

Tableau Server 10.0

高可用性：大规模提供任务关键型分析

Kitty Chou, 产品经理

Mike Klaczynski, 产品营销部

目录

自助式分析事关成败	3
了解高可用性	3
Tableau Server 可扩展性	4
直接可用的高可用性	4
了解 Tableau Server 高可用性	5
常规介绍	5
网关	5
应用程序服务器	6
协调服务	6
群集控制器	7
存储库	7
后台程序	8
数据服务器	8
缓存服务器	8
数据引擎	8
文件存储	9
搜索和浏览	9
VizQL Server	10
与第三方监视工具集成	10
主服务器节点的故障转移	11
配置备用主节点	11
监视群集状态	12
架构注意事项	13
选择最佳配置	13
三节点 HA 部署（最低级别）	14
3+ 节点部署	15
Tableau Server 灾难恢复	16
最简单的 DR 策略	16
第三方 DR 解决方案	16
Tableau DR 功能	16
超越高可用性	17

自助式分析事关成败

如今，自助式分析和数据驱动型决策正在全球范围内成为各组织的一种常态，这一趋势比以往任何时候更为明显。用户和决策者日益依靠可以即时访问的数据和自助式工具来实时回答问题。高管们认识到数据驱动型决策在公司中的重要性，并依赖相应的系统开展日常工作。对数据的这种依赖让基础系统的高可用性变得不可或缺。现有的团队和企业工具必须能够更加方便地访问和配置平台的功能。

Tableau Server 10.0 开创了任务关键型自助式分析的未来。它支持快速的自助式数据探索，通过可靠的管控来提高内容和数据的可信度，而且在任何企业中都易于部署、管理和扩展。本文将探讨 Tableau Server 10.0 如何大规模提供具有高可用性的自助式分析。

了解高可用性

高可用性系统旨在最大程度缩短系统的停机时间。可用性通常表示为“**9 的个数**”，它真正测量的是实际正常运行时间与预期正常运行时间的百分比。下表显示 9 的个数与每年停机时间之间的对应关系。

9 的个数	可用性百分比	年度合计停机时间
1	90%	36.5 天
2	99%	3 天, 15 小时
3	99.9%	8 小时, 45 分
4	99.99%	52 分, 34 秒
5	99.999%	5 分, 15 秒

图 1：典型的可用性指标与对应的每年停机时间。

系统管理员常常与其商业用户达成服务级别协议 (SLA) 以定义可接受的停机时间阈值。他们将根据此 SLA，选择可以实现这些目标的架构来进行部署。大多数系统管理员会为维护、升级和补丁安装安排停机时间。此外，还可能发生意外故障，这种情况被称为计划外停机时间。当然，管理员需要执行计划内维护来更新硬件或软件，目标是最大程度减少计划外停机时间。

我们知道，让用户可以随时查看和了解数据是十分重要的。我们也认识到，总是会有各种事件威胁商业智能系统的可用性，无论这些事件是与硬件、软件、网络还是人为失误相关。作为最基本的保障，Tableau Server 10.0 进程会在发生组件故障时自动重启，使您的系统保持运行。正确配置的多节点部署使用冗余进程来实现服务器高可用性 (HA)。然而，与大多数系统不同，Tableau 让您可以轻松地对自己的分析环境进行高可用性设置和配置。

Tableau Server 可扩展性

Tableau Server 的架构支持纵向扩展和横向扩展。它向大型组织提供企业级部署稳定性，同时又保留了深受小型团队青睐的简单易用的特点。根据您的环境，Tableau Server 可以在一台或多台计算机上运行，并可在同一个节点上运行一个或多个组件进程，从而针对您的用户需求以及高可用性要求提供最佳服务。

在内部，Tableau 管理若干云级别的 Tableau Server 部署，这些部署包括 Tableau Public 和 Tableau Online。Tableau Public 是一个自定义的 Tableau Server 部署，每周可在全球范围内提供数百万个公共视图。作为我们的工程设计和发布过程的一部分，我们将 Beta 版的 Tableau Server 部署到 Tableau Public，以便对稳定性和质量进行精调，然后再向企业客户发布。

直接可用的高可用性

Tableau Server 的安装和高可用性配置工作非常简单。默认安装仅需几分钟。可以将 Tableau Server 安装在主计算
机以及群集中的其他所有工作节点上，然后使用配置实用工具进行配置，组成高度可用的群集。

