

文章

[jieliang liu](#) · 一月 7, 2021 阅读大约需 28 分钟

精华文章---面向 Google Cloud Platform (GCP) 的 InterSystems IRIS 示例参考架构

Google Cloud Platform (GCP) 为基础架构即服务 (IaaS)

提供功能丰富的环境，其作为云提供完备的功能，支持所有的 InterSystems 产品，包括最新的 InterSystems IRIS 数据平台。

与任何平台或部署模型一样，必须留心以确保考虑到环境的各个方面，例如性能、可用性、操作和管理程序。本文将详细阐述所有这些方面。

以下概述和详细内容由谷歌提供，可在[此处](#)找到。

概述

GCP 资源

GCP 由一系列物理资产（例如计算机和硬盘驱动器）和虚拟资源（例如虚拟机（VM））组成，它们分布于谷歌遍布全球的数据中心。每个数据中心的位置都是一个泛区域。

每个区域都是地区的集合，这些地区在该区域内彼此分离。

每个地区都通过一个名称标识，名称由字母标识符和相应区域的名称组成。

这种资源分配带来众多优势，包括发生故障时提供冗余，以及通过将资源配置在客户端附近来减少延迟。这种分配也引入一些有关如何统筹资源的规则。

访问 GCP 资源

在云计算中，物理硬件和软件变成了服务。这些服务提供对基础资源的访问。在 GCP 上开发基于 InterSystems IRIS 的应用程序时，您可混合和匹配这些服务，组合它们来提供您所需的基础架构，然后添加您的代码来实现您要构建的方案。有关可用服务的详细信息，可在[此处](#)找到。

项目

您分配和使用的任何 GCP 资源必须属于一个项目。项目由设置、权限和其他描述应用程序的元数据组成。

根据区域和地区规则，单个项目中的资源能够轻松协作，例如通过内部网络进行通信。

每个项目包含的资源在项目边界上保持独立；您只能通过外部网络连接来互连它们。

服务交互

GCP 提供 3 种与服务和资源交互的基本方法。

控制台

Google Cloud Platform 控制台提供基于 web 的图形用户界面，供您管理 GCP 项目和资源。使用 GCP 控制台，您可创建新项目，或选择现有项目，可使用您在项目环境中创建的资源。您可以创建多个项目，因此您可以使用项目以任何对您有意义的方式分开您的工作。例如，如果您想确保只有某些团队成员可以访问项目中的资源，而所有团队成员可以继续访问另一个项目中的资源，则可以开始一个新项目。

命令行界面

如果您喜欢在终端窗口中工作，Google Cloud SDK 提供 gcloud 命令行工具，让您访问所需的命令。gcloud 工具可用于管理您的开发工作流程和 GCP 资源。有关 gcloud 的详细内容可在[此处](#)找到。

GCP 也提供 Cloud Shell，一种基于浏览器的 GCP 交互 shell 环境。可从 GCP 控制台访问 Cloud Shell。Cloud Shell 提供：

- 临时计算引擎虚拟机实例。
- 从 web 浏览器通过命令行访问实例。
- 内置代码编辑器。
- 5 GB 持久性磁盘存储。
- 预安装 Google Cloud SDK 和其他工具。
- 支持 Java、Go、Python、Node.js、PHP、Ruby 和 .NET 语言。
- Web 预览功能。
- 访问 GCP 控制台项目和资源的内置授权。

客户端库

Cloud SDK 拥有客户端库，让您轻松创建和管理资源。GCP 客户端库公开 API 有两个主要目的：

- 应用 API 提供对服务的访问。应用 API 针对支持的语言（例如 Node.js 和 Python）进行了优化。客户端库围绕服务隐喻而设计，因此您可以更自然地使用服务，并编写更少的样板代码。客户端库还提供身份验证和授权助手。有关详细信息可在[此处](#)找到。
- 管理 API 提供资源管理功能。例如，如果您想构建自己的自动化工具，可以使用管理 API。

您还可以使用 Google API 客户端库来访问产品的 API，如 Google Map、Google Drive 和 YouTube。有关 GCP 客户端库的详细信息可在[此处](#)找到。

InterSystems IRIS 示例体系结构

本文部分内容阐述了面向 GCP 的 InterSystems IRIS 部署示例，旨在为特定应用程序的部署抛砖引玉。这些示例可用作很多部署方案的指南。此参考体系结构拥有非常强大的部署选项，从最小规模的部署，到满足计算和数据需求的大规模可扩展工作负载，不一而足。

本文还介绍了高可用性和灾难恢复选项以及其他建议的系统操作。个体可对这些进行相应的修改以支持其组织的标准实践和安全策略。

针对您的特定应用，就基于 GCP 的 InterSystems IRIS 部署，您可联系 InterSystems 进一步探讨。

示例参考体系结构

以下示例体系结构按照容量和功能逐步升级的顺序讲述了几种不同的配置，分别为小型开发/生产/大型生产/分片集群生产。先从中小型配置讲起，然后讲述具有跨地区高可用性以及多区域灾难恢复的大规模可扩展性解决方案。此外，还讲述了一个将 InterSystems IRIS 数据平台的新分片功能用于大规模处理并行 SQL 查询的混合工作负载的示例。

