# WAFS 数据引接系统架构分析

石磊 民航局空管局航空气象中心 北京市 100122

**摘要：**本文分别从网络安全、系统架构和数据引接三个方面对 WAFS 数据引接系统所涉及的信息安全、高可用性服务以及数据获取及处理流程进行全面介绍，其中着重对四类 WAFS 数据（OPMET、BUFR、PNG、GRIB）的获取、传输和解码处理过程进行深入解析。

**关键词：**WAFS 高可用性服务 共享存储 信息安全 数据解码

## 1. 引言

随着世界区域预报系统（WAFS）数据下发方式由卫星广播转为 Internet 方式，为能够安全、稳定、高效的引接 WAFS 系统航空气象数据及产品，设计并实现了这一系统。在充分满足业务需求的前提下，即兼顾总成本又达到一定高可用性的同时，并且对整个系统采取了多种安全防护措施，最终达到信息安全相关要求，同时对获取的各类数据进行加工处理，使之能够被民航气象数据库系统识别和使用。

作者简介
石磊（1980/09/21），男，河北，工程师，主要从事民航气象信息系统日常管理和运行维护工作。

## 2. 网络安全架构

### 2.1. 网络拓扑

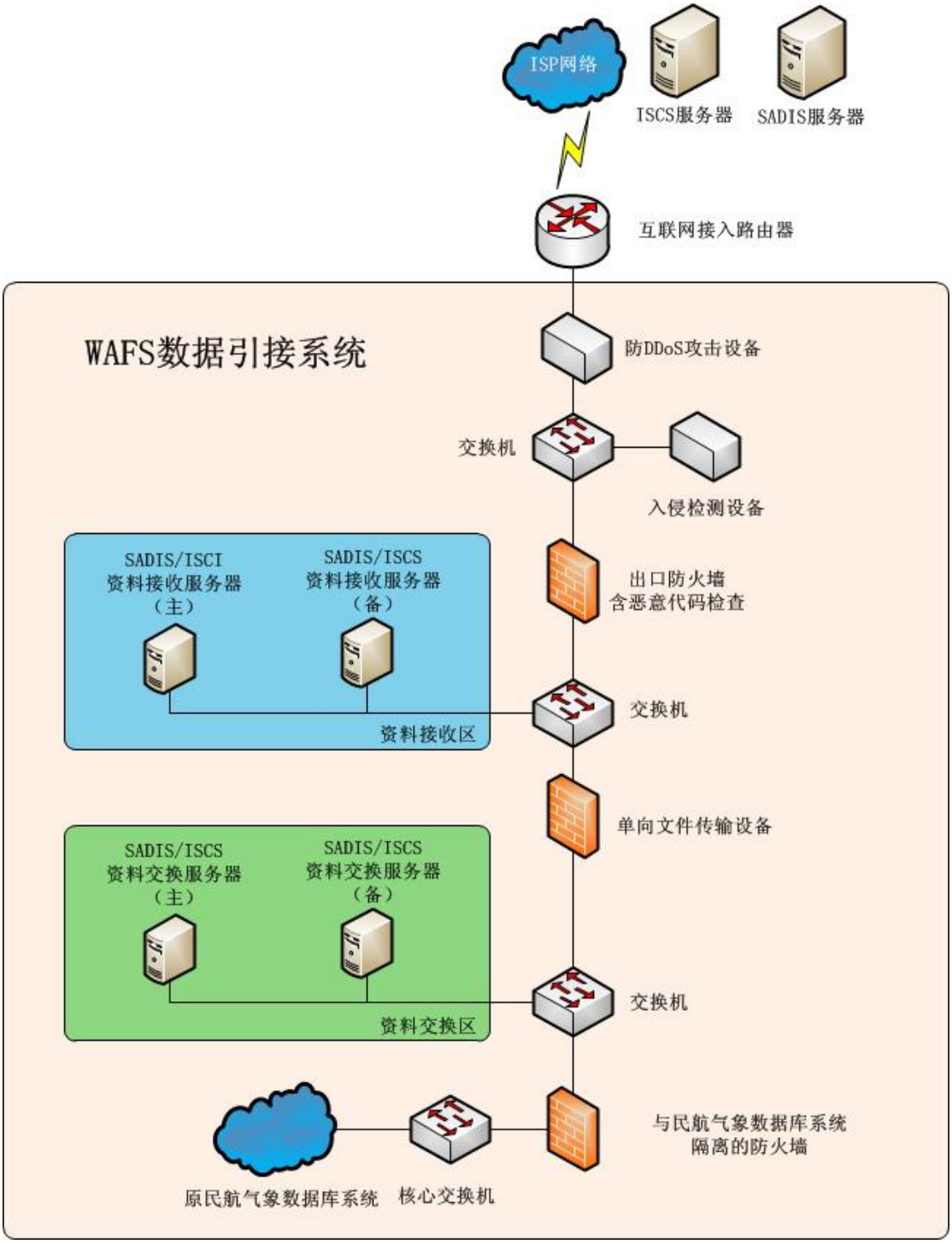


图 1

系统网络结构主要分为三个区域：外网区、资料接收区、资料交换区。其中配备的网络安全设备主要有：防 DDoS 攻击设备、IPS（入侵防御系统）、单向传输设备、防火墙。整个系统应用纵深防御策略，从网络、操作系统、应用和数据等方面，分层多处部署防护措施。以系统承载的气象数据为核心，注重从局部计算环境防止内部的威胁，从各种不同类型的边界区域，联合采用多种安全技术整体防御外部威胁。采取减轻风险的技术途径，尽可能减少

安全漏洞、尽可能阻止安全威胁、尽可能减小安全事件造成的损失，尽可能加强安全防范手段，通过保护、检测、响应和强化安全管理等行为把安全风险减轻到最低程度。

基于以上设计原则，构建系统时选用了以下信息安全设备。

