

# 紫光云 服务器操作系统技术白皮书

Copyright © 2022 紫光云技术有限公司 版权所有，保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

除紫光云技术有限公司的商标外，本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。

本文档中的信息可能变动，恕不另行通知。

## 目录

<b>1 产品概述</b> .....	<b>3</b>
<b>2 系统架构</b> .....	<b>4</b>
2.1 架构概述 .....	4
2.2 主要技术特性 .....	5
2.2.1 稳定可靠 .....	5
2.2.2 易维护性 .....	8
2.2.3 高性能 .....	10
2.2.4 高安全 .....	12
2.3 技术指标与规格 .....	14
<b>3 系统主要功能和服务</b> .....	<b>15</b>
3.1 支持虚拟化 .....	15
3.2 绿色节能 .....	15
3.3 邮件服务 .....	15
3.4 文件服务 .....	16
3.5 WEB 服务 .....	17
3.6 代理服务 .....	17
3.7 DDE 图形桌面 .....	18
3.8 内存分级扩展 .....	18

# 1 产品概述

伴随信息化技术的飞跃发展，数据中心管理中存在的诸多矛盾被进一步激化。在业务发展和数字化转型的过程中，越来越多的企业和组织也期望部门能够敏捷应对持续演变的业务需求。操作系统作为软硬件系统的核心部件，向上支持应用与服务，包括提升用户体验、提升业务竞争力和生态入口，向下抽象并管理各类硬件，统筹系统资源，在整个信息化系统中具有重要地位。

紫光云在操作系统研发方面具有丰富的经验，掌握操作系统内核研发相关核心技术。紫光云 UnisLinux（以下简称 UnisLinux）是紫光云集团基于 Linux 内核开发的国产化操作系统发行版本。UnisLinux 是一款出色的服务器操作系统，集成了大数据、虚拟化、云计算、边缘计算等多种技术能力并支持多样化应用场景，借助强大算力与海量存储，依托数据智能分析手段，UnisLinux 帮助用户在复杂且多样的环境中及时交付应用程序和功能，并为容器化、微服务等重要 IT 举措提供支持，助力百行百业用户实现数字化转型。

UnisLinux 目前已在紫光云多个产品中大规模使用，在技术方案中被政府、运营商、金融及教育等行业客户广泛采用。紫光云 UnisLinux 操作系统一直以其稳定，安全，高效，对企业级核心应用的良好支持等优点广受用户信赖。

UnisLinux 系统主要具有以下技术特点：

- 多算力支持

UnisLinux 支持国内飞腾、鲲鹏、海光等多个国产化平台，同时支持国外 intel 以及 AMD 平台，相比国外操作系统支持更多的 CPU 架构体系，并在国产基础硬件平台上有出色的性能表现；

- 性能优异

UnisLinux 对单线程、多线程的调用和回收能保证无阻塞、无障碍、无锁、少等待。对底层资源（如 CPU、IO、RAM）能进行统一调度和管理，对上层应用请求能使用最优方案分配运算资源，以及快速反馈运算后响应。

