Interoute Virtual Data Centre and Security

1 Cloud Computing – Creating an on-tap secure ICT service platform

The allure of cloud computing and cloud services is by now well documented. It offers a direct relationship between consumption and payment and can respond to demand curves in real time. It looks set to radically change the way in which IT managers think about their computing and services in future. Unfortunately most of the gains in Cloud Computing and mainly its convenience have been achieved outside longstanding IT security practices and the trusted enclave of the private network or virtual private networks based on MPLS WAN. This 'compromise' the trade-off and sacrifice of the trusted security model, in return for a far more cost effective and dynamic model has set the tone for the debate on cloud computing thus far. This debate however has largely taken place without consideration of how 'the network' can influence the development of a flexible computing capability without the need to compromise security.

Ignorance or avoidance of the influence of the network as a method of securing services comes largely from a traditionally held belief in the "stupid network" concept, i.e. the Internet and maintaining its simplicity and therefore accessibility. The alternative many believe is a costly private network and inflexible private computing offer which looses it flexibility and inefficiency through its securing. This perception is often a result of the genesis of the public cloud computing suppliers many who have very limited control or desire to control their networking infrastructure and have little or no real experience in delivering WAN services to customers. This Amish like preservation of tradition of separating computing and connectivity ignores a fundamental principle that even the most technologically illiterate would recognise. In the really physical world of everyday we define security around buildings in terms of their access, roads big or small, private, public or toll. i.e. control access and you control security.

This whitepaper describes a model whereby the long trusted, flexible and cost effective MPLS VPN forms the foundation of a platform from within which high agile ICT services can be created and securely used. It combines the convenience, immediacy and flexibility of the public proponents of cloud computing or communication with the trusted, well established and auditable model of MPLS VPN security model.

It goes onto describe in detail the underlying platform that offers secure separation of traffic and data and how services like utility cloud-computing can be securely added into this environment. It further describes how this trusted security model is extended through to the computing fabric creating a hybrid cloud-computing platform with all the attributes normally associated with typical public cloud services.

2 Principles of Providing Secure Private Communications Networks and the services built within them

Interoute Unified ICT set of services serves to offer the customer an on-tap scalable infrastructure that they can adapt and use. In the provision of its Unified ICT suite, including private networking, computing and communications, Interoute recognises the demands of customers with regard to information security. To satisfy these demands, Interoute places significant focus on three main tenets associated with information security, and how the available computer and network technologies enable a service provider to provide assured service in that regard.

- **Confidentiality/Privacy:** Data is available only to those authorised.
- **Integrity:** It is not possible to manipulate data without authorisation.
- **Availability:** Authorised access to data within the customer domain is available.

2.1 Principles

To achieve a realistic, practical assurance of these fundamental tenets, Interoute establishes some basic principle functionality that is sought across all of the technology domains it operates:

- Logical Separation
- High Availability

Separation

Confidentiality and **Integrity** are both addressed by isolating and separating traffic and allowing it to exist only within the scope of the owning customer organisation[1]. This separation makes it impossible for traffic from one customer domain to enter another customer domain. This prevents data leakage, and it also prevents interference by entities outside of the organisation.

Network separation can be achieved at many layers of the network infrastructure stack, but some technologies are more amenable, practical and cost-effective than others. For example, it may not make economic sense to dedicate physical optical fibres to different customers.

High Availability

Availability in line with customer expectation is addressed by the engineering of underlying platforms so that they meet the necessary standards of operation, longevity and failure tolerance.

Challenges to continuous operation may include:

- Operator error, or software error: a lack of diligence or quality control;
- Deliberate subversion: the malicious and motivated exploit of vulnerable systems,
- Natural physical events: component age, weather, acts of God, or other uncontrollable inputs,
- Unintended coincident activities: physical fibre network breakage.

All of these threats may result in failure and compromise of the provided service. The engineering of a resilient system involves the consideration of each of these factors and the selection of a technology and operating method that mitigates the risk of each these threats.

Sometimes individual systems and components cannot satisfy the availability requirements in isolation, and in these cases multiple systems and resources must be combined in parallel, along with mechanisms to detect a working sub-system in order to deliver the necessary service levels.

