Withdrawn Draft

Warning Notice

The attached draft document has been withdrawn and is provided solely for historical purposes. It has been superseded by the document identified below.

Withdrawal Date	November 17, 2022
Original Release Date	August 5, 2022

Superseding Document

Status	Final
Series/Number	NIST Special Publication (SP) 800-215
Title	Guide to a Secure Enterprise Network Landscape
Publication Date	November 2022
DOI	https://doi.org/10.6028/NIST.SP.800-215
CSRC URL	https://csrc.nist.gov/publications/detail/sp/800-215/final
Additional Information	

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Special Publication
NIST SP 800-215 ipd

Guide to a Secure Enterprise Network Landscape

Initial Public Draft

Ramaswamy Chandramouli

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-215.ipd

NATIONAL INSTITUTE OF
STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Special Publication
NIST SP 800-215 ipd

Guide to a Secure Enterprise Network Landscape

Initial Public Draft

Ramaswamy Chandramouli
Computer Security Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-215.ipd

August 2022

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology (NIST), nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

There may be references in this publication to other publications currently under development by NIST in accordance with its assigned statutory responsibilities. The information in this publication, including concepts and methodologies, may be used by federal agencies even before the completion of such companion publications. Thus, until each publication is completed, current requirements, guidelines, and procedures, where they exist, remain operative. For planning and transition purposes, federal agencies may wish to closely follow the development of these new publications by NIST.

Organizations are encouraged to review all draft publications during public comment periods and provide feedback to NIST. Many NIST cybersecurity publications, other than the ones noted above, are available at https://csrc.nist.gov/publications.

Authority

This publication has been developed by NIST in accordance with its statutory responsibilities under the Federal Information Security Modernization Act (FISMA) of 2014, 44 U.S.C. § 3551 et seq., Public Law (P.L.) 113-283. NIST is responsible for developing information security standards and guidelines, including minimum requirements for federal information systems, but such standards and guidelines shall not apply to national security systems without the express approval of appropriate federal officials exercising policy authority over such systems. This guideline is consistent with the requirements of the Office of Management and Budget (OMB) Circular A-130.

Nothing in this publication should be taken to contradict the standards and guidelines made mandatory and binding on federal agencies by the Secretary of Commerce under statutory authority. Nor should these guidelines be interpreted as altering or superseding the existing authorities of the Secretary of Commerce, Director of the OMB, or any other federal official. This publication may be used by nongovernmental organizations on a voluntary basis and is not subject to copyright in the United States. Attribution would, however, be appreciated by NIST.

NIST Technical Series Policies

Copyright, Fair Use, and Licensing Statements
NIST Technical Series Publication Identifier Syntax

Publication History

Approved by the NIST Editorial Review Board on YYYY-MM-DD [will be added upon final publication]

How to Cite this NIST Technical Series Publication:

Chandramouli R (2022) Guide to a Secure Enterprise Network Landscape. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-215 ipd.
https://doi.org/10.6028/NIST.SP.800-215.ipd

NIST Author ORCID iDs

Ramaswamy Chandramouli: 0000-0002-7387-5858

Public Comment Period

August 5, 2022 – September 19, 2022

Submit Comments

sp800-215-comments@nist.gov

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

All comments are subject to release under the Freedom of Information Act (FOIA).

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in federal information systems. The Special Publication 800-series reports on ITL's research, guidelines, and outreach efforts in information system security, and its collaborative activities with industry, government, and academic organizations.

Abstract

Access to multiple cloud services, the geographic spread of enterprise IT resources (including multiple data centers), and the emergence of microservices-based applications (as opposed to monolithic ones) have significantly altered the enterprise network landscape. This document is meant to provide guidance to this new enterprise network landscape from a secure operations perspective. Hence, it starts by examining the security limitations of the current network access solutions to the enterprise network. It then considers security feature enhancements to traditional network appliances in the form of point security solutions, network configurations for various security functions (e.g., application security, cloud access security, device or endpoint security, etc.), security frameworks that integrate these individual network configurations, and the evolving wide area network (WAN) infrastructure to provide a comprehensive set of security services for the modern enterprise network landscape.

Keywords

cloud access security broker (CASB); firewall; microsegmentation; secure access service edge (SASE); secure web gateway (SWG); security orchestration, automation, and response (SOAR); software-defined perimeter (SDP); software-defined wide area network (SD-WAN); virtual private network (VPN); zero trust network access (ZTNA).

Call for Patent Claims

This public review includes a call for information on essential patent claims (claims whose use would be required for compliance with the guidance or requirements in this Information Technology Laboratory (ITL) draft publication). Such guidance and/or requirements may be directly stated in this ITL Publication or by reference to another publication. This call also includes disclosure, where known, of the existence of pending U.S. or foreign patent applications relating to this ITL draft publication and of any relevant unexpired U.S. or foreign patents.

ITL may require from the patent holder, or a party authorized to make assurances on its behalf, in written or electronic form, either:

a) assurance in the form of a general disclaimer to the effect that such party does not hold and does not currently intend holding any essential patent claim(s); or

b) assurance that a license to such essential patent claim(s) will be made available to applicants desiring to utilize the license for the purpose of complying with the guidance or requirements in this ITL draft publication either:

    i. under reasonable terms and conditions that are demonstrably free of any unfair discrimination; or

    ii. without compensation and under reasonable terms and conditions that are demonstrably free of any unfair discrimination.

Such assurance shall indicate that the patent holder (or third party authorized to make assurances on its behalf) will include in any documents transferring ownership of patents subject to the assurance, provisions sufficient to ensure that the commitments in the assurance are binding on the transferee, and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding each successor-in-interest.

The assurance shall also indicate that it is intended to be binding on successors-in-interest regardless of whether such provisions are included in the relevant transfer documents.

Such statements should be addressed to: sp800-215-comments@nist.gov.

Table of Contents

List of Figures

Acknowledgments

The author would like to express his thanks to Isabel Van Wyk of NIST for her detailed editorial review.

Executive Summary

The enterprise network landscape has undergone tremendous changes in the last decade due to the following three drivers:

1. Enterprise access to multiple cloud services,
2. The geographical spread of enterprise-based (on-premises) IT resources (e.g., multiple data centers and branch offices), and
3. Changes to application architecture from being monolithic to a set of loosely coupled microservices.

The impact of these drivers on the security of the enterprise network landscape include:

- Disappearance of the concept of a network perimeter that can be protected and the necessity to protect each endpoint (device or service) that treats it as a perimeter
- Increase in attack surface due to sheer multiplicity of IT resources (computing, networking, storage) and components
- Sophistication of the attackers in their ability to escalate attacks across several network boundaries and leverage connectivity features

This document is meant to provide guidance to this new enterprise network landscape from a secure operations perspective. The adopted methodology considers the security challenges that the network poses and then examines the limitations of current network access technologies and how solutions have evolved from being security function-specific to a security framework to a comprehensive security infrastructure that provides a holistic set of security services. Specific areas addressed include:

- Feature enhancements to traditional network security appliances
- Secure enterprise networking configurations for various scenarios
- Security frameworks that integrate individual network configurations
- Evolving wide area network (WAN) infrastructure that provides a comprehensive set of security services

What is termed as the enterprise network in this document encompasses the various local networks on enterprise premises and that portion of wide area network that is used to connect its various geographically dispersed locations and cloud service access points.

1. Introduction

The enterprise network landscape has undergone a significant transformation in the last decade. The drivers for this transformation are (a) enterprise access to multiple cloud services, (b) the geographic spread of enterprise-owned (on-premises) IT resources (e.g., in a central office, multiple branch offices, and data centers), and (c) changes to application architecture from being monolithic to a set of loosely coupled microservices – often with a dedicated infrastructure (called the service mesh) that provides all application services, including security. The high-level impact of these drivers on the security of the current enterprise network landscape are (a) disappearance of the concept of a perimeter associated with the enterprise network; (b) an increase in attack surface due to the sheer multiplicity of IT resource components associated with application services, storage, and network appliances; and (c) sophistication of the attackers in their ability to escalate attacks across several network boundaries and leverage connectivity features. This document will consider these impacts by identifying the structural components of the new network landscape as well as specific security threats that they have opened up.

