Solid Freeform Fabrication 2023: Proceedings of the 34th Annual International
Solid Freeform Fabrication Symposium – An Additive Manufacturing Conference

Development of a Testbench for Additive Manufacturing Data Integration, Management, and Analytics

Chen-Wei Yang[1], Alexander Kuan[1], Sheng-Yen Li[1], Yan Lu[1], Jaehyuk Kim[1], Fan-Tien Cheng[2], Haw-Ching Yang[3]

[1]Systems Integration Division, National Institute of Standards and Technology (NIST)
[2]Institute of Manufacturing Information and Systems, National Cheng Kung University (NCKU)
[3] Department of Electrical Engineering, National Kaohsiung University of Science and Technology (NKUST)

Abstract

The NIST Additive Manufacturing (AM) Data Integration Testbench is a platform designed to evaluate data models, communication methods, and data analytics for AM industrialization. This paper describes a reference framework for AM data integration, named AMIF, and the design of the testbench based on AM Integration Framework (AMIF) for testing the integration of in-process data acquisition, real-time feature extraction, process control, and predictive models under a data management system. A specification of this testbench is developed to manage and stream voluminous data captured by high-speed cameras and performing data analytics using common information models and functional interfaces. The integration of the data, models, and computer tools sends operational decisions to an AM machine in real time. On top of the real-time control functions, AM data integration with MES and ERP systems is also included using a high-performance data warehouse for long-term data archiving and metadata management. The architecture of this testbench is illustrated in this work. AMIF can guide AM practitioners and system integrators to build their integrated AM manufacturing systems for production. The NIST AM testbench's plug-and-play features allow both internal and external researchers and developers to assess the effectiveness of their individual data models, data analytics, and decision-making algorithms on the systems engineering level.

1. Introduction

Additive Manufacturing (AM) demonstrates its capability of building complex geometry and customized products. AM is a strong contender to the current manufacturing methods with many successful applications, including fabrication of biomedical implants and aerospace engine parts [1]. Laser Powder Bed Fusion (LPBF) is a kind of metallic AM technology that employs lasers to fuse layers of metallic powder at specific energy levels and speeds. This technique builds near-net-shape parts using a layer-by-layer approach.

LPBF can significantly reduce processing steps to build complex parts; however, a long list of variables may affect the quality of AM parts, such as process parameters, environmental conditions, and material status. For example, the microstructure and mechanical features of the AM build parts are deeply dependent upon the process parameters, such as laser power, and scan speed. By fine-tuning the process parameters, the mechanical characteristics of the building samples can be enhanced [2]; however, when multiple build samples are created using the same process parameter conditions, the quality of each finished sample is dissimilar with each other. This result can be due to different factors, such as coating homogeneity, airflow stability, and fusion capability [1]. Also, building a delicate AM part usually takes a long time to complete the process. Evaluating the quality of the building process in real-time can assist the in-process decision-making and improve productivity.

To better evaluate an LPBF process, the in-situ measurements can provide rich information in different length and time scales. Different types of sensors are added to AM machines and these sensors generate large volumes of structured or unstructured measurement data at high frequencies. Successful integration of data and their analysis will enable effective in-process monitoring and real-time control [3].

At the same time, the maturation of AM into broad production and industrialization requires an expanded notion of integration for both AM systems and AM data. System and data integration includes other types of traditional manufacturing systems, manufacturing operation functions, and broader business processes across AM value chains, which are defined by ISA 95 Level 3 and 4 functions. Figure 1 shows the function hierarchy of manufacturing systems defined in ISA 95.

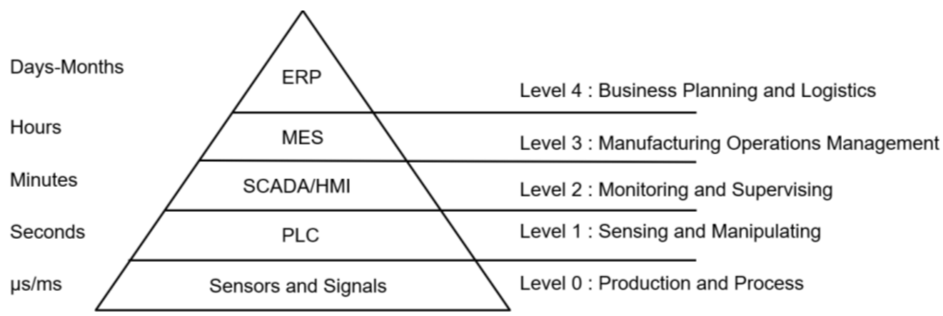

Figure 1. ISA 95 Standard [4]

