Towards a Cloud-based Machine Learning for Health Monitoring and Fault Diagnosis

Samir Khan[1], Takehisa Yairi[1] and Mariam Kiran[2]

[1]*Department of Aeronautics and Astronautics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
khan@ailab.t.u-tokyo.ac.jp

[2] *Energy Sciences Network, Lawrence Berkeley National Laboratory, Berkeley CA 94720, USA*

ABSTRACT

Large complex engineered systems collect large amounts of varied data sets, making it often difficult to process and analyze these for diagnosing, isolating, and predicting faults during operation. To recognize symptoms with standard testing tools, infer potential faults and eventually diagnose causes needs constant maintenance support. This problem is particularly faced in the aerospace industry, where it is essential to analyze and maintain assets to prevent potential failures or loss both technological and human. Recent usage of Cloud computing provides infinite computing resources to quickly process and troubleshoot, reducing 'time-to-fix' problems. Exploiting artificial intelligence (AI) algorithms, with Cloud resources, can help build an integrated fault diagnostic platform to provide resilient and scalable resources for data acquisition, processing and decision making. This paper presents an industrial perspective and problems when using machine learning methods for fault diagnosis, particularly using Cloud resources in the aerospace industry. Special attention is paid to the benefits; with potential future research on technical diagnosis being enumerated.

1. INTRODUCTION

Faults detected in modern technological services and associated data processing can immensely impact the correct functioning of technical systems. Any disruption is of major concern, not only to system users, but also to manufacturers, suppliers, operators, and maintainers of the system. For safety critical applications, such as in the aerospace industry, any potential faults can have adverse effects on safety, operation and directly reduce the profitability of all elements of the value chain. Therefore, it is important to detect and resolve any such disruptions that can influence customer requirements (Womack et al, 2015). As the number of assets and novel technological innovations diffuse within the service industry, together with an aggressive operating environment; a variety of failure modes appear to manifest themselves (Cao et al 2012, Zio 2009 and Khan et al 2014a). Although failure modes can be diagnosed and isolated, the growing maintenance costs in today's engineering industry has prompted a need for further research in novel methods to reduce maintenance, repair, and overhaul of complex high value assets (Lightfoot et al, 2013). As a consequence, recent efforts are being concentrated on the integration of anomaly detection, diagnostics and prognostic technologies across systems and platforms. Such capabilities can help maintain system performance in a cost-effective manner, whilst identifying ongoing issues to mitigate potential risks, and providing data exchange (and processing) within diagnostic technologies as a high priority research topic. However, the size of data exchange has continued to increase, along with disparate information sources (Xu, 2012). Coupled with complexities of contextual components for correlating information, existing approaches appear limited to deal with issues of during the design phase of the system lifecycle. Furthermore, modern day system complexities bring various challenges of system health data storage, its availability, interpretability, interoperability, time to process and more. Considering higher levels of interdependencies between assets, it becomes difficult (if not impossible) to identify root causes of system failures and to make real time decisions to compensate for them. Yet, the underlying engineering environment is expected to support both the technological platforms as well as system availability requirements (Khan, 2015). Additionally, fault prediction algorithms for real-time data processing, visualization and high volume big data transfers bring unprecedented demands on underlying networks to support their high capacity and rapid provisioning requirements for distributed end-to-end connectivity. With increasingly complex and diverse large-scale applications, the computing infrastructure needs to quickly process and make intelligent predictions on the system health data.

The Cloud computing paradigm is a promising environment delivering IT-as-a-service for industries and researchers to deploy their applications. However, as an enabler, it has also led to increased challenges in Big Data and Internet of Things applications, as the need for diverse data volumes and devices grows (Buyya et al, 2009 and Foster et al, 2008). Research challenges in finding effective means on how to manage massive data volumes has had an impact on space availability and driving up the processing costs and maintenance of results. Particularly in the aerospace sector, data from long term archiving and retrieving processes are used in combination with computer aided design and computer oriented languages to help reduce and find the possibility of faults. Aerospace companies have been collecting health and usage monitoring, operational information, stock levels and supply chain data for a long time (Khan et al, 2014b). However, now with big data applications, more data patterns can help manage and efficiently collate data behavior, which was previously a challenging task due to lack of infrastructure or software services. This had led to a 40% increase in the job market from 2009 to 2011 (Foster et al, 2008).