1.1. Structural Implication of Drivers on Enterprise Network Landscape

In order to have a good structural view of the current enterprise network landscape, it is necessary to look at the current enterprise IT environment in general. The IT environment now consists of:

- Subscription to multiple cloud services, such as IaaS for computing, SaaS for software, PaaS for an application development platform, and other cloud services (e.g., IDaaS for authentication)
- Enterprise IT applications (on-premises) located in corporate headquarters and geographically distributed branch offices and data centers
- IT applications range from being monolithic to ones that are made up of loosely coupled microservices, each of them hosted on heterogeneous platforms
- Presence of edge computing devices, such as IoTs, in some environments

The above scenarios call for widespread connectivity between IT systems that now defines the current enterprise network landscape. Connectivity, in turn, involves:

- Connectivity between IT resources (servers for computing and storage) in data centers (network fabric)
- Connectivity between IT resources within a corporate office or branch office (Wi-Fi, LAN, VLAN)
- Connectivity for users to remotely access to IT resources from home, travel locations, branch offices, and corporate offices using WANs, which use multiple networks such as the internet, MPLS, and – in some instances – cellular networks (e.g., 4G/LTE, 5G, etc.)
- Connectivity to cloud services through a cloud service provider (CSP), virtual private metworks (VPN), or subscription to WAN services (premises based equipment licenses or cloud-based)

1.2. Security Implication of Drivers for the Enterprise Network Landscape

The beginning of this section stated the following as the drivers for the state of the current enterprise network landscape:

- Subscription to multiple cloud services
- Geographically distributed IT resources
- Changes in application architecture

Now consider the immediate security implications of these drivers.

Subscription to multiple cloud services: Accessing cloud services from multiple cloud providers has become the norm for many enterprises. This trend is motivated not only by the need to avoid a cloud-vendor locked-in situation but also by different CSPs offering different value-added functions for different services (e.g., IaaS, SaaS). The consequence of this trend is that – from an enterprise point of view – the following networks have become extensions of the enterprise network and, thus, come under the scope of enterprise network management with attendant responsibilities for ensuring security protections becoming a critical function.

- Network used for accessing the cloud services
- Inter-cloud network (since communication between one CSP and another may be inevitable)
- The network inside the cloud provider that needs to be navigated to access the subscribed services (e.g., VPC, VNET, etc.)

Geographically distributed IT resources: The implication of distributed IT resources is that the users are also geographically distributed. Applications are now accessed by users not only from the enterprise premises, such as the corporate office and branch offices (through the enterprise network), but also from home and public locations (e.g., hotels and cafes) through multiple devices, such as desktops, laptops, and mobile phones. Ensuring secure access from these multiple locations and devices becomes the responsibility of the enterprise.

Changes in application architecture: Application architectures – especially those of cloud-native applications – have changed from being monolithic to being microservices-based, with the distributed nature increasing the communication channels between the components across a network (instead of just being local procedure/function calls). These applications have enlarged the threat and attack surface due to:

- Inherent architectures (multiple independent microservices and APIs),
- Automation tools used during software development and deployment, and
- Agile development and deployment methodologies, such as DevSecOps, that contain CI/CD pipeline code (workflows).

Attacks include data breaches, distributed denial of service (DDoS), account takeover (ATO) due to credential theft, and insider threats.

1.3. The Need for a Security Guide

Based on these considerations for security implementation, the arguments for the need for a security guide to the current enterprise network landscape are:

- Ubiquitous access locations, ubiquitous hosting locations of the application components, and multiple WAN transport protocols have caused shifts in security focuses, goals, and principles.
- The security focus has enlarged from being network-centric (i.e., internal/corporate network versus external/public internet) to user- and device/endpoint-centric.
- The new trust relationship has to be based not just on identity or the location of the access but enhanced to include validation of each access request (not just at the beginning of an access session), as well as the applicable set of contextual information associated with the user, device, or service.

1.4. Scope

The scope of this document includes:

- A structural view of the enterprise network landscape based on the distribution of IT resources and the consequent security challenges it poses
- Emerging and state-of-practice solutions in terms of feature sets and requirements to address the security challenges; solutions discussed will focus on the functional and operational levels

1.5. Target Audience

This guidance is intended for network design architects and network security solution architects in organizations with a hybrid IT environment (consisting of both on-premises and cloud-based applications) with a combination of legacy and microservices-based (i.e., cloud-native) applications.

1.6. Organization of This Document

The organization of this document is as follows:

- Chapter 2 considers traditional network access principles and technologies and their limitations in the context of the current enterprise network landscape.
- Chapter 3 provides a brief functional description of network security appliances – some new, some traditional (e.g., firewall) – that have enhanced capabilities to meet the security needs of the current network landscape.
- Chapter 4 outlines various network configurations that have evolved specifically for meeting the current network landscape (e.g., secure cloud access). It then considers the frameworks that integrate two or more of these stand-alone configurations in terms of their conceptual underpinnings and overall architectures.

- Chapter 5 focuses on the evolution of the WAN portion of the enterprise network landscape and enhanced offerings of the WAN services with global spread with a built-in security service infrastructure.
- Chapter 6 provides the summary and conclusions.

2. Traditional Enterprise Network Access Approaches and Their Limitations

Section 1 outlined the drivers for the current enterprise network landscape. Both drivers (change in application architectures and access to cloud-based applications) have impacted the mechanics of secure access to those applications through the network. Now consider the security limitations of the traditional enterprise network access approaches in the current enterprise network landscape context.

- Limitation of network perimeter-based protections
- Limitations of VPN-based access
- Limitations of MPLS technology as enterprise WANs
- Limitation of user identity-based controls

2.1. Limitations of Network Perimeter-based Protections

Early solutions for secure enterprise network access were geared toward environments with well-defined network perimeters. All enterprise IT resources were endpoints of enterprise LANs (usually defined as a floor in a large enterprise, building, or small campus), and multiple LANs connected together inside a defined building or campus constituted the internal corporate network. Entry points into this corporate network were protected using devices called firewalls, which were initially implemented as hardware appliances and later used software. In this environment, all devices and users within firewalls were totally trusted and, hence, considered safe for accessing application resources. However, the following factors have annulled the concept of that perimeter:

- Distributed nature of the application into ones located within a corporate data center, remote branch offices, and multiple cloud locations
- Perimeter approach based on the premise that the threat originates outside of the network, which is why most perimeter security solutions (e.g., IPS, IDS, firewalls) focus only on north-south traffic. However, over 75 % of network traffic is now east-west or server-to-server (due to applications being microservices-based), which is largely invisible to security teams. Any threat that is already inside of a network can move laterally and remain undetected for days or even months.
- Edge computing [1], where much of the computing takes place close to the location of multiple IoT devices
- Users located both within and outside of the corporate network, such as in homes, remote branch offices, and public locations (e.g., hotels, pubs, etc.). Some enterprises must also provide access to ecosystem partners, who may be on their own corporate networks.

The above scenarios have greatly expanded the attack surface.

2.2. Limitations of VPN-based Access

The increase in teleworking employees due to the pandemic has necessitated a means for secure access to IT resources inside an enterprise network in the form of virtual private networks (VPNs). A VPN allows organizations to extend a perimeter-based security across a public network. Security is enabled by setting up a secure tunnel in the public network using protocols such as IPSEC and TLS.

However, there are some limitations and security risks associated with VPNs.

- An increasing trend involves the movement of corporate resources to the cloud and the use of mobile devices. The VPN connections that remote users establish terminate at the VPN concentrators located at the edge of the corporate network. Hence, using a process called hair pinning, the traffic that lands at the corporate internet edge is routed back to the internet to access the cloud resources. This extra path increases network latency and has the potential to cause traffic bottlenecks.
- The mobile devices used by many employees, such as smartphones and tablets, can connect directly to software-as-a-service (SaaS) applications and data in the cloud. These mobile devices are especially prone to phishing attacks that steal credentials or deliver malware. Thus, the VPN becomes an entry point by which a bad actor could compromise a device and enter an organization's infrastructure.
- Two recent vulnerabilities were discovered in some VPNs [2]. One was "session hijacking," where malicious actors access a valid session ID through brute-force attacks or reverse engineering. The second vulnerability involved pulling a unique ID for an account, leveraging web browser development tools to manually set a value to the ID, and using that to obtain unauthenticated access to the VPN administrator console. That access was then used to remotely connect to internal systems, harvest passwords, move laterally in the network, and – in many cases – deploy ransomware.

