Virtualising Railway Control Centres: Can Virtualisation and Cloud Computing Deliver Increased Resilience?

Matt Slade, MEng (Hons) CEng MIET, CPC Systems

SUMMARY

Hardware virtualisation is a method of simulating elements of computer hardware on a shared host server or a group of servers configured as a cluster; this includes processing, storage and network resources. Use of virtualised hardware is becoming commonplace in Industrial Control Systems (ICS). Amongst other benefits this approach can afford greater efficiency, availability and maintainability. In 2017, 30% of ICS systems used some form of virtualisation within their architecture; this is expected to increase significantly (GE, 2016).

Despite the increased prevalence in ICS, adoption of virtualisation within the UK Rail industry for control system applications has been slow. Rail suppliers are starting to provide virtualised systems; however, these are limited and often hosted in isolation.

The fragmented adoption of virtualisation in railways is partly due to the extended operational lifetime of rail systems, the reliance on a diverse supply chain and the focus on integrity and availability. Greater adoption of virtualisation and increased integration of control systems may provide benefits.

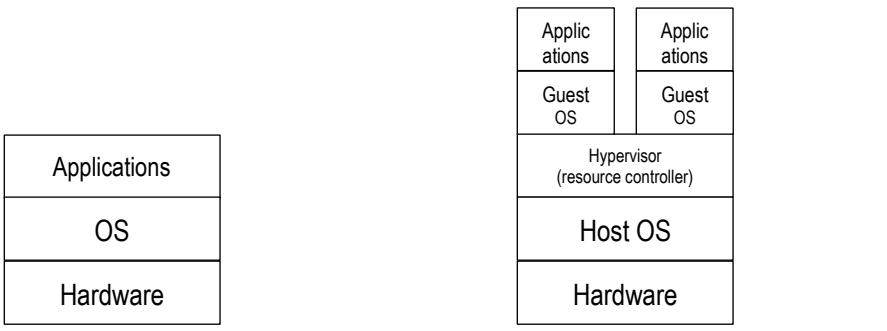
'Cloud computing' refers to the use of host machines operated by a third party and accessed via the internet. This approach is often cited as providing additional flexibility and reliability to normal business users but has not been widely adopted for industrial control system or rail applications.

The objective of this paper is to stimulate discussion around adoption of virtualisation and the use of cloud computing. The paper explains the concept of virtualisation and how it can be applied to rail control centres. The paper explains the potential benefits of virtualisation and cloud computing and identifies the challenges to the industry in adopting these technologies.

1 INTRODUCTION

Virtualisation is the process of running a simulation of something rather than an actual version. In computing the concept has been around since the 1960's, when there was a need to split the resources of a single powerful computer between multiple tasks. Since then, the use of the term virtualisation has become broad and can be ambiguous; this paper adopts the term virtualisation to mean the deployment of Virtual Machines (VM).

Figure 1 demonstrates the difference between a traditional standalone machine deployment and a Virtual Machine.



Standard Server/Workstation Deployment with Specific Hardware

Virtualised Server/Workstation: Deployment of 2 Machines on Shared Hardware

Figure 1: Representation of standalone and virtualised machine deployment

Traditionally machines have discrete hardware, whereas VMs share common hardware. The hypervisor software allows each VM to run independently, utilising the shared resources of a host server or cluster of servers. The VM is referred to as a guest and the hardware as the host machine. The number of VMs deployed on a host, or group of hosts known as a cluster, is limited by the pool of resources available and hypervisor software (Macek, 2019).

The VMs are not normally aware that they are running in a virtual configuration. The hypervisor layer simulates the hardware resources the machine requires to operate, e.g. RAM, processing power or network interface etc and allocates more resources as required. This approach can be more efficient at using resources, under normal operation standalone machines will use a fraction of their available resources but are specified to account for times of high load. Suppliers often claim virtualisation provides improved reliability and maintainability through the additional features and functionalities it offers.

Within the field of computing where large numbers of machines are hosted, the energy and resource efficiency benefit of virtualisation easily create the business case for its adoption. In rail where we only require a limited number of machines and our primary focus is the safe and efficient operation of the network, the business case is not as evident. Therefore, the business case must be made in terms of the emergent properties important to the railway, such as reliability, availability, maintainability, security and safety.

The term ‘cloud’ is used broadly within computing, as with virtualisation, this term has become ambiguous. ‘Cloud computing’ refers to the deployment of the guest VMs and storage on remote host machines provided by a third party. Users access the machines using thin clients, low-cost machines with minimal processing and memory resources that connect to the VMs via a network such as the internet or a company’s intranet. This approach affords companies flexibility without the cost of the hardware and allows greater numbers of machines to be hosted collectively; hence, the resource efficiency saving is greater.

Railways can use virtualisation and cloud computing in a variety of ways to provide benefit, but this also introduces challenges. This paper investigates the use of Virtual Machines within our control centres, providing a roadmap from the virtualisation of isolated systems, through to an integrated approach hosting all our systems on the same hardware or the cloud. The paper discusses the benefits and challenges of each stage of the roadmap; making recommendations on how to addresses some of the challenges.

Section 2 of this paper introduces a road map of virtualisation maturity levels for adoption within the rail industry. Section 3 discusses the benefits and challenges introduced with the adoption of each maturity level. Section 4 presents conclusions and makes recommendations on the adoption of virtualisation and use of the cloud for use in railway control centres.

2 VIRTUALISATION IN RAILWAYS

The railway control room contains multiple systems that interface together and are used by operators and maintainers to operate the railway, examples are:

- Signalling Control Systems
- Traffic Management Systems (TMS)
- Supervisory Control and Data Acquisition (SCADA):
    - Traction Power
    - Ancillary Systems – lifts, fire etc.
- Condition Monitoring
- Passenger Information
- Voice Communications
- Closed Circuit Television (CCTV)

The level of integration and virtualisation of these systems is dependent on the railway's age and the system suppliers. Greenfield railways procured recently may use a single supplier to provide a range of sub-systems concurrently, allowing the integration of the systems to a greater degree than if they were bought or renewed independently over a longer period by multiple suppliers.

2.1 Control System Model

To aid the explanation of how control centres are becoming virtualised, this paper uses a pyramid model. Each of the systems and its components is mapped to a traditional control pyramid hierarchy, shown in Figure 2.

