H3C SDN DFW技术白皮书

目 录

1 概述

1.1 产生背景

传统的集中式防火墙一般用来做边界防护，但是云数据中心不仅仅要求边界保护，还要求对内网虚机之间的流量进行安全防护，如果这些流量全部绕行到集中式防火墙，随着网络的升级和扩容，这种方式就难以满足大容量、高性能、可扩展的要求和挑战，容易形成性能瓶颈。因此，把安全功能嵌入数据中心内主机节点的分布式防火墙应运而生。

当前业界 SDN 控制器提供的嵌入式安全主要通过安全组的 ACL 功能实现。ACL 功能是通过在虚拟交换机上下发 ACL 规则来控制虚拟机的流量。在企业的网络规划日渐复杂的情况下，这种方式已经不能满足企业的需求，主要表现为：

- 仅检查当前报文的信息，不关心连接状态，安全性低。
- 部署方式复杂，不容易维护。

DFW（Distributed Firewall，分布式防火墙）是一种分布式状态监测防火墙，可记录并跟踪各种网络连接（如 TCP 连接等），并对各种类型的报文进行检查和处理。

1.2 技术优点

DFW 具有以下技术优点：

- 为整个数据中心提供了无所不在的安全防护，让安全机制既具有广泛性，又具有精确度；虚机之间的流量无须因为安全防护绕行，可扩展性好。
- 支持状态防火墙，不仅可对数据流量进行检查和处理，还可在 vSwitch 上建立状态连接表，根据数据包的状态执行相应的操作，提高安全性的同时，安全部署有效精简且易于实施。
- 支持更丰富和高效的匹配项，如支持匹配 TCP 标志位、支持匹配 IP 地址集，极大的提高了数据包的匹配效率。
- 同时支持白名单和黑名单。

2 技术实现

2.1 概念介绍

2.1.1 状态

报文在 vSwitch 上经过 DFW 处理后，会根据具体的协议类型生成以下四种状态：

- NEW：表示报文属于一条新的连接。
- ESTABLISHED：表示报文属于已建立的连接。
- RELATED：表示报文和一条已建立的连接相关联，比如 ICMP 目的不可达或者 FTP 数据报文。
- INVALID：表示不能识别报文状态，通常丢弃该报文。

规则可以匹配这些状态对虚机流量实施预定义的动作。

2.1.2 分布式防火墙

分布式防火墙由以下两个部分组成：

- VCF 控制器，作为 DFW 的管理平面，DFW 配置可通过登录 VCF 控制器的 Web UI 创建，或者由云管理平台调用 VCF 控制器提供的 Restful API 接口来创建。
- KVM 主机，作为 DFW 的数据平面，它从 VCF 控制器接收策略配置，并立即生效。KVM 主机负责对虚机的网络流量监测并执行相应的策略。

2.1.3 分布式防火墙策略

分布式防火墙策略是一条或者多条规则的集合。通过在虚拟端口或者端口组上引用策略可将分布式防火墙应用到指定的虚拟端口上。

2.1.4 规则

规则包含匹配条件和执行动作两部分。匹配条件可以是报文的源地址、目的地址、源端口号、目的端口号、源 IP 地址集、目的 IP 地址集等。动作包含允许、拒绝、返回或者跳转到子策略。分布式防火墙根据规则的匹配条件对报文进行匹配，并对匹配成功的报文执行相应的动作。

2.1.5 分布式防火墙子策略

与分布式防火墙策略相比，分布式防火墙子策略不能直接绑定到虚拟端口或者端口组上。只能被策略中规则的动作引用来间接生效。

2.1.6 IP 地址集

IP 地址集为一组 IP 地址的集合，或 IP 地址和协议端口号的集合。分布式防火墙策略或子策略可在添加规则时引用 IP 地址集，可实现一个 IP 地址集中的所有 IP 地址和端口号共用一个规则，从而简化了规则的匹配条件，提供了更高效的查找和匹配。

