distributed learning has edge nodes that lack access to historian data but only perform analytics as a processing unit. The main purpose of developing federated learning was to create a framework of private ML models by building training models locally on edge devices. Other advantages include the ability to keep raw data on the edge devices and to centralize locally computed model updates in order to reduce latency. Federated learning is thus recommended for future research, as it can improve the scalability of the NPP sensor fleet's PdM model predictions by minimizing training errors, reducing bandwidths, and capitalizing on data privacy on the cloud platform.

Federated aggregation is currently the most efficient global update method used in federated learning [32]. Through this approach, the NPP site edge devices transfer a step of local gradient updates to the cloud sever. The different data partition categories of federated learning are horizontal, vertical, and federated transfer learning [33]. In horizontal federated learning, the features of the sensor datasets from each NPP edge device are identical but reflect different training samples. Training samples can also be shared in vertical federated learning with different features. Federated transfer learning is used when neither the features nor the training samples are shared across edge devices. The weighted average of the transmitted weights is used to compute new weights, which are then sent back to privately connected edge devices located at various NPP sites.

The federated learning process, which is made up of several communication steps, can occur once the global aggregated model on the cloud reaches the desired level of accuracy. A specific DL/ML model can be trained in a distributed manner among several edge devices, the goal being to minimize a loss function *f(w)* on the entire datasets from various plant sites. The edge device objective optimization function is mostly non-convex [31]. The federated aggregated weight method can be expressed as follows:

$$w_{i,t+1} = w_{i,t} - \alpha \sum_{i=1}^{k} \frac{D_i}{D} \nabla F_i\big(w(i,j)\big) \qquad (1)$$

where w_i is the weight parameter for a specific edge server i, with the number of connected sensors j from NPP sites, α is the learning rate, t is the time step, D is the number of data points used for training the aggregated global model on the cloud server.

D_k is the number of data points used for training by a single edge server i, $\nabla F_i\big(w(i,j)\big)$ is the average gradient of the edge device's local data [32]. This equation is usually used to update the model weights on the edge device.

The entire training process on the cloud is periodical, with an arbitrary number of communication iterations between each of the connected edge servers. The federated learning algorithm for the NPP edge-to-cloud integration is summarized as follows (adapted from [31]):

- *Cloud server broadcast*: The cloud server broadcasts the aggregated global weight parameters, w , to all local edge devices connected via the wireless network.
- *NPP edge device updating stage*: The NPP edge devices use their private individual training data to optimize the broadcasted global model weights. Stochastic gradient descent is used to compute updates in minibatches over multiple epochs, minimizing the bandwidth when using Equation (1).
- *NPP edge device broadcast*: When training completes, each edge device broadcasts optimized parameters to the cloud server.
- *Cloud aggregation*: The cloud server performs the global weight aggregation using different weights, A_j, based on dataset size. The updated model is then shared to the edge device, and the iterations continue until optimum accuracy is achieved:

$$w_{i,t+1} = \sum_{j=1}^{N} A_j \, w_{i,t+1}^{j} \qquad (2)$$

Next, the proposed advanced federated DL and ML algorithms will be discussed.

Federated Support Vector Machine: For enabling the real-time fault detection classification algorithm to provide efficient results for alerts and alarm notifications, a federated SVM algorithm is recommended. This technique would aggregate all the separated failure/healthy sensor data collected from all NPP sites on the cloud, using the federated averaging technique described in [32]. The federated SVM algorithm would determine the federal hyperplane margin on the cloud without sharing the independent data from the various NPP sites.

Federated Deep Transfer Learning: The federated deep transfer learning approach for NPPs, using data from circulating water system, was implemented in [34]. This method can be optimized and implemented on the cloud platform by using the DL strategies proposed in [35]. With this method, a sparse autoencoder network with sensor failure history can be initially trained on the edge server, then the network weights can be updated and aggregated in real-time on the cloud. The federated deep transfer learning technique would adapt and enhance the feature extraction characteristics from historical fault data in order to effectively estimate RULs despite missing or limited operational data.

Federated Advanced Long Short-Term Memory: This method is recommended for time series data from NPP sensors in each privately connected edge device. Federated LSTM (FLSTM) could be used for predicting sensor equipment RULs. Conventional LSTM models tend to use high bandwidths in the edge-cloud distribution computations, as they require high-latency updates of large number of weights and bias parameters. Network latency is improved in the FLSTM model by using the moving average strategy [32] to minimize the number of consecutive LSTM blocks and reduce the training time at the edge level. The FLSTM random topology proposed in [32] can be initially implemented on the cloud, then further applied to the simulation on the edge devices, as shown in Figure 17. In addition, a hybrid model that incorporates a convolutional neural network with a stacked bi- and uni-directional LSTM, as implemented in [36], could also be developed on the cloud platform to extract failure features in the sequential time series data by using the forward and backward dependency stacked LSTM layers to enhance the RUL predictions. An autoencoder deep neural network could be utilized for dimensionality reduction and feature extraction on the NPP edge devices, while the stacked LSTM model could be used to denoise and encode the temporal information [36] on the cloud to enhance prediction accuracy.

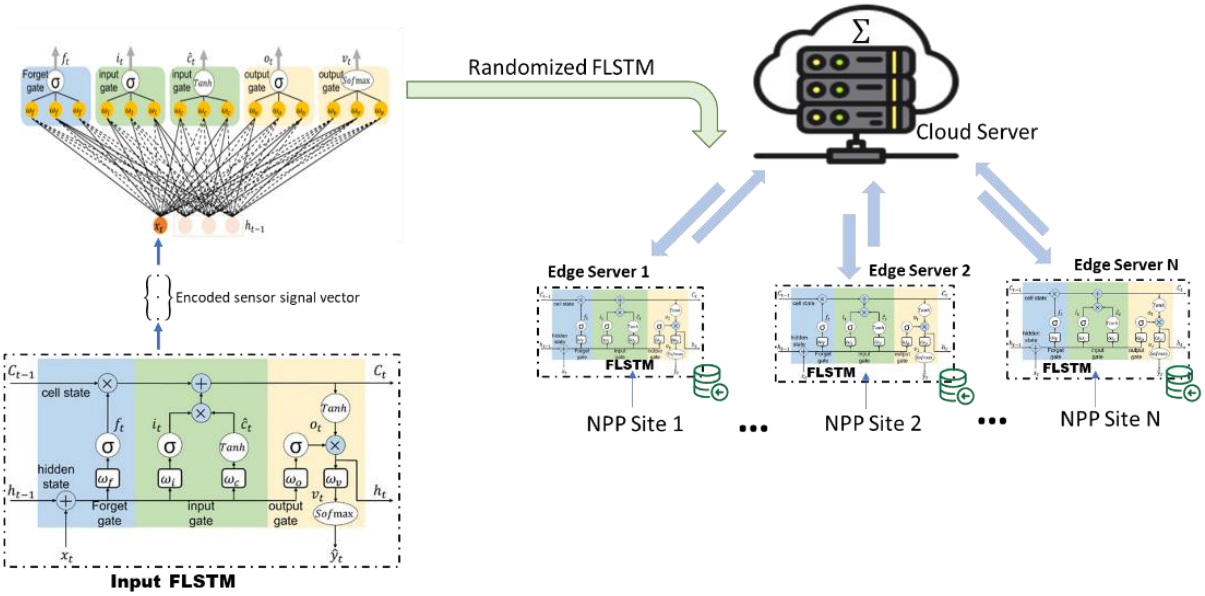

Figure 17. Proposed real-time FLSTM topology for NPP PdM.

3. REGULATIONS

This section covers regulations and other relevant topics related to cloud-based servers. The Office of Management and Budget (OMB) has issued guidance documents to help improve the dissemination of information regarding the operation of federal programs. Over two-thirds of federal agencies are now moving to cloud computing, and this includes the Nuclear Regulatory Commission (NRC).

Although the Federal Risk and Authorization Management Program (FedRAMP) is a government-wide program that promotes the adoption of secure cloud services across the federal government by providing a standardized approach to security and risk assessment both for cloud technologies and federal agencies, the requirements for federal agencies can be helpful in identifying how this technology can be used by non-government agencies.

Finally, licensees will have to address cybersecurity risks as they relate to cloud computing, including the protection of digital assets and the determination of whether any export control issues are associated with the storing and accessing of non-public information on servers outside the control of the licensees.

3.1 The Federal Government's Move to Cloud Computing

OMB Circular A-130 was first issued in December 1985 to meet the information resource management requirements detailed in the Paperwork Reduction Act of 1980. Specifically, the Paperwork Reduction Act assigned the OMB director the responsibility to develop and maintain a comprehensive set of information resource management policies for use across the federal government, and to promote the application of information technology to improve the use and dissemination of information in the operation of federal programs. The initial release of the Circular provided a policy framework for information resource management across the federal government.

Since the initial release of Circular A-130, Congress has enacted several additional laws—and OMB has issued several guidance documents—related to information technology management in federal agencies. To account for these new laws and guidance, OMB revised the Circular in 1994 [37], 1996 [38], and 2000 [39].

FedRAMP was established in 2011 as a government-wide program promoting the adoption of secure cloud services across the federal government by providing a standardized approach to security and risk assessment for cloud technologies and federal agencies [40].

The latest revision of Circular A-130 was released on July 27, 2016 [41,42]. The revised Circular represents a shift from viewing security and privacy requirements as compliance exercises to understanding them as crucial components of a comprehensive, strategic, and continuous risk-based program. With this change, federal agencies wanted an answer to the following question: "How do we secure it, how do we procure it, and where do we get the talent to actually help us make the successful transition to cloud technologies?" To address this, OMB published its strategy to accelerate agency adoption of cloud-based solutions in Cloud Smart [43].