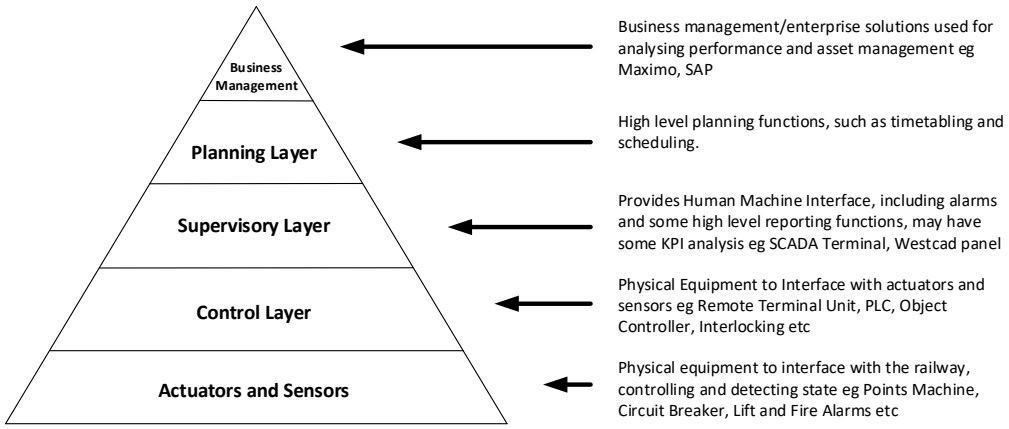


Figure 2: Control System Pyramid Model

Note - Modern systems in fields such as SCADA are reducing the equipment within the control layer, choosing to increase the intelligence of the actuation and sensor layer instead; however, this is outside the scope of this paper.

A similar mapping can be performed for each rail subsystem. The systems in the control, action and sensing layers are generally responsible for the detection and direct actuation of controls to equipment on the railway. These require low latency and safety-critical functions require higher Safety Integrity Levels (SIL). Using signalling control systems as an example, the bottom two layers are usually categorised as SIL4 as they are responsible for control of points and interlocking, whereas the supervisory level are usually required to be SIL2 to avoid requesting unsafe conditions.

The locations of components can also be generalised in the model. Equipment in the actuating and sensing layer is usually required to be located trackside or within stations/platforms. Supervisory and planning systems are frequently located within service control centres and their equipment rooms. Control layer equipment varies depending on system architecture and application being in remote equipment rooms; within location cases or

centralised equipment rooms. Business management systems are typically hosted by company IT providers, or third-party companies, so are not bound to a specific location.

This paper considers only the virtualisation and migration to the cloud of the planning and supervisory level systems traditionally hosted on standalone machines within railway control centres.

2.2 Roadmap of virtualisation and integration of systems

When considering the application of virtualisation, a range of elements require consideration; the number of systems virtualised, the adoption of advanced features that virtualisation provides and the level of integration between the virtualised systems. Figure 3 demonstrates a roadmap of the adoption of virtualisation, defining four levels used for comparison within this paper. The levels have been selected to demonstrate the benefits and challenges greater use of Virtual Machines will introduce, in practice, there are shades of grey between the levels.

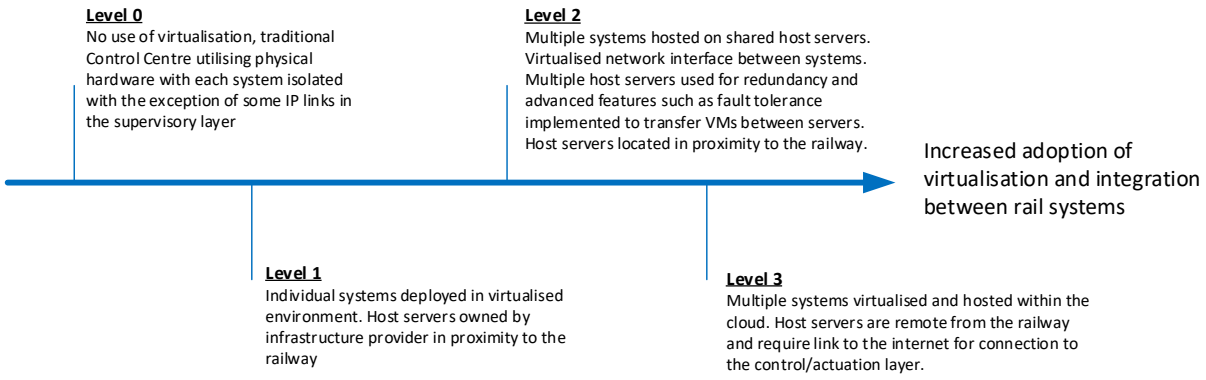


Figure 3: Virtualisation Roadmap

Depending on the country, rail network, route or line, the level of adoption will vary. Within the UK, most systems are around level 0 or 1, with only a minority of systems such as condition monitoring and Traffic Management Systems (TMS) using virtualised components. Within the Netherlands, the TMS has been running on a cluster for over a decade, with some sub-components virtualised, whereas the UK is just starting to use virtualisation for mainline TMS.

Suppliers such as Alstom and Siemens have integrated solutions that utilise virtualisation. Deployment of these systems is reliant on infrastructure owners requesting them which will not ordinarily be the case for brownfield railways renewing life expired systems at different times. Even in the case of greenfield railways, procurement mechanisms for systems often split the supply systems into separate contracts making integration and the adoption of the level 2 approach more difficult.

We are making progress; in 2017 Siemens deployed its Iltis system on the Gornergrat Railway in Switzerland. This application utilised hardware located 170Km away from the railway to host servers and workstations via dedicated links. This is the world's first control system provided as a service. Despite this, major high capacity metro projects currently being introduced are using standalone servers, potentially missing out on the advantages of hardware virtualisation.

Sections 2.2.1 to 2.2.4 summarise the architecture of each level within the roadmap. TMS and SCADA systems are used as an example, although the approach can be applied to most control centre systems.

2.2.1 Level 0 – No Virtualisation

Level 0 represents the traditional railway control systems; no systems utilise virtualisation and systems are deployed separately using standalone hardware, with minimum levels of integration. Where systems interface and events on one system require actions on another; the interaction is the responsibility of operators. An example could be when a fire alarm requires trains stopped and tunnel ventilation systems to be activated.

