

## 虚拟机环境准备

在开始架设 **hadoop** 环境之前，需要准备一定的运行环境，**hadoop** 可以运行在多种不同环境中，常见的环境可以是安装运行在多个物理机器上，也可以在虚拟化环境中运行，还可以运行在云环境中，包括公有云和私有云。

物理机和虚拟化环境中运行 **hadoop** 各有优缺点，两者选择要根据实际使用场景来进行判断。一般来说使用 **hadoop** 是因为单机资源不够，想使用分布式环境来利用更多的资源快速解决问题，而虚拟机环境则是因为单机资源充足，有多余资源未能充分利用，为提高使用效率而发展起来的技术。随着硬件技术的发展，资源未能充分利用的情况越来越多，所以现在虚拟环境的使用也越来越普遍，随着虚拟化技术的发展，使用虚拟化环境对计算资源的管理也越来越方便，同时还提供了更多的其它特性，比如高可靠性，节约能源等等。

使用物理机（有时又称为裸机）环境能够避免因虚拟软件的故障问题和虚拟化带来的性能损失，如果是一个需要长期高负荷运行的 **hadoop** 环境，建议考虑物理机环境，这样可以提供更多的资源给使用者，物理环境的部署与虚拟环境的部署基本相同，只是少了虚拟化环境的部署过程。如果部署的 **hadoop** 系统较少进行高负载的运行，则可以考虑部署在虚拟化环境中，平常可将资源分配给其它业务使用，在需要进行 **hadoop** 高强度运算时再灵活为 **hadoop** 分配足够资源即可，这样可提高资源的综合利用率。

虚拟化环境在这里我们简单分为桌面级虚拟化和企业级虚拟化，两者基本原理相差不多，企业级虚拟化软件提供了更加高效的运行和管理支持。为方便学习，我们将在虚拟环境中安装 **hadoop** 环境。

在个人电脑上，我们可以利用桌面虚拟化软件进行虚拟化操作，常见的桌面虚拟化软件有开源的 **virtualbox**，非开源的有 **VMware Workstation**(收费软件)，**VMware Player**(免费)，这三款软件 都有 **Windows**，**Linux**，**MacOS** 平台的对应版本，**Windows** 环境下，还可使用微软的 **Hyper-V** 技术进行虚拟化(桌面系统建议使用 **Win10**)。企业级虚拟化在 **x86** 架构上目前主要的虚拟化软件有 **VMware vsphere**、**Microsoft Hyper-V**、**Citrix XenServer** 以及基本 **linux** 发行版的 **kvm** 虚拟化(比如 **REDHAT** 的 **RHEV**)。

企业级虚拟平台安装部署相对比较复杂，这里我们仅以 VMware Workstation 试用版为例进行介绍，VMware Workstion 试用版提供了跟正式版相同的功能，只是有 60 天的使用期限。

首先下载 VMware Workstation 安装文件，可到 VMware 官方申请下载最新版本，网址为 [www.vmware.com](http://www.vmware.com)。安装过程如下：