2.3. Limitation of MPLS Technology as Enterprise WANs

Multi-protocol label switching (MPLS) technology is used for enterprise WANs, but the wide geographic span of an enterprise network due to multiple data centers and cloud services has imposed some limitations on its use.

- The geographic span of enterprise IT resources and subsequent networking connections have made traversal through internet inevitable for many portions of its enterprise's access network. Since MPLS is a different network, it provides access to the internet only through designated and limited access points. This increases latency for time-sensitive corporate applications.
- Given the different networking technology, the appliances and subsequent configuration procedures are different, making networking management a complex task.

2.4. Limitation of User Identity-based Controls

In traditional monolithic applications, all invocations of applications are either directly from the user or through scripts written and programmed to run by the user. Hence, the only parameters for access validation are the user identity or attributes associated with the user.

Changes in application architectures expand the validation parameters beyond user identity and attributes. The initial changes to application architectures are found in web-based and API-based applications where access can take place from any device located in any network (e.g., home, public WiFi, etc.). The latest changes are found in microservices-based applications (often called cloud-native applications because this architecture is the predominant one among cloud-hosted applications). This class of application consists of loosely coupled microservices that require the generation of multiple interservice requests to complete a business process or transaction. The limitations of identity-based controls can be seen from the following expanded security requirements for microservices-based applications:

- Validation is required not only for the identity of the users initiating the transaction but also for the identity of each service (service identity) making the request and the device on which the service is hosted (authorized device).
- The location of the service and device may change due to the virtualized nature of the application hosting environment (e.g., migration to VMs located in a different subnet, more powerful hosting devices and storage mechanisms, etc.), necessitating the need for validating an application request based not only on user identity and attributes but also on attributes associated with the device, network, geolocations, etc.
- The validation of identity (authentication) and authorization need to be done continuously (and not just at the beginning of an application invocation session) as the risk profile of an access may change due to there being multiple entities involved or changes in behavioral patterns that need to be included as a validation parameter (and monitored).

3. Network Security Appliances in Enterprise Network Landscape

This section will consider some new network security appliances as well as enhanced features in established appliances for meeting the security needs of the current landscape. These can be viewed simply as point security solutions, but evaluating their functions and features will provide an understanding of the effectiveness of network configurations and technologies that form part of the integrated solutions that are going to be discussed in Sections 4 and 5, respectively.

3.1. Cloud Access Security Broker (CASB)

Given the increasing subscription to multiple clouds in many enterprises, one of the most important pieces of software is the cloud access security broker (CASB). Just like IAM systems, a CASB can be run either on-premises or as a cloud-based service. It sits on the network between the cloud service customers (CSC) and the cloud service providers (CSP). The evolution of CASB functionality can be traced as follows [3]:

- The primary function of the first generation of CASBs was the discovery of resources. They provided visibility into all of the cloud resources that the enterprise users accessed, thus preventing or minimizing the chances of shadow IT. Shadow IT is the practice of some users using cloud applications that are not authorized by the enterprise IT management from home or the office using enterprise desktops. An example of this is the use of unapproved software-as-a-service (SaaS) applications for file sharing, social media, collaboration, and web conferencing by some enterprise users [4]. This generation of CASBs also provides some statistics, such as software-as-a-service (SAAS) utilization.
- The current generation of CASBs enforces security and governance policies for cloud applications, thus enabling enterprises to extend their on-premises policies to the cloud. Specific security services provided by CASBs include:
    - Protection of enterprise data that lives in cloud service providers' servers (due to SAAS or IAAS subscriptions), as well as data inflow and data outflow from those servers
    - Tracking of threats, such as account hijacking and other malicious activities. Some can detect anomalies in users' cloud access behavior (through robust user and entity behavior analytics, or UEBA, functionality) and stop insider threats and advanced cyberattacks [5].
    - Detection of misconfigurations in the enterprise's subscribed infrastructure as a service (IaaS) and cloud servers. These misconfigurations pose serious security risks such as data breaches. Alerts generated by CASB due to misconfigurations in the enterprise's IaaS deployments direct the enterprise to follow guidelines, such as the Center for Internet Security's (CIS) benchmarks for public cloud services, thus improving the overall security profile of the enterprise for cloud access [4].

3.2. Enhanced Firewall Capabilities

The security functions in firewalls have enlarged alongside the changing network landscape. Firewalls started as hardware appliances that prevented network packets from a device with a particular network location (e.g., combination of IP address and port) in one subnet (e.g., external network or internet) from accessing a device on another network location or subnet (e.g., intranet or DMZ or corporate network). In that setup, it primarily secured a network perimeter. The evolution of firewall functions can be traced based on the following feature sets [6]:

- Packet filters and network address translation: Packet filtering and NAT are used to monitor and control packets moving across a network interface, apply predetermined security rules, and obscure the internal network from the public internet.
- Stateful inspection: Stateful firewalling, also known as dynamic packet filtering, monitors the state of connections and makes determinations on what types of data packets belonging to a known active connection are allowed to pass through the firewall.
- Threat detection and response: Modern firewalls can gather and analyze enough data across multiple packets and sessions to detect threats and security incidents targeted at a particular system or a family of systems. The data from multiple firewalls can also be directed toward security information and event management (SIEM) and correlated with data from other security tools and IT systems to detect enterprise-wide attacks that span multiple systems and network layers. In addition, this data can be used to understand evolving threats and define new access rules, attack patterns, and defensive strategies [6].
- Logging and auditing capabilities: Logging and auditing capabilities result in the construction of network events that can be used to identify patterns of performance and security issues.
- Access control functions: Access control functions enforce granular sophisticated access control policies.
- Multiple locations and functions: Firewalls reside at different locations to perform different functions. Firewalls at the network edge perform the network perimeter protection function by filtering disallowed sources and destinations and blocking packets of potential threats. Firewalls inside a data center can create segmentation of the internal network to prevent the lateral movement of traffic and isolate sensitive resources (e.g., services and data stores). Device-based firewalls prevent malicious traffic in and out of endpoints.
- Open APIs integrate with many networking products.
- Some features centrally define or merge policies so that consistent policies are applied to different class of users (e.g., those on-premises and on private and public clouds).
- Web application firewalls (WAF): This class of firewalls has been used ever since web applications accessed through web protocols, such as HTTP, came into existence. A feature advancement in this class of firewalls is advanced URL filtering. This is the ability to detect traffic from malicious URLs and thus prevent web-based threats and

attacks by receiving real-time data analyzed by machine learning algorithms [7][8]. Specifically, this class of firewalls can inspect threat vectors for SQL Injection, OS command injections, and cross-site scripting attacks, as well as prevent inbound attacks. They are used in content delivery networks (CDN) and to prevent distributed denial-of-(DDoS) attacks. Some additional features found in this class of firewalls are:

1. Ability to specify allowable list of services (control at the application level)
2. Traffic matches the intent of allowed ports
3. Filtering of some unwanted protocols

3.3. Appliance-set with Integrated Functions

- Unified threat management (or UTMs): UTM devices combine many of the most critical security functions – firewall, intrusion prevention system (IPS), VPN concentrator, gateway antivirus, content filtering, and WAN load balancing – into a single device, usually with a unified management console.
- Next-generation firewall (NGFW): This all-in-one security appliance is based on the UTM model but is combined with enterprise-class scalability and performance and a focus on the granular inspection of Layer 7 application traffic. NGFWs have added capabilities to facilitate internal segmentation, integration with sandboxing products, secure sockets layer (SSL) inspection, and SD-WAN. Processing at the edge benefits from on-premises firewalls, which apply processing on-site. They are more energy-efficient than virtual machines and reduce latency because they avoid the "round trip" to the cloud. NGFWs come with high-performance threat protection (e.g., intrusion prevention, web filtering, anti-malware, application control) for known attacks, SSL/TLS inspection, and antivirus [9].
- Web application and API protection (WAAP): This is a comprehensive security approach and an enhancement over web application firewalls (WAF). WAF is an integral component for API security, BOT defense, and DDOS protection.
- These can be offered as a product suite or as a cloud-based service [10][11].
- Secure web gateway (SWGs): SWGs are appliances utilized for policy-based access to and control of cloud-based applications for enterprise users in ubiquitous locations (e.g., headquarters, branch offices, home, remote locations). A SWG is fundamentally a web filter that protects outbound user traffic through HTTP or HTTPS inspection [12]. They also protect user endpoints from web-based threats that can occur when users click on links to malicious websites or to websites infected with malware. They centralize control, visibility, and reporting across many locations and types of users. They are not a replacement for WAFs, which protect websites housed in enterprise data centers and large headquarter sites from inbound attacks.