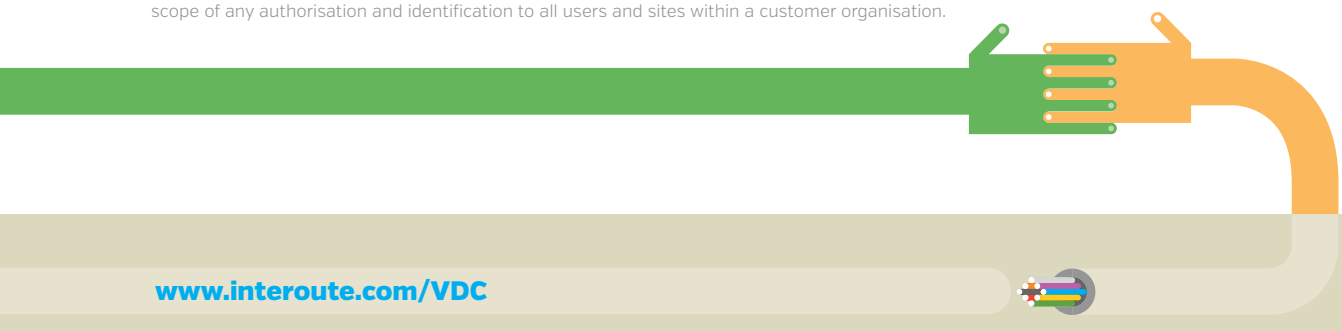
For the over-provision of redundant components in this manner to achieve the expected results, it is usually necessary for different sub-systems to be implemented in different ways, using different technologies, or underlying physical resources, or to be otherwise made diverse from each other so that a single root input event is unable to prevent the system from operating.

Different technologies provide different options for achieving availability through resiliency and redundancy.

2.2 Implementation

Interoute addresses the fundamental tenets of Information Security for managed network and compute services that it operates by successfully implementing Logical Separation and High Availability across **all of the technology domains that it operates**. These principles, applied to a technology, make use of a design and an available implementation. Confidence in both is required.

1. While computer systems provide many levels of authorisation, authentication and identification, a service provider typically has a very clear relationship with a customer organisation, rather than entities or departments or individual users within an organisation. It is reasonable in this situation to define the scope of any authorisation and identification to all users and sites within a customer organisation.



Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

Ethernet Access Aggregation Platforms

Within metro-area networks, Interoute operate IEEE 802.3 standard Ethernet switching networks. Cost-effective layer-2 switching provides an assured forwarding performance within a metropolitan area domain (where exact routing path is not significant since delay and loss characteristics are comparable)

Logical Separation: Implemented through the use of:

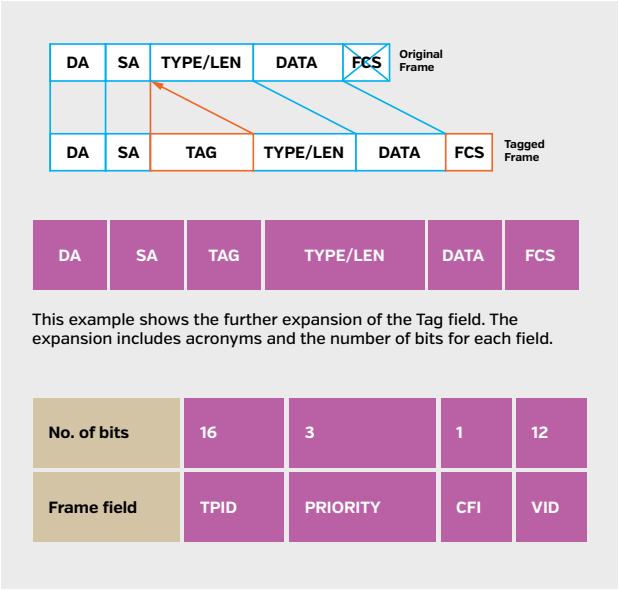
- IEEE 802.1Q VLAN tagging of Ethernet frames;
- VLAN-aware TCAMs for Ethernet switching.

The IEEE 802.1Q standard establishes the common practice of encapsulating multiple logical Ethernet networks on shared infrastructure. Originally pioneered by Cisco Systems in the form of the Inter-Switch Links for large-scale Ethernet networks, the IEEE standard provides a vendor independent framework for supporting multiple logical networks on shared Ethernet switch devices and shared link media.

