A Reference Architecture for IIoT and Industrial Control Systems Testbeds

B Craggs*, A Rashid*, C Hankin†, R.Antrobus, O Șerban†, N Thapen†

*Department of Computer Science, University of Bristol, UK
{barney.craggs, awais.rashid} @bristol.ac.uk
†Institute for Security Science and Technology, Imperial College London, UK
{c.hankin, o.serban, n.thapen} @imperial.ac.uk

Keywords: Industrial Control Systems, Industrial Internet of Things, Testbeds, Reference Architecture

Abstract

Conducting cyber security research within live operational technology and industrial Internet of Things environments is, understandably, not practical and as such research needs to be undertaken within non-live mimics or testbeds. However, testbeds and especially those which are built using real-world infrastructure are expensive to develop and maintain. Moreover, such testbeds tend to be representative of a single industry vertical (often based upon the skill set or research focus) and built in isolation.
In this paper we present a reference architecture, developed whilst designing and building the Bristol Cyber Security Group ICS/IIoT testbed for critical national infrastructure security research.

1 Introduction

Industrial control systems (ICS) underpin a number of societal services that are deemed critical national infrastructure. Examples include electricity generation and distribution systems, water and waste-water treatment, gas distribution networks and so on. Recent years have seen a number of high profile attacks on such infrastructures with examples ranging from Stuxnet [1] to more recent attacks against power grids [2] and manufacturing plants [3].

With the emergence of Internet of Things (IoT), there are a number of drivers for incorporating such devices into ICS environments – often referred to as Industrial Internet of Things (IIoT) – such as, enhanced visibility of industrial processes leading to improved integration with enterprise systems. This enhanced visibility and integration promises more efficient and effective business processes that take account of real-time intelligence from ICS environments leading to reduction in costs and fine-tuning of the physical processes and/or manufacturing operations controlled by the ICS.

As shown in Figure 1, IIoT leads to a convergence of IoT and Operational Technology (OT) used in ICS, hence blurring the boundaries between legacy ICS environments and contemporary IoT sensors and actuators. Cyber security issues arising from legacy components in ICS environments are well known, both in academic literature, e.g., [4, 5, 6, 7] and real-world attacks, e.g., [1, 2, 3].

In IIoT environments, legacy devices such as Programmable Logic Controllers (PLCs) and Remote Telemetry Units (RTUs) – often utilising protocols such as Modbus/TCP, Ethernet/IP, DNP3, and OPC DA where security was not a core consideration – interact and exchange data with IoT devices through IIoT data acquisition gateways. Such a scenario is depicted in Figure 1, with PLCs and RTUs residing on the same local IP network as IIoT gateways, one of which is directly polling PLCs for operational data. This same gateway is then communicating with two IIoT analytic platforms via a public communication medium (i.e. the Internet).
Yet little is understood about the attack surfaces of such convergent OT/IoT environments, their potential vulnerabilities and the kind of security architectures that may mitigate risks of attacks and disruption of the physical processes controlled by such environments. Experimentation on real-world infrastructures is not feasible due to the risks of inducing failures that may bring about the very disruption the experimentation is aiming to study and avoid in the first instance.

Within traditional ICS environments, physical testbeds have been shown to provide a good approximation of such real-world environments and have become a key tool for researchers to: A) explore and model vulnerabilities. And, B) produce viable datasets to enable the development and testing of solutions for ranging from security architectures and intrusion detection systems to novel protocols. Several extensive testbed infrastructures exist both within the UK and abroad and research has also distilled best practice guidelines for the design of such testbeds [8].

This report aims to provide similar guidance for the design of IIoT testbeds by proposing a reference architecture. We first review current ICS testbed infrastructures nationally and inter-