小型开发配置

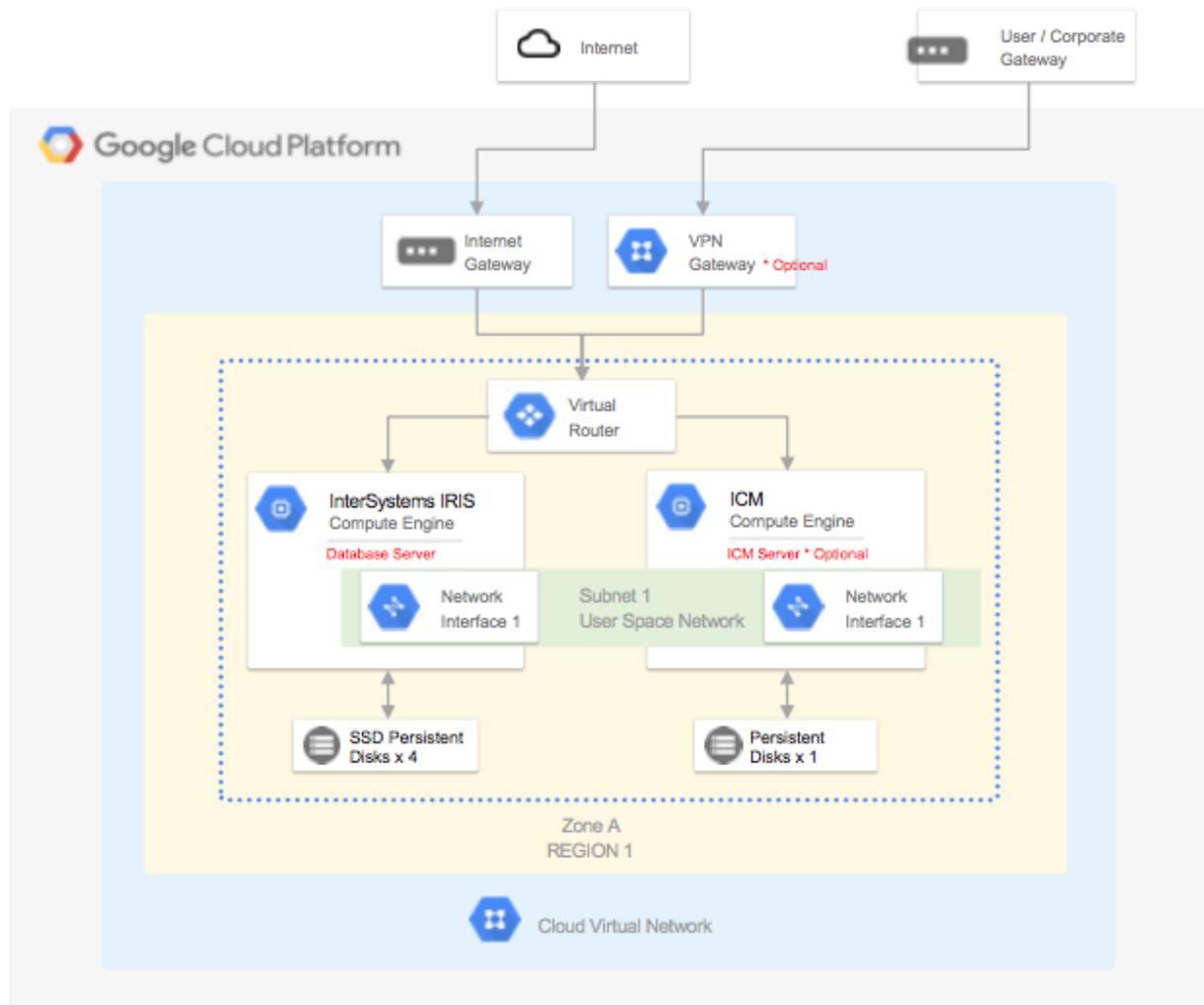
在本示例中，显示了一个能支持 10 名开发人员和 100GB 数据的小型开发环境，这基本是最小规模配置。只要适当地更改虚拟机实例类型并增加持久性磁盘存储，即可轻松支持更多的开发人员和数据。

这足以支持开发工作，并让您熟悉 InterSystems IRIS 功能以及 Docker 容器的构建和编排（如果需要的话）。小型配置通常不采用具有高度可用性的数据库镜像，但是如果需要高可用性，则可随时添加。

小型配置示例图

示例图 2.1.1-a 显示了图表 2.1.1-b 中的资源。其中包含的网关只是示例，可做相应地调整以适应您组织的标准网络实践。

Figure-2.1.1-a: Sample Small Development Architecture



下列 GCP VPC 项目资源是针对最小规模的配置提供的。 可根据需求添加或删除 GCP 资源。

小型配置 GCP 资源

?????????? GCP ????????