The on-wire Ethernet frame format is augmented to include an additional composite field that allows the encapsulation of the following extra information:

Virtual LAN identifier: used by service providers to differentiate logical functions, or customers;

Priority: a frame priority class, which allows traffic conditioning and queuing at congestion points in the network.



Modern switch architectures such as the following include implementations of IEEE 802.1Q frame switching using hardware-based TCAMs (Ternary Content-Addressable-Memory) that are VLAN aware. This means that the same Ethernet-layer MAC address can co-exist in multiple Virtual LANs (or customer domains) without any compromise of forwarding performance or ability.

Historically, the implementation of Ethernet networks based upon VLAN-based logical separation has attracted concerns from critics concerned with VLAN hopping, the practice of Ethernet frame injection with payloads crafted in such a way that it is possible to "jump" between VLANs (or customer domains).

Modern Ethernet switch implementations include many counter-measures to defy this as a feasible attack vector, however. At Interoute, the following features are used to ensure confidence in the ability of Ethernet Aggregation platforms to maintain logical separation of customer domains:

- 802.1Q "tunnelling" functionality, on Cisco switches, which results in mandatory, not conditional VLAN encapsulation, eg. a port associated with VLAN 300 is always tagged as such on ingress, even if already possesses a VLAN tag.
- Mandatory tagging on all trunk links, which outlaws any untagged frames (erroneous by design) on links within the service provider network.

High Availability: implemented through the use of:

- IEEE 802.1s (Multiple) Spanning Tree Protocol
- Constrained/Partitioned VLAN-based MAC tables ("Port security")
- Traffic rate policing

Native Ethernet networks include no built-in mechanism for loop-detection or hop count, and combined with the default forwarding action of sending a frame to all ports, this results in an inability to operate Ethernet networks with deliberate redundant links. But such redundant links are essential for providing continuous, resilient and fault-tolerant service in the face of real physical-layer events.

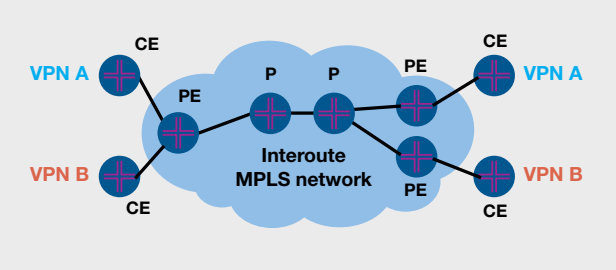
Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

The use of the IEEE 802.1s specification provides a modified version of the traditional Spanning Tree protocol defined in the original IEEE 802.1d Ethernet bridging specification. It builds upon the functionality of the original protocol that provides a practical loop-free topology from a set of redundant links and enhances it by providing faster failure detection and convergence and allowing operation in a VLAN transparent manner.

Partitioned VLAN tables, or Port-Security, is a vendor-specific enhancement implemented by Cisco which involves the artificial capping of learned MAC addresses within a certain port or VLAN. The most significant benefit is that the service provider is then able to divide up the delicate CAM resources of the switching network to multiple customers with the assurance that no single customer can over-burden the switching platform by connecting excessive MAC addresses to the service.

Traffic rate policing ensures that the service provider is also able to partition sections of bandwidth to different customers, dividing up the physical bandwidth resources in a meaningful manner.

WAN Routing Platforms



Over wide-area networks, where latency characteristics can vary, Interoute makes use of IP/MPLS based router devices in order to forward customer traffic to common wide-area destinations, while retaining logical separation.

Logical Separation: Implemented through the use of:

- MPLS encapsulation, IETF RFC 3031
- BGP-based MPLS VPNs, IETF RFC 4364
- Virtual Routing/Forwarding Tables

MPLS VPN technology introduces the notion of a Control Plane and a Data Plane (or Forwarding Plane). The interaction of both components is necessary for the operation of a customer's service and the application of logical separation to both must be considered.

Customer networks are separated into groups of sites. All sites that are permitted to communicate with one another (typically all sites within a single customer organisation) are associated with a Virtual Routing/Forwarding table.