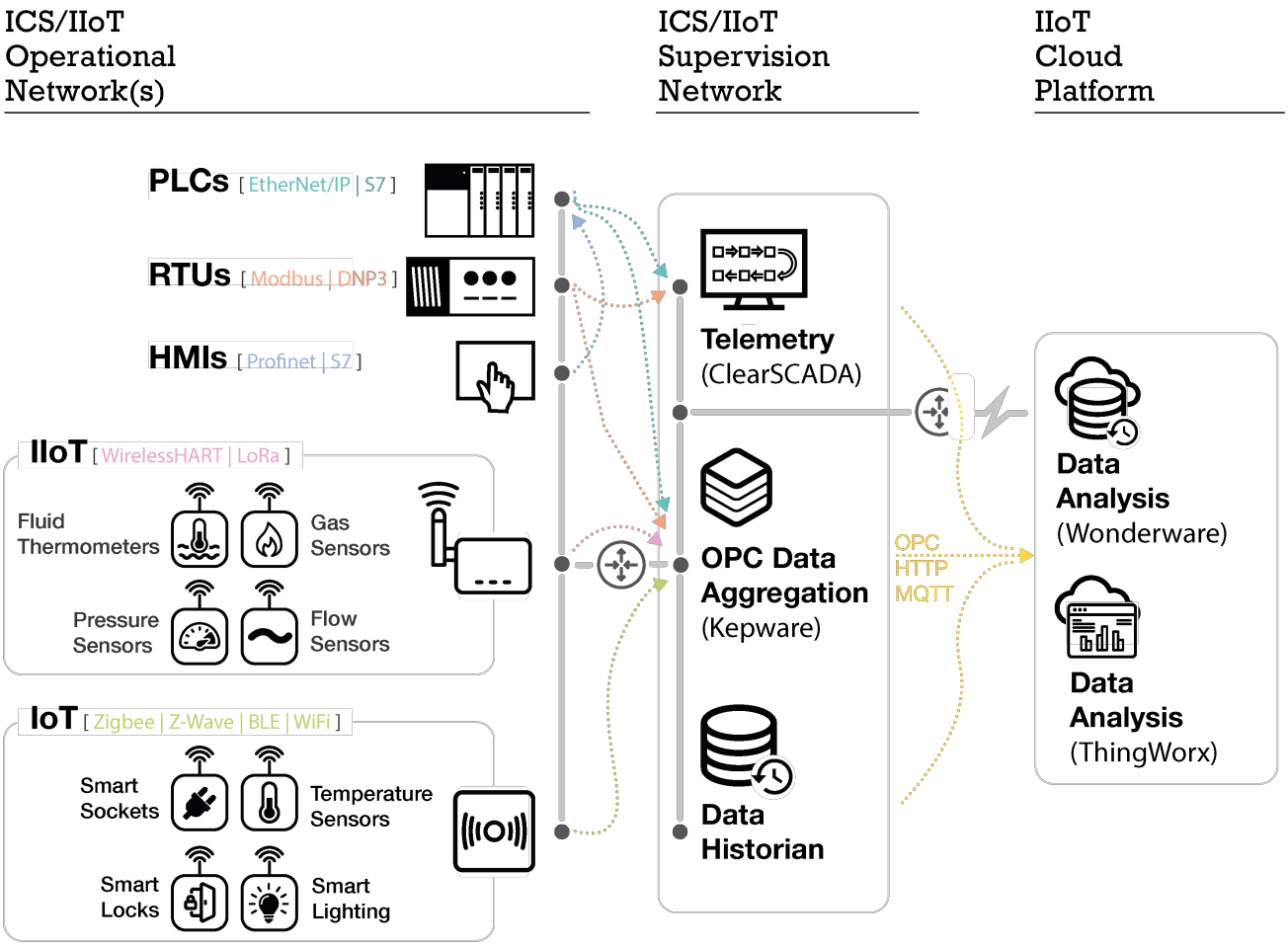

Fig. 1: A typical IIoT environment with a mix of legacy and non-legacy devices and protocols interacting with one another.

nationally (Section 2), followed by a summary of benchmark ICS and IoT datasets (Section 3). Next, based upon the converged environments portrayed in Figure 1, we detail design considerations and requirements for modelling vulnerabilities and attacks as well as data capture within testbeds (Section 4) followed by a proposed reference architecture (Section 5). We conclude (Section 6) by presenting a roadmap for future research both with regards to design of IIoT testbeds and research to be facilitated by such testbeds.

2 Map of current key ICS/IoT testbed facilities

Within the UK there are a number of ICS testbed facilities, the most extensive of which until recently was held at Lancaster University. Built as part of the MUMBA project [17], the laboratory testbed is based upon the Purdue Reference Architecture [18] and is reconfigurable to represent differing experimental contexts. Split across six Manufacturing Zones, an ICS Demilitarised Zone, and an Enterprise Zone (with its own separate Demilitarised Zone), the Lancaster ICS testbed has focused on the development of systems and devices across Levels 0, 1, 2, 3, DMZ and 4 of the Purdue model. More specialised facilities, focused on the power grid and particularly the interfacing of renewable energy sources into the grid, have been developed at Queen's University Belfast. De Montfort University is an Airbus Centre of Excellence and runs a small-scale testbed suite focused on manufacturing, water and electricity distribution.

The iTrust Water testbeds at the Singapore University of Technology and Design (SUTD) are small-scale networks within a controlled laboratory environment, composed of a small-scale water distribution network (WADI) and a treatment plant (SWaT). The testbeds are used for security analysis for water treatment and distribution networks, to assess detection mechanisms for cyber and physical attacks, as well as to understand cascading effects to other connected systems.