- 安全可靠

UnisLinux 的高安全性能支撑安全级别所需的访问控制、权限分隔、密码功能、可信验证等能力，并支持国密算法模块、可信执行环境、可信计算等安全基础设施。

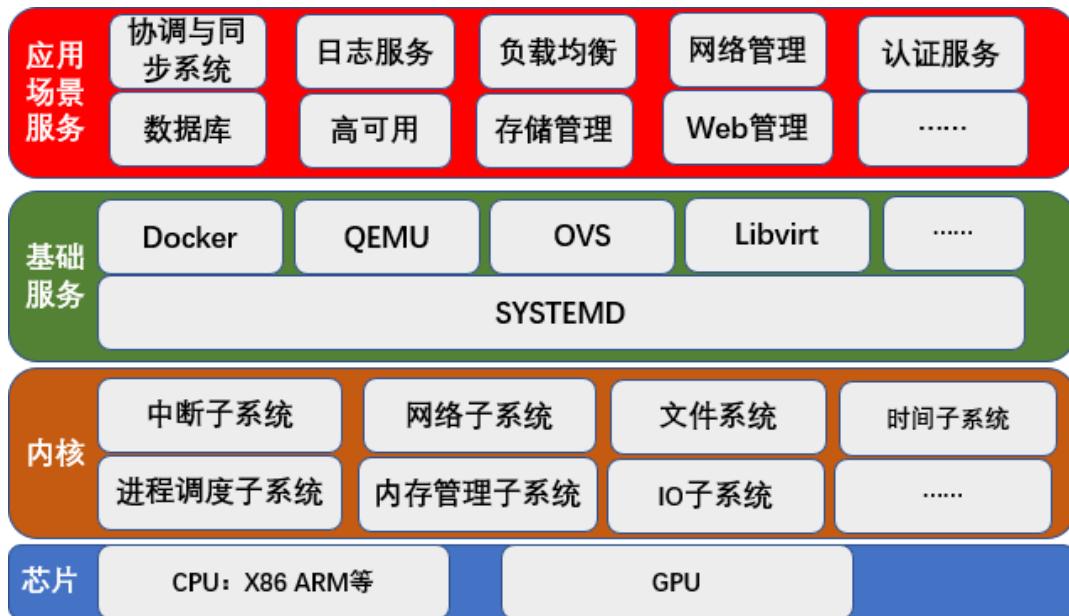
## 2 系统架构

### 2.1 架构概述

UnisLinux 2.0 基于 Linux 内核和多个开源社区软件做了二次开发和增强。紫光云操作系统研发团队对系统内核功能模块进行梳理和重新设计，并且对安全体系结构的底层支撑进行了统一安全增强设计，最终形成稳定统一的调用接口。此外根据服务器运行场景的实际需求，在性能、稳定性等方面对系统组件的基础函数库、工具和图形环境等方面进行了增强和优化。

UnisLinux 2.0 由操作系统内核（kernel）、基础服务（文件系统和 shell）、应用场景服务等多个部分组成。内核是操作系统的核心，内核不仅实现了进程调度、

内存管理、中断处理、异常处理，而且还实现了进程管理、进程通信机制、虚拟内存管理、文件系统驱动和网络等各类设备驱动子系统。操作系统基础服务主要面向虚拟化和容器两大主流行业应用场景，提供操作系统层的服务组件（不含集群平台组件）；应用场景服务主要面向各行业场景，提供基础软件包与依赖库。