3.4. Requirements for Network Automation Tools

Network automation tools automate the entire life cycle processes involved in deployment, observability/monitoring, threat intelligence gathering/reporting (e.g., generating alerts of security violations for security personnel to take timely action), and – in some instances –

automatic remediation. These automated tools are an indispensable part of a complex enterprise network landscape. The requirements for these tools can be broadly classified into generic and functional requirements. These requirements are described below. Each generic requirement is tagged with the abbreviation NAUT-GR-x, while each functional requirement is tagged with NAUT-FR-x, where x in both types of tags stand for the numerical sequence.

- NAUT-GR-1: Scale to meet the volume, velocity, and variety of today's application development deployment and maintenance paradigms [13]. This requirement is critical in environments where DevSecOps is used to deploy not only applications but also infrastructures, the latter using infrastructure-as-code (IaC) tools. These tools are made an integral part of the smart automated workflows called CI/CD pipelines, which invoke these tools to deploy servers (computing), networking, and storage infrastructure. Hence, this class of network automation tools can be seamlessly integrated into the corresponding CI/CD pipelines.
- NAUT-GR-2: They should have the capability to minimize human intervention for security remediation, which is slow and prone to error. In other words, the more automated remediation features built into the tool, the better.

The minimum functional requirements of network automation tools should be:

- NAUT-FR-1 (enhanced threat intelligence and protection): The tools should have advanced threat intelligence, real-time threat prevention capabilities for known and zero-day vulnerabilities, and sandboxing features for isolating malicious traffic.
- NAUT-FR-2 (leveraging knowledge of previous events): The tools should have features for matching current events to past ones and for leveraging the remediation measures performed for those instances in the current solution. This brings about reduction to the average outage time [14].

The network monitoring and observability tools and IaC tools are important classes of network automation tools, and the requirements and feature set are discussed in the following subsections.

3.4.1. Network Monitoring and Observability Tools

This class of tools gathers the data for obtaining visibility into the entire network. The data is then used to generate a dashboard that presents the topography of the enterprise network by showing all connections and presenting key operating parameters (e.g., latency, network traffic level, etc.). Some of the data generated by this class of tool and their uses are:

- Identification of interfaces: Monitoring tools identify the interfaces for defining the parameters for network resources provisioning and help the IaC generate the relevant code for invoking those interfaces.
- Measurement of drift: Despite using IaC to deploy the network infrastructure, unauthorized or ad hoc changes in network configuration can alter the performance and security parameters for application execution (called the drift). Monitoring tools should have the ability to monitor these parameters (e.g., bandwidth availability, unwanted traffic, etc.) and alert for corrective action.

- Secure overlay designs for cloud service access: Monitoring tools can generate data to enable centralized network management tools to perform security functions, such as building a virtual network segmentation on top of the native network segmentation features offered by CSP, provided that suitable APIs are available.
- Support for incidence response process: Sophisticated network monitoring tools generate network security alerts and threat intelligence feeds. Handling these alerts and feeds is part of the incidence response (IR) process in an enterprise and is carried out by members of a security operations center (SOC). A security strategy that has evolved in recent years to automate the IR process is called security orchestration, automation, and response (SOAR). Some of the state of practice applications of SOAR include threat detection and response, vulnerability prioritization, compliance checks, and security audits with potential applications in many emerging areas, such as IoT management [15].

3.4.2. Automated Network Provisioning Tools

As already stated, automated network resource provisioning is enabled by infrastructure as code (IaC) tools. The code that describes the networking infrastructure (in addition to the computing and storage infrastructure) is stored in a code repository. The process of initial deployment of the networking infrastructure and subsequent upgrade is automated by defining a workflow that invokes the IaC (e.g., GitOps workflow) as part of a CI/CD pipeline definition [16]. The advantages of this approach for managing the enterprise networking infrastructure for multi-cloud deployment are the following:

- Enables the enterprise to have tight version control (tracking changes) so that unauthorized networking devices and unauthorized changes in associated configurations do not open up security vulnerabilities.
- Enables the enterprise to have a uniform infrastructure across all environments – development, testing, staging, and production.
- Monitoring the drift (the unintended changes) between the defined infrastructure (as found in IaC) and the operational infrastructure (as measured by monitoring tools described in Section 3.4.1) and taking corrective action to address the drift help to maintain the necessary security posture for the enterprise networking environment.
- The DevSecOps paradigm consisting of CI/CD pipelines invokes the network provisioning tool (IaC code generator) to automate the initial deployment and subsequent re-configuration of the networking infrastructure. Since the pipelines have a built-in audit process, the changes in network configuration are automatically captured in the audit, which in turn enables the enterprise to demonstrate corporate security policy compliance and regulatory policy compliance for their networks where applicable.
- Testing the code (IaC code) generated by IaC tools (and invoked by the CI/CD pipeline code that deploys the infrastructure using IaC) ensures that security policies are consistently and uniformly applied across the entire enterprise networking infrastructure (i.e., multiple cloud services).
- The advantage of having plug-ins for defining network provisioning for different public cloud provider environments is that they can be used to customize the observability tools

used for networking monitoring for each of those cloud services that the enterprise has subscribed to [17].

3.5. Networking Appliances as Services

Another trend in the enterprise network landscape is that a portion of network infrastructure can be obtained as a leased service called a network as a service (NaaS) from third-party providers. This service is offered using technologies such as enterprise 5G and edge computing. The advantages of NAAS are as follows:

- Just like subscriptions to SaaS and IaaS, it reduces capex costs for the enterprise.
- Being software-defined and virtualized, it is flexible and scalable.
- As a consequence of the previous advantage, QoS requirements of diverse applications can be met by creating customized traffic flow for each application type [18].
- New applications that require an increased network footprint can be quickly introduced to the enterprise (agility), thus facilitating business diversification.

4. Enterprise Network Configurations for Hybrid Application Environments

Since the enterprise context in this document refers to enterprises that consist of on-premises and cloud-hosted applications (i.e., hybrid application environments), this document describes the network configurations or designs (and network communication exchanges based on them) that have emerged as state of practice in those enterprises.

The state of practice network configuration features (NCF) found in enterprises with hybrid application environments can be classified under the following areas [19]:

- Network configuration for Device Management
- Network configuration for User Authentication
- Network configuration for Device Authentication and Health Monitoring
- Network configuration for Authorizing Application Access
- Network configuration for Preventing Attack Escalation (Microsegmentation)

Each of the network configuration features are enumerated using the identifier of the form HAE-NCF-x, where HAE denotes a hybrid application environment, NCF denotes the network configuration feature, and x stands for the sequence number of the feature.

4.1. Network Configuration for Device Management

With the disappearance of the network perimeter (Section 2.1) and the distribution of the application targets (being a hybrid application environment), enterprises should adopt an "endpoint is the perimeter" paradigm and have a device management system in place.

HAE-NCF-1: All endpoints that will be accessing on-premises and cloud-based applications should be managed using a dedicated management network. Minimal managed tasks should include:

a) Installation and maintenance of device and service authentication certificates
b) Installation and maintenance of device health applications
c) Updates of patches on the devices
d) Creation and maintenance of white pages that contain device-service mappings to prevent service hijacking (preventing malicious or compromised servers posing as legitimate hosts for the services)

4.2. Network Configuration for User Authentication

HAE-NCF-2: The network should be configured to route the user access request to different destinations for authenticating the user, depending on the target application accessed.

a) When the user access request is for a cloud-based application (e.g., SaaS), the user should be routed to the enterprise IdP. When the user access request is for an on-premises web application, the user should be directed to a web gateway (reverse proxy). This redirection can be affected through a process called split DNS. If a digital certificate is used as the first authentication factor, the IdP should check the validity of the user certificate (right status and not expired) through mechanisms such as CRL, OCSP, or Active Directory calls.

b) A minimum of two authenticator factors must be used to authenticate users. If possession of a valid certificate is the first factor, then the acknowledgement of a push message (using technologies such as DuoMobile, TouchID or Yubikey) or OTP to the cell phone can be used as the second factor.

