

ZENTRIX GATEWAY

企业 AI 产品文档

配置 · 使用 · 结果核对

Zentrix Gateway POC 部署手册

从下载安装包到交付可用环境：规划、取包、环境准备、安装、初始化、验收与运维的完整操作说明。

操作与使用 | 权限与边界 | 结果核对

文档版本

2026.10

核对日期

2026 年 10 月 22 日

目录

关于本手册

1. 部署前必须先定的四件事

2. 从下载站取得安装物料

3. 环境准备

4. 交互式安装

5. 非交互安装

6. 首次登录与初始化

7. 安装验收

8. 安装后必做事项与日常运维

9. 增加远端网关节点

10. 备份、恢复、升级与回滚

11. 清空重装与更换产品

12. 故障排查与提交支持

> 版本说明：本手册对应 2026.10 版本，安装器行为、提示文字与诊断代码按 2026-10-22 主线代码核对，下载站截图取自 2026-10-22。实际命令输出以当次交付的安装器为准。

> 安装器行为、提示文字与诊断代码按 2026.10 主线代码核对。实际命令输出与界面以当次交付的安装器和所在环境为准。

关于本手册

适用范围

本手册适用于 Zentrix Gateway 2026.10 版本的 POC 部署：从下载站取得安装器与离线包，到在客户主机上完成安装、初始化、验收，并把可用环境交给客户。产品功能的日常使用见《Zentrix Gateway 用户手册》。

读者	主要阅读章节
实施工程师（负责安装与交付）	全部章节
客户方系统管理员（主机、网络、签字确认）	第 1、3、6、7、8、10 章
客户方数据库与存储管理员（使用现有服务时）	第 3.5 节、第 5 章
客户方安全负责人	第 3.3、3.4、3.7 节，第 8.1 节

阅读约定

- 界面上的菜单、按钮、字段名称用「」标出，菜单路径用「›」分隔。
- 命令块第一行注释标明执行位置、身份与影响，例如 `# 安装主机 · root · 只读`；标为「写入」「删除」的命令执行前先读懂注释。
- `<离线包>.zpkg`、`<访问地址>`、`<YOUR_API_KEY>` 等为占位符，请替换为实际值。
- 「注意」用于提示生效时机、影响范围与容易误解之处；每章末尾的「能力说明」列出产品不提供的能力与可用的替代做法。

部署流程一览

1. 与客户确认产品、形态、数据依赖与访问地址（第 1 章）。
2. 在下载站定制并下载安装器与离线包，核对摘要（第 2 章）。
3. 准备主机、时间同步、权限、网络与现有依赖（第 3 章）。
4. 预检通过后执行安装，等待装后验收通过（第 4、5 章）。
5. 首次登录、修改初始密码、创建租户、导入许可证（第 6 章）。
6. 完成一次真实业务调用并填写验收记录（第 7 章）。
7. 离机备份密钥材料，完成首次平台备份，清理临时文件后交接（第 8 章）。

1. 部署前必须先定的四件事

本章内容：安装器一旦开始写入，产品、部署形态和大部分数据依赖选择就不能再改。本章把这几个决定和它们的影响一次说清，读者照本章填完 1.6 节的确认表之后，后面的章节才有唯一答案。读者是负责本次 POC 的实施工程师，以及需要签字确认的客户方系统管理员。

1.1 选哪个产品

定制站第一步就要选产品，安装完成后不能更换。

产品	包含什么	不包含什么
完整版	企业 AI 平台：员工对话、知识库、文档解析、智能体、模型与工具网关	—
Zentrix 网关	模型与工具网关及治理面：供应商接入、模型服务、调用方与密钥、路由策略、护栏、用量与审计	托管智能体、文档解析

两个产品用同一个安装器、同一套命令，区别只在离线包里带了哪些能力。换成另一个产品只能清空重装（第 11 章），原有数据全部删除。

怎么定：本次只验证网关能力（模型接入、密钥与调用、路由与护栏、用量与审计）就选 Zentrix 网关；要一并验证知识库、文档解析、智能体就选完整版。拿不准时按完整版选，未启用的能力保留在离线包内，安装后可在控制台启用，不用重新下载。

注意：离线包文件名里带产品字段（`zenova-zentrix_gateway-…` 或 `zenova-zenova-…`），安装器文件名不带。同一版本两个产品的安装器文件名相同，下载后分目录存放，否则上传时极易拿错。

1.2 选哪种形态

当前版本提供两种形态，适用场景完全不同。

形态	定制站上的选项	适用	本手册范围
单机原生	Linux · x86 / Linux · ARM「服务器直接安装（systemd）」	一台主机跑控制面、中心网关节点和数据依赖；POC 与单机生产	第 2～8 章
Kubernetes	「Kubernetes Helm 部署」	客户已有集群	不在本手册范围，按专项实施方案执行

单机加远端网关节点（业务调用需要在别的网络区域就近接入）在第 9 章。

注意：本手册的验收、备份、回滚、故障排查命令全部针对原生安装。Helm 形态的离线包内容、安装方式和日常操作都不同，不要混用。

1.3 主机与容量

安装器在写入任何内容之前先做主机检查，不满足要求会停在检查这一步。

项目	要求	不满足的后果
系统	Linux（systemd），x86_64 或 aarch64	阻断

项目	要求	不满足的后果
内存	不低于 8 GiB	阻断
磁盘	按能力累加，见下表	默认阻断（显式指定安装目录时只告警）
时间	与时间源偏差不超过 30 秒	阻断
权限	安装与运维命令需要 root	阻断
CPU 指令集	本机部署分析数据库时需支持 AVX、AVX2、BMI1、BMI2、FMA、F16C、MOVBE、LZCNT	阻断，可改用现有分析数据库
容器运行时	仅在选择托管智能体、文档解析，或多机方案由安装器部署高可用数据层时需要	阻断

磁盘按所选组件累加，检查对象是安装目录（默认 `/opt/zenova`）所在文件系统的可用空间。

组成	增加的空间
基础安装	20 GiB
本机部署平台数据库（PostgreSQL）	30 GiB
本机部署分析数据库（ClickHouse）	50 GiB
自带对象存储	10 GiB
托管智能体	20 GiB
知识库	5 GiB
文档解析	5 GiB
语义缓存	5 GiB

三种典型组合：

组合	合计
Zentrix 网关，核心 + 语义缓存，三项依赖都由安装器部署，自带对象存储	20 + 30 + 50 + 10 + 5 = 115 GiB
完整版，除高可用数据层外全部能力，三项依赖都由安装器部署，自带对象存储	20 + 30 + 50 + 10 + 20 + 5 + 5 + 5 = 145 GiB
核心能力，三项依赖全部使用现有服务	20 GiB

注意：上表是安装检查的门槛，不含业务数据增长、日志与本机备份。POC 期间调用量小，按门槛准备即可；转正式运行前按并发量与数据保留期重新评估。

磁盘不足时的两种行为：不传 `--install-dir` 时，空间不足直接阻断，并输出诊断代码 `HOST_DISK_LOW`，Linux 上同时列出本机可用空间最大的三个挂载点供改用。显式传了 `--install-dir`（包括显式写成默认值 `/opt/zenova`）时，空间不足只告警并继续。因此不要为了绕过检查而随手加 `--install-dir`——加了之后这项检查就只剩提示。

1.4 访问地址与端口

访问地址是浏览器和调用方使用的地址，本机监听端口是安装主机实际接收请求的端口。两者在有代理、NAT 或 TLS 入口时可以不同。

场景	访问地址	本机监听端口	安装时怎么填
内网直连，用 80	`http://10.20.30.40`	80	只填访问地址
内网直连，用非 80 端口	`http://10.20.30.40:8080`	8080	只填访问地址
前置代理或 NAT	`http://gw.example.com:18080`	8080	问「前面是否有代理或 NAT 转发」时选「有」，填本机监听端口；或加 `--business-entry-port 8080`
前置 TLS 入口	`https://ai.example.com`	80	直接询问 TLS 入口转发到的本机端口，回车为 80

注意：填写 `https://` 不会让本机提供 HTTPS。证书与加密终结必须在客户的负载均衡或反向代理上完成，转发到本机的仍是明文 HTTP。另外，访问地址为 `http://` 时登录页不提供通行密钥，两步验证只能用验证器应用。

管理入口（控制台）默认与业务入口共用一个端口。需要把控制台单独放一个端口时加 `--admin-entry-port`。业务入口承载模型、MCP、A2A 调用与健康检查，端口被占用时安装停止且不自动更换；管理入口端口被占用时安装器顺延到下一个空闲端口并提示。

开火墙：入站只需放行业务入口端口（与启用时的 443）。安装器在 Linux 上只自动放行业务入口与 443，不放行分离后的管理入口；ufw 与 nftables 不处理，需要运维方手工配置。

1.5 数据依赖

三项数据依赖各自选择「由我们部署」或「使用现有服务」。

依赖	由安装器部署	使用现有服务时的要求
平台数据库（PostgreSQL）	本机初始化，只监听回环	PostgreSQL 16，超级用户账号；扩展 `vector` 0.7 及以上、`pg_trgm`、`uuid-ossp`、`pgcrypto`、`btree_gist` 可用；`max_connections` 不低于 400

依赖	由安装器部署	使用现有服务时的要求
分析数据库（ClickHouse，必装）	本机初始化，只监听回环	26.7 及以上；HTTP 或 HTTPS 接口端口（常见 8123 / 8443，不是原生 9000）；`default` 账号带密码；`cluster`、`shard`、`replica` 三个宏已配置；`default` 有建库权限
对象存储	自带对象存储（推荐生产），或本机目录（单机试用）	S3 兼容；预先建好存储桶；访问密钥对该桶有读、写、删权限；地址、区域、桶、寻址方式四项齐全

选择时按以下约束判断：

约束	影响
备份范围	`zenova backup` 只备份由安装器部署的平台数据库、分析数据库与本机对象数据。使用现有 PostgreSQL 或现有 S3 时这两部分不纳入，由对应管理员负责
清空重装	只适用于三项依赖都由安装器部署（对象存储为自带或本机目录）的环境
对象存储方式	安装后不能在线更换。更换需要备份、按新选择重装、再恢复数据
现有平台数据库	库中已有本平台历史数据时，普通安装会拒绝继续，改用接管方式（不在本手册范围）

POC 建议：三项全部由安装器部署。这样备份、清空重装、回滚都能在本机闭环，不受客户数据库管理员的窗口约束。

云数据库的坑：云上「高权限账号」通常不是超级用户（`rolsuper` 为假），安装器检测会拒绝。选现有 PostgreSQL 前先在客户侧确认 `SELECT rolsuper FROM pg_roles WHERE rolname = current_user;` 返回 `t`。

1.6 部署参数确认表

安装前填完下表，由客户确认。口令和密钥不填在表上，只登记保管位置与保管人。

项目	现场值
产品（完整版 / Zentrix 网关）、版本、通道	
能力清单	
主机名、IP、系统版本、架构	
CPU 核数、内存、安装目录所在磁盘可用空间	
访问地址、本机监听端口	

项目	现场值
管理访问地址、管理入口端口（分离时）	
证书与 TLS 终结位置、负责人	
平台数据库：部署方式、地址、管理员账号	
分析数据库：部署方式、地址、库类型	
对象存储：方式、地址、区域、桶、寻址方式	
运维通知：渠道与地址、接收人	
时间源地址	
模型供应商接入地址与凭据保管人	
备份转存位置、密钥材料保管人	
实施负责人、客户负责人、维护窗口	

注意：这张表是后续所有步骤的唯一依据。第 6 章初始化完成后，把实际值与这张表核对一遍，不一致的地方记进交付记录——POC 转正式运行时，这张表就是交接材料的第一页。

能力说明

不提供的能力	替代做法
容量规划工具，安装前对内核参数与操作系统发行版做识别	安装器只检查 CPU 架构、内存、磁盘、服务托管、时间、端口与安全模块。容量按 1.3 节的门槛加业务估算准备，发行版满足主机检查即可安装
安装前自动选择磁盘与自动扩容	空间不足时输出 `HOST_DISK_LOW` 并列出候选挂载点，由现场决定改用哪个目录；也可以减少由本机部署的组件
在 macOS 与 Windows 上安装服务端	服务端只提供 Linux（systemd）平台。定制站的平台列表以实际提供的发布为准
安装后更换产品	只能清空重装，原数据全部删除

2. 从下载站取得安装物料

本章内容：安装只需要两个文件——安装器与离线包。本章按下载站的五步操作说明如何选出与本次部署完全一致的这两个文件，以及下载后怎样在不装任何东西的前提下核对它们没拿错、没下坏。读者是实施工程师。

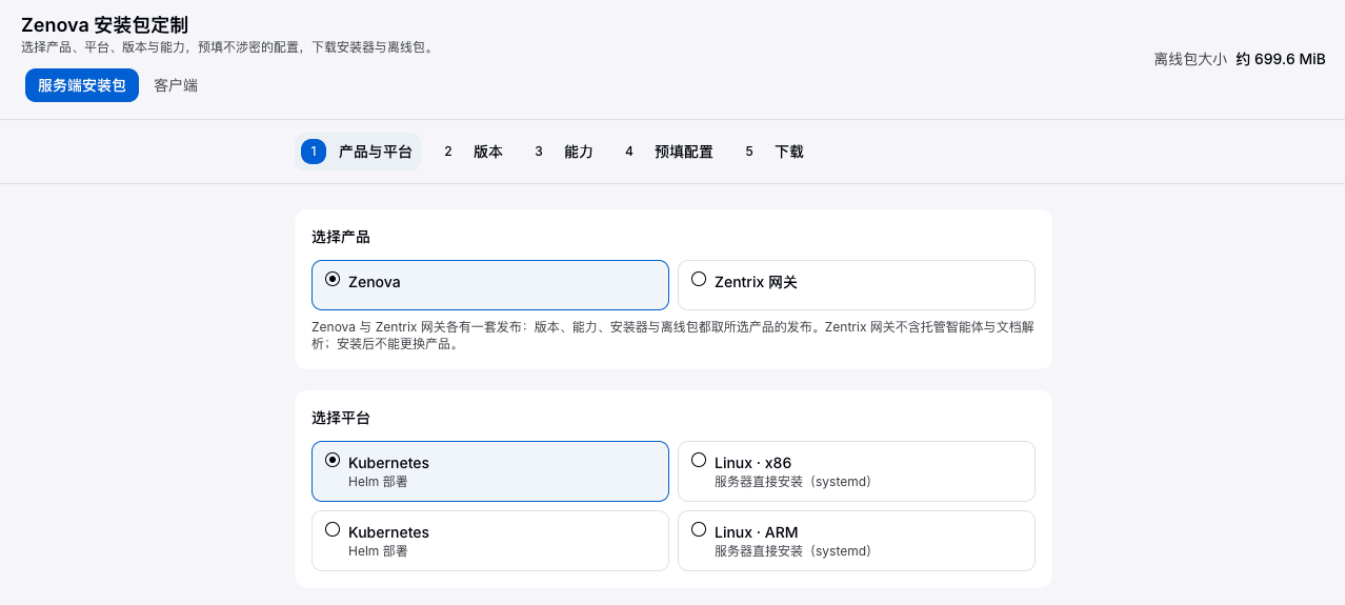
2.1 下载站上有什么

下载站（地址由交付方提供，页面标题为「安装包定制」）是一个离线物料定制页，不是软件商店。它做三件事：按你选的产品、平台、版本和能力，现场组装出对应的离线包；把不涉密的部署参数预填进离线包；给出可以直接复制执行的现场安装命令。

页面顶部有两个页签：

页签	内容
服务端安装包	本次要下载的安装器与离线包，五步选择流程
客户端	与本版本配套的 ZenSwitch 与婉儿客户端安装包

右上角的「离线包大小」随选择实时变化，这是判断有没有多勾能力的直接依据。



注意：定制站不收集任何凭据和许可证。数据库口令、对象存储访问密钥、许可证一律不在这里填，只在现场安装时录入。页面上的预填配置全部可以留空。

2.2 第一步：产品与平台

操作步骤

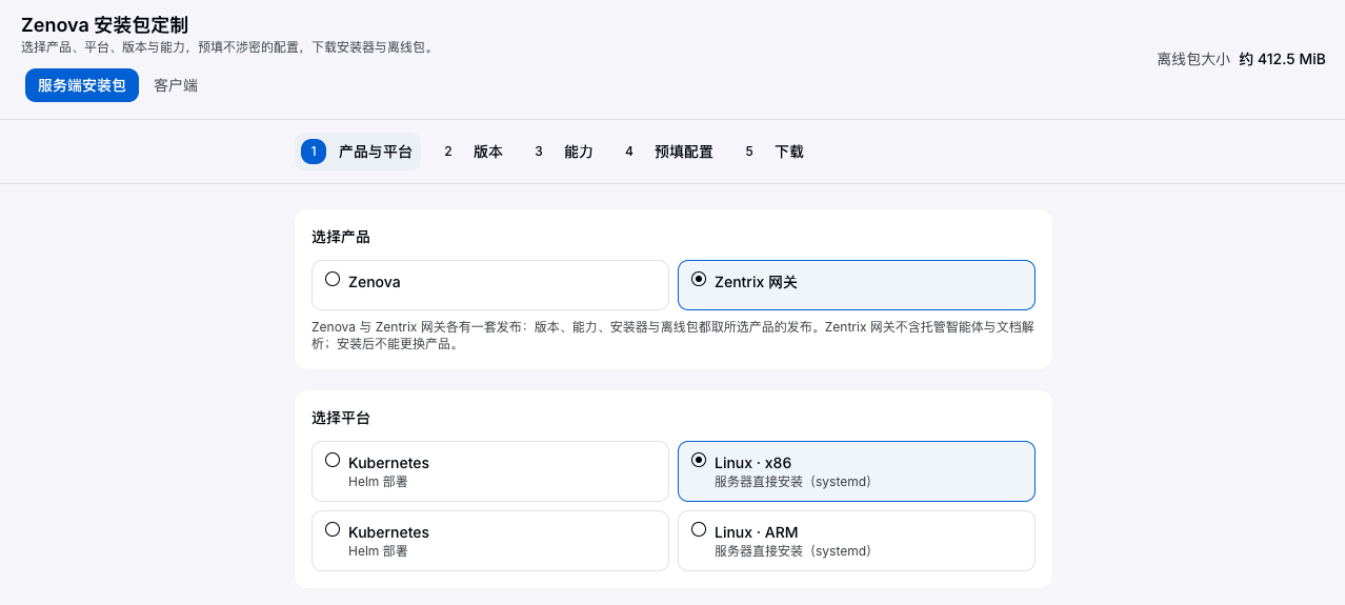
1. 在「选择产品」里点选 Zentrix 网关（另一个选项为完整版）。页面默认选中第一个产品，进页面后先核对这一项。
2. 在「选择平台」里点选平台。每一项的标题是「系统 · 架构」，副标题是部署形态。

界面上看到的	含义
Linux · x86 服务器直接安装（systemd）	单机原生安装，本手册使用这一项
Linux · ARM 服务器直接安装（systemd）	ARM 服务器的单机原生安装

界面上看到的	含义
Kubernetes Helm 部署	部署到客户已有的 K8s 集群

3. 点「下一步」。

你会看到：右侧「离线包大小」随之变化；已完成的步骤在页签上显示对勾，当前步骤高亮。



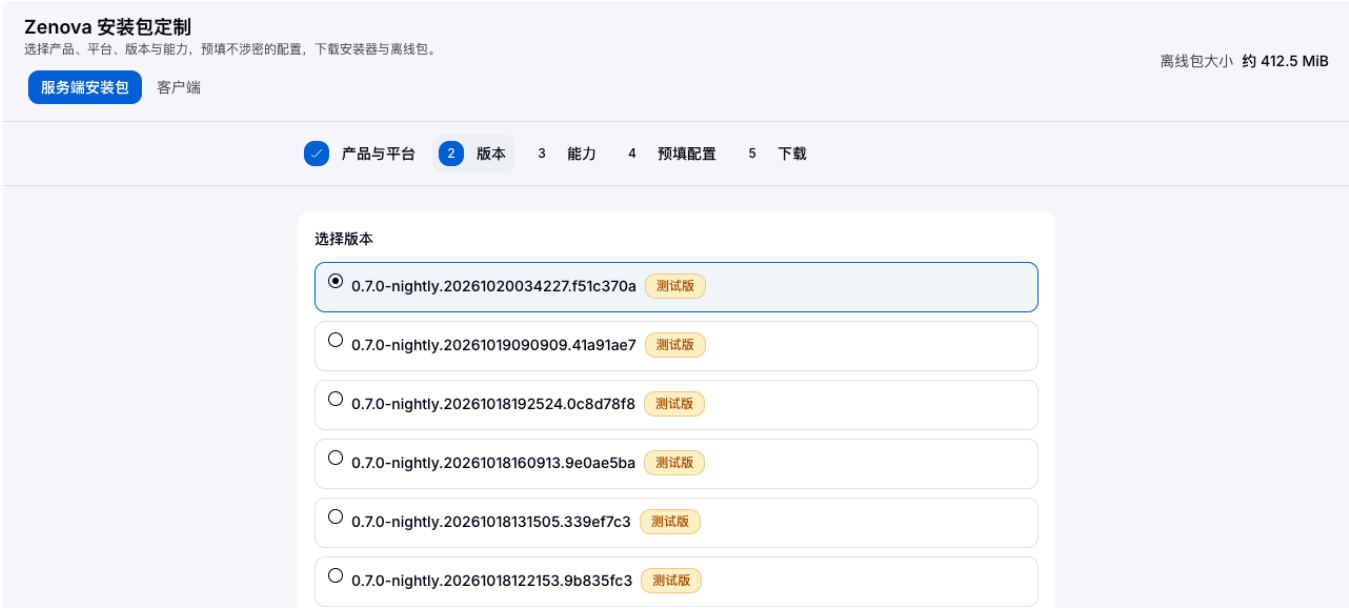
注意：产品选错会导致整个离线包不可用，安装时会被拒绝，不是换个选项就能补救。安装完成后更换产品只能清空重装，原有数据全部删除。

注意：界面上的平台列表由发布流水线实际产出的包决定。某个平台没出现在列表里，说明该版本没有为它构建，不是页面出错。

2.3 第二步：选择版本

操作步骤

1. 点第 2 步「版本」。
2. 在版本列表里选择。每项后面标注通道：`稳定版` 或 `测试版`。页面默认选中最新稳定版；该版本没有稳定版时，选中列表最上面的一项。
3. 记录选中的完整版本号，后面核对物料要用。



注意：列表里只有测试版时（版本号形如 `0.7.0-nightly.20261020034227.f51c370a`），说明该产品当前没有发布稳定版。POC 阶段可以用测试版，但要在交付记录里写明版本号与通道，后续升级与问题定位都以它为准。版本号的三段含义：`0.7.0` 是产品版本，`20261020034227` 是构建时间，末尾 8 位是构建提交。

2.4 第三步：选择能力

操作步骤

1. 点第 3 步「能力」。
2. 勾选本次要启用的能力。「核心」固定勾选，不能取消。每项后面显示该组件占用的体积。
3. 需要全部能力时点右上角「全家桶」，一次性选中所有可勾选项。
4. 点「下一步」。

能力	说明
核心	必装。控制面、入口代理、中心网关节点、数据依赖
知识库	知识库的向量检索服务
托管智能体	需要主机已有容器运行时
文档解析	解析 PDF 与 Office 文档；勾选时会一并勾选知识库
语义缓存	模型调用的语义缓存
数据高可用	多机方案使用；需要容器运行时



能力之间的依赖会自动联动：勾选「文档解析」时，「知识库」一并选中；取消时一并取消。能力的勾选只决定离线包带哪些组件，安装时还能再选一次本次实际启用哪些——未启用的保留在本机，装完在控制台「系统更新」页可直接启用，不需要重新下载。

注意：某些能力会置灰并标注原因，例如「该平台暂未提供此能力」「需要本机有容器运行时」。置灰的项勾不上，这是发布阶段就确定的事实，不要试图绕过。

注意：勾选越多，离线包越大，磁盘门槛也越高。Zentrix 网关在 x86 上选「核心 + 语义缓存」，离线包约 425 MiB；全选会显著更大。POC 阶段建议只勾本次真要验证的能力。

2.5 第四步：预填配置

操作步骤

这一步可以整页跳过。填了的话，安装时会逐项展示并要求确认。

可预填项	说明
访问域名	浏览器与调用方使用的地址
管理访问地址	留空时由访问域名推导
平台数据库（PostgreSQL）	现场再定 / 由我们部署 / 使用现有服务
分析数据库（ClickHouse）	同上
对象存储	现场再定 / 自带对象存储 / 本机目录 / 现有 S3 兼容存储
链路导出地址	可选
运维通知渠道	钉钉 / 飞书 / 企业微信 / 回调地址 / 现场再定
主机清单	仅多机部署需要；单机留空



你会看到：选择「使用现有服务」时，才会展开地址、端口、库名、登录账号等字段；选择「现有 S3 兼容存储」时展开地址、区域、存储桶。数据库口令与访问密钥在任何选项下都不会出现输入框。

注意：页面明确标注「数据库口令、访问密钥等凭据和许可证一律不在本站填写，只在现场录入」。预填的地址类配置在离线包里标记为「来自定制站，未签名」，安装时必须逐项输入 `y` 确认，拒绝的当场改填。

POC 建议：这一页整页留空。POC 的主机与数据库往往到现场才确定，预填反而多一层确认；留空后安装器按第 4 章逐屏询问，答案更贴近实际。

2.6 第五步：下载

操作步骤

1. 点第 5 步「下载」。
2. 点「下载安装器（Linux · x86）」，保存安装器归档。
3. 点「下载离线包」，保存 `.zpkg`。下载中断时重新点击即可从断点继续。
4. 核对页面显示的信息，抄进交付记录。
5. 复制「现场安装命令」备用。也可以复制「配置链接」保存本次全部选择，日后在另一台机器上打开可复现同样的包。



页面上能读到的信息，逐项的含义：

页面项	用途
离线包	.zpkg 的完整文件名，含产品、版本、平台与 8 位摘要
离线包大小	与下载后的实际大小比对，判断是否下完整
安装器	安装器归档文件名与大小
安装器摘要（SHA-256）	下载后本地校验用
现场安装命令	解压安装器 + 执行安装的一条命令，安装器与离线包放在同一目录后执行
配置链接	保存本次全部选择，链接中不含凭据

文件名里各段含义：

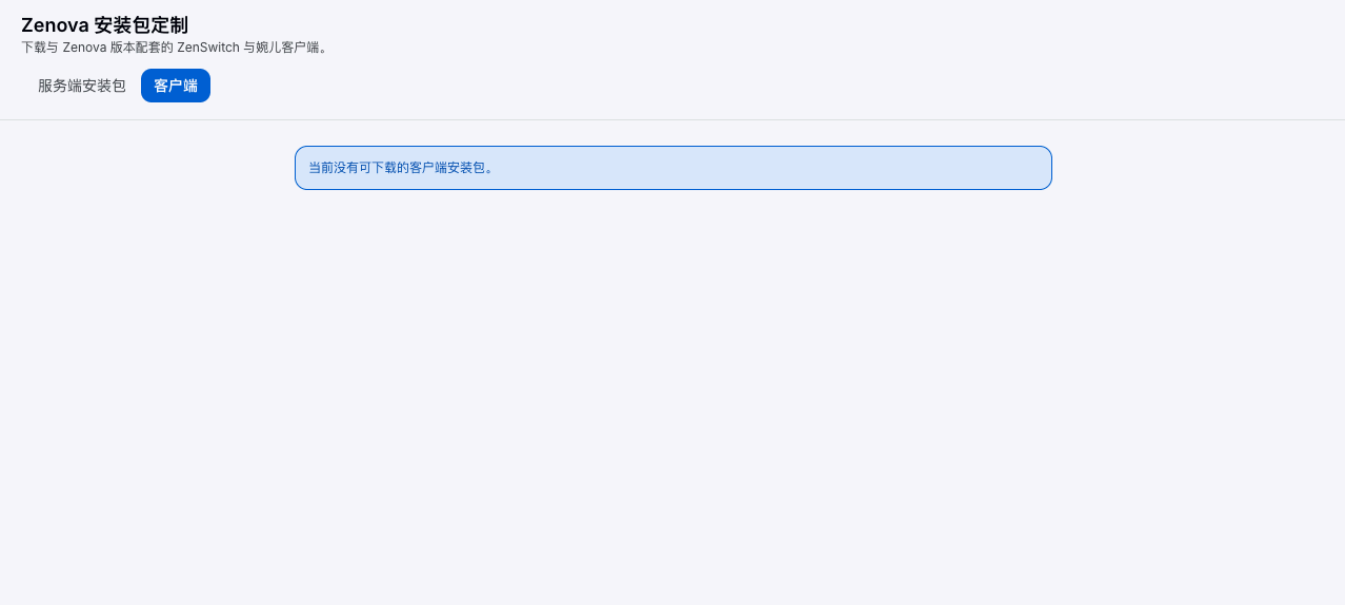
字段	取值示例	含义
产品	zentrix_gateway / zenova	只出现在离线包文件名里
版本	0.7.0-nightly.20261020034227.f51c370a	与第 2 步选的版本一致
系统-架构	linux-amd64 / linux-arm64	与第 1 步选的平台一致
8 位摘要	bee743d0	由版本、平台、能力组合与预填答案算出