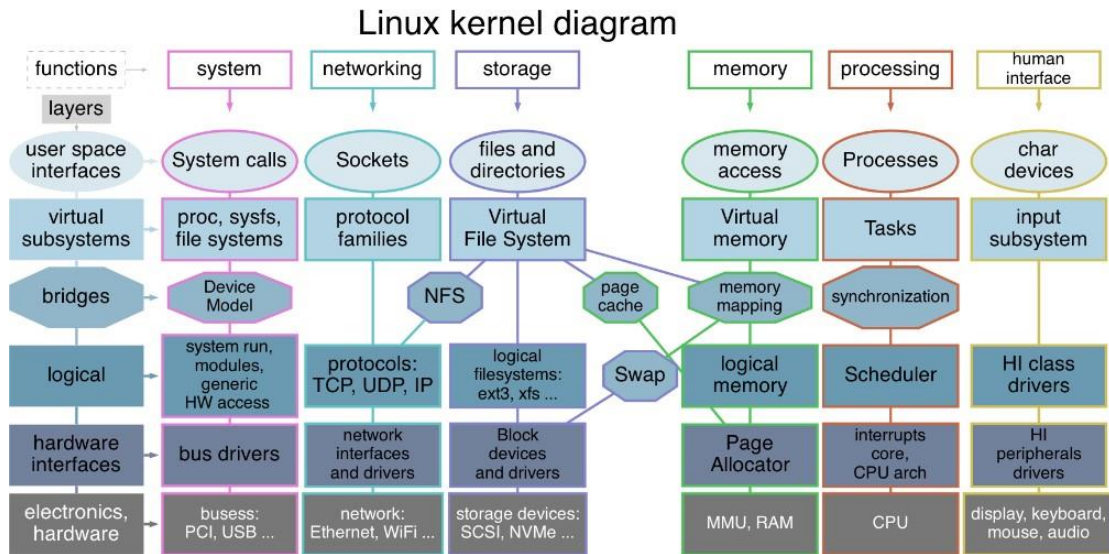


## 2.2 主要技术特性

### 2.2.1 稳定可靠

- 稳定内核

选取上游稳定版 Linux 5.10 内核构建，对外接口遵循 POSIX 标准，并根据客户业务功能需求和特性，提供内核定制、系统服务裁剪与扩充、性能调优、应用迁移指南等服务。



- 网络性能优化

DMA 全称是 Direct Memory Access。DMA 可以同时应用于网络数据的发送和接收。DMA 本身是一个通用的技术，它有一个独立于 CPU 的 DMA 控制器。在数据拷贝的时候，CPU 只需要告诉 DMA 控制器，拷贝数据的起始地址，数据长度，之后将总线控制权交给 DMA 控制器，就可以不需要 CPU 的介入，完成数据拷贝。

使用 DMA，在网卡从内存拷贝数据（发送），和网卡向内存拷贝数据（接收）时，只需要很少的 CPU 介入。

- IPv4/IPv6 双栈支持

2017 年 11 月 26 日，中共中央办公厅、国务院办公厅印发《推进互联网协议第六版（IPv6）规模部署行动计划》，计划在 5 到 10 年的时间里，形成下一代互联网自主技术体系和产业生态，建成全球最大规模的 IPv6 商业应用网络，实现下一代互联网在经济社会各领域深度融合应用，成为全球下一代互联网发展的重要主导力量。因此，IPv6 已经上升到了网络强国的国家战略层面。

与 IPv4 相比，IPv6 具有以下几个优势：

- (1) 海量 IP 地址

IPv6 具有更大的地址空间。IPv4 中规定 IP 地址长度为 32，最大地址个数为  $2^{32}$ ；而 IPv6 中 IP 地址的长度为 128，即最大地址个数为  $2^{128}$ 。与 32 位地址空间相比，其地址空间增加了  $2^{128}-2^{32}$  个。

#### (2) 高效的转发能力

IPv6 使用更小的路由表。IPv6 的地址分配一开始就遵循聚类 ( Aggregation ) 的原则，这使得路由器能在路由表中用一条记录 ( Entry ) 表示一片子网，大大减小了路由器中路由表的长度，提高了路由器转发数据包的速度。

#### (3) 敏捷接入

IPv6 加入了对自动配置 ( Auto Configuration ) 的支持。这是对 DHCP 协议的改进和扩展，使得网络 ( 尤其是局域网 ) 的管理更加方便和快捷。

#### (4) 端到端安全

IPv6 具有更高的安全性。在使用 IPv6 网络中用户可以对网络层的数据进行加密并对 IP 报文进行校验，在 IPv6 中的加密与鉴别选项提供了分组的保密性与完整性。极大的增强了网络的安全性。支持支持端到端 IPsec 加密，克服传统的 IPsec 与 NAT44 冲突问题。

#### (5) 增强对流的控制

IPv6 增加了增强的组播 ( Multicast ) 支持以及对流的控制 ( Flow Control )，这使得网络上的多媒体应用有了长足发展的机会，为服务质量 ( QoS, Quality of Service ) 控制提供了良好的网络平台。

#### (6) 便于溯源

所有终端都可以采用全球唯一 IPV6 地址标识并被记录，可对网络上的不法犯罪活动做到有效监控和打击。

- 高可靠性优化

UnisLinux 可以对数据中心 IT 基础设施进行基于集群的集中化管理，由多台独立服务器主机聚合形成的集群不仅降低了管理的复杂度，而且具有内在的高可靠性，从而为用户提供一个经济、有效、适用于所有应用的高可靠性解决方案。