4.3. Network Configuration for Device Authentication and Health Monitoring

HAE-NCF-3: Device authentication can be performed through certificate validation using appropriate protocols. A device health check can be performed by invoking the resident application.

HAE-NCF-4: Microservices-based applications (on-premises or in cloud) should have service proxies installed with each service to provide the necessary connectivity for inter-service communication in addition to performing authentication and authorization services for each service request.

4.4. Network Configuration for Authorizing Application Access

HAE-NCF-4: Standardized protocols, such as OAuth 2.0 [20], should be used to issue access tokens to the validated user, device, or service to enable access to cloud-based applications.

4.5. Network Configuration for Preventing Attack Escalation (Microsegmentation)

Microsegmentation is a security design practice where an internal network (e.g., in the data center, cloud provider region) is divided into isolated segments so that the traffic in and out of each segment can be monitored and controlled [21].

Things enabled by microsegmentation are:

- Segments being isolated and relatively small enables close monitoring of the traffic because of better visibility.
- The consequence of the above capability is that granular access control is possible by defining associated policies.

The enablement of the above capabilities restricts the unauthorized lateral movement of a user or application that has either (a) breached the perimeter to enter the internal network or (b) been initiated by users within the internal network itself.

4.5.1. Prerequisites for Implementing Microsegmentation

a) Creation of application identity: The fundamental requirement to enable this is the assignment of a unique identity to each application or service, just like how each user carries a unique identity (e.g., userid). Prior to the era of cloud-based applications, the application requests were validated based on the IP subnet or IP address from which they originated. Ubiquitous access and multi-clouds have eliminated the concept of network perimeters. Hence, authentication and authorization based on those parameters are neither feasible nor scalable. Further, the presence of proxies, network address translation, and load balancers make it impossible for the called application to know the IP address of the calling application in order to make authentication or authorization decisions. A unique application identity is inevitable.

b) Establishment of trust in application identity: The created application identity should not be subject to spoofing and should be continuously verifiable. Hence, a cryptographic identity in the form of a public key carried in a certificate issued by a trusted source is required to meet these criteria. Verification of the authenticator associated with this identity is done by the authenticating party by sending a challenge, and insurance against replay attack for the authentication process is ensured by sending a nonce attached to the challenge. A secure directory that provides a mapping of the service to the hosting server should be maintained to ensure that applications or services are hosted only on authorized servers and that spurious versions of services do not exist.

c) Discovery of application resources: There should be a robust means for discovering all application resources (e.g., services, networks, etc.).

d) Segmentation of workloads: Security requirements for all applications and services must be identified and groupings established based on identical security requirements.

e) Mapping of logical application groupings to physical or virtual infrastructures: Application-centric groupings must be mapped to physical or virtual infrastructures that constitute the data center topology to facilitate actual applications and services deployment.

4.5.2. Microsegmentation – Implementation Approaches

The following approaches are employed to implement microsegmentation [22]:

a) Segment-based approach: In this approach, the applications and services resources with similar security requirements are grouped into a unique segment, and firewall rules are created to block or allow communication with each group or segment. The segments are created using network layer abstractions, such as VLAN IDs or some other tagging approaches, while policies are defined using network address constructs (e.g., IP addresses and ports). Policies apply to subnets (e.g., VLANs) and not to individual hosts. Each segment is protected by gateway devices, such as intelligent switches and routers or next-generation firewalls (NGFWs), which should have the capacity to react and adapt in response to the threats and changes in the application workflows. Segmentation gateways monitor traffic, stop threats, and enforce granular access across east-west traffic (rarely for north-south traffic) within on-premises data centers or cloud regions. The main difficulty with this approach is the difficulty in mapping the applications security requirements-based segments created to corresponding network segments [23].

A schematic diagram of the segment-based microsegmentation is shown in Figure 1. Each numbered microsegment in the figure is a unique VLAN identified by a VLAN ID. The group of applications that will run in that particular VLAN segment can be defined using different criteria. One of the criteria is “all applications with similar security requirements." Another is that "all tiers (web frontend, application logic servers, and database servers) associated with a particular application" should run in a single microsegment, as shown in the figure.

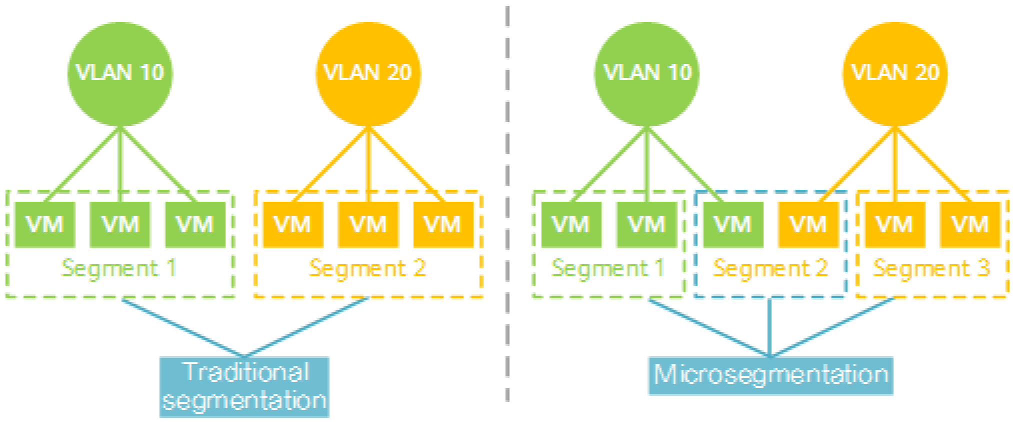

Fig. 1. Segment-based Microsegmentation

b) Virtualized server-based approach: This approach is only applicable to networks that contain virtualized servers since it is implemented in the hypervisor. There are two possible mechanisms:

1. Using virtual firewalls inside a hypervisor to isolate traffic destined for different VMs inside the hypervisor
2. Using encapsulation techniques to create overlays (e.g., VXLAN) that run on top of an underlay network consisting of IP addresses designations; access control policies are enforced on the hypervisor itself outside of the workload (application or microservice)

c) Host-based microsegmentation: Alternatively (or additionally), host-based microsegmentation can be implemented using software agents on the endpoint artifacts (e.g., servers). It leverages native firewall functionality built into the host. Software agents can overlay a software-defined segmented network across data centers, cloud, bare metal, and hybrid environments. The agent provides context awareness and visibility for each workload and, hence, enables the definition and enforcement of fine-grained policies.

d) Identity-based microsegmentation: Identity-based microsegmentation policies use contextual, application-driven identifiers (e.g., order processing front-end service can communicate with inventory back-end service) instead of network parameters (permit calls from 192.168.10.x subnet to 10.0.0.31) [24]. The identifiers assigned to services are cryptographic identities, which they use for mutual authentication and authorization during each service request and response.

The advantages of this type of microsegmentation are:

- Policies based on service/application identities do not use any infrastructure-related variables (e.g., IP addresses, subnets, etc.), so these policies are environment-agnostic and provide the freedom for the services/applications to be migrated to different environments and still maintain the same policies.

- Policies being independent of infrastructure enables them to be tested by merely exercising the application and observing the outcomes (e.g., trace of the sequence of service calls and requests/responses instead of configuring the infrastructure correctly for test runs).
- With the availability of tools for the declarative specification of policies through "policy as code" tools, microsegmentation policies can be defined/implemented by incorporating the code into automated workflows, such as CI/CD pipelines.
- Microsegmentation enables granular (fine-grained) access control by providing visibility to application call sequences/interdependencies and data flows through host-level tracking, thus enabling the enforcement of security policies for application traffic that is both north-south and east-west, irrespective of the environment (e.g., corporate data center or cloud infrastructure).

The reason that identity-based microsegmentation is studied under the enterprise network landscape is that it enables only valid network traffic between the various component services of the application due to the mutual authentication and authorization using service identities, thus enabling the goals of zero trust network access (ZTNA) to be met [25].

4.6. Security Frameworks Governing Network Configurations

The network configurations described in Sections 4.1 through 4.5 are each for specific functions (e.g., user authentication, preventing attack escalation, etc.). In many enterprise environments, these network configurations are not ad hoc but are driven by some conceptual underpinning (e.g., contextual information) and/or some evolving enterprise network security framework that the enterprise has chosen to implement. Examples of such frameworks are software-defined perimeter (SDPs) and zero trust network access (ZTNA).