This can be thought of as a portion of a router's routing table, and the IP-layer equivalent of an Ethernet switch's VLAN-aware TCAM forwarding table. The VRF retains routing information, **specific to a single customer**, and a single service provider PE ("Provider-Edge") router may accommodate a large number of VRFs, each one associated with different customers.

Network Destination	Netmask	Gateway	Interface	Metric
0.0.0.0	0.0.0.0	192.168.0.1	192.168.0.100	10
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
192.168.0.0	255.255.255.0	192.168.0.100	192.168.0.100	10
192.168.0.100	255.255.255.255	127.0.0.1	127.0.0.1	10
192.168.0.255	255.255.255.255	192.168.0.100	192.168.0.100	10

Crucially, each of these customer-specific VRFs may contain overlapping address space. This is a fundamental enabler in allowing a service provider to provide private network connectivity for multiple customers' private IP networks without clumsy NAT or explicit tunnelling technology[2].

Customer interfaces are associated with the customer's own VRF, and this prevents any traffic crossing customer domains (shown as blue and red in the diagram) by way of the service "Provider-Edge" router.

Beyond the customer-service provider links, it is necessary to preserve the logical separation of customer traffic on the service provider's own internal links. For instance, both the Blue VPN A customer and the Red VPN B customer may be communicating with an IP host 10.0.0.1. Conventional IP routing forwarding strategies would not preserve the logical separation and an alternative mechanism is required: MPLS encapsulation of customer traffic.

2. Since MPLS is an encapsulation technology, it can be considered a tunneling technology but the fact that the encapsulations are dynamically signaled using co-operating protocols usually blurs this fact.

Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

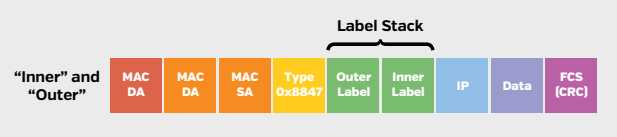
By encapsulating customer traffic in MPLS, the service provider is able to take several benefits:

- The removal of the customer's own routing identifiers from influencing forwarding decisions on the backbone
- The encapsulation of end-point routing information, after a single routing table lookup.

The first benefit is significant since it also assures us that the logical separation is providing Confidentiality and Integrity. If the format of a packet sent by Blue VPN A customer cannot influence the service provider's backbone beyond routing decisions made in the Blue VPN's VRF, there is no way that a user on the Blue VPN can direct traffic outside of his own customer domain, and in to that of other customers.

Typical MPLS encapsulation on MPLS VPN enabled backbones includes the use of two MPLS "labels":

- Outer label: specifies the egress PE to which traffic should be routed,
- Inner label: specifies the VRF table on the egress PE.



High Availability: implemented through the use of:

- IS-IS and LDP interior routing protocols,
- LACP Ethernet control protocol,
- BGP exterior routing protocol.

IP/MPLS networks have many features that facilitate a large-scale network to operate at high bandwidth and with a level of redundancy and resilience appropriate to wide-areas where the risk of physical transmission link compromise is significant. Interoute makes use of the necessary minimal subset of these features in order to provide an assured level of availability to private network customers while still making the economic gains associated with a shared network asset.

IS-IS and LDP operate in order to propagate essential information about network topology and MPLS label assignments respectively throughout the whole service provider network. LACP provides fast detection of interface failure on a link-by-link basis.

In the face of an internal link compromise, the failure is usually detected by LACP if actual absence of optical laser is not obvious. Using information available with the IS-IS and LDP databases, any affected routing devices are able to make immediate recalculations in order to forward traffic across alternate paths to their destinations.

This internal reroute is virtual transparent to end-users whom, at worst, may notice a small change in latency as traffic takes on the alternate route.