ISA 95 is an international standard from the International Society of Automation for developing integration interfaces between enterprise and control systems. Level 0 defines the physical production processes. Level 1 defines the activities included in sensing and manipulating the physical processes. Level 2 defines the activities of monitoring, supervising, and controlling the physical processes. Level 3 defines the activities of the workflow, such as the process of manufacturing products from raw materials to finished products. Level 4 defines the business-related activities required for Level 3.

For AM process manufacturing systems, integration is needed to automate engineering workflow and improve decision-making across all the function levels of ISA 95 covering a complex set of engineering domains, such as design, material, process, machine, and manufacturing. However, the ISA 95 function architecture is not a sufficient guide to model the system and data integration for AM industrialization. The desired system requires high-speed data processing and integration for advanced data analytics, including the use of machine learning or deep learning for process planning, control, and part qualification. Figure 2 captures the system integration scope for AM and highlights its specific need for data integration (in blue fonts).

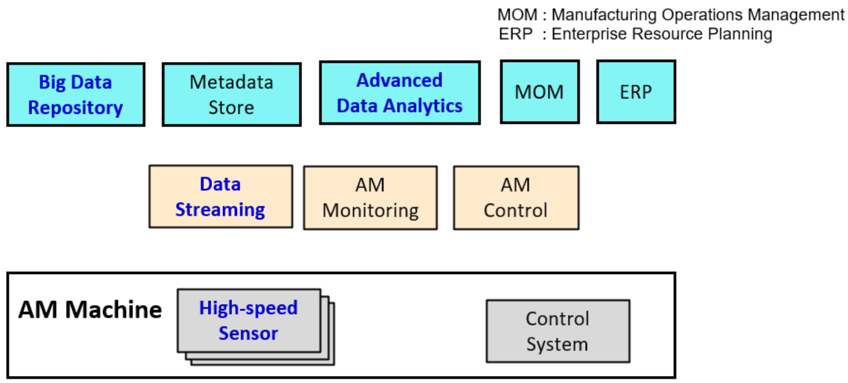

Figure 2. AM Specific Integration Requirements

In this research, we developed an AM integration framework that leverages ISA 95 addressing the specific high-speed data integration and analytics needs. Based on this reference architecture, we designed an AM integration testbench that can be used to test various communication protocols, information models, and advanced AI functions for AM development production and deployment. This testbench covers data collection, management, and analysis to automate offline engineering functions and data flow in real-time or near real-time. The testbench is designed for both internal and external use.

This paper is organized in three sections. In section 2, we describe the AM Integration Framework (AMIF) as a reference architecture for any modern manufacturing systems including AM processes. In section 3, we present an implementation architecture for the NIST AM integration testbench, which employs an AM emulator instead of a real AM machine. We also listed some potential communication protocols, information models, and advanced analytical functions. In section 4, we explain the use of AM big data storage and display the test results for determining the current data streaming capabilities.

2. AMIF Leverage on ISA 95

We leverage ISA 95 to define an AMIF for specifying the common AM data exchange scenario between various functions or applications to offer a generic reference framework for AM data integration.

The AMIF we propose is to leverage the ISA 95 function architecture but only define three different levels for integration purposes**.** In Figure 3, the AMIF displays the process relationship of how data, information, and events work between each function. **Level 1** defines the main functions provided to AM machines, which corresponds to ISA 95 Level 0 and 1. Between the AMIF Levels 1 and 2, we have a stage called **“real-time” monitoring & control**. This stage is necessary to support third-party advanced monitoring add-ons and functions faster than edge computing can offer; for example, FPGA is required to support this kind of high processing speed. **Level 2** defines the activities including near real-time monitoring and control, which corresponds to the ISA 95 Level 2. This level relies mostly on edge computing. **Level 3** defines activities, including data management, manufacturing operation, and enterprise application, which corresponds to ISA 95 Levels 3 and 4.