Cloud Smart aims to help government agencies "fully actualize the promise and potential of cloud-based technologies while ensuring thoughtful execution that incorporates practical realities." Five years after the OMB went to Cloud Smart, a new survey revealed that two-thirds of federal information technology (IT) leaders said their agencies are now using, or are starting to use, the cloud for "mission critical applications."[44] The increased willingness to host mission-critical applications—and not just business or research data—in the cloud reflects the growing trust that agency IT executives say they have in the security of federally approved cloud services.

In February 2022, FedScoop conducted a survey of federal agencies, government contractors, and system integrators. Of the federal agency IT leaders polled in the survey, a greater number (27%) believed they could "maintain the greatest security control over their strategic data" via a federally approved, FedRAMP-authorized cloud service than said they put their trust in their own on-premises data centers (23%) [44,45]. Managed by the U.S. General Services Administration, FedRAMP provides a standardized approach to security assessment, authorization, and continuous monitoring of cloud products and services. But when it comes to handling cloud workloads, respondents said they put greater trust their in-house staff than in third-party service providers. Nevertheless, long-standing concerns about protecting data in the cloud are giving way to a broadening belief that data can be managed more securely in a cloud than in traditional agency-built/managed data centers.

In the same survey, roughly half of all respondents said they trusted government-approved cloud service providers to securely store, process, and analyze data for a significant range of use cases, including for research (51%), business and financial operations (47%), and employee/human-resource information (47%). More than 4 in 10 respondents similarly trusted federally approved cloud services to handle mission-critical operating data.

Many federal agencies, including NRC, are moving to cloud-based computing. Significantly, more than a third (36%) of federal IT executives say they now trust their classified and sensitive data in government-authorized clouds (albeit using more exclusive cloud services that have demonstrated a higher standard of security controls). Although OMB Circular A-130 is not directly applicable to NPPs, the federal agencies' concerns regarding barriers to cloud adoption (i.e., security, workforce, and procurement) are equally applicable to NPPs.

3.2 NRC Adoption of Cloud Smart

Many offices within NRC are already participating in cloud-based services (e.g., adjudicatory processes for high-level waste) [46]. NRC's objectives in adopting Cloud Smart are to [47]:

- Improve security, cost effectiveness, efficiency, agility, and scalability in delivering IT services
- Align with OMB's Cloud Smart policy and the Federal Cloud Computing Strategy
- Accomplish appropriate system and application migrations to cloud services as part of complying with Federal Data Center Optimization Initiative mandates

- Establish consistent cloud solution planning and migration practices
- Reduce the risks to IT delivery, availability, and performance through a more distributed and consistent infrastructure and platform environment.

NRC is evaluating several potential options for replacing its licensing service network [48]. One involves utilizing cloud-based technology to store, index, search, and retrieve a collection of high-level waste headers/documents. In its simplest terms, cloud-based solutions use on-demand resources that are part of a provider's shared cloud environment or infrastructure. Prominent examples of cloud-based services include Azure, Gmail, Microsoft Office 365, and Dropbox.

NRC recognizes that any cloud provider must follow the guidelines set by FedRAMP: "…They must be FEDRamp certified. There is no ifs and buts about it" [49]. That is, federal regulations require sites administered by federal agencies to be located on a .gov subsite of the agency and to be evaluated for a cloud solution, thus ruling out physical computers and an offsite data center. Note that the requirement that the cloud provider be FedRAMP certified applies to government agencies, not private companies. However, using FedRAMP-certified cloud providers may reduce the burden of regulatory review and uncertainty for licensees.

FedRAMP certification may limit the options of cloud-based service providers. In addition, the providers must have platforms that are authorized to be run there. Any platform to be used by NRC must not only be authorized by NRC (i.e., it must have an Authority to Operate granted by NRC), but because it will impact the infrastructure of the cloud provider, the cloud provider must be NRC-approved as well. This is because, in the event of an intrusion, other customers in the cloud could potentially be impacted.

NRC recognizes that using shared systems and environments introduces shared risks, vulnerabilities, and exposures, and that untrusted parties may have access to the same shared resources. All risks must be carefully considered, then mitigated or accepted before using cloud computing and hosted solution providers. NRC capabilities may be hosted by another federal agency or a third-party service provider, but that hosting agency or provider must meet all NRC cybersecurity requirements.

NRC requires that applications for electronic commerce services, including cloud storage services, shall use encryption to protect sensitive data at rest and data in motion, and shall provide protections commensurate with the level of risk associated with the value of the transaction and the application [50]. It is assumed that any necessary NRC review of cloud-based servers would require similar protection.

The major components of NRC's cloud strategy are [51]:

- Azure Commercial (IaaS and PaaS)
- AWS, NRC RES-managed
- Other SaaS.

3.3 Regulatory Requirements for Cloud-based Servers

Regardless of how data are stored and shared, licensees must still be able to protect safety-related and important-to-safety functions. To accomplish this, licensees must analyze digital computer and communication systems/networks, identify which assets require protection against cyberattacks, and establish, implement, and maintain a cybersecurity program for the protection of those assets.

3.3.1 Protection of Digital Computer and Communication Systems and Networks

10 CFR 73.54 [52] requires, in part, that licensees provide a high level of assurance that digital computer and communication systems/networks are adequately protected against cyberattacks—up to and including the design-basis threat.

In particular, 10 CFR 73.54(a)(1) requires licensees to protect from cyberattack those digital computer and communications systems/networks associated with the following categories of functions:

- Safety-related and important-to-safety functions
- Security functions
- Emergency preparedness functions, including offsite communications
- Support systems and equipment that, if compromised, would adversely impact safety, security, or emergency preparedness functions.

10 CFR 73.54(a)(2) requires licensees to protect the systems and networks from any cyberattacks that would:

- Adversely impact the integrity or confidentiality of data and/or software
- Deny access to systems, services, and/or data
- Adversely impact the operation of systems, networks, and associated equipment.

To accomplish this, 10 CFR 73.54 (b) requires the licensee to:

1. Analyze digital computer and communication systems and networks, and identify those assets that must be protected against cyberattacks to satisfy paragraph (a) of this section.
2. Establish, implement, and maintain a cyber security program for protecting the assets identified in paragraph (b)(1) of this section.
3. Incorporate the cyber security program as a component of the physical protection program.

Furthermore, digital plant systems are typically isolated from other devices by firewalls, digital diodes, etc., in order to prevent cyberattacks or unwanted communications.

The cyber-security incident at Davis-Besse Nuclear Power Station in January 2003 shows how connecting the corporate network to the plant network can be susceptible to attack [53]. First Energy Nuclear Operating Company (i.e., the licensee)'s corporate network, which is linked to Davis-Besse's plant network, is also connected to external networks. Among the many access control policies enforced by Davis-Besse's corporate firewall was the policy of disallowing any data using the User Datagram Protocol into the network by closing port 1434 of the firewall. The Slammer worm propagates itself by making 376-byte packets and then sending them to randomly chosen IP addresses in User Datagram Protocol port 1434. Closing port 1434, as was required by policy, would have protected Davis-Besse's networks from being infected by the MS SQL slammer worm. However, behind the firewall, a consultant had created a connection to the consultancy's office network. This allowed the Slammer worm to bypass the firewall and infect First Energy's corporate network. From there, the worm infected the plant's process control network. The traffic generated by the worm clogged both the corporate and control networks.

Although the operators were burdened by these losses, the event was not deemed significant, as the plant control and protection functions remained unaffected. This event:

- Occurred at the Level 0 to Level 1 communication boundary (internet to corporate wide area network)
- Did not impact protection or control systems
- Represented a loss of electronic and physical access control

This event shows that any data transmitted to a cloud-based server must come from a network outside the plant's internal network.

3.3.2 Export Control

NRC, the Department of Commerce (DOC), and the Department of Energy all have requirements regarding the transfer of technology. More specifically:

- 10 CFR 110 (NRC) — §110.8 provides a list of nuclear facilities and equipment that fall under NRC export licensing authority; Appendix A (in the CFR) provides an illustrative list of nuclear reactor equipment falling under NRC export licensing authority.
- 15 CFR 730-774 (DOC) — "Software" that is "specially designed" or modified for the "development," "production," or "use" of items controlled by 2A290 or 2A291 is export controlled.
- 10 CFR 810 (Department of Energy) — This restricts the transfer of technology for the development, production, or use of equipment or material especially designed or prepared for any of the activities listed in 10 CFR 810.2(b). This part does not apply to exports authorized by NRC, the Department of State, or the DOC.

If the operational data transferred to a cloud-based server may entail restrictions on the access to and management of the server, such restrictions should be evaluated prior to utilization of that cloud-based server. Such restrictions would likely include that the data remain in the U.S. and that the providers/administrators be U.S. citizens. If the data are not export controlled, users of cloud-based servers would likely face similar data protection requirements.

NRC has incorporated the NIST Risk Management Framework into the terms and conditions of its contracts and service-level agreements with FedRAMP-authorized cloud service providers. The Agency requires its external providers to present appropriate evidence demonstrating that they have complied with the Risk Management Framework by protecting federal information. This includes independent third-party assessments and continuous monitoring. The Agency maintains responsibility for granting external service providers the authority to operate. It is expected that the regulatory uncertainty for licensees would be reduced were they to fall under the same requirements and guidance that apply to external NRC servers and other federal agencies.

4. ECONOMICS

The primary purpose of this section is to analyze the costs/benefits of creating a scalable risk-informed PdM strategy across the nuclear fleet. This envisions moving all PdM and M&D activities, including data and applications, to the cloud. Transition to a cloud-based infrastructure additionally assumes that most or all equipment diagnostic sensors will be upgraded, replaced, or retrofitted with wireless sensors to enable a more efficient and cost-effective data collection and transmission process. The anticipated benefits include a less expensive and more secure IT environment; easier, more-unified user access; and enhanced data integrity and useability. The project contemplates considerable onsite cost savings. Such savings will accrue through the retirement of disparate single-purpose systems, reduced O&M staffing requirements and effort, decreased IT expenses in maintaining systems at the NPP site, and the retirement of maintenance contracts.