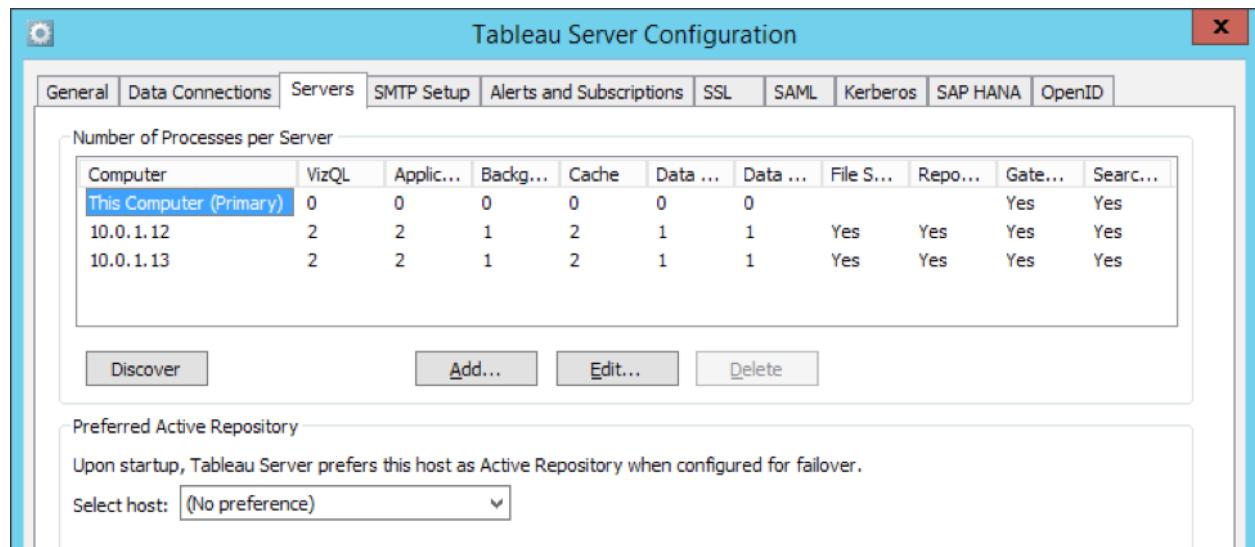


图 2：Tableau Server 配置实用工具可在节点的计数和配置以及群集中每个节点的进程数方面提供灵活性。

了解 Tableau Server 高可用性

Tableau Server 通过一系列专用进程来确保最终用户可以访问整个系统。此部分假设读者熟悉 Tableau Server 组件及其功用。如果您对此并不熟悉，我们建议您首先阅读 [Tableau Server 管理指南](#)。

了解如何使 Tableau Server 高度可用其实也就是了解如何使其每个组件高度可用。换言之，必须确保每个组件的高可用性才能让整个 Tableau Server 群集高度可用，为单点故障提供冗余性。让我们依次考虑每个组件。

常规介绍

Tableau Server 具有用于重新启动其所有服务器进程的内置自动化功能。此自动化功能可以自动重启发生故障的服务器进程并通知管理员，从而确保高可用性。运行 Tableau Server 的硬件或虚拟机必须处于正常状态才能执行此自动化功能。

为了避免整个节点发生故障，务必对唯一的服务器进程进行配置，使其在群集中的不同节点上存在冗余。可以为所有进程配置冗余，但许可服务进程除外，该进程只能在主节点上运行。我们会在主服务器节点的故障转移部分更详细地讨论如何应对此进程限制。

网关

从 Tableau Server 8.1 起，网关进程可以在 Tableau Server 群集的任何（所有）节点上运行，而不只是在 Tableau Server 主节点上运行。这对于高可用性具有重大意义。

在 Tableau Server 8.1 之前，只能配置一个节点来运行网关进程。这个进程发生故障时，Tableau Server 无法与群集之外的任何人通信，还会导致许多内部通信中断。在 Tableau Server 10.0 中，配置多个节点来运行网关进程时，只要至少一个网关进程在运行，与服务器的通信就保持完好。

缓解网关故障的风险

实现网关高可用性的关键是在一个 Tableau Server 群集中拥有多个节点并配置多个节点来运行网关进程。事实上，我们建议您在每个节点上配置一个网关进程。这样就可以缓解网关进程成为单一故障点并导致服务不可用的风险。

如果网关进程发生故障，会出现什么情况？正如前文所述，如果没有任何网关进程在运行，则整个 Tableau Server 群集将不可用。如果其他网关进程保持运行状态，向这些工作网关提出的请求将得到正常处理。然而，虽然存在其他正常运行的网关，但是故障网关接收的任何请求不会被重定向，而是继续失败。发生故障的网关进程将自动重新启动；所以只要计算机本身在正常运行，发生故障的网关进程将重新启动并继续为请求提供服务。

为了让您的安装更能耐受网关故障，我们建议将您的 Tableau Server 群集置于外部负载均衡器 (ELB) 之后，从而确保请求仅路由到正常运行的网关进程。[Tableau Server 管理员指南](#)为上述设置提供了更多指导。

应用程序服务器

应用程序服务器的高可用性很容易实现。您只需在 Tableau Server 群集中的每个节点上配置应用程序服务器的实例。

如果应用程序服务器进程发生故障，会出现什么情况？由该实例处理的请求将会失败，但后续请求将路由到处于运行状态的其他应用程序服务器进程。如果包含发生故障的应用程序服务器的节点仍然在运行，发生故障的进程应该在数秒内自动重新启动。

协调服务

协调服务是 Tableau Server 基础安装的一部分。安装协调服务的节点数量取决于群集中的节点数量，如下表所述：

群集节点数	协调服务节点数
1 - 2	1
3 - 4	3
5+	5

协调服务进程将依次安装在群集的前 n 个节点上（包括主节点），其中 n 是安装协调服务的节点数（基于上表）。

了解法定数量

如果正在运行的协调服务进程数量没有达到法定数量（基于配置的协调服务进程总数），Tableau Server 运行会完全中断。法定数量其实就是绝对多数的另一种说法。所以具有三台或四台计算机的群集最多可以容许一个节点（协调服务的一个实例）停止运行。具有五台或六台计算机的群集最多容许两个协调服务进程停止运行。