UnisLinux 定时对集群内的主机和虚拟机状态进行监测，当服务器发生故障的时候，受影响的虚拟机将在集群中留有备用容量的其它主机上自动重启，从而将停机时间和服务中断降低到最低，提供不同故障情况下的高可靠性保证。

此外，UnisLinux 还吸收了大量上游社区高版本的安全新特性，增加系统的高可靠性：

**支持 PAC(Pointer Authentication Code) 特性：**在使用寄存器的值作为指针访问数据或代码之前验证其内容，抵御 ROP/JOP 攻击。

**支持 BTI ( Branch Target Identifiers) 特性：**对间接跳转的目标进行限制。与 PA 结合使用减少控制流攻击。

**支持 SGX 特性：**SGX (software guard extensions)是 Intel 推出的指令集扩展，旨在以硬件安全为强制性保障，不依赖于固件和软件的安全状态，提供用户空间的可信执行环境，通过一组新的指令集扩展与访问控制机制，实现不同程序间的隔离运行，保障用户关键代码和数据的机密性与完整性不受恶意软件的破坏。

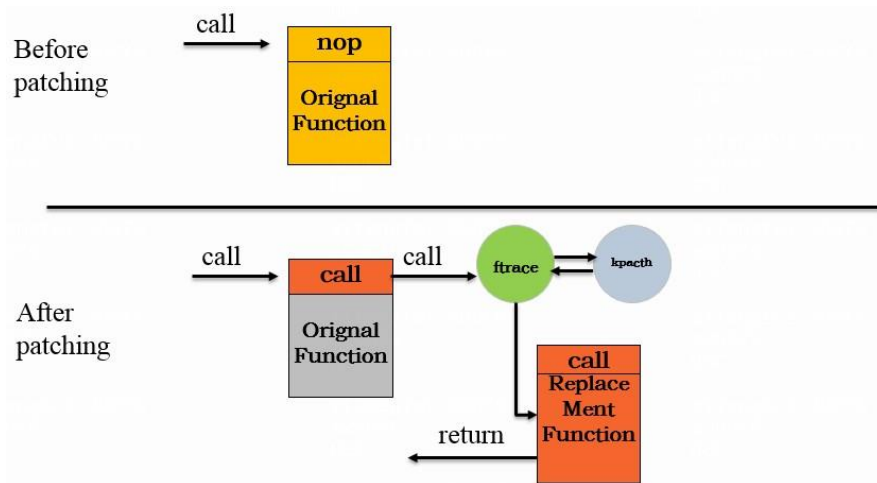
## 2.2.2 易维护性

- 热补丁升级机制

在升级最新系统补丁、修复内核缺陷和安全漏洞时，无需重启系统或中断业务应用，最大程度上减少系统关机和业务中断时间，增加系统可靠性和稳定性。

热补丁实现的底层原理图如下：





- 调度问题分析工具 sched\_delay

CPU 资源利用和调度情况在很大程度上影响着系统的性能。当 CPU 资源得不到合理的利用或者有个别进程长时间占用过多的资源，对上层产品业务、用户体验等造成极大影响。因此对于 CPU 资源调度事件的监控和准确地性能分析，并及时发现性能瓶颈，是十分必要的。

sched\_delay 是 UnisLinux 团队开发的调度系统子诊断工具之一。用于分析调度系统和监控 CPU 资源整体使用情况并追踪热点进程；提供 CPU 运行时序图功能。使用该工具可以清晰的得到 CPU 资源使用情况和 CPU 运行时序分析图。并可以实现对所有进程调度情况的分析统计，包括切换次数、平均（最大）调度延迟、最大延迟时间点等信息。如下较为直观的可视化图表，对 CPU 资源使用和性能分析提供了极大的便利。

- 内存分析工具

mem\_probe 是 UnisLinux 团队开发的内存子系统诊断工具之一。可统计出内核内存占用字节数，用户态内存占用字节数，以及内核直接通过 alloc\_page 申请无法在 meminfo 中显示的内存占用字节数。其中用户态内存占用从 LRU, Page