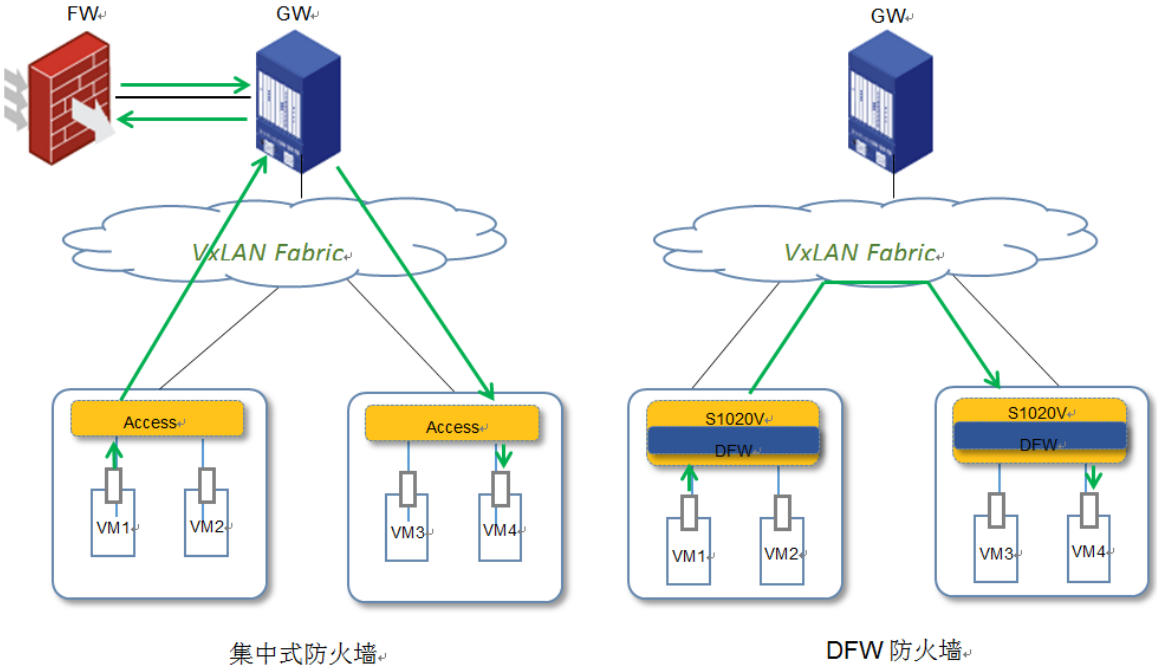
IP 地址集支持两种类型，Net 和 Net&Port。

- Net 类型仅包含 IP 地址。
- Net&Port 类型同时包含 IP 地址和协议端口号。

2.2 运行机制

相对于传统的集中式防火墙，DFW 为云数据中心提供了内部虚机之间的安全防护，虚机之间的流量无需再绕行网关至防火墙处理，消除了传统边界安全机制的低效流量模式，如图 1 所示。

图1 DFW 机制示意图



通过在虚拟端口或端口组的入、出方向绑定分布式防火墙策略，可将分布式防火墙应用到虚拟端口的指定方向。虚拟端口或端口组的每个方向都可绑定多个分布式防火墙策略。

如图 2 和图 3 所示，以 VM4 为数据中心的其他 VM 提供访问服务为例，图 2 和图 3 分别为静态安全组 ACL 方式和 DFW 方式的实现。二者的差异主要表现在：

- 静态安全组 ACL 方式配置复杂，不但需要在 VM4 的入方向配置规则放行源 IP 为 VM1~VM3 的虚机，还需在 VM4 的出方向配置规则放行目的 IP 为 VM1~VM4 的虚机。DFW 方式的实现较为简单，利用状态连接表，在出方向仅需添加规则放行已经建立连接的报文。
- 静态安全组 ACL 方式不能阻止 VM4 主动访问其他的虚机，DFW 方式通过只放行已经建立连接的报文，实现了 VM4 不能主动访问其他虚机，提高了安全性。