值得指出的是，仅有两个节点的群集不容许任一协调服务进程停止运行，这可以解释完全 HA（包括自动故障转移）为何需要至少三个节点。

如果协调服务进程发生故障，会出现什么情况？只要剩余的协调服务进程仍然达到法定数量，就不会有任何问题。如果仍然正常运行的协调服务进程少于法定数量，则整个 Tableau Server 群集变为不可用，这是为了保护底层 Postgres 数据库的引用完整性。

只要计算机本身状态正常，发生故障的协调服务进程将自动重新启动。

群集控制器

群集控制器也是 Tableau Server 10.0 基础安装的一部分。在安装期间，群集的每个节点上都会配置一个群集控制器进程，无需显式配置。

如果群集控制器进程发生故障，会出现什么情况？同一节点上的所有其他 Tableau Server 组件将变为不可用并在 Tableau Server 状态页面上显示为“不可用”。在该节点上运行的任何存储库进程也将不可用。（要全面了解这种影响，请阅读下文的存储库部分。）为了在群集控制器进程发生故障时提供保护，请确保每个唯一的服务器组件都具有冗余性且在群集中的至少两个不同节点上运行。

只要计算机本身状态正常，发生故障的群集控制器进程将自动重新启动。群集控制器重新启动时，它还会重新启动该节点上配置的任何存储库进程。

存储库

存储库是 Tableau Server 执行其功能所需的数据库。如果没有完全正常运行的存储库，整个 Tableau Server 群集将不可用。在任何时间点，只能有一个完全正常运行的存储库，此存储库称为“主动”存储库。涉及存储库的所有操作都在此主动存储库上运行。

为了提高可用性，可以在群集的其他节点上为 Tableau Server 配置一个额外的“被动”存储库。主动存储库的内容不断流入被动存储库。如果主动存储库发生故障，具有高可用性配置的群集会自动将被动存储库提升到主动状态，从而使服务器继续可用。需要高可用性的客户当然希望配置被动存储库。在整个群集中，最多只能有两个存储库（一个为主动，另一个为被动），而且不能置于同一节点上。

群集控制器管理存储库的启动、关闭以及从主动到被动的任何故障转移。这可以解释群集控制器进程故障为何会导致存储库问题。如果启动存储库的群集控制器进程发生故障，则该存储库也会发生故障。

如果存储库进程发生故障，会出现什么情况？这不能一概而论。以下列表概括了各种情况：

- 如果被动存储库发生故障，用户不会受到任何影响。由于主动存储库仍然在正常运行，所有功能都将保持正常。在后台，被动存储库将重新启动，数据复制将会恢复，但是在被动存储库与主动存储库重新完全同步之前，可能会存在一些延迟。

- 如果主动存储库发生故障，且不存在完全同步的被动存储库，则 Tableau Server 将不可用，直到主动存储库重新启动。系统将尝试自动执行此操作，但是对于由某些原因导致的故障，操作可能会失败。

注意：这意味着 - 如果目前仅有主动存储库可用，则不能将 Tableau Server 视为高度可用。如果不存在已同步的被动存储库，则主动存储库就成了整个系统的单一故障点

- 如果主动存储库发生故障，但存在完全同步的被动存储库，且群集采用了高可用性配置，则将自动触发转向被动存储库的故障转移。在故障转移之后，先前的被动存储库将变为新的主动存储库。系统会将发生故障的主动存储库重新启动为新的被动存储库，然后开始进行同步。它还将自动重新启动其他相关进程，以便它们能感知新提升的主动存储库并进行重新连接。在此短暂的重新启动窗口期间，用户会遇到服务中断。然而，管理员无需为了确保持续可用性而进行干预，因为这是一个自动执行的序列。如果希望手动将被动存储库提升为主动存储库，可以使用 tabadmin failoverrepository 命令来执行此操作。

后台程序

为了使后台程序服务高度可用，您应该配置多个后台程序进程，使之在群集中的多个节点上运行。在决定后台程序的数量和运行位置时，请根据其他服务器进程，考虑每个计算机的可用容量。

如果后台程序进程发生故障，会出现什么情况？后台程序正在处理的作业将失败且不会重试。然而，大多数后台作业会被计划为定期运行。在下一个计划运行时间，正常运行的后台程序进程将选取并正常执行相同的后台任务。

只要计算机本身状态正常，发生故障的后台程序进程将自动重新启动，但是它们不会重试已失败的作业。

数据服务器

要使数据服务器高度可用，请配置一个或多个数据服务器进程，使之在群集中的多个节点上运行。

如果数据服务器进程发生故障，会出现什么情况？由数据服务器进程通过代理运行的查询将会失败，导致视图呈现失败。只要有正常工作的数据服务器可以接受重新路由的请求，后续请求（包括重试的失败操作）应该可以成功。