注意：安装器文件名里没有产品字段，同一版本两个产品下载到的安装器文件名完全相同。两个产品都要准备时，下载后分目录存放。

注意：8 位摘要是「选择内容」的摘要，不是校验和。同样的能力组合改一次预填配置，摘要就会变，文件名也随之改变。所以不能只凭摘要判断两次下载是不是同一个版本。

2.7 客户端页签

顶部「客户端」页签列出与本版本配套的 ZenSwitch 与婉儿客户端安装包，显示产品、版本、文件、摘要。客户端与平台同一次发布，配套使用。



你会看到：当前没有可下载的客户端安装包时，页面显示「当前没有可下载的客户端安装包」。这表示该版本没有随发布产出客户端包，不是页面故障。

注意：POC 阶段如果只需要验证网关的模型调用与治理能力，可以不装客户端——用 curl 即可完成第 7 章的全部验收。客户端用于验证成员侧的使用入口，不影响服务端功能。

2.8 下载后核对

两个文件上传到安装主机之前，先在能上网的机器上核对。这一步只读，不改任何东西。

核对安装器摘要

```
# 联网机器 · 普通用户 · 只读
sha256sum zenova-<版本>-linux-amd64.tar.gz
```

你会看到：输出的摘要与下载站「安装器摘要（SHA-256）」一栏完全一致。macOS 上用 shasum -a 256。

核对离线包内容

离线包是标准 tar 归档，可以不解包直接看清单：

```
# 联网机器 · 普通用户 · 只读
tar -tf <离线包>.zpkg | head -20
tar -xOf <离线包>.zpkg release.json | grep -oE '"(edition|version|channel)": ?"[^"]*"'
```

你会看到：第一条列出 release.json 及各组件条目；第二条输出三行，产品（zenova 或 zentrix_gateway）、版本号、通道（stable 或 test）。

核对一致性：以下三处必须一致，任何一处不符就重新下载。

1. 安装器与离线包文件名里的版本号；
2. release.json 里的 version；
3. 下载站第 2 步选中的版本。

离线包没有单独的校验文件。它的完整性由安装器在启动时按 release.json 里记录的摘要逐个核对组件来保证，因此只需核对文件大小与下载站显示的一致即可。

能力说明

不提供的能力	替代做法
在下载站填写凭据、导入许可证	凭据与许可证只在现场录入。许可证在安装完成后于控制台导出授权请求、换取签名文件后导入（第 6.4 节）
下载单独的答案文件	预填答案写在离线包内的 answers.toml 中，不单独产出。需要答案文件时按第 5.2 节手工编写
对离线包做签名校验	摘要核对只用于发现下载损坏，不是信任机制。物料来源的可信性由交付渠道保证
在定制站上直接看到已装环境的版本并一键取升级包	从控制台「系统更新」页「复制获取链接」，在联网机器上打开，定制站已预选本机组合（第 10.5 节）

3. 环境准备

本章内容：安装器在确认安装之前会做一轮主机检查，但有几类问题只有到写盘之后才会暴露，到那时再找客户的主机、安全、数据库和网络管理员，往往要等很久。本章把这些准备工作提前到安装日之前：先用只读命令自查主机，再分别与主机安全管理员、数据库管理员、存储管理员、网络管理员确认各自负责的事项，最后把物料传上去并核对。读者是实施工程师和客户方的各类管理员。

3.1 主机检查项与安装前自查

适用场景：拿到安装主机的登录权限后、正式安装前一天或更早。

安装器的主机检查项如下。「何时检查」一栏说明这一项在哪个环节暴露，用来判断第 4 章的预检（--check-only）能不能替你提前发现它。

检查项	要求	不满足时	何时检查
系统与服务托管	Linux，以 systemd 托管服务	阻断，提示「本机没有 systemd，无法托管服务」	主机检查（预检包含）
CPU 架构	x86_64 或 aarch64，与离线包平台一致	阻断（HOST_ARCH_MISMATCH），需换用本机平台的离线包	主机检查（预检包含）

检查项	要求	不满足时	何时检查
内存	不少于 8 GiB	阻断，提示「内存不足：需要 8 GB，本机 N GB」	主机检查（预检包含）
磁盘空间	按本次部署的组件累加，见下表	未传 `--install-dir` 时阻断（`HOST_DISK_LOW`）；显式传了 `--install-dir`（含显式写 `/opt/zenova`）时只告警	主机检查（预检包含）
系统时间	已与时间源同步，偏差不超过 30 秒	阻断（`HOST_CLOCK_UNSYNCED` / `HOST_CLOCK_SKEW`），见 3.2 节	主机检查（预检包含）
CPU 指令集	本机部署 ClickHouse 时需要 AVX、AVX2、BMI1、BMI2、FMA、F16C、MOVBE、LZCNT	阻断（`HOST_CPU_FEATURES`），可改为使用现有 ClickHouse	主机检查（预检包含）
容器运行时	仅托管智能体、文档解析、多机部署高可用数据层时需要	阻断（`HOST_CONTAINER_RUNTIME`），见 3.4 节	主机检查（预检包含）
本机网络地址	本机至少有一块已启用、配置了有效 IP 的非回环网卡	阻断（`HOST_RUNTIME_HOST`）	主机检查（预检包含）
安装目录上级可遍历	服务用户 `zenova` 能进入 `/opt` 等上级目录	阻断（`HOST_ROOT_TRAVERSE`），见 3.3 节	主机检查（预检包含）
安装目录允许执行	安装目录所在挂载点没有 `noexec`	阻断（`HOST_NOEXEC_MOUNT`），见 3.3 节	主机检查（预检包含）
服务用户属性	已有同名用户 `zenova` 时，主组必须是同名组 `zenova`	阻断（`HOST_SERVICE_USER`），见 3.3 节	主机检查（预检包含）
权限	以 root 执行	阻断，提示加 `sudo`	启动时
业务入口端口	访问地址对应的本机监听端口空闲，且不与平台内部端口相同	阻断（`HOST_ENTRY_PORT_BUSY` / `HOST_ENTRY_PORT_RESERVED`），不会自动更换	主机检查之后、录入凭据之前；预检不包含
程序执行权限	SELinux、fapolicyd、麒麟 KYSEC 等不拦截服务用户执行 `/opt/zenova` 下的程序	阻断（`HOST_SELINUX_DENIED` / `HOST_FAPOLICYD` / `HOST_KYSEC_DENIED` / `HOST_EXEC_DENIED`）	确认安装、解包之后；预检不包含

磁盘需求按本次部署的组件累加（GiB）：

组件	需要空间
基础	20
本机部署 PostgreSQL	+30
本机部署 ClickHouse	+50
自带对象存储	+10
托管智能体	+20
知识库	+5
文档解析	+5
语义缓存	+5

典型组合：

组合	合计
Zentrix 网关（核心 + 语义缓存，三项依赖都由安装器部署，自带对象存储）	115 GiB
完整版全部能力（不含数据高可用）	145 GiB
只装核心，三项依赖都使用现有服务	20 GiB

操作步骤

1. 以 root 登录安装主机，逐条执行下面的只读命令，把输出记入交付记录。

```
# 安装主机 · root · 只读（系统、架构、内存、磁盘）
uname -m
free -g
df -h /opt
```

```
# 安装主机 · root · 只读（安装目录的挂载选项与上级目录权限）
findmnt -T /opt -o TARGET,OPTIONS
ls -ld / /opt
```

```
# 安装主机 · root · 只读（是否已有同名服务用户）
id zenova
```

```
# 安装主机 · root · 只读（CPU 指令集，仅 x86_64 且本机部署 ClickHouse 时需要）
grep -o -w -E 'avx|avx2|bmi1|bmi2|fma|f16c|movbe|abm' /proc/cpuinfo | sort -u
```

```
# 安装主机 · root · 只读（是否以 systemd 启动、是否有非回环地址）
ps -p 1 -o comm=
ip -br addr
```

2. 对照下面的预期逐项判断；不符合的项，按对应小节处理后再执行一次。

你会看到

- `uname -m` 输出 `x86_64` 或 `aarch64`，且与下载站选的平台一致（`linux-amd64` 对应 `x86_64`，`linux-arm64` 对应 `aarch64`）。
- `free -g` 的 `Mem:` 一行 `total` 不小于 8。注意 `free -g` 向下取整，物理内存恰为 8 GiB 的主机可能显示 7，以 `free -m` 的数值（不小于 8192）为准。
- `df -h /opt` 的 `Avail` 不小于上表算出的合计。不够时，选一个空间更大、可执行的持久数据盘，安装时用 `--install-dir <目录>` 指定（第 4 章）。
- `findmnt` 的 `OPTIONS` 一列里没有 `noexec`。
- `ls -ld / /opt` 两行权限的最后一位都是 `x`（例如 `drwxr-xr-x`）。
- `id zenova` 输出 `no such user` 表示不存在，安装器会自己创建，这是最常见的情况；如果已存在，输出里 `gid=` 后面括号中的组名必须是 `zenova`。
- 指令集命令输出 8 行：`abm`、`avx`、`avx2`、`bmi1`、`bmi2`、`f16c`、`fma`、`movbe`（`abm` 即 LZCNT）。少任何一项，本机都不能部署 ClickHouse。
- `ps -p 1 -o comm=` 输出 `systemd`。
- `ip -br addr` 至少有一行不是 `lo` 且带 IP 地址。

注意：以下两类问题，预检（`--check-only`）发现不了，必须在安装日之前单独确认：

- 执行权限：SELinux、fapolicyd、麒麟 KYSEC 对程序的拦截，只有在确认安装、离线包解开之后，安装器以服务用户身份实际执行一次随包程序时才能发现。启用了这类强制访问控制的主机，提前按 3.3 节与主机安全管理员约定放行方式。
- 业务入口端口：访问地址依赖的本机监听端口是否被占用，在主机检查之后才检查，预检不包含。提前用 3.7 节的 `ss` 命令确认端口空闲。

注意：虚拟机上指令集不全，多半是虚拟化平台给虚拟机配的 CPU 型号偏低。处理办法是把 CPU 型号改为 host 直通，或 Haswell 及以上型号；做不到时，把分析数据库改为「使用现有的」ClickHouse（3.5 节）。

3.2 时间同步

适用场景：所有安装。时间不同步是安装日最常见的阻断项之一。

前提：安装主机能访问客户内网的 NTP 服务器，或客户指定的时间源。

安装器按以下任一情况判定时间检查失败：

情况	屏幕上的代码
与时间源的偏差超过 30 秒	`HOST_CLOCK_SKEW`
chrony 只把本地时钟作为时间源	`HOST_CLOCK_UNSYNCED`

情况	屏幕上的代码
时间源层级（Stratum）不在 1～15	`HOST_CLOCK_UNSYNCED`
`chronyc` 与 `timedatectl` 都给不出结果	`HOST_CLOCK_UNSYNCED`

操作步骤

1. 先看当前状态。

```
# 安装主机 · root · 只读
chronyc tracking
chronyc sources -v
timedatectl
```

2. 如果没有配置时间源，或只配了本地时钟，编辑 chrony 配置文件（RHEL、麒麟、openEuler 一般为 `/etc/chrony.conf`，Ubuntu、Debian 为 `/etc/chrony/chrony.conf`），写入客户提供的 NTP 服务器，删除或注释掉 `local stratum` 一类以本地时钟为源的配置行。

```
server <NTP服务器地址> iburst
```

3. 重启时间服务并等待同步。

```
# 安装主机 · root · 写入（重启时间服务）
systemctl restart chronyd    # Ubuntu 上服务名为 chrony
chronyc waitsync 30
chronyc tracking
```

4. 偏差已经很大时（例如几分钟），可以让 chrony 立即跳变校准，再执行一次 `chronyc tracking`。

```
# 安装主机 · root · 写入（立即校准系统时间）
chronyc makestep
```

5. 主机没有 chrony、使用 systemd-timesyncd 时：在 `/etc/systemd/timesyncd.conf` 的 `[Time]` 段写入 `NTP=<NTP服务器地址>`，然后执行：

```
# 安装主机 · root · 写入（启用 systemd-timesyncd 同步）
timedatectl set-ntp true
systemctl restart systemd-timesyncd
timedatectl show -p NTPSynchronized
```

你会看到

- `chronyc tracking` 中 `Leap status` 为 `Normal`，`System time` 的偏差远小于 30 秒，`Stratum` 在 1～15 之间。
- `chronyc sources -v` 中至少有一个源行首为 `^*`（当前选用的源）。
- 使用 systemd-timesyncd 时，最后一条命令输出 `NTPSynchronized=yes`。
- 安装时主机检查的「系统时间」一项显示 ✓，下面一行写明「已同步，时间偏差 N 秒（安装前上限 30 秒）」。

注意：在安装器的主机检查页处理时间问题时，不需要退出安装器。另开一个终端执行上面的命令，回到安装器按回车重新检查；安装器重新检查前会先等待最多 30 秒，让刚重启的时间服务完成同步。

注意：时间同步不只是安装前的门槛。登录的两步验证码、证书校验、调用记录的时间顺序都依赖准确的时间，安装后也要保持时间服务开机自启。

3.3 目录权限与安全模块

适用场景：/opt 权限被收紧、/opt 挂在带 noexec 的分区、主机上已有 zenova 用户，或主机启用了 SELinux（Enforcing）、fapolicyd、麒麟 KYSEC。

平台服务以系统用户 zenova 身份运行（安装器创建，不可登录）。它需要能进入安装目录的每一级上级目录，并能执行安装目录下的程序。

情况	屏幕上的代码	处理
/opt 或更上级目录权限为 0700 等，服务用户进不去	HOST_ROOT_TRAVERSE	setfacl -m u:zenova:--x /opt，或把 /opt 设为 0755
/opt 所在挂载点带 noexec	HOST_NOEXEC_MOUNT	调整挂载选项，或为 /opt/zenova 单独挂一块允许执行的数据盘
已有同名用户 zenova，但主组不是 zenova	HOST_SERVICE_USER	usermod -g zenova zenova 修正主组，或确认该用户无其他用途后删除
SELinux 拦截执行	HOST_SELINUX_DENIED	由主机安全管理员放行
fapolicyd 拦截执行	HOST_FAPOLICYD	由主机安全管理员放行
麒麟 KYSEC 拦截执行	HOST_KYSEC_DENIED	由主机安全管理员放行
执行被拒绝，原因无法自动判定	HOST_EXEC_DENIED / HOST_EXEC_DENIED_UNKNOWN	按屏幕给出的命令处理；仍不行时查看 dmesg 与审计日志，执行 sudo zenova doctor 后联系技术支持

操作步骤

1. 上级目录不可遍历时，二选一执行：

```
# 安装主机 · root · 写入（只给服务用户增加进入权限）
setfacl -m u:zenova:--x /opt
```

```
# 安装主机 · root · 写入（把 /opt 设为常规权限）
chmod 0755 /opt
```

setfacl 要求服务用户已经存在。用户还没创建时（首次安装前的常见情况），用 chmod，或先与主机管理员约定在安装器报出 HOST_ROOT_TRAVERSE 后再执行 setfacl。

2. 挂载点带 noexec 时，由主机管理员去掉该挂载选项并同步修改 /etc/fstab；或另挂一块允许执行的数据盘，安装时用 --install-dir 指向它。

```
# 安装主机 · root · 写入（重新挂载为允许执行，需同步修改 /etc/fstab）
mount -o remount,exec <挂载点>
```

3. 已有同名用户但主组不对时：

```
# 安装主机 · root · 写入（把 zenova 用户的主组改为同名组）
usermod -g zenova zenova
```

4. 启用了强制访问控制的主机，安装前与主机安全管理员确认放行方式。安装器报出拦截时，屏幕会直接给出针对本机路径的处理命令，形式如下（`<安装根>` 默认为 `/opt/zenova`），由安全管理员审核后执行：

机制	屏幕给出的处理命令
SELinux	`sudo chcon -R -t bin_t <安装根>/releases/`
fapolicyd	`sudo fapolicyd-cli --file add <安装根>/releases/ --trust-file zenova && sudo fapolicyd-cli --update`
麒麟 KYSEC	`sudo kysec_set -rn exectl -v verified <安装根>/releases/`（或 `sudo setstatus softmode -p`）

5. 放行后，重新执行同一条安装命令。

你会看到

- 目录与挂载问题在主机检查页以 `✗` 标出，处理后按回车重新检查即可变为 `✓`。
- 执行权限问题出现在确认安装之后，屏幕逐条列出 `✗ [<代码>] <原因> <处理办法>`，安装停止。此时安装器已写入部分内容，按第 4.8 节续装，不要手工清理。

注意：安全模块的拦截在解包之后才能发现，预检通过不代表不会被拦截。启用了强制访问控制的主机，最稳妥的做法是安装前就把放行方式与主机安全管理员谈定，并约好安装当天的联系人。

注意：`chcon` 设置的标签在系统执行全盘重新标记后可能丢失；需要长期保留时，由主机安全管理员按客户自己的 SELinux 策略管理规范处理。升级会在 `releases/` 下新增版本目录，放行策略要覆盖整个 `releases/` 目录而不是某一个版本目录。

3.4 容器运行时

适用场景：本次要启用「托管智能体」「文档解析」，或多机方案中由安装器部署高可用 PostgreSQL。Zentrix 网关不包含以容器运行的能力，单机安装 Zentrix 网关不需要容器运行时，可以跳过本节。

前提：主机已安装 Docker 并处于运行状态。

操作步骤

1. 确认 Docker 可用。

```
# 安装主机 · root · 只读
docker info --format '{{.ServerVersion}}'
```

2. Docker 未启动时：

```
# 安装主机 · root · 写入（启动 Docker 并设为开机自启）
systemctl enable --now docker
```

3. 取得客户安全负责人的书面同意，见下方「注意」。
4. 正式安装时，安装器会把服务用户 `zenova` 加入 `docker` 组，并以该用户身份实际连接一次 Docker，无需手工操作。

你会看到

- 第 1 步输出一个版本号。输出报错表示 Docker 未安装或未运行。
- 安装器主机检查页的「容器运行时」一项显示 `✓`；不可用时为 `✗` 并提示「所选能力需要容器运行时，但本机 Docker 不可用」。

注意：加入 `docker` 组等同于授予宿主机 root 权限。 启用这类能力前，必须取得客户安全负责人的书面同意，并把同意记录归入交付材料。

注意：安装时没选这类能力、装完后在控制台「系统更新」页再启用时，控制台会重新检查容器运行时。检查失败时提示「所选能力需要容器运行时，但平台服务用户无法访问本机的 Docker」，按本节准备好 Docker 后重试。

3.5 准备现有数据库与对象存储

适用场景：平台数据库（PostgreSQL）、分析数据库（ClickHouse）或对象存储中有任何一项要「使用现有的」。三项都由安装器部署时跳过本节。

三项依赖在安装时逐项选择：

依赖	由安装器部署	使用现有服务
平台数据库（PostgreSQL）	本机初始化，只监听回环地址	见下文「PostgreSQL」
分析数据库（ClickHouse，必装）	本机初始化，只监听回环地址	见下文「ClickHouse」
对象存储	「自带对象存储（推荐，生产）」或「使用本机目录（单机试用）」	见下文「S3 兼容对象存储」

注意：选择「使用现有的」会改变后续运维边界，事先与客户讲清楚：

- `zenova backup` 不备份现有 PostgreSQL 与现有 S3，这两部分由客户数据库管理员、存储管理员按其备份机制保护。ClickHouse 无论自建还是现有，都经 HTTP 接口纳入平台备份。
- 三项依赖都由安装器部署时才能使用清空重装（`--fresh`）；只要有一项使用现有服务，就不能清空重装。
- 对象存储方式安装后不能在线更换。

PostgreSQL

要求

项目	要求
版本	PostgreSQL 16
账号	超级用户（`rolsuper` 为真）。安装器用它创建平台所需的角色、库与扩展，数据库管理员不需要另外执行脚本；该账号的密码只在安装过程中使用，不写入任何文件与日志
扩展	服务器上可用：`vector`（0.7 及以上）、`pg_trgm`、`uuid-ossp`、`pgcrypto`、`btree_gist`
连接数	`max_connections` 不低于 400
库名	不需要提供。平台的库由安装器创建，库名固定
网络	安装主机能访问数据库的地址与端口

操作步骤（由客户数据库管理员执行）

1. 用准备交给安装器的那个管理员账号登录。下面的命令可以在任何装有 `psql` 且能访问数据库的主机上执行。

```
# 数据库管理员主机 · 数据库管理员 · 只读
psql "host=<数据库地址> port=<端口> user=<管理员账号> dbname=postgres"
```

2. 在 `psql` 中逐条执行：

```
-- 只读：版本应为 16
SHOW server_version;

-- 只读：当前账号是否超级用户，应返回 t
SELECT rolsuper FROM pg_roles WHERE rolname = current_user;

-- 只读：所需扩展是否可用，应返回 5 行，vector 的版本不低于 0.7
SELECT name, default_version
  FROM pg_available_extensions
 WHERE name IN ('vector', 'pg_trgm', 'uuid-ossp', 'pgcrypto', 'btree_gist')
 ORDER BY name;

-- 只读：连接数上限，应不低于 400
SHOW max_connections;
```

3. 把以下信息交给实施工程师：访问地址（主机:端口）、管理员账号、加密方式（见下文「加密方式与 CA」）、CA 证书文件（如需要）。密码在安装现场由数据库管理员本人录入，或经凭据文件提供（第 5 章），不经即时通讯或邮件传递。

你会看到

- 四条查询分别返回 `16.x`、`t`、5 行扩展、不小于 400 的数字。

注意：云数据库提供的「高权限账号」通常不是超级用户，`rolsuper` 返回 `f`，安装器的依赖检测会拒绝。此时需要客户换用自建 PostgreSQL，或请云厂商开通超级用户；做不到时改为由安装器部署。

注意：扩展少一项时，由数据库管理员在数据库服务器上安装对应的扩展包（vector 即 pgvector）。pg_available_extensions 里能查到即可，不需要预先 CREATE EXTENSION，安装器会自己创建。

ClickHouse

要求

项目	要求
版本	26.7 及以上
接口	HTTP 或 HTTPS 接口端口（常见 8123 / 8443），不是原生协议端口 9000
账号	default 账号，并设置了密码
宏	已配置 cluster、shard、replica 三个宏
建库权限	default 账号可以建库
库类型	单机库（普通 MergeTree）或复制库（库引擎 Replicated）。安装时要选择，安装器会核对选择与库的实际引擎是否一致

操作步骤（由客户数据库管理员执行）

1. 在能访问 ClickHouse HTTP 接口的主机上执行。用 HTTPS 接口时把 http:// 换成 https://、端口换成对应端口。

```
# 数据库管理员主机 · 数据库管理员 · 只读（版本与宏）
CH='http://<ClickHouse地址>:8123/'
curl -sS -u 'default:<密码>' "$CH" --data-binary 'SELECT version()'
curl -sS -u 'default:<密码>' "$CH" --data-binary 'SELECT macro, substitution FROM system.macros'
```

```
# 数据库管理员主机 · 数据库管理员 · 写入（建一个临时库再删除，验证建库权限）
curl -sS -u 'default:<密码>' "$CH" --data-binary 'CREATE DATABASE zenova_selfcheck'
curl -sS -u 'default:<密码>' "$CH" --data-binary 'DROP DATABASE zenova_selfcheck'
```

2. 告诉实施工程师本次使用单机库还是复制库，以及访问地址、端口、加密方式与 CA 证书（如需要）。

你会看到

- 第一条返回不低于 26.7 的版本号。
- 第二条至少返回 cluster、shard、replica 三行。
- 建库、删库两条无输出、无报错。

注意：用 9000 端口填写访问地址，依赖检测无法通过。ClickHouse 是所有部署形态的必装组件，调用明细、用量统计与策略命中都存放在这里；不能跳过，只能在「由我们部署」与「使用现有的」之间选。

S3 兼容对象存储

要求

项目	要求
协议	S3 兼容
存储桶	预先建好
访问密钥	对该存储桶有读、写、删权限
需要提供的四项	访问地址（`http://…` 或 `https://…`）、区域、存储桶、寻址方式（路径寻址或虚拟主机寻址）

操作步骤（由客户存储管理员执行）

1. 建好存储桶，生成访问密钥并授予读、写、删权限。
2. 用存储管理员熟悉的 S3 客户端验证三种权限。以 AWS CLI 为例：

```
# 存储管理员主机 · 存储管理员 · 写入（上传、读取、删除一个临时对象）
export AWS_ACCESS_KEY_ID='<访问密钥 ID>'
export AWS_SECRET_ACCESS_KEY='<访问密钥>'
echo selfcheck > /tmp/zenova-selfcheck.txt
aws --endpoint-url <访问地址> --region <区域> s3 cp /tmp/zenova-selfcheck.txt s3://<存储桶>/zenova-selfcheck.txt
aws --endpoint-url <访问地址> --region <区域> s3 cp s3://<存储桶>/zenova-selfcheck.txt -
aws --endpoint-url <访问地址> --region <区域> s3 rm s3://<存储桶>/zenova-selfcheck.txt
```

3. 把访问地址、区域、存储桶、寻址方式交给实施工程师。访问密钥 ID 与访问密钥在安装现场录入或经凭据文件提供。

你会看到

- 三条命令依次显示上传成功、打印 `selfcheck`、显示删除成功。

注意：对象存储没有单独的加密选项。访问地址以 `https://` 开头即为加密连接，此时证书须能被安装主机的系统信任库校验。

加密方式与 CA

使用现有 PostgreSQL 或 ClickHouse 时，安装器询问加密方式，三选一，缺省为明文：

选项	含义	需要准备
「明文（不加密）」	不加密	无
「加密（校验交给系统信任库）」	加密，并校验服务器证书与主机名，CA 取自安装主机的系统信任库	数据库证书由系统信任库中的 CA 签发；访问地址里的主机名与证书一致
「加密并用指定 CA 证书校验」	加密，并校验服务器证书与主机名，CA 取自指定文件	把 CA 证书文件提前放到安装主机上，记下绝对路径，安装时要填

注意：后两种都会校验证书与主机名，区别只在 CA 从哪里来。用 IP 访问而证书里只有域名时，校验会失败，要么改用证书里的域名作为访问地址，要么由数据库管理员重签证书。生产环境建议加密。

注意：安装器的依赖检测会创建并删除临时库（zenova_probe_<随机串>）或临时对象（zenova-probe/<随机串>），不写入业务数据。如果客户的数据库审计会对建库操作告警，提前告知数据库管理员。

3.6 上传物料与校验

前提：已按第 2 章取得安装器归档与离线包，并在联网机器上核对过。

暂存目录的选择规则：

规则	说明
不放在安装根里	不要放在 /opt/zenova 或 --install-dir 指定的目录下；清空重装时安装器会拒绝处理位于待清理目录内的安装输入
空间足够	至少容纳安装器归档与离线包两个文件。离线包在安装时解开到安装根，不占暂存目录的额外空间
仅 root 可读	暂存目录之后还会放凭据文件 secrets.toml，权限设为 0700
保留到交付验收结束	安装器与离线包在续装、清空重装、远端节点安装时还要用

本手册统一使用 /root/zenova-install 作为暂存目录。

操作步骤

1. 在安装主机上建暂存目录。

```
# 安装主机 · root · 写入（创建暂存目录）
install -d -m 0700 /root/zenova-install
```

2. 在存放物料的机器上，把两个文件传到安装主机的暂存目录。登录账号无权写 /root 时，先传到该账号的家目录，再在安装主机上以 root 移入。

```
# 物料所在机器 · 普通用户 · 写入（上传到安装主机）
scp zenova-<版本>-linux-amd64.tar.gz <离线包>.zpkg <账号>@<安装主机地址>:/root/zenova-install/
```

3. 在安装主机上核对安装器摘要。<安装器摘要> 取自下载站「安装器摘要（SHA-256）」一栏，两个空格分隔摘要与文件名。

```
# 安装主机 · root · 只读（核对安装器摘要）
cd /root/zenova-install
echo '<安装器摘要>  zenova-<版本>-linux-amd64.tar.gz' | sha256sum -c -
```

4. 核对离线包大小与下载站显示的「离线包大小」一致。

```
# 安装主机 · root · 只读
ls -l /root/zenova-install
```

5. 解开安装器，查看安装器版本与产品。

```
# 安装主机 · root · 写入（在暂存目录解出 zenova 程序）
cd /root/zenova-install
tar -xzf zenova-<版本>-linux-amd64.tar.gz
./zenova --version
```

6. 查看离线包的产品、版本与通道。

```
# 安装主机 · root · 只读
tar -xOf <离线包>.zpkg release.json | grep -oE '"(edition|version|channel)": ?"[^"]*"'
```

7. 为后面的命令设置离线包变量。

```
# 安装主机 · root · 只读（仅设置当前会话的变量）
ZPKG=/root/zenova-install/<离线包>.zpkg
```

你会看到

- 第 3 步输出 `zenova-<版本>-linux-amd64.tar.gz: OK`。输出 `FAILED` 表示文件损坏或拿错，重新下载。
- 第 5 步输出一行，以版本号开头，后跟产品名（应为 `Zentrix 网关`）与括号内的构建提交。
- 第 6 步输出三行：产品（`zentrix_gateway` 或 `zenova`）、版本号、通道（`stable` 或 `test`）。

注意：以下三处必须一致，任何一处不符就不要继续，换用正确的文件：

1. `./zenova --version` 的版本号与产品名；
2. `release.json` 中的 `version` 与 `edition`；
3. 下载站上选择的版本与产品。