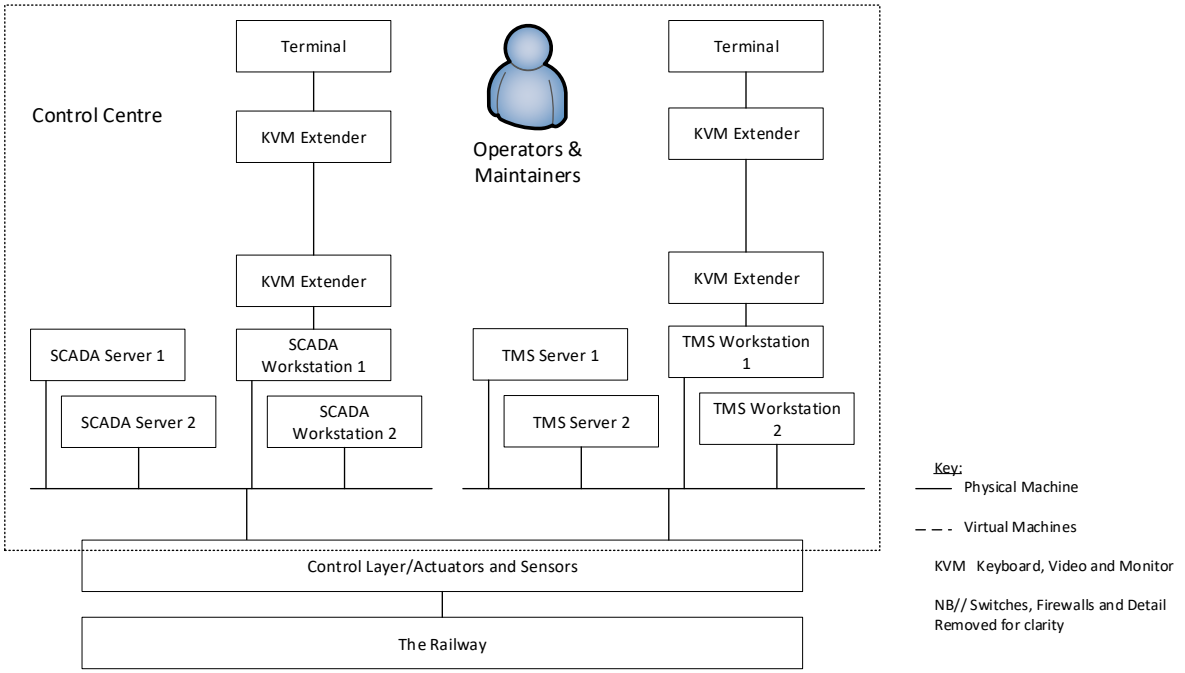


Figure 4: Basic Overview of Level 0 System – Traditional deployment with standalone machines

The reliability of one system does not influence another. To achieve high reliability, servers are deployed in hot-standby configurations; application layer software manages the failover between machines should a fault be detected. The control centre network for sub-systems may be physically separate, or maybe separated on virtual LANs. All network hardware such as firewalls and switches are physical devices housed in the control centre.

2.2.2 Level 1 – Virtualisation of Isolated Systems

Level 1 considers the virtualisation of one individual system, but where the systems are deployed independently, hosted on separate servers and located in the control centre. Figure 5 displays a basic architecture showing the SCADA servers and workstations virtualised. The TMS and other control centre systems are unchanged, elements of the SCADA system within the control layer are also unchanged.

The infrastructure owner deploys the host machines within the control centre. The VMs are unaware they are running in virtual configuration, and changes to the application software are avoided. Advanced virtualisation facilities are not used to manage failovers between machines. Instead, redundancy and failover between primary and standby machines is managed at the application layer by the guest systems, as it would be with the servers hosted in a standalone configuration.

The VMs for the primary and standby server may be split over to separate host machines to meet reliability targets should a host machine fail.

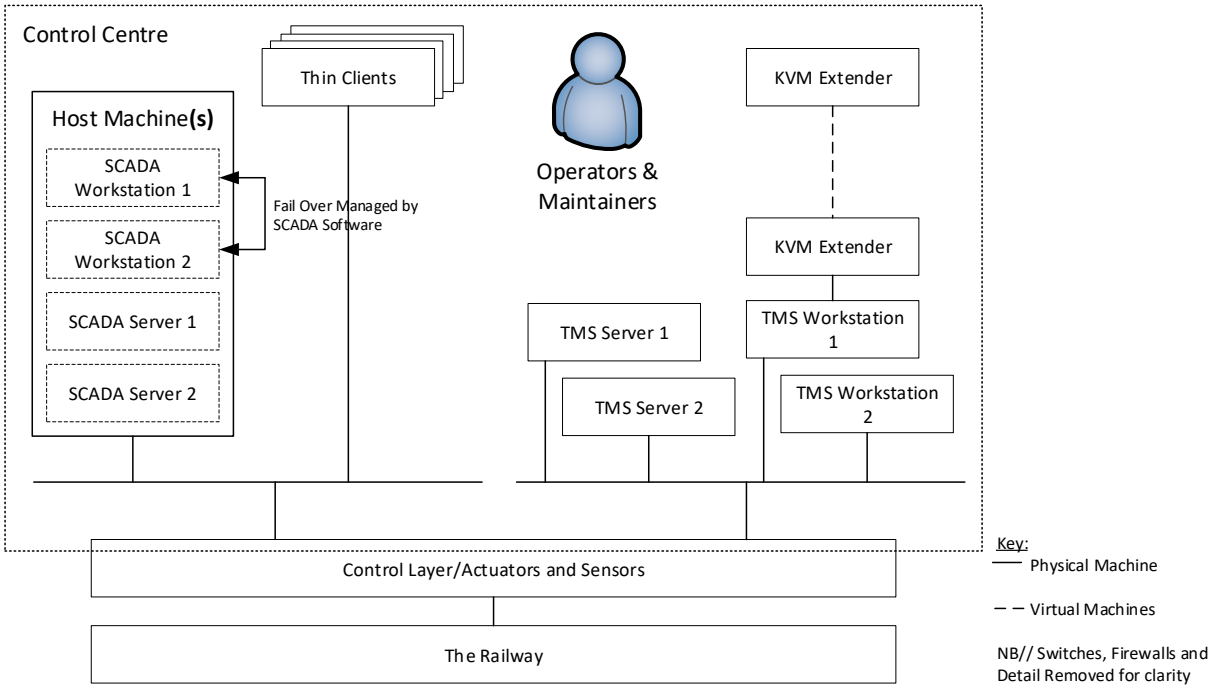


Figure 5: Basic Overview of Level 1 – SCADA system shown running using VMs

Virtualised systems that require a Human Machine Interface (HMI) interface use thin-client machines to provide terminals for inputs or interfaces to overview screens. Interfaces between systems use hardwired connections. Maintenance of the servers maybe be by different teams or organisations.