4.6.1. Conceptual Underpinnings – Contextual Information

Section 2.4 discussed the limitation of using user identity alone to authorize application access. This, however, does not mean that identity verification can be relegated to a secondary requirement. It has been widely recognized that identity validation is the entry point (may be a highly vulnerable point of entry into the system) to an application request [26] since all requests – whether coming from a service (or microservice), user, or device – come with a claimed identity. This identity must be verified using robust, phishing-resistant multi-factor authentication.

However, other attributes associated with the user and the information associated with other entities involved in an application access request, such as devices and services, are required in current enterprise IT environments and are collectively called contextual information. This contextual information set may vary from one enterprise to another and is also based on the level of trust that the organization requires for a particular access request. Since the role of contextual information in potential attacks may not be known, the set to be included in the access decision is a risk-based decision. Contextual information may broadly belong to the following five key areas [27]:

1. Information about the user requesting access – Apart from user identity, attributes associated with the user, such as their role in the organization, current assignments, and status (cross verification of identity in the enterprise IDM vs enterprise directory)
2. Information about the device from which access is being requested – Establishing trust in the device through a combination of health and risk profiles of the device. For example, the risk profile of the device can be obtained through an out-of-the-box posture check (risk of the device [28]) with or without integrating with an endpoint protection tool for the device. Other crucial information (provided by telemetry data) needed to assess the security status of the endpoint devices [29] include (a) device support label (the device is managed or corporate-owned) and (b) device posture information (whether it has been compromised). All of these factors go into a policy evaluation for determining the level of trust and must be channeled into authentication and monitoring decisions [30].
3. Information about real-time contextual data – Date, time, and geolocation at which the access request occurs
4. Information about IT services (e.g., app, data, etc.) being accessed
5. Information about the security of the environment hosting the IT services being accessed

The requirements for contextual information [27]:

- Should include not only that which is collected by the native platform (the platform on which the application is hosted) but also that which can be obtained from third-party platforms and can provide more detailed information
- Should be available in real time so that user experience with access is not affected
- Should be prioritized based on the value each provides
- Should be consistent with the level of risk associated with each access request

No application and/or data access in the modern enterprise network context can be deemed secure by ignoring relevant contextual information when the access scenario involves allowing a user, device, or service from any network channel (e.g., corporate network, home network, public network, or branch office) to access a resource located anywhere (on-premises or cloud).

4.6.2. Network Security Framework – Software-defined Perimeter (SDP)

One conceptual underpinning for secure network access to IT resources is the software-defined perimeter (SDP) [31]. In SDP, the separation between networks is not defined by network address group or VLANs, making it network-agnostic. It is logically and dynamically defined for each user and each particular request. In other words, for each user request, the subset of IT resources to which the user has access is dynamically allocated irrespective of the location of the resource (e.g., corporate data center, branch office, private or public cloud, etc.). The salient principles of SDP include:

- The SDP concept involves making all IT resources invisible (e.g., ports, workloads, and applications) and making them known and accessible only after the user is authenticated and authorized. Only a network connection between the user and the allowed IT resources is established, thus following the least privilege principle.

- The access level determined by the previous process is continuously reevaluated during the user session and recalibrated if required. In other words, as the context surrounding the identity changes in real time, so can the user's entitlements [31].
- Reduce the attack surface by preventing lateral movement [32] through techniques like microsegmentation, as described in Section 4.4. With the increasing deployment of microservices, the inter-services resource requests (generator of east-west traffic) dominate external application requests (north-south traffic). Application of this principle thus secures east-west traffic.

4.6.3. Network Security Framework – Zero Trust Network Access (ZTNA)

ZTNA is the consequence of a zero trust architecture, which in turn is a realization of zero trust principles. NIST defines zero trust and zero trust principles as [33]:

- Zero trust (ZT) is the term for an evolving set of cybersecurity paradigms that move defenses from static, network-based perimeters to focus on users, assets, and resources. It is a set of security primitives rather than a particular set of technologies. Zero trust assumes that there is no implicit trust granted to assets or user accounts based solely on their physical or network location (i.e., local area networks versus the internet) or based on asset ownership (enterprise or personally owned). Zero trust focuses on protecting resources (e.g., assets, services, workflows, network accounts, etc.) rather than network segments, as the network location is no longer seen as the prime component to the security posture of the resource.

- A zero trust architecture (ZTA) uses zero trust principles to plan industrial and enterprise infrastructure and workflows.

NIST's guidance on ZTA [33] contains an abstract definition of zero trust architecture (ZTA) and gives general deployment models and use cases where zero trust could improve an enterprise's overall information technology security posture. From the NIST vision of ZTA and state of practice implementations [34], the following have emerged as the three building blocks of ZTA:

1. **Client or Browser**: The point of entry for all users to access any resources hosted in multi-cloud and on-premises environments

2. **The Controller**: The policy decision engine, which manages the policies, conditions, and entitlements that grant access for all users, devices, and workloads from a single dashboard or via API
3. **The Gateway**: The policy enforcement point. Gateways control the flow of access to protected resources. It dynamically builds micro-segmentation rules based on granted entitlements.

In all security frameworks for current enterprise network environments, the common principles that underly application-specific requirements – such as low latency, high data transfer rates, and high reliability – that were applicable in previous network landscapes remain the same.

5. Secure Wide Area Network Infrastructure for an Enterprise Network

The wide area network (WAN) became an integral part of the enterprise network when organizations needed to connect their local area networks (LANs) across multiple geographically distributed locations (within the country and, in some cases, globally) starting in the 1980s. The initial WAN technology involved point-to-point (P2P) leased lines followed by Frame Relay. The first IP-based network was multiprotocol label switching (MPLS), which enabled multiple types of traffic – such as voice, video, and data – to travel on the same line.

With the advent of technologies such as virtualization and increasing enterprise access to cloud services, enterprises have begun to adopt a new WAN technology called the software-defined wide area network (SD-WAN). SD-WAN technology removes the tight coupling between the control plane and data plane functions of the network and enables the centralized specification of various policies, such as access control, routing, and application traffic prioritization.

Another development involved integrating all of the point security solutions provided by various network security appliances (Section 3) into a network security services infrastructure. Industry and industry consortiums use the term secure access service edge (SASE) [35] to refer to a comprehensive framework that offers wide area networking and various security services. SASE can be looked upon as the networking counterpart of the application's service mesh, which provides a comprehensive set of application services, including security for cloud-native applications.

Based on the above discussion, this section will focus on the following topics:

- Requirements for a secure SD-WAN
- Requirements for an integrated security services architecture for SD-WAN

5.1. Common Requirements for a Secure SD-WAN

In addition to CSP-provided VPNs, a networking technology that provides network connectivity for accessing cloud-based services for enterprises is the software-defined wide area networking (SD-WAN).

The design goals and common features in all SD-WAN offerings include:

- <u>Extensive connectivity</u>: To securely connect users located anywhere (e.g., home, public location, branch office, corporate office, etc.) to applications and resources hosted anywhere (e.g., data center, single or multiple cloud services) using any WAN transport (e.g., MPLS, Broadband Internet, 4G, LTE, 5G wireless)
- <u>Application awareness</u>: To monitor the network traffic and dynamically choose the best path available based on (a) the type of network traffic, (b) network load conditions, and (c) the application's business priority. This capability is enabled using techniques such as bandwidth utilization, load balancing, and the optimization of speed by reducing jitters, latency, and packet loss. Addressing application's business priority is only possible if the SD-WAN solution has the ability to identify different types of applications (e.g., messaging/email application, social media application, general storage-related applications, supply chain applications) and allocate routing priorities and WAN resources accordingly.

- Integration of security and networking functions: Use of appliances that contain a combination of networking and security functions (e.g., the presence of a firewall and secure web gateway [SWG] functions in a WAN router) [36]
- Centralized visibility and management capabilities: Includes the ability to recognize and authenticate newly connected appliances and bring them under the defined management workflows as nodes so as to configure a uniform set of policies that cover all components
- Integration with remote LAN locations: An additional preferred but non-essential feature is the integration of WAN and LAN functions in a single appliance (the latter going by the name SD-Branch), which can be managed using a single management console, thus providing better visibility into both components. This feature enables the connectivity of SD-WAN into the local LAN at the remote branch offices.