安装器与离线包版本不一致时，安装器会直接拒绝（「安装器版本 … 与离线包版本 … 不一致」）；产品不一致时提示「安装包与安装程序不属于同一版本」。提前核对可以省下一次失败。

注意：离线包没有单独的校验文件。安装器启动后会按 `release.json` 记录的摘要逐个校验组件，摘要校验只用于发现下载损坏，不是签名校验。

3.7 网络与防火墙准备

适用场景：所有安装。与客户网络管理员一起完成，网络策略变更通常需要提前走审批。

入站放行项

端口	用途	放行范围	安装器在 Linux 上是否自动放行
业务入口端口（访问地址对应的本机监听端口，缺省 80）	浏览器访问控制台、调用方访问模型接口	用户与调用方所在网段，或前置代理/负载均衡	是
443	网关域名站点的 TLS 监听	按需	是

端口	用途	放行范围	安装器在 Linux 上是否自动放行
管理入口端口（仅在用 `--admin-entry-port` 分离时存在）	控制台单独入口	管理员所在网段	否，需手工放行
SSH	运维登录	运维堡垒机或运维网段	不涉及

不得对外开放的端口

以下端口监听在全部网卡上，供本机智能体容器回连宿主服务使用，不得在主机防火墙、云安全组或网络边界上对外放行：

端口	用途
3010	内部服务
3080	中心网关节点数据面
3300	内部服务
3320	入口代理内部入口
3325	入口代理内部入口

其余内部服务只监听回环地址，不需要也不应从主机外访问。

出站需要放行的目标

安装过程本身不需要外网，安装所需内容都在安装器与离线包里。出站需求来自安装后要连接的对象，按交付范围逐项列出并请网络管理员放行：

目标	何时需要
现有 PostgreSQL、ClickHouse、S3 兼容存储的地址与端口	选择「使用现有的」时，安装前就必须通
运维通知渠道（钉钉、飞书、企业微信或 Webhook 地址）	安装时配置了运维通知；安装器会发送测试消息，装后验收会检查真实出网
模型供应商的接口地址	配置模型时
MCP、API 服务、外部智能体等工具地址	交付范围含工具调用时
SSO、目录同步、SMTP	交付范围含这些集成时
证书颁发机构与 DNS 服务商接口	使用入口证书自动签发时；不允许出网时改为手动上传证书
在线获取升级包	使用「系统更新」页「检查更新」时；离线环境改用「复制获取链接」

远端网关节点主机还需要能访问控制面回连地址；该地址以 https:// 开头，放行的端口以回连地址中的端口为准。

操作步骤

1. 确认业务入口端口在安装主机上空闲（以 80 为例，换成你的本机监听端口）。

```
# 安装主机 · root · 只读
ss -lntp | grep -w ':80'
```

2. 确认主机使用哪种防火墙。

```
# 安装主机 · root · 只读
systemctl is-active firewalld
systemctl is-active ufw
systemctl is-active nftables
```

3. 主机使用 ufw 或 nftables 时，安装器不会处理，由主机管理员在安装后手工放行业务入口端口（及 443、分离后的管理入口端口）。
4. 把上面的出站清单交给网络管理员，确认现有依赖的地址在安装前已经可达。

```
# 安装主机 · root · 只读（检查到现有依赖的 TCP 连通性，以 PostgreSQL 为例）
timeout 5 bash -c '</dev/tcp/<数据库地址>/<端口>' && echo reachable
```

5. 安装完成后，核对防火墙与监听情况（第 4.6 节的「监听边界」验收也会检查本产品进程的监听范围）。

```
# 安装主机 · root · 只读
firewall-cmd --list-ports
ss -lntp
```

你会看到

- 第 1 步无输出，表示端口空闲；有输出时，最后一列是占用的进程，安装前先与客户确认能否停止它，或改用其他端口。
- 第 2 步中处于 active 的那一项就是本机使用的防火墙。
- 第 4 步输出 reachable。
- 安装完成后，firewall-cmd --list-ports 中出现业务入口端口与 443。

注意：安装器只在 Linux 首次安装时，把业务入口端口与 443 写入 firewalld 或 iptables。它不放行分离后的管理入口端口，不放行安装后在控制台修改过的入口端口，不处理 ufw 与 nftables。云主机还要另外在云安全组中放行。

注意：业务入口端口被占用时，安装会停在端口检查，不会自动更换；平台内部服务的端口被占用时，安装器自动顺延到下一个空闲端口并逐条说明，不需要处理。

能力说明

不提供的能力	替代做法
在预检中发现安全模块拦截与业务入口端口占用	安装前按 3.3 节与主机安全管理员约定放行方式，按 3.7 节用 ss 确认端口空闲
自动安装或配置时间服务、Docker、数据库扩展	由客户主机管理员、数据库管理员按 3.2、3.4、3.5 节准备
使用非超级用户的 PostgreSQL 账号	换用自建 PostgreSQL 或开通超级用户；做不到时改为由安装器部署
自动放行 ufw、nftables、云安全组和分离后的管理入口端口	由主机管理员、网络管理员按 3.7 节手工放行
备份现有 PostgreSQL 与现有 S3	纳入客户数据库管理员与存储管理员的备份体系
安装后在线更换对象存储方式	安装前定好；需要更换时备份后按新选择重新安装，再恢复数据

4. 交互式安装

本章内容：本章按安装器在终端里的实际顺序，从预检、启动、逐屏问答，到安装执行、装后验收和完成输出，一步步说明每一屏看到什么、该填什么、出了例外怎么处理，最后说明中途中断后怎样继续。读者是在安装主机上操作的实施工程师；涉及数据库、存储与网络的屏，需要对应的客户管理员在场或提前按第 3 章准备好。

4.1 安装流程总览

安装器的全部步骤如下。确认安装之前不写任何文件、不创建任何服务；任一步输入 q 退出，本机不留任何内容（退出码 130）。问答没有「返回上一步」，填错了只能输入 q 退出后重来，或在确认页之前对相应问题按提示修改。

顺序	阶段	是否写盘	预检（--check-only）是否包含
1	读取并校验离线包、核对安装器与离线包的版本和产品，显示「将安装：…」	否	包含
2	选择能力	否	包含
3	确认预填配置（离线包内有定制站预填时）	否	包含
4	补充配置：访问地址与本机监听端口、平台数据库、分析数据库、对象存储、运维通知渠道	否	包含

顺序	阶段	是否写盘	预检（--check-only）是否包含
5	主机检查	否	包含
6	业务入口端口检查与内部端口分配	否	不包含
7	录入凭据（仅「使用现有的」依赖）	否	不包含
8	依赖检测（仅「使用现有的」依赖）	否（只创建并删除临时探测对象）	不包含
9	运维通知测试（配置了通知渠道时）	否	不包含
10	确认页	否	不包含
11	写入部署配置、解开离线包、以服务用户身份检查程序执行权限	是，从这里开始写盘	不包含
12	安装执行：数据层 → 控制面 → 已选能力 → 装后验收	是	不包含
13	生成 root 初始密码，打印完成信息	是	不包含

注意：第 11 步以后，安装器在 `/opt/zenova/state/` 下记录进度。此后中断或失败，重新执行同一条命令即可从断点继续（4.8 节），不要手工删除任何文件。

4.2 先跑预检

适用场景：正式安装前，在安装主机上确认主机条件与问答内容。预检可以反复执行，不留任何内容。

前提：已按第 3.6 节把物料放到 `/root/zenova-install` 并核对。

操作步骤

1. 进入暂存目录，设置离线包变量，执行预检。

```
# 安装主机 · root · 只读（预检，不写任何内容）
cd /root/zenova-install
ZPKG=/root/zenova-install/<离线包>.zpkg
sudo ./zenova install "$ZPKG" --check-only
```

2. 按 4.4 节回答能力选择、预填确认与补充配置各屏的问题。预检里填的内容不会保存，正式安装时还要再答一次，可以借预检把答案先理顺并记下来。
3. 主机检查有 `✗` 项时，按屏幕提示与第 3 章处理，按回车重新检查。

4. 需要把安装目录放到其他数据盘时，预检也要带上同样的参数，例如 `--install-dir <目录>`，否则检查的是默认的 `/opt/zenova`。

你会看到

- 最后一行输出「预检通过：未写入任何内容」，命令退出。

注意：预检不包含以下检查，预检通过不代表正式安装一定通过：

预检不含的检查	怎样提前排除
业务入口端口是否被占用	第 3.7 节用 `ss -lntp` 确认
现有依赖检测（数据库登录、权限、扩展、版本、宏、对象存储读写删）	第 3.5 节由客户管理员自查
解包后的程序执行权限（SELinux、fapolicyd、麒麟 KYSEC）	第 3.3 节与主机安全管理员约定放行方式
运维通知测试	确认安装主机能访问通知渠道地址

4.3 启动安装

前提：预检已通过；客户方管理员已就位（使用现有依赖时需要数据库管理员、存储管理员到场录入密码，或已准备好凭据文件）。

操作步骤

1. 建议在 `tmux` 或 `screen` 会话中执行（主机上已有时），避免 SSH 断开打断安装。
2. 执行安装命令。

```
# 安装主机 · root · 写入（确认后开始安装）
cd /root/zenova-install
sudo ./zenova install "$ZPKG"
```

3. 核对屏幕第一行「将安装：<产品> <版本>」，不对就输入 `q` 退出。

主要参数（交互式安装通常一个都不用加，按需选用）：

参数	说明
`--install-dir <目录>`	安装根，默认 `/opt/zenova`。必须是绝对路径，只能含字母、数字、点、下划线、短横线、斜杠；不能放在 `/tmp`、`/var/tmp`、`/home`、`/root`、`/run/user`、`/usr`、`/etc`、`/boot`。显式传了它，磁盘不足只告警不阻断
`--log-dir <目录>`	独立日志目录，必须在安装根之外，首次使用时为空
`--business-entry-port <端口>`	业务入口的本机监听端口。访问地址前面有代理、NAT 或 TLS 入口时，用它指定转发到的本机端口

参数	说明
`--admin-entry-port <端口>`	管理入口（控制台）端口。缺省与业务入口共用；指定不同端口即分离
`--with <能力,…>`	本次启用的能力，逗号分隔；核心总包含。给出后不再显示能力选择页
`--answers <文件>`	答案文件，给出后不再交互提问（非交互安装见第 5 章）
`--secret-file <文件>`	凭据文件，权限 0600；凭据只能通过它或终端输入提供
`--reconfigure`	保留数据，回到「选择能力」，带入上次的答案重新确认
`--check-only`	只做预检，不写入

你会看到

- 第一行形如「将安装：Zentrix 网关 0.7.0-nightly.20261020034227.f51c370a」。

注意：第一行的产品名由安装器决定，版本号来自离线包。核对两点：产品是否是本次要交付的产品（安装后不能更换产品，只能清空重装）；版本号是否与交付记录一致。安装器与离线包版本或产品不一致时，安装器在这一行之前就会拒绝并说明原因。

4.4 逐屏说明

下面按出现顺序说明每一屏。定制站已经预填的项不会再问，而是在「确认预填配置」屏统一确认。

4.4.1 能力选择

你会看到

- 标题「本次要安装的能力（核心必装；未勾选的会保留在本机，之后可在控制台直接启用）」，下面是离线包中包含的能力列表，带编号与勾选状态。
- 提示「输入编号切换，回车完成」。
- Zentrix 网关首次安装默认勾选「核心」与「语义缓存」；完整版默认勾选离线包中的全部能力。

操作步骤

1. 输入编号切换某项的勾选状态，可以多次输入。
2. 核对勾选结果后直接回车。

注意：「核心」始终安装，不能取消。依赖关系自动联动：勾选「文档解析」会一并勾选「知识库」。离线包里有但本次没勾的能力保留在本机，装完后在控制台「系统更新」页直接启用，不需要重新下载。勾选会影响磁盘需求（第 3.1 节），磁盘紧张时只勾本次真要验证的能力。

4.4.2 确认预填配置

适用场景：下载站第四步填了预填配置。整页留空时没有这一屏。

你会看到

- 标题「以下配置来自定制站」，分两级列出：
  - 一般项（访问地址、依赖的部署方式、对象存储模式、通知渠道类型）合并列出，提示「请直接回车确认，或输入 n 修改」。
  - 地址类项（现有 PostgreSQL、ClickHouse、对象存储的地址，通知渠道地址）逐项显示并标注「来自定制站」，提示「这一项需要输入 y 才能确认」。

操作步骤

1. 一般项核对无误直接回车；需要改时输入 `n`，安装器随后重新询问。
2. 地址类项逐项输入 `y` 确认；不对时拒绝，安装器当场要求改填。

注意：预填内容在离线包里标记为未签名，地址类项必须逐项输入 `y`，直接回车不算确认。这是为了防止离线包在传递途中被改了地址。

4.4.3 访问地址与本机监听端口

你会看到

- 「访问地址（回车使用本机 <本机地址>；也可填写 http://… 或 https://…；地址里写的端口即本机监听端口，不写端口为 80）」。

操作步骤

1. 按下表填写访问地址。直接回车使用本机地址 `http://<本机地址>`。
2. 访问地址是 `http://` 且带非 80 端口时，安装器再问「访问地址前面是否有代理或 NAT 转发」：
   - 选「没有，本机直接监听访问地址里的端口」，本机就监听访问地址里的端口；
   - 选「有，由代理或 NAT 转发到本机的另一个端口」，再填写「本机监听端口（代理、NAT 或 TLS 入口转发到的端口，回车为 80）」。
3. 访问地址是 `https://` 时，安装器直接询问 TLS 入口转发到的本机监听端口，回车为 80。

场景	访问地址	本机监听端口	怎么填
内网直连	`http://<主机地址>`	80	只填访问地址
内网直连，非 80 端口	`http://<主机地址>:8080`	8080	只填访问地址；「是否有代理或 NAT 转发」选「没有」
前面有代理或 NAT	`http://<对外域名>:18080`	8080	「是否有代理或 NAT 转发」选「有」，填 8080
前置 TLS 入口	`https://<对外域名>`	80	填 TLS 入口转发到的本机端口，回车为 80

注意：

- 访问地址是用户和调用方最终使用的地址，填 `https://` 不会让本机提供 HTTPS，证书由客户的负载均衡或反向代理终结。

- 访问地址的端口与本机监听端口不同时，必须由前置代理、NAT 或 TLS 入口把访问地址的端口转发到本机监听端口，否则访问地址打不开。
- `http://` 访问地址下，登录页不提供通行密钥，两步验证只能用验证器应用。客户要用通行密钥时，访问地址必须是 `https://`。
- 地址格式不对时提示「访问地址必须是 http:// 或 https:// 开头的完整地址」，当场重填。填了不带点的单段主机名且解析不到时，安装器会说明风险并要求输入 `y` 才继续使用。

4.4.4 平台数据库（PostgreSQL）

你会看到

- 「平台数据库（PostgreSQL）」，两个选项：「由我们部署」「使用现有的」。

操作步骤

1. 无特殊要求选「由我们部署」：安装器在本机初始化 PostgreSQL，只监听回环地址，后续不再提问。
2. 选「使用现有的」时，依次填写：
   - 「平台数据库访问地址（主机:端口）」，例如 `<数据库地址>:5432`；
   - 「平台数据库管理员账号（须为超级用户，安装器用它创建平台的角色与库）」；
   - 「平台数据库加密方式」：「明文（不加密）」「加密（校验交给系统信任库）」「加密并用指定 CA 证书校验」，缺省为明文；选第三项时再填「CA 证书文件路径」，必须是安装主机上已存在的文件。

注意：不问库名，平台的库由安装器创建、库名固定。管理员账号的密码在后面的「凭据」一屏录入。各项要求与数据库管理员自查方法见第 3.5 节。

4.4.5 分析数据库（ClickHouse）

你会看到

- 「分析数据库（ClickHouse，必装）」，两个选项：「由我们部署」「使用现有的」。

操作步骤

1. 无特殊要求选「由我们部署」。
2. 选「使用现有的」时，依次填写：
   - 「分析数据库访问地址（主机:端口）」，填 HTTP 或 HTTPS 接口端口（常见 8123 / 8443），不是 9000；
   - 「分析数据库库名（回车使用 zenova）」；
   - 「分析数据库登录账号」，按第 3.5 节为 `default`；
   - 「分析数据库加密方式」，三个选项同 PostgreSQL；
   - 「分析数据库的库类型」：「单机库（普通 MergeTree）」或「复制库（库引擎 Replicated，建表带 ON CLUSTER）」，缺省为单机库，按数据库管理员告知的选择。

注意：ClickHouse 是必装组件，没有「不安装」选项。本机 CPU 指令集不满足要求时，主机检查会提供「重新选择分析数据库」，回到这一屏改为「使用现有的」（4.4.8 节）。

4.4.6 对象存储

你会看到

- 「对象存储」，三个选项：「自带对象存储（推荐，生产）」「使用本机目录（单机试用）」「使用现有的 S3 兼容存储」，缺省为第一项。

操作步骤

1. 单机安装推荐选「自带对象存储（推荐，生产）」，安装器在本机部署并自动建桶。
2. 只做短期试用时可以选「使用本机目录（单机试用）」。
3. 选「使用现有的 S3 兼容存储」时，依次填写「对象存储访问地址（http://… 或 https://…）」「区域」「存储桶」「对象存储寻址方式」（「路径寻址」或「虚拟主机寻址」，缺省路径寻址）。访问密钥在「凭据」一屏录入。

注意：对象存储方式安装后不能在线更换，需要更换时只能备份后按新的选择重新安装，再恢复数据。POC 之后要转生产的环境，这里直接按生产的选择填。

4.4.7 运维通知渠道

你会看到

- 「运维通知渠道」，选项：「钉钉」「飞书」「企业微信」「Webhook」「暂不配置」。

操作步骤

1. 有渠道时选择类型，再填「渠道地址」（机器人或回调的完整地址，必须以 `http://` 或 `https://` 开头）。
2. 暂时没有渠道时选「暂不配置」，安装完成后在控制台补配。

注意：配置了渠道时，安装器在确认页之前会发一条测试消息（4.4.11 节）。选「暂不配置」时，装后验收中的「真实出网」「运维渠道绑定」两项按通过处理，但平台的运维告警在补配之前发不出去。

4.4.8 主机检查

你会看到

- 标题「主机检查」，逐项列出检查结果：`✓` 通过、`⚠` 提示、`✗` 阻断。检查项包括 CPU 架构、内存、磁盘空间、系统服务托管、系统时间，以及按本次选择适用的处理器指令集、容器运行时、目录遍历权限、挂载执行权限、服务用户属性等。
- 「系统时间」通过时，下面一行显示「已同步，时间偏差 N 秒（安装前上限 30 秒）」。
- 未通过的项逐条列出原因与处理办法，不会因第一项失败就停止其余检查。

操作步骤

1. 全部为 `✓`，或只有 `⚠`：自动继续。
2. 有 `✗`：另开一个终端按屏幕提示与第 3 章处理，回到安装器按回车重新检查；输入 `q` 退出。
3. 只有「处理器指令集」阻断时，可以选「重新选择分析数据库」，把分析数据库改为「使用现有的」。

常见阻断	处理
内存不足：需要 8 GB，本机 N GB	扩容
磁盘空间不足（HOST_DISK_LOW）	屏幕会列出可用空间最大的 3 个挂载点。输入 q 退出，加 --install-dir <目录> 指向其中一个重新执行；或扩容、减少本机部署的组件
HOST_CLOCK_UNSYNCED / HOST_CLOCK_SKEW	按第 3.2 节处理时间同步
HOST_CPU_FEATURES	选「重新选择分析数据库」改用现有 ClickHouse，或调整虚拟机 CPU 型号后重新检查
HOST_ROOT_TRAVERSE / HOST_NOEXEC_MOUNT / HOST_SERVICE_USER	按第 3.3 节处理
HOST_CONTAINER_RUNTIME	按第 3.4 节处理

注意：显式传了 --install-dir（包括显式写 /opt/zenova）时，磁盘不足显示为 ⚠ 并直接继续。这只是不阻断，空间实际不足时仍可能在解包、安装或运行时失败。

4.4.9 端口

你会看到

- 平台内部服务的端口被占用时，安装器自动改用下一个空闲端口，并逐条说明改用了哪个端口，不需要处理。
- 业务入口端口被占用时，显示「本机监听端口 <端口> 已被进程 <进程名>（PID <进程号>）占用。访问地址依赖这个端口，安装不会自动更换；…」，并提示「释放端口后按回车重新检查，输入 n 重新填写访问地址与本机监听端口，输入 q 退出」。
- 业务入口端口与平台内部端口相同时，提示该端口是平台内部服务使用的端口，不能作为访问入口。

操作步骤

1. 业务入口端口被占用时，与客户确认后停止占用进程，回到安装器按回车重新检查。
2. 不能停止占用进程时，输入 n 重新填写访问地址与本机监听端口。命令行带了 --business-entry-port 时不能在问答里改，输入 q 退出后换端口重新执行。
3. 管理入口分离时，其端口被占用会自动顺延到下一个空闲端口，完成输出中的「本机直接访问」按实际端口给出。

注意：这一屏在预检中不出现。改了本机监听端口后，访问地址的端口与本机监听端口可能不再一致，此时必须有前置转发（4.4.3 节）。

4.4.10 凭据与依赖检测

适用场景：至少一项依赖选了「使用现有的」。三项都由安装器部署时没有这一屏。

你会看到

- 依次询问（输入内容不显示）：「平台数据库管理员账号 <账号> 的密码」「分析数据库的登录密码」「对象存储访问密钥 ID」「对象存储访问密钥」，只问本次选了「使用现有的」那几项。
- 录入后立即检测。全部通过时直接进入下一屏；有未通过项时显示标题「依赖检测」，逐条列出「<依赖>：<原因>，请联系<负责人>」，并提示「修正后按回车重新检测，输入 q 退出」。

操作步骤

1. 由数据库管理员、存储管理员本人在终端录入密码与密钥；不要把口令写在聊天记录或纸面上转交。
2. 检测未通过时，按屏幕列出的原因与负责人处理（对照第 3.5 节自查），修正后按回车重新检测。

检测内容：

依赖	检测项
PostgreSQL	管理员账号能否登录、是否为超级用户、所需扩展与版本、max_connections
ClickHouse	版本、cluster/shard/replica 宏、建库权限、库类型与实际引擎是否一致
S3 兼容存储	对存储桶的读、写、删

注意：检测会创建并删除临时库（zenova_probe_<随机串>）或临时对象（zenova-probe/<随机串>），不写入业务数据。PostgreSQL 管理员密码只在安装过程中使用，不写入任何文件与日志；安装完成后平台只用自己的角色访问数据库。

注意：如果现有 PostgreSQL 的平台库中已经有平台数据，普通安装会在这里停止并提示改用接管安装（--adopt-existing）。接管不在本章范围，联系技术支持确认后再操作。

4.4.11 运维通知测试

适用场景：4.4.7 节配置了通知渠道。

你会看到

- 发送成功：「测试消息已发送，请确认已收到」。
- 发送失败：「运维通知测试：<失败原因>」，并提示「输入新的渠道地址后回车重发，输入 s 跳过，输入 q 退出」。

操作步骤

1. 到对应的群或接收端确认收到测试消息。
2. 发送失败时，核对渠道地址与安装主机的出网，输入正确的地址回车重发；暂时解决不了时输入 s 跳过，等同于「暂不配置」。

注意：「已发送」只表示渠道接口接受了请求。务必到群里确认真的收到了，例如钉钉机器人设置了关键词或加签校验时，消息可能被渠道一侧拦下。

4.4.12 确认页

你会看到

- 标题「即将开始安装，请确认以下内容」，逐项列出：
  - 访问地址，紧跟「本机监听端口」。两者不同时提示「访问地址的端口与本机监听端口不同，需由前置代理或 NAT 把访问地址的端口转发到本机监听端口」；访问地址为 `https://` 时提示「访问地址使用 https，需由前置 TLS 入口终结加密后转发到本机监听端口」。
  - 平台数据库、分析数据库、对象存储、运维通知渠道的选择与地址。
  - 「能力：…」。
  - 「安装目录：/opt/zenova（需要 N GB 磁盘空间）」；显式传了 `--install-dir` 时显示「建议」。
  - 「服务账户 zenova：安装根外的祖先目录须可遍历，系统目录权限由系统管理员管理」。
- 最后一行「确认开始安装」。

操作步骤

1. 逐项与交付记录核对，特别是访问地址、本机监听端口和对象存储方式。
2. 无误后回车确认；有误则输入 `q` 退出，重新执行安装命令。

注意：这是最后一个不写盘的点。回车之后安装器开始写入部署配置、解开离线包，之后再退出就会留下部分安装，只能按 4.8 节续装或按清空重装流程处理。

4.5 安装执行

你会看到

- 先解开离线包，并以服务用户身份检查随包程序能否执行。被安全模块拦截时，逐条显示「✗ [<代码>] <原因> <处理办法>」并停止（处理见第 3.3 节，之后按 4.8 节续装）。
- 随后依次执行四个阶段，每个阶段开始时显示「正在安装：<阶段>」，结束时显示「<阶段> 已完成」。

阶段	内容
数据层	初始化由安装器部署的 PostgreSQL、ClickHouse、自带对象存储；三项都使用现有服务时跳过初始化
控制面	创建平台角色与库、生成运行配置、安装服务单元、执行数据库迁移、启动服务并等待就绪、放行防火墙、创建存储桶。中心网关节点 `office-primary` 随控制面一起启动
已选能力	启用本次勾选的能力
装后验收	执行 4.6 节的验收项

- 四个阶段结束后显示「各阶段耗时」表，以及「逐项计时与迁移分步耗时已写入安装日志」。

操作步骤

1. 等待安装执行，不要中断。三项依赖都由安装器部署的干净主机上，全程约 4 分钟，其中数据库迁移约占一半，控制面阶段停留较久属于正常现象。
2. 需要观察细节时，另开一个终端查看安装日志：

```
# 安装主机 · root · 只读（跟踪安装日志）
tail -f /opt/zenova/state/install.log
```

使用 `--install-dir` 时，把 `/opt/zenova` 换成实际安装根。

注意：某一阶段失败时，屏幕停在失败的步骤并给出原因与下一步。处理原因后重新执行同一条安装命令，已完成的阶段会跳过（4.8 节）。安装期间控制台不可用；首次安装尚无网关业务流量，不影响任何现有业务。

4.6 装后验收

装后验收是安装的最后一个阶段，自动执行，不需要输入。全部阻断级验收项通过，安装才记为完成。

验收项	级别	适用条件	检查内容
配置完整性	阻断	全部	已发布的部署配置完整，满足本部署形态的要求
工作负载就绪	阻断	全部	本机每个启用的服务都已就绪并应答
告警引擎	阻断	全部	平台告警引擎最近 60 秒内完成过一轮判定，且内置告警规则已全部载入
原生事件投递	阻断	全部	平台内部事件能正常投递
真实出网	阻断	全部（未配置通知时通过）	本机能连上已绑定的运维通知渠道的接收端
密钥服务金丝雀	阻断	全部	以控制面身份写入、读取、吊销一条测试密钥，吊销后不能再读取
模型网关健康	阻断	全部	本机网关数据面的业务端口存活并能服务
智能体容器可达性	警告	启用托管智能体	当前显示「检查项尚未实现」
监听边界	阻断	全部	本产品进程的监听端口与绑定范围符合端口表，内部端口没有越界监听
运维渠道绑定	阻断	全部（未配置通知时通过）	安装时填的运维通知渠道已登记到平台，并绑定到运维告警路由

验收项	级别	适用条件	检查内容
存量渠道合规	警告	全部	平台中已有的通知渠道是否合规
数据依赖可达性	阻断	全部	以平台运行期账号逐项连接 PostgreSQL、ClickHouse、对象存储
网关节点健康	阻断	全部	中心网关节点就绪，最多等待 150 秒

你会看到

- 验收通过：进入 4.7 节的完成输出，其中包含验收结果。
- 验收失败：屏幕显示「装后验收失败：未通过的检查：<验收项>（<依据>）」，安装不记为完成。
- 警告级项未通过只提示，不阻止安装完成。

操作步骤（验收失败时）

1. 按「依据」中的原因处理。常见情形：

失败项	常见原因	处理
工作负载就绪 / 模型网关健康	某个服务启动失败	`sudo zenova status` 查看状态，`sudo zenova logs <组件> --since '30 min ago'` 查看日志
真实出网 / 运维渠道绑定	安装主机到通知渠道的出网被拦截	按第 3.7 节放行出站
数据依赖可达性	现有数据库或对象存储在安装过程中变得不可达、账号或网络策略被改动	与数据库管理员、存储管理员核对
监听边界	有端口监听范围超出预期	按依据中列出的端口与进程处理，必要时联系技术支持
网关节点健康	150 秒内中心网关节点未就绪	查看 `sudo zenova logs` 中网关相关组件的日志
其他	以依据中给出的原因为准	执行下面的信息采集命令后联系技术支持

2. 处理后重新执行同一条安装命令，安装器从装后验收继续，不重做前面的阶段。

```
# 安装主机 · root · 写入（从装后验收继续）
cd /root/zenova-install
sudo ./zenova install "$ZPKG"
```

3. 需要远程支持时，采集信息：

```
# 安装主机 · root · 只读
sudo zenova status --json > /root/zenova-install/status.json
sudo zenova doctor --all --save /root/zenova-install/doctor
sudo cp /opt/zenova/state/install.log /root/zenova-install/
```