2.1 AMIF Diagram

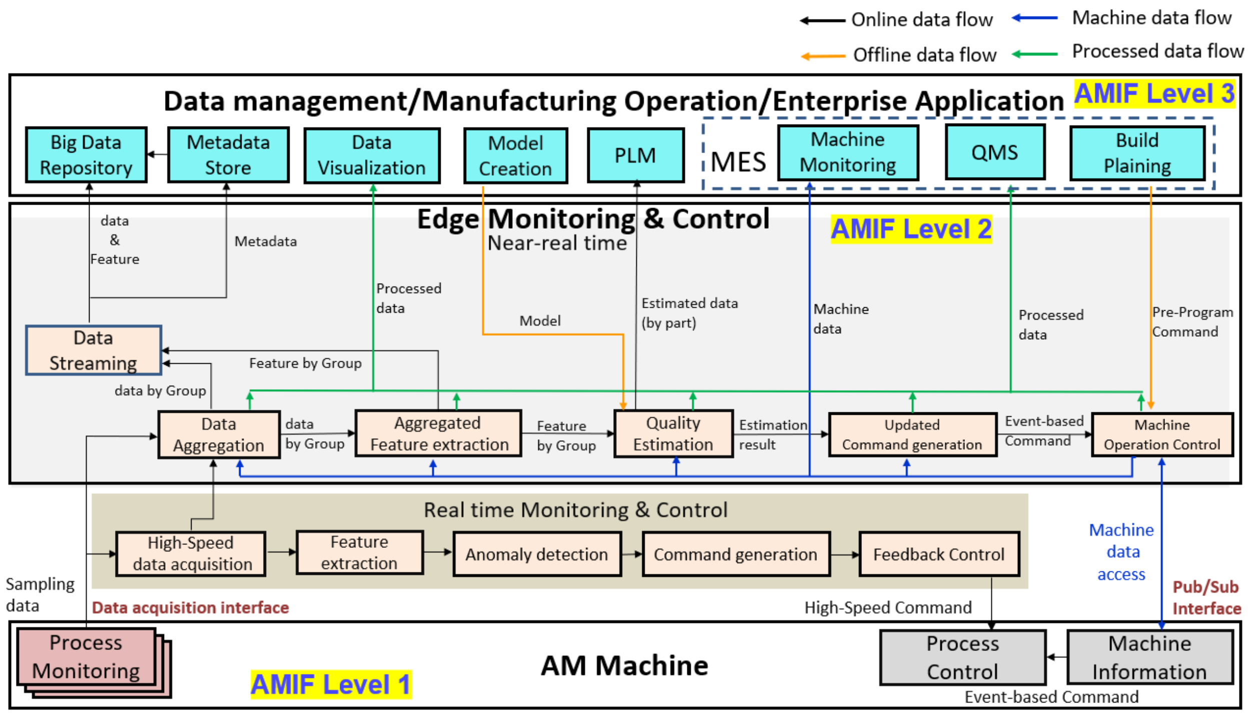

Figure 3. AMIF Diagram

2.1.1 Functions in AMIF Level 1 - AM Machine/Emulator

The AMIF Level 1 captures common functions the AM machines are designed with. Process Monitoring is for monitoring the process with high and low speed to collect the raw data for process analysis. **Process Control** is for controlling machine operation or monitoring devices with control systems based on the original process design command or command generation function during feedback control. **Machine Information** is for outputting or updating the information of the current machine status (Event-based). Each edge-calculation function should activate according to the current process status, so an operating system dedicated to function communication must be installed for near real-time feedback control.

2.1.2 Real-Time Monitoring & Control Stage

High-Speed Data Acquisition is needed for capturing and processing data from advanced in-situ monitoring systems in AM machines, facilitating both high-speed operation and real-time control. **Feature Extraction** is for analyzing the data from high-speed data acquisition and discovering useful features for anomaly detection. **Anomaly Detection** is based on the feature extraction results. This is used to detect abnormalities in each sampling data. **Command Generation** is based on the anomaly detection result. This is used to generate the command for high-speed feedback control. **Feedback Control** is for transmitting the command to the process control function located in the AM machine level for high-speed process compensation.