Tableau Server 的运行并不依赖于数据服务器；然而，如果数据服务器未运行，群集就无法通过代理将工作簿连接到外部数据源。没有将数据服务器用于其数据源之一的任何视图应该仍能正常运行。

缓存服务器

缓存服务器提供共享的外部查询缓存。这是键/值对缓存，可以保留先前查询的信息并加快后续请求的处理速度。这意味着，如果一个缓存服务器进程变为不可用（或者即使它们全部不可用），相对而言，后果也不会很严重。Tableau Server 仍将继续运行，但是执行操作可能需要更长时间，因为它们没有预缓存结果可用。重新运行查询时，重新启动的缓存服务器将重新填充，最终加快最终用户的查询速度。缓存服务器对可用性基本没有影响；然而，它对最终用户的各种性能场景会产生影响。就像其他所有进程一样，只要计算机本身运行状态正常，缓存服务器也会自动重新启动。

数据引擎

数据引擎组件负责在使用内存中分析时加载并查询数据提取。要使数据引擎高度可用，只需配置一个或多个数据引擎进程，使之在群集中的多个节点上运行。所有数据引擎都以主动/主动模式运行并执行完全相同的功能。务必注意，配置用于运行数据引擎进程的任何节点也会被配置来运行文件存储进程。文件存储进程管理对数据提取文件的存储和复制，我们接下来将对其进行说明。

如果数据引擎进程发生故障，会出现什么情况？该数据引擎进程当前运行的查询将失败，导致视图呈现失败或数据提取刷新失败。重新运行同一操作的任务会被重新分配到正常运行的其他数据引擎。

发生故障的数据引擎进程将自动重新启动；只要计算机本身状态正常，数据引擎进程将重新启动。

文件存储

如前文所述，文件存储进程将在运行一个或多个数据引擎进程的每个节点上自动实例化。文件存储进程管理节点之间数据提取文件的存储和复制。

文件存储的工作原理

用户首次将数据提取文件发布到 Tableau Server 或者对其进行数据提取刷新时，系统中将创建该数据提取文件。一旦发生上述任何事件，数据提取就会存在于单个节点的单个文件存储中。该数据提取文件尚不能被称为高度可用，因为它没有冗余性，因此是单个故障点。文件存储进程相互通信，将本地数据提取迅速复制到群集中的所有其他文件存储节点。文件存储进程用于在群集网络资源允许的限度内迅速复制文件，但是根据数据提取的大小，需要的时间有所不同。一旦在群集内的多个节点上创建可用副本，数据提取文件将被视为高度可用。

如果文件存储进程发生故障，会出现什么情况？有两个后果：

- 将数据提取文件复制到受影响节点以及从受影响节点复制数据提取文件的操作停止。
- 从受影响节点移除不再需要的数据提取文件的操作暂停。此移除过程通常称为“数据提取收割”。

数据提取收割暂停时，没有直接影响。它只是意味着不需要的数据提取文件将开始累积，从而消耗该节点上的磁盘空间。这最终会造成问题，但是正确设置大小的节点应该具有足够的磁盘空间缓冲容量。

无法复制文件意味着，添加到发生故障的文件存储节点的新数据提取文件在群集中的其他任何文件存储节点上均不可用，反之亦然。在文件存储进程重新启动之后，系统会确保在所有节点上的所有文件存储之间进行同步，从而自我修正。

只要计算机本身状态正常，发生故障的文件存储进程将自动重新启动。文件存储进程应该会迅速恢复其功能，包括对所有文件的同步，不论这些文件是在故障期间还是之后添加的。

搜索和浏览

搜索和浏览进程的高可用性很容易实现。应该配置系统，使搜索和浏览进程在多台计算机上运行。

如果搜索和浏览进程发生故障，会出现什么情况？Tableau Server 的大部分功能将无法使用，虽然用户仍然可以登录系统，但工作簿内容会表现为缺失状态。内容实际上并未缺失，只是没有返回在搜索结果中。这些内容将在搜索和浏览进程重新启动之后重新显示。在发生故障时，即使已经配置了在多个节点上运行的多个搜索和浏览进程，向发生故障的搜索和浏览进程提出的请求也会失败，但后续请求将路由到正常运行的搜索和浏览进程。每个搜索和浏览进程都会跨群集中的所有节点编制索引，因此即使所有搜索和浏览进程中仅有一个进程未发生故障，仍然会跨所有节点返回结果。

VizQL Server

要为 VizQL Server 进程实现高可用性，只需配置一个或多个实例，使之在多台计算机上运行。

如果 VizQL Server 进程发生故障，会出现什么情况？如果仅有一个 VizQL Server 进程且该进程发生故障，则 Tableau Server 不能再呈现任何视图。实现高可用性需要配置冗余的 VizQL 进程。一个相当典型的配置是在每个节点上包含 2 到 4 个 VizQL Server 进程。这可以同时满足高可用性和可扩展性需求。如果多个 VizQL Server 进程正在运行，则单个进程故障将导致请求失败，同时导致进程故障时的会话数据丢失。后续的任何请求都将路由到 Tableau Server 群集中正常运行的其他 VizQL Server 进程。