图2 静态安全组 ACL 运行机制图

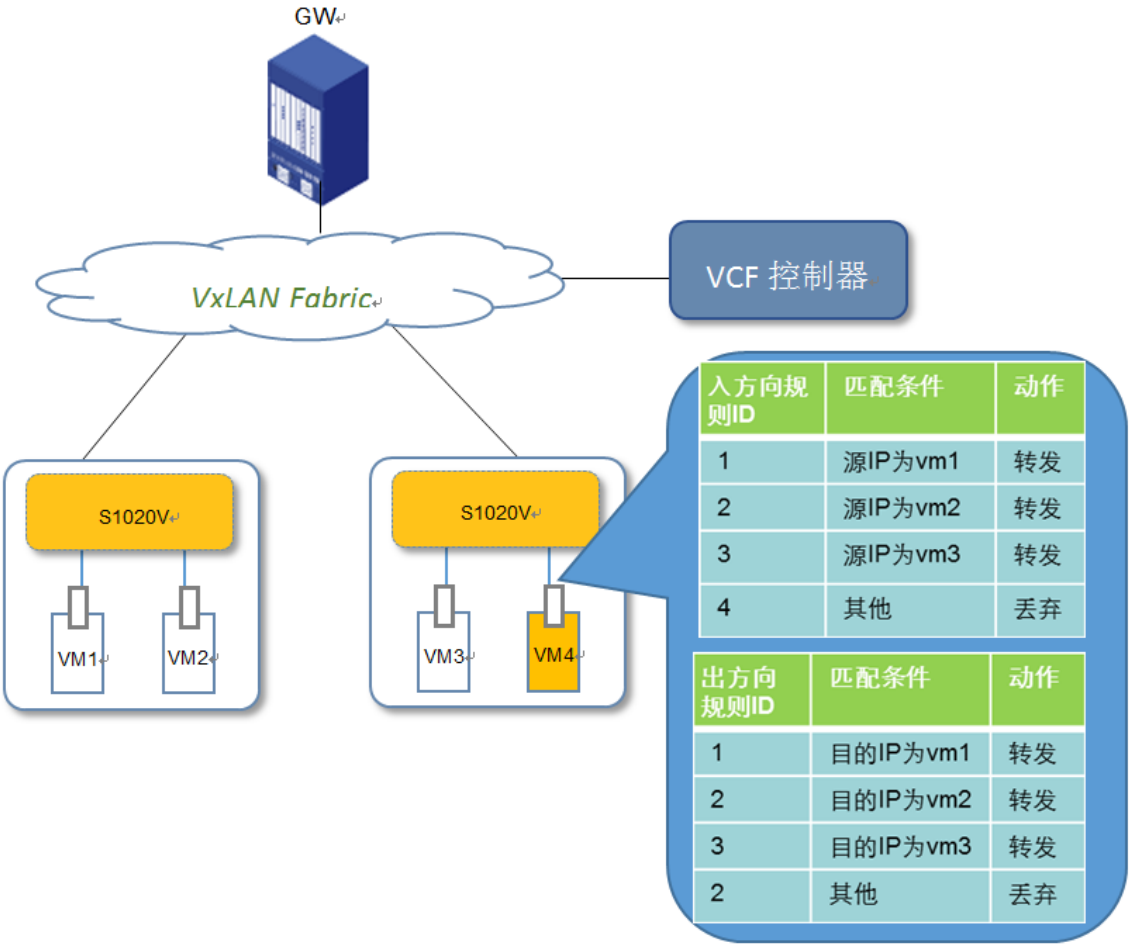
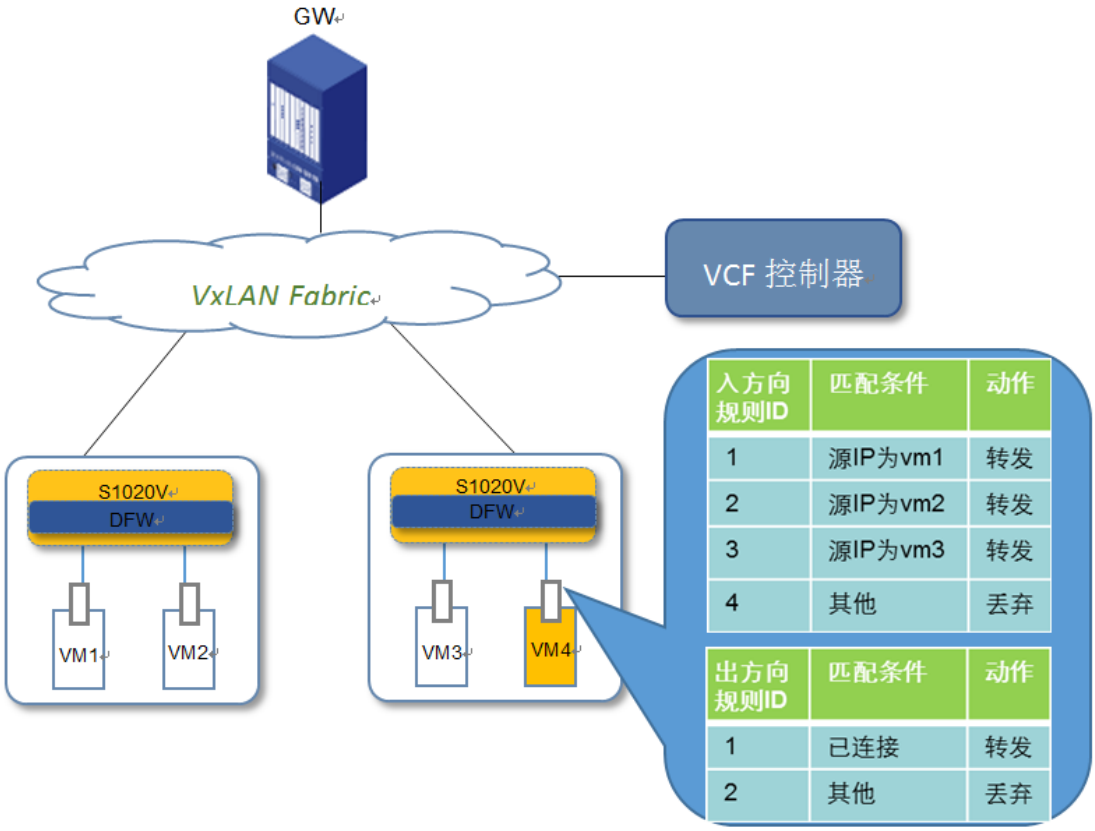