This paper explores the use of cloud-based machine learning framework to allow a cost-effective intelligent fault monitoring systems for aerospace maintenance. We argue that a novel framework is required for reconfiguring applications, as well as mechanisms for making better decisions at the system-level (may this be achieved through partial or full autonomy). In the nominal environment, such problems require advanced capabilities to monitor in-service operations, record and share expert knowledge, and address critical aspects of on-board software. To highlight the importance of advanced intelligent decision-making, recent industry efforts, have begun investigating machine learning (ML) to improving and analyzing telemetry data (Kwon et al, 2016). In this context, the authors present their perspective on identifying open research problems in how ML efforts can help improve fault discovery for the aerospace industry.

The rest of the paper is structured as follows: Section 2 highlights some key industrial requirements in terms of maintenance of high value assets and the ever-increasing health management gap due to the nature of the problem. Section 3 presents an industrial perspective and advocates the benefits of cloud based solutions, followed by the need to use of artificial intelligence (AI) concepts for improvements. Section 4 discuss the cloud-based decision support system and notes the benefits and limitations of using such a platform for fault diagnosis purposes; followed by discussion and conclusions.

2. THEORETICAL BACKGROUND

2.1. Big Data Analytics

Data analytics are essential to plan and create decision support systems for optimizing underlying infrastructure. This involves not only processing of online real time data, in search for certain events, but also historical data sources, previously saved, to help find data patterns. Cloud providers are paramount for availability and durability for resources, where data is replicated across multiple servers in different geographical locations. Elasticity can help allocate more resources on-the-fly to handle increased demand. Big data processing gives companies a competitive advantage for efficient analysis and predicting costs through system-life (Yasumoto et al, 2016). This also brings new challenges to processing large amount of data in a bandwidth-limited, power-constraint, unstable and dynamic environment (Sharma et al, 2013). When failures are too complex to diagnose and isolate based on operating organizations. This allows using cloud platforms to carry out the analysis with the help of other participating organizations and maintain a knowledge based system. Further work used data processing toolkits to forecast and redistribute resources on the fly (Kiran et al, 2015). But there is still lack of research in providing multiple users from varying backgrounds to write and deploy optimized data processing applications is still needed. Tailored solutions for online and batch data processing can keep non-functional attributes such as cost and network complexities satisfied. Current industry focus of using Spark SQL have aided in faster processing counteracting the Hadoop processing model weaknesses.

Cloud computing provides stakeholders with means through which various applications requirements from computing resources, infrastructure, business processes, and related dependencies can be offered as a service. These services can be accessed from anywhere and be deployed wherever they are required. For health management and fault analysis, a cloud service can overcome limitations of handling large data sets, often located in various repositories. However, this introduced challenges of utilization, network bandwidth, resource provisioning, improving application-network interaction and performance characteristics. Furthermore, such technology can be used to produce models of the decision-making process. Traditionally, experts used process variables to make recommendations. However, with a cloud solution, stakeholders can collect communication, component information and data analysis tools, incorporate device status data, such as condition, performance, utilization, and degradation information, for decision making.

Such analytics are not just description of the data when it gets "big" (e.g. in terabytes a day). It comprises of many facets such as how to organize this data, how to label different kinds of it (structured, unstructured, semi-structured, internal and external), which technologies that are used to store it and retrieve it. Therefore, “Big Data” is broader concept then just data that happens to be big.

