

OpenScaler – передовые открытые технологии из Китая

Дмитрий Варенов

Технический комитет сообщества
OpenScaler

<https://t.me/openscaler>

Дата: 08.10.2024



Российское сообщество разработчиков ОС с открытым исходным кодом

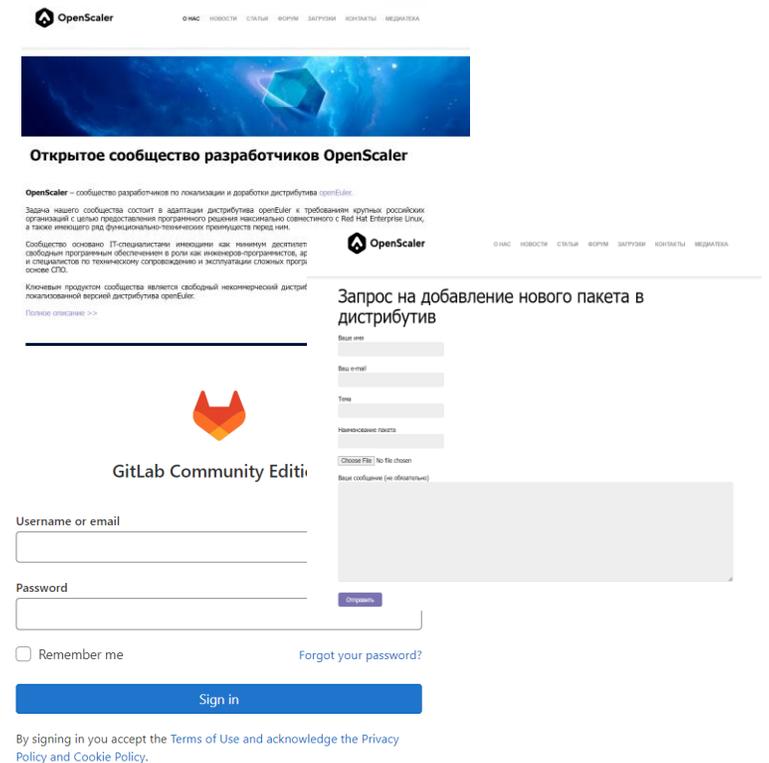


Мы стремимся создать эффективную платформу для разработки современной и надежной отечественной операционной системы с открытым исходным кодом, чтобы обеспечить лидирующие позиции России на мировом рынке информационных технологий

В конце 2022 года было создано российское сообщество разработчиков операционной системы OpenScaler

3

- ▶ В основе решений сообщества OpenScaler лежат решения openEuler – операционной системы №1 в Китае и крупнейшего open-source сообщества разработчиков ОС
- ▶ Мы стремимся обеспечить максимальную совместимость с RedHat Enterprise Linux, а также предоставить ряд функционально-технических преимуществ
- ▶ Основным продуктом сообщества является OpenScaler – свободный некоммерческий дистрибутив операционной системы
- ▶ В сообществе полностью открыты GIT репозитории с исходным кодом программных пакетов
- ▶ Сообщество максимально открыто к идеям и сотрудничеству