图3 DFW 状态机制图



KVM 主机作为 DFW 的数据平面，负责监控虚机的流量，并对符合 DFW 策略的报文执行相应的处理，如(3)图 4 所示，具体的处理流程如下：

(1) 收到报文后，使用第一条策略进行匹配，按照该策略中配置的规则进行逐条匹配：

- 如果在第一条策略中匹配到某条规则，则立即执行规则中的动作：
  - 动作为允许，则转发该报文，并结束匹配流程。
  - 动作为拒绝，则丢弃该报文，并结束匹配流程。
  - 动作为返回，则结束当前策略的匹配流程，由下一条策略继续匹配。
  - 动作为跳转至子策略，则结束当前策略的匹配流程，由指定的子策略继续匹配。
- 如果在第一条策略中未匹配任何规则，则执行该策略的默认动作：
  - 动作为允许，则转发该报文，并结束匹配流程。
  - 动作为拒绝，则丢弃该报文，并结束匹配流程。
  - 动作为返回，则结束当前策略的匹配流程，由下一条策略继续匹配。

(2) 报文跳转至子策略匹配后，则按照子策略中配置的规则进行逐条匹配：

- 如果在子策略中匹配到某条规则，则立即执行规则中的动作：
  - 动作为允许，则转发该报文，并结束匹配流程。
  - 动作为拒绝，则丢弃该报文，并结束匹配流程。

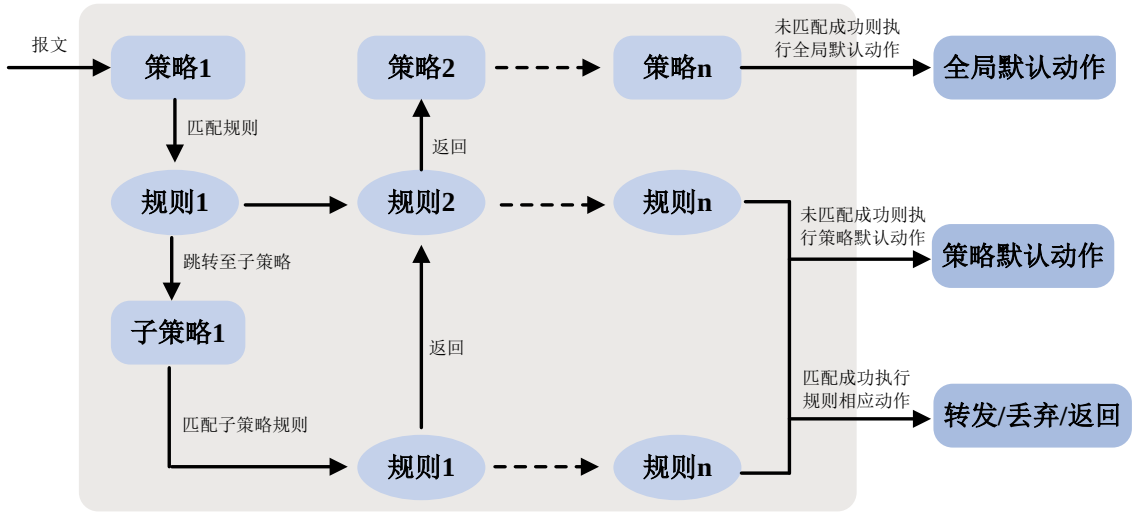
- 动作为返回，则结束当前子策略的匹配流程，由跳转前的下一条规则继续匹配。

- 如果在子策略中未匹配任何规则，则执行该策略的默认动作：
  - 动作为允许，则转发该报文，并结束匹配流程。
  - 动作为拒绝，则丢弃该报文，并结束匹配流程。
  - 动作为返回，则结束当前策略的匹配流程，由跳转前的下一条规则继续匹配。