需要考虑适当的网络安全和防火墙规则，以防止对 VPC 的不必要访问。谷歌提供网络安全最佳做法供您入门使用，可在[此处](#)找到。

注意： VM 实例需要公共 IP 地址才能访问 GCP 服务。谷歌建议使用防火墙规则来限制这些 VM 实例的传入，尽管这种做法可能会引起一些问题。

如果您的安全策略确实需要内部 VM 实例，则您需要在网络上手动设置 NAT 代理和相应的路由，以便内部实例可以访问互联网。务必要明确，您无法使用 SSH 直接完全连接到内部 VM 实例。要连接到此类内部机器，必须设置具有外部 IP 地址的堡垒实例，然后建立隧道通过它。可以配置堡垒主机，以提供进入 VPC 的外部入口点。

有关堡垒主机的详细信息，可在[此处](#)找到。

产品配置

在本示例中，展示了一个规模较大的产品配置，其采用 InterSystems IRIS 数据库镜像功能来支持高可用性和灾难恢复。

此配置包括 InterSystems IRIS 数据库服务器同步镜像对，该镜像服务器在区域 1 内分为两个地区，用于自动故障转移，在区域 2 内的第三个 DR 异步镜像成员用于灾难恢复，以防万一整个 GCP 区域脱机。

InterSystems Arbiter 和 ICM 服务器部署在单独的第三个地区，以提高弹性。此示例体系结构还包括一组可选的负载均衡 web 服务器，用于支持启用 Web 的应用程序。这些使用 InterSystems 网关的 web 服务器可以根据需要进行缩放。

产品配置示例图

示例图 2.2.1-a 显示了图表 2.2.1-b 中的资源。其中包含的网关只是示例，可做相应地调整以适应您组织的标准网络实践。

建议将以下 GPC VPC 项目中的资源作为分片集群部署的最低配置。可根据需求添加或删除 GCP 资源。

产品配置 GCP 资源

?????????? GCP ????????

大型产品配置

在本示例中，提供了一个大规模可缩放性配置。该配置通过扩展 InterSystems IRIS 功能也引入使用 InterSystems 企业缓存协议 (ECP : EnterpriseCacheProtocol) 的应用程序服务器，实现对用户的大规模横向缩放。本示例甚至包含了更高级别的可用性，因为即使在数据库实例发生故障转移的情况下，ECP 客户端也会保留会话细节。多个 GCP 地区与基于 ECP 的应用程序服务器和部署在多个区域中的数据库镜像成员一起使用。此配置能够支持每秒数千万次的数据库访问和数万亿字节数据。

大型产品配置示例图

示例图 2.3.1-a 显示了图表 2.3.1-b 中的资源。其中包含的网关只是示例，可做相应地调整以适应您组织的标准网络实践。

此配置中包括一个故障转移镜像对，四个或更多的 ECP 客户端（应用程序服务器），以及每个应用程序服务器对应一个或多个 Web 服务器。故障转移数据库镜像对在同一区域中的两个不同 GCP 地区之间进行划分，以提供故障域保护，而 InterSystems Arbiter 和 ICM 服务器则部署在单独的第三地区中，以提高弹性。

灾难恢复扩展至第二个 GCP 区域和地区，与上一示例中的情况类似。如果需要，可以将多个 DR 区域与多个 DR 异步镜像成员目标一起使用。

建议将以下 GPC VPC 项目中的资源作为大型生产部署的最低配置。 可根据需求添加或删除 GCP 资源。

大型产品配置 GCP 资源

???????????? GCP ??????

采用 InterSystems IRIS 分片集群的生产配置

在此示例中，提供了一个针对 SQL 混合工作负载的横向缩放性配置，其包含 InterSystems IRIS 的新分片集群功能，可实现 SQL 查询和表跨多个系统的大规模横向缩放。 本文后面将详细讨论 InterSystems IRIS 分片集群及其功能。

采用 InterSystems IRIS 分片集群的生产配置

示例图 2.4.1-a 显示了图表 2.4.1-b 中的资源。
其中包含的网关只是示例，可做相应地调整以适应您组织的标准网络实践。

此配置中包括四个镜像对，它们为数据节点。每个故障转移数据库镜像对在同一区域中的两个不同 GCP 地区之间进行划分，以提供故障域保护，而 InterSystems Arbiter 和 ICM 服务器则部署在单独的第三地区中，以提高弹性。

此配置允许从集群中的任何数据节点使用所有的数据库访问方法。大型 SQL 表数据在物理上跨所有数据节点进行分区，以实现查询处理和数据卷的大规模并行。将所有这些功能组合在一起，就可以支持复杂的混合作业负载，比如大规模分析 SQL 查询及引入的新数据，所有这一切均在一个 InterSystems IRIS 数据平台中执行。

注意，上面图表中以及下表“资源类型”列中的术语“计算[Engine]”是一个表示 GCP（虚拟）服务器实例的 GCP 术语，将在本文的 3.1 节中做进一步介绍。
它并不表示或暗示本文后面所描述的集群体系结构中对“计算节点”的使用。

建议将以下 GPC VPC 项目中的资源作为分片集群部署的最低配置。可根据需求添加或删除 GCP 资源。

使用分片集群配置 GCP 资源的产品

???????????? GCP ??????