As this is a precursor to a lengthier, more in-depth study, several simplifications will be made to the analysis while still laying the framework for the detailed analysis to come. These simplifications include:

1. Estimation of the number of wireless sensors required for a single system in the NPP feedwater and condensate system. This is a surrogate for a larger facility-wide sensor network.
2. Computation of the costs/benefits of this program. Sensor costs will not be included, as they are not a requirement for migrating to a cloud computing paradigm. Moreover, much of the costs associated with adding new sensors for enabling online monitoring will be present irrespective of selecting a traditional M&D center or cloud-based solutions.

3. Assumption of a distribution of sensor types (e.g., temperature, pressure, vibration) and a distribution of sensor environmental hardening levels.
4. Estimation of networking costs by first considering the costs of acquiring the equipment, then factoring in the ongoing maintenance costs. The ongoing maintenance costs will be directly used in the cost-benefit analysis, and the acquisition costs will be depreciated over 5 years.
5. Assumption of a 5G *active* DAS system as the backbone of the facility-wide network. This is most appropriate, given the number of buildings and locations in the network. As this is a preliminary study, we compute the DAS costs by using the cost per square foot in-building wireless costs and employing approximate square footage measurements of a representative facility.
6. Estimation of the yearly data generated and allocating them evenly over the year to estimate cloud and network data rates. In future work, we will examine actual daily data production rates as well as anticipated user activity in accessing the data in the cloud.

4.1 Sensor Costs

The first task is to estimate the cost of purchasing and installing approximately 600 wireless sensors at a generic NPP site—specifically for the feedwater and condensate system. The benefits of wireless sensing technology within a NPP extend well beyond cloud applications. Equipment is generally spread over a large footprint, with much of the cost in adding new instrumentation being attributable to the cabling for wired sensors. The exact function of each individual sensor is not explicitly defined but is expected to conform largely to the functions listed in Table 2. Furthermore, the simplified architecture depicted in Figure 18 will be adopted for estimating the sensor configuration. This section focuses on estimating the cost of the sensors (shown in red in Figure 18) and the necessary labor to install, configure, and provision the sensors onto the onsite network. Onsite wireless infrastructure and computing are calculated separately, as are any costs associated with transmitting data into the cloud. Value-added cloud services such as dashboards for monitoring, analytics, and value-added ML are likewise calculated separately.

4.1.1 Sensor Configuration

For the purpose of this report, "wireless sensors" is meant to include conventional sensing devices (e.g., thermocouples) as well as the electronics required to convert the sensor output into an electrical signal, filter that signal, digitize it using an analog/digital (A/D) converter, and then transmit it to a receiver (wireless module). Depending on the data acquisition speeds and application requirements, it may be desirable for the data to be preprocessed at the sensor prior to transmission. For example, averaging, down sampling, and fast Fourier transform can be performed at the sensor, potentially benefitting data egress and decreasing the throughput demand placed on the wireless network. Other considerations include offsetting the increased battery consumption due to faster data rates and increasing the battery demand by conducting preprocessing at the sensor.

As depicted in Figure 18, sensor paths may come in one of at least three configurations. Some sensors—especially those specifically designed for NPP applications and/or that are already installed—may not be integrated with an A/D converter or a wireless transmitter. Those three components (i.e., sensor + A/D + wireless module) will need to be acquired separately and then integrated at the expense of additional parts and labor. It is believed that most sensors with analog outputs can be interfaced directly to wireless devices that include A/D conversion and provide down sampling and other logic (see the thermocouple and A/D converter in Figure 19). Finally, some sensors may be fully integrated to provide sensing, A/D, logic, and a wireless transmitter. Such is the case with the vibration sensor shown in Figure 20.

Table 2. Range of sensor types required for diagnostic monitoring of NPP equipment.

Nuclear Power Plant Sensor Types			
1	Accelerometer	9	Infrared
2	Electrical current	10	Optical
3	Electrical voltage	11	Pressure
4	Fluid level	12	Proximity
5	Fluid velocity	13	Radiation
6	Fluid volume flow	14	Smoke
7	Gyroscope	15	Temperature
8	Humidity	16	Vibration

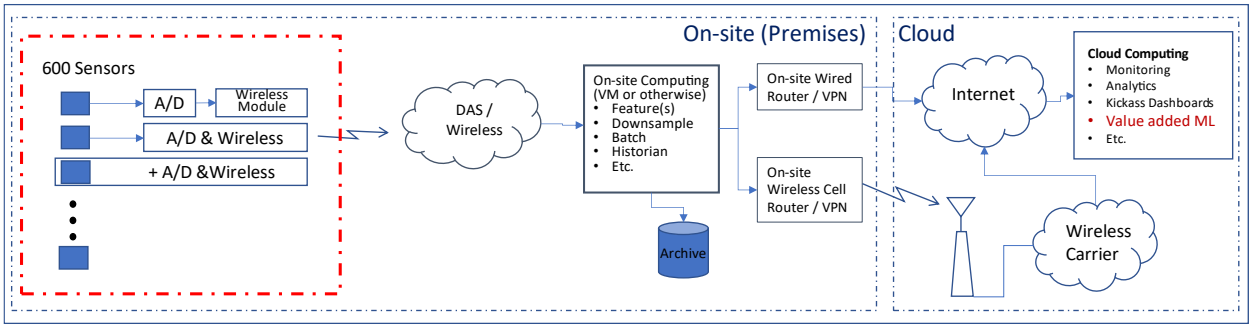

Figure 18. Abstracted sensor data pipeline, with a focus on wireless sensors (in red, on the left-hand side of the diagram).

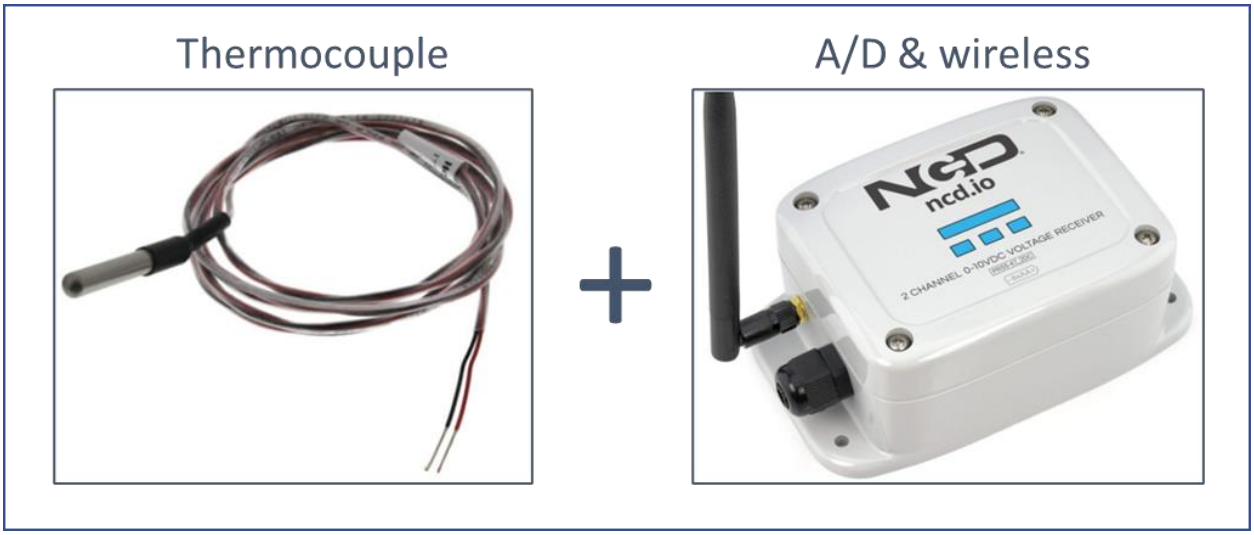

Figure 19. A sensor (thermocouple) married to a combined A/D converter and wireless transmitter.

Figure 20. An integrated vibration sensor with A/D conversion, down-sampling logic, and a wireless transmitter (also includes a magnetic mount) [55].

4.1.2 Costs

With these constraints in mind, an estimate regarding the purchase, installation, configuration, and provisioning of 600 wireless sensors is given in Table 3. This estimate includes sensors and intermediate equipment necessary to wirelessly transmit digital measurement data. A Gaussian distribution was applied to the sensor grade, as some sensors may require additional environmental requirements (e.g., hardened for radiation-intensive environments) or greater precision. Downstream onsite DAS equipment and edge computing devices are not included herein. It is estimated that each sensor integration will require approximately 3 hours, at an average cost of ~$100/hour (600 × 3 hours × $100 = $180,000).

Ongoing maintenance of sensors, including the resetting of components and the replacement of batteries and devices, is estimated at 10% of the total costs per year. With this methodology, the costs associated with maintaining the wireless sensor installation equate to an annual cost of approximately $76,000.

Table 3. Estimate of sensor and installation costs associated with instrumenting the NPP feedwater and condensate system for online M&D [55-56].

Sensor Grade	Number of Sensors	Price ($)	Subtotal ($)	Labor/Install Costs ($)
Low	70	$300	$21,000	$21,000
Low/Mid	120	$500	$60,000	$36,000
Mid-Range	220	$800	$176,000	$66,000
Mid/High	120	$1,200	$144,000	$36,000
High Grade (hardened)	70	$2,500	$175,000	$21,000
—	—	—	$576,000	$180,000
Grand Total	—	—	—	$756,000

4.2 Distributed Antenna System Costs

This section addresses the costs associated with the installation and configuration of a DAS used by the wireless sensors to transmit and communicate to local (i.e., onsite) servers. As was mentioned, the exact nature of the DAS configuration and associated engineering cost estimates will be the focus of future work. Here, we use publicly available cost estimates based on the square footage of the coverage area in order to establish an approximate cost for bringing a DAS online. To estimate total square footage, we will use the estimates below for the DC Cook facility, then multiply the square footage values by 0.75 to obtain a single-unit site estimate. This assumes that facilities (e.g., the warehouse space and IT infrastructure) are somewhat more efficient for a two-unit site than for a single-reactor site.