至此，我们已了解每个服务器进程在故障后的行为方式以及您可以如何缓解故障风险，确保 Tableau Server 群集整体上高度可用。除了针对每种故障场景进行规划，您还应该积极监视群集以查看过去的进程故障。

与第三方监视工具集成

除了使用 Tableau Server 的内置机制来监视系统运行状况外，您还能以机器可读的 XML 版本形式远程接收每个进程的状态。您将需要启用对服务器的远程访问并使用正确的实例 URL: http://<my_tableau_server>/admin/systeminfo.xml

这将返回一个 status.xml 文件，可以解析该文件以确定服务器状态，还可以将其集成到其他系统或监视工具：

通过远程访问服务器状态 URL 提供的 status.xml 文件的机器可读 XML 示例

```
<systeminfo xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <machines>
        <machine name="my_tableau_server">
            <repository worker="my_tableau_server:8060" status="Active"
preferred="false"/>
            <dataengine worker="my_tableau_server:27042" status="Active"/>
            <applicationserver worker="my_tableau_server:8600" status="Active"/>
            <apiserver worker="my_tableau_server:8000" status="Active"/>
            <vizqlserver worker="my_tableau_server:9100" status="Active"/>
            <dataserver worker="my_tableau_server:9700" status="Active"/>
            <backgrounder worker="my_tableau_server:8250" status="Active"/>
            <gateway worker="my_tableau_server:80" status="Active"/>
            <searchandbrowse worker="my_tableau_server:11000" status="Active"/>
            <cacheserver worker="my_tableau_server:6379" status="Active"/>
            <filestore worker="my_tableau_server:9345" status="Active"
pendingTransfers="0" failedTransfers="0" syncTimestamp="2015-02-
27T20:30:48.564Z"/>      <clustercontroller worker="my_tableau_server:12012"
status="Active"/>
            <coordination worker="my_tableau_server:12000" status="Active"/>
        </machine>
    </machines>
    <service status="Active"/>
</systeminfo>
```

主服务器节点的故障转移

安装第一个 Tableau Server 实例的服务器会被认定为 Tableau Server 主节点。主服务器是唯一的，它包含唯一的许可证管理进程以及其他管理功能，还可能成为完善的服务器安装。在更大的部署中，您可以部署仅具有唯一最简管理功能的主服务器，省略在工作节点上冗余的进程。虽然 Tableau Server 可以提供很大的部署灵活性，但务必记住，其进程会每 72 小时轮询一次主服务器以检查许可状态。如果在许可检查期间无法访问主服务器，检查将会失败且您的 Tableau Server 部署将变为“未经许可”并被禁用。

然而，在现实世界中，您的主服务器可能由于 Tableau Server 软件之外的任意数量的原因（例如，硬件、操作系统或网络故障）而发生故障。在此情况下，如果您将群集部署在外部负载均衡器之后并将网关安装在群集中的所有节点上，则群集的其余部分将继续可用于处理服务请求。然而，如果在许可检查窗口到期时主节点仍然被禁用，则整个群集将变为未经许可。为了应对此类故障，请确保为主节点准备随时可用的温备份。

配置备用主节点

如果主服务器发生故障，管理任务将会停止运行，直到导致服务器故障的问题解决为止。因此，应该分配一台专用计算机作为“备用”主节点，以防主服务器节点发生故障。一旦备用主节点已配置并就绪，它就不应开启或连接到群集。这样可以确保许可和管理功能仍然针对主服务器节点运行。遵循 [Tableau Server 管理员指南](#) 中的指示，了解如何设置和布设备用主服务器。

[Tableau Server 管理员指南](#) 还介绍了在主服务器发生故障时如何切换到备用主服务器。这些步骤会提示群集的其余部分注意此改变，并自动将其管理请求和许可请求重新路由到现在的的新主服务器。这些步骤也可以自动执行。

监视群集状态

系统管理员可以在 Tableau Server 状态页面上监视 Tableau Server 10.0 进程的状态。该页面会显示所有节点的服务器进程以及主动和被动存储库的状态，使管理员了解群集的总体运行情况。

Server Status			
Process Status			
Process	Primary 10.0.1.11	Worker 1 10.0.1.12	Worker 2 10.0.1.13
Cluster Controller	✓	✓	✓
Gateway	✓	✓	✓
Application Server		✓✓	✓✓
VizQL Server		✓✓	✓✓
Cache Server		✓✓	✓✓
Search & Browse	✓	✓	✓
Backgrounder		✓	✓
Data Server		✓	✓
Data Engine		✓	✓
File Store		⌚ Synchronizing	⌚ Synchronizing
Repository		✓	✓
Refresh Status	✓ Active	⌚ Busy	✗ Passive
	⚠ Unlicensed	✗ Down	◻ Status unavailable

图 3. 服务器状态管理页面显示每个节点中每个进程的状态。

还可以使用 Tableau Server 配置实用工具对 Tableau Server 进行配置，使其在发生系统故障时向服务器管理员发送电子邮件通知。此实用工具还可以用于针对磁盘空间容量问题提供早期预警。