注意：装后验收通过只证明平台基础服务可用。模型与工具的配置、真实调用、拒绝路径等，按第 7 章逐项完成。

4.7 完成输出

你会看到

安装完成后，屏幕依次打印：

- 「各阶段耗时」表与验收结果；
- 「安装完成」；
- 「管理台登录账号（初始密码只保存在本机，日志中不记录）」，下面是登录账号 root 与 32 位随机字符的初始密码（具体行格式以实际输出为准）；
- 「本机直接访问：http://<本机地址>[:<端口>]」；
- 「访问地址：<访问地址>」；
- 「下一步：用上面的账号登录控制台，按提示修改初始密码并绑定两步验证，然后完成初始化向导」。

两个地址的区别：

地址	怎么来的	用途
本机直接访问	按管理入口的实际监听端口拼出，始终是 http://，不经任何前置转发	在内网直接打开控制台；访问地址打不开时，用它判断问题出在前置设备还是本机
访问地址	安装时填写的地址	日常使用与对外提供服务的地址

操作步骤

1. 立即按客户的口令管理规范保管初始密码。
2. 首次修改密码前需要再次查看时，在安装主机上执行：

```
# 安装主机 · root · 只读（查看初始密码）
sudo zenova root-password show
```

3. 用浏览器打开「访问地址」登录，按第 6 章完成首次登录与初始化。打不开时先试「本机直接访问」。

注意：初始密码不得通过即时通讯、邮件或工单传递，不要录屏或截图外传。 初始密码同时保存在 /opt/zenova/.env.runtime/root-password-initial（权限 0600），首次修改密码后失效，root-password show 也随之不再可用。root 忘记密码或两步验证全部丢失时，用 sudo zenova root-password set 重置（第 8.7 节）。

注意：完成输出之后，务必按第 8.1、8.2 节完成密钥材料离机备份与首次平台备份，这两项不做，平台出故障时无法恢复。

4.8 中断与续装

确认安装之后，安装器把每个已完成的阶段记录在 `/opt/zenova/state/` 下。重新执行同一条安装命令，已完成的阶段会跳过，从断点继续。

情形	本机状态	处理
确认页之前输入 `q` 退出	没有写入任何内容（退出码 130）	重新执行安装命令，从头作答
确认页之前 SSH 断开或终端被关闭	没有写入任何内容	重新执行安装命令，从头作答
确认后，解包后的执行权限检查失败	已写入部署配置并解开离线包	按第 3.3 节放行后，重新执行同一条命令
确认后，某个阶段失败	已完成的阶段已记录	按屏幕原因处理后，重新执行同一条命令，从失败的阶段继续
装后验收失败	前三个阶段已完成	按 4.6 节处理后，重新执行同一条命令，从装后验收继续
安装执行中 SSH 断开	视断开时进行到哪一步	重新登录，重新执行同一条命令；用 `tail /opt/zenova/state/install.log` 查看断开前的进度
续装时屏幕显示「上次安装的答案：」	安装器沿用上次的答案	核对后回车「以上答案继续使用」。上次的答案中已不合法的项会提示「以下上次的答案不再合法，将重新询问：」并重新询问
续装时 `--with` 与上次不同	—	交互时列出「新增」「移除」，输入 `y` 确认后按本次 `--with` 继续；非交互时直接拒绝
续装时答案与上次不一致	—	提示「答案与上次安装不一致；如需修改配置，请加 --reconfigure 重新执行」。恢复原答案继续，或加 `--reconfigure`
已安装完成后又执行安装命令	安装已完成	提示「本机已完成安装」。升级用 `sudo zenova upgrade <包.zpkg>`；修改安装配置加 `--reconfigure`；需要删除旧数据重新首装按清空重装流程
想改安装时填写的配置	—	`sudo ./zenova install "$ZPKG" --reconfigure`，保留数据，回到「选择能力」并带入上次的答案
清空重装时，旧部署已清理、后续安装失败	旧数据已删除，新安装进度已记录	处理后对同一个包执行普通安装命令，不再加 `--fresh`

操作步骤

1. 查看失败原因：屏幕上的提示，或安装日志。

```
# 安装主机 · root · 只读
tail -n 100 /opt/zenova/state/install.log
```

2. 按原因处理。
3. 重新执行与上次完全相同的命令（参数不变）。

```
# 安装主机 · root · 写入（从断点继续安装）
cd /root/zenova-install
sudo ./zenova install "$ZPKG"
```

你会看到

- 已完成的阶段不再显示「正在安装」，安装器直接从未完成的阶段开始。

注意：不要删除 `/opt/zenova/state/` 下的进度文件（`install.json` 等），也不要手工删除 `/opt/zenova` 下的任何目录或停用服务来「重来」。进度文件被删后，安装器无法判断哪些已完成，现场会处于无法续装、也无法安全清理的状态。确实需要推倒重来时，只能在满足条件的环境中用清空重装流程（`--fresh --dry-run` 预览后再执行），严禁手写 `rm -rf`、`docker rm`、`systemctl` 等手工清理命令。

注意：续装时参数要与首次一致，特别是 `--install-dir`、`--log-dir`、`--business-entry-port`。换了参数，安装器会认为是另一套安装或拒绝继续。

能力说明

不提供的能力	替代做法
问答中返回上一步	确认页之前输入 `q` 退出后重新执行；安装后用 `--reconfigure` 带着上次的答案重新确认
预检覆盖端口、依赖、执行权限与通知	按 4.2 节表格在安装前分别排除
业务入口端口被占用时自动更换	停止占用进程，或重新填写访问地址与本机监听端口，或用 `--business-entry-port` 配合前置转发
本机直接提供 HTTPS	由客户的负载均衡或反向代理终结 TLS，访问地址填 `https://`
安装完成后再次查看已修改过的初始密码	用 `sudo zenova root-password set` 重置
手工清理后重装	满足条件时使用清空重装流程；不满足时联系技术支持

5. 非交互安装

本章内容：非交互安装用一份答案文件（`--answers`）和一份凭据文件（`--secret-file`）代替问答，安装器全程不提问，适合批量部署、远程执行和无人值守。本章给出两份文件的全部字段、写错时的后果，以及两套可以直接改值使用的完整示例。读者是需要用脚本或远程会话完成安装的实施工程师；只装一台、现场有人值守时，用第 4 章的

交互安装即可。

5.1 什么时候用非交互安装

适用场景：同一套参数要装多台主机；只能通过远程会话或自动化工具执行；安装过程中无人在终端前回答问题。

非交互安装与交互安装走的是同一个安装流程、同样的主机检查、依赖检测、四个安装阶段和装后验收，区别只在「答案从哪里来」和「遇到问题时怎么处理」。交互安装遇到问题会停下来让你当场修正；非交互安装遇到问题直接报错退出，修正文件后重跑同一条命令。

规则一览

规则	说明
进入非交互的条件	命令行给出 `--answers <文件>`。同时加了 `--reconfigure` 时仍会进入问答，不是非交互
必填段	`[postgres]`、`[clickhouse]`、`[object_store]` 三段必须写；运维通知必须二选一：写 `[notify]` 段，或在顶层写 `notify_skipped = true`。缺任何一项都会报「答案不合法：缺少 <项目>」。缺 `[clickhouse]` 时单独报分析数据库为必装
访问地址缺省	不写 `public_url` 时，安装器取本机出口地址写成 `http://<本机地址>`（加了非 80 的 `--business-entry-port` 时带上该端口）。有代理、NAT、TLS 入口或域名时必须显式写
答案优先级	`--answers` 给出的文件 > 上一次安装保存的答案（续装与 `--reconfigure`）> 离线包内的预填配置。`--answers` 是整份替换，不与离线包预填合并
凭据只走凭据文件	口令与访问密钥只能写在 `--secret-file` 指定的文件中，不得写进答案文件。非交互时安装器不会在终端补问凭据，缺凭据即停止
凭据文件权限	组和其他用户不得有任何权限（0600 或更严），否则报「凭据文件权限必须是 0600：<路径>」
能力	用 `--with <能力,…>` 指定，答案文件里没有能力字段。不给 `--with` 时：Zentrix 网关启用核心与语义缓存，完整版启用核心与知识库
主机检查	全部检查项照常执行并逐条打印，带 ⚠ 的是提示项；只要有阻断项，安装器报出第一个阻断项后停止，不进入后续步骤
通知测试	写了 `[notify]` 时，安装器发送一条测试消息；发送失败即安装失败，屏幕显示「运维通知测试失败：<原因>」。非交互时不提供「跳过」
未定义字段	答案文件与凭据文件都只接受本章列出的字段。多写一个字段（包括手工混入的 `password`、`license`）直接报「答案文件里有不认识的字段：<字段名>」

规则	说明
确认页	非交互时不显示确认页，检查与检测全部通过后直接开始写入

操作步骤

1. 按 5.2 节写好答案文件 answers.toml。
2. 有任何一项依赖使用现有服务时，按 5.3 节写好凭据文件 secrets.toml 并设为 0600。三项依赖都由安装器部署时不需要凭据文件。
3. 先加 --check-only 跑一次预检，通过后去掉 --check-only 正式执行。完整命令见 5.4、5.5 节。

你会看到

- 预检通过时输出「预检通过：未写入任何内容」。
- 正式执行时屏幕依次打印主机检查结果、依赖检测结果、通知测试结果（如配置）、四个安装阶段与装后验收，最后打印登录账号与初始密码（见第 4 章 4.7 节）。

注意：安装中断或失败后，修正原因并重跑同一条命令，安装器从断点继续。续装时答案文件内容必须与第一次一致，否则报「答案与上次安装不一致；如需修改配置，请加 --reconfigure 重新执行」；--with 与上次不同也会被拒绝并列出差异。--reconfigure 会进入问答，需要有人在终端前操作。

5.2 答案文件字段说明

答案文件是 TOML 格式。顶层字段必须写在所有 [段] 之前——写在某个段之后的顶层字段会被当作该段的成员，随即报「不认识的字段」。

顶层字段

字段	必填	取值	写错的后果
public_url	可选（强烈建议写）	访问地址，http://… 或 https://…，可带端口，例如 "http://<主机地址>:8080"	格式不合法时报「访问地址地址不合法」并停止。不写时取本机出口地址，有代理、NAT、域名时会得到错误的访问地址
business_listen_port	可选	1～65535 的整数。只在访问地址前面有代理、NAT 或 TLS 入口时填写，表示转发到本机的监听端口	写 0 报「本机监听端口必须是 1 到 65535 之间的数字」；写了它却没写 public_url 报「填写本机监听端口时必须同时填写访问地址」。与命令行 --business-entry-port 作用相同，两者都给时以命令行为准
notify_skipped	与 [notify] 二选一	true 表示暂不配置运维通知	既不写它也不写 [notify] 时报「缺少」通知渠道。写 false 等于没写

另有两个顶层可选字段，POC 一般不用：`admin_public_url`（管理访问地址，只在控制台单独走反向代理或独立域名时填写，留空时由访问地址的主机与管理入口端口推导）、`otlp_endpoint`（链路导出地址）。两者都必须是合法的 `http://` 或 `https://` 地址。

`[postgres]` 平台数据库

只能写下面两种之一。

写法	含义
`managed = {}`	由安装器在本机部署，只监听回环
`external = { host = "<地址>", port = 5432, user = "<超级用户账号>" }`	使用现有 PostgreSQL

`external` 内的字段：

字段	必填	取值	写错的后果
`host`	必填	主机名或 IP	不合法时报「平台数据库 地址不合法」
`port`	必填	端口，常见 5432	写 0 同上
`user`	必填	超级用户账号（`rolsuper` 为真），安装器用它创建平台的角色与库	不写报「使用现有平台数据库时必须填写管理员账号」；不是超级用户时依赖检测不通过
`tls`	可选	省略为明文；`"required"` 要求加密，用系统信任库校验证书与主机名；`{ verify_full = { ca_file = "<CA 文件绝对路径>" } }` 加密并用指定 CA 校验	证书由内部 CA 签发却写 `"required"` 时，依赖检测会因证书校验失败而不通过
`database`	不得填写	—	平台的三个库由安装器创建、库名固定。写了报「平台数据库的库由安装器创建、库名固定，答案里不能填写库名」
`replicated`	不得填写	—	写了报「复制库（replicated）只对分析数据库 ClickHouse 有意义，平台数据库不能填写」

`[clickhouse]` 分析数据库（必装）

写法	含义
managed = {}	由安装器在本机部署，只监听回环。需要 CPU 支持第 1 章 1.3 节所列指令集
external = { host = "<地址>", port = 8123, database = "<库名>", user = "default" }	使用现有 ClickHouse

external 内的字段：

字段	必填	取值	写错的后果
host、port	必填	HTTP 接口端口（常见 8123）或 HTTPS 接口端口（常见 8443），不是原生端口 9000	写成 9000 时依赖检测连不上
database	建议填写	平台使用的库名，例如 "zenova"	省略时的行为以实际输出为准
user	建议填写	登录账号，通常为 default，密码写在凭据文件 ch_password	账号无建库权限时依赖检测不通过
tls	可选	同 [postgres]	同 [postgres]
replicated	可选	库为复制库（库引擎 Replicated）时写 true；省略即单机库	声明与库的实际引擎不一致时，安装前检测会拒绝

[object_store] 对象存储

只能写下面三种之一。安装后不能在线更换。

写法	含义
bundled = {}	自带对象存储，推荐生产单机使用
local_dir = {}	本机目录，仅用于单机试用
external_s3 = { endpoint = "<地址>", region = "<区域>", bucket = "<桶>", addressing = "path" }	使用现有 S3 兼容存储

external_s3 内的字段（四项全部必填）：

字段	取值	写错的后果
endpoint	http://… 或 https://…，可带端口。是否加密由地址开头决定，external_s3 内没有 tls 字段	地址不合法时报「对象存储 地址不合法」；多写 tls 报不认识的字段

字段	取值	写错的后果
region	区域，例如 "us-east-1"	与存储端不一致时依赖检测读写失败
bucket	预先建好的桶名	桶不存在时依赖检测不通过
addressing	"path"（路径寻址）或 "virtual-hosted"（虚拟主机寻址）	缺这一项时答案文件被拒绝；写成其他值报格式错误

[notify] 运维通知渠道

字段	必填	取值	写错的后果
kind	必填	"dingtalk"（钉钉）、"feishu"（飞书）、"wecom"（企业微信）、"webhook"（Webhook）。不提供邮件	写其他值报格式错误
target	必填	机器人或 Webhook 的 http:// / https:// 地址	地址不合法时报「地址不合法」；地址合法但发不出去时安装失败（「运维通知测试失败」）

[[hosts]] 主机清单

单机安装不要写这一段。写了 [[hosts]] 即进入多机安装，每一行包含 address（地址）、role（control_plane 控制面、gateway 网关节点、postgres 平台数据库、clickhouse 分析数据库、arbiter 仲裁）与 site（站点）。多机安装只能使用现有 S3 兼容存储，写 bundled 或 local_dir 会报「多副本部署必须使用现有的 S3 兼容存储，不能使用自带对象存储或本机目录」。多机安装不在本手册范围。

注意：答案文件不含任何口令、密钥和许可证，可以随交接材料归档。安装完成后，本次实际使用的答案保存在 /opt/zenova/state/answers.toml。

5.3 凭据文件

凭据文件也是 TOML 格式，只接受下表五个键，按所选依赖填写需要的几项。

键	何时需要	内容
pg_admin_password	[postgres] 为 external	user 所写超级用户账号的密码。只在安装过程中用来创建平台的角色与库，不写入任何文件与日志
ch_password	[clickhouse] 为 external	user 所写账号的登录密码
s3_access_key_id	[object_store] 为 external_s3	访问密钥 ID
s3_access_key_secret	[object_store] 为 external_s3	访问密钥

键	何时需要	内容
master_key	仅在已有平台数据上安装（接管）时	原部署的主密钥。单机首装不要写

操作步骤

1. 在安装暂存目录创建只有 root 可读写的空文件。不要用 echo 写入口令，避免口令进入命令历史。

```
# 安装主机 · root · 写入（创建空的凭据文件，权限 0600）
cd /root/zenova-install
umask 077
touch secrets.toml
chmod 600 secrets.toml
```

2. 用文本编辑器打开 secrets.toml，按需要写入，例如：

```
pg_admin_password = "<PostgreSQL 超级用户密码>"
ch_password = "<ClickHouse 账号密码>"
s3_access_key_id = "<访问密钥 ID>"
s3_access_key_secret = "<访问密钥>"
```

3. 核对权限与属主。

```
# 安装主机 · root · 只读
ls -l /root/zenova-install/secrets.toml
```

你会看到

- 权限列为 -rw-------，属主为 root。

注意：缺少某项依赖所需的凭据时，安装器提示「使用现有服务时，非交互安装必须用 --secret-file 提供凭据」并以依赖检测未通过停止，不会改为在终端询问。权限不对时报「凭据文件权限必须是 0600：<路径>」。安装完成后按第 8 章 8.3 节删除该文件，或按客户凭据管理要求移交。

5.4 示例一：三项依赖全部由安装器部署

适用场景：POC 推荐组合。平台数据库、分析数据库、对象存储全部由安装器在本机部署，不需要凭据文件。

操作步骤

1. 在安装暂存目录写入答案文件 answers.toml：

```
# 顶层字段写在所有段之前
public_url = "http://<主机地址>"
# 暂不配置运维通知；要配置时删掉这一行，改写下面的 [notify] 段
notify_skipped = true

[postgres]
managed = {}

[clickhouse]
managed = {}

[object_store]
bundled = {}

# [notify]
# kind = "dingtalk"
# target = "<机器人地址>"
```

2. 预检（不写入任何内容）：

```
# 安装主机 · root · 只读（预检，不写任何内容）
cd /root/zenova-install
sudo ./zenova install ./<离线包>.zpkg --answers answers.toml --with core,semantic_cache --check-only </dev/null
```

3. 预检通过后正式安装：

```
# 安装主机 · root · 写入（开始安装，不再提问）
cd /root/zenova-install
sudo ./zenova install ./<离线包>.zpkg --answers answers.toml --with core,semantic_cache </dev/null
```

`</dev/null` 把标准输入接到空设备。非交互安装正常情况下不会提问；万一某一步仍需要输入，安装器立即读到输入结束，按输入 `q` 处理并退出（退出码 130），而不会在无人值守的会话里一直卡住等待。远程会话或自动化工具调用时建议始终加上。

你会看到

- 预检输出「预检通过：未写入任何内容」。
- 正式安装依次完成数据层、控制面、已选能力、装后验收，干净主机上约 4 分钟。
- 最后打印各阶段耗时、验收结果、`root` 账号与初始密码、「本机直接访问」与「访问地址」。

注意：`--with` 要与本次离线包所属产品匹配。Zentrix 网关不提供托管智能体与文档解析，写了会报「…不含「…」能力，请从 --with 中去掉」。能力标识为 `core`、`knowledge_base`、`agent_runtime`、`doc_parse`、`semantic_cache`、`ha_data`。

5.5 示例二：使用现有数据库与对象存储

适用场景：客户要求平台数据库、分析数据库与对象存储都使用现有服务，且数据库连接需要加密。

前提：现有服务已满足第 1 章 1.5 节的全部要求（PostgreSQL 超级用户与扩展、ClickHouse 版本与宏、预先建好的存储桶与读写删权限）；如使用内部 CA 签发的证书，已把 CA 文件放到安装主机上。

操作步骤

1. 把客户提供的 CA 证书放到安装暂存目录，例如 /root/zenova-install/customer-ca.pem（必须用绝对路径引用）。
2. 写入答案文件 answers.toml：

```
public_url = "https://<访问域名>"
business_listen_port = 80

[postgres]
# 证书由公共 CA 签发时可写 tls = "required"
external = { host = "<PostgreSQL 地址>", port = 5432, user = "<超级用户账号>", tls = { verify_full = { ca_file =
"/root/zenova-install/customer-ca.pem" } } }

[clickhouse]
external = { host = "<ClickHouse 地址>", port = 8443, database = "zenova", user = "default", tls = { verify_full = {
ca_file = "/root/zenova-install/customer-ca.pem" } } }

[object_store]
external_s3 = { endpoint = "https://<对象存储地址>", region = "<区域>", bucket = "<桶名>", addressing = "path" }

[notify]
kind = "wecom"
target = "<企业微信机器人地址>"
```

[postgres] 中不写 database。现有 ClickHouse 为复制库时，在 external 内加 replicated = true。

3. 按 5.3 节创建凭据文件 secrets.toml（0600），写入：

```
pg_admin_password = "<PostgreSQL 超级用户密码>"
ch_password = "<ClickHouse 账号密码>"
s3_access_key_id = "<访问密钥 ID>"
s3_access_key_secret = "<访问密钥>"
```

4. 预检，再正式安装：

```
# 安装主机 · root · 只读（预检，不写任何内容）
cd /root/zenova-install
sudo ./zenova install ./<离线包>.zpkg --answers answers.toml --secret-file secrets.toml --with core,semantic_cache -
-check-only </dev/null
```

```
# 安装主机 · root · 写入（开始安装，不再提问）
cd /root/zenova-install
sudo ./zenova install ./<离线包>.zpkg --answers answers.toml --secret-file secrets.toml --with core,semantic_cache
</dev/null
```

加密方式写法对照

场景	tls 写法
内网明文连接	省略 `tls`
服务端证书由系统信任的 CA 签发	`tls = "required"`
服务端证书由客户内部 CA 签发	`tls = { verify_full = { ca_file = "<CA 文件绝对路径>" } }`
对象存储加密	不写 `tls`，把 `endpoint` 写成 `https://` 开头

你会看到

- 依赖检测逐项通过：PostgreSQL 账号可登录、为超级用户、扩展齐全；ClickHouse 版本、宏与建库权限满足；对象存储读、写、删正常。检测会创建并删除临时库或临时对象，不写入业务数据。
- 通知测试显示「测试消息已发送，请确认已收到」，到企业微信群确认已收到。
- 其后与示例一相同。

注意：

- 依赖检测不通过时，屏幕显示「依赖检测未通过：<原因>，请联系<负责人>」，例如需要数据库管理员或存储管理员处理。修正后重跑同一条命令。
- 现有 PostgreSQL 中已有本平台数据时，普通安装会停止并提示改用接管方式（不在本手册范围）。
- 本例的访问地址为 `https://`，证书由客户前置的负载均衡或反向代理终结，`business_listen_port` 写的是 TLS 入口转发到本机的端口。本机不提供 HTTPS。
- 使用现有 PostgreSQL 与现有 S3 时，`zenova backup` 不包含这两部分数据，由对应管理员负责；该环境也不支持清空重装。

5.6 从下载站预填答案改为非交互

适用场景：离线包是在下载站填写了「预填配置」后下载的，包内已带一份答案，现在要改为非交互执行。

下载站的预填配置会随离线包带到现场。交互安装时，安装器把这些预填项列出来请你确认：一般项回车确认，访问地址类（现有服务地址、通知渠道地址）标注「来自定制站，未签名」，逐项输入 `y` 才算确认。

改为非交互时：

操作步骤

1. 按 5.2 节写一份完整的 `answers.toml`。可以参照下载页「配置链接」中的同一份预填内容（以实际页面为准），但必须补齐全部必填段——`--answers` 是整份替换，离线包里预填过、而答案文件里没写的项目不会被沿用，会报「缺少 <项目>」。
2. 需要凭据时按 5.3 节准备凭据文件。下载站不收集凭据，这一步不能省。
3. 按 5.4 或 5.5 节的命令执行，带上 `--answers`。

你会看到

- 安装器不再显示「确认预填配置」这一步，也不要求逐项输入 `y`，直接进入主机检查。

注意：`--answers` 的优先级最高，给出后离线包内的预填配置不再参与本次安装。答案文件由实施工程师提供，视为已确认，因此地址类字段务必在写入前与客户核对一遍（可对照第 1 章 1.6 节的确认表）。

能力说明

不提供的能力	替代做法
在答案文件里写口令、密钥或许可证	口令与访问密钥写在 0600 的凭据文件中，用 `--secret-file` 传入；许可证在安装后由租户管理员在控制台导入（第 6 章 6.4 节）
非交互时遇到阻断项、依赖检测失败或通知测试失败后自动跳过	按屏幕提示修正主机、依赖或通知渠道，或修改答案文件后重跑同一条命令；暂时不配置通知时改写 `notify_skipped = true`
`--answers` 与离线包预填配置自动合并	在答案文件里写全所有必填段
续装时直接改用另一份答案或另一组能力	加 `--reconfigure` 回到问答重新确认（需要有人在终端前），或保持与第一次完全一致的答案与 `--with`
答案文件中为现有 PostgreSQL 指定库名	平台库由安装器用超级用户创建、库名固定；只需给出地址、端口、超级用户账号与加密方式

6. 首次登录与初始化

本章内容：安装完成只代表服务已经跑起来。本章带读者完成从第一次登录到平台可以交给业务方使用之间的全部动作：修改 root 初始密码并绑定两步验证、核对平台状态、创建首个租户、离线导入许可证、补配运维通知，最后把第 1 章确认表里的计划值与现场实际值逐项核对。读者是实施工程师；6.4 节许可证导入需要客户方的租户管理员一起完成。

6.1 修改初始密码与绑定两步验证

前提：手中有安装完成时打印的访问地址、「本机直接访问」地址和 root 初始密码（32 位随机字符）。

操作步骤

1. 在能访问安装主机的办公终端上，用浏览器打开「访问地址」。打不开时改用「本机直接访问」地址：本机直接访问能打开而访问地址打不开，问题在前置代理、NAT 或负载均衡；两者都打不开，问题在本机，按第 12 章故障排查处理。
2. 账号填 `root`，密码填初始密码，登录。
3. 页面进入「修改初始密码」。按页面列出的口令规则输入新密码，点「修改并继续」。
4. 页面进入「绑定两步验证」，选择一种方式：
   - 验证器应用：用手机上的身份验证器应用扫描页面二维码，输入应用中显示的 6 位验证码完成绑定。
   - 通行密钥：按浏览器提示用本机的安全密钥或系统生物识别完成注册。
5. 使用验证器应用时，页面会给出一组恢复码。把恢复码抄写或打印后离线保存，确认保存后点「我已保存」。

6. 进入平台管理控制台，即为完成。

你会看到

- 第 3 步提示「当前密码为初始密码或已被重置，请先修改后再进入」。
- 第 4 步页面标题为「绑定两步验证」，说明「所属范围要求使用两步验证，请选择一种方式完成绑定后再进入」。
- 访问地址以 `http://` 开头时，第 4 步没有通行密钥选项，只能用验证器应用。
- 以后每次登录，在密码之后要求输入验证器中的 6 位验证码；验证器不在手边时可选「使用恢复码」。

注意：

- 修改密码与绑定两步验证各有 10 分钟时限，超时后重新登录、从头再来即可，不会留下半完成的状态。
- 修改密码之后，安装主机上的 `sudo zenova root-password show` 不再能查看密码，`/opt/zenova/.env.runtime/root-password-initial` 中的初始密码同时失效。
- 恢复码每个只能使用一次，与验证器应用分开保管。不要截图存放在同一台手机上，也不要通过即时通讯、邮件或工单传递新密码与恢复码。
- root 忘记密码，或验证器与恢复码同时丢失时，在安装主机上执行 `sudo zenova root-password set`。该命令生成新的随机密码，清除两步验证与恢复码，使现有会话全部失效，并写入审计记录；之后用新密码重新执行本节步骤。

6.2 核对平台状态

适用场景：完成 6.1 节后、创建租户之前，确认安装结果在控制台里的状态与安装器输出一致。

操作步骤

1. 打开「网关与网络 › 网关 › 节点」，找到中心网关节点 `office-primary`。
2. 打开「系统维护 › 部署配置」，查看「最近一次验收」。
3. 在同一页核对部署配置各分组：「访问地址」「对象存储」「运维通知渠道」「入口监听」「数据依赖」。
4. 打开「系统维护 › 系统更新」，核对当前版本与各能力状态。

你会看到

位置	核对什么	正常状态
网关节点 `office-primary`	节点在线、版本与本次安装一致	在线。中心网关节点随控制面一起安装，单机上不需要另装节点
最近一次验收	与安装器打印的装后验收结果一致；各检查项状态为「通过」「未通过」「已跳过」	阻断级检查项全部「通过」。未配置通知时，「真实出网」「运维渠道绑定」同样显示通过；启用托管智能体时，「智能体容器可达性」可能提示检查项尚未实现，属于警告级，不影响使用
部署配置	业务访问地址、入口端口与绑定网卡、对象存储方式、数据依赖的部署方式与第 1 章确认表一致	由安装器部署的依赖显示「由本平台部署」，使用现有服务的显示「使用现有服务」

位置	核对什么	正常状态
系统更新	当前版本号、产品、本机形态、已启用与未启用的能力	版本与离线包文件名中的版本一致。未导入许可证前，受许可证控制的能力显示「许可证未包含此能力」，导入后恢复（见 6.4 节）

注意：

- 「最近一次验收」页提供「立即验收」按钮，处理完问题后可以在控制台重新触发一次，不必重跑安装命令。
- 还没有租户和模型路由时，中心网关节点数据面的就绪检查返回 503、原因为 `no_tenant_routes`，这是正常现象，创建租户并配置模型后恢复。
- 控制台的状态与命令行一致。需要留存时，在安装主机上执行 `sudo zenova status`。