Cached 以及进程占用的 Pss 内存总合，3 个不同的维度呈现用户态的内存占用

情况，高效定位内存不足问题。

```
.SH EXAMPLES
启动mem_probe, 分析系统内存, 并保存此次分析结果

% h3diag mem_probe enable -r
MemTotal:      3872164 kB
MemFree:       2644780 kB
MemAvailable:  3164968 kB
...
HugePages_Free:      0
HugePages_Rsvd:      0
HugePages_Surp:      0
Hugepagesize:        2048 kB
DirectMap4k:         98156 kB
DirectMap2M:        4096000 kB
-----
Kernel mem:         105 MB
LRU mem:            1055 MB X: 38 MB --> 以LRU的占用情况计算X的值
Cached mem:         1054 MB X: 38 MB --> 以Page Cached的占用情况计算X的值
Pss mem:            410 MB X: 38 MB --> 以所有进程占用的Pss的情况计算X的值
```

### • IO 性能分析工具

I/O 子系统是操作系统中负责和 I/O 设备打交道的子系统，其性能直接影响上层的业务吞吐量和读写延迟。io\_tmout 是 UnisLinux 团队开发的 IO 系统子诊断工具之一，用于分析内核态 IO 超时问题和定位 IO 性能瓶颈。报告中会打印 IO 在内核态不同阶段消耗的时间(s)，包括最小值、平均值、最大值，以及数量。且支持 IO 级别的 q2g |g2i | q2d | i2d | d2c | q2c 阶段分析，通过分析报告可以快速确认 IO 性能瓶颈。

```
h3diag io_tmout抓取到的信息形式如下:

Tue Sep 14 09:37:48 CST 2021      抓取 sda 持续 10 s
===== All Devices =====

```

	ALL	MIN	AVG	MAX	N
Q2Q	0.000005142	1.408171224	5.632660189	4	
Q2G	0.000001279	0.000002372	0.000005403	5	
G2I	0.000004845	0.000046054	0.000166371	5	
I2D	0.000002847	0.000006069	0.000009765	5	
D2C	0.000054135	0.000073214	0.000107230	5	
Q2C	0.000086414	0.000127709	0.000282011	5	

```
===== Device Overhead =====

```

DEV	Q2G	G2I	Q2M	I2D	D2C
( 8, 0)	1.8572%	36.0619%	0.0000%	4.7522%	57.3287%
Overall	1.8572%	36.0619%	0.0000%	4.7522%	57.3287%

```
===== Device Merge Information =====
.....
```

## 2.2.3 高性能

UnisLinux 2.0 作为高性能的服务器操作系统平台，持续对场景化应用提供性

能优化，针对系统微架构层、系统服务层和应用接口层做了大量针对性的优化。

包括但不限于：

- 云原生调度增强：针对云场景在线和离线业务混合部署场景，创新 CPU 调度算法，保障在线业务对 CPU 的实时抢占及抖动抑制，创新业务优先级 OOM 内存回收算法，保障在线业务安全可靠运行。

- 内存分级扩展 etMem：新增用户态 swap 功能，策略配置淘汰的冷内存交换到用户态存储，用户无感知，性能优于内核态 swap。

- 内存 RAS 增强：内存可靠性分级技术，可以指定内核、关键进程等对内存故障敏感的数据优先使用高可靠内存，降低宕机率，提升可靠性（技术预览特性）。UnisLinux 2.0 基于开源 Linux 5.10，吸收了大量社区高版本的高性能特性：

- 进程调度优化：优化进程负载均衡算法，减少负载均衡过程中的开销，提升性能。

- 内核动态抢占：新增启动选 preempt=none/voluntary/full，允许内核动态切换抢占模式。

- mremap 性能优化：通过移动 PMD/PUD 级别的表项，加速映射大块内存的速度。

- per memcg lru lock：采用 per memcg lru\_lock，减少云原生容器实例锁竞争，提升系统性能。

- 大页内存管理优化：通过共享映射方式将 HugeTLB 管理页中无实际作用的 tail 页释放掉，降低大页内存管理结构的开销。

- TLB 并发刷新支持：本地 TLB 和远端 TLB 刷新并行，优化 TLB shutdown 流程加速 TLB 刷新，提升业务性能。

- 大页 vmalloc 性能优化：对于超过 huge page 的最小 size 的空间进行 vmalloc() 分配时，将会尝试使用 hugepage 而不是 base page 来映射内存，改善 TLB 的利用，降低 TLB miss。