This approach minimises infrastructure and application software changes from level 0, therefore is more applicable where brownfield railways are renewing hardware and seeking to take advantage of some of the benefits of virtualisation.

2.2.3 Level 2 – Virtualisation & Integration of Multiple Systems Including the Use of Advanced Virtualisation Functionality

Level 2 takes advantage of advanced virtualisation functionality and requires architecture and application software changes to support operation in the virtual environment. This approach is suited to greenfield railways or railways renewing multiple systems simultaneously.

Multiple systems are virtualised down to the supervisory layer and hosted on shared host servers, owned and maintained by the infrastructure owner. The host servers are split over diverse location to maintain geographic diversity. A Wide Area Network is used for the communications link, and the reliability is suitable for the systems being hosted. *Figure 6* demonstrates a basic overview of the architecture.

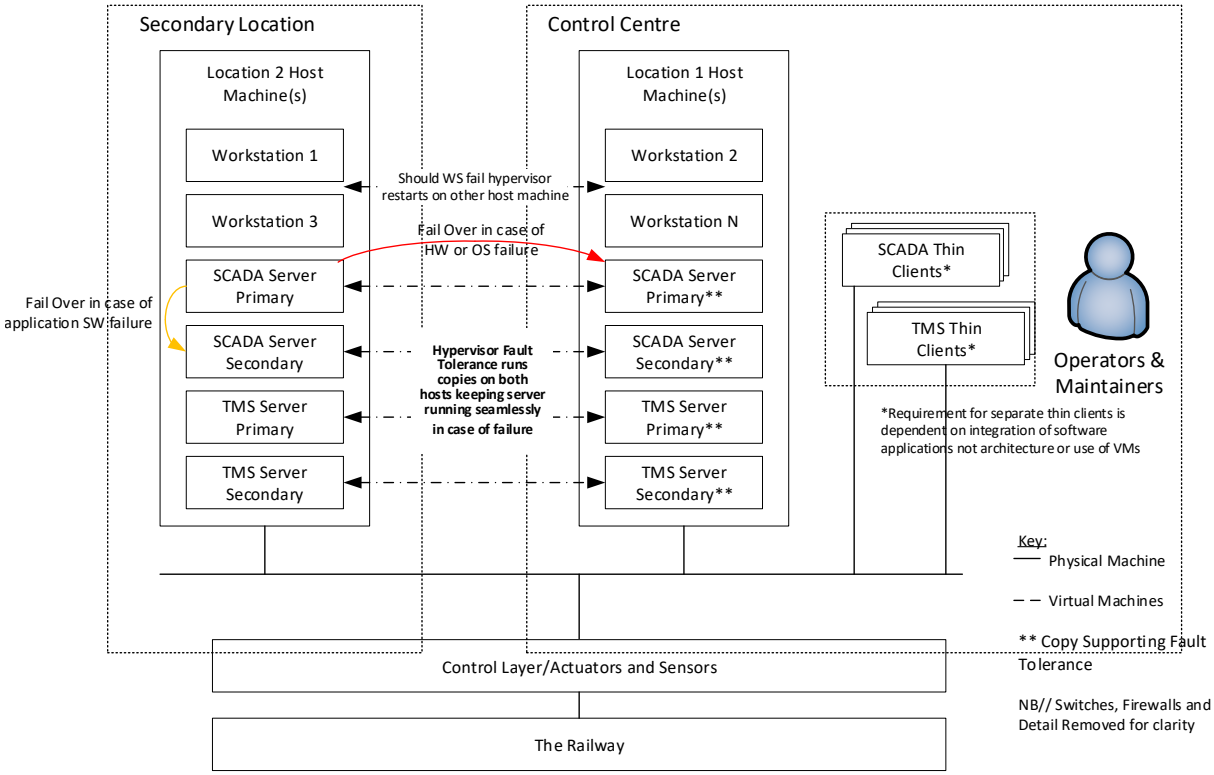


Figure 6: Basic Overview of Level 2 Architecture.

The host servers are configured as a cluster, using advanced hypervisor functions. 'Fault Tolerance' is a function performed by the hypervisor, it continually monitors the VMs and runs a mirrored copy of the VM on a separate host machine, should the one host machine fail the VM will remain operational on the other host and the user will not notice the failure, shown by the red arrow in *Figure 6*. While 'Fault Tolerance' manages issues with host hardware or guest operating systems, issues may still occur with the application software. Additional servers are provided for software diversity and allow failover should an issue occur with the application software on the primary server, shown by orange arrow in *Figure 6*.

Workstations use a hypervisor function called 'High Availability', if the hypervisor detects a fault with the VM, or host machine, it automatically restarts the workstation VM on another host.

The communications interface between the systems is virtualised, for example, communications between the SCADA system and TMS is via a virtualised IP network. This can facilitate integration between systems but is reliant the software applications being updated to take advantage of this.

Note: When new systems are deployed, it may be desirable that the HMI provides an integrated overview of the railway via a single workstation summarising information from all subsystems. This is possible with a single supplier providing all systems at the same time, but for brownfield railways may not be possible. This paper is considering

just the benefits and challenges of virtualisation on the same host server and does not consider the integration due to use of a single supplier using one integrated software platform.

2.2.4 Level 3 – Migrating to the 'Cloud'

Level 3 considers the migration of the guest machines to the 'Cloud'. The host machines are not owned and maintained by the infrastructure owner. Servers are hosted by a third-party, a Service Level Agreement (SLA) is used to manage the service provided. Advanced virtualisation features such as High Availability and Fault Tolerance are implemented.

Thin clients are provided within control facilities to provide the HMI interface for operators and maintainers. The thin clients connect to the workstation VMs in the cloud via dedicated leased line fibre links provided by an Internet Service Provider (ISP). An SLA is used for the provision of leased line links; dual redundant links are provided to separate exchanges to access the ISP's core network.

Figure 7 provides and basic overview of architecture. The communications network between VMs in the cloud is virtualised.