The iTrust *Internet of Things Automatic Security Testbed* is a small-scale laboratory composed of GPS simulator, Wi-Fi localisation simulator, time simulator, and movement sensor, to simulate the different environmental conditions in which IoT devices operate [19]. The testbed supports standard and

#	Dataset Name	General Details			Machine Learning Details		
		Format	Protocol	Ref	Time Series	Classes	Features
1	Power System †	ARFF	MODBUS	[9]	N	· 2 (normal vs attack)	128
6						· 3 (normal vs natural vs attack)	
						· 41 (attack scenarios)	
2	Gas Pipeline †	CSV	MODBUS	[10]	Y	2 (normal vs attack)	12
3	Gas Pipeline and Water Storage Tank †	ARFF	MODBUS	[11]	Y	8 (attack scenarios)	28
4	New Gas Pipeline †	ARFF	MODBUS	[12]	Y	8 (attack scenarios)	20
5	Energy Management System †	CSV	MODBUS	[13]	Y	*unknown*	9
6	Network Forensics Lessons for ICS ‡	*unknown*	EtherNet/IP (EN/IP)	[14]	Y	7 (normal vs attack types)	1-3 PLCs
7	Secure Water Treatment (SWaT) ∗	CSV	CIP over EN/IP	[15]	Y	2 (normal vs abnormal)	51 sensors
8	S317 & Blaq_0 ‡	CSV	CIP over EN/IP	[15]	Y	2 (normal vs abnormal)	51 sensors
9	Detection of IoT Botnet Attacks ⋆	CSV	Network traffic (IP packets)	[16]	Y	10 class of attacks plus 1 class of benign	115

† Dataset available at https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets
‡ Dataset not publicly available but can be requested from authors
∗ Dataset available at https://itrust.sutd.edu.sg/dataset/
⋆ Dataset available at https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_ N_BaIoT/

Table 1: List of benchmark ICS datasets available for anomoly detection

context-based security testing and analysis for IoT devices under real conditions against a set of security requirements.

Critically, missing from the current testbed landscape are testbeds that look to nascent converged ICS/IIoT environments, and most importantly the inherent risks in implementing such diverse and complex systems. Section 4 details design considerations for such testbeds to enable the modelling of vulnerabilities and attacks with this convergence in mind.

3 Benchmark Datasets

In addition to testbed facilities, and often a direct result of experimentation within those facilities, there are a number of benchmark datasets available, see Table 1).

As with current testbed facilities (see Section 2) these key datasets provide parts of the environmental puzzle, but none directly capture data from converged ICS/IIoT environments. Section 4.2 provides two key areas of data capture required for converged testbeds.

4 Design Considerations

Development of ICS testbeds is a costly, labour- and time-intensive activity that must balance a range of design considerations. Previous work [8] has highlighted the need to consider and balance three key factors:

- **Diversity**, in terms of a range of devices and software to replicate real-world scenarios in sufficient depth,
- **Scalability** of the experimental infrastructure to represent realistic scenarios,
- **Complexity** of the infrastructure and ease of deploying experiments (both locally and remotely).

These three factors also apply to convergent ICS/IIoT environments, such as that shown in Figure 1. However, to apply them to these more complex systems we need to address seven aspects that ensure coverage for both of the primary purposes of such testbeds - A) modelling vulnerabilities and attacks, and B) ensuring the production of viable datasets such as those in Table 1 for the development of detection and mitigation studies.

4.1 Modelling vulnerabilities and attacks

As discussed in Section 2, the lack of converged ICS/IIoT testbeds presents a challenge for studying such complex environments, especially where they are vulnerable and how they might be compromised. To address this we present five key aspects for the design of future converged testbeds:

1) Diverse physical processes interacting with each other: Previous research has argued that process diversity is not crucial in ICS testbeds and simpler processes and/or simulation provide suitable alternatives for process diversity [8, 20]. However, the challenges are more complex in IIoT-based ICS environments where, say, a building management system (BMS) may potentially interact with processes controlling manufacturing or processing. As a result, realistic physical processes (on a scale replicable in a laboratory setting) are essential not only to capture such interactions but also realistic environments with PLCs controlling different processes (e.g., building management, water treatment) within the same cyber-physical infrastructure.