2.1.3 Functions in AMIF Level 2 – Near Real-Time Monitoring & Control

Data Aggregation is for carrying out corresponding data grouping according to different analysis requirements of scale. **Aggregated Feature Extraction** is based on the aggregated data performing feature calculation allowing the acquisition of useful features at a certain scale/resolution. **Data Streaming** is for uploading the aggregated data to an online data archive service located in the data management level**. Quality Estimation** is for using an AI/ML model for online quality estimation. The model is created offline in a model creation function located in the AMIF Level 3. **Updated Command Generation** is based on quality estimation results. This level generates the command for event-based feedback control. **Machine Operation Control** is for getting the AM machine information and controlling the process according to the different feedback control requirements. The function to update the machine information, located in the AM machine/emulator level, is based on the event-based command of feedback control**.**

2.1.4 Functions in AMIF Level 3 - Data Management/Manufacturing Operation/Enterprise Application

Big Data Repository is for archiving AM raw data. **Metadata Store** is for saving and managing the metadata of big data objects or building pedigree information. **Data Visualization** is for visualizing the processed data of functions in AMIF Level 2. **Model Creation Offline** is for creating the AI/ML model for online quality estimation, including feature selection, model training, and validation for different prediction targets or purposes. **Product Lifecycle Management (PLM)** is an information management system that integrates data, processes, and business systems. By estimating AM part quality, we can optimize the AM product geometry design. **Machine Monitoring** is for monitoring the machine information, including process operation, parameters, and sensors. **Quality Management System (QMS)** is a collection of business processes focused on consistently meeting customer requirements and enhancing their satisfaction. **Build Planning** uses data analysis results to optimize the building strategy and design. After build optimization, it generates the pre-program command.

3. Testbench Design for AM Operation and Control

The functional components in Figure 4 are essential to facilitate comprehensive testing of AM data integration scenarios. The NIST AM Data Integration Testbench implements the AMIF function architecture at three levels. However, instead of working directly on AM machines, which may result in the disruption of AM processes, the AM testbench at NIST utilizes an AM emulator. The **AM Emulator level** focuses on AM process event generation and sends the in-situ monitoring images out with a high-speed camera protocol. **The Edge Computing** level focuses on high-speed image acquisition and other necessary functions, such as streaming or decision-making. This level also implements a machine operation control. The **Cloud** level is for saving and managing raw data and metadata. These data can be used by manufacturing operations and management functions or enterprise functions for decision-making. In this section, we will introduce the main test modules for different levels of data integration for AM operation and control.