云概念简介

Google Cloud Platform (GCP) 为基础架构即服务 (IaaS) 提供功能丰富的云环境，使其具备完备的功能，支持所有的 InterSystems 产品，包括支持基于容器的 DevOps 及最新的 InterSystems IRIS 数据平台。与任何平台或部署模型一样，必须留心以确保考虑到环境的各个方面，例如性能、可用性、系统操作、高可用性、灾难恢复、安全控制和其他管理程序。本文档将介绍所有云部署涉及的三个主要组件：计算、存储和网络。

计算引擎（虚拟机）

GCP 中存在数个针对计算引擎资源的选项，以及众多虚拟 CPU 和内存规范及相关存储选项。在 GCP 中值得注意的一点是，对给定机器类型中 vCPU 数量的引用等于一个 vCPU，其是虚拟机监控程序层上物理主机中的一个超线程。

就本文档的目的而言，将使用 n1-standard * 和 n1-highmem * 实例类型，这些实例类型在大多数 GCP 部署区域中广泛可用。但是，对于将大量数据缓存在内存中的大型工作数据集而言，使用 n1-ultramem * 实例类型是不错的选择。除非另有说明，否则使用默认实例设置（例如实例可用性策略）或其他高级功能。有关各种机器类型的详细信息，可在[此处](#)找到。

磁盘存储

与 InterSystems 产品最直接相关的存储类型是持久性磁盘类型，但是，只要了解并适应数据可用性限制，本地存储可以用于高水平的性能。还有其他一些选项，例如云存储（存储桶），但是这些选项更特定于单个应用程序的需求，而非支持 InterSystems IRIS 数据平台的操作。

与大多数其他云提供商一样，GCP 对可与单个计算引擎关联的持久性存储施加了限制。这些限制包括每个磁盘的最大容量、关联到每个计算引擎的持久性磁盘的数量，以及每个持久性磁盘的 IOPS 数量，对单个计算引擎实例 IOPS 设置上限。此外，对每 GB 磁盘空间设有 IOPS 限制，因此有时需要调配更多磁盘容量才能达到所需的 IOPS 速率。

这些限制可能会随着时间而改变，可在适当时间与谷歌确认。

磁盘卷有两种类型的持久性存储类型：“标准持久性”磁盘和“SSD 持久性”磁盘。SSD 持久性磁盘更适用于那些要求低延迟 IOPS 和高吞吐量的生产工作负载。标准持久性磁盘对于非生产开发和测试或归档类型的工作负载，是一种更经济的选择。

有关各种磁盘类型及限制的详细信息，可在[此处](#)找到。

VPC 网络

强烈建议采用虚拟私有云 (VPC) 网络来支持 InterSystems IRIS 数据平台的各个组件，同时提供正确的网络安全控制、各种网关、路由、内部 IP 地址分配、网络接口隔离和访问控制。本文档中提供了一个详细的 VPC 示例。

有关 VPC 网络和防火墙的详细信息，可在[此处](#)找到。

虚拟私有云 (VPC) 概述

GCP VPC 与其他云提供商略有不同，其更加简单和灵活。可在[此处](#)找到各概念的比较。

在 GCP 项目中，每个项目允许有数个 VPC（目前每个项目最多允许 5 个），且创建 VPC 网络有两个选项——自动模式和自定义模式。

[此处](#)提供每个类型的详细信息。

在大多数大型云部署中，采用多个 VPC 将各种网关类型与以应用为中心的 VPC 进行隔离，并利用 VPC 对等进行入站和出站通信。有关适合您的公司使用的子网和任何组织防火墙规则的详细信息，强烈建议您咨询您的网络管理员。本文档不阐述 VPC 对等方面的内容。

在本文档提供的示例中，使用 3 个子网的单一 VPC 用于提供各种组件的网络隔离，以应对各种 InterSystems IRIS 组件的可预测延迟和带宽以及安全性隔离。

网络网关和子网定义

本文档的示例中提供了两种网关，以支持互联网和安全 VPN 连接。
要求每个入口访问都具有适当的防火墙和路由规则，从而为应用程序提供足够的安全性。
有关如何使用路由的详细信息，可在[此处](#)找到。

提供的示例体系结构中使用了 3 个子网，它们专与 InterSystems IRIS 数据平台一起使用。
这些单独的网络子网和网络接口的使用为上述 3
个主要组件的每一个提供了安全控制、带宽保护和监视方面的灵活性。有关各种用例的详细信息，可在[此处](#)找到。

有关创建具有多个网络接口的虚拟机实例的详细信息，可在[此处](#)找到。

这些示例中包含的子网：

用户空间网络用于入站连接用户和查询
分片网络用于分片节点之间的分片间通信
镜像网络通过同步复制和单个数据节点的自动故障转移实现高可用性。

注意：仅在单个 GCP 区域内具有低延迟互连的多个地区之间，才建议进行故障转移同步数据库镜像。
区域之间的延迟通常太高，无法提供积极的用户体验，特别是对于具有高更新率的部署更如此。