## 2.2. 抗拒绝服务系统（ADS）

DDoS（分布式拒绝服务）通常是指黑客通过控制大量互联网上的机器（通常称为僵尸机器），在瞬间向一个攻击目标发动潮水般的攻击。大量的攻击报文导致被攻击系统的链路被阻塞、应用服务器或网络防火墙等网络基础设施资源被耗尽，无法为用户提供正常业务访问[1]。而 ADS 能够及时发现背景流量中各种类型的攻击流量，针对攻击类型迅速对攻击流量进行拦截，保证正常流量的通过。

- DDoS 攻击与 DoS 攻击的区别：

DDoS 的攻击策略侧重于通过很多“僵尸主机”（被攻击者入侵过或可间接利用的主机）向受害主机发送大量看似合法的网络包，从而造成网络阻塞或服务器资源耗尽而导致拒绝服务，分布式拒绝服务攻击一旦被实施，攻击网络包就会犹如洪水般涌向受害主机，从而把合法用户的网络包淹没，导致合法用户无法正常访问服务器的网络资源。而 DoS 则侧重于利用主机特定漏洞进行攻击导致系统网络栈失效及崩溃、主机死机等，进而无法提供正常的网络服务功能，从而造成拒绝服务。就这两种拒绝服务攻击而言，危害较大的主要是 DDoS 攻击，原因是很难防范，至于 DoS 攻击， 通过给主机服务器打补丁或安装防火墙软件就可以很好地防范。

## 2.3. 入侵防御系统（IPS）

IPS 是英文“Intrusion Prevention System”的缩写，中文意思是入侵防御系统。随着网络攻击技术的不断提高和网络安全漏洞的不断发现，传统防火墙和 IDS 技术已经无法应对一些安全威胁。在这种情况下，IPS 技术应运而生。对于部署在数据转发路径上的 IPS，可以根据预先设定的安全策略，对流经的每个报文进行深度检测(协议分析跟踪、特征匹配、流量统计分析、事件关联分析等)，如果一旦发现隐藏于其中网络攻击，可以根据该攻击的威胁级别立即采取抵御措施，这些措施包括(按照处理力度)：向管理中心告警；丢弃该报文；切断此次应用会话；切断此次 TCP 连接。

- IPS 与 IDS 的区别：

IDS 是英文“Intrusion Detection Systems”的缩写，中文意思是“入侵检测系统”。就是依照一定的安全策略，对网络、系统的运行状况进行监视，尽可能发现各种攻击企图、攻击行为或者攻击结果，以保证网络系统资源的机密性、完整性和可用性。