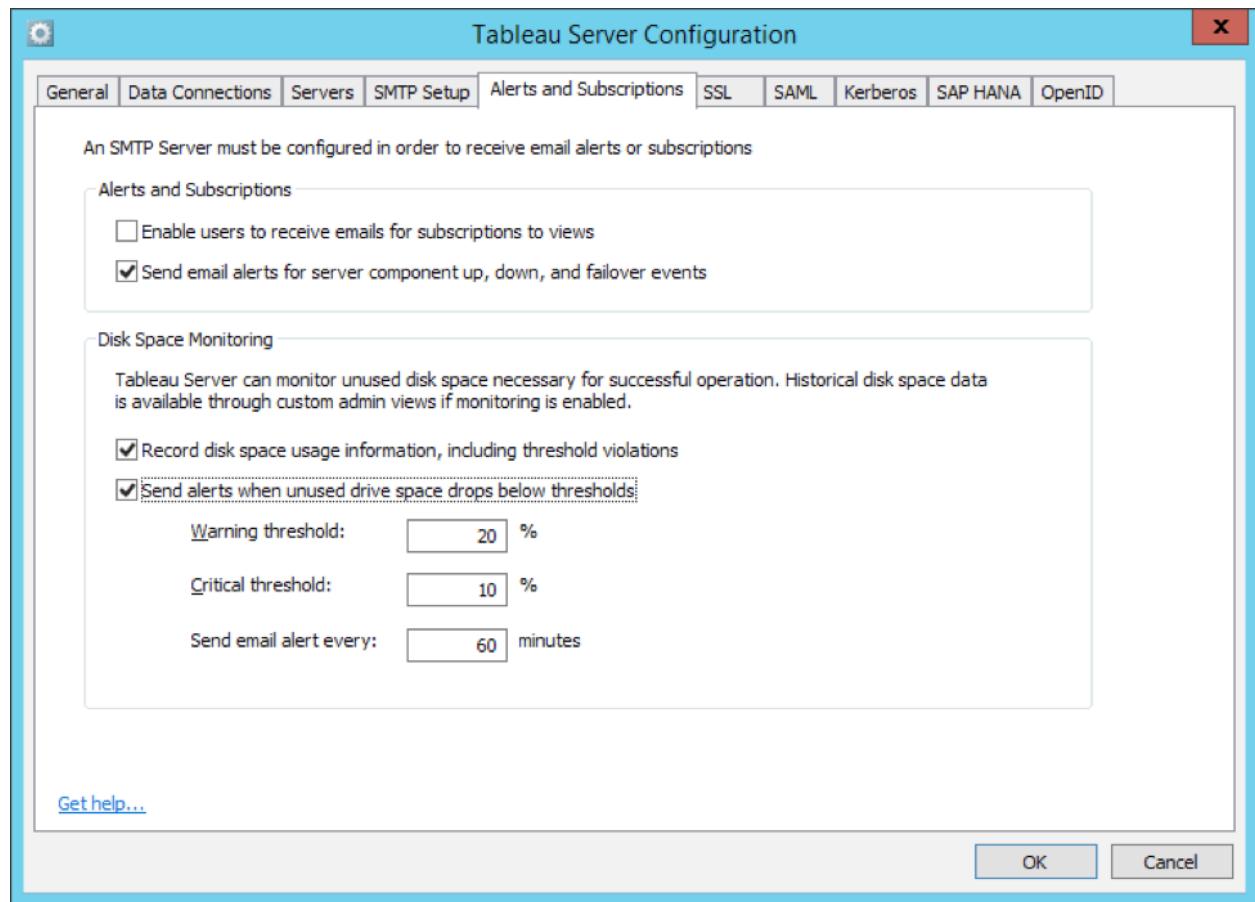


图 4. Tableau Server 配置实用工具可以设置关于系统故障的通知。

架构注意事项

为了成功实现高可用性，应该理解正常运行时间目标以及组织的期望服务级别，然后相应地规划冗余。

选择最佳配置

应根据企业的需求和组织中可用的资源来确定最佳群集配置。虽然有若干可能的成功配置可实现高可用性，但重要的是了解各种规模的群集的利弊，这样才能针对您的环境做出最佳选择。此外，您可以在对称（每个工作节点具有相同的全套组件）和非对称计算机配置（每个节点的组件数量和类型不同）之间进行选择。对称的工作节点配置常常可以简化克隆操作以及将其他节点附加到群集的操作。然而，在规划您的配置时，请牢记单个主动和单个被动存储库的限制。

三节点 HA 部署（最低级别）

如上文所述，群集中必须至少有三个节点，才能完全以 HA 模式运行。一个三节点的群集也是更大部署的良好起点。如果一个节点发生故障，剩余的两个主动节点仍然能够达到法定数量。下方图 5 显示主节点在管理功能和许可功能方面的职责。该图还展示了支持 Tableau Server 的数据和可视化功能的另外两个工作节点。这些组件的确切组合方式可能各不相同，这不仅取决于 HA 需求，还取决于可扩展性需求。例如，虽然图中只为每个工作节点显示了一个 VizQL 进程，但在现实中，可能有两个或更多 VizQL Server 进程在一个工作节点上运行以支持 HA 需求和可扩展性需求。此外，如果主节点有更多核心容量，可以选择包含更多服务器进程，例如 VizQL Server 或数据引擎/文件存储组合，从而提高可扩展性。