2) Legacy and non-legacy ICS and IIoT software platforms and devices: There are more than a hundred vendors of OT and IIoT devices that provide hardware and communication services to ICSs hosted in 170 countries. Any test environment must support a diversity of such devices – with device and

technology selection market-driven – as well as a variety of physical processes so that both small-scale (one PLC and one physical process) as well as larger distributed control systems (multiple PLCs controlling one or more physical, interacting, processes) are supported.
Further, whilst IIoT devices are fairly recent in their introduction to the market and built upon modern technology and standards, the same cannot be said to be true of all ICS products. In order to model realistic industrial contexts – where legacy ICS devices interact with modern IIoT devices and gateways – any testbed must incorporate both legacy and non-legacy devices, and the ability to configure multiple field sites into a distributed IIoT-based control system managing one or more physical processes. This is essential to model a wide range of attacks in such convergent environments especially where attackers may pivot from IIoT to ICS or vice versa.

3) Support for communication protocols: Similar to device and software diversity, a testbed infrastructure must support typical ICS communication protocols such as Modbus/TCP, Ethernet/IP, DNP3, and OPC DA as well as IoT specific protocols, for example, WiFi, Bluetooth, ZigBee, Z-Wave and WirelessHart, as shown in Figure 1.
By its very inclusion, this diversity of communication protocols will introduce a number of gateways, protocol convertors and software-based aggregation platforms into the testbed, something which will ensure that learnings from the test environment can be transferred to real-world industrial contexts and applications where these are commonplace.

4) Virtualisation for scalability and ease of management: Previous work [8] has shown that deploying server and workstation instances across physical hardware is both time consuming and costly. The use of virtualisation in conjunction with virtual local area networks (VLANs) provides an easy and cost effective way to integrate new systems, scale up existing instances, and provide support for whole-network data capture. It also reduces the technical knowledge required when scaling up experiments while providing clean backups of known good systems should damage be caused during experimentation.

5) Ease of deployment of local and remote experiments: Given the intensive activity required to create a testbed, it is important that such deployments are utilised as much as possible, and not duplicated within other research groups without due consideration as to functionality already available. Furthermore, shared testbeds provide valuable opportunities for cross-institution collaboration and knowledge exchange. To facilitate this, deployments should support remote access to testbeds via secure data-communications links such that external project partners are able to utilise functionality within the testbed as though it were located in their own premises. Considerations as to resource time/cost sharing, physical reconfiguration for experimentation, the integrity of the testbed and wider cyber security issues would need to be agreed upon by all project partners.

4.2 Optimsed data capture

As with modelling vulnerabilities, the current lack of converged ICS/IIoT datasets (see Section 3 makes the development of attack detection and mitigation more complicated. Unlike an enterprise IT system will typically be disabled if an attack is detected and then restored afterwards this is not possible in an OT environment which is safety-critical and designed to protect against accidents. Nodes in an ICS network are typically resource and energy constrained, so resource-demanding cryptographic security can conflict with safety by downgrading system performance [21]. In general ICS must often provide continuous uptime and meet demanding real-time requirements for network traffic and processing as violations of these constraints can lead to safety breaches. The reference architecture should capture data on network performance so that Denial of Service attacks can be modelled and studied. Recent international standards organisations such as IEC have begun to look at safety and security together. At the end of 2017, NCSC [21] issued information about the Triton malware which targets safety controllers in safety instrumented systems. The reference architecture should include safety instrumented system controllers to enable the broader study of interactions between safety and security.
To address this we present two further aspects for the design of future converged testbeds which need to be considered to generate viable datasets:

6) Network traffic capture across IT/OT boundaries: Given the converged environments already discussed it is critical, for the identification of anomalies, that network traffic be captured across both the IT and the OT VLANs of the testbed, as well as any external (i.e. Internet facing - real or simulated) boundary routers.

On the OT (ICS) side the most common data packages used by current benchmark datasets (as shown in Table 1) are MODBUS and Ethernet/IP. MODBUS is a low level serial communication protocol used by PLCs, and is often used to pass telemetry data between instrumentation and control devices. In practice, the MODBUS packages are stored in their raw form along with the timestamps and CRC signatures. Ethernet/IP is more commonly available with systems employing modern digital infrastructures and is an extension of the Common Industrial Protocol (CIP) over Ethernet. On the IT (IIoT) side of the testbed most commonly will be TCP/IP and UDP/IP - with protocols such as OPC DA being mapped to TCP/IP and often referred to as *"OPC/TCP"* or *"OPC Binary"*.

7) Telemetry data capture: Recent work based on the use of Generative Adversarial Networks (GANs) [20] shows that attackers can achieve different effects depending on the parts of

the system they are able to access. The high success rates reported in [20] highlight that analysis of network data is unlikely to be sufficient, on its own, to identify sophisticated attacks. There are clearly other (physical) indicators, such as power consumption, pressure and temperature, that provide additional evidence of malfunction. Some of this data (or normal ranges of values) are potentially encoded in the control model for the system. Normal ranges of values can be pre-defined, extracted from the historical running of the system, or generated by an appropriate mathematical model describing the operation of the system. Sub-networks within the reference architecture should be accompanied by such a control model.