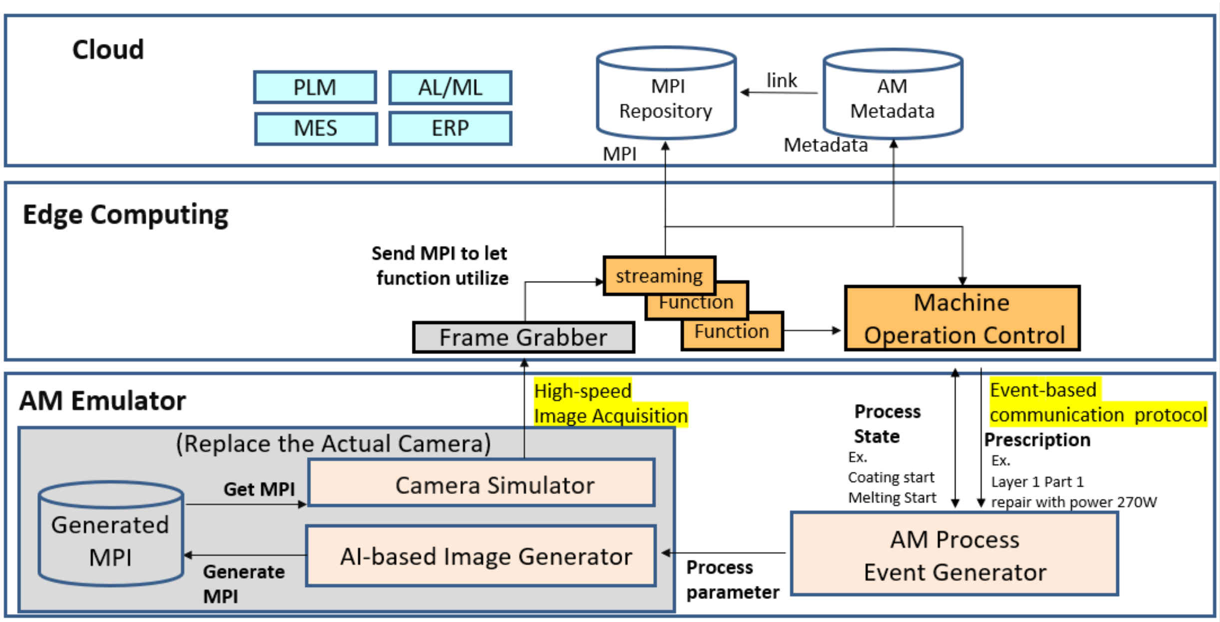

Figure 4. Testbench Design to Optimize AM Data Integration Performance

3.1 AM Process Event Generator

The event generator is for simulating the occurrence of events and utilizing a publish/subscribe interface to communicate with the functions of AM operation control. A working AM operation control follows the AM process step allowing the function in edge computing to work at the right time. If the decision-making needs to do a certain AM process for feedback control, the system utilizes this event generator to generate that event.

For example, when getting the ending event of the powder coating in Figure 5, this mechanism will inform the global camera to take the global image. After the function analyzes the image, a decision is sent back to the event generator to decide if we need to perform a recoat to control the coating quality. After the laser melting process, we estimate if the surface roughness is under the desired specification and send the decision to the event generator for the next step [1][5].

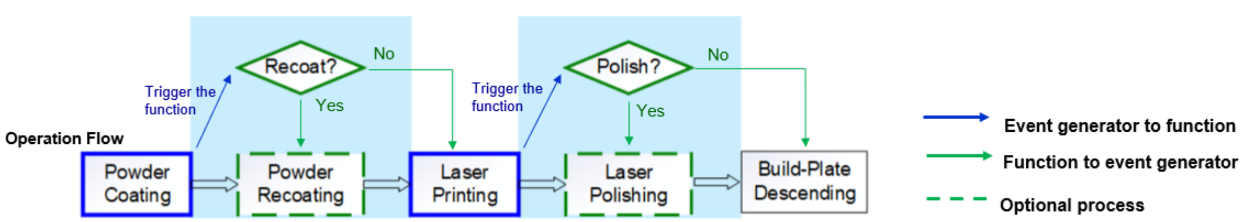

Figure 5. AM with Feedback Control

3.2 AI-Based Image Generator

Figure 6 illustrates the input and output of the AI-based image generator. The generator takes one-dimensional process parameters, such as laser power, scan speed, and direction along with random noise as its inputs, to produce the corresponding melt-pool image (MPI). This generator is built upon deep learning generative models, including generative adversarial networks and the Diffusion Model. Within the workflow of the AM emulator, the image generator emulates the MPI based on prescriptions of the AM process event generator.

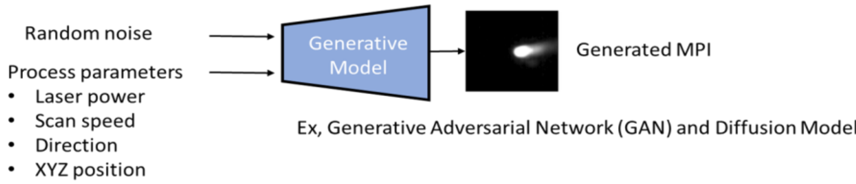

Figure 6. AI-Based Image Generation

3.3 Camera Simulator with High-Speed Communication Protocol

The purpose of using a high-speed camera in AM is to study material behavior, process monitoring, and quality control. By recording the process in real-time at high frame rates, manufacturers can closely observe potential defects and perform in-process feedback control.

The camera simulator can generate a high-speed video stream or test patterns for testing frame grabbers or vision/imaging systems. It allows developers to quickly prototype their vision and imaging systems without the need for physical cameras. It also enables the testbench to evaluate the performance of the high-speed camera interface, such as CoaXPress and CameraLink, in transmitting image data. CoaXPress reaches 50 Gbps by using four lanes which enables a higher data rate than CameraLink. Table 1 shows the comparison between CoaXPress and CamerLink [3].