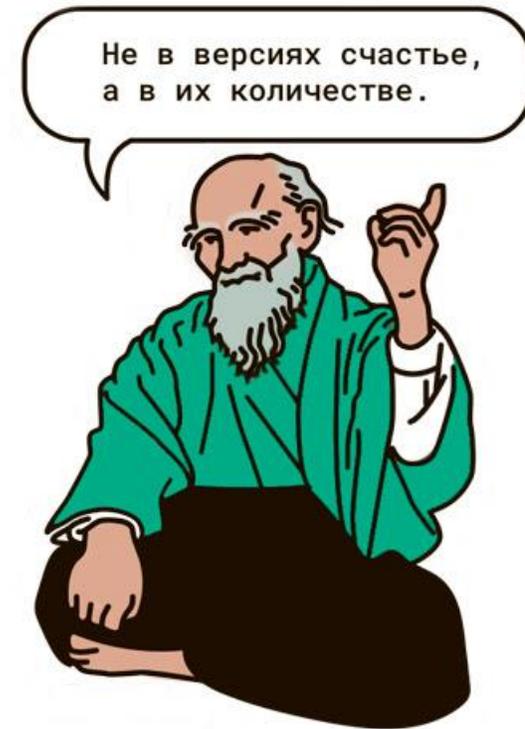


The screenshot displays the OpenScaler website interface. At the top, there is a navigation menu with links for 'О нас', 'Новости', 'Статьи', 'Форум', 'Загрузки', 'Контакты', and 'Медиацентр'. Below the navigation is a blue banner with a 3D cube graphic. The main heading reads 'Открытое сообщество разработчиков OpenScaler'. The page contains introductory text about the community's goals and a form titled 'Запрос на добавление нового пакета в дистрибутив'. The form includes fields for 'Your name', 'Your e-mail', 'Title', 'Package name', 'Choose File' (with a note 'No file chosen'), and a text area for 'Your package (in Russian)'. There are also fields for 'Username or email' and 'Password', along with a 'Remember me' checkbox and a 'Forgot your password?' link. A blue 'Sign in' button is at the bottom of the form. Below the button, there is a small disclaimer: 'By signing in you accept the Terms of Use and acknowledge the Privacy Policy and Cookie Policy.'

Я разрабатываю «отечественную ОС» и у нас все хорошо (?)

Практически все российские дистрибутивы основаны и зависят от кодовых баз зарубежных дистрибутивов – Debian, Ubuntu, CentOS (Alma/Rocky) и др.

- Но не входят в состав технических комитетов и не имеют своих представителей в органах управления сообществом
- Не имеют возможности влиять на путь развития «дистрибутива-родителя»
- Риск получить «трагедию Baidu»
- Технологически сильно зависят от профильных сообществ – больше занимаются «адаптацией для российских реалий» нежели развитием и разработкой

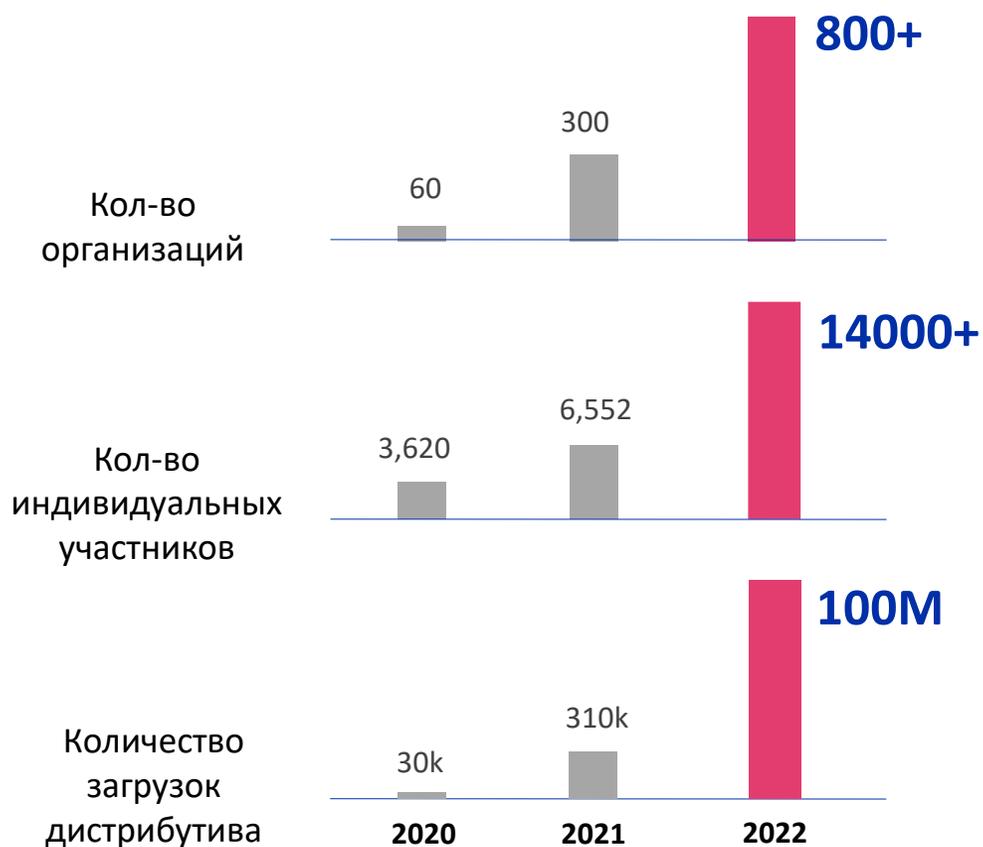


И зачем нам новый Linux и какая-то там кодовая база?