4.2.1 DAS Considerations

Several factors must be considered in choosing a DAS for this type of application. A DAS consists of a distribution system and a signal source. The signal source determines whether the DAS is *active* or *passive*. In either case, the signal source is connected to the telecom provider network, thus allowing signals from the covered area to be transmitted across the entire network. The signal source can be the local cell signal from the carrier (passive DAS) or can come from an on-premises Base Transceiver Station (BTS). BTS, NodeB, eNodeB, and gNodeB each refer to a type of technology used inside cell phone towers to generate a cellular signal. For simplicity, these technologies are often simply referred to as a BTS signal source.

Generally, any space comprising less than 100,000 square feet can be serviced by a passive DAS. Anything over 500,000 square feet will almost certainly require an active DAS. Figure 21 shows the basic configuration of a passive DAS. Item # 1 receives the signal from the existing cell tower and distributes it throughout the facility via antennas (Item # 3). While convenient and less expensive, this arrangement does not afford the site extra capacity. Consequently, the cost estimates within this section are based on an *active* DAS.

Ordinarily, multi-carrier installations are utilized for a DAS that provides mobile services for voice users. This adds significant cost and complexity as a result of multiple carriers needing to connect to the in-building system. For NPP applications, a single carrier for the data transmission to the cloud provider can easily be implemented. Finally, the choice of data transmission technology must be made. While 5G is newer and faster than 4G, it has significant penetration drawbacks when utilized in a building with high-density walls and floors. 5G is not a simple speed enhancement applied to existing networks. Rather, it behaves very differently from 4G/LTE, as 5G signals do not pass through structural materials (e.g., steel, wood, glass, concrete) effectively. This issue will be detailed in future work.

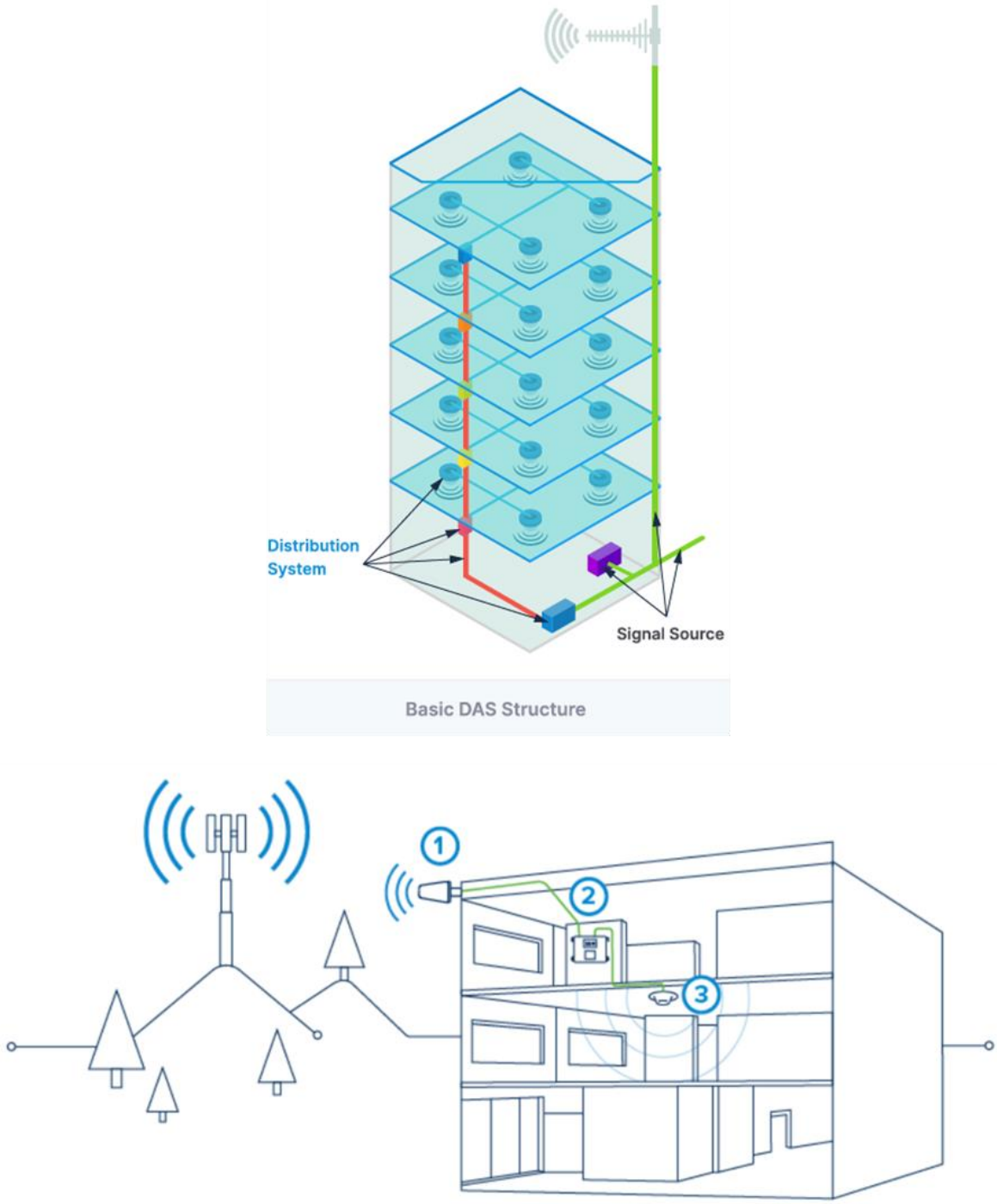

Figure 21. Top: Basic DAS structure. Bottom: Components of a passive DAS system.

4.2.2 Model Facility

Cost estimates will be based on the square footage of the coverage area required within a NPP to support wireless online monitoring. To estimate the total square footage, the DC Cook facility in Berrien County, Michigan (see Figure 22), was used [57]. This site was chosen to serve as a model, due to the extensive knowledge that exists about the building contents and functionality. For square footage calculations, The DC Cook footprint will be reduced by 0.75 to obtain a single-unit site estimate (This

assumes that facilities such as warehouse space and IT infrastructure are somewhat more efficient for a two-unit site than for a single-reactor site.)

Figure 22. The DC Cook facility in Berrien County, Michigan, features two functioning power plants [58].

Figure 23. DC Cook facility in Berrien County, Michigan, highlighting the six warehouses (W1–W6) [58].

4.2.3 Costs

As a simple rule of thumb, the costs associated with DAS equipment and the installation and commissioning thereof can be estimated at $4.00 per square foot for large installations [58]. Utilizing the information presented in Figure 22, a breakdown of the costs for full DAS coverage at a NPP is given in Table 4.

In computing the DAS cost, we considered only the square footage of one containment building and one turbine building associated with a single unit. It may be reasonable to exclude full coverage within warehouses, as they can be handled by the standard internet connection to the cloud and thus do not require DAS service per se. The RFID system networks and aggregates all the signals from the individual RFID tags, and most RFID systems conveniently connect to the internet.

Table 4. Cost estimates based on coverage requirements and square footage for DAS equipment and the installation and commissioning thereof [58].

DAS Cost Calculation	Building	Square Footage	Subtotal
A	Turbines	81,500	$326,000.00
B	Safeguards Building	32,370	$129,480.00
E	Containment Buildings	6,000	$24,000.00
F	Circulating Water Pumps	48,000	$192,000.00
W1	Warehouse	20,000	$80,000.00
W2	Warehouse	20,000	$80,000.00
W3	Warehouse	20,000	$80,000.00
W4	Warehouse	10,000	$40,000.00
W5	Warehouse	10,000	$40,000.00
W6	Warehouse	10,000	$40,000.00
—	—	With Warehouse	$1,031,480
—	—	Without Warehouse	$671,480

4.3 Onsite Edge Computing Costs

Data will undergo onsite processing and storage prior to being transmitted to the cloud. The amount of processing and the degree to which the onsite facility will replicate the cloud environment shall be explored in future work. For the present study, we considered a minimal amount of processing and storage. Integrity checks and packaging will be conducted onsite in addition to temporary storage, prior to cloud backup. This enables the highest number of labor and hardware savings in moving the maintenance data and software to the cloud. We estimated the cost of the edge computing needed to transport the output of 600 wireless sensors into the cloud from a yet-to-be-determined NPP location.

For the purpose of estimation, we assumed the simplified architecture depicted in Figure 24. This section focuses on estimating the cost of the edge computing infrastructure (boxed in red), as well as the labor to install, configure, and maintain it.

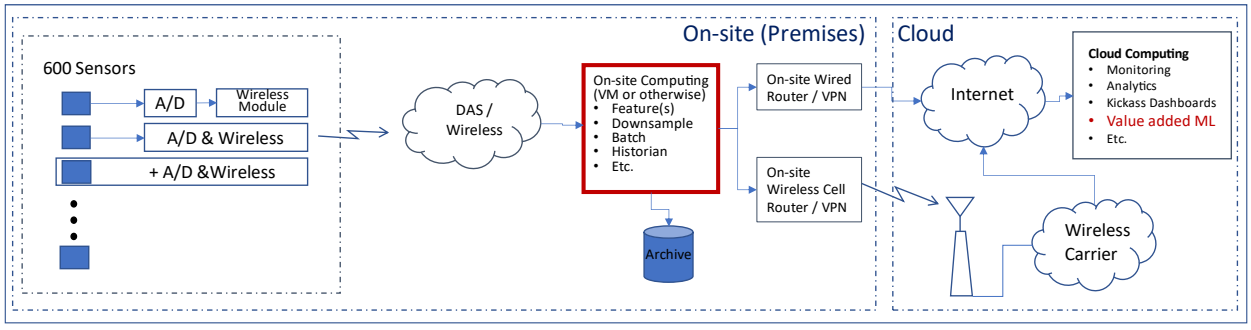

Figure 24. Abstracted sensor data pipeline, with a focus (in red) on onsite edge computing.