IPS 和 IDS 是两类产品，并不存在互相替代的可能。IPS 注重的是网络安全状况的监管。IDS 关注的是对入侵行为的控制。IDS 的核心价值在于通过对全网信息的分析，了解信息系统的安全状况，进而指导信息系统安全建设目标以及安全策略的确立和调整，而 IPS 的核心价值在于安全策略的实施—对黑客行为的阻击；IDS 需要部署在网络内部，监控范围可以覆盖整个子网，包括来自外部的数据以及内部终端之间传输的数据，IPS 则必须部署在网络边界，抵御来自外部的入侵，对内部攻击行为无能为力。

## 2.4. 单向传输设备

网闸（GAP）全称安全隔离网闸。它是一种由带有多种控制功能专用硬件在电路上切断网络之间的链路层连接，并能够在网络间进行安全适度的应用数据交换的网络安全设备。安全隔离网闸是由软件和硬件组成。 隔离网闸分为两种架构，一种为双主机的 2+1 结构，另一种为三主机的三系统结构。2+1 的安全隔离网闸的硬件设备由三部分组成:外部处理单元、内部处理单元、隔离安全数据交换单元。安全数据交换单元不同时与内外网处理单元连接，

为 2+1 的主机架构。隔离网闸采用 SU-Gap 安全隔离技术，创建一个内、外网物理断开的环境。三系统的安全隔离网闸的硬件也由三部分组成：外部处理单元（外端机）、内部处理单元（内端机）、仲裁处理单元（仲裁机），各单元之间采用了隔离安全数据交换单元。

- 单向传输设备与防火墙的区别：

防火墙一般在进行 IP 包转发的同时，通过对 IP 包的处理，实现对 TCP 会话的控制，但是对应用数据的内容不进行检查。这种工作方式无法防止泄密，也无法防止病毒和黑客程序的攻击。防火墙是保证网络层安全的边界安全工具（如通常的非军事化区），而安全隔离网闸重点是保护内部网络的安全。因此两种产品由于定位的不同，因此不能相互取代。

## 3. 双机系统架构

WAFS 数据引接系统数据下载和处理部分均分别采用 2 台 IBM X3650 服务器组成两节点高可用集群系统，每台主机分别绑定两块网卡并以主备工作模式用于对外网络服务和 DRBD 数据通信，使用主机内部管理接口作为 Fence 设备，从而实现 RHCS+DRBD 架构的双节点软件共享存储 HA 集群系统。