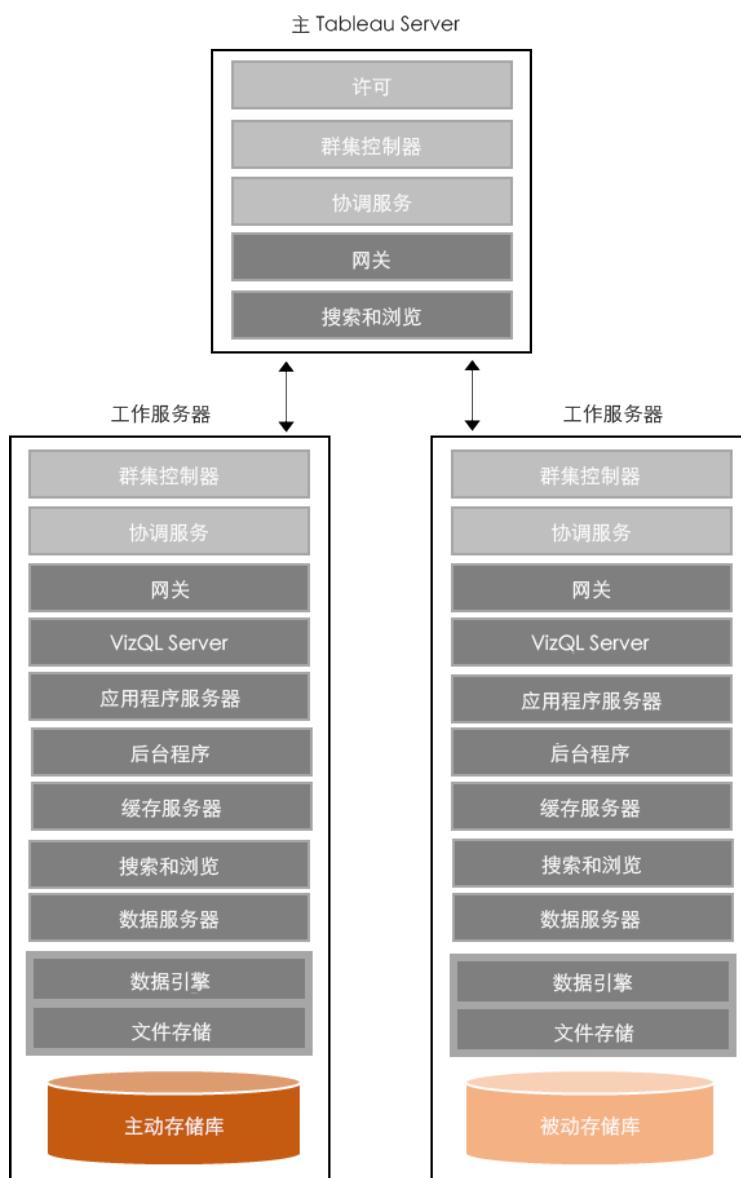


图 5. 三节点 HA 部署（最低级别）示例

通过将数据引擎安装在不同的计算机上，可以针对硬件故障提供额外的冗余并实现快速备份恢复。如果您还需要更多冗余，可以在每个工作节点上部署一个额外的数据引擎。

如果任一工作节点发生故障，主服务器上的进程不会受到影响。然而，工作节点依赖于主节点。在此配置中，如果具有主动存储库的工作节点发生故障，具有被动存储库的工作节点将自动变为主动状态。然而，如上文所述，最好配置和布设一个备用主节点，以便确保从主服务器故障中迅速恢复。此外，如果您的组织具有外部负载均衡器的访问权，您应该将群集部署在该 ELB 之后，从而提高最终用户可用性。如果您刚刚接触高可用性，那么这种配置是很好的入门配置。

3+ 节点部署

在使用三个或更多的节点时，在群集中部署奇数个节点是首选拓扑，可改善高可用性。在满足法定数量方面，具有偶数个节点的群集与节点数量少一个的群集具有相同的能力。例如，如果两个节点停止运行，具有三个节点和四个节点的群集会受到同等影响，即无法达到法定数量并导致关闭。

向三节点群集添加第四个节点（数量变为偶数）将跨更多节点分布工作负载，有助于实现可扩展性，而且可以通过提供额外的冗余进程来减少发生群集全面故障的风险。

撇开其他注意事项不谈，如果您只想最大限度缩短停机时间，应该考虑采用具有五个或更多节点的部署架构。虽然本文建议在开始阶段优先选择三节点架构，但是在大规模的企业级部署中，由于能否满足高可用性要求事关成败，因此应该考虑将五个或更多节点视为基准。

对于高度依赖数据提取和数据提取刷新的组织，将后台程序进程部署到专用工作节点上是常用的部署架构。数据提取刷新工作负载有时会干扰 VizQL Server 进程支持的可视化工作负载，所以您可以将后台程序进程划分到专用工作计算机，以便避免这两个不同工作负载之间争用资源。为了确保您在群集的多个节点中具有进程冗余，仅限后台程序的工作节点常常成对部署。