The current NPP data collection and retention convention is on the order of a single 64-bit floating point number for each sensor, at collection intervals of 1 second. For 600 sensors, this represents a data generation rate of approximately 40 kbps, which would pose no strain on any modern network throughput (equivalently ~150 GB/year, two orders of magnitude smaller than a $300 hard drive). This line of reasoning suggests the throughput and data storage needs to be relatively small. A very basic estimate would consider a single server blade and redundant storage, duplicate equipment for redundant hardware, and fault tolerance and failure switchover architecture.

Finally, additional cost components that include (1) labor for the hardware setup and the configuration of the platform and (2) a relatively larger expense for software development on the platform should be expected. Data featurization, down sampling, compression, and archival should be considered for each of the independent data streams.

The total fixed costs associated with onsite edge computing are given in Table 5.

Table 5. Costs pertaining to the onsite edge computing infrastructure for handling approximately 600 data streams associated with the feedwater and condensate system.

Item	Cost
Server Blade	$2,000
Redundant Server Blade	$2,000
Failover Infrastructure	$1,000
Setup Labor	$1,000
Software Development	$20,000
Total	$26,000

4.4 Cloud Costs

4.4.1 Cloud Infrastructure

An estimate of the necessary cloud-based network and compute infrastructure, along with the associated costs for the feedwater system, is given herein. The myriad architecture and product choices and interdependencies mean that (1) the final design will evolve as the determination and details of the sensors, data inclusion, security requirements, and end-use cases progress, and (2) there is significant opportunity for cost and performance optimization. For this study, Microsoft Azure was selected, owing to its ease of integration with existing Microsoft services already adopted by some industry partners.

4.4.2 Architecture Overview

Transitioning to cloud-based M&D reduces the necessity for onsite infrastructure and associated administration, while also ensuring that systems and security are kept up to date. The onsite hardware required for the design shown in Figure 24 consists of a gateway for uploading wireless sensor data and a server for collecting wired sensor data. Figure 25 shows a breakdown of the data types that are expected to be generated. The server pushes these data to the organization’s existing cloud storage (e.g., OneDrive) and syncs the wireless data to the local historian, which serves as an important onsite repository. This architecture allows for the preprocessing of raw sensor data (e.g., down sampling) and the performing of other analytics prior to sending the data to the final cloud endpoints.

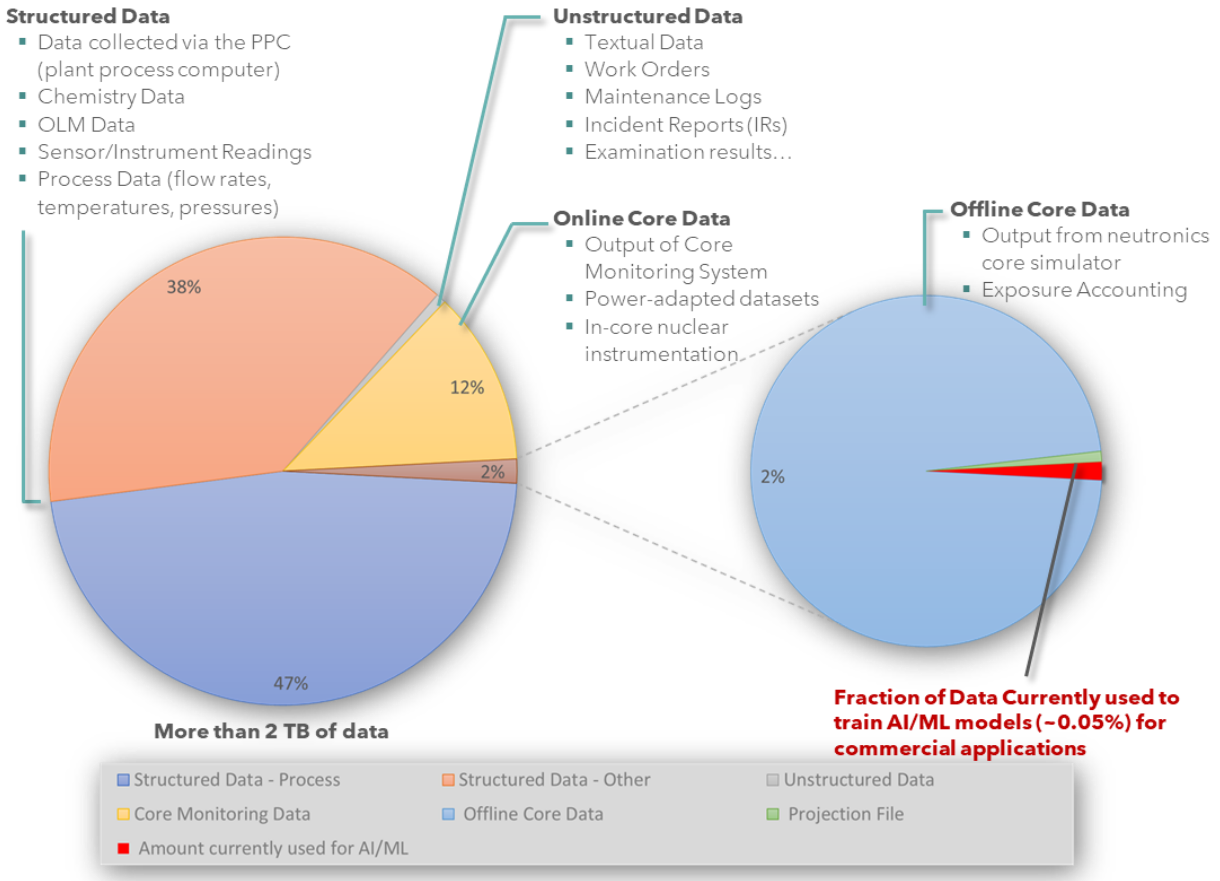

Figure 25. Breakdown of data generated during a full 2-year cycle at a commercial boiling-water reactor.

4.4.3 Cost Estimates for the NPP Feedwater and Condensate System

Records from our partner's plant data historian were used to generate an initial estimate of the number of sensors currently used to support the plant feedwater and condensate system. Our method led to the identification of approximately 600 measurement points, serving as the basis for the current analysis. A collection rate of 1 Hz, with 4 bytes for each datum, yields around 70 GB/yr.

Alternatively, the data throughput for this system can be estimated by utilizing a known ~2 TB complete (2-year cycle) dataset from the plant process computer, along with the ~3% incidence rate of feedwater-related descriptions. This method generated an estimate of 30 GB/yr. While likely generous, these numbers provide guidance on selecting data throughput and storage capabilities in the cloud cost estimates.

In applying the Azure pricing calculator (https://azure.microsoft.com/en-us/pricing/calculator) to the estimated data throughput and resource utilization, the "pay as you go" service costs of the Azure real-time analytics scenario example (Figure 26)—hosted in the U.S. Gov. Virginia region—are estimated at $6K/mo. (see Table 6). A more conservative estimate in the non-government "East U.S." region came to $3K/mo.

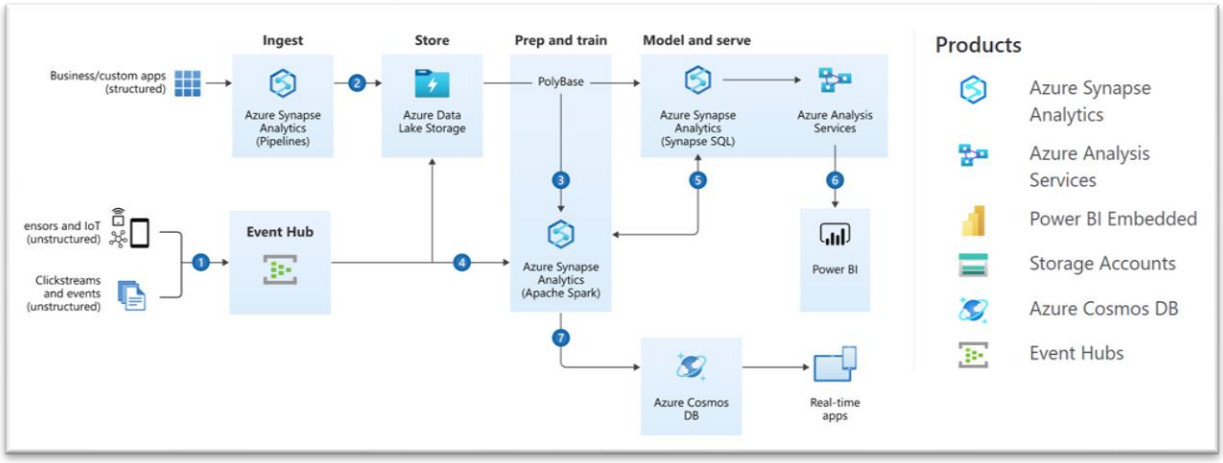

Figure 26. Microsoft Azure real-time analytics scenario used for cost estimation [60].

Table 6. Cloud cost estimates obtained utilizing Azure [60].