2.2. Cloud-based Machine Learning

As Big Data represents content, Cloud Computing is more concerned about the infrastructure. It is a paradigm for computing “on-the-fly”, where almost everything can be dematerialized with the user not having to worry a lot about the infrastructure and process optimizations. Since intelligence is a vast discipline with multiple algorithms, each can be optimized for particular application domains. Collection and analysis of health monitoring data and fault diagnosis can also consider several well researched intelligence methods; ranging from classical statistical methods – such as linear and logistic regression – to neural network and tree-based techniques[1]. Other options include making use of hybrid and adaptive systems, that make use of fuzzy controllers and network-based predictors. Exhaustive reviews on various developments can be found in published books by Pascual 2015 and Michalski et al, 2013. These techniques can be recognizing patterns, help cluster and classify data to extract features to perform regression (or reinforcement learning) for anomaly detection problems. In sudden fluctuating conditions, such as the ones considered in pervasive computing, systems are expected to adapt their behavioral models according to current conditions, anytime and everywhere, e.g., a mobile device can be context-aware if it can acquire, process and use this information depending on the operational functionality to the current context of use – like sending location, video feeds, away status message, etc. Such information is then openly shared over the network with other devices, thus improving not only their situational awareness but also providing opportunities to interact with other devices. However, the main issue with these systems rests with that way the information is presented and conveyed to others. Many applications have moved to the cloud for more efficient context awareness (Wan et al, 2014). This is due to the diversity in the types of information. Therefore, any fault diagnostic activity will require recognition and context awareness whilst being hosted on the cloud solution. Recent advances in parallelization using GPUs, virtual machine and containers have reduced these issues considerably (Navarro et al, 2014). But there is a need adapt these methods or cloud implementation with distributed data sources and to optimize network provisions. One option is the parallelization of workloads which can improve memory consumption and reduce the complexity of decision processing.

Using cloud based ML solution can help with a wide range of analytics, giving designers and maintainers a great opportunity to investigate the symptoms and possible root cases that lead to failure events. This also allows compensating for failures until the next maintenance activity takes place. Although geographically distributed solution can be applied to a variety of activities for maintenance technology development, currently their use is limited to design (Khan et al, 2015). As preventative maintenance strategies are becoming more mission-critical, the collections of condition monitoring and environmental data in faulty situations has become indispensable, and hence such solutions will have a wider industry application. These bring new challenges but do ultimately improve system reliability and safety. The use of AI methods has become increasingly extended, using them will enrich decision support through means as coordinating data delivery, analyzing data trends, prognosis, quantifying uncertainties, predictive user needs, presenting appropriate data forms and decision making. Even though current maintenance research is focusing on implementing such concepts on a local platform, a cloud-based implementation can provide more processing flexibility, increased system collaboration and overall improve competitiveness. Such concepts would move away from reactive maintenance concepts into more proactive practices, providing vital information on root causes, that are unavailable from traditional tests and aid in overall maintenance decision making. In general, this decade has seen an incredible level of investment to enable AI capabilities in cloud platforms – organizations such as Amazon, Google, Microsoft and IBM are at the forefront of many platform service (PaaS) solutions. Currently, there seem to be predominately two groups of cloud AI technologies:

- Cloud ML platforms such as Azure Machine Learning, AWS Machine Learning.
- Google Cloud ML which helps to machine learning models using a specific technology: Also supports open source libraries such as TensorFlow (Abadi et al, 2015), but not other frameworks such as Theano, Torch and Caffe.

Therefore, there is a need to develop interfacing mechanisms to integrate AI capabilities without having to invest in sophisticated AI infrastructures. As AI technologies evolves, cloud platforms should shift from this level of basic support for AI capabilities to a model in which AI programs are as widely supported as web and databases are today. This is not only required for health management and fault diagnosis, but in all technological disciplines.

2.3. Intelligent Fault diagnosis

Machine learning algorithms can be used to predict behavior such as ‘which component fault will cause what failure X with what probability P’. Detecting anomalies cuts down costs and troubleshooting time in complex infrastructures. Also, actively predicting failures enables engineers to anticipate and proactively perform better maintenance scheduling. Making these decisions in real-time requires massive data processing power and time, such as to digest all relevant datasets to recognize multiple assets that affect the system. Various intelligence methods can be used to classify specific activities, the nature of the fault (i.e. soft, hard,