A note on data processing

Network Data can be viewed as *Big Data* as, due to its volume and velocity, the data logs cannot be processed with traditional techniques. It has recently been shown that Deep Learning approaches such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can achieve state of the art performance on anomaly detection in network data from ICS systems. Long Short-Term Memory Networks are a subtype of RNNs with very good results in this field [19]. More recently, Generative Adversarial Networks (GANs) have been used to improve these results by increasing the robustness of the models [20]. Traditionally, this problem was also approached using Support Vector Machines (SVM), Bayesian Networks (BN) or Hidden Markov Models (HMM), which may still be considered state-of-the art for certain classes of problem.

The problem of anomaly detection and intrusion for ICS can be approached in either a supervised or an unsupervised manner. In the supervised context, the dataset would usually need to contain both normal and abnormal samples, so the algorithm can derive general rules from the data patterns. When only normal data are available an anomaly detection system can be trained to predict the value of the next sample, and raise an error if the actual sample is too different from the prediction. In the unsupervised context, the algorithm does not have any information regarding the type of behaviour (e.g., normal or abnormal) at training time.

In order to use Machine Learning techniques, the data usually needs to be preprocessed. Some algorithms, such as Random Decision Trees or HMMs can work with categorical data. Most of the other models, such as Deep Learning or SVMs, require numerical data.

In order to use these architectures successfully training data must be pre-processed and normalised. When normalising the data, each numerical field presented to the network must be on the same scale. This allows gradient descent to function more efficiently. To achieve this each numerical data field should be normalised to have zero mean and unit variance. All categorical data must be converted to either a one-hot vector or into a vector embedding space.

5 Reference Architecture for IIoT and ICS Testbeds

A high-level representation of the IIoT/ICS reference architecture is given in Figure 2. Broadly, the architecture is split into five main *zones*, which together address the seven key design considerations outlined in Section 4. The zones are connected through a network topology analogous to that often found within organisations maintaining an IT/OT infrastructure.

Zone 1 - Experimental

The *Experimental* zone provides the underlying infrastructure utilised by the testbed, and is essentially a utility to enable design considerations 4—7. At its core lays a series of virtualised network attached servers, workstations and storage arrays which are accessed and used across the zones 2 (Control) and 5 (Management). The choice of virtualised operating systems should be based upon those which are deployed in the wild by organisations utilising such IT/OT environments. Primarily these would tend to be a mixture of Microsoft Windows workstations and servers, although care should be taken to introduce alternatives where observed (for example, an email or web server might actually be based upon Linux within the Management Zone).

As discussed briefly in Section 4, the volumes of network and telemetry data from within such an architecture can be significant, as such network attached storage, or even a storage area network, should be considered as the default storage mechanisms for all virtualised workstations and servers.

It can be seen that the Experimental zone connects to the main *Corporate WAN/LAN Gateway Router*. It is this top level connectivity, along with further discrete network port taps at all

Zone	1: Experimental	2: Control	3: Production	4: Process	5: Management
Design Consideration					
i) Diverse physical processes		✓	✓	✓	
ii) (Non) legacy sw & devices		✓	✓	✓	
iii) Multiple comms protocols		✓	✓	✓	
iv) Virtualised environment	✓	•			•
v) Easy experimental deployment	✓	•	•		•
vi) Network data capture	✓	•	•	•	•
vii) Telemetry data capture	✓	•	•	•	•

Key: ✓ enables | • used here

Table 2: Reference architecture zone mapping to design considerations.