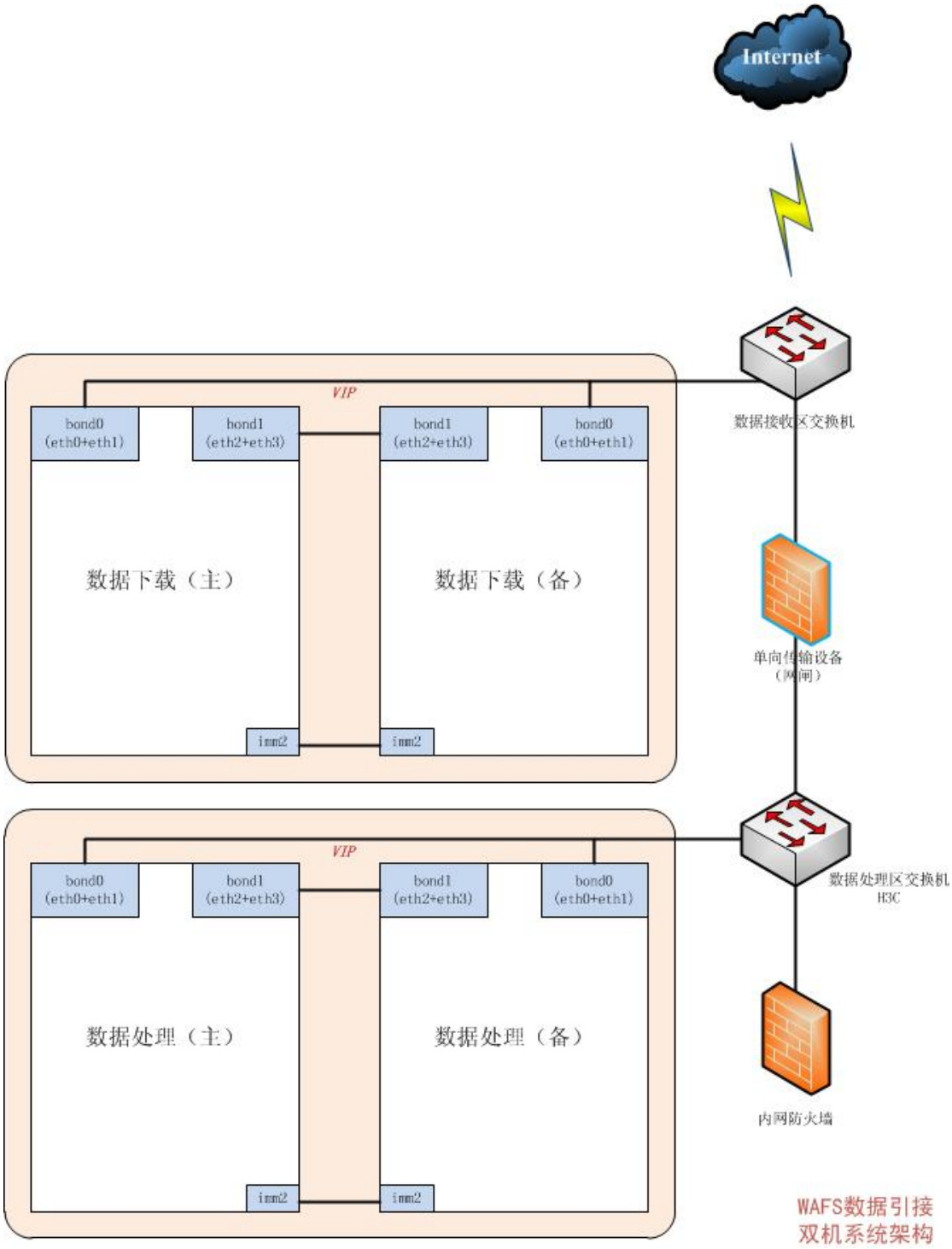


图 2

## 3.1. RHCS

RHCS 即 REDHAT CLUSTER SUITE,中文意思即红帽集群套件。它是一套综合的软件组件，可以通过在部署时采用不同的配置，以满足用户对高可用性，负载均衡，可扩展性，文件共享和节约成本的需要。RHCS 主要由以下几部分组成：

- 集群管理器（cman）

Cluster manager 简称 CMAN（集群管理器的缩写），运行在集群的各个节点上，为 RHCS 提供集群管理任务。集群管理负责集群仲裁和成员资格管理。CMAN 在高可用性附加组件中为

Red Hat Enterprise Linux 6 执行集群管理。CMAN 是发布的集群管理器，它在每个节点中运行。集群中的所有节点中都会部署集群管理。

CMAN 通过监控集群节点数目来了解集群仲裁。如果多过半数的节点是活跃的，那么集群就具有仲裁。如果一半（或更少）节点是活跃的，集群就没有仲裁，所有的集群活动将停止。集群仲裁会阻止 "split-brain" 条件的发生 — 此时同一集群的两个实例在运行。"split-brain" 条件将允许每个集群实例访问集群资源而无需了解另外一个实例，这样会破坏集群的完整性。

- 锁管理（DLM）

Distributed Lock Manager，简称 DLM（分布式锁管理器），锁管理是常见的集群基础结构服务，它提供一种同步集群基础结构组件对共享资源的访问的机制。在 Red Hat 集群中，DLM 是锁管理器。顾名思义，DLM 是一个分布式的锁管理器，它运行在集群的每个节点中；锁管理分布于集群的所有节点中。GFS2 和 GLVM 都使用锁管理器提供的锁。GFS2 用它来同步对文件系统元数据（在共享存储中）的访问。CLVM 使用它来同步对 LVM 卷和卷组（也在共享存储中）的更新。另外，rgmanager 使用 DLM 同步服务状态。

- Fence

Fencing(保护)将节点从集群的共享存储里断开。Fencing 切断和共享存储之间的 I/O，因此保证了数据的完整性。集群基础结构通过守护进程 fenced 来执行保护。

当 CMAN 决定某个节点已经失败后，它将通知其它集群基础结构组件该节点已失败。在收到通知后，fenced 将保护（fence）故障节点。其它集群基础结构组件将决定采取什么行动 — 也就是说，它们执行所有必须执行的恢复。例如，当被告知节点故障时，DLM 和 GFS2 将暂停活动，直到它们检测到 fenced 已经完成对故障节点的保护。当确认故障节点已经被保护时，DLM 和 GFS2 会执行恢复。DLM 释放对失败节点的锁定；GFS2 恢复故障节点的日志。