- SVA ( Shared Virtual Addressing ) 支持：进程虚拟地址在主机进程和设备间共享，实现资源跨主机与设备免拷贝复用，提升跨主机和设备业务通讯性能。

- TCP 压缩特性：hbase 等分布式数据库节点间数据传输量大，网络传输是性能瓶颈；在 TCP 层对指定端口的数据进行压缩后再传输，收包侧把数据解压后再传给用户态，从而提升分布式场景节点间数据传输的效率。

- XDP ( eXpress Data Path ) 支持：基于 ebpf 的一种高性能、用户可编程的网络数据包传输路径，在网络报文还未进入网络协议栈之前就对数据进行处理，提升网络性能。可用于 DDOS 防御、防火墙、网络 QOS 等场景。

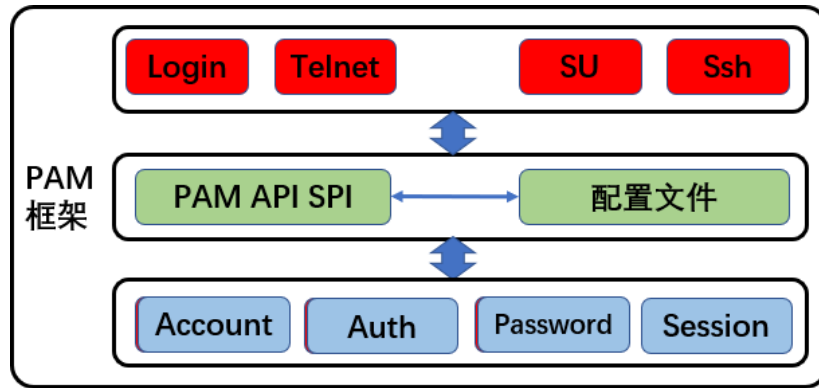
## 2.2.4 高安全

- 安全审计：实现基于内核级别的审计，记录整个系统活动以及相关修改。



- 身份鉴别与认证机制：保证用户身份在系统全生命周期内的唯一性。

- 多因子认证：保障用户身份的真实性鉴别，加强安全保护，采用统一的 PAM 认证。



- 支持国密算法

密码算法是保障信息安全的核心技术，在一些重要的行业例如：银行业、游戏业、互联网业以前长期以来都是沿用 3DES、SHA-1、RSA 等国际通用的密码算法体系及相关标准。UnisLinux 2.0 支持加密角度更高、响应时间更快的国家商密算法 SM2、SM3 和 SM4，即非对称算法、哈希算法和对称算法。



- 安全漏洞管理机制

UnisLinux 团队开发了完善的漏洞追踪方案，自动同步各开源软件 CVE 漏洞，方便用户查看各个 UnisLinux 产品的漏洞修复情况。

记录日期	漏洞编号	概述	发布日期	影响
2021-12-16 15:52:53	CVE-2021-25291	RedHat: A flaw was found in python-pillow. Invalid tile boundaries could lead to an OOB Read in TiffReadRGBATile in TiffDecode.c.	2021-02-28T08:00:00+08:00	中
2021-12-16 15:52:53	CVE-2021-25292	RedHat: A flaw was found in python-pillow. The PDF parser has a catastrophic backtracking regex that could be used as a DOS attack.	2021-02-28T08:00:00+08:00	中
2021-12-16 15:52:54	CVE-2021-25293	RedHat: A flaw was found in python-pillow. There is an Out of Bounds Read in SGIRleDecod e.c.	2021-02-28T08:00:00+08:00	中
2021-12-16 15:52:54	CVE-2021-25289	RedHat: A flaw was found in python-pillow. TiffDecode has a heap-based buffer overflow when decoding crafted YCbCr files because of certain interpretation conflicts with LibTIFF in RGB A mode. The previous fix for CVE-2020-35654 was insufficient due to incorrect error checking in TiffDecode.c. The highest threat from this vulnerability is to data confidentiality and integrity as well as system availability.	2021-02-28T08:00:00+08:00	高
2021-12-16 15:52:54	CVE-2021-25290	RedHat: A flaw was found in python-pillow. In TiffDecode.c, there is a negative-offset memcpy with an invalid size which could lead to a system crash.	2021-02-28T08:00:00+08:00	中
2021-12-16 15:52:54	CVE-2021-30178	RedHat: A flaw was found in the Linux kernel. A NULL pointer dereference occurs for certain accesses to the SynIC Hyper-V context. The highest threat from this vulnerability is to system availability.	2021-02-26T08:00:00+08:00	中