内部负载均衡器

大多数 IaaS 云提供商缺乏提供虚拟 IP (VIP) 地址的能力，这种地址通常用在自动化数据库故障转移设计中。
为了解决这一问题，InterSystems IRIS 中增强了几种最常用的连接方法，尤其是 ECP 客户端和 Web
网关，从而不再依赖 VIP 功能使它们实现镜像感知和自动化。

xDBC、直接 TCP/IP 套接字等连接方法，或其他的直接连接协议，均需要使用类 VIP 地址。
为了支持这些入站协议，InterSystems 数据库镜像技术使用称作 *mirrorstatus.cxw* 的健康检查状态页面为 GCP
中的这些连接方法提供自动化故障转移，以与负载均衡器进行交互，实现负载均衡器的类 VIP
功能，仅将流量重定向至活动的主镜像成员，从而在 GCP 内实现完整且强大的高可用性设计。

[此处](#)提供了使用负载均衡器实现类 VIP 功能的详细信息。

示例 VPC 拓扑

下图 4.3-a 中的 VPC 布局组合了所有组件，具有以下特点：

利用一个区域内的多个地区实现高可用性
提供两个区域进行灾难恢复
利用多个子网进行网络隔离
包括分别用于互联网和 VPN 连接的单独网关

使用云负载均衡器进行镜像成员的 IP 故障转移

持久性存储概述

如简介中所述，建议使用 GCP 持久性磁盘，尤其 SSD 持久性磁盘类型。之所以推荐 SSD 持久性磁盘，是由于其拥有更高的读写 IOPS 速率以及低的延迟，适合于事务性和分析性数据库工作负载。在某些情况下，可使用本地 SSD，但值得注意的是，本地 SSD 的性能提升会在可用性、耐用性和灵活性方面做出一定的权衡。

可在[此处](#)找到本地 SSD 数据持久性方面的详细信息，您可了解何时保存本地 SSD 数据以及何时不保存它们。

LVM 条带化

与其他的云提供商一样，GCP 在每个虚拟机实例的 IOPS、空间容量和设备数量方面都施加了众多存储限制。有关当前的限制，请查阅 GCP 文档，可在[此处](#)找到。

由于这些限制的存在，使用 LVM 条带化实现数据库实例的单个磁盘设备的 IOPS 最大化变得非常必要。在提供的此示例虚拟机实例中，建议使用以下磁盘布局。与 SSD 持久性磁盘相关的性能限制可在[此处](#)找到。

注意：目前，每个虚拟机实例最多有 16 个持久性磁盘，但 GCP 近期的方案则增至 128 个（测试），这将是令人欣慰的提高。

LVM 条带化的优势在于可以将随机的 IO 工作负载分散到更多的磁盘设备并继承磁盘队列。以下是如何在 Linux 中将 LVM 条带化用于数据库卷组的示例。本示例在一个 LVM PE 条带中使用 4 个磁盘，物理盘区 (PE) 大小为 4MB。或者，如果需要，可以使用更大的 PE 容量。

- 步骤 1：根据需要创建标准性磁盘或 SSD 持久性磁盘
- 步骤 2：使用 “lsblk -do NAME,SCHED” 将每个磁盘设备的 IO 调度器设置为 NOOP
- 步骤 3：使用 “lsblk -do KNAME,TYPE,SIZE,MODEL” 识别磁盘设备
- 步骤 4：使用新的磁盘设备创建磁盘卷组
 - `vgcreate s 4M`
 - **示例：** `vgcreate -s 4M vgirisdb /dev/sd[h-k]`
- 步骤 4：创建逻辑卷
 - `lvcreate n -L -i -l 4MB`
 - **示例：** `lvcreate -n lvirisdb01 -L 1000G -i 4 -l 4M vgirisdb`
- 步骤 5：创建文件系统
 - `mkfs.xfs K`
 - **示例：** `mkfs.xfs -K /dev/vgirisdb/lvirisdb01`
- 步骤 6：装载文件系统

- 使用以下装载条目编辑 /etc/fstab
 - /dev/mapper/vgirisdb-lvirisdb01 /vol-iris/db xfs defaults 0 0
 - 装载 /vol-iris/db

使用上表，每个 InterSystems IRIS 服务器将具有以下配置：2 个 SYS 磁盘、4 个 DB 磁盘、2 个主日志磁盘、2 个备用日志磁盘。

为了增长，LVM 允许在需要的情况下不中断地扩展设备和逻辑卷。有关持续管理和扩展 LVM 卷的最佳做法，请查阅 Linux 文档。

注意 0>：强烈建议同时为数据库和写入映像日志文件启用异步 IO。有关在 Linux 上启用的详细信息，请参阅下列社区文章：<https://community.intersystems.com/post/lvm-pe-striping-maximize-hyper-converged-storage-throughput>

配置

InterSystems IRIS 新增了 InterSystems Cloud Manager (ICM)。ICM 执行众多任务，并提供许多用于配置 InterSystems IRIS 数据平台的选项。ICM 作为 Docker 映像提供，其拥有配置强大的、基于 GCP 云的解决方案所需的一切。

ICM 当前支持以下平台上的配置：

- Google Cloud Platform (GCP)
- Amazon Web Services，包括 GovCloud (AWS / GovCloud)
- Microsoft Azure Resource Manager，包括 Government (ARM / MAG)
- VMware vSphere (ESXi)