Service Category	Service Type	Region	Description	Estimated Monthly Cost
Databases	Azure Synapse Analytics	U.S. Gov Virginia	Tier: Compute Optimized Gen 1, DWU 100 × 720 Hours, 1 TB of storage	$2,248.70
Analytics	Azure Analysis Services	U.S. Gov Virginia	Basic B2 (Hours), 1 Instance(s), 720 Hours	$774.00
Analytics	Power BI Embedded	U.S. Gov Virginia	1 node(s) × 720 Hours, Node type: A2, 2 Virtual Core(s), 5GB RAM, 301–600 Peak renders/hour	$1,807.29
Storage	Storage Accounts	U.S. Gov Virginia	Data Lake Storage Gen 2, Standard, LRS Redundancy, Hot Access Tier, Flat Namespace File Structure, 1,000 GB Capacity - Pay as you go, Write operations: 4 MB × 1,000 operations, Read operations: 4 MB × 1,000 operations, 1,000 Iterative read operations, 1,000 Iterative write operations, 1,000 Other operations. 1,000 GB Data Retrieval, 1,000 GB Data Write	$187.60
Databases	Azure Cosmos DB	U.S. Gov Virginia	Autoscale provisioned throughput, Always-free quantity disabled, Multiple Region Write (Multi-Master) - US Gov Virginia (Primary Write Region); 5,000 RU/s × 730 Hours × 100% Average Utilization; 100 GB transactional storage, analytical storage enabled; Dedicated Gateway not enabled	$789.25
Analytics	Event Hubs	U.S. Gov Virginia	Standard tier: 1 Throughput unit(s) × 730 Hours, Capture enabled, 1 million Ingress events	$94.93
Support	—	—	—	$0.00
Total	—	—	—	$5,901.77/mo.

4.5 RFID Network Costs

Inventory management is an important function in all businesses but especially for nuclear power providers. The amount of inventory is easily in the tens of millions which we estimate to be worth somewhere between $10M and $50M per reactor in Constellation's fleet (for the purpose of this report, $40M is used as an estimate). The principal issues are the expense of frequent cycle counting, inventory shrinkage, and stockouts. While inventory shrinkage can be a high dollar issue, stockouts can exacerbate problems as they can extend outage duration until the part(s) can be acquired leading to potentially significant losses in revenue.

RFID networks provide a way to carefully track and manage inventory. Each item in inventory is tagged with either a "passive" or "active" tag. Passive tags receive energy from the RFID reader and retransmit that energy with their tag ID which is associated with a particular item in inventory. The existence of the item is either initially entered or continuously verified in the inventory database. If the tag is no longer detected, the item is flagged as removed from inventory. Active tags have a small battery that provides a stronger transmission signal, giving longer range between the tag and the RFID receiver. The disadvantage of active tags is the batteries have a finite five-year life, while passive tags have no need for battery replacement. The diagram represented in Figure 24Figure *27* shows the general RFID system layout.

4.5.1 System Components

There are three main components in the system which are described below with representative pricing: The RFID Reader, The RFID Antenna, The RFID tags.

RFID Reader

There are numerous options for RFID readers. The readers have ports that connect to four RFID antennas. The antennas may each be located up to thirty feet away. Some readers support up to sixteen antennas and can read up to 750 tags per second. The tag information will be transmitted via cellular network directly to the cloud or a locally connected system to update the inventory system. A typical reader is shown in Figure *28*. The readers and antennas are typically arranged in a star configuration near the middle of the warehouse aisle to give maximum coverage. This particular reader operates in the ultra-high frequency range which allows greater distance between the tag and the antenna – up to twenty feet. While not being considered for this implementation, RFID systems can be configured to track the movement of items within, into, and out of the warehouse. The cost differential from static tracking is minimal.

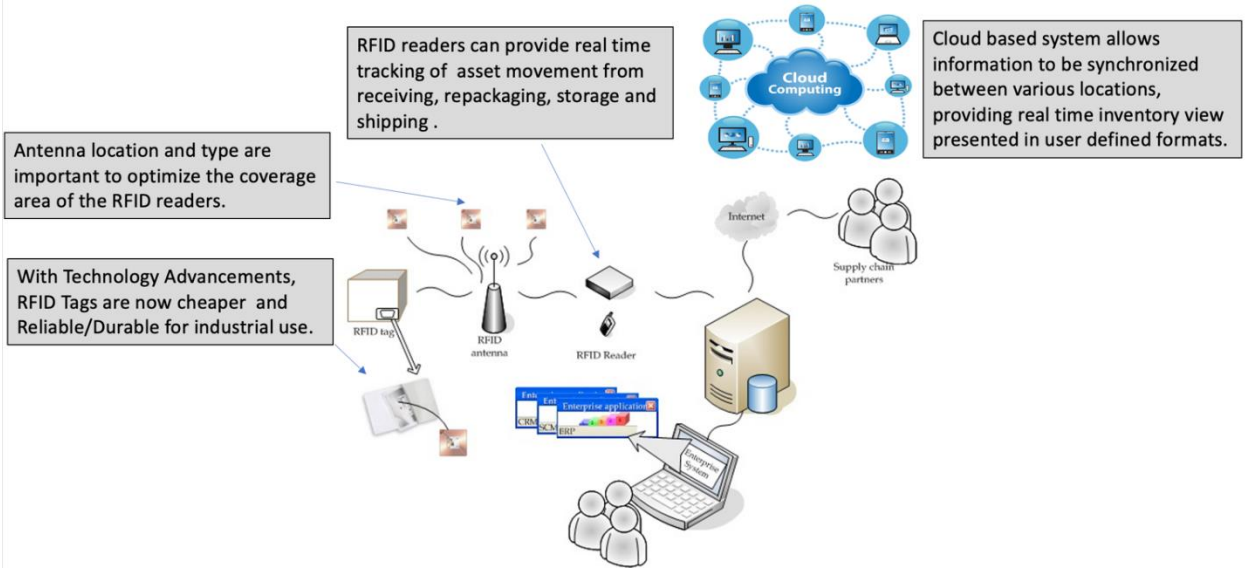

Figure 27. General layout for an RFID System.

Figure 28. Impinj Speedway Revolution R420 ultra-high frequency RFID Reading (left), RFMAX Indoor RFID Antenna (middle), and Vulcan RFID printable labels (right) [61-63].

RFID Tags

RFID tags or transponders are small devices that utilize low-power radio waves to receive, store, and transmit data to nearby readers. RFID tags are comprised of the following main components: a microchip or integrated circuit, an antenna, and a substrate or protective material layer that holds all the components together. Tags come in a very large variety of sizes, shapes, and configurations. Many of the tags are printable which allows initializing blank tags with information that makes identification of warehouse elements simple. RFID tags operate in three frequency ranges: Ultra-High Frequency, High Frequency, and Low Frequency. RFID printable labels have two parts: a face label for printing barcode and human readable information, and an RFID chip which will be encoded by the RFID printer. The RFID labels shown in Figure *28* are blank for customer printing and encoding when using their own RFID printers.

RFID Antennas

Finally, RFID antennas come in a variety of sizes, frequency ranges and network interfaces. The RFMAX antenna, shown in Figure *28*, is representative of a class of antennas we might use.

4.5.2 RFID System Cost

The layout in Figure 22 is used for the purpose of cost estimation, where there are three 20,000 square-foot warehouses and three 10,000 square-foot warehouses. Given the warehouse size for the larger buildings, it is assumed that there are 7 double aisles that are 12 feet wide. The aisles will be approximately 200 feet long allowing for administrative space. This gives 1400 linear feet of aisle space with two sides to cover. Each assembly (consisting of a reader, four antennas, and four cables) can cover approximately 40 linear feet. In total, 35 assemblies will be required for a large warehouse and 18 assemblies for the smaller 10,000 square-foot warehouses.

The cost per assembly is shown in Table 7, and the total system costs are presented in Table 8.

Table 7. Cost per RFID Assembly [61-63].

RFID Assembly	Quantity	Unit Price
Reader	1	$1,100
Antennas	4	$141
Cables	4	$94
Total		**$2040 per assembly**

Table 7. Total RFID system cost for model facility with 3 large (20,000 square-foot) and 3 small (10,000 square-foot) warehouses.

Item	Quantity	# of Assemblies required	Total
20,000 ft^2 Warehouse	3	35	$214,200
10,000 ft^2 Warehouse	3	18	$110,160
RFID Tags	4,000		$1,240
Installation	1		$150,000
Total			**$ 475,600**

4.6 Cost-Benefit Analysis

Numerous computer systems and associated data must be hosted and maintained to measure, track and manage the performance of critical components in the NPP. Additionally, there is staff overhead in extracting data from these disparate systems, some of which require walking through the plant and reading and recording charts through physical rounds. In other cases, data must be located, formatted, and downloaded prior to any application of predictive analytics. This is a time-consuming process that requires significant resources across the entire organization. The new cloud-based environment can aggregate and automate of many of these manual tasks while centralizing the analytics in one place for a risk-informed predictive maintenance program. The cost-benefit analysis that we consider is the base case of all current practices and activities which incur a cost that can be eliminated by transitioning to a cloud-based paradigm.

4.6.1 Present Costs (Base Case)

Present costs fall into five primary categories. Below is a list of these categories along with assumptions made for cost estimation.

1. **Hardware** – servers, data storage devices (for on-demand access and archiving), network equipment.
    a. **Assumptions:** It is assumed that twelve (12) servers and thirty (30) network elements are retired as a result from transition to the cloud. Furthermore, it is assumed that the cost of replacing these systems every five years, as well as the cost of maintaining the systems while in service, is eliminated with a transition to the cloud.
2. **Software** – Database software, purpose-built data collection and retrieval, analytical software. The costs associated with software are primarily maintenance fees for commercial software and programmers and maintenance contracts to maintain purpose-built software.
    a. **Assumptions:** It is assumed that commercial software licenses are cancelled, and the maintenance fees are no longer required with transition to the cloud. We assume that purpose-built software is retired and contract programming to maintain and update this software is discontinued.
3. **IT Staff –** IT staff maintains plant maintenance computers and systems. There are as many as twelve distinct systems in a single nuclear power plant.
    a. **Assumptions:** It is assumed to require a four (4) Full-time equivalent (FTE) headcount to maintain servers and perform all support functions for hardware and those elements of software not covered under third-party contracts.
4. **Operational Staff** – Personnel required to find, record, log, download, preprocess, and analyze plant and maintenance data to support present maintenance capabilities.
    a. **Assumptions:** It is assumed that line operations management as a group spend at least 1% more time than necessary on reporting and analytics due to inefficient data organization and non-centralization of access to data and analytical software. Moreover, it is assumed that four headcounts are needed to manually walk through the plant, read and record sensor values, and inputting the data into the plant maintenance computer.
5. **Inventory Management** – Warehouse employees perform quarterly cycle counts to align inventory system records with physical inventory. Additionally, there is "shrinkage" of inventory due to miscounting, and mistakes in recording inventory withdraws as its consumed.
    Assumptions: It is assumed that fewer cycle counts are needed with RFID tracking and there will be 0.5 % less shrinkage.