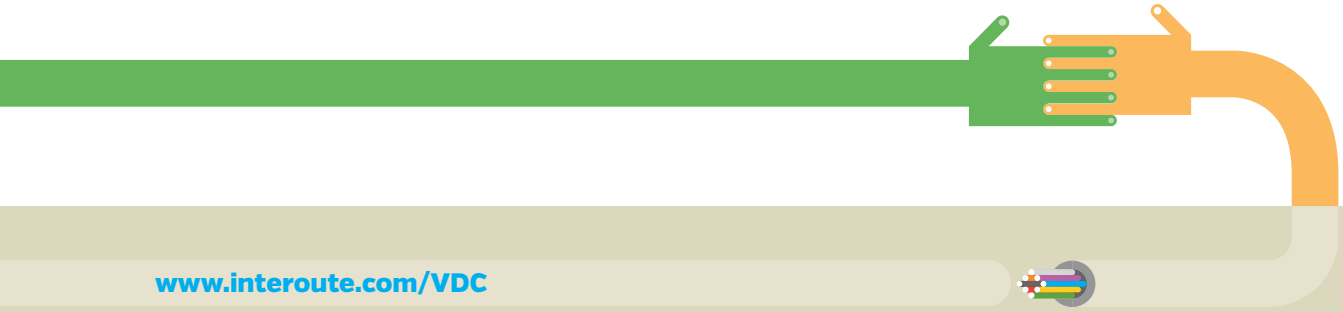
BGP brokers communications between customers' private networks and the shared service provider backbone that connects them. In the face of external (customer) link compromise, BGP propagates the news of the failure to all other routers in the service provider network, which in turn inform other customer routers. This allows customer site routers to make alternate routing decisions to complete their transfer.

The net effect of both events is that a customer's private network service remains unaffected long-term in the face of:

- Physical equipment failure: whether due to natural failure, or intentional abuse,
- Interior link failure,
- Exterior link failure – if a customer has alternate connectivity options.

2.3 Securing the Interoute Virtual Data Centre

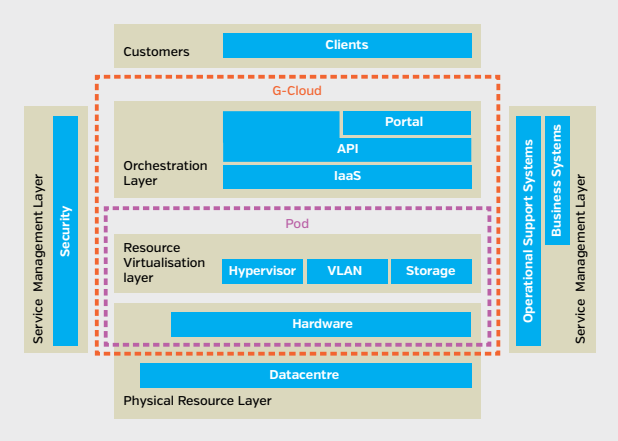
Unlike most cloud computing services Interoute's Virtual Data Centre (VDC) is neither solely a private or public cloud computing solution. Interoute has built the capability as a service of Interoute's MPLS/IP backbone that allows it to serve as either a public Internet based cloud or replace internal private clouds. In the same way that Interoute customers buy either or both Internet or MPLS VPN services Interoute has applied the same approach to its cloud-computing offer. Instead of it being simply a virtualised computing or storage capability it includes virtualisation at the network layer. It therefore means that "it's the network" that defines the audience not the way in which the capability was built. Interoute has literally made the network the computer.



Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

Integration of VDC into the Interoute MPLS/IP network

VDC works by allocating unique VLANs (within an Interoute specified range) to a customer. These VLANs extend down to the partitioned resources within the physical infrastructure of VDC called a pod, see illustration below (pod = switch infrastructure connecting compute and storage hardware). South of the switch is another separate range of VLANs that the customers can acquire on-demand to build their own network topology. At no part of the process are customers VLANs exposed to the customer these are managed by the VDC orchestration software.



Customers manage their virtual appliances/machines through one of these customer VLANs, most commonly, an MPLS VPN. Customers then have the choice of adding another VLAN, through the abstraction of adding "VDC Networks". These can provide access to the Internet or else just a "Private" network that only appliances and machines in the customer's VDC can utilise. In this way VDC mimics exactly the creation of physical DC where the switching infrastructure defines whether computing is exposed to public addressing or it is concealed.

VDC integration into the network - Guest process separation

Having concluded that the integrity of the network preserves data often discussions regarding cloud-computing shift the focus to the integrity of the hypervisor and the relationship between the hypervisor and the physical hardware.

Interoute for most of its primary IaaS services uses the XEN Type 1 hypervisor[3]. One of the main benefits of XEN is it ensures guest process separation by utilizing virtualization specific processor instructions. These ensure the isolation of guests from the hypervisor and each other.