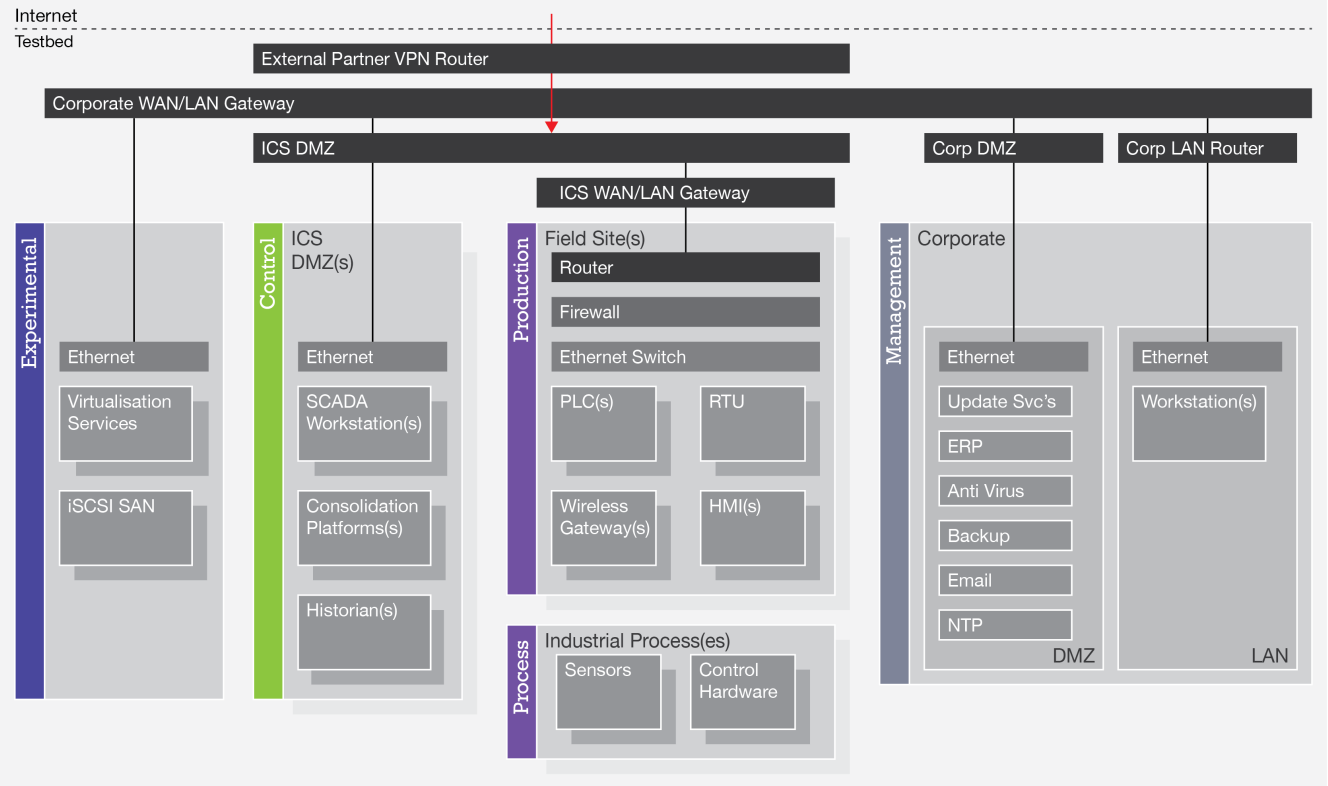

Fig. 2: A reference architecture for converged IIoT and ICS testbeds.

key router boundaries, which enables the Experimental zone to not only provide services across the architecture but, most importantly, to gather the necessary data for analysis.

Zone 2 - Control

The *Control* zone is part of the testbed operational technology (OT), and provides the *top-end* services that oversee zones 3 (Production) & 4 (Process). This would typically include SCADA workstations, data & protocol aggregation platforms and time series data historians - these would normally be virtualised from Zone 1. Together these services may be delivered as part of an operations centre.

In the reference architecture, the Control zone is given as being connected to the ICS demilitarised zone (DMZ), as this is often where it would reside in organisational networks. It should be noted that aspects of the Control zone may also be located elsewhere. It is not unheard of to find SCADA workstations located within a corporate IT zone (Zone 5 Management). Similarly, and more recently with the advent of IIoT, there are commercialised deployments of Control services outside the corporate auspices in the Cloud. Examples of this would include SCADA and IIoT aggregation services (including historians) in the Cloud. Whilst these are not expressly represented in the reference architecture a number of experimental configurations can cater for these:

- Firstly, through the use of VLANs and virtualisation it is readily possible to home any Control service in a pseudo-Cloud VLAN external to the Corporate WAN/LAN gateway.
- Secondly, and made possible by the design consideration for ease of experimental deployment (see Section 4), it is possible to use actual Cloud-based services (live or preferably experimentally sandboxed) running on data from the testbed.

Zones 3 - Production & 4 - Process

At the core of the reference architecture are the physical processes (Zone 4) and the equipment that provides direct control of them (Zone 3). Together these two zones represent the diverse interacting ICSs required by the design considerations.

Zone 3 (Production) - sometimes also referred to as *field-site(s)* - provides support for multiple devices, for example, programmable logic controllers (PLCs), Remote Terminal Units (RTUs), wireless gateways and Human Machine Interfaces (HMIs). Furthermore, these devices and the consolidation platforms, in the Control zone, support a wide variety of different communication protocols used in ICS including Modbus/TCP, Ethernet/IP, DNP3, and OPC DA; as well as IoT specific protocols, for example, ZigBee and WirelessHart.