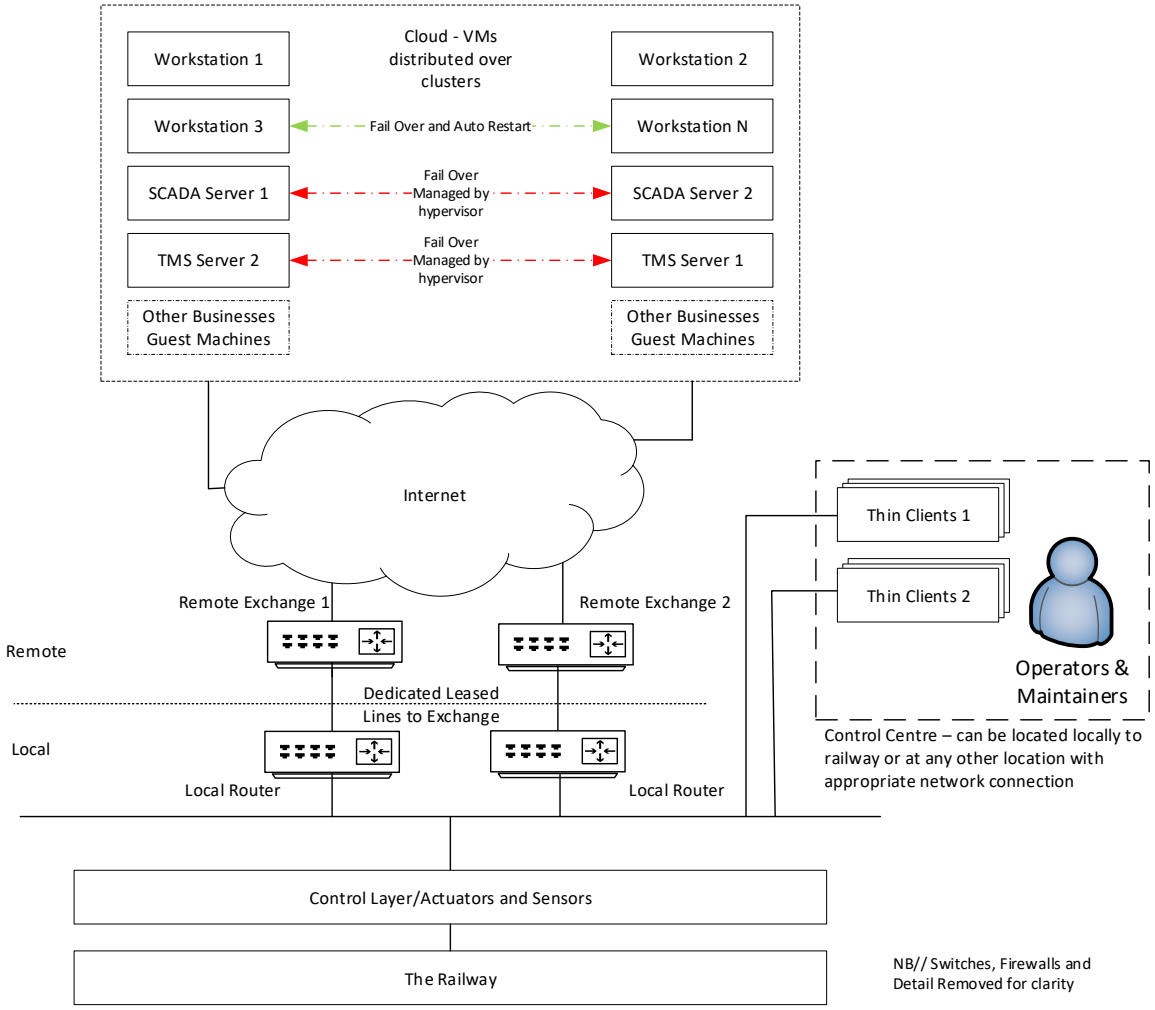


Figure 7: Basic Overview of Level 3 Implementation

The cloud computing service is either provided by a cloud service provider, such as Amazon AWS, or by the signalling or SCADA supplier, hosting the VMs on a cluster in their factory. This approach paves the way for systems as a service, e.g. 'Signalling as a Service', I believe this is a similar approach used to the Gornergrat Railway implemented by Siemens.

3 BENEFITS AND CHALLENGES

The benefits and challenges of levels 1 through 3 have been consider relating to the following areas:

1. **Business and Organisational Impact**
2. **Functionality Impact**
3. **RAMS Impact**
4. **Security Impact**

Table 1 summarises the benefits and challenges related to these categories at a high level, selected points underlined in Table 1, are expanded on in sections 3.1 through 3.7 below.

3.1 Testing and Commissioning

Virtualisation allows encapsulation of machines, this is the ability to record frequent images of a VMs state, these can be copied and moved between host machines. This facility provides several benefits, a test system that closely mirrors the live system can be maintained. This clone system can be used to develop and test software in a close approximation of the target environment.

Encapsulation simplifies testing and commissioning, the over backing of supervisory level systems can be completed almost instantaneously by stopping the live system and starting the test system, preconfigured with the software under test. This approach allows test windows to be used to maximum efficiency and fast reversion following commissioning if necessary.

This approach also allows operators to maintain a low-cost clone of the revenue system for testing and training. Security patches can be tested offline then quickly commissioned without risk to the revenue system.

3.2 System Reliability

Reliability is stated as a key benefit of virtualisation; however, railway systems are already designed to be highly reliable, with servers often operating in hot-standby configurations. Simply switching to the use of VMs does not guarantee improved reliability and if not implemented carefully, could reduce it. The reliability achieved by migrating the servers, workstations and other supporting supervisory equipment to a virtual environment will be dependent on the architecture adopted.

For all levels, consideration needs to be given to:

- VM Location - Hosting of VMs to remove single points of failure, i.e. don't host redundant servers on the same host machine. Use of hypervisor functions to allow fault tolerance.
- Communication Diversity – Ensuring diverse paths are maintained for physical and virtual network connections. Where level 3 is concerned diverse connections must be maintained between remote servers and local control equipment, this needs to include the route via national and international internet network.
- Thin Client Reliability – Thin clients present a reliability weak point as they may fail; therefore, spare terminals must be provided to allow for this. Control centres utilising thin clients are susceptible to some of the same issues as a conventional control centre, e.g. fire and loss of power.
- Interface to Legacy Equipment – This is a major challenge, legacy equipment within the control layer, such as interlockings or Remote Terminal Units may not support IP based communications. Examples of this are the use of serial communications (e.g. DNP3) or the use of Time Domain and Frequency Domain Multiplexing Systems (TDM/FDM). IP converters can resolve this; however, these impact reliability and may require alteration to software to incorporate drivers. Where possible control/actuation equipment capable of IP based communications should be used to avoid this.
- Hypervisor reliability – The hypervisor software must be highly reliable, as a failure could result in multiple VMs running on a host machine failing simultaneously.