保护程序根据集群配置文件决定使用哪种保护方法。集群配置文件用两个关键元素定义了保护方法：保护代理（fencing agent）和保护设备（fencing device）。保护程序调用这里定义的保护代理。而保护代理则通过保护设备来保护节点。当保护过程结束时，保护程序将通知集群管理器。

- 高可用性服务管理

高可用性服务管理提供在高可用性附加组件中创建和管理高可用性集群服务（cluster service）的功能。高可用性附加组件中高可用性服务管理的关键组件是 rgmanager，它实现了 off-the-shelf 应用程序的冷故障切换（cold failover）。在高可用性附加组件中使用其它集群资源配置应用程序来组成高可用性的集群服务。高可用性集群服务可以从一个节点故障切换到另外一个节点，而不会对集群客户端有明显的影响。当某个集群节点发生故障或集群系统管理员将服务迁移（例如，需要对这个节点进行预定的维护时）到另外一个节点时会发生集群服务的故障切换。

要创建高可用性服务，您必须在集群配置文件中进行配置。集群服务由集群资源组成。集群资源是您在集群配置文件中创建和管理的构建块（building block） — 例如，IP 地址、应用程序初始化脚本或者 Red Hat GFS2 共享分区。

您可使用故障切换域（failover domain）关联集群服务。故障切换域是有资格运行特定的集群服务的集群节点的一个子集

集群服务在某个时间只能在一个节点中运行，这样可以维护数据的完整性。您可以在故障效切换域里指定故障切换的优先级。指定故障切换的优先级是通过为故障切换域里的每个节点分配优先级别实现的。优先级决定故障切换的顺序 — 决定集群服务应该切换到哪个节点。如果您没有指定优先级，集群服务将可能切换至故障切换域里的任意节点。而且，您可

以指定是否限制集群服务在其相关的故障切换域里的节点中运行。（如果与非限制性故障切换域相关联，在故障切换域里的成员都不可用的时候，集群服务可以在任意节点中启动。）

## 3.2. DRBD

Distributed Replicated Block Device(DRBD)是一个基于软件的、彼此独立的主机之间块设备内容镜像存储复制解决方案。镜像数据以实时、透明、同步或异步的方式被复制。DRBD 的核心功能通过 Linux 内核实现，它靠近并工作于系统 I/O 栈的底层。