如前所述，高可用性可减少严重故障造成的停机时间。尽管达到了服务级别要求并部署了支持高可用性的拓扑，您仍然可能遇到需要迅速恢复的故障情况。为了防止灾难和/或硬件故障，请根据灾难恢复目标制定系统恢复计划。

Tableau Server 灾难恢复

针对 Tableau 环境制定灾难恢复 (DR) 计划时，要考虑两个主要因素：(1) 您的企业对恢复过程中停机时间的承受度，以及 (2) 您的企业对数据丢失的承受度。这是按照“恢复时间目标” (RTO) 和“恢复点目标” (RPO) 来定义的，RTO 指标表示系统经过多少时间可以完全恢复，RPO 指标表示企业对数据丢失的承受度。

在确定 RTO 和 RPO 阈值时，必须了解两者的利弊得失（通常是从财务方面），这一点很重要。无论具体情况如何，考虑到使系统迅速恢复联机所需的基础设施和人员，支持极短的恢复时间目标会产生高昂的成本。如果选择减少在基础设施以及员工的响应速度方面的投入，您需要能够耐受长得多的 RTO（停机时间更长）。这种权衡并非特定于 Tableau Server，而是适用于任何服务器软件的部署。

对数据丢失的承受度将取决于各个应用程序以及这些应用程序中的数据变化速率，还有其他因素。例如，24 小时的 RPO 意味着，最多只会丢失过去 24 小时的数据。

最简单的 DR 策略

在许多情况下，最简单且最有效的 DR 策略是在夜间运行备份，并在恢复数据中心内布设好随时可用的备用服务器计算机（温或热）。与传统的商业智能工具相比，Tableau Server 可迅速安装、配置和运行，而且在大多数情况下，您可以通过在备用计算机上安装干净的服务器并从备份进行恢复，实现企业目标。安装服务器以及恢复备份所需的时间取决于备份的大小以及服务器的复杂性，但不管怎样，都可以在几小时内轻松完成。

第三方 DR 解决方案

有很多可提供 15 分钟或更长 RPO 的第三方解决方案，您的 IT 团队可能已在使用这些解决方案。虽然第三方备份恢复和 DR 解决方案不在本白皮书的范畴内，但是在与此类恢复系统一起使用时，应该像对待系统上的任何其他任务关键型应用程序一样对待 Tableau Server，并且应相应地配置第三方备份恢复技术。

Tableau DR 功能

如果您无法使用第三方 DR 技术，Tableau Server 包括一项内置的本机备份和恢复功能。Tableau Server 可以在运行时进行备份，因此您可以创建完全热备份。此外，这些热备份与 Tableau Server 一样智能，它们只备份数据提取文件、存储库数据库以及相关配置。您还可以在 Tableau Server 未运行时（例如，计划停机期间）运行备份。

如果 RPO 为 24 小时，那么可以选择每 24 小时备份一次服务器，也可以按照 RPO 更改备份计划。您的恢复时间将取决于您遭遇的是软件还是硬件故障，以及您在现场配置新硬件的能力。如果有另一个节点备用，且其上已安装并配置了相同的 Tableau Server 实例，那么要使 Tableau Server 恢复生产，只需使用保存的备份文件执行一个简单的命令行管理操作即可。

对于大多数企业，灾难恢复策略还规定了作为服务器主要运行地的数据中心与作为恢复目标的数据中心之间的距离要求（例如在美国，DR 站点必须至少相距 250 英里）。在这些情况下，Tableau 可以视同为了灾难恢复而管理的任何其他业务关键型应用程序。利用 Tableau 的内置备份和恢复技术来满足您的 DR 需求。

超越高可用性

要实现高可用性，仅利用 Tableau Server 中用于提供大规模自助式分析的内置功能是不够的。故障可能在 Tableau Server 软件范围之外发生，例如硬件或网络故障。您可以在对这些类型的故障更有复原能力的基础设施（例如，公共云或私有云中的虚拟化硬件）中部署 Tableau Server，从而改善高可用性。有些供应商甚至提供实时迁移功能，例如 VMware vMotion。

真正的高可用性源自于对用户需求的理解，以及遵循最佳做法和过程来满足这些需求。虽然 Tableau 凭借其强大功能使高可用性更易于实现，但是这些不应替代最佳做法，例如定期的维护和计划的备份。Tableau 认识到商业智能应用程序在帮助您的企业进行决策方面的重要性。Tableau Server 已准备好为您的组织解答最关键的问题。

关于 Tableau

Tableau 帮助人们将数据转化为可以付诸行动、发挥重大作用的见解。轻松连接到以任何形式存储在任意地点的数据。快速执行临时分析，发现隐藏的见解。通过拖放操作，创建包含高级可视化分析的交互式仪表板。然后在整个组织共享，让其他团队成员能够从自己的数据视角进行探索。从全球性企业到早期初创企业和小企业，使用 Tableau 的分析平台来查看和理解数据的人无处不在。

资源

[Tableau for the Enterprise: An IT overview](#) (适用于企业的 Tableau: IT 概述)

[Tableau Server Scalability: A Technical Deployment Guide for Server Administrators](#) (Tableau Server 可扩展性: Server 管理员技术部署指南)

[Server 管理指南](#)

[Tableau on Amazon Web Services](#) (Amazon Web Services 上的 Tableau)