Guest isolation is a key security feature, particularly where multiple clients maybe serviced by a single software instance, the hypervisor. The hypervisor must be protected against security breaches involving guests operating on top of the hypervisor. These security issues include:

- Guests bypassing security controls to access either the host or other guests in ways that violate the host's security policy
- Guests intercepting client data or host resources to which they are not authorized
- Guests attempting or becoming the victim of security attacks, which could possibly take down the entire platform.

Additionally client data must be protected from unnecessary access from the hypervisor itself, and guests need the capability to create controlled shared storage for collaboration purposes.

XEN guest processes are subject to all the usual user space process separation that is integral to the Linux kernel's operation. However, the most basic protection mechanisms have existed since early in the development of the Linux kernel, and are well tested and certified. On x86 systems, the kernel, at the lowest level, uses the central processing unit (CPU) chip set hardware to achieve separation between user space mode and kernel (privileged or supervisor) mode.

Hardware-based isolation

XEN uses virtualization-specific processor instructions to ensure isolation of guests from the hypervisor and from each other. Intel's virtual machine extensions (VMX) and AMD Secure Virtual Machine (SVM) instructions add a third level of isolation and protection by running guests in a restricted (guest) mode.

3. Interoute can support multiple hypervisors its default hypervisor unless specified otherwise is XEN

Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

Type 1 Hypervisor bare metal virtualisation

XEN is classified as a Type 1 hypervisor since it translates physical to virtual resources once. XEN runs in privileged mode and directly uses hardware instructions to virtualize the guest. Benefits are that guest code executes most of the time directly on hardware at full speed and most importantly, the virtual-to-physical resource translation occurs just once, unlike in type 2 hypervisors. Regarding these features, XEN is equal to other bare metal hypervisors.

Aside from the inherent separation hypervisor integrity and availability is tracked by Interoute's server monitoring, any events that occur on the hypervisor are logged off-server to a centralised logging facility and immediately visible to the Interoute operations centre. Together with the hardening process that is implemented on all physical machines, any change in hypervisor integrity or availability will be identified, isolated and remediated in a short time.

Storage Isolation

Another potential concern for users of a cloud-computing platform is the level of storage isolation and the subsequent customer data security. Within VDC there are two storage types that may contain customer data. Both storage types are built on top of the TCP/IP protocol allowing the networking security principles and design to apply, i.e. VLAN isolated. Furthermore, both the iSCSI and NFS networks are isolated by presentation to each hypervisor on separate physical network interface cards (NICs). Consequently, any virtual machine created within the VDC has no direct means to access or address the underlying NFS or iSCSI networks.

Within Interoute's VDC, customers are provided and recommended to use the external block storage to host data sets. This is provided to each virtual machine as a volume, which is in fact, a unique, customer allocated, iSCSI LUN.

When discussing security of iSCSI, as with the network discussion, 3 similar principles apply; Authentication, Authorisation and Data Integrity, and these can be applied to data at rest (DAR) and data in flight (DIF).

For DIF, iSCSI authentication refers to the login phase authentication of the hypervisor initiator, or the mutual authentication of the initiator and target. Interoute has implemented Challenge Handshake Authentication Protocol (CHAP) to secure the connection. CHAP uses one-way, three-phase authentication for targets authenticating initiators and for initiators authenticating targets.

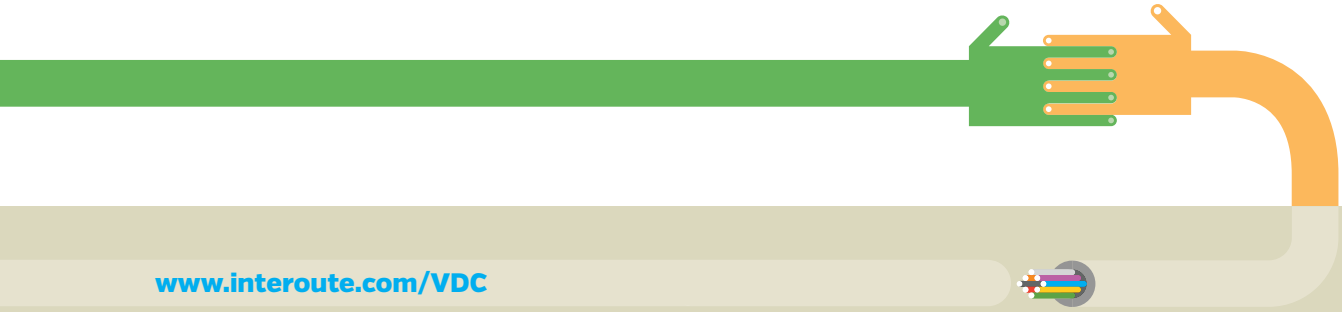
DIF Authorisation is achieved, indirectly by the VLAN isolation of the iSCSI network. Only clients connected into the iSCSI VLAN can connect to the target initiator, thereby