6.3 创建首个租户

适用场景：平台状态核对无误后，为客户创建业务使用的租户及其默认管理员。

前提：已与客户确认租户名称、组织标识、默认管理员人选，以及结算货币。

操作步骤

1. 以 root 登录平台管理控制台，打开「租户管理」，点「创建租户」。
2. 填写四项内容：

字段	填写要求
租户名称	企业或部门的展示名称，例如客户公司简称
组织标识	仅小写字母、数字和连字符，用作身份源的组织标识，例如 `<客户简称>`
默认管理员	管理员姓名、管理员邮箱、管理员密码。该账号即本租户的租户管理员，6.4 节由他导入许可证
结算货币	本租户的售价、额度与账单都用这个货币

3. 点「创建」。
4. 把默认管理员的邮箱和初始密码通过客户认可的安全方式交给租户管理员本人，不要经即时通讯、邮件或工单明文传递。

你会看到

- 提示「<租户名称> 创建成功（管理员: <邮箱>）」，租户出现在列表中，状态为活跃。

注意：

- 结算货币创建后不能修改。 创建前必须与客户确认；选错后无法在页面上更正。结算货币与平台记账货币不同时，页面会要求填写汇总汇率。

- root 只用于平台级设置（网关节点、部署配置、系统更新、租户管理等），不要作为日常业务账号使用。模型供应商、模型服务、成员与 API 密钥等业务配置由租户管理员在租户内完成。
- 租户管理员第一次登录时，如页面要求修改初始密码或绑定两步验证，按 6.1 节同样的方式完成。

6.4 导入许可证

适用场景：交付范围包含受许可证控制的能力时。许可证由租户管理员导入，root 在平台管理控制台看不到「授权许可」页签。

前提：6.3 节已创建租户；租户管理员已完成首次登录；客户方已确定把文件送出与取回所用的介质或通道（例如经审批的移动存储介质、内网文件交换系统）。

整个过程分三步：导出授权请求 → 送给供应商换取签名后的 `.lic` 文件 → 导入。全程不需要安装主机访问外网，也不需要访问数据库。

操作步骤

1. 导出授权请求。租户管理员在办公终端上登录控制台，打开「系统设置 › 全局设置 ›「授权许可」页签」，点「导出授权请求 (.json)」。浏览器把请求文件下载到办公终端。
2. 在办公终端上计算请求文件的摘要并记录，便于对方核对文件在传递中没有被改动：

```
# 办公终端 · 普通用户 · 只读（Linux / macOS；Windows 可用 certutil -hashfile <文件> SHA256）
sha256sum <授权请求文件>.json
```

3. 把请求文件送出。按客户的数据外发流程，把请求文件和摘要拷到可以联网的机器上，再通过与供应商约定的渠道发送。请求文件只包含租户与本次部署的绑定信息，不含口令、密钥或业务数据。
4. 换取签名后的许可证。供应商返回一个 `.lic` 文件，同时请对方给出该文件的 SHA-256 摘要。
5. 把 `.lic` 文件取回。按客户的数据导入流程，把 `.lic` 文件拷回能访问控制台的办公终端，计算摘要并与供应商给出的摘要比对：

```
# 办公终端 · 普通用户 · 只读
sha256sum <许可证文件>.lic
```

6. 导入。租户管理员回到「系统设置 › 全局设置 ›「授权许可」页签」，点「导入 .lic 文件」，选择该文件。
7. 打开「系统维护 › 系统更新」（或请 root 查看），确认此前显示「许可证未包含此能力」的能力已恢复可用。

你会看到

- 导入成功提示「许可导入成功」，校验通过后立即生效，无需重启任何服务。
- 「授权许可」页签的许可来源由「社区版（未导入许可）」变为「已导入商业许可」，并显示授权对象、许可 ID 与导入时间。
- 导入失败时提示「许可导入失败」，常见原因是文件在传递中被改动，或许可证与本租户、本次部署不匹配；核对摘要后重新取回，或重新导出授权请求后再次申请。

注意：

- 许可证与租户和本次部署绑定。在另一个租户导入、或在另一套部署上导入都不会通过。

- 清空重装（第 11 章）会生成新的部署标识，原许可证随之失效，需要按本节重新导出请求、重新申请。
- 页面上的「移除许可」会让部署回到未导入许可证的状态，不要误点。
- 未被许可证覆盖的能力在「系统更新」页显示「许可证未包含此能力」。此时核对的是许可证内容，不是安装问题，不要为此重装。

6.5 补配运维通知

适用场景：安装时运维通知选了「暂不配置」（或答案文件写了 `notify_skipped = true`），现在补上。

前提：客户已建好接收告警的钉钉、飞书或企业微信群机器人，或提供了 Webhook 接收地址；安装主机能访问该地址。

操作步骤

1. 以 root 登录，打开「系统维护 › 部署配置」，找到「运维通知渠道」分组。
2. 点「修改」，填写渠道地址并保存。
3. 发送一条测试消息。
4. 请客户方告警接收人到对应群或接收端确认已收到测试消息，并在交付记录里登记接收人姓名与确认时间。
5. 在同一页「最近一次验收」点「立即验收」，确认「真实出网」「运维渠道绑定」两项为「通过」。

你会看到

- 保存后提示「已保存，相关服务将自动重新加载」，随后提示修改已生效。
- 测试消息送达时提示「测试消息已送达」；失败时提示「测试消息投递失败」。
- 重新验收后，两项检查显示「通过」。

注意：

- 不提供邮件作为运维通知渠道。
- 测试消息投递失败时，先在安装主机上用 `sudo zenova doctor --target <渠道地址>` 分层检查 DNS、TCP、TLS、HTTP 连通性；报告注明「直连，可能与生产出口不一致」时，结果不代表网关出口的连通性。
- 只有「发送成功」不够，必须由接收人确认收到，才能证明告警真的能送到人。
- 各渠道在页面上具体显示哪些字段、测试按钮的位置，以实际页面为准。

6.6 初始化完成后的核对表

把第 1 章 1.6 节确认表中的计划值与现场实际值逐项核对。不一致的项目要么修正现场，要么经客户确认后修改计划值，并记入交付记录。

项目	计划值（1.6 节）	现场实际值	在哪里看
产品、版本、通道			「系统维护 › 系统更新」
能力清单			「系统维护 › 系统更新」
主机名、IP、系统版本、架构			安装主机 `hostnamectl`、`uname -m`

项目	计划值（1.6 节）	现场实际值	在哪里看
安装目录与可用空间			安装主机 `df -h /opt/zenova`（或实际安装目录）
访问地址、本机监听端口			「部署配置 › 访问地址」「入口监听」
管理访问地址、管理入口端口（分离时）			「部署配置 › 入口监听」
证书与 TLS 终结位置			客户负载均衡或反向代理配置
平台数据库部署方式与地址			「部署配置 › 数据依赖」
分析数据库部署方式与地址			「部署配置 › 数据依赖」
对象存储方式、地址、区域、桶、寻址方式			「部署配置 › 对象存储」
运维通知渠道与接收人确认			「部署配置 › 运维通知渠道」，6.5 节交付记录
时间源与时钟偏差			安装主机 `chronyc tracking`
网关节点 `office-primary`	在线		「网关与网络 › 网关 › 节点」
最近一次验收	阻断项全部通过		「部署配置 › 最近一次验收」
root 已改密并绑定两步验证、恢复码保管位置	已完成		6.1 节
首个租户名称、组织标识、结算货币			「租户管理」
许可证导入时间、许可来源	已导入商业许可		租户管理员「系统设置 › 全局设置 ›「授权许可」页签」
密钥材料归档保管位置、保管人、摘要			第 8 章 8.1 节登记
首次平台备份标识与转存位置			`sudo zenova backup list`
实施负责人、客户负责人、维护窗口			交付记录

```
# 安装主机 · root · 只读（生成核对时可附上的状态快照）
sudo zenova status
sudo zenova backup list
chronyc tracking
```

注意：这张表核对完成、客户签字后，与第 1 章确认表、答案文件副本（不含口令）一起归入交接材料。之后按第 7 章完成模型供应商接入与真实调用的验收。

能力说明

不提供的能力	替代做法
root 跳过修改初始密码或两步验证直接进入控制台	按 6.1 节在 10 分钟内完成；访问地址为 http:// 时用验证器应用绑定
root 在平台管理控制台导入许可证	由租户管理员在「系统设置 › 全局设置 ›「授权许可」页签」导入
安装主机或控制台在线申请、自动获取许可证	导出授权请求文件，经客户数据外发流程送给供应商，换回 .lic 文件后导入
修改租户的结算货币	创建前与客户确认，并把确认结果记入交付记录
清空重装后沿用原许可证	重新导出授权请求并重新申请
以邮件作为运维通知渠道	使用钉钉、飞书、企业微信或 Webhook

7. 安装验收

本章内容：本章回答一个问题——装完的环境到底能不能用。验收分四层：主机侧服务正常；就绪状态读得懂；一次真实业务调用从调用方经网关到模型再回来，并能在控制台按请求标识查到；拒绝路径按预期拒绝且错误码可解读。四层都完成并填好 7.6 节的记录表，才能把环境交给客户开始 POC。读者是实施工程师与客户方验收人。

安装器自带的「装后验收」（第 4.6 节）只证明平台基础服务可用，不代表模型、授权与业务链路已经打通。本章的验收以「装后验收结果 + zenova status + 真实业务调用」三者为准，不以直接访问某个内部端口返回什么为准。

7.1 主机侧检查

适用场景：安装完成后第一次验收；此后每次升级、修改部署配置、重启主机之后也照做一遍。

前提：以 root 登录安装主机；安装已显示完成且装后验收无阻断项。

操作步骤

1. 查看各服务状态与当前版本。

```
# 安装主机 · root · 只读
sudo zenova status
```

2. 查看不以 zenova- 开头的基础服务。zenova status 只列 zenova- 开头的服务，下面这些要用 systemctl 单独看（未由安装器部署的项会显示单元不存在，属正常）。

```
# 安装主机 · root · 只读
systemctl status postgres clickhouse caddy secret-service spicedb --no-pager
systemctl status rustfs --no-pager      # 仅对象存储选「自带对象存储」时存在
```

3. 执行全量体检并保存报告，随后查看退出码。

```
# 安装主机 · root · 写入（仅在 /root/zenova-install 下生成报告文件）
sudo zenova doctor --all --save /root/zenova-install/doctor
echo "doctor 退出码：$?"
```

4. 把报告文件与 `zenova status` 的输出一起归入交付材料。

```
# 安装主机 · root · 写入（仅在 /root/zenova-install 下生成文件）
sudo zenova status --json > /root/zenova-install/status.json
ls -l /root/zenova-install/doctor* /root/zenova-install/status.json
```

你会看到

- `zenova status`：列出各 `zenova-` 服务的运行状态、当前版本、各实例，以及进行中与最近一次升级。控制面、后台任务、运行时存储这类分槽服务会显示当前接流的槽位（如 `zenova-control-plane@a`）。各列的具体名称以实际输出为准。
- systemctl：每个已部署的单元显示 `active (running)`，且为开机自启（`enabled`）。
- `zenova doctor --all`：列出全部检查项，包括已通过与跳过的项；未通过的项带稳定的诊断代码（如 `HOST_…`、`SVC_…`）和处理建议。`--save` 后面给的路径没有扩展名时，报告自动保存为 `.txt`。
- 退出码：没有阻断项为 `0`；有阻断项为 `2`。

doctor 结果	含义	处理
退出码 0，无告警	本机与连通性全部正常	进入 7.2 节
退出码 0，有告警项	不影响使用，但需记录	在 7.6 节记录告警项与诊断代码，确认是否属于本次范围
退出码 2	有阻断项	按报告中的诊断代码与处理建议处理后重跑；代码含义查产品文档站「安装与运维诊断代码」页

注意：`zenova start/stop/restart` 只接受 `zenova-` 开头的服务名；`postgres`、`clickhouse`、`caddy`、`secret-service`、`spicedb` 不经 `zenova` 命令启停。服务未运行时先看日志（第 8.5 节），不要反复重启。

注意：需要排查到某个模型或工具地址的连通性时，加 `--target <URL>`，按 DNS → TCP → TLS → HTTP 分层检查。报告注明「直连，可能与生产出口不一致」时，结果不代表网关出口的连通性。

7.2 端口与就绪状态的含义

适用场景：验收过程中看到 503、看到某列显示「未核验」、或在 `zenova status` 里找不到某个服务时，先对照本表再决定要不要处理。

操作步骤

1. 遇到下表中的现象，按「含义」列判断是否属于正常状态。
2. 不属于正常状态的，回到 7.1 节用 `zenova status` 与 `zenova doctor` 定位。

现象	含义
中心网关节点数据面（3080）的就绪检查返回 503，原因 `no_tenant_routes`	正常。还没有租户或还没有配置模型路由。创建租户（第 6.3 节）并完成 7.3 节的模型配置后恢复
控制面切换槽位后几秒内，内部入口返回 503	正常。入口代理会暂存新请求，最长 5 秒
「模型供应商」列表的「模型真伪」列显示「未核验」	与健康状态无关。离线环境无法完成核验时保持「未核验」
`zenova status` 里看不到 postgres、clickhouse、caddy	正常。它只列 `zenova-` 开头的服务，其余用 `systemctl status <名称>` 查看
「装后验收」中「智能体容器可达性」显示「检查项尚未实现」	正常。该项是警告级，当前版本未实现，不影响安装完成
「网关节点健康」一项等待较久	正常。该项最多等 150 秒

监听端口的边界：

端口	监听范围	是否允许对外开放
业务入口端口（默认 80）、443	全部网卡（或安装时指定的绑定网卡）	允许，安装器在 Linux 上已自动放行
分离出的管理入口端口	同上	按客户管理网段要求自行放行，安装器不放行
3010、3080、3300、3320、3325	全部网卡，供本机智能体容器回连	不得对外开放
其余内部服务端口	仅回环地址	不涉及

你会看到：装后验收中的「监听边界」一项已通过；从业务网络只能访问到业务入口（与分离后的管理入口）。

注意：不要用「直接 curl 某个内部端口」来判断平台能不能用。内部端口的返回受租户、路由、槽位切换等状态影响，503 往往是正常状态。能不能用，以装后验收结果、`zenova status` 与 7.3 节的真实业务调用为准。

注意：3010、3080、3300、3320、3325 监听全部网卡是为了让本机容器回连，安装器不会替你在网络边界上屏蔽它们。请在主机防火墙或云安全组上只对外放行入口端口。

7.3 真实业务调用

适用场景：每个 POC 环境交付前必须完成一次；这是证明「装完能用」的唯一直接依据。

前提：已完成第 6 章（改密、两步验证、创建租户、导入许可证）；手上有一个可用的上游模型地址与凭据；租户内已有一名测试成员账号（在租户的「组织与权限」中创建）。模型供应商、模型服务、可用范围由租户管理员操作，平台 root 没有租户归属，不能试跑。

操作步骤

1. 添加模型供应商。租户管理员进入「管理控制台 → 搭建 → 模型供应商」，选「添加模型供应商」，选择供应商类型、核对「请求地址」与「模型请求格式」、填写凭据，点「验证并获取可用模型」，在模型表中「启用」需要的模型，点「保存模型供应商」。
2. 添加模型服务。进入「管理控制台 → 搭建 → 模型服务」，选「添加模型服务」，填写名称，「添加模型」加入刚启用的模型，点「保存模型服务」。
3. 试跑。打开该服务详情的「试跑」页签，选模型、输入一句话发送，确认返回成功摘要。
4. 授权测试成员。在同一服务的「可用范围与配额」页签选「编辑可用范围」，勾选测试成员，有效期选「长期有效」或覆盖 POC 周期，保存。
5. 测试成员自己创建 API 密钥。测试成员登录，进入「能力中心 → 首页 → API 密钥」，选「创建 key」，填写名称、选择上一步授权的模型服务与计费归属，「确认」后立即复制「API 密钥（仅本次可见）」。管理员不在控制台为成员创建密钥。
6. 在业务网络主机上发起真实调用。

```
# 业务网络中的主机 · 普通用户 · 写入（产生一次真实调用，会计费）
export API_KEY='<测试成员的 API 密钥>'
curl -sS -i "<访问地址>/api/v1/agent-gateway/openai/v1/chat/completions" \
  -H "Authorization: Bearer $API_KEY" \
  -H "Content-Type: application/json" \
  -d '{"model":"<模型名>","messages":[{"role":"user","content":"你好"}],"max_tokens":32}'
```

7. 记下响应头中的 `x-request-id`，管理员进入「管理控制台 → 概览」的「网关请求」页签，在搜索框按该请求 ID 搜索，打开详情核对。
8. 执行拒绝路径，见 7.4 节。

三种协议的 Base URL（`<访问地址>` 为安装时填写的访问地址）：

协议	Base URL	鉴权头
OpenAI	`<访问地址>/api/v1/agent-gateway/openai/v1`	`Authorization: Bearer <API 密钥>`
Anthropic	`<访问地址>/api/v1/agent-gateway/anthropic`	`x-api-key: <API 密钥>`，也接受 `Authorization: Bearer`
Gemini	`<访问地址>/api/v1/agent-gateway/gemini`	`x-goog-api-key: <API 密钥>` 或查询参数 `?key=`，也接受 `Authorization: Bearer`

你会看到

- 第 1 步：供应商列表中该供应商的「可用」开关为开启，「可用模型」列显示已启用的模型；保存后提示「已保存，模型路由稍后同步」属正常。
- 第 3 步：试跑成功时显示形如「通 · 首字 … · 共 … · … tok · …」的摘要。失败时页面会写明失败发生在本平台侧还是供应商侧，以及按错误码给出的中文原因。
- 第 6 步：预期结果为 `HTTP/1.1 200`，响应体含模型回复；响应头带 `x-request-id`。

- 第 7 步：「网关请求」中能查到同一次调用，模型、结果、耗时与 Token 用量与本次调用一致。
- 业务方需要其他协议时，用上表的 Base URL 各调一次，预期结果相同。

注意：请求里的 model 必须是该密钥所绑定模型服务中的模型名（模型广场或 GET <OpenAI Base URL>/models 返回的名称），否则返回 404，见 7.4 节。

注意：试跑与真实调用都会真实计费。最近的调用在用量页显示「待计费」，属计费短暂延迟。

注意：交付范围包含 MCP 服务或 API 服务时，再补一次「取一次工具列表 + 一次无副作用的工具调用」，并在 7.6 节单列一行记录。

注意：API 密钥只显示一次，复制后保存在测试机的密钥管理中，不要贴进工单、即时通讯或截图。

7.4 拒绝路径与错误码核对

适用场景：7.3 节正常调用完成之后立即进行。网关「该拒绝的能拒绝、拒绝时说得清原因」与「能调通」同等重要。

操作步骤

1. 用不在可用范围内的模型名调用一次：沿用 7.3 节的命令，把 <模型名> 换成一个不属于该密钥所绑定模型服务的模型名。
2. 用已禁用的密钥调用一次：租户管理员在「管理控制台 → 搭建 → 模型服务」的「API 密钥管理」页签找到测试密钥，打开详情点「禁用」（也可由测试成员在「能力中心 → 首页 → API 密钥」中「撤销」一把专门用于测试的密钥），然后用这把密钥再调用一次。
3. （可选）把测试成员移出可用范围后调用一次：在「可用范围与配额」中「移除」该成员，再用其有效密钥调用。
4. 每次都记录：HTTP 状态码、响应体中的 reason 与 code、响应头 x-request-id 与 x-should-retry。
5. 对能关联到成员的拒绝（如第 1、3 步），按 x-request-id 在「网关请求」中查到对应记录，核对结果为失败、原因与响应一致。

```
# 业务网络中的主机 · 普通用户 · 只读（只看状态行与错误体关键字段）
curl -sS -i "<访问地址>/api/v1/agent-gateway/openai/v1/chat/completions" \
  -H "Authorization: Bearer $API_KEY" \
  -H "Content-Type: application/json" \
  -d '{"model":"<不在该模型服务中的模型名>","messages":[{"role":"user","content":"你好"}],"max_tokens":8}'
```

OpenAI 协议的错误体形如 {"error":{"message","type","param","code","reason","request_id"}}，其中 reason 是稳定错误码（大写），code 是 OpenAI 兼容码（小写），request_id 与响应头 x-request-id 相同。Anthropic 协议错误体为 {"type":"error","error":{"type","message","reason"},"request_id"}；Gemini 协议错误体为 {"error":{"code","message","status","details"}}，code 与 HTTP 状态相同。

你会看到（以产品文档站「错误参考」页为准核对）

场景	HTTP 状态	reason	code（OpenAI 协议）	预期处理
模型名不在密钥绑定的模型服务中	404	LLM_MODEL_NOT_AVAILABLE	model_not_found	修改请求中的模型名

场景	HTTP 状态	reason	code（OpenAI 协议）	预期处理
密钥已禁用、已撤销、无效或缺失	401	常见为 UNAUTHENTICATED；401 组内的具体 reason 以实际输出为准	invalid_api_key	检查密钥状态、到期时间与鉴权头
已认证但不在可用范围或授权已过期	403	PERMISSION_DENIED	permission_denied	管理员核对可用范围与授权有效期
内容被护栏拦截	403	CONTENT_POLICY_BLOCKED	content_policy_violation	调整内容后重试
额度池用尽	429	POOL_EXHAUSTED	insufficient_quota	错误信息说明对象、周期与恢复时间；管理员调整额度
命中流量规则限速或并发上限	429	TRAFFIC_RATE_LIMITED、TRAFFIC_CONCURRENCY_LIMITED	rate_limit_exceeded	稍后重试；仍持续时管理员检查流量规则
上游限流	429	LLM_RATE_LIMITED	rate_limit_exceeded	稍后重试
租户本月 Token 上限用尽	402	MONTHLY_TOKEN_EXCEEDED	insufficient_quota	平台管理员在「租户管理」中「编辑配额」
模型路由未配置	503	MODEL_ROUTE_NOT_CONFIGURED	model_route_unavailable	管理员检查模型供应商与模型服务配置
流量保护熔断打开	503	TRAFFIC_CIRCUIT_OPEN	circuit_open	稍后重试；持续时管理员检查流量规则与上游
上游过载或不可用	503	LLM_OVERLOADED、LLM_PROVIDER_UNAVAILABLE	server_is_overloaded、upstream_unavailable	稍后重试
上游服务错误	502	LLM_UPSTREAM_SERVER_ERROR	upstream_server_error	稍后重试

本次验收必须出现的是前两行：第 1 步预期 404 LLM_MODEL_NOT_AVAILABLE，第 2 步预期 401（code 为 invalid_api_key）。其余行是 POC 期间可能遇到的状态，列出供现场判断。

注意：同一个 HTTP 状态码下，不同 reason 的处置完全不同。 例如 403 既可能是可用范围问题（找管理员授权），也可能是护栏拦截（改内容）；429 既可能是额度池用尽（重试无用，需调额度），也可能是上游限流（稍后重试即可）；503 可能是熔断，也可能是上游过载或模型路由未配置。判断一律看 reason，不看状态码、不解析 message 文本（message 会随版本调整，多数为英文）。

注意：是否值得重试看响应头 x-should-retry：为 true 时可按 Retry-After 等待后重试；为 false 时重试不会改变结果。

注意：401 表示网关没有认出调用方，这类请求不一定能在「网关请求」中按发起人检索到；401 场景以响应本身（状态码、reason、x-request-id）为准记录，控制台能否查到以实际为准。404、403 等已认证的拒绝应能按请求 ID 查到。

注意：「禁用」的密钥不能在控制台重新启用。不要禁用 7.3 节交付给客户使用的那把密钥，拒绝路径请用专门创建的测试密钥。

7.5 可选能力验收

适用场景：本次安装启用了核心之外的能力。只验收本次已采购并启用的能力；产品本身不含、未采购或未启用的能力，在 7.6 节记为「不适用」。

前提：先在控制台「系统维护 › 系统更新」核对能力状态。能力显示「本机已有，未启用」表示未启用；显示「许可证未包含此能力」表示许可证未覆盖。这两种情况都记为「不适用」，不要为了验收临时启用。

操作步骤

按下表逐项核对。

能力	Zentrix 网关	完整版	验收方法	预期结果
语义缓存	可选（默认预选）	可选	① 在「路由策略 → 流量规则」中给一条不指定转发目标的规则添加「语义缓存」动作；② 用同一句话、OpenAI Chat Completions 单轮请求调用两次；③ 打开「流量规则 → 缓存」页签	「缓存概况」出现该模型的条目；「命中占比」大于 0；第二次调用在「网关请求」详情的上游各段显示「命中语义缓存，未请求上游」
知识库	可选（需许可证包含知识库）	可选	① 管理员在「知识库管理」中创建知识库并设置分发目标；② 维护者上传一份文档，等待状态变为「可用」；③ 在库详情「检索」页签输入文档中的关键词	能命中刚上传的内容并查看原文片段；顶部页签出现「知识库」
文档解析	不适用（网关版不含）	可选	在知识库中上传 PDF 或 Office 文档	文档状态变为「可用」，检索能命中文档正文；解析失败的文档标注原因

能力	Zentrix 网关	完整版	验收方法	预期结果
托管智能体	不适用（网关版不含）	可选	创建一个智能体、发布，在对话中调用一次	智能体正常回复；具体步骤以完整版用户手册为准
数据高可用	多机方案使用	多机方案使用	单机 POC 不涉及	单机 POC 记为「不适用」

你会看到

- 已启用的能力在「系统更新」的能力列表中显示为已启用。
- 语义缓存未配置缓存服务时，「缓存」页签提示「本部署未配置语义缓存服务，无法清除」，说明该能力未启用。
- 知识库：许可证不含知识库或部署未安装时，能力中心的「知识库」页签和管理控制台的「知识库管理」菜单都不显示。

注意：语义缓存只对 OpenAI Chat Completions 格式、只有一条用户消息（可另带系统或开发者消息）、请求体不超过 256 KiB 的单轮请求生效。规则链中带转发目标、应用限额、单次请求 Token 上限、挂载插件、响应改写、输出检测或改写请求头时，这次请求不查也不写缓存。验收时缓存不命中，先核对规则是否满足这些条件。语义缓存属于动态路由许可模块，许可证未包含时记为「不适用」。

注意：知识库更换嵌入模型会触发全部文档重建向量索引，重建期间语义检索暂停。验收前先确定嵌入模型。

注意：启用托管智能体或文档解析需要容器运行时，且服务用户加入 `docker` 组等同于授予宿主机 root 权限，启用前须取得客户安全负责人的书面同意（第 3.4 节）。

7.6 验收记录表

适用场景：验收全程边做边填，完成后由实施方与客户方共同签字，随交接材料归档。

操作步骤

1. 先填写环境信息。
2. 逐行执行并填写「结果」：`完成`、`未完成`（附原因与诊断代码或 reason）、`不适用`。
3. 请求类条目在「结果」中写明 `x-request-id`。

环境信息：

项目	填写
产品	Zentrix 网关 / 完整版
版本号（完整）	
平台与部署形态	
已启用能力	
访问地址	

项目	填写
安装根	默认 /opt/zenova
数据依赖部署方式（PG / CH / 对象存储）	
许可证 License ID 与到期时间	

验收项：

验收项	方法	结果	记录人	时间
装后验收	安装完成输出与「部署配置 › 最近一次验收」无阻断项			
服务状态	sudo zenova status，全部 zenova- 服务运行中			
基础服务	systemctl status 查看 postgres、clickhouse、caddy、secret-service、spicedb（及 rustfs）			
本机体检	sudo zenova doctor --all --save …，退出码 0；告警项已记录			
监听边界	3010/3080/3300/3320/3325 未对外开放；仅入口端口对外			
网关节点	控制台中 office-primary 在线			
模型供应商	「验证并获取可用模型」成功，已启用模型			
模型服务试跑	「试跑」返回成功摘要			
可用范围	测试成员已加入「可用范围与配额」			
成员自建密钥	测试成员在「能力中心 → 首页 → API 密钥」创建成功			

验收项	方法	结果	记录人	时间
真实调用（OpenAI）	curl 返回 200 + 模型回复；记录 `x-request-id`			
真实调用（Anthropic）	按需；返回 200			
真实调用（Gemini）	按需；返回 200			
调用记录	「网关请求」按请求 ID 查到同一次调用，模型/结果/耗时/用量一致			
拒绝：模型不在范围	404 `LLM_MODEL_NOT_AVAILABLE`；控制台可查到			
拒绝：密钥已禁用	401，`code` 为 `invalid_api_key`；记录实际 `reason`			
拒绝：成员移出范围（可选）	403 `PERMISSION_DENIED`			
工具调用（含 MCP / API 服务时）	取一次工具列表 + 一次无副作用工具调用			
语义缓存	「缓存」页签命中占比大于 0			
知识库	上传文档状态「可用」，检索命中			
文档解析	网关版记「不适用」			
托管智能体	网关版记「不适用」			
数据高可用	单机 POC 记「不适用」			
运维通知	通知渠道测试消息已收到（未配置时记「未配置」）			
密钥材料离机备份	第 8.1 节完成并登记			
首次平台备份	第 8.2 节完成，记录备份标识与转存位置			

签字：

角色	姓名	签字	日期
实施工程师			
客户方系统管理员			

注意：表中出现「未完成」的条目，必须写明原因、诊断代码或 reason、责任方与计划完成时间，不得留空。依赖客户外部条件（如上游模型出网、许可证尚未签发）的条目，写明条件后记为「未完成」，不要记为「不适用」。