5.2. Specific Requirements for WANs for Cloud Access

Enterprises can gain cloud access in two ways: (1) through the VPN services provided by the cloud providers or by (2) integrating their own SD-WAN with cloud providers' private networks, often called the cloud WAN. The advantage of the second approach is that enterprises can extend their existing WANs into and across a cloud provider's private network, enabling consistent enterprise networking and security policy enforcement. Two of the advantages of this extension are:

1. Complete end-to-end visibility between "access endpoint" and IT resource (application or data) endpoint even though the latter is located in a cloud provider's network
2. Application of the network segmentation logic deployed for accessing on-premises resources to the cloud-based resources [37]

This orchestration of the cloud provider's private network can be achieved by designing a customized overlay network on top of the cloud provider's network as the underlay network. This feature is contingent upon CSPs offering API integrations for different SD-WAN offerings [38][39][40].

An architecture has emerged for managing enterprise networks that are connected to multiple CSPs. A portion of the industry calls the collection of appliances in this architecture a cloud network platform. The requirements for this multi-cloud networking platform are [41]:

- It should deliver common operational visibility and control across native network access provided by multiple cloud providers. The big challenge is that public cloud providers have different proprietary architectures using their own "constructs." In order to provide a networking architecture that can "cross clouds," one needs to leverage the cloud-native functionality (especially native cloud networking constructs) of each cloud; abstract that functionality with APIs; add advanced data plane features for high-availability, security, and operational visibility/control; and provide the tools to manage these features dynamically or automatically [42].
- It should deliver a common ingress and egress security policy for application environments (e.g., VPCs, VNETs, VCNs, etc.) across clouds.

- It should enable end-to-end encryption inside of the cloud as well as high-performance encryption from the data center to the cloud.
- It should support automation for deployment and configuration.

Based on the above requirements, multi-cloud networking platform offerings have emerged with the following architectural elements:

- An abstraction layer that sits on top of the native network access offered by individual CSPs to their services. This layer enables the enterprise to manage the entire enterprise network – consisting of connectivity to multiple clouds, intra-cloud connections, and the on-premises data center network fabrics – as one unit. To enable this, complete visibility into the entire enterprise network landscape is needed. Hence, this layer needs input from sophisticated observability and monitoring tools to carry out its functions.
- Choosing an infrastructure configuration (e.g., virtual private cloud configuration with isolated network segments) for hosting applications in the network infrastructure provided by the CSP is facilitated by a class of tools called IaC tools, which have features with network configuration definitions of major CSPs built in as plug-ins. This facilitates initial networking resource provisioning and subsequent modification of networking configuration and resources for hosting enterprise applications in clouds.

There are four industry trends [43] that may have security implications with regard to SD-WAN [44]:

1. SD-WAN access is acquired as a cloud-based service under the umbrella of network as a service (NaaS), just like IaaS and SaaS.
2. AI-based algorithms are used for monitoring networks for security-related conditions; for resiliency-improving measures, such as throttling for certain destinations; and for dynamic routing decisions to maintain QoS parameters, such as latency and bandwidth.
3. Wireless networks are used for last mile connectivity using a 5G Radio Access Network (RAN).
4. Secure remote access functionalities provided by technologies such as VPN are combined into SD-WAN [45].

5.3. Requirements for an Integrated Security Services Architecture for SD-WAN

An integrated security services architecture for SD-WAN has integrated within it both networking and security functions. The network access and security functions capabilities are offered as a cloud service that are accessible for enterprises through strategic network locations spread over a wide area called Point of Presence (PoP). The term coined by Gartner in 2019 to denote an architecture that converges networking and security functions and delivered at a global scale as a cloud service is called Secure Access Service Edge (SASE) [46]. The networking and security services delivered by a service called SASE are not new but just delivered together as a single package instead of through point security solutions (chapter 3). The various points of connectivity from the enterprise to SASE PoPs are called enterprise edges. The enterprise edges can be either:

- Clients (Users accessing through desktops, laptops and mobile devices either from branch offices or remote locations such as Home, or IoTs)

- IT Resources (Internal Apps hosted in data centers or branch offices, Cloud-based Apps (SaaS, IaaS))

The SASE network infrastructure thus becomes an integral part of the enterprise network whenever one or more of the enterprise edges get connected to various PoPs of SASE cloud service.

The three primary functions delivered by SASE are [46]:

- Optimization of Network Traffic for different types of Traffic – Reduce Latency and Improve Availability
- Access Control for accessing different types of IT resources -Applications, Databases etc.
- Threat Prevention – Monitoring, Gathering threat and attack information, remedial action

Some of the structural features in SASE offerings are:

- Globally distributed point of presence (PoP): A global SD-WAN service with its own private backbone network consisting of worldwide points of presence (PoPs) intended to minimize latency problems. In some instances, major cloud vendors' PoPs may also be leveraged.
- Security agent on devices: The security agent on the end user's device undertakes networking decisions and directs traffic from different applications. Specific capabilities include dynamically allowing or denying connections to services and applications based on an organization's defined business rules.

The following are the minimal security services found in an integrated architecture:

- Firewall services
- Secure web gateway services
- Anti-malware services
- IPS services
- CASB services
- DLP services

Some of the advanced security features found in SASE offerings include:

- <u>Browser isolation technology</u>: This is often combined with secure web gateway solutions and provides improved web activity security to tackle threats in real time.
- <u>Continuous adaptive risk and trust assessment (CARTA) strategy</u>: This strategy involves constantly monitoring sessions and performs adaptive behavior analysis on monitoring parameters to dynamically change security levels and permissions if the trust profile (e.g., trust deficit) of a device changes.

6. Summary and Conclusions

The purpose of this document is to provide insights into the current enterprise network landscape in terms of topology, traffic flows, and security threats. It takes the view that changes in application architecture and technologies (monolithic to microservices-based, bare metal to virtualization/containers) and increased subscriptions to various types of cloud services (e.g., IaaS, SaaS) are drivers of the current state of enterprise networks.

It outlines the limitations of existing network access security assumptions and technologies due to changes in network topologies in modern enterprise networks. The emergence of new network appliances (e.g., CASB), enhanced features in existing appliances (firewalls), network automation tools for gathering data for visibility/monitoring, threat detection and remedial actions, and tools for automated network provisioning for different public CSP environments (enabled by IaC tools invoked as part of the smart workflows called CI/CD pipelines defined under the DevSecOps paradigm) are all discussed under point security solutions. Various networking configurations for user, device, and service authentication and authorization as well as microsegmentation to prevent the escalation of attacks are also discussed.

Finally, this document discusses the latest WAN technologies that form part of the current enterprise network landscape, as well as the features of WAN offerings with global PoP and integrated security services called SASE.

References

[1] Craven C (2019) *What is the Difference between Edge Computing and MEC*. Available at https://www.sdxcentral.com/edge/definitions/whats-the-difference-between-edge-computing-and-mec/

[2] The Monitor Issue 13 (2020) *VPN Vulnerabilities Tied to Raising Data Exposure, Ransomware*. Available at https://www.kroll.com/en/insights/publications/cyber/monitor/vpn-vulnerabilities-rising-data-exposure-ransomware

[3] Hardcastle JL (2018) *Why CASB Is the Fastest Growing Security Category*. Available at https://www.sdxcentral.com/articles/news/casb-fastest-growing-security-category-ever/2018/02/

[4] Proofpoint (2021) *Getting Started with CASB*. Available at https://www.proofpoint.com/us/resources/white-papers/getting-started-with-casb

[5] Lookout (2021) *Embracing Zero Trust: A Guide for Agencies to Address the Cybersecurity Executive Order*. Available at https://www.govexec.com/media/embracing-zero-trust-guide-agencies-address-cybersecurity-executive-order.pdf

[6] CATO Networks (2022) *Network Firewall: Components, Solution Types, and Future Trends*. Available at https://www.catonetworks.com/network-firewall/

[7] Palo Alto Networks (2022) *Advanced URL Filtering*. Available at https://www.paloaltonetworks.com/network-security/advanced-url-filtering

[8] Oswal (2022) *Cloud NGFW: Managed Next-Generation Firewall Service for AWS*. Available at https://www.paloaltonetworks.com/blog/2022/03/next-generation-firewall-service-for-aws/