If the camera requirement generates less than 125 MB/s (1 Gbps), both Gig-E and USB3 Vision can be used. For example, each grayscale image is 120 by 120 pixels, and the bit depth is 8 bits. To capture 10,000 images per second, the data rate needs to achieve 137 MB/s (1.07Gbps) [6][7].

Table 1. Comparison between CoaXPress and CamerLink [6][7]

Interface	CoaXPress	Camera link
Connector	BNC	MDR or SDR
Cable Length	30 to 212 meters	1 to 10 meters

Data Rate	12.5 Gbps per lane (CoaXPress 2.0). Uses four lanes which can reach 50 Gbps.	255 Mbps to 850 Mbps, depending on the specific implementation mode. (Base, Medium, Full, Extended-Full)
Real Time Trigger	Yes	Yes

3.4 Event-Based Communication Protocol

Machine monitoring is crucial for a variety of applications. Preventive maintenance, optimizing productivity, quality estimation, or feedback control all rely on process and sensor data transmitted by machines. OPC UA and MQTT are common communication protocols used in industrial settings for data exchange and interoperability. Table 2 shows the comparison between them. If the message includes structure data, the OPC UA is more suitable. If the message is similar to a trigger signal, MQTT is the better choice for the lightweight case.

Table 2. Comparison between OPC UA and MQTT

Protocol	**OPC UA**	**MQTT**
Use case	Provide security and reliability. Widely used in complex industrial automation, PLC communication, etc.	Lightweight messaging protocol. For resource-constrained environments, like IoT devices, telemetry applications, or various low-bandwidth conditions.
Communication Models	Client-Server Architecture Publish-Subscribe mechanism. (Introduced in OPC UA 1.04).	Publish-Subscribe mechanism.
Overhead	Higher overhead	Lower overhead
Read/Write Operations	Read and write	Read and write

3.5 AM Data Streaming and Management

Automating the process of uploading data to the cloud in real time during AM process is important. We can utilize the cloud services to set up cloud storage for data integration. NIST has already developed a collaborative AM data management system, providing data storage in clouds structured by an AM lifecycle data schema [8]. In this data integration testbench, we will also build a similar AM data management system to simulate the data streaming and archiving application as well as explore the different methods of storage and evaluate their capabilities.

The process of AM produces large amounts of data for long-term archiving and future data analysis. AM big data may include CAD model files and in-process melt-pool images. The new framework desires two different types of storage systems for archiving AM big data and metadata separately. Currently, the Additive Manufacturing Materials Database (AMMD) [8], a NoSQL MongoDB document data warehouse developed by NIST, is the data storage used to store structured and unstructured AM metadata. However, there are limitations with the current storage system that may affect the potential speed and performance of the AM testbench.

3.5.1 AM Metadata Store

As the quantity of generated AM data grows, searching for a specific dataset in a big data repository becomes increasingly difficult. This system requires cooperation with a database capable of holding structured metadata that provides details for data objects stored in the cloud. An acceptable option for a metadata store would be a document database, as it allows highly structured documented data architecture for immutable AM metadata. Another option worth exploring is a graph database, which focuses on graphing node entities and edge relationships.

3.5.1.1 Document Storage

The current AMMD [8] storage system stores AM metadata in XML format and follows an XML Schema Definition (XSD) data structure which partitions XML documents into 9 types. Each document type has its own validated structure layout and are linked together by identification values. These databases can handle large quantities of data storage and are flexible to small changes to the schema structure. What makes the system limited is the difficulty in establishing complex relationships between metadata located in different XML documents. Each separate document establishes a connection by the document's identifier; however, the relationships between subsets of metadata within each document are not clearly specified. This makes data query a slightly more tedious task.

3.5.1.2 Graph Storage

Graph databases serve to be an efficient AM data analysis tool with easily navigable node and edge relationships. In comparison to document-based databasing, the graph makes use of relationships between node entities as the primary way to draw connections between data in the entire graph view. This method is great for data analytics because it consolidates data, enabling quicker traversal between the node entities and edge relationships; however, the database's performance is reliant on a well-defined data ontology. An unnecessary or undefined relationship in the ontology may result in a separation of important information between two related nodes.