## 2.3 技术指标与规格

类别	技术指示与规格
CPU	支持 Intel, AMD, 海光, 飞腾, 鲲鹏
内存	不小于 4GB (为了获得更好的应用体验, 建议不小于 8GB)
硬盘	不小于 32GB (为了获得更好的应用体验, 建议不小于 120GB)
标准符合度	符合POSIX, 符合 LSB5.0
核心参数	Kernel: 5.10  Glibc: 2.34  Gcc: 10.3.1  Systemd: 249
板卡支持	存储设备: 支持 SATA, IDE, SCSI, RAID 等  高速板卡: 支持 HBA 卡, Infiniband 卡等  低速接口: 支持 LAN 卡、USB 接口等
安装方式支持	支持通过光盘、USB 等物理存储介质安装  支持PXE 网络安装  支持 kickstart 自动化安装
升级方式支持	支持镜像包本地升级, 支持单包/补丁包升级

## 3 系统主要功能和服务

### 3.1 支持虚拟化

UnisLinux 为客户应用提供了多种灵活的部署方案，支持提供 KVM 虚拟机方案，也支持用 Docker 和 Kubernetes 进行容器化应用的生命周期管理，支持管理大规模容器集群能力，提供多种应用发布方式和持续交付能力并支持微服务架构；满足行业合规监管和安全要求；为客户业务应用提供更好的灵活性、扩展性、高可用性、更高效的业务上线效率。适用于已搭建虚拟化平台，且对 PaaS 平台有需求的政府、新型互联网或金融等创新型企业，希望从技术层面更有效地支持业务创新。

### 3.2 绿色节能

UnisLinux 通过系统级的设计和重构，对现有 Linux 内核中的能耗管理部分进行了重构，优化系统能效，提升整机资源利用率，减少单位算力的电能消耗，整体减少系统运行时的碳排放。

CPU 是服务器功耗的主要来源，通过降电压使 CPU 进不同深度睡眠，可降低 CPU 的功耗，当前大部分主流 CPU 支持 CPU 深度睡眠状态(C1-C6)，能效提升空间达 5-30%。

### 3.3 邮件服务

Sendmail 是 Internet 上最流行的邮件传输代理 ( MTA )，它处理互联网上绝大部分电子邮件的传送。Sendmail 功能强大、遵从互联网标准，高可配置，允

许用户控制电子邮件几乎每一个处理过程。它的主要任务是在主机之间安全地传送电子邮件，通常使用简单邮件传输协议（SMTP）。Sendmail 作为邮件路由器它获取信件、检查收件人地址并确定发送邮件的最好路径，并支持邮件转发、邮件过滤提供安全认证机制。SSLSMTP 和 SSLPOP 是在 SSL 基础上建立的安全传输通道上运行 SMTP 和 POP 协议，同时又对这两种协议作了一定的扩展，以便更好地支持加密的认证和传输。提供 smtp 认证，支持 SASL、SSL/TLS 等加密机制，防止 Dos 攻击。Sendmail 可以与 LDAP 结合使用户可以迅速、快捷的找到大量的特定用户的特定信息。除 Sendmail 外，UnisLinux 还可支持 postfix、dovecot、imap4 等邮件服务器服务。