[1] e.g., these include feed-forward networks, such as multilayer perception, Radial-Basis Function networks, Self-Organizing Map, Kohonen-Networks or Support Vector Machines.

incipient, low, high) and detect critical situations or failures. However, these algorithms often come with a heavy price on computation and memory requirements to evaluate models during test and run-time. Consider the following scenarios of unsuccessful fault diagnosis during component repairs (Khan et al 2014a):

- A fault cannot be reproduced with real conditions: the fault is considered as a one off and the system is declared serviceable. The fault will reappear later because the origin has not been identified.
- The maintainer decides to replace a component because they consider that it is the root-cause of a fault. After a few tests on the new component, the system is declared serviceable. Nevertheless, the fault still reappears after a while, thus the cause has not been clearly identified.
- The same fault reoccurs, but this time in another part of the system.

Repair challenges require assets to be operationally available for a maximum amount of time. If they are undergoing maintenance, they are not earning revenue and are consuming resources such as spares and man-hours. This introduces pressure on the aerospace technicians and maintenance operators to be as efficient and effective in delivering best possible availability and operational performance. To meet these demands, several possible suspect units might be replaced in order to ensure that the fault has been removed, although only one of the units may be faulty but the several units will now need to be bench tested or repaired with the obvious associated costs.

Industries are using big data analytics to help predict faults before they compromise the systems. Due to the high cost of the equipment maintenance due to their complexity, it is necessary to 'simplify' modern maintenance management systems. The conventional condition-based maintenance (CBM) to reduce maintenance activities and operate according to the indication of an equipment condition. Khan (2015) demonstrated that the major problems facing modern aerospace engineering are high inventory cost for spare parts, pre-planning maintenance work for complex equipment under a complex environment and avoiding the risk of major failure and eliminating unforeseen circumstances of equipment or systems. Analyzing raw data, subcomponent behaviors and fault detection techniques the process can be automated reducing the cost and improving machine performance.

3. AN INDUSTRIAL PERSPECTIVE

For safety-critical applications, there is a need to implement an effective fault monitoring system to collect (relevant) data from various sensor sources and carry out the necessary signal processing including the extraction of key features, fault diagnosis and prediction. Based on this analysis, the system capable to recommend further actions according to user requirements. This phase plays an important role in adding resilience to the overall setup and for regulating availability during service operation. During diagnosis, a number of recommended actions might be issued including fault alarms, alternatives to maintain availability, in-service feedback, etc. Depending on the recommendation, the human operator may either choose to delay any action – if the failure can be tolerated until the next scheduled maintenance, or take an immediate action e.g. in the case of failures that can affect safety.

Traditionally, visual inspection routines were carried out during scheduled maintenance. These practices relied heavily on expert knowledge and experience of the maintenance personnel (Khan, 2015). However, with the drive towards industry 4.0 concepts, information systems such as internet of things and cloud computing, have become instrumental technologies for enabling improved system performance and resilience (Lee et al, 2014, Lee et al, 2015 and Jazdi, 2014). Yet, no matter how well a maintenance system is designed, there is always the possibility that it will contain deficiencies (due to decisions and trade-offs in design) that can lead to difficulties in the quality of maintenance in service. Also, most fault diagnosis systems operate independently for each other and not sharing any information[2].

Fortunately, some technologies such as Web Service Description Languages (WSDL), ontologies, Service Oriented architecture languages (SOA), have been developed to enable interoperability and knowledge sharing (Jung, 2011). The ultimate responsibility for recognizing, interpreting, and compensating for deficiencies in the diagnosis process, rests with human maintainers (Campbell and Reyes-Picknell 2015). These maintainers are fallible and arising subsequent issues[3] have been shown to be statistically significant, i.e. "they do not get it right all of the time". Considering the size of assets in modern industrial domains, even trying to understand the physical behavior of these 'large' systems, it seems unrealistic to believe that ubiquitous and integrated system level decisions can be made. Especially, when the operating conditions, and even the maintenance environment, are always subjected to unpredictable fluctuations which can have unforeseen consequences.

[2] It is therefore necessary when any maintenance system is designed, and before it becomes operational, to thoroughly test it, in order to identify any potential problems.