Much remains unknown about integrating AM data with graph databases. We are currently investigating the metadata storage and query capabilities of Deep-Lynx [9], an open-source graph data warehouse developed by Idaho National Laboratory (INL), to obtain a closer understanding of the compatibility of node and edge architecture along with the AMIF.

4. AM Big Data Streaming and Archiving

To determine if a storage system fits our criteria, we emulate the AM data generation and streaming processes to test a database's features and capabilities. The original system utilizes batch uploads to archive all types of AM data, but this method is an inefficient use of time. By streaming AM data during the AM process, we can improve the speed of archiving data, making data readily available for analysis at any time.

4.1 Utilize S3 Cloud Storage

Streaming and archiving AM big data requires a high-volume file storage system compatible with fast and continuous upload. Amazon Simple Storage Service (S3) is the selected cloud object data lake storage which offers scalable file storage, configurable security access, and several API software development kits. The concern of using this service is the lack of understanding about the S3 capabilities and internet data transfer rate. A vital component of near real-time data streaming is a consistent high-speed connection between local AM machines and cloud services. Slower connections result in a growing backlog of data queued for upload. To ensure timely processes, an AM in-process simulation is constructed to emulate the procedure of generating image data, compressing each package, and streaming the data to the designated S3 bucket destination.

4.2 Upload Speed Test Dataset

The AM in-process simulation consists of artificial parameters representing the desired and testing upload speed configurations. The build is a 12-part model each constructed with 250 layers. One package contains sample image data for each part of each layer resulting in a total of 3000 packages produced per test run. Regarding the internet connection, the simulation uses an RJ45 ethernet cable to connect local AM processors to the AWS cloud services. The simulation compresses each package into a ZIP file since compressed data results in a quicker internet transfer rate. In Table 3, if we divide the data size per package by the compressed data size per package, the compressed package is estimated to be ¼ of the full package size.

Table 3. Upload Speed Test Dataset Base Parameters

Configuration	Data
# of Layers	250
# of Parts	12
Total # of Packages	3000
# of Images per Package	1500
Single Image Size	43.3 KB
Data Size per Package	**64.881 MB**
Compressed Data Size per Package (ZIP)	**16.92 MB**

The setting for the camera frame rate is around 4000 FPS. During the time frame of layer scanning and recoating, image data is captured and streamed to the cloud.

Table 4. Upload Speed Test Scan and Recoat Time Parameters

Configuration	Data
Scan Time per Layer	4.8 seconds
Recoat Time per Layer	5 seconds
Total Time for Minimum Buffer	9.8 seconds

4.3 Desired Speed Calculation

Estimating the minimum upload speed requirement involves creating mock configuration parameters that imitate an actual AM process. Results of the upload speed simulation that fall short of the minimum requirement could end up with an undesirable stockpile of data files queued for cloud transfer.

Table 5. Minimum Desired Upload Speed Base Parameters

Configuration	Data
# of Layers	250
# of Parts	12
Data Size per Package	**60 MB**
Scan Time per Layer	4.8 seconds
Recoat Time per Layer	5 seconds

In Table 6, the calculated desired upload speed for one package is obtained by dividing data size per layer by total time per layer. The actual and compressed data size from the upload speed testing dataset is rounded to a ¼ compression ratio, so the estimated desired upload speed for the compressed package is acquired by dividing the desired upload speed by 4.

Table 6. Minimum Desired Upload Speed Calculated

Configuration	Data
Total Time per Layer	9.8 seconds
Data Size per Layer	720 MB
Desired Upload Speed	**73.469 MB/s**
Desired Compressed Upload Speed (ZIP)	**18.367 MB/s**

4.4 Upload Speed Test Results

The test results will be compared to the estimated desired upload speed. If the upload speed test results exceed the desired speed, we can say the AM data transfer to the S3 bucket will be able to keep pace with AM machine data generation. In Table 7, the average data upload speed displays the theoretical transfer speed of the data size within a compressed package, not the upload speed of each individual data file.