ICM 和 Docker 可以从台式机/笔记本电脑工作站运行，也可以具有中央专用的适度“配置”服务器和中央存储库。

ICM 在应用程序生命周期中的作用是“定义->配置->部署->管理”

有关安装和使用 ICM 及 Docker 的详细信息，可在[此处](#)找到。

注意：任何云部署都非必须使用 ICM。完全支持传统的 tar-ball 分布式安装和部署方法。但是，建议使用 ICM，以简化云部署中的配置和管理。

容器监视

ICM 包含基本的监视工具，其使用 Weave Scope 进行基于容器的部署。默认情况下不会部署该工具，需要在默认的文件中使用 *监视器* 字段指定它。

有关使用 ICM 进行监视、编排和调度的详细信息，可在[此处](#)找到。

有关 Weave Scope 的概述和该文档，可在[此处](#)找到。

高可用性

InterSystems 数据库镜像可在任何云环境中提供最高级别的可用性。有一些选项可以直接在实例级别提供虚拟机弹性。有关 GCP 中可用的各种政策的详细信息，可在[此处](#)找到。

上文中已讨论了云负载均衡器如何通过数据库镜像为虚拟 IP（类 VIP）功能提供自动化 IP 地址故障转移。云负载均衡器使用 `mirrorstatus.cwx` 健康检查状态页面，上文 *内部负载均衡器* 部分提到过该页面。数据库镜像有两种模式——自动故障转移同步镜像、异步镜像。在本示例中，将介绍同步故障转移镜像。有关镜像的详细信息，可在[此处](#)找到。

最基本的镜像配置是仲裁器控制配置中的一对故障转移镜像成员。仲裁器放置在同一区域内的第三个地区中，以防止潜在的地区中断影响仲裁器和其中一个镜像成员。

在网络配置中，有多种方法专供设置镜像。在本示例中，我们将使用本文档前述 *网络网关和子网* 定义部分中定义的网络子网。下一部分内容将提供 IP 地址方案示例，并且基于本部分内容，将仅描述网络接口和指定的子网。

灾难恢复

InterSystems 数据库镜像将支持灾难恢复的高可用性功能扩展到另一个 GCP 地理区域，以在整个 GCP 区域万一脱机的情况下支持操作弹性。应用程序如何耐受此类中断取决于恢复时间目标 (RTO) 和恢复点目标 (RPO)。这些将为设计适当的灾难恢复计划进行的分析提供初始框架。以下链接中的指南提供了您在为自己的应用程序制定灾难恢复计划时要考虑的事项。
<https://cloud.google.com/solutions/designing-a-disaster-recovery-plan> 和
<https://cloud.google.com/solutions/disaster-recovery-cookbook>

异步数据库镜像

InterSystems IRIS 数据平台的数据库镜像提供强大的功能，可在 GCP 地区和区域之间异步复制数据，以帮助支持您的灾难恢复计划的 RTO 和 RPO 目标。有关异步镜像成员的详细信息，可在[此处](#)找到。

与上述高可用性部分中讲述的内容相似，云负载均衡器也使用上文 *内部负载均衡器* 部分中提到过的 `mirrorstatus.cwx` 健康检查状态页面为虚拟 IP（类 VIP）功能提供自动化 IP 地址故障转移，以进行 DR 异步镜像。

在本示例中，将介绍 DR 异步故障转移镜像，并介绍 GCP 全局负载均衡服务，以便为上游系统和客户端工作站提供单个任播 IP 地址，不分您的 InterSystems IRIS 部署是否在哪个区域或地区中运行。

GCP 的其中一个发展就是负载均衡器的诞生，这是一种软件定义的全局资源，并且不受限于特定的区域。由于其不是基于实例或设备的解决方案，因此它具有跨区域利用单个服务的独特功能。有关通过单个任播 IP 进行 GCP 全局负载均衡的详细信息，可在[此处](#)找到。

在上述示例中，所有 3 个 InterSystems IRIS 实例的 IP 地址都提供给了 GCP 全局负载均衡器，它会将流量仅定向到承担主要镜像的镜像成员，而不论其位于哪个地区或区域。

分片集群

InterSystems IRIS 拥有一系列全面的功能来缩放您的应用程序，您可以根据自己工作负载的性质以及所面临的特定性能挑战来单独或组合应用这些功能。分片功能是这些功能中的一种，可跨多个服务器对数据及其关联的缓存进行分区，从而为查询和数据引入提供灵活、高性价比的性能扩展，同时通过高效的资源利用最大化基础架构的价值。

InterSystems IRIS 分片群集可以为各种应用提供显著的性能优势，尤其对于工作负载包括以下一项或多项的应用更是如此：

- 大容量或高速数据引入，或两者的组合。
- 相对较大的数据集、返回大量数据的查询，或两者。
- 执行大量数据处理的复杂查询，例如扫描磁盘上大量数据或涉及大量计算工作的查询。

这些因素分别都会影响分片的潜在优势，但组合起来使用它们可能会增加优势。例如，这 3 个因素的组合——快速引入大量数据、大型数据集、检索和处理大量数据的复杂查询——使得当今的许多分析性工作负载非常适合进行分片。