Within an implementation of the reference architecture we would recommend that Production *field-sites* be associated one-to-one with physical processes for ease of initial setup as

the cabling requirements can be extensive and highly inefficient to re-configure for different experiments. However, to enable ease of experimental setup, we suggest that *field-sites* are designed to be as device-manufacturer agnostic as possible as to enable, for example, PLCs from multiple manufacturers to be readily swapped out for any given process.

Zone 4 (Process) contains the actual physical industrial processes and their associated sensors and control hardware. As the reference architecture is designed to assist in studying converged interacting systems it is necessary to have multiple processes, each representing a different part of an organisation's OT. For example, it would be common to have a building management system (BMS) interacting with another such as a smart-factory process or lift-control system.

Zone 5 - Management

The Management zone represents what might be considered an organisation's *normal* information technology (IT) and can consist of any number of VLANs, services (ideally also virtualised), workstations etc. The intention of including a representation of an IT zone is to provide an environment that is often considered separate from OT - both physically and support-wise. It is not uncommon to find OT environments have considerably stricter change control than a corporate IT network yet, and especially with the creep of IoT and use of shared facilities such as BMS, there is potential for network bridging and attack pivoting. The inclusion of an IT zone affords a place where these can be studied in an IT/OT mixed environment.

Secure Federated Experimentation

Within the reference architecture is the inclusion of an *External partner VPN router*. As mentioned, testbeds are expensive to build and maintain. They can also sit idle between experiments. To better take advantage of such often publicly funded resources, and the domain specific expertise that often resides with a vertical sector testbed, the reference architecture facilitates the secure federation of testbed resources between partner organisations in three main configurations:

1) Remote desktop access to virtualised testbed
The most straightforward method of remote access is facilitated by the extensive use of virtualisation, allowing external project partners to connect to remote desktop sessions via a dedicated virtual private network (VPN). This affords the external partner access to the SCADA workstations with the same view into the production level as if they were physically located in the testbed operations centre.

Advantages
+ Simple to setup and manage
+ Does not require partners to hold licenses for consolidation platforms such as ThingWorx or ClearSCADA
+ Secure, with no ability to push data into / pull data from testbed enforced by configuration of RDP sessions

Disadvantages
- Requires cooperation as to what experimental setup (within production level) is physically enabled at any one time.

2) Control zone access to local production zone
The second use case for external partner testbed access is where the partner already has their own control zone setup – an operations centre, for example, but not their own production or process deployment. Using the same secure data-communications an external partner can connect their own control zone to the production/process zone they wish to run experiments upon.

Advantages
+ Provides external partner more flexibility, for example running a security information and event management (SIEM) system over multiple disparate/remote production zones at a number of sites.

Disadvantages
- Requires careful planning and setup to ensure testbed integrity

3) Federation of external production zone into local testbed
The third use case allows for testbeds to federate production zones, allowing for interoperability and data sharing between themselves. For instance, where an implementation of the reference architecture lacks, for example, an agriculture production/process deployment but the researchers want to look at the interplay between this and a local IIoT deployment. By co-opting a remote agriculture testbed as though it was part of the local testbed this becomes possible.

Advantages
+ Incredibly flexible, allowing for reduced duplication of testbeds
+ Allows production level deployments to remain within locus of expertise

Disadvantages
- Complex to set up
- Due consideration to cyber security needed at planning stage

6 Conclusion

The challenges of capturing data through instrumentation can only be fully understood as testbeds, developed on the basis of the reference architecture, come online and are used over a period of time for a variety of experiments. This will yield fundamental insights into patterns of required instrumentation as well as new research into optimisation of such instrumenta-

tion. Research is also required into data curation approaches to ensure that testbeds instrument for and capture data pertinent to research in cyber security and that the captured data is fit for purpose.

We strongly encourage further research utilising and interconnecting such testbed infrastructures with a focus upon tackling three key gaps relating to cyber security in convergent ICS/IIoT environments:

1. A set of reference attack scenarios that to be utilised by the community at-large to evaluate cyber security solutions for such environments. This will support comparability between solutions.
2. Studying whether current attack modelling techniques can effectively capture how attacks propagate in such converged environments and developing techniques to study the impact of such attacks and their propagation.
3. Design principles and guidelines to mitigate the cyber security risks arising from the convergence of OT and IoT.