(3) 如果报文未匹配虚拟端口绑定的所有的策略，则执行全局默认动作：

- 动作为允许，则转发该报文，并结束匹配流程。
- 动作为拒绝，则丢弃该报文，并结束匹配流程。

图4 报文转发流程图



另外，分布式防火墙还具备以下机制：

- 在虚拟机迁移时，分布式防火墙策略会同步迁移；在虚拟机删除时会自动移除策略。
- H3C VCF 控制器与 H3C S1020V 的连接中断重连时，H3C VCF 控制器会与 H3C S1020V 进行配置的平滑同步。
- 分布式防火墙功能和安全组 ACL 功能可以同时配置，同时生效。

2.3 应用限制

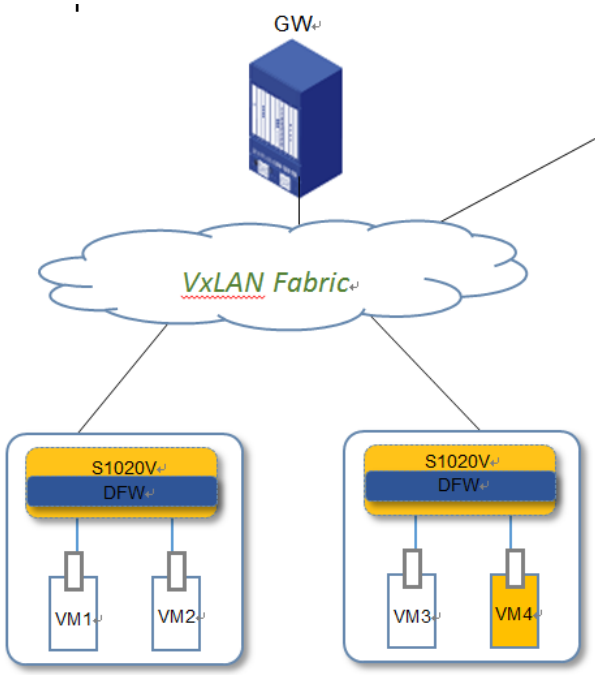
- 网络 Overlay 不支持分布式防火墙功能。
- 主机 Overlay 下 ESXi 虚拟化主机不支持分布式防火墙 DFW。
- 不支持应用层的检测。

3 典型组网应用

3.1 云数据中心部署DFW的典型应用

如图 5 所示，使 VM4 作为数据中心的 Sever 为 VM1~3 提供访问服务，但是 VM4 不可主动访问 VM1~3；VM1~3 只能访问 VM4，不能访问其他站点。

图5 典型组网应用示意图



VM4	规则ID	匹配条件	动作
入方向	1	源IP为VM1	转发
	2	源IP为VM2	转发
	3	源IP为VM3	转发
	4	其他	丢弃
出方向	1	已连接	转发
	2	其他	丢弃

VM1~3	规则ID	匹配条件	动作
入方向	1	已连接	转发
	2	其他	丢弃
出方向	1	目的IP为VM4	转发
	2	其他	丢弃

如图 5 所示，可通过以下步骤为 VM4 配置 DFW 策略：

(1) 在 VCF 控制器上配置入方向分布式防火墙策略，在策略中配置规则 1，定义源 IP 地址为 VM1，动作为转发。配置规则 2，定义源 IP 地址为 VM2，动作为转发。配置规则 3，定义源 IP 地址为 VM3，动作为转发。配置规则 4，动作为丢弃。

(2) 在 VCF 控制器上配置出方向分布式防火墙策略，在策略中配置规则 1，定义匹配条件为已连接，动作为转发。配置规则 2，动作为丢弃。

(3) 将配置的分布式防火墙策略分别绑定至 VM4 对应虚拟端口的入出方向。

通过以下步骤为 VM1~3 配置 DFW 策略：

(1) 在 VCF 控制器上配置入方向分布式防火墙策略，在策略中配置规则 1，定义匹配条件为已连接，动作为转发。配置规则 2，动作为丢弃。

(2) 在 VCF 控制器上配置出方向分布式防火墙策略，在策略中配置规则 1，定义目的 IP 地址为 VM4，动作为转发。配置规则 2，动作为丢弃。

将配置的分布式防火墙策略分别绑定至 VM1~3 对应虚拟端口的入出方向。