[9] Fortinet (2021) *Fortigate Next Generation Firewall*. Available at https://www.fortinet.com/content/dam/maindam/PUBLIC/02_MARKETING/02_Collateral/SolutionBrief/sb-fortigate-network-firewall.pdf

[10] F5 Networks (2022) *WAAP Buying Guide*. Available at https://media.bitpipe.com/io_15x/io_158522/item_2439191/EBOOK-SEC-798086545-waap-buying-guide_FNL%20%281%29.pdf

[11] F5 Networks (2022) *Choose the WAF That's Right for You*. Available at https://media.bitpipe.com/io_15x/io_158522/item_2439191/EBOOK-SEC-798087620-which-waf-is-right-for-you-refresh-FNL%20%281%29.pdf

[12] AT&T (2020) *The essential guide to secure web gateway*. Available at https://cybersecurity.att.com/resource-center/white-papers/essential-guide-to-secure-web-gateway

[13] Itential (2020) *Redefining Network Configuration Management*. Available at https://www.itential.com/resource/ebook/redefining-network-configuration-compliance-across-hybrid-infrastructure/

[14] McGillicuddy S (2022) *Taking a Strategic Approach to Network Operations*. Available at https://media.bitpipe.com/io_16x/io_161947/item_2553630/NBT002b_NetBrain-WP_Final%20%281%29.pdf

[15] Palo Alto Networks (2020) *The State of SOAR Report, 2020*. Available at https://media.bitpipe.com/io_15x/io_154375/item_2268964/the-state-of-soar-report-2020.pdf

[16] Aviatrix (2021) *DevOps Guide to Multi-cloud Networking*. Available at https://media.bitpipe.com/io_15x/io_158772/item_2444655/devops-guide-to-multi-cloud-networking%20%281%29.pdf.

[17] Itential (2021) *Automating Multi-Cloud Networking*. Available at https://www.itential.com/solutions/automation-use-cases/multi-cloud-network-automation/#:~:text=Automating%20Multi%2DCloud%20Networking&text=By%20leveraging%20the%20right%20automation,automate%20the%20Network%20of%20Clouds

[18] Verizon (2021) *The future of networking is here*. Available at https://media.erepublic.com/document/Network-as-a-Service_Solution_Brief.pdf

[19] Miller LC (2021) *Data Center and Hybrid Cloud Security – E-Book.* https://www.paloaltonetworks.com/resources/ebooks/data-center-and-hybrid-cloud-security-for-dummies

[20] Vertocci B (2021) *JSON Web Token (JWT) Profile for OAuth 2.0 Access Tokens*. (Internet Engineering Task Force (IETF) Network Working Group), IETF Request for Comments (RFC) 9068. https://datatracker.ietf.org/doc/html/rfc9068

[21] ColorTokens (2022) *What is Micro-segmentation?* Available at https://colortokens.com/micro-segmentation/

[22] Mandal A (2020) *Microsegmentation – the quintessential architecture for Zero Trust*. Available at https://medium.com/@anandadip/microsegmentation-the-quintessential-architecture-for-zero-trust-344715990c8e

[23] Kollimarla S (2021) *How Micro-Segmentation for Data Centers Works*. Available at https://colortokens.com/blog/data-center-micro-segmentation/

[24] Palo Alto Networks (2021) *Prisma Cloud Identity-based Microsegmentation*. Available at https://media.bitpipe.com/io_15x/io_157597/item_2439737/prisma-cloud-identity-based-microsegmentation.pdf

[25] Slattery T (2022) *How to implement network segmentation for better security*. Available at https://www.techtarget.com/searchnetworking/tip/How-to-implement-network-segmentation-for-better-security

[26] Frazier S (2021) *Why the cyber EO made zero trust no longer a suggestion.* Available at https://federalnewsnetwork.com/federal-insights/2021/09/why-the-cyber-eo-made-zero-trust-no-longer-a-suggestion/

[27] Brasen S (2020) *Contextual Awareness: Advancing Identity and Access Management to the Next Level of Security Effectiveness.* Available at https://dbac8a2e962120c65098-4d6abce208e5e17c2085b466b98c2083.ssl.cf1.rackcdn.com/advancing-identity-access-management-to-next-level-security-effectiveness-pdf-7-w-7727.pdf

[28] Appgate (2020) *SDP and Risky Devices*. Available at https://www.appgate.com/blog/sdp-and-risky-devices-dynamic-controls-for-secure-access

[29] Srinivas S (2020) *Democratizing Zero Trust with an expanded BeyondCorp Alliance.* Available at https://cloud.google.com/blog/products/identity-security/google-cloud-announces-new-partners-in-its-beyondcorp-alliance

[30] Tanium (2021) *Tanium Insights: It's Time to Ditch the VPN for Zero Trust*. Available at https://site.tanium.com/rs/790-QFJ-925/images/EB-ZeroTrust.pdf

[31] Scheels C (2021) *VPN VS. ZTNA VS. SDP VS. NAC: What's the Difference?* Available at https://www.appgate.com/blog/vpn-vs-ztna-vs-sdp-vs-nac

[32] QTS (2020) *Driving Data Center Innovation with Microservices*. Available at https://media.bitpipe.com/io_15x/io_155464/item_2314862/QTS_Whitepaper_SDP.pdf

[33] Rose S, Borchert O, Mitchell S, Connelly S (2020) Zero Trust Architecture. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-207. https://doi.org/10.6028/NIST.SP.800-207

[34] Appgate (2021) *5 Steps for Successful VPN to ZTNA Migration*. Available at https://d3aafpijpsak2t.cloudfront.net/docs/VPN_to_ZTNA_migration_ebook-6.pdf

[35] Shread P (2020) *What is SASE and How does it Work?* Available at https://www.esecurityplanet.com/networks/sase/

[36] Fortinet (2022) *Required Capabilities for Effective and Secure SD-WAN: The Network Leader's Guide*. Available at https://www.fortinet.com/content/dam/maindam/PUBLIC/02_MARKETING/02_Collateral/eBooks/eb-network-leaders-guide-to-secure-SD-WAN.pdf.

[37] Mann T (2021) *AWS Cloud WAN Parries Google, Microsoft*. Available at https://www.sdxcentral.com/articles/news/aws-cloud-wan-parries-google-microsoft/2021/12/

[38] Mann T (2021) *Is Multi-Cloud SD-WAN's Final Destination?* Available at https://www.sdxcentral.com/articles/news/is-multi-cloud-sd-wans-final-destination/2021/12/

[39] Mann T (2021) *Google Cloud Drives SD-Underlays into Cisco SD-Wan*. Available at https://www.sdxcentral.com/articles/news/google-cloud-drives-sd-underlays-into-cisco-sd-wan/2021/03/

[40] Mann T (2021) *Fortinet Fortifies Microsoft Azure vWAN With SD-WAN Firewalls*. Available at https://www.sdxcentral.com/articles/news/fortinet-fortifies-microsoft-azure-vwan-with-sd-wan-firewalls/2021/11/

[41] Aviatrix (2021) *The Security Architect's Guide to Multi-Cloud Networking*. Available at https://media.bitpipe.com/io_15x/io_158772/item_2444655/security-architects-guide-multi-cloud-networking-v2%20%281%29.pdf

[42] Aviatrix (2020) *Multi-Cloud Networking*. Available at https://aviatrix.com/wp-content/uploads/2020/07/Multi-Cloud-Networking-by-Futuriom-July2020.pdf

[43] Robb D (2022) *Top Software-Defined SD-WAN Trends*. Available at https://www.enterprisestorageforum.com/networking/sd-wan-trends/

[44] TechTarget (2022) *4 Key SD-WAN Trends to Watch in 2022*. Available at https://media.bitpipe.com/io_14x/io_148038/item_2494980/4%20key%20SD-WAN%20trends%20to%20watch%20in%202022.pdf

[45] Doyle L (2020) *The pros and cons of SD-WAN and remote access*. Available at https://www.techtarget.com/searchnetworking/tip/The-pros-and-cons-of-SD-WAN-and-remote-access

[46] CATO (2021) *5 Questions to Ask Your SASE Provider*. Available at https://go.catonetworks.com/rs/245-RJK-441/images/5_Questions_to_Ask_Your_SASE_Provider.pdf