[3] Such as poor design of human tasks, poorly perceived maintenance operating procedures and inadequate training, as well as the pressures of the job.

As a result, there is a lot of emphasis on collecting vast amounts of maintenance datasets – which is expected to be stored, processed and optimize the operation. It should be noted that these datasets are useless unless some meaning for analytics can be extracted from them. Therefore, many organizations aim to invest for processing large datasets (also termed as big-data) for fault diagnostic purposes to gain better understanding of their systems failures and how to deal with them (Chen and Zhang, 2014). This indicates the development of online technology which allows to send and receive data, whilst building an ecosystem. Here, the aim is to continuously feed performance data in order to build an ecosystem with built-in sensors, diagnostic and prognostic monitoring software routines which would train themselves depending on their operating and performance data to the original equipment manufacturers (OEM) central data warehouse for processing. For example, it is possible to collect flight data (and related statistics) from an aircraft and transmit it via satellite networks to a central data repository (Cope and Kaufman, 2003). In this way, a manufacturer will be able to predict asset failures by recognizing any early indications of incipient faults and help in maintenance scheduling. The overall process hence can be optimized to reduce warranty costs, maintain spare parts and system availability, and fulfil stakeholder requirements. E.g. such information can be utilized to identify individual components' performance from the rest of the fleet. This can help to quantify a components' remaining useful life or its reliability going forward. Therefore, data plays a significant role in influencing the next generation of products by identifying existing issues with current implementation across the fleet, feeding this information back to design to improve quality.

Another requirement is concerned with the fact that assets might be located in various geographical locations around the world, which may often communication with each other e.g. aircrafts. Each asset would produce condition monitoring data depending on its use, environmental conditions and predefined user requirements. Depending on the communications architecture, a centralized system will collect all this data in order to process and analyze it accordingly. Depending on the maintenance granularity, the central system will be able to make system level decisions about the overall health condition of the application and related cost implications (Khan, 2015). This information could also be shared with on-site maintenance personnel or even the operating customer if required to help facilitate the diagnosis process or to correlated information with expert knowledge to investigate unknown failure incidents. Likewise, it is also required to identify and order replacement components (or other resources) from a strategic point of view to ensure availability is maintained. Within this whole process, there are various other factors such as reliability requirements, maintenance levels, built in tests results, environmental and condition monitoring data; which could be used to provide decision makers with a more complete picture about the health of the system (Cai et al, 2014).

4. CLOUD-BASED DECISION SUPPORT SYSTEMS

Figure 1 is an illustration of a typical the data flow and communication routes that are relevant within a maintenance environment. The illustration includes a data collection system, that receives data from various sources, which may include component monitoring, diagnostic information, alarming conditions – which can help identify false alarms. Data may also be generated depending on models developed from a priory information. Finally, cost models can also be developed and included (Othman et al, 2014).

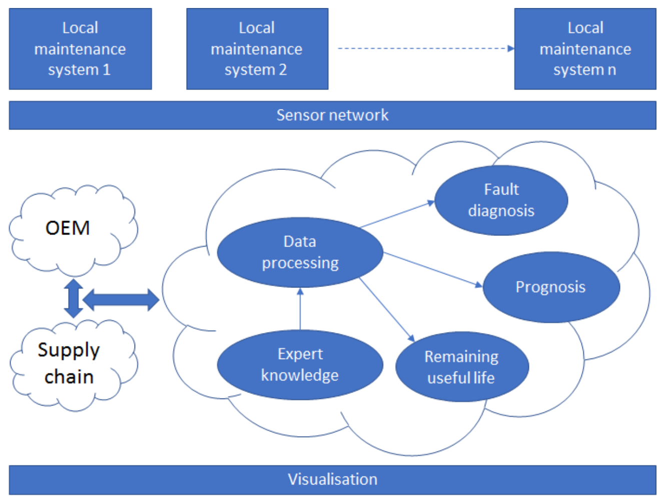

Figure 1: Initial Cloud setup/architecture

When data is collected, it needs to be conditioned to be compatible with the common format. It is converted before it can be stored or be used by for any 'training' or application control to be available for any authorized user who wishes to access the processed information. This includes details of any alarms, maintenance schedules, fault analysis results, prognostic results, costings, and more, hence enabling users to work more efficiently. Considerations associated with data collection through various sources include:

- Providing on-field expert knowledge: it is important to keep updating this database as novel failure modes appear during service.
- Typical sensors collect environmental/operating data.
- Past decision history.
- Equipment monitoring data collects data from the predesigned testability mechanisms such as built in tests and related equipment monitoring equipment. This includes degradation profiles and a priori datasets.
- Costing information can be correlated with current cost models as part of the performance monitoring data.

- Data provided by third party vendors. This could be the component history, its manufacturing details, its testability profile, etc.

The local maintenance system is expected to collection all health-related data. Acquisition is an important part of fault diagnosis. These consist of various sensors networks that is being used to record performance and health of the system, along with the environmental information it is operating in. A central repository is expected to store all this information and carry out necessary preprocessing before communication with other services on the platform. Once the data is ready, the engineer (or an autonomous system) can request services, e.g., to use particular signal processing algorithms. But more importantly, the platform must focus on maintaining seamless communication and collaboration services – where system designers, OEMs, system operators and maintenance related organizations are all linked together (See Figure 2). They are expected to share system information, report faults/failures, FMEA, novel root causes, on to a central data repository. Other factors which must be considered include anticipating bottlenecks and accounting for such instances can help maintain health management services, even if there is network congestion and component failures. Since there is increased data movement as compared to traditional methods, this growing demand must warrant dedicated network performance.

5. DISCUSSION

Collected data can be provided to various personnel, processed in various formats, and be used by a diverse range of applications for different purposes. As a consequence, some of this information might be used by maintenance support organizations to develop proprietary software which may not be able to recognize others applications – due to a lack of a standard in the industry. Similarly, if information is processed of cost/budgeting purpose, it might not be in a format which is compatible with health monitoring applications. A maintenance personnel and diagnostic monitoring equipment being used often have access to a priori data stored through experience, process models or budgeting applications. Finally, the time taken to monitor large systems, such as a fleet of aircraft, and to monitor individual failures on each one, makes it difficult to make decisions with regards to availability requirements. An AI cloud based platform can overcome these challenges by accessing data from various geographic locations and its subsequent real-time implementation. It can request and process information, and store this knowledge in a common format that can be accessed and used by other collaborating organizations. Such integration promises improved personnel safety, higher process throughput and equipment uptime, reduction in false alarms, cost reductions, improved availability and the ability to carry out operations according to design and manufacturing warranty limits. The ability to locate failure root causes also improves the quality of troubleshooting activities. Some other notable research issues associated with the concept includes:

- Knowledge Gathering
- Loss of connectivity
- Lack of real-time data
- Extrapolation of data
- Cost of analysis
- Appropriate visualization methods

AI has already influenced a generation of cloud computing infrastructure. An exciting proposition about this technology is the use of the internet of things (IoT) also termed as industry 4.0. From this perspective, IoT capabilities should be materialized as backend services that can be used from mobile applications (and other IoT devices) that use to provide services that enable the backend capabilities. On the other hand, AI applications are required to not only provide sophisticated backend services but specific runtime optimized for processing-intensive requirements of AI solutions. Despite these momentarily limitations, AI cloud based concepts have a lot of potential for the maintenance industry and practices, with more research opportunity to enable its implementation.