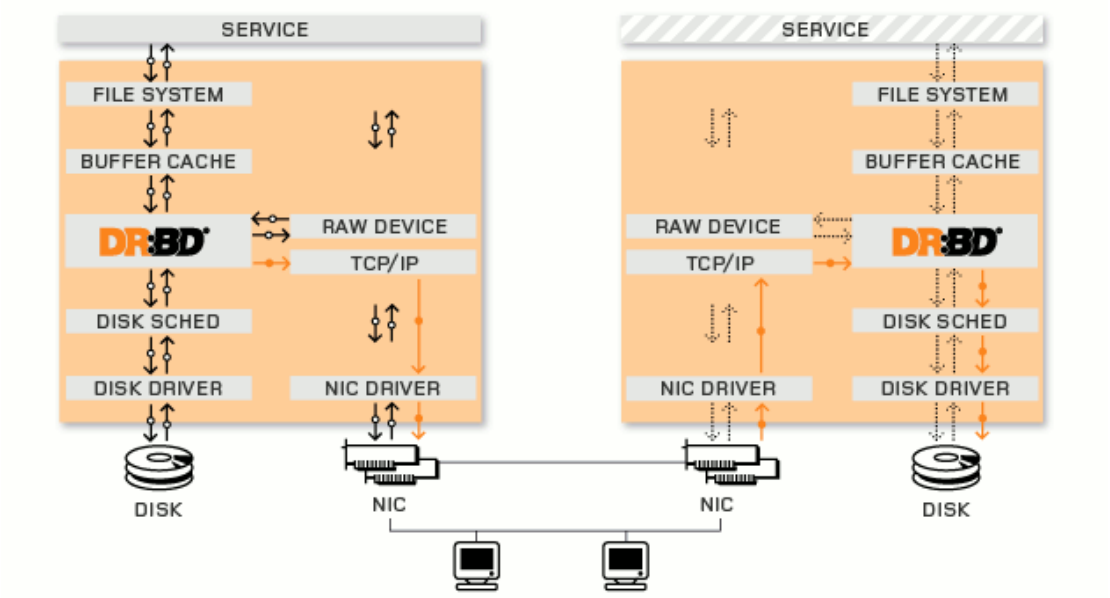


图 4

DRBD 的主要概念：

- 资源

在 DRBD 中，资源是一个集合术语，指与某个复制数据集相关联的所有方面。包括：

- 资源名：任意除空格外的 US-ASCII 字符。
- 卷：每个资源都是由一个或多个共享一个公共复制流的卷组成的复制组，DRBD 确保数据可以真实写入到资源里的所有卷。卷号从 0 开始，在一个资源里最大可以到 65535 个卷，每个卷都包含有复制数据集和由 DRBD 使用内部使用元数据集。
- DRBD 设备：是一个由 DRBD 管理的虚拟块设备。
- 连接：是指两台主机之间共享一个复制数据集的通信链路。

- 资源角色

在 DRBD 中每一个资源都具备一个角色，可以是“主”或“从”两种。

- 处于“主”角色中的 DRBD 设备可以无限制的进行“读”和“写”操作。可以创建、挂载文件系统，可以以 raw 或直接 I/O 方式访问块设备。
- 处于“从”角色中的 DRBD 设备接受来自对等结点的所有更新，但不能获得对其的完全访问权限，不能被应用程序访问，只能进行“读”访问但不能进行“写”访问。

DRBD 的主要功能：

- 单主模式：同一时刻，集群中只能有一个成员拥有。一般应用于典型 HA 集群中。
- 双主模式：同一时刻，具备“主”角色的资源可以被集群中的两个结点同时拥有，它们可以并行访问数据，这种模式下，必须使用诸如 GFS 或 OCFS2 等具备分布式锁管理器的共享集群文件系统。一般应用于负载均衡集群中。

复制模式：

- 协议 A：异步复制协议。在主结点本地写成功后会将数据包放到本地 TCP 发送缓存中，在集群故障切换时，数据可能出现丢失。
- 协议 B：内存同步（半同步）复制协议。在主结点本地写成功后会将数据会同时发送到对等结点并保证其确实到达，在一般情况下，集群故障切换时，数据不会出现丢失。
- 协议 C：同步复制协议。在主结点本地的写操作会等待本地和远程磁盘写操作成功确认。这种方式不会出现任何数据丢失。

一般都选用协议 C。

# 4. 应用程序架构

WAFS 数据引接系统应用程序分为数据获取和数据处理两部分，使用 BASH 脚本配合 AWK、SED 和 WGET 等开源工具实现，通过操作系统 Crontab 方式定时启动，对英、美 WAFS 网站中的数据进行实时镜像，并根据不同种类数据特点及生成时间进行特别检测，自动扫描下载程序运行和数据获取状态，智能判断程序各部分运行情况，中止僵死进程，提交新任务作业，保障在各种极端情况下也能及时获取最新数据。

数据获取并传输至传输区后，数据处理部分程序对各种数据进行分类处理，对文件名和文件格式进行规范，使之符合民航气象数据库系统要求；对 OPMET 进行初步过滤；对 BUFR、GRIB 数据进行解码和转换处理等。

## 4.1. 数据获取部分

数据下载程序，OPMET 下载为 1 或 5 分钟运行一次，PNG、BUFR、GRIB 下载为 15 分钟运行一次；各数据下载程序在运行前都会事先检测操作系统内是否存在已运行程序，防止多个下载程序同时运行占用系统资源和网络带宽情况的出现；系统专门配有 WGET 进程检测程序，1 分钟运行一次，进一步检测操作系统中是否存在超时或死锁数据下载程序，保证整个系统运行的持续和稳定及数据的及时性和完整性；文件服务器状态检测程序 1 小时运行一次；文件系统清洁程序每日运行一次。

- OPMET 数据获取

英、美 OPMET 数据分别取自 OPMET_SET_OF_5MIN_FILES 目录和 OPMET-MINUTE 目录，针对航空气象报文的实时和及时性要求，程序启动后，仅会下载最近 30 分钟内的数据文件。