From these categories, the itemized breakdown of current costs that can be eliminated by transitioning to a cloud-based predictive maintenance strategy is presented in Table 9. As evident in the table, over 55% of the savings come from reduced IT support and equipment recurring costs. Another 26% of the savings are expected to come from increased staff efficiency. This does not include the benefit of anticipated savings due to better predictive maintenance and troubleshooting practices.

Table 9. Total RFID system cost for model facility with 3 large (20,000 sq. ft.) and 3 small (10,000 sq. ft.) warehouses.

Category	Number of Items saved	Item Cost	Total Cost	Annualized Cost
Hardware (assumed that servers and network elements replaced every five years)				
Servers	12	$4,500	$54,000	$10,800
Network Elements	30	$800	$24,000	$4,800
Software (annual support costs estimated at 15% total software costs)				

Commercial Software	Base Cost	$200,000	$200,000	$30,000
Purpose-built Software	Base Cost	$500,000	$500,000	$75,000
Operational Staff				
Group (excess fraction 1%-time savings)	50	$150,000	$7,500,000	$75,000
Individual (manual rounds, and data preprocessing)	4	$85,000	$340,000	$340,000
IT Support Staff	5	$150,000	$750,000	$750,000
Facilities Costs (e.g., electricity)	12	$850	$10,200	$10,200
Inventory Management				
Cycle Counting (labor for 10 employees 3 days per quarter)	120 (staff days per year)	$400	$48,000	$48,000
Inventory Shrinkage (0.5%)	0.5%	$50M	$250,000	$250,000
Total				**$1,593,800/year**

Table 8. Summary of cost categories and amortized costs associated with cloud-based implementation of a risk-informed predictive maintenance program.

Item	Cost	Annualized Cost
Sensors		
Acquisition Cost (amortized over five years)	$756,000	$151,200
Annual Maintenance (10% of acquisition cost)	$76,000	$76,000
DAS		
Acquisition Cost (with warehouse coverage, amortized over ten years)	$1,031,480	$103,148
Maintenance Cost (10% of acquisition cost)	$103,148	$103,148
Network Aggregation Equipment		
Acquisition Cost (amortized over five years)	$50,000	$10,000
Maintenance Cost (10% of acquisition cost)	$5,000	$5,000
On-Site Edge Computing		
Acquisition Cost (amortized over five years)	$26,000	$5,200
Maintenance Cost (10% of acquisition cost)	$2,600	$2,600
RFID Network		
Acquisition Cost (amortized over ten years)	$475,600	$47,560
Maintenance Cost (10% of acquisition cost)	$32,560	$32,560
Cloud Computing		
Acquisition Cost	$0	$0
Recurring Cost	$5,902	$70,822
Total		**$602,238/year**

4.6.2 Cloud-Based Cost Summary

Table 10 summarized the cost categories and costs associated with transitioning to a cloud-based implementation of a risk-informed predictive maintenance program. Equipment acquisition costs are

spread over their assumed useful life, being amortized to either five or ten years. Recurring maintenance fees are broken out, separately, as a fixed percentage of the acquisition costs.

4.6.3 Cost-Benefit Analysis

As can be seen by comparison of the cost savings (Table 9) and costs incurred (Table 10) by transitioning to a cloud-based predictive maintenance program, there is a significant advantage to moving the predictive maintenance-related data and applications to the cloud. A large percentage of savings results from lower IT costs both in staff and equipment. With a savings of approximately $1M per year the case is compelling. The intangible benefits are important as well. There is increased cybersecurity protection, better uptime (reduced unplanned downtime due to equipment failure), and load flexibility due to the virtualization features of the cloud.

5. SUMMARY AND PATH FORWARD

As seen in this report, cloud computing can be a cost-effective alternative to onsite storage and diagnostics. By leveraging the cloud's elastic processing power, high availability, and pay-as-you-go methodology, a data shift can be made from plant site computers to cloud servers with integrated security management. There are no major technological hurdles blocking such a transition, though choices must be made concerning deployment strategies and the number of extra wireless sensors that would be added to support this change.

As far as regulations go, many of the external technical assets on which NRC depends (and which include cloud-based services [i.e., infrastructure, platform, software], data center operations, and telecommunication circuits) are not within its direct control [54]. External providers can be public or private sector entities, either domestic or international. More specifically, Federal Information Security Modernization Act and OMB policies require that federal agencies using external service providers offer assurance that those providers meet the same security requirements that federal agencies are required to meet. NRC currently uses FedRAMP requirements for most of its commercial and non-commercial cloud services, and plans to use FedRAMP requirements for all its future commercial and non-commercial cloud services.

Future work will serve to enhance this study through collaboration with a specific NPP. Many of the estimates in this report (e.g., number, type, and placement of wireless sensors; DAS requirements; and cost estimates) were made for a generic plant. By working closely with a NPP, a more realistic account of the transition from onsite to cloud-based resources can be made.

To increase the value of cloud computing, future work will also explore how the data can be shared and evaluated. The focus will be on identifying how data searches can provide the information needed without excessive data provided that requires review by human assets to remove irrelevant data.

6. REFERENCES

1. IAEA, "Artificial Intelligence for Accelerating Nuclear Applications, Science and Technology, Non-serial Publications, IAEA, Vienna (2022). https://www.iaea.org/publications/15198/artificial-intelligence-for-accelerating-nuclear-applications-science-and-technology.
2. Smith, J. A., C. Xu, Y. Deng, K. Araseethota Manjunatha, V. Agarwal. 2020. "Wireless Sensing and Communication Capability from In-Core to a Monitoring Center," INL/EXT-20-59435, Rev 0, Idaho National Laboratory.
3. Al Rashdan, A. Y. 2018. "Development of a Technology Roadmap for Online Monitoring of Nuclear Power` Plants." INL/EXT-18-52206, Rev 0, Idaho National Laboratory.

4. Zhang, S. et al. 2022. “Practical Adoption of Cloud Computing in Power Systems—Drivers, Challenges, Guidance, and Real-World Use Cases,” in *IEEE Transactions on Smart Grid*, vol 13, no. 3, pp 2390–2411. May 2022. https://doi.org/10.1109/tsg.2022.3148978.
5. Mell, P., T. Grance. 2011. “The NIST Definition of Cloud Computing, Special Publication (NIST SP 800-145).” National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-145.
6. Cloudways. n.d. “IaaS vs. PaaS vs. SaaS: How Are They Different?,” Digital Ocean. Accessed November 15, 2022. https://www.cloudways.com/blog/iaas-vs-paas-vs-saas/.
7. Walker, C. M., V. Agarwal, N. J. Lybeck, M. Taylor. 2022. “Development of an End State Vision to Implement Digital Monitoring in Nuclear Plants,” INL/RPT-22-66542, Rev 0, Idaho National Laboratory.
8. NetApp. n.d. “Azure migration planning: Optimize to save time and money.” NetApp. Accessed November 27, 2022. https://www.netapp.com/blog/azure-migration-planning-optimize-save-time-money.
9. Ansley, Bill. n.d. “Wireless in Nuclear at Exelon.” Energiforskmedia. Accessed October 21,2022. https://energiforskmedia.blob.core.windows.net/media/24312/11-exelon_ansley.pdf
10. Agarwal, V., et al. 2021. “Machine Learning and Economic Models to Enable Risk-Informed Condition Based Maintenance of a Nuclear Plant Asset,” INL/EXT-21-61984, Rev 0, Idaho National Laboratory.
11. Microsoft. n.d. “Big data architectures.” Azure. Accessed November 27, 2022. https://learn.microsoft.com/en-us/azure/architecture/data-guide/big-data.
12. Al Rashdan, A. Y., C. J. Krome, S. W. St. Germain, J. Rosenlof. 2019. “Method and Application of Data Integration at a Nuclear Power Plant,” Light Water Reactor Sustainability Program report, INL/EXT-19-54294, Rev 0, Idaho National Laboratory.
13. Microsoft. n.d. “Azure storage redundancy.” Azure. Accessed October 20, 2022. https://learn.microsoft.com/en-us/azure/storage/common/storage-redundancy.
14. Microsoft. n.d. “Azure page blobs pricing.” Azure. Accessed September 21, 2022. https://azure.microsoft.com/en-us/pricing/details/storage/blobs/.
15. Microsoft. n.d. “Storage account overview.” Azure. Accessed September 14, 2022. https://learn.microsoft.com/en-us/azure/storage/common/storage-account-overview.
16. Microsoft. n.d. “Hot, cool, and archive access tiers for blob data.” Azure. Accessed October 7, 2022. https://learn.microsoft.com/en-us/azure/storage/blobs/access-tiers-overview.
17. Microsoft. n.d. “Virtual Machine series.” Azure. Accessed October 17, 2022. https://azure.microsoft.com/en-us/pricing/details/virtual-machines/series/.
18. Küfner, T., S. Schönig, R. Jasinski, A. Ermer. 2021. “Vertical data continuity with lean edge analytics for industry 4.0 production,” Computers in Industry, 125: 103389. https://doi.org/10.1016/j.compind.2020.103389.
19. AVEVA. n.d. “AVEVA™ Insight - Predictive Analytics on AVEVA Connect.” 2021, from https://www.aveva.com/content/dam/aveva/documents/legal/service-documents/AVEVA-Insight-Predictive-Analytics-on-AVEVA-Connect-v1.1.pdf.
20. AVEVA. n.d. “AVEVA releases a new version of AVEVA™ Predictive Analytics.” AVEVA. Accessed October 25, 2022. https://industrial-software.com/community/news/aveva-releases-a-new-version-of-aveva-predictive-analytics.