以管理员身份运行安装文件。

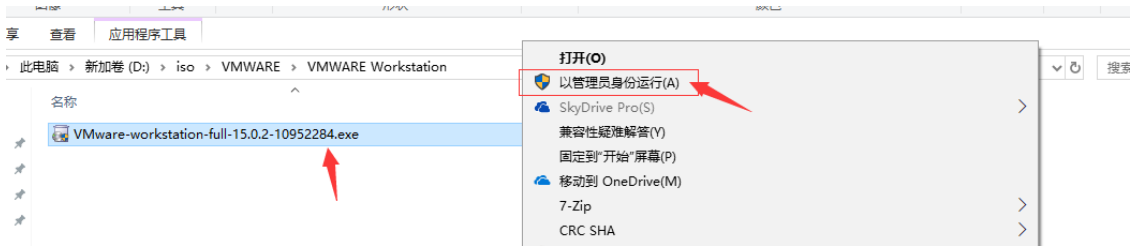


图 3-1 运行虚拟机安装软件



图 3-2 安装向导图

接下来基本一直点击下一步，即可正常安装好软件。

请输入序列号：

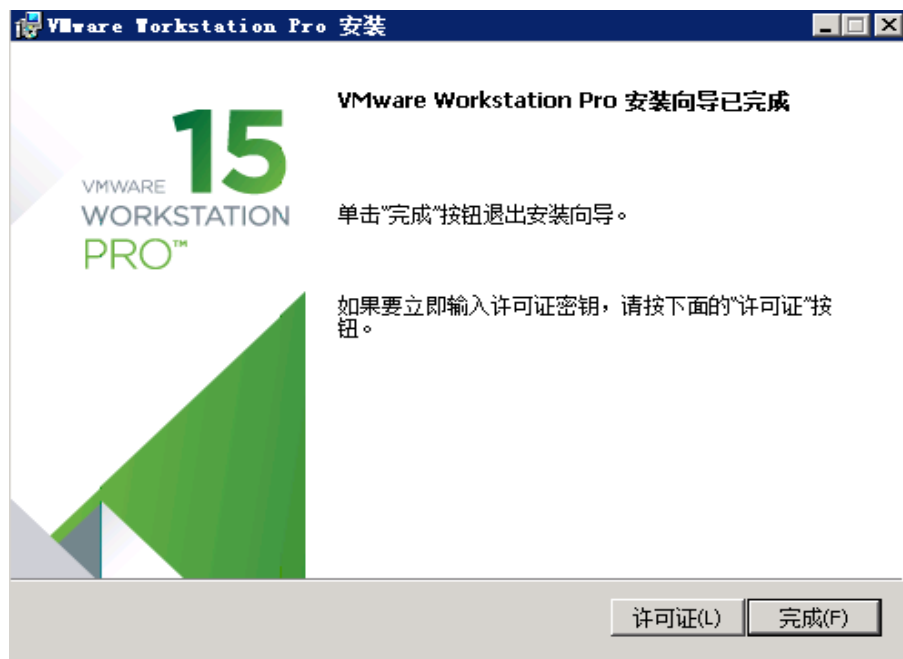


图 3-3 序列号输入

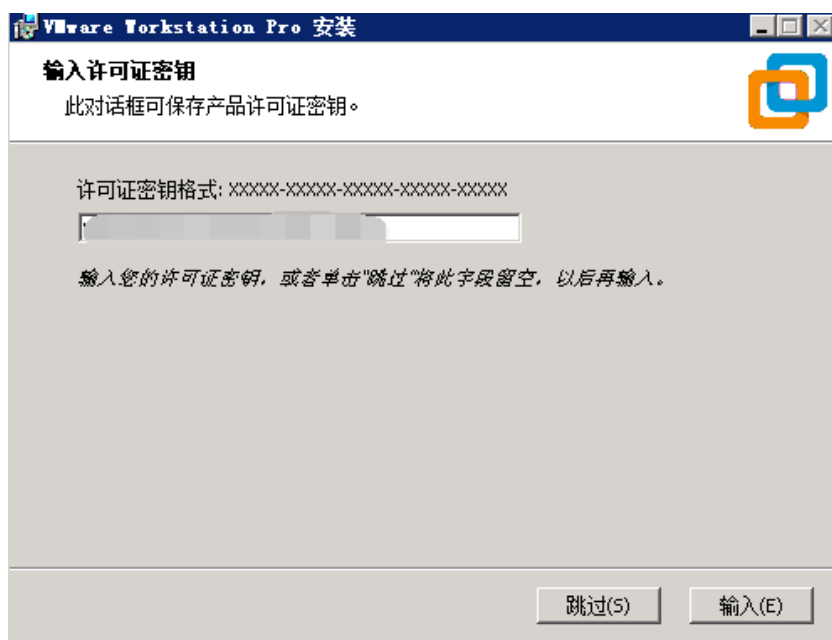


图 3-4 许可密钥验证