utilising access control, as implemented in the Interoute networks (ACLs). This provides authorisation of communication channel access based on information contained in packets - IP address, iSCSI Qualified Name (IQN), virtual LAN (VLAN), and other parameters.

iSCSI data integrity is provided at the Ethernet, TCP, IP, and iSCSI level. In Ethernet, the checksum is also called the frame check sequence (FCS) and is a 32-bit cyclic redundancy check (CRC). The iSCSI checksum is also called the iSCSI digest, and covers multiple iSCSI packets.

For DAR, the VDC design allows the customer complete control over the authentication, authorisation and integrity of the data. iSCSI LUNs are provisioned on demand for each customer, by each customer. The LUNs are separate volumes consisting of RAID volumes on the underlying storage hardware. Data stored on the external block storage is made available as volumes to the virtual machines. Once attached to the VM, the OS and application stack are then provided and managed solely by the customer. i.e. Interoute have no administrative access to an OS or any application within.

On this external block storage platform each volume created by the customer is attached to any virtual machine at boot time and seen as a block device. The customer can format this block device in the same way traditional sys-admins add additional drives/disks to physical machines. In this fact it is quite possible for the customer to use an encrypted file system (EncFS, eCrypt, etc.) for the data. This can therefore provide the customer with highest level of data integrity possible as keys to the encryption are created and held by the administrator of the virtual machine and that in turn is secured by normal means by the customer.



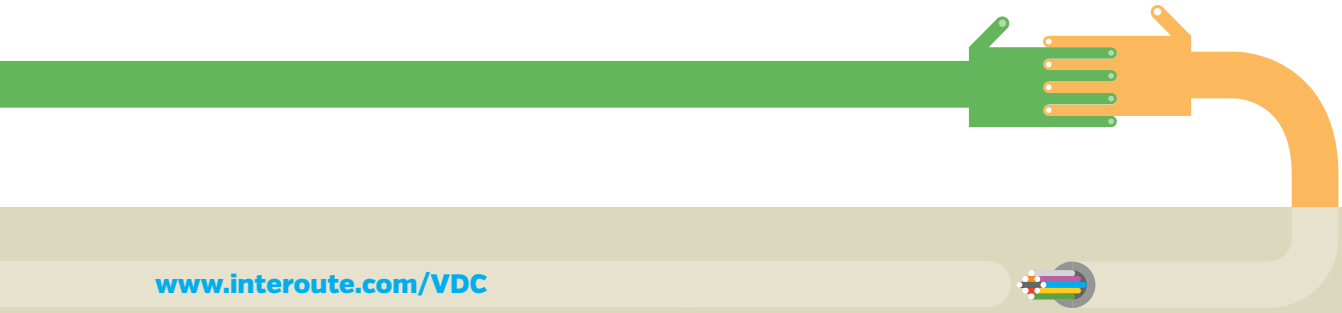
Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com

3 Conclusion

Interoute Virtual Data Centre maintains integrity from the VLAN through to storage virtualisation and the hypervisor. In this regard it is the network that secures the service. Interoute is currently unique in pre-providing the VPN integration directly into the VLANs of the Virtual Data Centre.

This extension of network separation extends therefore from the core of the network straight through to the hypervisor and the 'guest'.

By basing development and engineering efforts on establishing the crucial parts of technology necessary to implement core principles of logical separation and high-availability, Interoute creates a base platform upon which managed service customers can be assured of the **confidentiality**, **integrity** and **availability** of valuable private data.



Interoute Communications Limited 31st Floor 25 Canada Square Canary Wharf London E14 5LQ
Telephone +44 020 7025 9000 **Email** info@interoute.com