能力说明

不提供的能力	替代做法
一条命令完成端到端业务验收	装后验收与 zenova doctor 只覆盖平台与本机；模型、授权、调用与拒绝路径按 7.3、7.4 节在控制台与调用方手动完成
管理员在控制台替成员创建 API 密钥	由测试成员登录后在「能力中心 → 首页 → API 密钥」自己创建；管理员只能检索、轮换、禁用、删除
用内部端口的返回判断平台是否可用	以装后验收、zenova status 与真实业务调用为准（7.2 节）
被禁用的密钥在控制台重新启用	新建一把密钥替换；拒绝路径使用专门的测试密钥
平台 root 账号直接试跑模型服务	root 没有租户归属；以租户管理员身份试跑
「智能体容器可达性」装后验收	当前版本显示「检查项尚未实现」，属警告级；按 7.5 节用控制台功能核对对应能力

8. 安装后必做事项与日常运维

本章内容：安装完成不等于交付完成。本章前三节是装完当天必须做完的三件事——离机备份密钥材料、做首次平台备份并转存、清理临时文件；后四节是日常运维要用到的命令、日志查看、部署配置修改，以及 root 密码与管理面网段的应急恢复。读者是实施工程师与接手运维的客户方系统管理员。本章所有命令都在安装主机上以 root 执行，不依赖控制台是否可用。

8.1 离机备份密钥材料

适用场景：安装完成当天必须做；此后更换主密钥或密钥服务材料发生变化时重做。

前提：安装已完成；已与客户确定一个与安装主机隔离的受控保管位置（不同主机、不同存储，最好不同机房），以及保管人。

平台备份（8.2 节）包含 /opt/zenova/.env（含主密钥），但不包含密钥服务材料 /etc/zenova/secret-service/。密钥服务的 keys/ 用于加密平台保存的全部凭据（模型供应商密钥、对象存储密钥等）。缺了它，即使有平台备份，也无法在新主机上恢复出可用的平台。

操作步骤

1. 打包密钥材料并计算摘要。

```
# 安装主机 · root · 写入（在 /root 下生成密钥材料归档）
umask 077
tar -C / -czf /root/zenova-keys-$(hostname)-$(date +%Y%m%d).tar.gz \
  etc/zenova/secret-service opt/zenova/.env
sha256sum /root/zenova-keys-*.tar.gz
```

2. 按客户规定的受控方式，把归档转存到与安装主机隔离的保管位置。
3. 在保管位置重新计算 SHA-256，与第 1 步的输出逐字核对一致。
4. 填写下面的登记表，随交接材料归档（登记表本身不含密钥内容）。
5. 确认复制成功、摘要一致后，删除本机上的归档（见 8.3 节）。

登记项	填写
归档文件名	zenova-keys-<主机名>-<日期>.tar.gz
SHA-256 摘要	
保管位置	
保管人	
备份日期	
本机归档删除日期	

你会看到

- /root 下生成 zenova-keys-<主机名>-<日期>.tar.gz，权限只有 root 可读（由 umask 077 保证）。
- sha256sum 输出一行 64 位十六进制摘要加文件名。

注意：高危。该归档包含主密钥与密钥服务的全部加密密钥。泄露等同于泄露平台保存的全部凭据；丢失则无法从备份恢复平台。 按客户最高密级数据管理：不得通过即时通讯、邮件、工单或网盘分享链接传递；不得与平台备份文件（8.2 节）放在同一位置；保管人变更时在登记表上留痕。

注意：使用自定义安装根（--install-dir）时，命令中的 opt/zenova/.env 换成 <安装根>/.env（去掉开头的 /）。

8.2 首次平台备份与转存

适用场景：安装与第 7 章验收完成后立即做一次；此后每次升级前、每次重要配置变更前都做一次。

前提：安装主机上有足够的可用空间，见本节「注意」。

操作步骤

1. 立即备份，列出备份确认标识，再把它导出为单个归档文件。

```
# 安装主机 · root · 写入（生成本机备份与导出归档）
sudo zenova backup
sudo zenova backup list
sudo zenova backup export <备份标识> --output /root/zenova-install/first-install.zbackup
```

2. 计算导出文件摘要，转存到与安装主机不同故障域的受控位置（不同主机、不同存储设备，最好不同机房），在目标位置核对摘要一致。

```
# 安装主机 · root · 只读
ls -l /root/zenova-install/first-install.zbackup
sha256sum /root/zenova-install/first-install.zbackup
```

3. 在交接材料中记录：备份标识、导出文件名、摘要、转存位置、保管人、日期。

你会看到

- `zenova backup list` 列出本机备份，最新的在前。备份标识形如 `zb-20261020T020000Z-1a2b3c4d`（`zb-` + UTC 时间 + 8 位随机串）。
- 本机备份保存在 `/opt/zenova/data/backups/<备份标识>/`。
- 导出文件权限为 `0600`。

备份范围：

对象	是否纳入
由安装器部署的 PostgreSQL	纳入
现有 PostgreSQL	不纳入，由客户数据库管理员按其机制备份
ClickHouse（安装器部署与现有服务）	纳入，经 HTTP 接口导出
自带对象存储 / 本机目录	纳入
现有 S3 兼容存储	不纳入，由客户存储管理员负责
安装状态（`.env`、`.env.runtime`、`etc`、`state/install.json`）	纳入
密钥服务材料 `/etc/zenova/secret-service/`	不纳入，按 8.1 节单独备份

注意：导出文件未压缩、不另加密，且含主密钥。 按最高密级保管，与 8.1 节的密钥材料归档分开存放，不要放在同一台主机或同一个存储桶。

注意：所需空间约等于「本机 PG 数据目录 + 本机对象数据目录 + 1 GiB」。本机备份与导出文件各占一份，导出前确认 `/opt/zenova` 与 `/root/zenova-install` 所在磁盘都有余量。

注意：本机备份放在 `/opt/zenova/data` 下，与平台数据同盘。只留本机备份不算备份——主机或磁盘故障时会一起丢失；执行清空重装（第 11 章）时也会被一并删除。一定要转存。

注意：恢复要求版本、产品、部署形态、平台都相同，旧版本的备份不能恢复到新版本的安装。升级后要重新做一次备份并转存。

8.3 清理临时文件

适用场景：8.1、8.2 节完成之后；交付验收结束时再做最后一次。

前提：密钥材料归档与平台备份导出文件均已转存，并在目标位置核对摘要一致。

操作步骤

1. 按下表处置每一类文件。

文件	位置（以实际暂存目录为准）	处置
凭据文件 secrets.toml	/root/zenova-install/secrets.toml	安装完成后立即删除，或按客户凭据管理要求移交
答案文件 answers.toml	/root/zenova-install/answers.toml	保留一份副本随交接材料归档（不含口令）
密钥材料归档	/root/zenova-keys-<主机名>-<日期>.tar.gz	转存并核对摘要后删除本机文件
平台备份导出文件	/root/zenova-install/first-install.zbackup	转存并核对摘要后，按客户要求删除本机副本
安装器与离线包	/root/zenova-install/	保留到交付验收结束；网关节点安装、重新配置、升级排障时还要用
体检与状态报告	/root/zenova-install/doctor*、status.json	随交接材料归档

2. 删除凭据文件与已转存的归档。

```
# 安装主机 · root · 删除（凭据文件与已转存的密钥材料归档、备份导出文件）
rm -f /root/zenova-install/secrets.toml
rm -f /root/zenova-keys-*.tar.gz
rm -f /root/zenova-install/first-install.zbackup
ls -la /root/zenova-install/ /root/
```

你会看到：ls 的输出中不再有 secrets.toml、zenova-keys-*.tar.gz 与 first-install.zbackup；安装器、离线包与 answers.toml 仍在。

注意：只删除上表列出的文件。不要删除 /opt/zenova/state/answers.toml（安装记录，升级与重新配置会用到）、/opt/zenova/.env.runtime/ 下的任何文件；初始密码文件 root-password-initial 在修改 root 密码后自动失效，不需要手工删除。

注意：不要在转存与摘要核对完成之前删除密钥材料归档。删除后如发现转存的副本损坏，回到 8.1 节重新打包。

8.4 日常运维命令速查

适用场景：日常巡检、故障处理、变更窗口。所有命令以 root 在安装主机执行，不依赖控制台。

操作步骤

按「什么时候用」选择命令。

命令	作用	什么时候用
zenova status [--json]	各服务状态、当前版本、各实例、进行中与最近一次升级	每日巡检；控制台打不开时第一步；任何变更前后
zenova doctor [--target <URL>] [--json]	全量本机与连通性体检，输出诊断代码与处理建议；--all 显示全部项，--save 保存报告	服务异常但原因不明；向支持方反馈前；排查到某个模型或工具地址的连通性（--target）
zenova start <服务>	启动一个 zenova- 开头的服务	服务被手工停止后恢复
zenova stop <服务>	停止一个 zenova- 开头的服务	维护窗口内需要停服；恢复备份前停止控制面组
zenova restart <服务>	重启一个 zenova- 开头的服务；控制面、后台任务、运行时存储自动解析当前槽位	zenova status 显示某服务未运行，且已看过日志
zenova upgrade <包.zpkg>	升级（用新版本的安装器执行）	维护窗口内升级，与控制台「系统更新」是同一流程（第 10 章）
zenova rollback	回到上一次升级之前的版本	升级后业务异常且数据库迁移尚未提交（第 10 章）
zenova backup	立即备份	升级前、重要配置变更前、定期备份
zenova backup list	列出本机备份，最新在前	查备份标识
zenova backup export <备份标识> --output <文件>	导出为 .zbackup 归档	转存到其他故障域
zenova backup import <文件>	导入 .zbackup 并登记为本机备份	在新主机或重装后的主机上准备恢复
zenova backup delete <备份标识>	删除一个本机备份	本机空间不足，且该备份已转存
zenova restore <备份标识或归档>	从备份恢复，需输入 RESTORE 确认	数据损坏或误操作后恢复；会覆盖当前数据
zenova root-password show	查看初始 root 密码	首次修改密码之前忘记初始密码（8.7 节）
zenova root-password set	重置 root	root 忘记密码或两步验证全部丢失（8.7 节）
zenova infra upgrade <包>	升级由安装器部署的基础设施	升级前检查提示基础设施版本不满足时
zenova config set <键> [--value-stdin]	修改可在线变更的部署配置（共 18 项）	无法使用控制台，或需要脚本化修改时（8.6 节）
zenova logs <组件> [--since] [--follow]	查看组件日志	服务异常时（8.5 节）

命令	作用	什么时候用
zenova admin-network show/allow/disable	管理面访问网段	网段白名单配错导致进不了控制台（8.7 节）

你会看到：每条命令都在终端直接给出结果；zenova doctor 没有阻断项时退出码为 0，有阻断项时为 2。

注意：zenova start/stop/restart 只接受 zenova- 开头的服务名，例如 sudo zenova restart zenova-edge。postgres、clickhouse、caddy、secret-service、spicedb 用 systemctl 管理。

注意：恢复（zenova restore）前，先对当前环境再做一次备份；恢复条件是版本、产品、部署形态、平台都相同，没有升级在进行，控制面组服务已停止。恢复后、启动控制面组之前，先重启 spicedb、secret-service、rustfs（未用自带对象存储时去掉 rustfs）。

注意：严禁用手写的 rm -rf、docker rm、systemctl disable 等命令清理或"重装"平台。需要清空时只走第 11 章的清空重装流程。

8.5 日志怎么查

适用场景：zenova status 显示服务异常、安装或升级失败、需要向支持方提供日志。

操作步骤

1. 按组件查看日志。zenova logs 会解析出该组件当前对应的 systemd 单元（控制面会自动定位到当前接流槽位），再转调 journalctl。

```
# 安装主机 · root · 只读
sudo zenova logs control-plane --since '30 min ago'
sudo zenova logs edge --follow
```

2. 不在组件列表里的服务，直接用 journalctl 按单元名查看。

```
# 安装主机 · root · 只读
journalctl -u zenova-worker@a.service --since '30 min ago' --no-pager
journalctl -u spicedb.service --since '30 min ago' --no-pager
```

3. 安装、续装、重新配置的过程日志看 install.log。

```
# 安装主机 · root · 只读
tail -n 200 /opt/zenova/state/install.log
```

4. 远端网关节点的日志在节点主机上查看。

```
# 节点主机 · root · 只读
sudo zenova logs node-edge --since '30 min ago'
sudo zenova logs node-caddy --since '30 min ago'
```

5. 向支持方反馈时，按下面的命令一次采集齐全。

```
# 安装主机 · root · 写入（仅在 /root/zenova-install 下生成文件）
sudo zenova status --json > /root/zenova-install/status.json
sudo zenova doctor --all --save /root/zenova-install/doctor
sudo zenova logs control-plane --since '30 min ago' > /root/zenova-install/control-plane.log
sudo cp /opt/zenova/state/install.log /root/zenova-install/
```

组件与单元名对照（Linux，`zenova logs` 可用的组件共 8 个）：

`zenova logs` 组件	说明	对应 systemd 单元	在哪台主机上查
`control-plane`	控制面	`zenova-control-plane@<当前槽位>.service`	安装主机
`edge`	网关数据面（中心网关节点）	`zenova-edge.service`	安装主机
`caddy`	入口（Caddy）	`caddy.service`	安装主机
`postgres`	PostgreSQL	`postgres.service`	安装主机（仅由安装器部署时）
`clickhouse`	分析数据库（ClickHouse）	`clickhouse.service`	安装主机（仅由安装器部署时）
`secret-service`	密钥服务	`secret-service.service`	安装主机
`node-edge`	网关节点数据面	`zenova-node-edge.service`	远端节点主机
`node-caddy`	网关节点入口（Caddy）	`zenova-node-caddy.service`	远端节点主机

其他服务单元（用 `journalctl -u <单元名>` 查看）：`zenova-worker@<槽位>`、`zenova-updater`、`spicedb`、`rustfs`，以及能力组件如 `qdrant`、`docling`、`zenova-runtime-store@<槽位>`。实际存在哪些单元以本机安装的能力为准。

你会看到

- `zenova logs` 输出的就是 `journalctl -u <单元>` 的内容；不带 `--follow` 时一次输出完毕，带 `--follow` 时持续跟随，按 `Ctrl+C` 退出。
- 组件名写错时，提示「未知的日志组件 …」并列出全部可用组件。

注意：Linux 上的服务日志由 journald 保存，`--since` 原样传给 `journalctl --since`，可写 `'30 min ago'`、`'2026-10-20 09:00'` 等 journalctl 接受的格式。不存在 `/opt/zenova/<服务>/logs/` 这类目录，不要按这种路径去找。

注意：日志里可能出现请求内容与内部地址。外发给支持方前，按客户数据管理要求审阅；不要外发 `.env`、密钥材料与初始密码文件。

8.6 修改部署配置

适用场景：访问地址变更、对象存储参数调整、运维通知渠道更换、调整资源容量、修改入口端口等安装后的配置变更。

前提：没有升级正在进行；上一次修改已生效完毕。变更前先按 8.2 节做一次备份。

操作步骤

1. 首选控制台。平台管理员进入「系统维护 › 部署配置」，在目标项选「修改」，填写新值并保存。页面分为七个分组：

分组	可修改的内容
访问地址	访问域名（管理访问地址可用 ADMIN_PUBLIC_URL 修改）
对象存储	服务地址、区域、存储桶、寻址方式、访问密钥，以及报文归档、审计归档存储桶（对象存储方式不能改）
运维通知渠道	一般告警、严重告警与心跳通知地址
链路导出	外部 OTLP 接收端地址（只填地址不会开启导出）
资源容量	控制面数据库最大连接数（12～200）
入口监听	业务入口与管理入口的端口和绑定网卡
数据依赖	查看平台数据库、分析数据库、对象存储的部署方式与管理方

2. 控制台不可用或需要脚本化时，用命令行。普通项直接写值：

```
# 安装主机 · root · 写入（修改访问地址，相关服务会重新加载或重启）
sudo zenova config set PUBLIC_URL <新访问地址>
echo "退出码：$?"
```

3. 敏感项不能写在命令行上，按提示在终端输入（不回显），或用 --value-stdin 从标准输入读取：

```
# 安装主机 · root · 写入（修改对象存储访问密钥，按提示输入，不回显）
sudo zenova config set LP_OBJECT_STORE_ACCESS_KEY_SECRET
```

4. 修改完成后，sudo zenova status 核对服务状态，并按第 7.3 节做一次真实业务调用。

可用 zenova config set 修改的 18 个键：

键	对应内容
PUBLIC_URL	访问地址
ADMIN_PUBLIC_URL	管理访问地址
ZENOVA_OBJECT_STORE_MODE	对象存储方式（不能在线更换，见「注意」）
LP_OBJECT_STORE_ENDPOINT	对象存储服务地址
LP_OBJECT_STORE_REGION	对象存储区域
LP_OBJECT_STORE_BUCKET	平台数据存储桶

键	对应内容
LP_OBJECT_STORE_ADDRESSING	寻址方式
LP_OBJECT_STORE_ACCESS_KEY_ID	对象存储访问密钥 ID
LP_OBJECT_STORE_ACCESS_KEY_SECRET	对象存储访问密钥（敏感项）
LP_PAYLOAD_STORE_BUCKET	报文归档存储桶
LP_PAYLOAD_STORE_ACCESS_KEY_ID	报文归档访问密钥 ID
LP_PAYLOAD_STORE_ACCESS_KEY_SECRET	报文归档访问密钥（敏感项）
LP_AUDIT_ARCHIVE_STORE_BUCKET	审计归档存储桶
OTEL_EXPORTER_OTLP_ENDPOINT	链路导出地址
LP_DB_MAX_CONNECTIONS	控制面数据库最大连接数

键名以命令的实际提示为准：写了不在列表中的键时，命令提示「<键> 不是可修改的部署配置项」，并列出全部可修改的键。

你会看到

- 控制台：提示「已保存，正在让新值生效」，完成后显示已重启的服务数量；敏感项只显示「已设置」或「未设置」，不回显原值。
- 命令行：先输出「已保存，正在让新值生效…」，生效后可能输出「已重启：…」「已更新安装记录：…」，最后一行为「修改已生效」。
- 值为空表示清除该项。

退出码	含义	处理
0	已生效	按步骤 4 核对
1	新值已保存，但有服务没能重启，或控制台没能切换到新配置（仍按原配置运行）	sudo zenova status 查看，按 8.5 节查日志
2	键或值不合法（含敏感项写在命令行上、缺少新值）	按提示修正后重试
3	升级进行中	等升级完成后再修改

注意：对象存储方式安装后不能在线更换。 控制台提示「不支持在线修改：请备份后按新的选择重新安装，再恢复数据」。需要更换时，安排维护窗口按「备份 → 重新安装 → 恢复」处理。

注意：入口端口与绑定网卡只能在控制台「入口监听」中修改，不能用 zenova config set 改。 修改业务入口后，外部调用方与远端网关节点要改用新端口；修改管理入口后，当前页面地址失效，按页面给出的新地址访问。安装器只在首次安装时放行防火墙端口，改端口后由运维方自行调整防火墙、负载均衡与「网关 → 入口与证书」中的公开

访问地址。

注意：修改「资源容量」会让控制面重启，期间控制台暂时不可用，网关请求不受影响。

注意：不要手工编辑 /opt/zenova/.env.runtime/ 与 /opt/zenova/etc/ 下的文件。 它们由安装器与配置编译生成，手改的内容会在下一次修改、升级或重新配置时被覆盖，或导致服务无法启动。/opt/zenova/.env 含主密钥，同样不要手改。

注意：使用现有服务的数据库与对象存储，其凭据轮换由客户管理员负责，「部署配置」页不提供这些依赖的凭据轮换。

8.7 root 密码与两步验证恢复、管理面访问网段应急

适用场景：root 忘记密码、两步验证设备与恢复码全部丢失；或「管理面访问网段」白名单配错，导致无法从客户网络进入控制台。

操作步骤

一、首次改密前查看初始密码

1. 尚未修改初始密码时，可再次查看初始密码：

```
# 安装主机 · root · 只读
sudo zenova root-password show
```

二、重置 root

1. root 忘记密码，或两步验证与恢复码全部丢失时执行：

```
# 安装主机 · root · 写入（重置 root 密码，清除两步验证并使会话失效）
sudo zenova root-password set
```

2. 用终端打印的新密码登录，按页面要求修改密码并重新绑定两步验证（验证器应用或通行密钥），保存新的恢复码。改密与绑定各有 10 分钟时限，超时重新登录。

三、管理面访问网段应急

1. 查看当前白名单配置：

```
# 安装主机 · root · 只读
sudo zenova admin-network show
```

2. 追加一个允许访问管理控制台的网段：

```
# 安装主机 · root · 写入（追加管理面允许网段）
sudo zenova admin-network allow --cidr <网段>
```

3. 或直接关闭白名单：

```
# 安装主机 · root · 写入（关闭管理面网段白名单）
sudo zenova admin-network disable
```

4. 恢复访问后，以 root 打开「平台账号与登录 › 全局设置 ›「账号与会话安全」页签」，在「管理面访问网段」分区按正确的网段重新配置。

你会看到

- root-password show：首次改密前打印初始密码；改密后初始密码失效。
- root-password set：打印新生成的随机密码；root 原有的两步验证与恢复码被清除，root 的全部会话失效，操作写入审计记录。
- admin-network：修改约 10 秒内在所有实例生效。从安装主机本机回环地址访问控制台始终放行。

注意：root-password set 打印的新密码与初始密码一样，不得通过即时通讯、邮件或工单传递，不要录屏或截图外传。执行前确认终端没有被录屏或共享。

注意：http:// 访问地址下登录页不提供通行密钥，两步验证只能用验证器应用。

注意：这两组命令只恢复平台 root。租户管理员丢失两步验证时，由平台管理员在「租户管理」打开该租户详情，在「管理员」页签为其「重置两步验证」。

注意：「仅允许以下网段访问管理控制台」开启前，先把实施方与客户运维所在网段都加入；admin-network disable 只用于应急，恢复后应重新开启并核对网段。

能力说明

不提供的能力	替代做法
平台备份自动包含密钥服务材料	按 8.1 节单独打包 /etc/zenova/secret-service/ 与 .env，离机保管
备份导出文件自动加密或压缩	导出文件未压缩、不加密、含主密钥；转存时使用客户的加密存储或加密介质，按最高密级保管
备份现有 PostgreSQL 与现有 S3	由客户数据库管理员、存储管理员按其备份机制保护
在线更换对象存储方式	备份后按新选择重新安装，再恢复数据
用 zenova config set 修改入口端口与绑定网卡	在控制台「部署配置 › 入口监听」修改，并同步调整防火墙与负载均衡
用 zenova 命令启停 postgres、clickhouse、caddy 等非 zenova- 服务	用 systemctl status/restart <名称>；日志用 zenova logs <组件> 或 journalctl -u <单元>
通过控制台找回 root 密码	在安装主机执行 sudo zenova root-password set

9. 增加远端网关节点

本章内容：业务调用方位于另一个网络区域、控制面不在那里时，可以在该区域增加一台远端网关节点，让模型与工具调用在当地完成鉴权、路由与转发，不必绕回控制面。本章说明何时需要独立节点、开通前要准备什么、如何在控制台创建节点并在节点主机上安装、如何核对上线，以及重装、更换控制面与升级的做法。读者是实施工程师与客户方系统管理员。

9.1 什么时候需要独立节点

适用场景：业务调用需要在另一个网络区域就近接入，而控制面不在那里。

控制面所在的主机本身就是第一个网关节点，名为 office-primary（中心节点）。它随控制面一起安装、一起启动，单机部署时不需要再单独安装节点。只有当调用方与控制面不在同一网络区域、跨区域访问代价高或不被允许时，才需要增加远端节点。

部署方式	适用	代价
单机（只有中心节点 office-primary）	POC、调用方与控制面在同一网络区域	一台主机；主机故障时控制台与调用同时中断
单机 + 远端节点	部分调用方位于另一个网络区域，需要就近接入；或需要一个网络区域在控制面不可达时继续服务	每个节点多一台主机；节点版本必须与控制面一致，升级时多一步「先升节点」；需要为节点地址准备入口证书与 DNS
多机高可用	生产环境要求控制面与数据层无单点	主机数量与运维复杂度最高；需要容器运行时与「数据高可用」能力，按实施方案单独规划，不在本手册范围

远端节点的工作方式：

- 节点自己完成鉴权、路由、限流、护栏检测与计费预授权，请求不绕回控制面。
- 节点主动连接控制面拉取配置、上报心跳；控制面从不主动连接节点。
- 节点与控制面断开时，按最近一次拿到的配置继续服务；请求记录先暂存在节点本地，恢复连接后自动补传。

注意：节点上线后不会自动承接公开流量。需要让公开访问地址解析到该节点时，还要在节点详情中「加入入口池」并调整 DNS 或负载均衡，这一步通过入口自检后才生效。POC 阶段只验证就近调用时，业务方直接使用节点地址即可。

9.2 前提

前提：控制面已按第 4～7 章安装完成并通过验收，中心节点 office-primary 在线。

项目	要求
节点主机	不能是控制面主机。一台主机只允许一套中心安装，中心节点已随控制面安装

项目	要求
安装物料	与控制面同一产品、同一版本的安装器与离线包，按节点主机的系统与架构从下载站下载（节点主机与控制面主机架构不同时，要另下一套）
系统	Linux（systemd），x86_64 或 aarch64
权限	安装与运维需 root
节点 → 控制面	节点主机能通过 HTTPS 访问「控制面回连地址」（页面要求以 `https://` 开头，默认端口 443）。控制面使用自签或内部 CA 时，在创建节点时粘贴 CA 证书
调用方 → 节点	节点入口端口对调用方开放；节点地址需被「入口与证书」里的入口证书覆盖（通配符证书即可）
控制面 → 节点	不需要
安装令牌	一次性使用，有效期 24 小时
时间	偏差 ⩽ 30 秒，与控制面主机要求相同

节点安装器在写入任何文件前，会一次性执行完整的主机检查：CPU 架构、内存、磁盘、服务托管器、可执行权限与安全模块、时钟偏差、监听端口，以及到控制面的连通性。有阻断项时汇总打印诊断代码与处理建议，不写入任何内容。诊断代码中 `NODE` 组为节点相关代码，含义查产品文档站的「安装与运维诊断代码」页。

节点安装目录规则

规则	说明
默认目录	`/opt/zenova-node`；节点程序、数据（`<安装目录>/data`）与配置（`<安装目录>/etc`）都在该目录下，安装包解包暂存也在这里
自定义目录	节点主机 `/opt` 空间不足时，在创建节点时填写空间更大的卷上的目录，生成的命令会带上 `--install-dir <安装目录>`
路径格式	至少两级的绝对路径，只能包含字母、数字、点、下划线、短横线和斜杠
禁止位置	不能位于中心安装目录 `/opt/zenova` 之内，同机已有中心安装时也不能与其目录重叠；不能位于 `/tmp`、`/var/tmp`、`/home`、`/root`、`/run/user`、`/usr`、`/etc`、`/boot`
不符合时	控制台「安装目录」字段标红并说明原因；节点安装器在写入任何文件前拒绝

注意：安装目录首次安装后不能通过升级更换。 创建节点前按节点主机的磁盘规划好目录。

9.3 在控制台创建节点

前提：以 root 登录平台管理控制台。

操作步骤

1. 打开「网关与网络 › 网关 › 节点」，点「添加节点」。
2. 按四步弹窗填写：

步骤	字段	含义
基本信息	节点名	小写字母、数字和连字符，2～31 个字符，创建后不可修改
基本信息	地域	可选，便于识别节点所在区域
基本信息	启用的面	节点承接哪些类型的请求（模型服务为必选）；未启用的面，节点直接拒绝对应请求
基本信息	安装方式	选「Linux（systemd）」
基本信息	安装目录	可选；留空安装到 `/opt/zenova-node`，规则见 9.2 节
网络	节点地址	调用方访问这个节点用的域名或 IP，入口证书需要覆盖它
网络	控制面回连地址	节点主动连接控制面用的地址，以 `https://` 开头，节点主机必须能访问
网络	控制面 CA	控制面使用自签或内部 CA 时粘贴 CA 证书；使用公认 CA 时留空
本地暂存	配额（GB）	与控制面断开期间，节点本地暂存请求记录的空间
本地暂存	暂存满时	「拒绝新请求」：保证每一条请求都可计费、可审计；「丢弃请求内容，只保留计量」：优先保证服务不中断，请求正文不再保存
安装	副本数	节点上运行的实例数量

3. 点「创建并生成安装命令」，立即复制整条命令，保存到受控位置。

你会看到

- 弹窗显示完整的安装命令，其中包含一次性令牌。
- 弹窗进入「上线进度」，状态为等待节点首次上线；节点安装完成后无需刷新页面即可更新。

注意：安装命令只显示一次。关闭弹窗后无法再次查看。令牌过期（24 小时）、丢失或已使用后，在节点详情点「重新签发安装命令」获取新命令。

注意：令牌等同于节点接入凭据。不要通过即时通讯、邮件或工单明文传递，不要截图外传。

9.4 在节点主机安装

前提：节点主机满足 9.2 节要求；已从下载站取得与控制面同产品、同版本、匹配节点主机平台的安装器与离线包，并按第 2 章核对过 SHA-256。

操作步骤

1. 把安装器与离线包放到节点主机同一个暂存目录，解压安装器，并把离线包改名为 `zenova-node.zpkg`（控制台生成的命令使用这个固定文件名）。

```
# 节点主机 · root · 写入（解压安装器、离线包改名）
mkdir -p /root/zenova-install
cd /root/zenova-install
tar xzf <安装器>.tar.gz
mv <离线包>.zpkg zenova-node.zpkg
```