openEuler – операционная система №1 в Китае и крупнейшее open-source сообщество разработчиков ОС

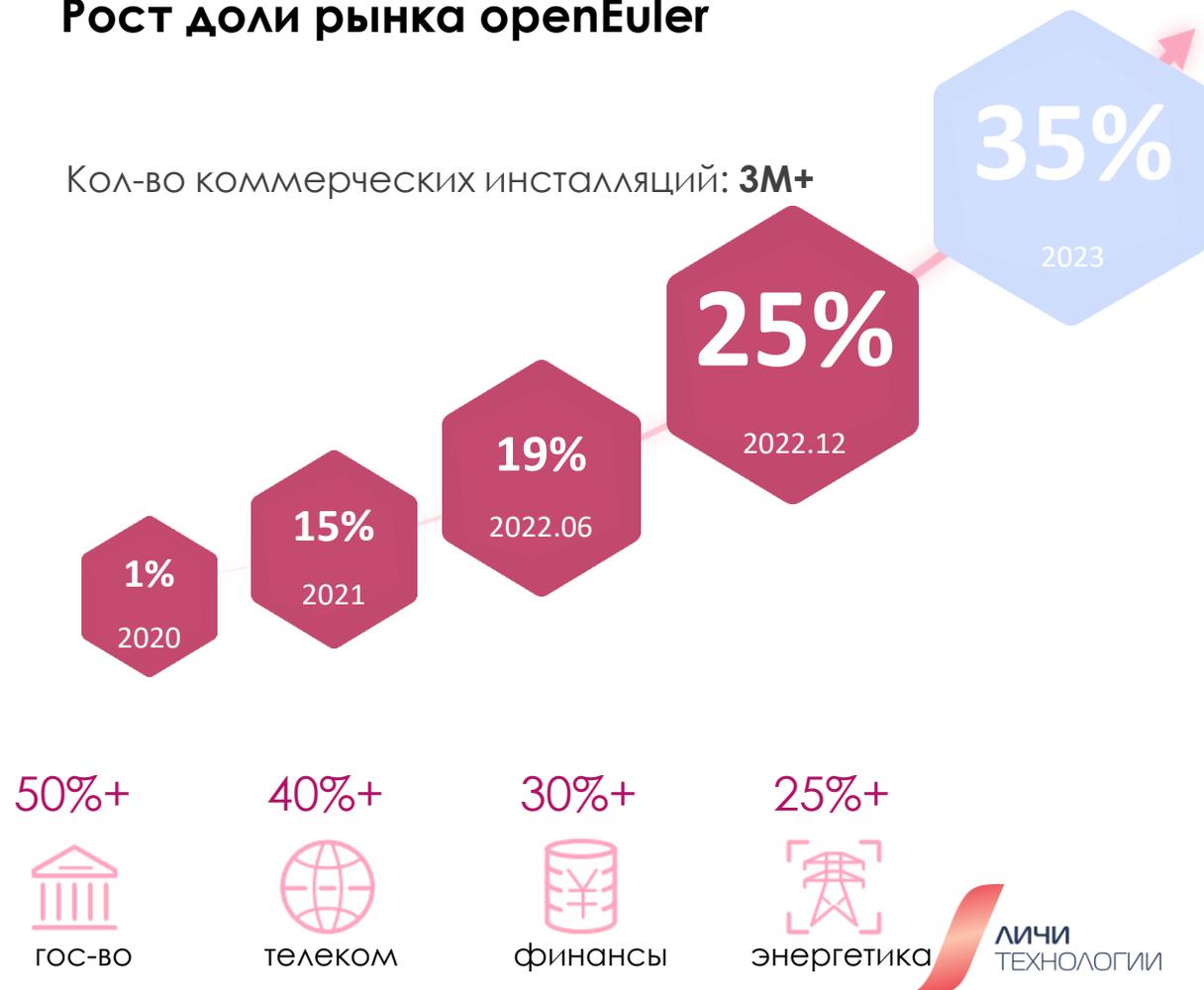
5

Рост сообщества openEuler



Рост доли рынка openEuler

Кол-во коммерческих инсталляций: 3M+

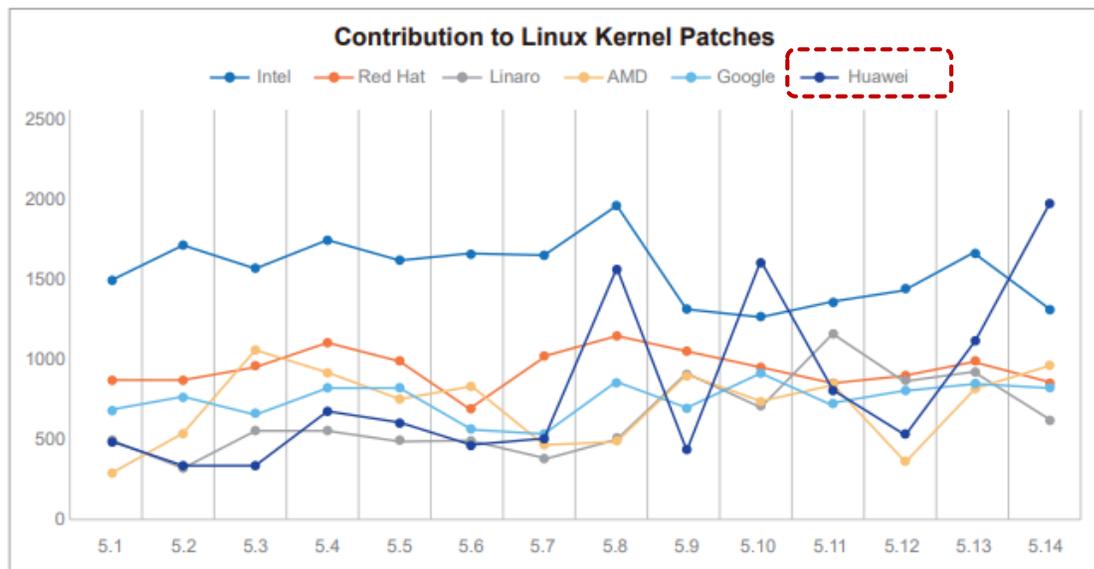


История развития openEuler уже насчитывает 13 лет

Корпоративная ОС для внутреннего использования



Большой технический вклад Huawei в развитие ядра Linux



Переход в Open Source в Июле 2019

全面开源，发展操作系统

- 总体策略：完全开源
- 基于Linux开源项目，持续投资开发EulerOS，满足华为云及其它ICT基础设施解决方案需要
- 基于EulerOS内部自版本，推出社区发行版
- 免费支持各行业主流应用和软件开发商把软件和应用迁移到基于鲲鹏的EulerOS
- 支持OS发行商服务客户

20+ коммерческих дистрибутивов



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud/Embedded



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Server/Cloud



Edge



Server/Cloud/Embedded



Server/Cloud



Server/Cloud



Embedded



Server/Cloud

Не в версиях счастье,
а в их количестве.



Крупнейшие компании Китая в различных отраслях используют операционные системы на базе openEuler

8



China Mobile



China Telecom



China Unicom



State Grid



Southern Power Grid



China Construction Bank



China UnionPay



CITIC Bank



Baidu Intelligent Cloud



China Post

Почему OpenScaler?



9

Linux корпоративного уровня «с востока», как зрелая альтернатива Redhat/CentOS, высочайшие стандарты безопасности и чистоты кода

Архитектурная и бинарная совместимость с Redhat/CentOS

Нативная поддержка китайского аппаратного обеспечения

Лучшая совместимость с перспективными отечественными чипами на ARM и RISC-V

Единая ОС для всей вычислительной инфраструктуры предприятия (сервера, рабочие станции, edge, IoT и др)

Тесное взаимодействие с крупнейшими open-source фондами и поддержка их ключевых проектов (OpenStack, Hadoop, Spark и др)

Эффективная площадка для экспансии в Китай и на другие «дружественные» рынки



"Разработка софта на основе открытого кода, мобильных операционных систем, а также запуск центров обработки данных станут приоритетными направлениями совместной работы России и Китая в IT-сфере".