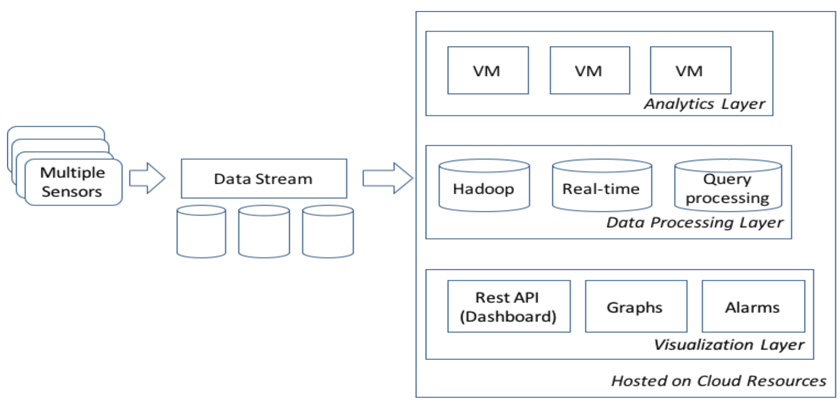

Figure 2: Cloud-based layers involved for fault diagnosis and health monitoring

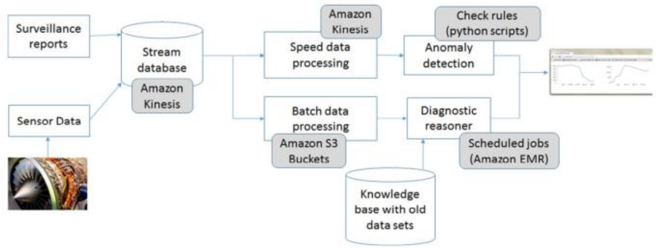

Figure 3: Architecture overview

Figure 3 shows an overall architecture for processing fictitious engine data. This architecture is being implemented over Amazon's Elastic Compute Cloud (EC2)[4] infrastructure to collect data and run diagnostics. Sensor data and failure monitoring reports into a single streaming database using the Amazon EC2 service - Kinesis. Kinesis allows data to be collected as a stream and perform real-time processing. Once processed, the data is divided into two subgroups: first for detecting anomalies and second for batch processing to find long term behavior patterns. Both actions can be done on separate Amazon services. Once calculations are done, the information can visually be analyzed in detail.

6. CONCLUSION

Traditionally, system maintenance and performance monitoring were carried out independently. Each solution attempted to optimize its own functional area, sometimes ignoring the effect its actions might have on the other functional areas. As a result, a low-priority equipment problem has the potential to cause a larger problem (or failure) whilst attempting to maintain availability requirements. Cloud Computing makes AI more accessible; even if there is a lack the computing power (to run many AI applications proficiently) by the end-user's hardware.

By using such services, various stakeholders can have access to a broader range of options to provide solutions for overall equipment data monitoring, process performance data, and process control monitoring data. Similarly, diagnostics performed on a high value asset can be considered during operational service to provide a better diagnostic analysis. Cloud based solutions also allow computation and data storage to multiple redundant off-site locations available on the network, presenting an opportunity for application software to be operated using internet-enabled devices through portable devices (such as smart phones).

Reducing the downtime for maintenance activities can help maintain availability requirements. Such solutions can be used to influence aerospace maintenance industry standards towards developing centralized maintenance regimes. However, there is a need to address further developments in existing infrastructure for large processing and frameworks where ongoing research trends indicate that AI capabilities will become a fundamental part of health monitoring.

ACKNOWLEDGEMENT

The authors would like to thank the Japanese Science Promotion Society for funding this research: P16079.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., & Ghemawat, S. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation computer systems*, *25*(6), 599-616.

Cai, G., Gross, G., Llinas, J., & Hall, D. (2014, May). A visual analytic framework for data fusion in investigative intelligence. In SPIE Sensing Technology+ Applications (pp. 91220A-91220A). International Society for Optics and Photonics.

Campbell, J. D., & Reyes-Picknell, J. V. (2015). Uptime: Strategies for excellence in maintenance management. CRC Press.

Cao, W., Mecrow, B. C., Atkinson, G. J., Bennett, J. W., & Atkinson, D. J. (2012). Overview of electric motor technologies used for more electric aircraft (MEA). IEEE Transactions on Industrial Electronics, 59(9), 3523-3531.

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 275, 314-347.

Cope, G. and Kaufman, G., (2003). Flight data transmission via satellite link and ground storage of data. U.S. Patent Application 10/441,441.