2. 先预检。预检执行主机检查与到控制面的连通性检查，不消耗令牌，不写入任何内容。

```
# 节点主机 · root · 只读（预检，不消耗令牌）
cd /root/zenova-install
sudo ./zenova install node --package zenova-node.zpkg --control-plane <控制面回连地址> --check-only
```

3. 预检通过后，执行控制台给出的安装命令。控制面 CA 相关参数以控制台生成的命令为准，完整复制、不要删改：控制面使用自建 CA 或填写了安装目录时，生成的命令会自动带上对应参数。命令形如：

```
# 节点主机 · root · 写入（安装节点并消耗令牌）
cd /root/zenova-install
sudo ./zenova install node --package zenova-node.zpkg --control-plane <控制面回连地址> --token <令牌>
```

4. 推荐改用 `--token-file`：把令牌写入一个仅 root 可读的文件，在控制台生成的命令中把 `--token <令牌>` 换成 `--token-file <文件>`，其余参数原样保留。`--token-file` 与 `--token` 二选一。

```
# 节点主机 · root · 写入（令牌写入文件后安装节点，消耗令牌）
cd /root/zenova-install
umask 077
printf '%s' '<令牌>' > /root/zenova-install/node.token
sudo ./zenova install node --package zenova-node.zpkg --control-plane <控制面回连地址> --token-file /root/zenova-install/node.token
```

5. 安装成功后删除令牌文件。

```
# 节点主机 · root · 删除（清理令牌文件）
rm -f /root/zenova-install/node.token
```

你会看到

- 预检通过时，屏幕给出通过提示且不写入任何内容；有阻断项时列出诊断代码与处理建议（以实际输出为准）。
- 安装成功时，安装器打印结果；失败时打印原因（例如令牌已过期或已使用、节点包校验失败、系统服务启动失败）并以非零状态退出。

注意：`--token-file` 的好处是令牌不出现在进程列表和 shell 历史中。使用 `--token` 时，同机其他用户在安装期间可能通过进程列表看到令牌。

注意：预检不校验令牌。 令牌只在正式安装时兑换，令牌过期或已被使用的错误只会在这一步出现；预检通过也不代表正式安装一定通过。出现令牌错误时，在节点详情「重新签发安装命令」后用新命令重跑。预检失败时不要直接执行带令牌的命令。

9.5 核对上线

操作步骤

1. 回到控制台「添加节点」弹窗或节点详情，查看「上线进度」。
2. 在节点主机上查看节点状态。

```
# 节点主机 · root · 只读
sudo zenova status
```

3. 在节点列表与节点详情核对节点信息（见下）。
4. 让调用方把 Base URL 指向节点地址，按第 7.3 节的方法完成一次真实调用，并在控制台「网关请求」中按响应头 `x-request-id` 查到这次调用。
5. 上线异常时查看节点日志。

```
# 节点主机 · root · 只读（查看节点日志）
sudo zenova logs node-edge --since '30 min ago'
sudo zenova logs node-caddy --since '30 min ago'
```

你会看到

- 「上线进度」依次显示三个状态：

状态	含义	卡在这一步之前时先查
节点已注册	令牌兑换成功，节点身份已创建	安装命令是否成功结束；屏幕上的失败原因（令牌过期或已使用时重新签发）
收到首次心跳	节点服务已启动并连上控制面	节点主机能否访问控制面回连地址（curl -v <控制面回连地址>）；控制面使用自签证书时是否粘贴了 CA；`node-edge` 日志
配置已同步	节点已加载控制面下发的配置，可以处理请求	节点版本是否与控制面一致；`node-edge` 日志

- `zenova status` 显示节点身份、所连控制面与就绪状态。
- 节点列表中可核对：「状态」「版本」「配置落后版本数」「最近心跳」「5 分钟请求数」「5 分钟错误率」「回传积压」「本地暂存」。
- 真实调用返回 HTTP 200 与模型回复，控制台「网关请求」中能按 `x-request-id` 查到这次调用。

注意：节点与控制面断开 2 分钟内，界面仍显示在线；超过 2 分钟标为「离线」，节点仍按最后一份配置继续服务。

注意：控制面使用自签证书而创建节点时没有粘贴 CA 证书时，节点无法连上控制面。处理办法是在节点详情「编辑网络」补上 CA 证书，再按需重新签发安装命令。

9.6 重装、更换控制面与升级

已有节点再安装

在已安装节点的主机上再次执行安装命令时，安装器提示「本机已存在网关节点安装」并停止，不会覆盖现有节点。

需要让该主机连接到另一个控制面、或改用新令牌重新接入时，先在目标控制面上为它创建节点（或重新签发安装命令），再执行：

```
# 节点主机 · root · 写入（重新接入，消耗新令牌）
cd /root/zenova-install
sudo ./zenova install node --package zenova-node.zpkg --reenroll --control-plane <控制面回连地址> --token <新令牌>
```

注意：不要手工删除节点目录（如 `/opt/zenova-node`）或系统服务单元来"重装"。节点目录里保存着节点身份、本地暂存与尚未回传的请求记录，手工删除会使这些记录永久丢失，且后续安装与升级无法正确识别本机状态。重新接入一律用上面的 `--reenroll`。

升级节点

前提：已取得目标版本、与节点主机平台匹配的安装器与离线包。

操作步骤

1. 先升级全部远端节点，再升级控制面。 控制面的升级前检查会拒绝仍在旧版本运行的在线节点。
2. 在节点主机上放好新版本的安装器，并把新版本离线包改名为 `zenova-node.zpkg`。
3. 在控制台节点详情点「签发升级命令」，复制命令。
4. 用新版本安装器在暂存目录执行该命令，命令带 `--upgrade`，形如：

```
# 节点主机 · root · 写入（升级节点，保留节点身份与数据）
cd /root/zenova-install
sudo ./zenova install node --package zenova-node.zpkg --control-plane <控制面回连地址> --token <新签发的令牌> --upgrade
```

5. 升级完成后按 9.5 节核对节点版本与上线状态，并通过该节点完成一次真实调用。

你会看到

- 节点切换到新版本；节点列表「版本」列与控制面目标版本一致。
- 节点切换到新版本后 60 秒内未就绪，会自动切回旧版本，节点继续以旧版本服务。此时查 `node-edge` 日志定位原因，处理后重新签发升级命令再执行，不要反复执行同一条命令。

注意：升级保留节点身份、暂存与账本数据；升级命令不带 `--install-dir`，沿用首次安装的目录。在升级命令上另指目录会被拒绝。

注意：节点版本低于控制面目标版本期间，不要开始控制面升级。全部节点升级完成后，再按第 10 章升级控制面。

能力说明

不提供的能力	替代做法
在控制面主机上再装一个远端节点	控制面主机已自带中心节点 `office-primary`；远端节点装在另一台主机上
节点上线后自动承接公开访问地址的流量	在节点详情「加入入口池」，通过入口自检后由运维方调整 DNS 或负载均衡
节点使用与控制面不同的产品或版本	按节点主机平台下载同产品、同版本物料；升级时先升节点再升控制面
安装命令关闭后再次查看	在节点详情「重新签发安装命令」
通过升级更换节点安装目录	首次创建节点时规划好安装目录
手工删除目录后重新安装节点	用 `--reenroll` 配合新令牌重新接入

10. 备份、恢复、升级与回滚

本章内容：本章说明平台备份覆盖哪些数据、如何创建与转存备份、如何从备份恢复，以及升级前要准备什么、如何在控制台或命令行升级、升级后出现问题时如何回滚。恢复与回滚都会改变当前运行的数据或版本，执行前务必读完对应小节的「注意」。读者是客户方系统管理员与实施工程师。

10.1 备份范围

`zenova backup` 与控制台「备份与恢复」页使用同一套备份，只备份由安装器部署和管理的内容。

内容	是否纳入	由谁负责
安装器部署的 PostgreSQL	纳入	平台管理员
现有 PostgreSQL	不纳入	客户数据库管理员，按数据库自身的备份机制保护

内容	是否纳入	由谁负责
ClickHouse	纳入（安装器部署与现有服务都经 HTTP 接口导出）	平台管理员
自带对象存储 / 本机目录	纳入	平台管理员
现有 S3	不纳入	客户存储管理员，按存储自身的机制保护
安装状态（`.env`、`.env.runtime`、`etc`、`state/install.json`）	纳入，含主密钥	平台管理员
密钥服务材料 `/etc/zenova/secret-service/`	不纳入	平台管理员按第 8.1 节单独离机备份

备份保存在 `/opt/zenova/data/backups/<备份标识>/`。所需空间估算：

> 所需空间 ≈ 本机 PG 数据目录 + 本机对象数据目录 + 1 GiB

操作步骤

1. 估算一次备份所需空间。

```
# 安装主机 · root · 只读（估算备份所需空间）
du -sh /opt/zenova/data/*
df -h /opt/zenova/data
```

2. 对照 `du` 输出中本机 PostgreSQL 与本机对象存储的数据目录大小（目录名以实际输出为准），加 1 GiB，与 `df` 的剩余空间比较。

注意：两类「不纳入」的数据必须在交付前与客户书面约定由谁负责、多久备份一次，并约定恢复时与平台备份恢复到同一时间点。只恢复平台备份而现有 PG 或 S3 不同步恢复，会出现数据不一致。

注意：密钥服务材料不在平台备份内。在另一台主机上恢复时，没有原主机 `/etc/zenova/secret-service/` 的离机备份，平台保存的全部凭据无法解密。

10.2 创建、导出与转存

操作步骤

1. 创建一次备份，记录屏幕输出的备份标识（形如 `zb-20261020T020000Z-1a2b3c4d`）。

```
# 安装主机 · root · 写入（创建备份）
sudo zenova backup
```

2. 列出本机备份，核对刚才的标识。

```
# 安装主机 · root · 只读
sudo zenova backup list
```

3. 导出为 .zbackup 归档。

```
# 安装主机 · root · 写入（导出归档）
sudo zenova backup export <备份标识> --output /root/zenova-install/<备份标识>.zbackup
sha256sum /root/zenova-install/<备份标识>.zbackup
```

4. 把归档与摘要转存到与安装主机不同故障域的受控位置（不同主机、不同机房或离线介质），登记保管位置、保管人、摘要与日期；确认复制成功后删除本机导出文件。
5. 不再需要的本机备份按标识删除。

```
# 安装主机 · root · 删除（删除一个本机备份，不可找回）
sudo zenova backup delete <备份标识>
```

命令行与控制台对照（控制台为平台管理控制台的「备份与恢复」页，菜单位置以实际界面为准）：

操作	命令行	控制台
立即备份	sudo zenova backup	「立即备份」
列出备份	sudo zenova backup list（最新的在前）	「备份列表」
导出归档	sudo zenova backup export <备份标识> --output <文件>	备份行的「下载」
导入归档	sudo zenova backup import <文件>	「上传备份」
删除备份	sudo zenova backup delete <备份标识>	备份行的「删除」

你会看到

- 命令行备份完成后打印备份标识。
- 控制台点「立即备份」后提示「备份已开始，完成后会出现在列表中。」，任务依次经过「准备」「安装状态」「平台数据库」「分析数据库」「对象存储」「校验」。
- 「备份列表」显示「创建时间」「版本」「发起人」「大小」「包含内容」；未纳入的组件显示为「（未纳入）」及原因。
- 导出文件权限为 0600。

注意：导出文件未压缩、不另加密、含主密钥。拿到它的人可以解开平台保存的凭据。按最高密级保管，不得通过即时通讯、邮件或工单传递。

注意：只保存在安装主机上的备份不算备份。主机磁盘损坏或执行清空重装（会删除全部本机备份）时，本机备份随之丢失。

注意：备份由页面或命令行手动触发，平台不提供定期自动备份与过期清理；定期备份、异地转存和清理由客户按运维流程安排。已有备份、恢复或升级在进行时，新的备份会被阻止。

10.3 恢复

适用场景：数据被误删或损坏，需要把平台回到某次备份时的状态；或在演练主机上验证备份可用。

前提：同时满足以下条件，否则恢复被拒绝。

条件	说明
版本相同	备份所在版本与本机已安装版本相同；版本不同时先把本机安装或升级到备份所在版本
产品、部署形态、平台相同	例如都是 Zentrix 网关、单机原生、Linux x86
没有升级在进行	等待升级结束
控制面组服务已停止	控制面、运行时存储、后台任务

注意：恢复会用备份覆盖当前数据，备份之后写入的全部数据（配置变更、调用记录、用量、审计等）将永久丢失，无法撤销。 恢复前必须先对当前环境再做一次备份并导出转存，作为恢复失败或选错备份时的退路。

注意：现有 PostgreSQL 与现有 S3 不在恢复范围内。使用它们时，由对应管理员按其备份机制把数据恢复到与平台备份相同的时间点，否则平台数据前后不一致。

操作步骤

1. 先为当前环境再做一次备份，并导出转存（见 10.2 节）。

```
# 安装主机 · root · 写入（恢复前的保护性备份）
sudo zenova backup
sudo zenova backup list
sudo zenova backup export <当前备份标识> --output /root/zenova-install/<当前备份标识>.zbackup
```

2. 停止控制面组服务。先用 `zenova status` 查看服务名，未启用的组件跳过。

```
# 安装主机 · root · 写入（停止控制面组，控制台与管理接口不可用）
sudo zenova status
sudo zenova stop zenova-worker
sudo zenova stop zenova-runtime-store
sudo zenova stop zenova-control-plane
```

3. 执行恢复。参数可以是备份标识，也可以是 `.zbackup` 归档文件（先导入再恢复）。提示时输入确认词 `RESTORE`。

```
# 安装主机 · root · 写入（用备份覆盖当前数据，不可撤销）
sudo zenova restore <备份标识或 .zbackup 文件>
```

4. 恢复完成后、启动控制面组之前，重启密钥服务、权限服务与自带对象存储，使它们加载恢复后的数据。未使用自带对象存储时去掉 `rustfs`。

```
# 安装主机 · root · 写入（重启依赖服务，加载恢复后的数据）
systemctl restart spicedb secret-service rustfs
systemctl status spicedb secret-service rustfs --no-pager
```

5. 启动控制面组。

```
# 安装主机 · root · 写入（启动控制面组）
sudo zenova start zenova-control-plane
sudo zenova start zenova-runtime-store
sudo zenova start zenova-worker
sudo zenova status
```

6. 核对：登录控制台；确认网关节点在线；在「部署配置 › 最近一次验收」执行一次验收；按第 7.3 节完成一次真实调用，并核对模型、路由、权限、审计与告警配置与备份时一致。
7. 记录可恢复到的时间点、恢复耗时、缺失内容与需要重建的外部依赖数据。

你会看到

- `zenova restore` 先要求输入 `RESTORE`，输入其他内容时不执行；完成时的输出以实际输出为准。
- 控制台发起恢复时（备份行「恢复」→「从备份恢复」→ 输入 `RESTORE` →「开始恢复」），页面显示「正在恢复」，依次经过「停止控制台服务」、各组件恢复和「启动控制台服务」，完成后提示「恢复已完成。」。
- 控制面仍在运行，或有升级、备份任务进行时，恢复被拒绝。

注意：恢复过程中旧数据在新数据完整就位前保留；但一旦恢复完成，被覆盖的数据只能靠步骤 1 的保护性备份找回。

注意：恢复到另一台主机时，先按相同版本、相同产品与形态完成安装，再把原主机 `/etc/zenova/secret-service/` 的离机备份放回原位，然后执行上述步骤。覆盖生产环境前，先在隔离的演练主机上完成一次恢复演练。

10.4 升级前准备

操作步骤

1. 阅读目标版本的发布说明，确认变更内容、已知问题与是否涉及数据库变更。
2. 确认可以从当前版本直接升级到目标版本。当前版本早于目标版本要求的「可直接升级的最低版本」时，先升级到中间版本，再升级到目标版本。
3. 确认当前已启用的全部能力。目标版本离线包必须包含当前已启用的全部能力，缺任何一项的组件，升级前检查不通过。

```
# 安装主机 · root · 只读（查看当前版本与服务）
sudo zenova status
```

也可在控制台「系统维护 › 系统更新」的「当前版本」核对版本与各能力状态。

4. 按当前产品、平台与已启用的全部能力，从下载站下载目标版本的安装器与离线包，按第 2 章核对 SHA-256。离线环境可在「系统更新」点「复制获取链接」，在联网机器上打开，定制站已预选本机组合与最新版本。
5. 有远端网关节点时，先按第 9.6 节把全部节点升级到目标版本。

6. 完成一次备份，导出并转存到不同故障域，记录备份标识（见 10.2 节）。
7. 与客户约定维护窗口，通知控制台使用者。

注意：升级期间控制台短暂不可用或变慢；网关调用不受影响，控制面切换的几秒内入口代理会暂存新请求，最长 5 秒。维护窗口主要是为控制台使用者与回退操作预留时间。

注意：升级前的备份是新版本数据库变更提交后唯一的退路（见 10.6 节），不能省略。

注意：当前版本启用了「运维观测」时，升级到不再提供该能力的版本前，须先在当前版本的「系统维护 › 系统更新」停用「运维观测」，否则升级前检查会拒绝升级：「部署配置」一项判为未通过，展开「查看技术详情」（命令行 `zenova upgrade` 为「详情：」一行）可见提示「能力『运维观测』已在新版本中退役，请先在当前版本的『系统更新』页停用该能力，再升级」；「能力包齐全」一项同时提示升级包缺少已启用的能力，二者是同一原因，按提示停用即可，不需要另找升级包。升级完成后，可手工删除 `<安装目录>/data/` 下的 `prometheus`、`alertmanager`、`grafana`、`loki`、`tempo` 目录以回收磁盘空间。

10.5 升级

前提：已完成 10.4 节全部准备，且没有其他升级在进行。

在控制台升级

操作步骤

1. 以 root 打开「系统维护 › 系统更新」。
2. 取得升级包：
   - 可访问外网：点「检查更新」，系统按本机当前版本、平台与能力组合获取升级包。
   - 离线环境：点「复制获取链接」→ 在联网机器上打开并下载 → 回到本页上传 `.zpkg`。上传中断后可「继续上传」。
3. 上传完成后系统自动执行升级前检查。逐项核对结果；未通过时按提示处理，处理后点「重新检查」。

检查项	未通过时
许可证	页面只提供「重新检查」；请租户管理员在「系统设置 › 全局设置 ›「授权许可」页签」更新许可证后重新检查
磁盘空间	清理或扩容后重新检查
平台匹配	重新下载本机平台的升级包
部署配置	按提示前往部署配置修正
能力包齐全	重新下载包含当前已启用全部能力的升级包
网关节点已升级	先按第 9.6 节升级仍在旧版本的在线节点
可直接升级的最低版本	先升级到中间版本
分析数据库	未配置分析数据库时按提示执行 `zenova install --reconfigure` 部署或填写现有 ClickHouse

控制台上显示的检查项名称与数量以实际界面为准，可能还包括基础设施版本、数据库变更耗时、版本顺序等项。由安装器部署的 PostgreSQL 或 ClickHouse 版本不满足要求时，在维护窗口执行 `sudo zenova infra upgrade <基础设施包>` 后重新检查；使用现有数据库时由数据库管理员升级。

4. 全部通过后点「升级」，在弹窗中输入 root 密码确认。
5. 观察进度，直到升级结束。不要关闭服务器。
6. 升级完成后：在「升级历史」核对版本、结果、用时与操作人；在「部署配置 › 最近一次验收」点「立即验收」；按第 7.3 节完成一次真实调用。

你会看到

- 升级依次经过九个阶段，不适用的阶段跳过：

序号	阶段
1	准备
2	控制台切换
3	前端
4	内置网关
5	扩展能力
6	智能体
7	网关节点
8	数据整理
9	收尾

- 控制台切换约数秒，期间操作稍慢但不报错；网关调用不受影响。
- 结果为「已完成」「已完成，部分项需处理」「未成功，已恢复到原版本」或「需要人工处理」之一（以实际界面为准）。

在命令行升级

操作步骤

1. 在安装主机暂存目录解压目标版本的安装器，用新版本安装器执行升级。

```
# 安装主机 · root · 写入（升级到目标版本）
cd /root/zenova-install
tar xzf <目标版本安装器>.tar.gz
sudo ./zenova upgrade <目标版本离线包>.zpkg
echo $?
```

2. 按退出码判断结果；完成后同控制台第 6 步核对。

你会看到

- 命令依次完成上传、升级前检查与切换，持续显示进度。与控制台升级是同一个流程、同一条升级记录。
- 退出码：

退出码	含义	处理
0	升级完成	核对版本、执行验收与真实调用
4	完成，但部分项目保持旧版本	按输出说明处理对应项目（例如某个网关节点），处理后再核对
1	升级失败	保留输出，按提示处理；必要时按 10.6 节回滚
2	升级前检查未通过或未开始	按检查结果处理后重新执行，本机未发生变化
3	已有升级在进行	等待当前升级结束，不要重复发起

注意：必须使用新版本的安装器执行 `zenova upgrade`。已安装的旧版本程序可能无法识别新版本的升级包。

注意：同一时间只能有一个升级在进行。控制台与命令行发起的升级互相可见，不要在一处升级进行中从另一处再次发起。

10.6 回滚

适用场景：升级完成后新版本出现无法接受的问题，需要切回升级前的版本。

前提：升级已经结束（无论成功与否），没有升级正在进行。

注意：回滚只切回上一次升级之前的版本（槽位），不是任意版本降级。 不能跨越多次升级回退，也不能回到更早的历史版本。回滚后在新版本上产生的配置变更以实际结果为准，回滚前先做一次备份并导出转存。

注意：新版本的数据库迁移一旦提交，回滚会被拒绝，这一结果不可逆。 此时只有两条路：升级到包含修复的版本，或按 10.3 节恢复升级前的备份（备份之后写入的数据全部丢失）。这也是 10.4 节要求升级前必须完成备份并转存的原因。

操作步骤

1. 确认没有升级正在进行。

```
# 安装主机 · root · 只读
sudo zenova status
```

2. 执行回滚。

```
# 安装主机 · root · 写入（切回上一个版本）
sudo zenova rollback
```

3. 回滚成功后，按第 7.3 节完成一次真实调用，并在「部署配置 › 最近一次验收」执行一次验收。

你会看到

- 成功时显示「已回到版本 <版本>」。
- 以下三种情况回滚被拒绝：

拒绝情况	处理
升级正在进行	等待升级结束后再判断是否需要回滚
新版本的数据库迁移已提交	升级到修复版本，或按 10.3 节恢复升级前备份
上一次升级没有生效，本机仍在旧版本运行	无需回滚，当前已是旧版本

注意：有远端网关节点时，回滚前先与交付方确认节点版本的处理方式，回滚后在节点列表核对各节点状态与版本，以实际输出为准。

能力说明

不提供的能力	替代做法
备份现有 PostgreSQL 与现有 S3	由客户数据库管理员与存储管理员按其备份机制保护，并与平台备份对齐时间点
平台备份包含密钥服务材料	按第 8.1 节单独离机备份 `/etc/zenova/secret-service/`
定期自动备份与过期清理	由客户按运维流程定期执行 `zenova backup`、导出转存并删除过期备份
跨版本、跨产品或跨形态恢复	先把本机安装或升级到备份所在版本与形态，再恢复
回滚到任意历史版本	只能回到上一次升级之前；更早的状态通过恢复对应版本的备份实现
数据库迁移提交后回滚	升级到修复版本，或恢复升级前备份

11. 清空重装与更换产品

本章内容：本章说明什么时候需要删除已有部署重新安装、什么时候不需要，以及确需清空时怎样先预览、再确认、出错后从哪里继续。更换产品（完整版与 Zentrix 网关互换）也只能走这条路，放在本章最后一节。读者是实施工程师与客户方系统管理员；执行本章任何写入命令前，须取得客户书面同意并安排停服窗口。

11.1 三个入口的区别

适用场景：已有一套安装，准备「重新装一下」之前，先判断要的到底是哪一种。

绝大多数情况你要的是前两个：换新版本用升级，改安装时填错的配置用重新填写配置。这两个入口都保留数据。只有确定要丢弃全部数据、从零开始时，才使用删除重装。

目的	入口	对已有数据的处理
升级到新版本	`zenova upgrade <包.zpkg>`	保留数据；同一条升级记录，可回滚到上一次升级之前的版本（见第 10 章）
重新填写安装配置	`zenova install <包.zpkg> --reconfigure`	保留数据；回到「选择能力」一步，带入上次的答案逐项确认
删除旧部署并重新首装	`zenova install <包.zpkg> --fresh`	删除旧部署的业务数据、配置、身份与本机全部备份，再按首装流程安装新包

操作步骤

1. 只是要换版本：用新版本的安装器执行 `sudo zenova upgrade <包.zpkg>`，或在控制台「系统维护 › 系统更新」上传升级包。
2. 只是要改访问地址、通知渠道、资源容量等：先看控制台「系统维护 › 部署配置」或 `zenova config set` 能否在线修改；不能在线修改的，再用 `--reconfigure`。
3. 以上都不能解决、且客户同意丢弃全部数据：继续阅读 11.2 节。

你会看到

- 已完成安装的主机上直接再跑一次普通安装命令时，安装器停止并提示本机已完成安装（诊断代码 `SVC_ALREADY_INSTALLED`），同时给出三条去向：升级用 `sudo zenova upgrade <包.zpkg>`，修改配置用 `sudo zenova install --reconfigure`，清空重装加 `--fresh`。普通安装命令不会覆盖已有安装。

注意：`--fresh` 与 `--reconfigure` 互斥，不能同用。`--adopt-existing`（在已有平台数据上接管安装）也不能与这两者同用。

11.2 适用条件

前提：以下三条须同时满足，缺一条安装器都会拒绝清空。

条件	为什么
单机原生安装（systemd）	清空流程由控制面完整安装命令执行，只处理本机上属于该安装的服务、目录与独占资源。多机部署的数据分布在多台主机上，K8s 部署的资源归集群管理，都不在本机安装器的清理范围内
PostgreSQL 与 ClickHouse 都由安装器部署	由安装器部署的数据库只监听回环、归本安装独占，安装器能证明它们属于这套部署。现有数据库通常与其他业务共用，安装器无法证明可以删，因此始终拒绝清空
对象存储为自带对象存储或本机目录	理由同上。现有 S3 兼容存储的桶与密钥归客户存储管理员，安装器不会删除也不会复用

以下环境不支持清空重装：

- 平台数据库或分析数据库使用现有服务；

- 对象存储使用现有 S3 兼容存储；
- 多机部署；
- Kubernetes Helm 部署；
- 远端网关节点主机（zenova install node 不使用 --fresh，节点换控制面或换令牌用 --reenroll，见第 9 章）。

注意：使用现有数据库或现有 S3 的环境确需重装时，需要客户的数据库管理员与存储管理员先自行处理原有库与桶，再按首装流程安装；这部分不由安装器执行，也不在本手册范围内。本机上的网关节点（含与控制面同机的 office-primary）会随清空一并清理；如果本机节点曾注册到另一个远端控制面，该控制面上节点会变为离线，需在其控制台退役。

11.3 先预览再确认

适用场景：任何一次清空重装的第一步。

前提：安装器、离线包、答案文件与凭据文件已放在旧安装根与业务目录之外的独立目录（如 /root/zenova-install），见 11.4 节。

操作步骤

1. 执行只读预览。

```
# 安装主机 · root · 只读（预览清理范围，不停服务、不删数据、不开始安装）
cd /root/zenova-install
sudo ./zenova install "$ZPKG" --fresh --dry-run
```

2. 逐行核对预览输出，确认列出的就是这台主机上要删除的那套部署。
3. 预览中出现「归属不明」或带风险提示的残留项时，按下面三种方式之一处理，不要手工清理：

情况	处理方式
现场有管理员值守	交互模式：直接运行 sudo ./zenova install "$ZPKG" --fresh，安装器逐项列出待确认资源并询问 y/N，由现场管理员逐项确认
远程支持或需要书面审核	先导出计划文件，审核后按计划执行（命令见下）
外部数据库或对象存储	安装器始终硬拒绝清空外部 PostgreSQL、ClickHouse 与 S3，没有任何参数可以放开

```
# 安装主机 · root · 只读（导出清理计划文件，供审核）
sudo ./zenova install "$ZPKG" --fresh --dry-run --plan ./reinstall-plan.toml
```

```
# 安装主机 · root · 删除（按审核后的计划清空旧部署并重新安装）
sudo ./zenova install "$ZPKG" --fresh --plan ./reinstall-plan.toml --confirm-data-loss DELETE_EXISTING_DATA
```

你会看到

- 预览一次性汇总全部待清理项与拒绝项：旧部署的身份、版本与部署形态，以及服务、业务路径、独占容器、卷和网络的清理范围。
- 遇到以下情况时，预览说明具体对象并停止，处理所列对象后重新预览：