注意，这些特征都与数据有关；InterSystems IRIS 分片的主要功能是**缩放数据量**。不过，当涉及某些或所有这些与数据相关的因素的工作负载也经历大量用户的超高查询量时，分片群集也能提供用户量缩放功能。分片也可以与纵向缩放相结合。

操作概述

分片架构的核心是跨多个系统对数据及其关联的缓存进行分区。分片集群跨多个 InterSystems IRIS

实例以行方式（称为**数据节点**

）对大型数据库表进行物理上的横向分区，同时允许应用通过任何节点透明地访问这些表，但仍将整个数据集看作一个逻辑并集。该架构具有 3 个优点：

- **?????**：
查询在数据节点上并行运行，然后将结果进行合并和组合后，由节点作为完整查询结果返回给连接的应用。许多情况下，这大大提高了执行速度。
- **分区缓存**：
每个数据节点都有自己的缓存，专用于它存储的分片表数据分区，再不是单个实例的缓存服务于整个数据集，这大大降低了缓存溢出的风险，并强制执行性能降低式磁盘读取。
- **并行加载**：
数据可以并行加载到数据节点，从而减少了引入工作负载和查询工作负载之间的缓存和磁盘争用，提高了两者的性能。

有关 InterSystems IRIS 分片集群的详细信息，可在[此处](#)找到。

分片元素和实例类型

分片集群包含至少一个数据节点，如果特定性能或工作负载有需要，则可添加一定数量的计算节点。这两种节点类型提供简单的构建块，从而实现简单、透明和高效的调整模型。

数据节点

数据节点存储数据。

在物质层面，分片表^[1]数据分布在集群中的所有数据节点上，非分片表数据仅物理存储在第一个数据节点上。这种区分对用户是透明的，唯一可能的例外是，第一个节点的存储消耗可能比其他节点略高，但是由于分片表数据通常会超出非分片表数据至少一个数量级，因此这种差异可以忽略不计。

需要时，可以跨集群重新均衡分片表数据，这通常发生在添加新的数据节点后。这将在节点之间移动数据的“存储桶”，以实现数据的近似均匀分布。

在逻辑层面，未分片的表数据和

所有分片的表数据的并集在任何节点上都可见，因此客户端会看到整个数据集，这与其连接哪个节点无关。元数据和代码也会在所有数据节点之间共享。

分片集群的基本架构图仅由在集群中看起来统一的数据节点组成。客户端应用程序可以连接到任何节点，并且可以像在本地一样体验数据。

^[1]为方便起见，术语“分片表数据”

”在整个文档中用于表示支持分片的任何数据模型的“盘区”数据（标记为已分片）。

术语“未分片表数据”和未分片数据

”用于表示处于可分片盘区但却未这样标记的数据，或表示尚不支持分片的数据模型。

数据节点

对于要求低延迟（可能存在不断涌入数据的冲突）的高级方案，可以添加计算节点以提供用于服务查询的透明缓存层。

计算节点缓存数据。每个计算节点都与一个数据节点关联，为其缓存相应的分片表数据，此外，它还根据需要缓存非分片表数据，以满足查询的需要。

由于计算节点物理上并不存储任何数据，其只是支持查询执行，因此可对其硬件配置文件进行调整以满足需求，例如通过强调内存和 CPU 并将存储空间保持在最低限度。

当“裸露”应用程序代码在计算节点上运行时，引入数据会被驱动程序 (xDBC, Spark) 直接或被分片管理器代码间接转发到数据节点。

分片集群说明

分片集群部署有多种组合。下列各图说明了最常见的部署模型。These diagrams do not include the networking gateways and details and provide to focus only on the sharded cluster components.

基本分片集群

下图是在一个区域和一个地区中部署了 4 个数据节点的最简单分片集群。GCP 云负载均衡器用于将客户端连接分发到任何分片集群节点。

在此基本模型中，除了 GCP 为单个虚拟机及其连接的 SSD 持久性存储提供的弹性或高可用性外，没有其他弹性或高可用性。建议使用两个单独的网络接口适配器，一则为入站客户端连接提供网络安全隔离，二则为客户端流量和分片集群通信之间提供带宽隔离。

具有高可用性的基本分片集群

下图是在一个区域中部署了 4 个镜像数据节点的最简单分片集群，每个节点的镜像在地区之间进行了划分。GCP 云负载均衡器用于将客户端连接分发到任何分片集群节点。

InterSystems 数据库镜像的使用带来了高可用性，其会在该区域内的第二地区中维护一个同步复制的镜像。

建议使用 3 个单独的网络接口适配器，一方面为入站客户端连接提供网络安全隔离，另一方面为客户端流量、分片集群通信、节点对之间的同步镜像流量之间提供带宽隔离。

此部署模型也引入了本文前面所述的镜像仲裁器。

具有单独计算节点的分片集群

下图采用单独的计算节点和 4 个数据节点扩展了分片集群，以此来应对大量的用户/查询并发。云负载均衡器服务器池仅包含计算节点的地址。更新和数据引入将像以前一样继续直接更新到数据节点，以维持超低延迟性能，并避免由于实时数据引入而在查询/分析工作负载之间造成资源的干扰和拥挤。