This paper has proposed a reference architecture to underpin realistic future testbed infrastructures for the support of rigorous studies of the cyber security challenges posed by the convergence of OT and IoT. Developed whilst designing and building the Bristol Cyber Security Group's ICS/IIoT testbed for critical national infrastructure (CNI) security research, the reference architecture aims to address a number of requirements for modelling attacks and vulnerabilities as well as those for capturing data pertinent to data analytics and machine learning techniques.

Acknowledgment

This work has been funded by the UK Engineering and Physical Science Research Council (EPSRC) as part of PETRAS: Cybersecurity of the Internet of Things Research Hub, grant no EP/N023234/1.

References

[1] D. Kushner, "The real story of stuxnet," *ieee Spectrum*, vol. 3, no. 50, pp. 48–53, 2013.

[2] T. Pultarova, "News briefing: Cyber security-ukraine grid hack is wake-up call for network operators," *Engineering & Technology*, vol. 11, no. 1, pp. 12–13, 2016.

[3] R. M. Lee, M. J. Assante, and T. Conway, "German steel mill cyber attack," *Industrial Control Systems*, vol. 30, p. 62, 2014.

[4] S. McLaughlin and P. McDaniel, "Sabot: specification-based payload generation for programmable logic controllers," in *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 2012, pp. 439–449.

[5] D. Hadžiosmanović, R. Sommer, E. Zambon, and P. H. Hartel, "Through the eye of the plc: semantic security monitoring for industrial processes," in *Proceedings of the 30th Annual Computer Security Applications Conference*. ACM, 2014, pp. 126–135.

[6] W. Jardine, S. Frey, B. Green, and A. Rashid, "Senami: Selective non-invasive active monitoring for ics intrusion detection," in *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. ACM, 2016, pp. 23–34.

[7] R. Antrobus, S. Frey, B. Green, and A. Rashid, "Simaticscan: towards a specialised vulnerability scanner for industrial control systems," in *Proceedings of the 4th International Symposium for ICS & SCADA Cyber Security Research 2016*. BCS Learning & Development Ltd., 2016, pp. 1–8.

[8] B. Green, A. T. Le, R. Antrobus, U. Roedig, D. Hutchison, and A. Rashid, "Pains, gains and plcs: Ten lessons from building an industrial control systems testbed for security research," 2017.

[9] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 3104–3113, 2015.

[10] J. M. Beaver, R. C. Borges-Hink, and M. A. Buckner, "An evaluation of machine learning methods to detect malicious scada communications," in *Proceedings of the 2013 12th International Conference on Machine Learning and Applications-Volume 02*. IEEE Computer Society, 2013, pp. 54–59.

[11] T. Morris and W. Gao, "Industrial control system network traffic data sets to facilitate intrusion detection system research," *Critical infrastructure protection VIII—8th IFIP WG*, vol. 11, pp. 17–19.

[12] T. H. Morris, Z. Thornton, and I. Turnipseed, "Industrial control system simulation and data logging for intrusion detection system research," *7th Annual Southeastern Cyber Security Summit, Huntsville, AL.*, 2015.

[13] Collated by Morris, T. (2017) Industrial Control System (ICS) Cyber Attack Datasets. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets [Accessed: Jul-18]

[14] T. D. Nguyen, "Network forensics lessons for industrial control systems," Naval Postgraduate School Monterey United States, Tech. Rep., 2016.

[15] J. Goh, S. Adepu, K. N. Junejo, and A. Mathur, "A dataset to support research in the design of secure water treatment systems," in *International Conference on Critical Information Infrastructures Security*. Springer, 2016, pp. 88–99.

[16] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-baiot—network-based detection of iot botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, 2018.

[17] RITICS: Research Institute in Trustworthy Inter-connected Cyber-physical Systems. (2017) Mumba - multi-faceted metrics for ics business risk analysis. [Online]. Available: https://ritics.org/mumba/ [Accessed: Nov-18]

[18] T. J. Williams, "The purdue enterprise reference architecture," *Computers in industry*, vol. 24, no. 2-3, pp. 141–158, 1994.

[19] S. Siboni, V. Sachidananda, Y. Meidan, M. Bohadana, Y. Mathov, S. Bhairav, A. Shabtai, and Y. Elovici, "Security testbed for internet-of-things devices," *IEEE Transactions on Reliability*, 2018.

[20] S. McLaughlin, C. Konstantinou, X. Wang, L. Davi, A.-R. Sadeghi, M. Maniatakos, and R. Karri, "The cybersecurity landscape in industrial control systems," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1039–1057, 2016.

[21] National Cyber Security Centre. (2017) Weekly Threat Report 15th December 2017. [Online]. Available: https://www.ncsc.gov.uk/report/weekly-threat-report-15th-december-2017 [Accessed: Jan-18]