3.3 Integrated Systems

Hosting multiple systems on shared machines simplifies the network interface, but without modification to the software, the systems will not be integrated and cannot take advantage of this. The cost of integrating existing software during midlife upgrades is prohibitive. Integration is easiest through entire system renewals; as suppliers offer new combined systems designed for integration which will be more cost-effective the altering existing systems.

Careful consideration is required when hosting systems on the same hardware. The loss of diversity between voice communications and signalling has implications on degraded modes and emergency operation. The deployment of systems, reliability requirements and failure/disaster recovery plans must be established to ensure intolerably safe scenarios do not occur.

3.4 Maintainer Competence

Maintainers will be required to be highly competent with the administration of the virtualised systems and networks. Within the IT field, a study found 37% system outages (Dimension Data, 2016) were the result of maintainer error. Given that virtualisation is not an area of competence traditionally associated with the railway, there is a greater potential for issues to effect service.

Maintenance and engineering teams will require significant upskilling to allow them to maintain, administer and fault find the virtual systems, jeopardising the business case for virtualisation. Where multiple systems are hosted on the same servers, there is the potential for clashes between maintenance teams that are traditionally separate, e.g. Signalling and SCADA.

Where systems rely heavily on virtualisation, creation of a network administration team to maintain the virtualised infrastructure may be beneficial. Existing rail maintenance teams can be limited to the control and actuation layers, avoiding the overlap between teams. This approach reduces the cost of upskilling existing maintainers and decreases the potential for errors, as the network administrators specialise in maintenance of the virtualised infrastructure.

3.5 Cybersecurity

A challenge for any new system is to ensure it is sufficiently secure. Historically air gapping systems from outside networks has been relied upon to help secure systems. However, the advent of STUXNET and the necessity for real-time access to data and integration with business management systems, means air gapping is becoming unviable. Air gapping also prevents systems using a level 3 style architecture, or an architecture where the host servers do not have a direct link to the control and actuation layer network.

Despite this, systems with a link to an external network can be secured and the adoption of a modern architecture allows us to take advantage of new countermeasures. Systems and sub-systems can be segregated in the virtual environment making the spread of intrusions more difficult. Thin clients provide an additional layer of segregation between the core systems and user terminals helping reduce the threat of attack from non-disabled USB ports. An external link to the internet allows the automatic download and application of updates to components such as firewalls, operating systems and antivirus tools. Also, the management of the system by skilled administrators provides advantages. They are capable of actively managing the systems, monitoring and acting on issues as they are detected, VM encapsulation allows an easy fallback position in case of attacks.

Virtualisation also provides challenges to security; issues such as VM sprawl, security of dormant VMs and lack of network traffic visibility must be actively managed. The security of the hypervisor software, as with any other Commercial Off The Shelf operating system, software or code must be assured.

Data security and security of systems in the cloud may require measures to be implemented by the cloud service provider and mandated within the SLA. System suppliers must be actively involved in the migration to a cloud-based platform as their propriety software will be hosted on third-party servers.

The challenges introduced by levels one through three are manageable threw risk assessment and deployment of appropriate mitigations.

3.6 Level 3 - Remote Links and Latency

Implementation of level 3 requires a link to the internet. In the UK dedicated high reliability leased line links are available from ISPs. These use two diverse routes to separate exchanges (provided two exchanges are in the proximity of the railway) (British Telecom, 2018). The dedicated links guarantee bandwidth and roundtrip latency times to the core network in the UK are <20ms, but the round-trip time to the host servers could be greater (≈100ms).

The UK core network is highly redundant and has a minimum of 4 potential communication routes between nodes. BT state that their core network has been 100% available over the last 20 years, therefore the vulnerable component of the architecture is the link to the core network from the exchanges and the cloud service provider's data centre. Often data centres are co-located with nodes known as Points of Presence (PoP) on the core network to account for this. When selecting the cloud service provider, infrastructure owners should ensure the data centre is located on the same core network as the control centre and is co-located with a PoP. An acceptable reliability should be achievable in locations with a robust core network such as the UK.

The introduction of the additional latency may cause issues, options are available for low latency links - usually used by financial services - but the cost is prohibitive. To account for this, systems may require alteration to allow for increased latency between the supervisory and control layer, i.e. alteration of timeout timers or movement of time-critical functionality to the control layer if the latency cannot be accommodated. For example, it is undesirable for Communication Based Train Control (CBTC) system to locate the sub-system responsible for issuing movement authorities to trains remotely, via a high latency link, as the delay will impact capacity and could introduce issues.

3.7 Level 3 - Service Provision and Third-Party Reliance

Provision of critical systems as a service presents a variety of challenges to infrastructure owners:

- Procurement model – Infrastructure owners are accustomed to procuring systems through a model that requires a large capital investment followed by an ongoing service contract. Use of services or a complete shift to systems as a service, alters the cost profile for systems. While this may seem beneficial, it will require a shift in the way railways manage their capital investment programmes and procurement mechanisms. This may hamper the introduction of the concept of systems as a service.
- Reliability/Availability - Standard SLAs provided by cloud service providers specify an availability figure of 99.99%, which equates to approximately 52 minutes for downtime per year. However, the planned downtime is not limited to 'Engineering Hours', therefore an agreement is required to limit planned maintenance to non-operational hours. This may require the provision of dedicated host machines within a data centre. The SLA should include requirements for the distribution of VMs over host machines to provide diversity should an issue occur.
- Time to Repair - Should issues occur with the cloud services or leased line links, the infrastructure owner is reliant on the third parties to resolve the issue quickly. While this can be accounted for within SLAs, the benefit of railway maintainers being responsible for infrastructure is the control and awareness of the criticality of resolving issues. Utilising suppliers who are not conversant with the operation railways may result in them not applying the urgency required resolving issues. It is not hard to imagine an issue where the control room can't connect to the cloud. Everything in the control room looks ok and the ISP denies any problem with the leased line as does the cloud host. In this situation, the railway maintainer has limited control to resolve the issue and is reliant on suppliers sending engineers to site for further investigation. BT provide a 5-hour response time to resolve issues with leased line links, Amazon AWS can provide a 15-minute response time. However, the financial penalties should these SLAs not be met are small, meaning the suppliers incentive to resolve issues quickly is not proportionate to the impact of a non-operational railway. It is unlikely service providers will accept a greater liability for downtime without a significant increase in service cost, specific SLAs will need to be created to support railway services but will be expensive jeopardising the business case for the use of the 'cloud'.