启动虚拟化软件

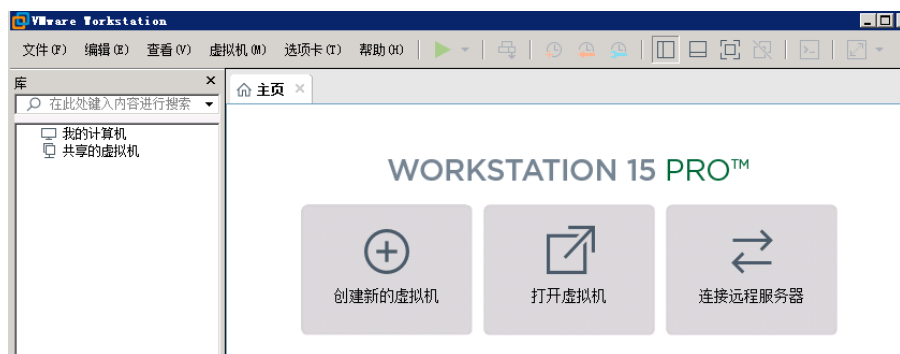


图 3-5 运行虚拟化软件

建议设置虚拟机默认存放位置，将后续创建的虚拟机存放到指定位置。

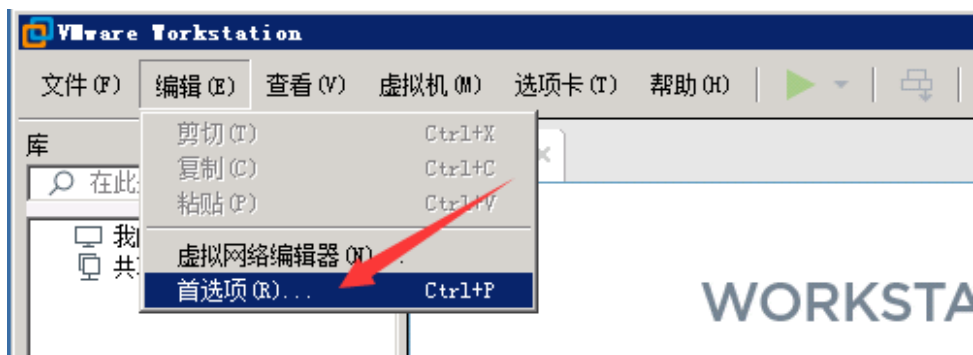


图 3-6 配置首选项

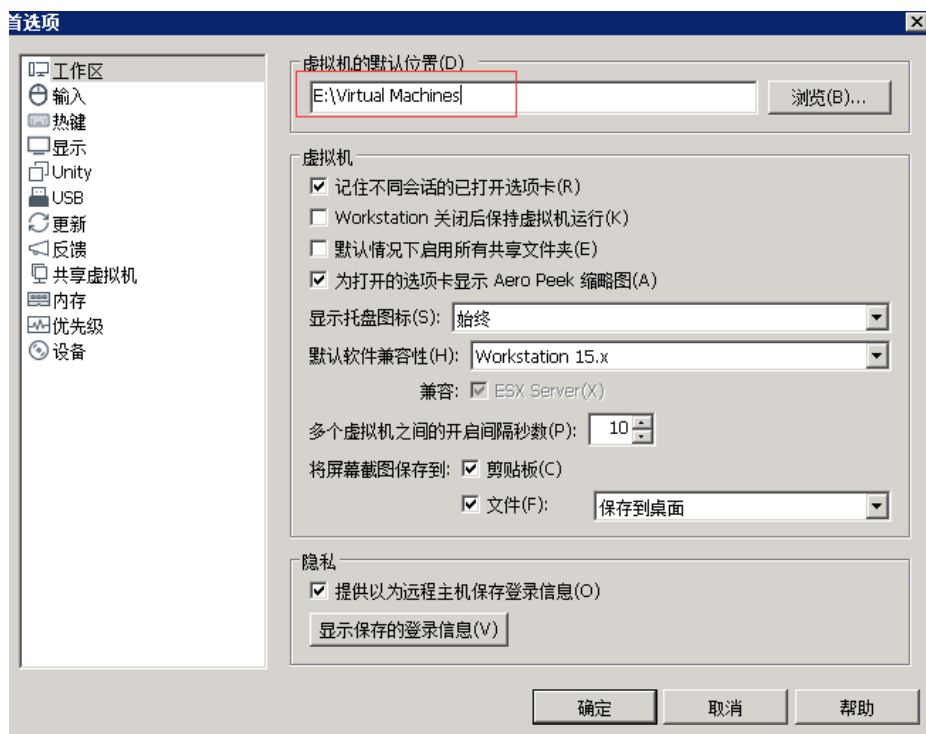


图 3-7 配置虚拟机默认路径

到这里虚拟机软件安装完成。