М.В.Мишустин



2022 - 2023: Российское сообщество OpenScaler

10

2022

- Создано локальное сообщество OpenScaler
- Официально выпущены 6 комьюнити-версий ОС для x86 и для ARM
- Начало разработки комм. версии ОС на базе OpenScaler
- Достигнуты договоренности о запуске пилотных проектов с рядом крупнейших клиентов

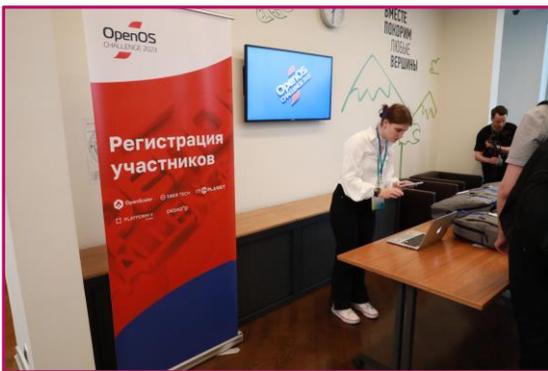


2023

- Весной 2023 года завершен образовательный курс с общим охватом 350+ студентов на базе НГУ и НГТУ
- На базе дистрибутива OpenScaler прошел всероссийский конкурс среди разработчиков системного ПО OpenOS Challenge, который привлек 1200+ участников
- Достигнуты соглашения о совместной работе с крупнейшими вендорами ОС и СИ Китая
- Запланировано 50+ активностей по развитию экосистемы



Ежегодные конкурсы Open OS Challenge и Open Education Cup



OpenScaler – в учебном процессе ВУЗов

12

- OpenScaler применяется в учебном процессе в:
 - Институте вычислительной математики и математической геофизики (ИВМиМГ СО РАН)
 - Новосибирском государственном университете
 - Новосибирском государственном техническом университете
- OpenScaler демонстрируется студентам в рамках курсов по операционным системам, разъясняются особенности решения.
- OpenScaler предлагается к установке для выполнения заданий по разработке программного обеспечения в рамках лабораторных работ, семестровых проектов в университетах в ходе курсов по организации высокопроизводительных вычислений, по операционным системам, а в ИВМиМГ СО РАН в рамках практики студентов.



Новосибирский государственный
технический университет

НЭТИ



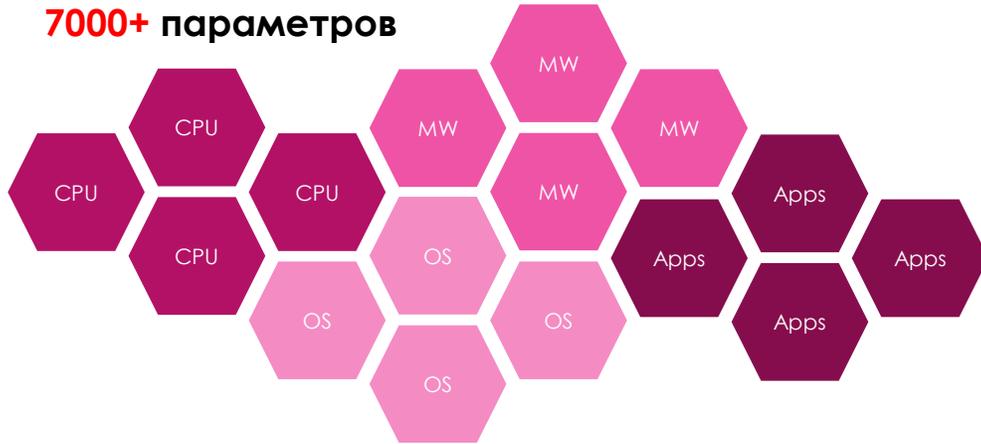
N* Новосибирский
государственный
университет
***НАСТОЯЩАЯ НАУКА**



A-Tune: оптимизация производительности с помощью ИИ

13

7000+ параметров



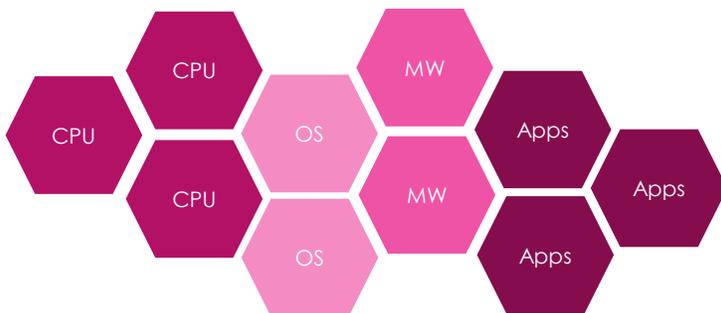
Определение оптимальных значений и комбинации параметров на разных уровнях стека (оборудование, ОС, промежуточный уровень, приложение)