		BUSINESS IMPACT	FUNCATIONALITY IMPACT	RAMS IMPACT	SECURITY
LEVEL 1 - VIRTUALISATION OF ISOLATED SYSTEMS	Benefits	**Limited capital investment** - Limited alterations required to software to migrate to the virtualised platform. **Reduced whole life cost** – The number of physical machines/hardware is reduced, reducing maintenance cost.	**Scalability** - VMs allow additional workstations/servers to be deployed easily. **Backup control facilities** - Thin clients and new VMs can be deployed cost-effectively to provide low cost back up control facilities. **Testing and commissioning** - Virtualisation allows fast switchover between test and revenue configurations.	**Obsolescence management -** VMs are hardware agnostic; host server replacement is simplified and can be performed independently from the railway. **Training system** - A clone of the revenue system can be taken to test system updates/security patches and train staff. **Reduced maintenance** – Fewer physical machines to maintain and manage spares for. **Availability improvement -** Reduced maintenance and testing, decreases downtime increasing the system availability.	**Security updates -** Security patches and anti-virus definitions can be rapidly tested on a clone system and commissioned quickly on the railway. **Roll-back –** Encapsulation allows regular images of the system to be taken; these can be used to roll-back the system in the case of ransomware attacks or other issues.
	Challenges	**Maintainer competence** – The skillset required to maintain virtualised systems is significantly different to traditional equipment.	**Functionality** – The migration to a virtualised environment should not impact the functionality.	**Reliability** -The system architecture requires careful consideration to ensure reliability is not degraded as a result of migration to the virtual environment. **Hypervisor reliability –** For all levels, the hypervisor software needs to be assured to be sufficiently reliable to support revenue operation and be supported over the life of the system.	**Configuration control –** As with software, the VM images require careful config management to avoid deploying incorrect versions. **System access points** - Deployment of many thin clients increases the number of access points to the system; these will require locking down to avoid security vulnerabilities.

		BUSINESS IMPACT	FUNCATIONALITY IMPACT	RAMS IMPACT	SECURITY
LEVEL 2 – VIRTUALISATION OF MULTIPLE SYSTEMS	Benefits	**Reduced whole life cost -** Resource pooling and reduction in hardware maintenance should reduce the whole life cost of supervisory hardware.	**Integrated systems** - Hosting control systems in the same cluster simplifies the interface. Should suppliers and or clients wish, systems can be integrated to provide shared HMIs and functionality based on the inputs from both systems. **Combined data warehouse** - Data from multiple systems can be combined easily for monitoring and reporting.	**Improved reliability** - Use of hypervisor functions such as 'Fault Tolerance' and 'High Availability' allows VMs to perform automatic restarts and instantaneous failovers to mirrored VMs. **Improved availability** – Resulting from the increased reliability and maintenance benefits outlined for level 1. **Network Reliability -** Virtualisation of the network reduces the opportunity for hardware failures.	**Network micro-segmentation -** Virtualisation of the network allows the use of micro-segmentation, helping to prevent the spread of security issues or intrusions between systems. **Reduction in attack surface** – Virtualisation of the network hardware (firewalls, switches, routers etc.) reduces the system's available attack surface.
	Challenges	**System lifecycle alignment** - Systems may not life expire at the same point, therefore staged migration is required, increasing testing requirements and hardware costs in the interim. **Overlap of maintenance teams** - Historically maintenance teams for systems have maintained separate hardware. Use of a shared host server introduces and overlap. The competence requirement to support the virtualised systems also increases.	**System integration cost** – To achieve functionality benefits from the combined hosting of systems clients must invest in the integration of the supervisory software. **Reliance on serial communication -** Applicable to all levels, to support the migration of the supervisory layer to a virtual/cloud environment, the systems must migrate away from the use of legacy equipment such as TDM and serial protocols and move toward IP based technology.	**Loss of diversity -** Virtualisation of systems requires careful consideration of the failure modes, to ensure safe recovery from degraded and emergency modes, e.g. Voice Comms and Signalling. **Host network connection reliability** – The host server's network links must fast enough to the support mirroring servers and the reliability requirements of the overall system. **Fault finding software-centric systems** - Fault finding requires the use of server and network diagnostics, see competency challenge.	**Lack of Encryption** - Some virtualisation platforms do not encrypt traffic between mirrored VMs hosted on different servers; therefore additional hardware is required to secure link and prevent man in the middle style attacks. **Increase in Users** – Hosting multiple systems increases the number of users and hence the security threat. System deployment needs to account for this, locking down thin client terminals and limiting access to servers.