21. AVEVA. n.d. “AVEVA to release a cloud-native solution for industrial operations data.” AVEVA. Accessed October 25,2022. https://www.aveva.com/en/perspectives/blog/aveva-to-release-a-cloud-native-platform-for-industrial-operations-data.

22. AVEVA .n.d. “Brochure AVEVA™ Predictive Analytics.” AVEVA. Accessed October 25, 2022. https://www.aveva.com/content/dam/aveva/documents/brochures/brochure-aveva-predictive-analytics-02-20-3.pdf.

23. Zhang, W., D. Yang, H. Wang, “Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey,” *IEEE Systems Journal*, vol. 13, no. 3, pp. 2213-2227, May 2019. https://doi.org/10.1109/JSYST.2019.2905565.

24. Damji, J. S., B. Waenig, D. Tathagata, D. Lee. 2020. “*Learning Spark - Lightning-Fast Data Analytics*,” O’Reilly Media, Inc. https://pages.databricks.com/rs/094-YMS-629/images/LearningSpark2.0.pdf.

25. Spark. n.d. “Spark-Classification and regression.” Spark. Accessed November 3, 2022. https://spark.apache.org/docs/latest/ml-classification-regression.html.

26. Microsoft n.d. “Real-time streaming in Power BI.” Azure. Accessed November 5, 2022. https://learn.microsoft.com/en-us/power-bi/connect-data/service-real-time-streaming.

27. Ran, Y, X. Zhou, P. Lin, Y. Wen, R. Deng. 2019. “A Survey of Predictive Maintenance: Systems, Purposes and Approaches,” *IEEE Communications Surveys & Tutorials*, arXiv: 1912.07383. (December). https://doi.org/10.48550/arXiv.1912.07383.

28. Microsoft n.d. “Connected factory signal processing pipeline.” Azure. Accessed October 25, 2022. https://learn.microsoft.com/en-us/azure/architecture/example-scenario/iot/connected-factory-signal-pipeline.

29. Scikit-learn n.d. “sklearn.feature_selection.RFECV.” scikit-learn. Accessed October 28, 2022. https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFECV.html.

30. Microsoft n.d. “How to select algorithms for Azure Machine Learning.” Azure. Accessed October 28, 2022. https://learn.microsoft.com/en-us/azure/machine-learning/how-to-select-algorithms.

31. McMahan, H. B., E. Moore, D. Ramage, S. Hampson, B. Aguera y Arcas. 2017. “Communication-Efficient Learning of Deep Networks from Decentralized Data.” In Artificial intelligence and statistics, arXiv: 1602.05629v3 [cs.LG], pp. 1273-1282. PMLR. https://doi.org/10.48550/arXiv.1602.05629.

32. Bemani, A., N. Björsell. 2022. “Aggregation Strategy on Federated Machine Learning Algorithm for Collaborative Predictive Maintenance.” Sensors (Basel, Switzerland), vol. 22 no 16, 6252. https://doi.org/10.3390/s22166252.

33. Yang, Q., Y. Liu, Y. Cheng, Y. Kang, T. Chen, H. Yu. (2020). “Federated learning.” Morgan & Claypool Publishers. https://doi.org/10.2200/S00960ED2V01Y201910AIM043.

34. Manjunatha, K. A., V. Agarwal, H. Palas. 2022. “Federated-Transfer Learning for Scalable Condition-based Monitoring of Nuclear Power Plant Components.” Idaho National Laboratory, Idaho Falls, ID, USA.

35. Sun, C. M. Ma, Z. Zhao, S. Tian, R. Yan, X. Chen. 2019. “Deep Transfer Learning Based on Sparse Autoencoder for Remaining Useful Life Prediction of Tool in Manufacturing.” in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2416-2425, April 2019. https://doi.org/10.1109/TII.2018.2881543.

36. An, Qinglong & Tao, Zhengrui & Xu, Xingwei & EL Mansori, Mohamed & Chen, Ming. 2019. "A Data-driven Model for Milling Tool Remaining Useful Life Prediction with Convolutional and Stacked LSTM Network." https://doi.org/10.1016/j.measurement.2019.107461.
37. The Office of Management and Budget (OMB) Circular No. A-130 Revised, "Memorandum for Heads of Executive Departments and Agencies (Transmittal Memorandum No. 2)," 1994.
38. The Office of Management and Budget (OMB) Circular No. A-130 Revised, "Memorandum for Heads of Executive Departments and Agencies (Transmittal Memorandum No. 3)," February 8, 1996. http://www.whitehouse.gov/omb/rewrite/circulars/a130/a130.html.
39. The Office of Management and Budget (OMB) Circular No. A-130 Revised, "Memorandum for Heads of Executive Departments and Agencies (Transmittal Memorandum No. 4)," November 28, 2000. https://obamawhitehouse.archives.gov/omb/circulars_a130_a130trans4.
40. "Securing Cloud Services for the Federal Government." https://www.fedramp.gov/.
41. The Office of Management and Budget (OMB) Circular No. A-130 Revised, "Memorandum for Heads of Executive Departments and Agencies (Transmittal Memorandum No. 4," July 27, 2016. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf.
42. Federal Register, vol. 81, no. 145. Thursday, July 28, 2016 / Notices, 81 FR 49689). https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf.
43. "Federal Cloud Computing Strategy." https://cloud.cio.gov/strategy/.
44. "Federal Perceptions of Cloud Security." https://cdn.fedscoop.com/federal-perceptions-of-cloud-security-report.pdf.
45. "Federal agencies reach new threshold of cloud adoption." May 31, 2022. https://www.fedscoop.com/federal-agencies-reach-new-threshold-of-cloud-adoption/.
46. SECY-07-0135, "Status Of Staff Information Technology/Information Management And Business Process Activities In Preparation For The High-Level Waste Repository Proceedings," August 10, 2007. https://www.nrc.gov/docs/ML0718/ML071800367.pdf.
47. Jayapal, P., S. Schrader, "NRC Cloud Infrastructure," Data Science and Artificial Intelligence Regulatory Applications Workshops, Workshop 2: Current Topics, August 18, 2021. (NRC ADAMS Accession No. ML21277A144). https://www.nrc.gov/public-involve/conference-symposia/data-science-ai-reg-workshops.html.
48. "Reconstitution/Replacement Options for the Licensing Support Network (Rev 4)," February 22, 2018. (NRC ADAMS Accession No. ML17347B671). https://www.nrc.gov/docs/ML1734/ML17347B671.pdf.
49. Official Transcript of Proceedings NUCLEAR REGULATORY COMMISSION, Licensing Support Network Advisory Review Panel, Rockville, Maryland, February 28, 2018. https://www.nrc.gov/docs/ML1806/ML18067A313.pdf.
50. MD 12.5, NRC Cybersecurity Program, DT-20-11, October 1, 2020. (NRC ADAMS Accession No. ML20258A098). https://www.nrc.gov/docs/ML2025/ML20258A098.pdf.
51. Jayapal, P., S. Schrader, "NRC Cloud Infrastructure," Data Science and Artificial Intelligence Regulatory Applications Workshops, Workshop 2: Current Topics Opening Remarks, August 18, 2021. (NRC ADAMS Accession No. ML21277A144). https://www.nrc.gov/public-involve/conference-symposia/data-science-ai-reg-workshops.html.

52. 10 CFR Part 73, “Physical Protection of Plants and Materials,” U.S. Nuclear Regulatory Commission, Washington, DC. https://www.nrc.gov/reading-rm/doc-collections/cfr/part073/full-text.html.

53. Kesler, B., “The vulnerability of nuclear facilities to cyber attack,” Strategic Insights, Spring 2011. https://www.kesler.us/portfolio/SI-v10-i1_Kesler.pdf.

54. Risk Management Strategy, September 22, 2020. (NRC ADAMS Accession No. ML20266G443). https://www.nrc.gov/docs/ML2026/ML20266G443.pdf.

55. National Control Devices. n.d. “NCD Store.” NCD. Accessed December 7, 2022. https://www.ncd.io

56. PCB PIEZOTRONICS. n.d. “PCB PIEZOTRONICS.” PCB PIEZOTRONICS. Accessed December 7, 2022. https://www.pcb.com

57. U.S.NRC. n.d. “Donald C. Cook Nuclear Plant, Unit 2.” U.S.NRC. Accessed December 7, 2022. https://www.nrc.gov/info-finder/reactors/cook2.html

58. Google Maps. n.d. “D.C. Cook Nuclear Power Plant.” Google Maps. Accessed December 7, 2022. https://goo.gl/maps/Td9Y7fb6Uve8rK349

59. WilsonPro n.d. “How much does a DAS system cost.” Accessed November 1, 2022. https://www.wilsonpro.com/blog/how-much-does-a-das-system-cost

60. Microsoft. n.d. “Pricing Calculator.” Azure. Accessed November 20, 2022. https://azure.microsoft.com/pricing/calculator

61. Impinj. n.d. “Impinj Speedway RAIN RFID Readers for Flexible Solution Development.” Impinj. Accessed December 7, 2022. https://www.impinj.com/products/readers/impinj-speedway

62. RFMAX. n.d. “UHF RFID (902-928 MHz).” RFMAX. Accessed December 7, 2022. https://www.rfmax.com/collections/uhf-rfid-902-928-mhz

63. Vulcan RFID. n.d. “Forge Innovation Through RFID.” Vulcan RFID. Accessed December 7, 2022. https://www.vulcanrfid.com/