### 3.4 文件服务

UnisLinux 除了支持 ext4，xfs 等本地文件系统外，还支持 EulerFS 和 XFS 等文件格式。EulerFS 是一种面向非易失性内存的新文件系统，采用软更新、目录双视图等技术减少文件元数据同步时间，提升文件读写性能。XFS 是一种高性能的日志文件系统。XFS 文件系统采用回写模式日志，提高了系统本身的性能，但实际数据并没有存进日志文件中，因此带来了一定的风险。

XFS 文件系统的特点如下：

**数据完全性：**无论文件系统中存储的文件与数据有多少，文件系统都可以根据所记录的日志在很短的时间内迅速恢复磁盘文件的内容。

**传输特性：**XFS 文件系统采用优化算法，日志记录对整体文件操作影响非常小。XFS 查询与分配存储空间非常快而且还能连续提供快速的反应时间。

**可扩展性：**可以支持上百万 T 字节的存储空间。对特大文件、小文件或较大



数量目录都支持。

## 3.5 Web 服务

Apache Web 服务器具有稳定性高，速度快，功能强，可扩展性好的特点，它可以完成普通 Web 服务如：虚拟主机，代理服务，安全控制等多种服务。Apache Web 服务器可以提供目录，文件和 URL 等级别的访问控制。并支持 HTML、PHP 等脚本语言、支持 MySQL、Postgresql 等数据库；提供基于安全套接字层的安全控制 OpenSSL 支持与保密协议——HTTPS 结合可用于加密网络传输的信息和数据；提供 Tux 与 Apache 兼容的基于核心的线程级高性能 Web 服务器；提供 PHP 嵌入脚本语言、Python 语言、Perl、CGI 等语言模块的支持；提供 ASP 到 PHP 脚本的转换工具。

## 3.6 代理服务

Squid 代理服务器是比较优秀的代理服务器软件，它可以在服务器上作一个很大的缓存，可以把好多常去的网站内容存储到缓存中，这样，内部网的机器再访问那些网站，就可以从缓存里调用了。这样一方面可以加快内部网浏览因特网的速度，这就是所谓的提高客户机的访问命中率，另一方面，Squid 不仅仅支持 HTTP 协议，而且还支持 FTP，GOPHER，SSL 和 WAIS 等协议。Squid 支持的功能包括基于 IP 的访问控制、基于 URL 的访问控制、提供查看内存、交换空间，高速缓存目录的位置，所接受的连接类型及接受连接的端口的日志文件，设置最大请求连接数，管理员邮箱地址等。

## 3.7 DDE 图形桌面

DDE 是由统信软件开发的图形桌面系统，已与 Redhat 领导的 Gnome、Suse 领导的 KDE 和 Ubuntu 领导的 Unity 一起，成为全球范围内主流的操作系统桌面环境，并被移植到 UnisLinux 版本。DDE 以便捷的使用为宗旨，充分考虑国内用户的操作习惯，提供了极简的逻辑层级和功能结构。为了保证所有平台用户获得相同且优质的使用体验，DDE 对 PC 用户的心智模型进行了大量的研究，遵循科学和设计美学相结合的设计理念，设计了符合国内用户使用习惯的 UI 风格与窗口交互方式。直观且充满活力的界面风格、便利流畅的系统操作、即时的消息反馈，给用户带来舒适、愉悦的使用体验。

## 3.8 内存分级扩展

当前内存制造工艺已经达到瓶颈，生态发展让每个 CPU 核的成本越来越低。数据库、虚拟机、大数据、人工智能、深度学习场景同时需要算力和内存的支持。内存容量成为了制约业务和算力的问题。

内存分级扩展通过 DRAM 和低速内存介质，如 SCM、AEP，以及 RDMA 远端内存等形成多级内存，通过内存自动调度让热数据在 DRAM 高速内存区中运行，让冷数据交换到低速内存区，从而增加内存容量，保证核心业务高效平稳运行。该特性适用于内存使用量大，且使用相对不频繁的应用进程上，在这些场景中的效果好、收益大，实测等成本条件下 MySQL 性能提升 40%。针对用户态存储框架和用户需求，新增用户态的内存交换机制。