Do, P., Voisin, A., Levrat, E., & Iung, B. (2015). A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. Reliability Engineering & System Safety, 133, 22-32.

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008, November). Cloud computing and grid computing 360-degree compared. In *Grid Computing Environments Workshop, 2008. GCE'08* (pp. 1-10). Ieee.

Jazdi, N. (2014, May). Cyber physical systems in the context of Industry 4.0. In Automation, Quality and Testing, Robotics, 2014 IEEE International Conference on (pp. 1-4). IEEE.

[4] is a web service that provides secure, resizable compute capacity in Amazon's Cloud computing environment.

Jung, J. J. (2011). Service chain-based business alliance formation in service-oriented architecture. Expert Systems with Applications, 38(3), 2206-2211.

Khan, S. (2015). Maintenance requirements in aerospace systems. 4th International Conference in Through-life Engineering Services, Cranfield, UK November 2015.

Khan, S., Phillips, P., Jennions, I., & Hockley, C. (2014a). No Fault Found events in maintenance engineering Part 1: Current trends, implications and organizational practices. Reliability Engineering & System Safety, 123, 183-195.

Khan, S., Phillips, P., Hockley, C., & Jennions, I. (2014b). No Fault Found events in maintenance engineering Part 2: Root causes, technical developments and future research. Reliability Engineering & System Safety, 123, 196-208.

Khan, S., Phillips, P., Hockley, C., Jennions, I. (2015). No Fault Found: The search for the root cause. SAE publishing.

Kiran, M., Murphy, P., Monga, I., Dugan, J., & Baveja, S. S. (2015, October). Lambda architecture for cost-effective batch and speed big data processing. In Big Data (Big Data), 2015 IEEE International Conference on (pp. 2785-2792). IEEE.

Kwon, D., Hodkiewicz, M. R., Fan, J., Shibutani, T., & Pecht, M. G. (2016). IoT-based prognostics and systems health management for industrial applications. *IEEE Access*, *4*, 3659-3670.

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. Procedia CIRP, 16, 3-8.

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23.

Lightfoot, H., Baines, T., & Smart, P. (2013). The servitization of manufacturing: A systematic literature review of interdependent trends. International Journal of Operations & Production Management, 33(11/12), 1408-1434.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). Machine learning: An artificial intelligence approach. Springer Science & Business Media.

Navarro, C. A., Hitschfeld-Kahler, N., & Mateu, L. (2014). A survey on parallel computing and its applications in data-parallel problems using GPU architectures. *Communications in Computational Physics*, *15*(02), 285-329.

Othman, M., Madani, S. A., & Khan, S. U. (2014). A survey of mobile cloud computing application models. IEEE Communications Surveys & Tutorials, 16(1), 393-413.

Pascual, D., G. (2015). Artificial Intelligence Tools: Decision Support Systems in Condition Monitoring and Diagnosis. April 22, 2015 by CRC Press, Pages - 188, ISBN 9781466584051.

Roemer, M. J., Nwadiogbu, E. O., & Bloor, G. (2001). Development of diagnostic and prognostic technologies for aerospace health management applications. In Aerospace Conference, 2001, IEEE Proceedings. (Vol. 6, pp. 3139-3147). IEEE.

Sharma, A. K., Patel, S. K., & Gupta, G. (2013). Mobile Sensor Networks: A Review. Research Journal of Science and Technology, 5(3), 295-299.

Wan, J., Zhang, D., Zhao, S., Yang, L. T., & Lloret, J. (2014). Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions. IEEE Communications Magazine, 52(8), 106-113.

Womack, J. P., & Jones, D. T. (2015). Lean solutions: how companies and customers can create value and wealth together. Simon and Schuster.

Xu, X. (2012). From cloud computing to cloud manufacturing. Robotics and computer-integrated manufacturing, 28(1), 75-86.

Yasumoto, K., Yamaguchi, H., & Shigeno, H. (2016). Survey of real-time processing technologies of iot data streams. *Journal of Information Processing*, *24*(2), 195-202.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. Reliability Engineering & System Safety, 94(2), 125-141.