		BUSINESS IMPACT	FUNCATIONALITY IMPACT	RAMS IMPACT	SECURITY
LEVEL 3 - MIGRATING TO THE ' CLOUD'	Benefits	**Environmental impact** – Host servers can be located in regions with sustainable power.	**Scalability –** The limit on scalability is removed, as the third party's datacentre is much larger than the railway system's requirements.	**Server Location** - Locating hardware away from rail environment reduces electromagnetic immunity, vibration and environmental requirements. **Maintenance reduction** - Supervisory layer hardware is removed; however, network hardware and converters to interface with equipment in control/actuation layer utilising serial comms will require maintenance.	**Security updates** - Layer 7 firewall, anti-virus and security patches can link to online databases and be kept updated in near Realtime.
	Challenges	**Service provision –** Infrastructure operators are required to invest in services, VM hosting and internet, for the life of the contract as opposed as an initial capital investment followed by ongoing maintenance. Should rail industry suppliers wish to switch to a service provision model, e.g. 'Signalling as a Service' this switch from capital to ongoing operational investment will be greater. **Third-party reliance** - Infrastructure operators need to ensure services are guaranteed for the life of the system. Service contracts must meet the availability and maintenance requirements to support revenue operation.	**Latency -** Latency between cloud-based supervisory layer and control/actuation layers will require testing and potential alterations to software to support additional lag. **System distribution** – Decentralising equipment, distributing the control/actuation level makes the provision of power backup facilities (Uninterruptable Power Supplies) difficult. Accommodation of hardware; maintenance and testing of fallback positions can become complex and time-consuming. **Third-party compatibility** – Systems may require redesign to support hosting in the cloud environment. All systems will require testing.	**Internet link** - A designated link will be required to connect to the 'cloud'. This introduces a challenge to ensure the availability and latency is sufficient to support revenue operation. **Time to repair** - Should failures occur with the third-party services the infrastructure owner is reliant on the third parties to resolve the issue quickly. The level of service required may not be available or may be prohibitively expensive. There is also a political question of whether it is acceptable to have the operation of a railway reliant on a third party. **Network redundancy** – Must be guaranteed over 40yr system life, largely reliant on a third party to ensure this is maintained.	**Increased attack surface** - Systems can no longer be air-gapped to internet increasing the systems attack surface; however, the design of the system can mitigate against intrusions. **Remote servers** - System suppliers must accept the location of software on third-party servers for cloud deployment. However, if the system is being provided as a service, this will not be an issue, as the supplier will provide the servers.

Table 1 – High-Level Consideration of Benefits and Challenges of Each Level

4 CONCLUSION

Virtualisation is commonplace within industrial control systems and is increasingly used within the rail industry for systems such as condition monitoring and TMS but is not the industry norm. The virtualisation of control centre supervisory systems can provide a wide array of benefits for the maintenance of the system; allowing frequent backups, simple creation of test system, accelerated development, testing and commissioning. It does introduce challenges, but these are surmountable. While virtualisation itself will not inherently increase the reliability, the reduced downtime arising from planned maintenance improves overall availability. Therefore, it is recommended when brownfield railways look to renew the hardware of supervisory systems virtualisation in a level 1 configuration is considered.

Virtualising multiple systems on the same host machines is not common, but suppliers do offer solutions. The main benefits are the reduction in physical hardware, the simplification of the interface between systems and the greater virtualisation of network components. The use of advanced virtualisation features such as 'Fault Tolerance' can also provide small reliability gains. Brownfield railways' systems are normally renewed individually, as they reach life expiry and a staged migration of systems is required to transition to a shared virtual platform. To take advantage of the shared platform software development is required to integrate the systems, e.g. a single HMI showing signalling, SCADA, CCTV and traffic management data. The cost to implement this migration is unlikely to justify the benefits for brownfield railways. However, greenfield railways may wish to consider this approach when developing their requirements and considering their approach to apportioning contracts to allow for the procurement of an integrated solution.

Reliable solutions for leased line internet connections and cloud computing are available meaning a highly available system could be created using this architecture. However, the increased latency is a challenge requiring the development of systems. Where equipment in the control and actuation layer does not support IP based communication the cost and reliability impact of deploying converters may make the solution unviable. The commercial challenges of a service-based approach are also large. Infrastructure owners have investment and procurement models suited to large capital investment, not ongoing service provision. Also, traditional SLAs for cloud and network providers are structured towards the IT industry, not the railway. The cloud-based architecture has significant challenges and limited benefits and therefore, this approach is not recommended for most railways. However, the model may work for smaller railways wishing to avoid large capital investment or individual systems such as condition monitoring that need to be deployed quickly and cheaply.

A significant challenge to any level of adoption of virtualisation is the competence and structure of the traditional maintenance workforce. Any sizeable deployment of virtualised infrastructure will require a dedicated team competent in the configuration and management of the infrastructure. The core competence required to maintain virtualised equipment lies further within the field of network administrators than rail maintainers. Adding network administrators to the existing maintenance teams to support the virtualised elements is likely to be less challenging than upskilling the existing maintainers. This challenge could also be surmounted by rail suppliers providing systems as a service and ensuring their team maintains the required competence.

Modernisation and increased connectivity inevitably raise the issue of security, virtualisation provides benefits in this regard through micro-segmentation of VMs, automatic updates and reversion in case of major incidents. Virtualisation aside, we can no longer rely on an air gap to protect against intrusions, connection of systems to the internet allowing access to their data and real-time security updates will become the norm.

Overall, virtualisation offers a range of benefits and can improve the availability and resilience of our systems, while the challenges it introduces can be managed. The decision whether to use virtualisation will be specific to each railway's individual needs. What is right for a small low capacity line will differ to that of a major metro system. The large-scale adoption of cloud computing appears to be a step too far for our critical systems at present, but as IT infrastructure changes with the rollout of 5G and increased intelligence of equipment in the actuation layer, the case for migrating systems to the cloud may change.

5 REFERENCES

The following references have been used in the creation of this paper:

1. Amazon AWS, 2019. *Amazon AWS Customer Support Info.* [Online]
   Available at: https://aws.amazon.com/premiumsupport/plans/
   [Accessed 20 May 2019].
2. Amazon AWS, 2019. *Amazon AWS Monthly Calculator.* [Online]
   Available at: https://calculator.s3.amazonaws.com/index.html#s=EMR
   [Accessed 20 May 2019].
3. Amazon AWS, 2019. *Amazon AWS Service Level Agreement.* [Online]
   Available at: https://aws.amazon.com/compute/sla/
   [Accessed 20 May 2019].
4. British Telecom, 2018. *BTNet Customer Service Description,* s.l.: British Telecom.
5. British Telecom, 2018. *BTnet Resilience,* s.l.: British Telecom.
6. British Telecom, 2019. *BTnet – Service Level Agreement,* s.l.: British Telecom.
7. Dimension Data, 2016. *Network Barometer Report.* s.l.:Dimension Data.
8. GE, 2016. *Tacoma Power Reduces Costs and Risk with HMI/SCADA Virtualization,* s.l.: GE.
9. Herrmann, P., 2014. *Fault-Tolerant Virtualization,* Prague: Charles University in Prague.
10. RUOTSALAINEN, J., 2017. *HARDENING AND ARCHITECTURE OF AN INDUSTRIAL CONTROL SYSTEM,* s.l.: Tampere University of Technology.
11. VM Ware, 2016. *Why Businesses Are Adopting Network Virtualisation,* s.l.: VM Ware.