- PNG

英、美 PNG 数据分别取自 SIGWX_PNG 目录和 PNG 目录，数据是一天四个时次更新。

- BUFR

英、美 BUFR 数据分别取自 BUFR 目录和 BUFR/KKCI 目录，数据是一天四个时次更新。

- GRIB

英、美 GRIB2 数据分别取自其服务器上 GRIB2/COMPRESSED/EGRR 目录（每个时次 11 个文件，约 16MB）、GRIB2/COMPRESSED/KWBC 目录（每个时次 11 个文件，约 22MB）和 GRIB2/EGRR 目录（每个时次 11 个文件，约 16MB）、GRIB2/KWBC 目录（每个时次 11 个文件，约 22MB）；

英、美 GRIB1 数据分别取其自服务器上 GRIB1/EGRR 目录（每个时次 6 个文件，约 10MB）、GRIB1/KWBC 目录（每个时次 6 个文件，约 12MB）和 GRIB/KWBC 目录（每个时次 6 个文件，约 20MB）。

GRIB 数据均一天四个时次更新（目前 GRIB2 数据也下载了其测试产品：CAT、CB、

ICE、INCLDTURB）。

### 4.2. 数据处理部分

数据处理程序对由网闸传送来的各类数据，OPMET 每 30 分钟处理一次；PNG、BUFR 和 GRIB 数据每 6 小时处理一次，持续 1 个小时，每 15 分钟再运行一次；

- OPMET

程序会对所获取的英、美 WAFS 网站上航空气象报文进行实时扫描处理，将其中由大陆各航站发出的民航报剔除后，以文本文件方式存储。

- PNG

程序会对所获取的英、美 WAFS 网站上 PNG 格式重要天气预告图文件名进行改名，使之符合民航气象数据库系统命名规范。

- BUFR

程序会对所获取的英、美 WAFS 网站上的 BUFR 数据源文件名进行改名并对源文件增加文件头，使之符合民航气象数据库系统命名和处理规范。

- GRIB

为保持使用原有 GRIB1 格式数据应用程序的正常运行，程序使用 NOAA 官方提供的 wgrib 工具程序对所获取的英、美 WAFS 网站上的 GRIB1 全球预报场数据文件进行拆解，将其中各物理量、区域、预报时效和高度层按照相关业务需求进行拆分（对其中 9 个物理量、8 个区域、11 个预报时效、22 个高度层进行拆解），最后，在对拆分出的各文件增加文件头后，根据民航气象数据库系统命名和处理规范使之能够为现有业务系统识别使用。

对于 GRIB2 格式数据，程序使用 NOAA 官方提供的 wgrib2 工具程序可以分别对英国 WAFS 网站提供的英国 WAFS GRIB2 数据及美国 WFAS GRIB2 和美国 WAFS 网站上提供的美国 WAFS GRIB2 数据及英国 WAFS GRIB2 数据进行实时处理，将其中各物理量（包括测试产品）、区域、预报时效和高度层按照相关业务需求进行拆分（对其中 11 个物理量、8 个区域、6 个预报时效、19 个高度层进行拆解），同时在对拆分出的各文件使用 NOAA 官方提供的 cnvgrib 工具程序转为 GRIB1 格式保存。最后，根据民航气象数据库系统命名和处理规范使生成的各产品能够为现有业务系统识别使用。

鉴于 WAFS GRIB 数据的重要作用，以上两个程序正常由系统计划任务自动运行，也可在特殊情况下，由系统管理人员手工启动，只需指定日期日间参数，程序即可将所需时次 GRIB 数据重新拆解生成。

## 5. 结论

WAFS 数据引接系统配备了多种网络安全设备，基于 RHCS 高可用性服务，配合 DRBD 分布式软件共享存储技术，以较低的成本在保证较高信息安全的环境下，实现系统运行及数据传输的可靠和高效。通过专门开发的一整套应用程序，对英美 WAFS 发布的航空气象数据产品文件的命名和格式进行规范化处理，使之符合民航气象数据库系统标准，使各数据可以顺畅进入现有业务系统，提供给各系统数据共享使用。尤其系统对于 WAFS GRIB2 数据的下载和分解转换工作的完成，既保证了现有业务应用的正常持续可靠运行，同时又引入了 WAFS GRIB2 格式数据及其测试产品，为航空气象各业务的进一步开展提供了坚实的基础。

参考文献

[1] 绿盟科技官方网站
[2] Red_Hat_Enterprise_Linux-6-Cluster_Suite_Overview-zh-CN
[3] The DRBD User’s Guide