诊断代码	含义	处理
FRESH_EXTERNAL_DATA	检测到外部数据服务，或请求删除外部数据	安装器不会删除外部数据库或对象存储；在计划文件中保持 remove = false
FRESH_SHARED_SECRET	存在与外部共享、不受本安装独占的资源	先解除外部引用或手工迁移该资源，再执行清空重装
FRESH_OWNER_UNPROVEN	部分对象无法自动证明由本安装创建	交互模式逐项 y/N 确认，或导出计划文件审核后按计划执行
FRESH_LEGACY_IGNORED	检测到不由本安装器管理的旧版目录	安装器保持原样，不处理
FRESH_INVALID_ARGS	清空重装的参数不完整或组合无效	改用交互模式 --fresh，或配合 --dry-run / 计划文件执行

注意：严禁手写 rm -rf、docker rm、systemctl 等命令清理旧安装残留。 现场运维与远程支持一律通过 --fresh --dry-run、交互确认或计划文件处理。手工删除会让安装器无法再证明资源归属，后续预览会被拒绝，问题只会更难处理。

注意：计划文件只对生成它时的现场有效。停服后安装器会再次核对旧部署身份，若与计划不一致会以 FRESH_STATE_CHANGED 停止，此时重新执行 --fresh --dry-run 生成新的计划文件。

11.4 执行与中断处理

前提：已完成 11.3 节的预览与核对；客户已书面同意丢弃全部数据；已安排停服窗口。需要保留旧数据时，先执行 zenova backup 并用 zenova backup export 导出，并把导出的归档移到清理范围之外——清空重装不自动备份，且会删除本机全部备份。

操作步骤

1. 准备物料目录。使用与新离线包同版本的独立 ./zenova，不要使用待删除安装根里的 zenova（/usr/local/bin/zenova 指向的就是旧安装）。安装器、.zpkg、--answers 答案文件、--secret-file 凭据文件都必须放在旧安装根（默认 /opt/zenova）与业务目录之外；放在待清理目录内时，安装器会说明并停止。

```
# 安装主机 · root · 只读（确认物料位置与版本）
cd /root/zenova-install
ls -l ./zenova "$ZPKG"
sha256sum "$ZPKG"
```

2. 执行清空重装。确认词必须完整填写 DELETE_EXISTING_DATA。

```
# 安装主机 · root · 删除（清空旧部署后重新安装）
sudo ./zenova install "$ZPKG" --fresh --confirm-data-loss DELETE_EXISTING_DATA
```

3. 安装完成后按第 6 章重新完成首次登录与初始化、重新申请并导入许可证（6.4 节），再重新离机备份密钥材料并做首次平台备份——旧的密钥归档与备份对新部署无效。

你会看到

- 删除任何内容之前，安装器先核对：新包全部组件的 sha256 与大小、安装器与离线包的版本和平台、安装答案、所选能力是否齐全。任一项未通过，不会进入删除阶段，旧部署保持原样。组件摘要用于发现下载损坏，不代表签名校验。
- 通过后依次执行：停止旧服务和容器并再次核对旧部署身份 → 清理属于该安装的服务、数据路径、容器、独占卷和网络（保留 Docker 镜像与系统账户）→ 按首装流程安装新包并完成装后验收。
- 完成输出与首装相同：新的 `root` 初始密码、访问地址与「下一步」。

中断后从哪里继续

中断位置	处理
旧部署已清理成功，后续安装失败	处理安装错误后，对同一个包执行普通安装命令 `sudo ./zenova install "$ZPKG"`，从已有安装进度继续，不再加 `--fresh`
清理阶段失败或中断	执行 `sudo ./zenova install "$ZPKG" --fresh --dry-run`，把输出的清单与诊断代码回传给支持，按提示用计划文件或交互确认继续

注意：在旧部署已清理、新安装未完成的状态下再次带 `--fresh` 并确认删除，表示重新清空本机数据，不是继续上次安装。

注意：清空重装会生成新的部署标识。原许可证、原 `root` 密码、原两步验证绑定、原密钥材料与备份全部失效；原有租户、成员、模型供应商、API 密钥都需要重新创建，调用方需要换用新的 API 密钥。

11.5 更换产品

适用场景：已安装 Zentrix 网关，要改装完整版；或反过来。

前提：两个产品使用同一套安装命令，但产品在安装时锁定，安装后不能原地更换。两者互换只能清空重装，11.2 节的适用条件同样适用。

操作步骤

1. 向客户说明后果并取得书面同意：原有数据全部删除；原许可证失效，需在新产品上重新导出授权请求、重新申请许可证。
2. 从下载站重新选择目标产品，下载目标产品的安装器与离线包（第 2 章）。离线包文件名里的产品字段应与目标产品一致。
3. 用目标产品的安装器按 11.3 节预览、11.4 节执行清空重装。
4. 按第 6 章重新完成初始化与许可证导入，按第 7 章重新验收。

你会看到

- 不做清空、直接用另一个产品的安装器执行普通安装时，安装器停止并提示「本机已安装<已装产品>，该安装程序为<另一产品>版，请使用<已装产品>版的安装包」。这表示产品已锁定：要么换回已装产品的安装包，要么按本节清空重装。
- 输出中带诊断代码 `ANS_EDITION_LOCKED`（版本形态已锁定）时含义相同，处理办法是执行 `zenova install <安装包> --fresh` 后重新选择。
- 安装器与离线包分属不同产品时，提示「安装包与安装程序不属于同一版本：安装程序为…版，安装包为…版」，须换用同一产品的安装器与离线包。

注意：只是想多启用或停用某项能力（如语义缓存、知识库）时，不需要换产品，也不需要清空重装：在控制台「系统更新」页启用已随离线包带来的能力，或用 `--reconfigure` 重新选择能力。目标能力不在本产品中（安装器提示「…不含「…」能力，请从 --with 中去掉」）时，才需要考虑更换产品。

能力说明

不提供的能力	替代做法
清空使用现有数据库或现有 S3 的环境	由客户数据库管理员与存储管理员自行处理原有库与桶，再按首装流程安装
多机部署、Kubernetes 部署、远端网关节点的清空重装	多机与 K8s 按各自部署方案处理；节点使用 `--reenroll` 重新注册或在控制台退役
清空前自动备份	执行前手工 `zenova backup` 并导出到清理范围之外
原地更换产品	清空重装为目标产品，并重新申请许可证
跨版本恢复旧备份到新安装	恢复只能在同一版本、同一产品、同一部署形态与平台上进行
手工清理旧安装残留	只用 `--fresh --dry-run`、交互确认或计划文件处理

12. 故障排查与提交支持

本章内容：本章是出现问题时的处理顺序：先自己确认最常见的三类原因，再一次性采集排查信息，然后按安装阶段、诊断代码、装后访问与调用三种场景对照处理；仍无法解决时，按 12.6 节准备材料提交支持。读者是客户方系统管理员与实施工程师。本章所有采集命令都是只读的，可以放心执行。

12.1 先自己确认三件事

适用场景：调用方反馈「调不通」「被拒绝」「突然不能用了」，但平台服务本身运行正常。

这类反馈里，大部分不是故障，而是平台按配置做出的拒绝。拒绝的结论写在两个地方，不需要查日志：

- 调用方收到的错误响应体里的 `reason` 字段（OpenAI、Anthropic 协议的错误体还带 `request_id`，与响应头 `x-request-id` 相同）；
- 控制台「网关请求」中按 `x-request-id` 查到的这次调用记录。

操作步骤

1. 向调用方要这次调用的 x-request-id，以及响应体中的 reason。
2. 在控制台「网关请求」中按 x-request-id 查到这次调用，核对调用人、模型与结果。
3. 按下表判断属于哪一类，并按「去哪里确认」核对配置。

类别	典型 reason（HTTP 状态）	去哪里确认
权限与范围	MODEL_NOT_AVAILABLE_FOR_WORKSPACE（403）、PERMISSION_DENIED（403）、LLM_MODEL_NOT_AVAILABLE（404）、UNAUTHENTICATED / ACCESS_KEY_REVOKED（401）	模型服务的「可用范围与配额」是否包含该成员；调用使用的模型名是否在该服务内；API 密钥是否为该成员本人创建且未禁用
额度	QUOTA_EXCEEDED（429）、USER_TOKEN_EXCEEDED / USER_COST_EXCEEDED（429）、MODEL_SERVICE_QUOTA_HARD_REJECT（429）、MONTHLY_TOKEN_EXCEEDED / INSUFFICIENT_CREDITS（402）	模型服务「可用范围与配额」中该成员或该服务的额度是否已用完；租户额度是否已用完
规则拦截	CONTENT_POLICY_BLOCKED（403）、DATA_FENCE_BLOCKED（403）、NETWORK_NOT_ALLOWED（403）、TRAFFIC_RATE_LIMITED / TRAFFIC_CONCURRENCY_LIMITED（429）	路由策略中命中的规则（内容安全、数据边界、来源网段、限流与并发）

你会看到

- 属于这三类时，平台服务状态正常，同一成员换一个在范围内、额度未用完、不触发规则的请求可以成功。
- 处理方式是调整配置（加入范围、调整额度、修改规则）或让调用方修改请求，不需要重启任何服务。

注意：这三类拒绝都会留下调用记录。如果「网关请求」里按 x-request-id 完全查不到这次调用，说明请求没有到达网关（地址错、网络不通、前置设备拦截），按 12.5 节排查。

12.2 一次性采集信息

适用场景：12.1 节无法解释的问题，或任何需要提交支持的问题。先采集，再动手处理——重启服务之后，部分现场信息就不在了。

操作步骤

1. 在安装主机上执行下面四条命令。四条都是只读的，不改变任何服务状态。

```
# 安装主机 · root · 只读（采集排查信息，输出写到暂存目录）
sudo zenova status --json > /root/zenova-install/status.json
sudo zenova doctor --all --save /root/zenova-install/doctor
sudo zenova logs control-plane --since '30 min ago' > /root/zenova-install/control-plane.log
sudo cp /opt/zenova/state/install.log /root/zenova-install/
```

2. 问题发生时间早于 30 分钟前时，把 `--since` 改为覆盖问题发生时刻的时间，例如 `'2 hours ago'`。`--since` 的值原样传给 journalctl。
3. 问题涉及某个外部地址（模型供应商、通知渠道、对象存储）的连通性时，补充一次定向体检：

```
# 安装主机 · root · 只读（对指定地址分层检查 DNS、TCP、TLS、HTTP）
sudo zenova doctor --target <目标地址>
```

命令	产出	用途
`zenova status --json`	各服务状态、当前版本、各实例、进行中与最近一次升级	判断哪个服务没在运行、当前是什么版本、是否卡在升级中
`zenova doctor --all --save <目录>`	全量本机与连通性体检报告，含诊断代码与处理建议	一次性列出所有阻断项与警告项；退出码 0 表示没有阻断项，2 表示有阻断项
`zenova logs control-plane --since <时间>`	控制面在指定时间段内的日志	定位控制台报错、配置下发、数据库访问类问题
`cp /opt/zenova/state/install.log`	安装与升级过程日志	定位安装、升级、清空重装阶段的失败

你会看到

- `status.json`、`doctor/` 目录、`control-plane.log`、`install.log` 四份材料出现在 `/root/zenova-install/` 下。
- `zenova doctor --target` 的报告注明「直连，可能与生产出口不一致」时，结果只代表本机直连该地址的情况，不代表网关实际出口（例如经出口代理）的连通性。

注意：`zenova status` 只列 `zenova-` 开头的服务。`postgres`、`clickhouse`、`caddy`、`secret-service`、`spicedb` 用 `systemctl status <名称>` 查看。远端网关节点上的问题，在节点主机上执行 `sudo zenova status`、`sudo zenova logs node-edge`、`sudo zenova logs node-caddy` 采集。

注意：以上材料可能包含内部地址、账号名与请求内容片段。对外提交前按 12.6 节脱敏。不要采集或提交 `/opt/zenova/.env`、`/etc/zenova/secret-service/`、`/opt/zenova/.env.runtime/root-password-initial`，这三处含主密钥、密钥服务材料与初始密码。

12.3 安装阶段失败对照表

适用场景：`zenova install`（含 `--check-only` 预检）在确认安装前后停止，屏幕给出提示。

安装器在确认安装之前不写任何文件、不创建任何服务。确认前停止的，处理原因后重新执行同一条命令即可，本机没有需要清理的东西。

操作步骤

1. 按屏幕提示的原文在下表中查找。
2. 按「处理」列处理后，重新执行原命令。
3. 表中没有的，记下屏幕提示原文与诊断代码，查 12.4 节。

屏幕提示	原因	处理
答案不合法：请用 sudo 运行安装器	未以 root 执行	加 `sudo` 重新执行
安装包无法读取或已损坏	离线包不完整或路径写错	核对文件路径；核对文件大小与下载站显示的一致；不一致时重新下载（离线包支持断点续传）
安装器版本 … 与离线包版本 … 不一致	安装器与离线包不是同一版本	用与离线包同一版本的安装器；两者应从下载站同一次选择中下载
安装包与安装程序不属于同一版本	安装器与离线包分属不同产品	用同一产品的安装器与离线包；核对离线包文件名中的产品字段
本机已安装…，该安装程序为…版	本机已装另一个产品	换用已装产品的安装包；确需更换产品按第 11 章清空重装
离线包不适用于本机的平台或部署形态	系统或 CPU 架构与离线包不符	用 `uname -m` 核对本机架构，下载对应平台的离线包
…不含「…」能力，请从 --with 中去掉	指定了本产品不提供的能力	从 `--with` 中去掉该能力；确需该能力时改选提供它的产品
内存不足：需要 8 GB，本机 N GB	内存低于 8 GiB	扩容内存后重新执行
磁盘空间不足：/opt/zenova 需要 N GB，剩余 M GB	安装根所在卷可用空间不足	扩容；或用 `--install-dir <目录>` 安装到提示中列出的候选挂载点；或把 PostgreSQL、ClickHouse、对象存储改用现有服务以减少本机占用
本机没有 systemd，无法托管服务	主机不是 systemd 启动	更换为 systemd 启动的 Linux 主机
`[HOST_CLOCK_UNSYNCED]` / `[HOST_CLOCK_SKEW]`	时间未同步，或偏差超过 30 秒	重启 chrony 并等待同步（3.2 节），`chronyc tracking` 确认 `Leap status` 为 `Normal` 后重试

补充几种确认安装后才会出现的情况：

屏幕提示或现象	原因	处理
入口监听端口被占用（`HOST_ENTRY_PORT_BUSY`）	业务入口端口已被其他进程占用；安装停止，不自动更换	停止占用进程；或修改访问地址里的端口；或用 `--business-entry-port` 指定前置转发到的本机端口
安装上级目录不可遍历、挂载点带 noexec、SELinux / fapolicyd / KYSEC 拦截执行	服务用户 `zenova` 无法进入或执行安装根下的程序；这类拦截在解包后才能发现	按 3.3 节处理，必要时由主机安全管理员添加放行策略

屏幕提示或现象	原因	处理
装后验收失败：未通过的检查：<验收项>（<依据>）	某项装后验收未通过，安装不记为完成	按括号中的依据处理后，重新执行同一条安装命令，从装后验收继续

你会看到

- 处理后重跑，安装器从上次停下的位置继续：确认前停止的从头问答，确认后停止的沿用已有安装进度。

注意：`--check-only` 预检通过不代表正式安装一定通过。预检不包含业务入口端口检查、现有依赖检测、解包后的执行权限检查与通知测试，上面第二张表中的问题都可能在预检通过后才出现。

注意：不要删除 `/opt/zenova/state/` 下的进度文件或手工删除目录来「重新开始」。需要从头安装时，使用第 11 章的清空重装。

12.4 诊断代码怎么读

适用场景：安装器、`zenova doctor`、装后验收、升级、清空重装、备份恢复的输出中带有形如 `[HOST_DISK_LOW]` 的代码。

诊断代码共 56 个，按前缀分为九组。代码是稳定的：同一个问题在不同版本、中英文界面下代码不变，可以直接把代码报给支持，不需要转述屏幕文字。完整列表与处理办法见产品文档站的「安装与运维诊断代码」页。

前缀	数量	范围
`HOST`	22	主机条件：架构、内存、磁盘、时间、CPU 指令集、容器运行时、入口端口、目录权限、安全模块、root 权限
`PKG`	6	安装器与离线包：完整性、平台、版本、产品所含能力
`FRESH`	6	清空重装：外部数据、共享资源、归属证明、现场变化、参数
`NODE`	6	远端网关节点：已安装、注册、配置下发
`ANS`	4	安装答案与配置值：不合法、无法读取、续装不一致、产品已锁定
`NET`	4	网络：访问地址格式、依赖与通知渠道探测、目标连通性、管理面网段
`UPG`	4	升级与回滚：升级前检查、回滚被拒、节点未升级
`SVC`	3	本机服务：已完成安装、服务控制失败、无法确定本机安装
`BAK`	1	备份与恢复被拒绝

操作步骤

1. 先看前缀，确定是哪一类问题。
2. 再看代码后面的参数（如 path、have_gib、need_gib），它们给出了具体对象与数值。
3. 按屏幕给出的处理建议或下表处理；仍不清楚时，把代码与完整输出一并提交支持。

常见代码举例：

代码	含义	处理
HOST_DISK_LOW	安装目录所在卷可用空间不足，安装已停止；输出列出候选挂载点	用 --install-dir <目录> 安装到可用空间更大的卷，或清理磁盘后重试
HOST_CPU_FEATURES	CPU 缺少本机部署 ClickHouse 所需的指令集	虚拟机把 CPU 型号设为 host（x86-64-v3）；或在安装问答中改用现有 ClickHouse
HOST_ENTRY_PORT_BUSY	访问地址依赖的本机监听端口已被其他进程占用	停止该进程，或修改访问地址里的端口，或加 --business-entry-port
PKG_INVALID	无法读取或校验离线包	确认文件为完整的 .zpkg 后重试
PKG_VERSION_MISMATCH	安装器版本与离线包版本不一致	使用与离线包同一版本的安装器
PKG_CAP_NOT_IN_EDITION	所选产品不提供该能力	从 --with 中移除该能力，或换用提供该能力的产品
FRESH_EXTERNAL_DATA	清空重装检测到外部数据服务	安装器不会删除外部数据库或对象存储；在计划文件中保持 remove = false
FRESH_OWNER_UNPROVEN	部分对象无法自动证明由本安装创建	交互模式逐项 y/N 确认，或导出计划文件审核后执行（第 11 章）
FRESH_STATE_CHANGED	生成清空计划后现场已发生变化	重新执行 --fresh --dry-run 生成最新计划文件
NET_PUBLIC_URL	访问地址不是合法的 http/https 地址	用 zenova config set PUBLIC_URL http://<IP或域名> 修正
NET_DEP_UNREACHABLE	数据依赖或通知渠道探测未通过	核对目标地址、网络连通性与访问凭据后重试
UPG_NODES_NOT_UPGRADED	仍有在线网关节点未升级到目标版本	先升级列出的节点，再升级控制面
UPG_ROLLBACK_REFUSED	当前状态不允许回滚	升级到修复版本，或从升级前的备份恢复
SVC_ALREADY_INSTALLED	本机已完成安装	升级用 zenova upgrade；改配置用 --reconfigure；清空重装加 --fresh
NODE_ALREADY_INSTALLED	本机已存在网关节点安装	同一控制面升级加 --upgrade；换控制面或换令牌加 --reenroll --token <令牌>

注意：`zenova doctor` 有阻断项时退出码为 2，没有阻断项为 0，可以据此在巡检脚本中判断。警告项不影响退出码，但应记入交付记录。

12.5 装完之后访问与调用异常

控制台打不开

操作步骤

1. 先用安装完成时打印的「本机直接访问」地址在内网打开控制台。它按管理入口实际端口拼出，不经前置转发。能打开，说明本机正常，问题在负载均衡、反向代理、NAT 或网络策略，转第 4 步。
2. 「本机直接访问」也打不开时，在安装主机上查看服务状态：

```
# 安装主机 · root · 只读（查看各服务状态）
sudo zenova status
systemctl status caddy postgres clickhouse secret-service spicedb
```

3. 对未运行的服务重启并查看日志：

```
# 安装主机 · root · 写入（重启指定服务）
sudo zenova restart <服务>
sudo zenova logs <组件> --since '10 min ago'
```

4. 本机正常但外部访问不通：检查主机防火墙、云安全组与域名解析；使用了前置 TLS 或代理时，确认前置设备转发到的本机端口与安装时填写的一致。

你会看到

- 重启后服务状态恢复为运行中，控制台可以打开。
- 控制面不可用期间，网关节点仍按最近一次配置继续服务，已有业务调用不受影响。

注意：`zenova restart` 只接受 `zenova-` 开头的服务。控制面切换槽位的几秒内，内部入口返回 503 属于正常现象，入口代理会暂存新请求最长 5 秒。中心网关节点数据面（3080）就绪检查返回 503、原因为 `no_tenant_routes`，表示还没有租户或模型路由，创建租户并配置模型后恢复，不是故障。

调用返回错误

网关的三种协议入口（OpenAI、Anthropic、Gemini）在失败时都返回统一的错误信封，错误体中带稳定的 `reason`。HTTP 状态由 `reason` 决定，三种协议一致；响应头 `x-should-retry` 说明重试是否有意义。

同一个 HTTP 状态码下，`reason` 不同，处置完全不同。 只看状态码会走错方向：同样是 429，可能是成员额度用完（要调额度），也可能是上游供应商限流（稍后重试即可）；同样是 503，可能是模型没配路由（要改配置），也可能是上游暂时不可用（稍后重试）。排查时一律以 `reason` 为准。

HTTP 状态	reason	含义	谁来处理 · 是否可重试
400	`LLM_REQUEST_BODY_INVALID` / `VALIDATION_FAILED`	请求体不合法	调用方修改请求 · 否

HTTP 状态	reason	含义	谁来处理 · 是否可重试
400	`LLM_CONTEXT_OVERFLOW`	超出模型上下文长度	调用方修改请求 · 否
401	`UNAUTHENTICATED`	未通过身份认证（未携带 API 密钥或密钥无效）	调用方修改请求 · 否
401	`ACCESS_KEY_REVOKED`	API 密钥已被吊销	调用方换用有效密钥 · 否
402	`MONTHLY_TOKEN_EXCEEDED` / `INSUFFICIENT_CREDITS`	月度用量或余额不足	管理员调整额度 · 否
403	`PERMISSION_DENIED`	无权访问	管理员核对权限与可用范围 · 否
403	`MODEL_NOT_AVAILABLE_FOR_WORKSPACE`	该模型不在调用人的可用范围内	管理员把成员加入模型服务的可用范围 · 否
403	`CONTENT_POLICY_BLOCKED`	请求内容被内容安全规则拦截	调用方修改请求内容 · 否
403	`NETWORK_NOT_ALLOWED`	来源网段不被允许	管理员核对网段策略 · 否
403	`CAPABILITY_NOT_LICENSED`	许可证未包含该能力	租户管理员核对许可证 · 否
404	`LLM_MODEL_NOT_AVAILABLE`	请求的模型名不存在或不可用	调用方改用可用的模型名 · 否
408	`REQUEST_TIMEOUT`	请求超时	上游问题，稍后重试 · 是
429	`QUOTA_EXCEEDED` / `USER_TOKEN_EXCEEDED` / `USER_COST_EXCEEDED`	额度已用完	管理员调整额度 · 否
429	`MODEL_SERVICE_QUOTA_HARD_REJECT`	模型服务配额硬拒绝	管理员调整模型服务配额 · 否
429	`TRAFFIC_RATE_LIMITED` / `TRAFFIC_CONCURRENCY_LIMITED`	触发平台限流或并发上限	管理员核对限流规则；稍后重试 · 是
429	`LLM_RATE_LIMITED`	上游供应商限流	上游问题，稍后重试 · 是
502	`LLM_UPSTREAM_CREDENTIAL_REJECTED`	上游拒绝了平台保存的供应商凭据	管理员更新模型供应商凭据 · 否
502	`LLM_UPSTREAM_MODEL_NOT_FOUND`	上游不存在该模型	管理员核对供应商侧模型名 · 否
502	`LLM_BILLING_EXCEEDED`	上游账号额度耗尽	管理员处理供应商账号额度 · 否

HTTP 状态	reason	含义	谁来处理 · 是否可重试
502	`LLM_UPSTREAM_SERVER_ERROR`	上游服务端错误	上游问题，稍后重试 · 是
502	`LLM_UPSTREAM_CONNECT_FAILED` / `LLM_STREAM_INTERRUPTED`	与上游建连失败或流中断	带同一幂等键重试 · 是
503	`MODEL_ROUTE_NOT_CONFIGURED` / `NO_MODEL_SERVICE_SELECTED`	模型路由未配置或未选择模型服务	管理员检查模型服务与路由配置 · 否
503	`LLM_PROVIDER_UNAVAILABLE` / `LLM_OVERLOADED`	上游暂时不可用或过载	上游问题，稍后重试 · 是
503	`CONTROL_PLANE_UNAVAILABLE` / `SERVICE_UNAVAILABLE`	平台服务暂时不可用	管理员按本节「控制台打不开」排查 · 是
504	`LLM_TIMEOUT`	等待上游超时	上游问题，稍后重试 · 是

你会看到

- 「谁来处理」为调用方的，调用方修改请求后即可成功；为管理员的，重试不会改变结果，须调整配置；为上游的，重试可能成功。
- 流式请求在已开始输出后失败时，网关发送一条与所选协议错误体同形的流内错误帧，随后正常结束流；已输出的内容与用量照常计费和展示，`reason` 同样按上表解读。

注意：上表只列出最常见的 reason。完整列表（含每个 reason 的 HTTP 状态、错误类型与是否可重试）见产品文档站「错误参考」页。上表没有的 reason，直接把 `reason` 原文与 `x-request-id` 一起提交支持。

注意：模型供应商测试或 `zenova doctor --target` 报告连通失败时，失败分为七类：`dnsFailed`（域名解析）、`proxyConnectFailed`（出口代理建连）、`connectFailed`（目标 TCP 建连）、`tlsFailed`（TLS 握手与证书）、`timeout`（超时）、`ssrfBlocked`（出站安全策略拦截）、`invalidEndpointUrl`（地址格式非法）。按分类找对应层的负责人，不要从头逐层排查。

12.6 提交支持需要准备什么

适用场景：按 12.1～12.5 节处理后问题仍在，需要交付方或供应商支持。

操作步骤

1. 按 12.2 节采集四份材料。
2. 按下表整理问题信息。能提供的尽量提供，没有的写「无」，不要留空让支持猜。

字段	说明	从哪里取
发生时间与时区	精确到分钟，注明时区（如 UTC+8）	调用方日志或现场记录

字段	说明	从哪里取
请求 ID	响应头 `x-request-id`，错误体中的 `request_id` 与其相同	调用方响应；控制台「网关请求」
链路 ID	调用记录中显示链路 ID 时一并提供	控制台调用记录（以页面实际显示为准）
发起人	调用所属成员的账号名	控制台「网关请求」
接入凭据前缀	API 密钥的前几位，用于区分是哪一把密钥	控制台密钥列表；不要提供完整密钥
客户端 IP	调用方出口地址	调用方或控制台调用记录
声明模型与实际路由模型	调用方请求里写的模型名，以及实际路由到的模型	调用方请求；控制台调用记录
HTTP 状态码与错误码	HTTP 状态、错误体中的 `reason`；安装与运维类问题提供诊断代码	调用方响应；屏幕输出
复现步骤	从哪台机器、用什么命令或界面操作、得到什么结果；能复现时附可执行的 `curl`（密钥用 `<YOUR_API_KEY>` 代替）	现场整理
截图	控制台报错页、「网关请求」记录详情、终端输出	现场截取

3. 对全部材料脱敏（见下）。

4. 按以下顺序组织材料，打包提交：

顺序	内容
1	一句话问题描述：现象、影响范围、开始时间
2	环境信息：产品、完整版本号与通道、平台、已启用能力、数据依赖部署方式（安装器部署 / 现有服务）、是否有远端网关节点
3	上表字段
4	已经做过的排查与处理，以及结果
5	12.2 节采集的四份材料，及 `zenova doctor --target` 报告（如有）
6	截图

你会看到

- 材料齐全时，支持可以直接按请求 ID 与诊断代码定位，不需要来回追问。

注意：以下内容必须脱敏或不得提交：

内容	处理
API 密钥、模型供应商凭据、数据库口令、对象存储访问密钥	只保留前缀，其余删除
root 密码与初始密码、两步验证恢复码	不得提交
/opt/zenova/.env、/etc/zenova/secret-service/、/opt/zenova/.env.runtime/ 下的文件，密钥归档与平台备份导出文件	不得提交；含主密钥与密钥服务材料
许可证文件内容	不得提交；需要时只说明许可证状态
请求与响应正文中的业务数据、个人信息	删除或替换为示例内容
内部 IP、域名、主机名	按客户信息安全要求替换，替换前后保持一一对应，便于支持比对

注意：初始密码与任何凭据不得通过即时通讯、邮件或工单传递；截图与录屏前确认画面中没有密码、完整密钥与恢复码。

能力说明

不提供的能力	替代做法
一键打包全部排查材料并自动脱敏	按 12.2 节四条命令采集，按 12.6 节手工脱敏
安装主机日志自动上报给供应商	由客户确认脱敏后，按 12.6 节顺序人工提交
zenova status 列出全部系统服务	postgres、clickhouse、caddy、secret-service、spicedb 用 systemctl status <名称> 查看
zenova doctor --target 代表网关实际出口的连通性	报告注明「直连」时，以模型供应商测试或经网关的真实调用结果为准
只凭 HTTP 状态码判断处置方式	以错误体中的 reason 为准，查 12.5 节或「错误参考」页