使用此模型，可以根据计算/查询和数据引入的规模单独微调资源分配，从而在“适时”需要的地方提供最佳资源，实现经济而简单的解决方案，而非只是进行计算或数据的调整，浪费不必要的资源。

计算节点非常适合直接使用 GCP

自动缩放分组（亦称自动缩放），允许基于负载的增加或减少自动从托管实例组添加或删除实例。自动缩放的工作原理为：负载增加时，将更多的实例添加到实例组（扩展）；对实例的需求降低时将其删除（缩减）。

有关 GCP 自动缩放的详细信息，可在[此处](#)找到。

自动缩放可帮助基于云的应用程序轻松应对流量增加的情况，并在资源需求降低时降低成本。只需简单地定义自动缩放策略，自动缩放器就会根据测得的负载执行自动缩放。

备份操作

备份操作有多个选项。以下 3 个选项可供您通过 InterSystems IRIS 进行 GCP 部署。

下面的前 2 个选项（下文详细说明）采用快照类型的过程，该过程会在创建快照之前将数据库写入操作挂起到磁盘上，然后在快照成功后恢复更新。

可采取以下高级别步骤通过任一快照方法来创建洁净的备份：

- 通过数据库外部冻结 API 调用暂停对数据库的写入。
- 创建操作系统和数据磁盘的快照。
- 通过外部解冻 API 调用恢复数据库写入。
- 将设施存档备份到备份位置

有关外部冻结/解冻 API 的详细信息，可在[此处](#)找到。

注意：本文档未包含备份示例脚本，但您可定期检查发布到 InterSystems 网站上开发者社区的示例。请访问 www.community.intersystems.com

第三个选项是 InterSystems 在线备份。这是小型部署的入门级方法，具有非常简单的用例和界面。但是，随着数据库的增大，建议将使用快照技术的外部备份作为最佳做法，因为其具有以下优势：备份外部文件、更快的恢复时间，以及企业范围的数据和管理工具。

可以定期添加诸如完整性检查之类的其他步骤，以确保洁净且一致的备份。

决定使用哪种选项取决于您组织的运营要求和策略。InterSystems 可与您详细讨论各种选项。

GCP 持久性磁盘快照备份

可以使用 GCP gcloud 命令行 API 以及 InterSystems 外部冻结/解冻 API 功能实现备份操作。这允许实现真正的 24x7 全天候操作弹性，并确保洁净常规备份。有关管理、创建和自动化 GCP

持久性磁盘快照的详细信息，可在[此处](#)找到。

逻辑卷管理器 (LVM) 快照

或者，可以在 VM 本身中部署单个备份代理，利用文件级备份，并结合逻辑卷管理器 (LVM) 快照，来使用市面上的许多第三方备份工具。

该模型的主要优点之一是能够对基于 Windows 或 Linux 的 VM 进行文件级恢复。此解决方案需要注意的几点是，由于 GCP 和大多数其他 IaaS 云提供商都不提供磁带媒体，因此所有的备份存储库对于短期归档均基于磁盘，并能够利用 Blob 或存储桶类型的低成本存储来进行长期保留 (LTR)。强烈建议您使用此方法来使用支持重复数据删除技术的备份产品，以最有效地利用基于磁盘的备份存储库。

这些具有云支持的备份产品的示例包括但不限于：Commvault、EMC NetWorker、HPE Data Protector 和 Veritas NetBackup。

注意：InterSystems 不会验证或认可一种备份产品是否优于其他产品。选择备份管理软件的责任由客户个人决定。

在线备份

对于小型部署，内置在线备份工具也是可行的选择。该 InterSystems 数据库在线备份实用工具通过捕获数据库中的所有块来备份数据库文件中的数据，然后将输出写入顺序文件。这种专有的备份机制旨在使生产系统的用户不停机。有关在线备份的详细信息，可在[此处](#)找到。

在 GCP 中，在线备份完成后，必须将备份输出文件和系统正在使用的所有其他文件复制到该虚拟机实例之外的其他存储位置。存储桶/对象存储是对其很好的命名。

使用 GCP 存储桶有两种选择。

- 直接使用 gcloud 脚本 API 来复制和操作新创建的在线备份（和其他非数据库）文件。有关详细信息可在[此处](#)找到。
- 尽管云存储桶属于对象存储，但将存储桶装载为文件系统，并将其用作类似持久性磁盘的功能足以。

有关装载云存储桶（使用云存储 FUSE）的详细信息，可在[此处](#)找到。

[#GCP #InterSystems 业务解决方案和架构 #IRIS Analytics Architect #云 #容器化 #平台 #系统管理 #高可用性 #InterSystems IRIS](#)

源

URL:

<https://cn.community.intersystems.com/post/%E7%B2%BE%E5%8D%8E%E6%96%87%E7%AB%A0-%E9%9D%A2%E5%90%91-google-cloud-platform-gcp-%E7%9A%84-intersystems-iris-%E7%A4%BA%E4%BE%8B%E5%8F%82%E8%80%83%E6%9E%B6%E6%9E%84>