Table 7. Local to S3 Upload Speed Test Results

Configuration	Data
Avg ZIP Upload Time per Package	0.446 seconds
Avg ZIP Upload Speed	37.937 MB/s
Avg Data Upload Speed	145.459 MB/s

After multiple simulations, the accumulated data averaged a ZIP package transfer rate of 37.937 MB/s meaning the data inside each package uploaded at a speed of around 145.459 MB/s. The test results more than doubled the minimum required upload speed.

Table 8. Comparison between Desired and Test Result Upload Speed

Upload Speed	Desired	Test
Compressed Package (ZIP)	18.367 MB/s	37.937 MB/s
Data inside Package	73.469 MB/s	145.459 MB/s

It is important to note that the upload speed test simulation does not contain each layer's time configurations for laser printing, powder filling, build plate descent, and other factors. Theoretically, adding these time parameters would significantly lower the estimated desired upload speed making the current simulation test results even more qualified for near real-time AM data streaming.

4.5 Combining Image and Metadata Streaming Processes

The AM data streaming simulation runs simultaneously with each other. When generated, data is recognized and uploaded to the specified endpoint. What makes this process problematic is how both streaming methods are run asynchronously from each other. This means data in Amazon S3 is not linked with metadata within the graph database. A potential solution could be to send an identifier metadata value within the S3 upload response back to the local computer. Then, a separate process will send an API request containing the S3 object identifier to the graph database and update the corresponding node metadata in the graph view.

5. Summary

AMIF is designed as a framework for AM system developers to integrate process data and sensor signals. The NIST AM data integration testbench is a use case that emulates a real AM system process for testing the interoperability of computer tools and evaluating the machine operation control. The testbench also demonstrates its capability to handle high-speed, high-volume in-process data in an autonomous workflow. This testbench can serve as a standard platform to test the integration of data flow and the pipeline of software operations. The goal is to integrate industrial cases on the testbench to evaluate the various communication protocols, information models, and numerical simulations. Also, utilizing this testbench can enable testing advanced LPBF process monitoring and predictive algorithms for near real-time or real-time data analysis and process control.

Disclaimer and Acknowledgement

Certain commercial equipment, instruments, or materials identified in this paper are not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

The authors would like to thank Dr. Ho Yeung for the support and information sharing for the testbench development. We also feel thankful for the support and the insightful discussions provided by Dr. Paul Witherell and Dr. Brandon Lane at NIST.

Reference

[1] M. Adnan, H. -C. Yang, T. -H. Kuo, F. -T. Cheng and H. -C. Tran, "*MPI-Based System 2 for Determining LPBF Process Control Thresholds and Parameters*," in IEEE Robotics and Automation Letters, vol. 6, no. 4, pp. 6553-6560, Oct. 2021.

[2] G.Tapia, A. H. Elwany, and H. Sang, “*Prediction of porosity in metal-based additive manufacturing using spatial Gaussian process models*,” Additive Manuf., vol. 12, pp. 282–290, 2016.s

[3] Perisic, Milica, Yan Lu, and Albert Jones. "In-Process Data Integration for Laser Powder Bed Fusion Additive Manufacturing." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 86212. American Society of Mechanical Engineers, 2022.

[4] ISA95, Enterprise-Control System Integration, https://www.isa.org/isa95/, Accessed on May 20, 2020.

[5] Yang, Haw-Ching, et al. "An intelligent metrology architecture with AVM for metal additive manufacturing." *IEEE Robotics and Automation Letters* 4.3 (2019): 2886-2893.

[6] CoaXPress Standard, https://www.coaxpress.com/; [accessed on August 29, 2023].

[7] CameraLink HS Standard, https://www.automate.org/a3-content/vision-standards-camera-link-hs; https://knowledge.ni.com/KnowledgeArticleDetails?id=kA00Z000000P7rmSAC; [accessed on August 29, 2023].

[8] Lu, Yan, Paul Witherell, and Alkan Donmez. "A collaborative data management system for additive manufacturing." *International design engineering technical conferences and computers and information in engineering conference*. Vol. 58110. American Society of Mechanical Engineers, 2017.

[9] Deep-Lynx, https://github.com/idaholab/Deep-Lynx; [accessed on August 31, 2023].