- **Статическая оптимизация:** анализ логов и подбор профиля настройки из предустановленных
- **Динамическая оптимизация:** выбор значимых параметров, подбор оптимальной конфигурации

Статическая
оптимизация



Динамическая
оптимизация



20+%
Postgres

10+%
Terasort, HDFS

10+%
HBase, HIVE, Spark

Поддержка nVidia GPU: Тестирование pyTorch AI Model на nVidia GPU

14

- **Тестовое окружение:**

- Тестирование проводилось на задаче многоклассовой классификации изображений из датасета CIFAR100 с помощью сверточной сети ResNet18 и transfer learning.

- **OS:** openScler 22.03 LTS SP1 (Linux 5.10.0-136.12.0.86.os2203sp1.x86_64)

- Nvidia GPU driver and CUDA toolkit version 12.1.1_530.30.02

```
NVIDIA-SMI 530.30.02          Driver Version: 530.30.02    CUDA Version: 12.1
+-----+-----+-----+-----+-----+
| GPU  Name                Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf           Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
|=====+=====+=====+=====+=====+
|  0   Tesla V100-PCIE-16GB      Off          | 00000000:5A:00:0 Off |            0         |
| N/A   57C    P0              149W / 250W |  3474MiB / 16384MiB |      58%    Default |
+-----+-----+-----+-----+-----+
|                                     | MIG M.         |
+-----+-----+-----+-----+-----+
|                                     | N/A           |
+-----+-----+-----+-----+-----+
```

```
val loss: 0.0003, val acc: 0.4127
Epoch 12/49
val loss: 0.0003, val acc: 0.4306
Epoch 13/49
val loss: 0.0003, val acc: 0.4515
Epoch 14/49
val loss: 0.0003, val acc: 0.4710
Epoch 15/49
val loss: 0.0003, val acc: 0.4865
Epoch 16/49
val loss: 0.0002, val acc: 0.5030
```

Kernel Live-Update: Смена ядра ОС без перезагрузки физического сервера или виртуальной машины

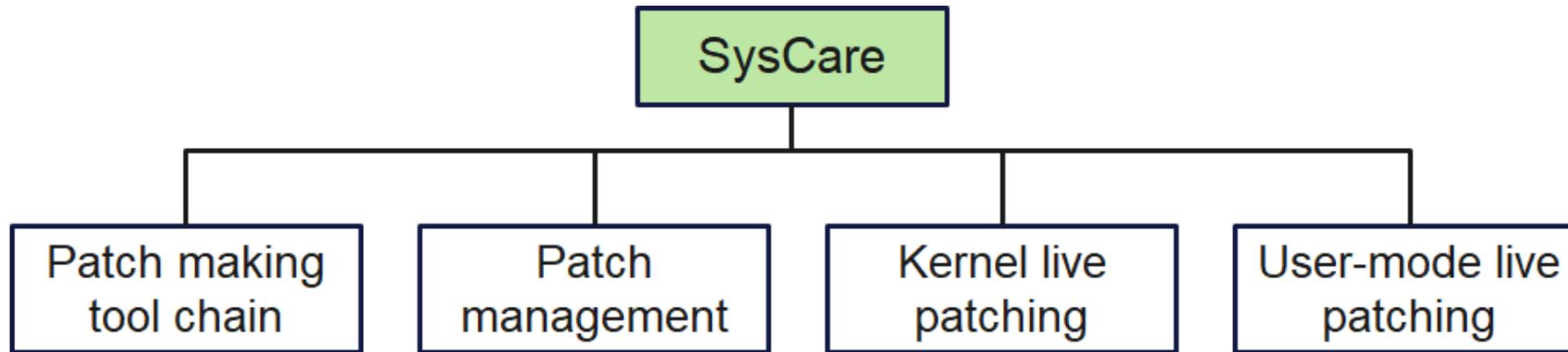
15



- Смена ядра ОС без физической перезагрузки сервера
- Возможность «заморозить» состояние отдельных процессов (при указание юнита system, PID)
- Поддержка архитектур ARM / X86

SysCare – Применение устранения уязвимостей

16



Возможности утилиты:

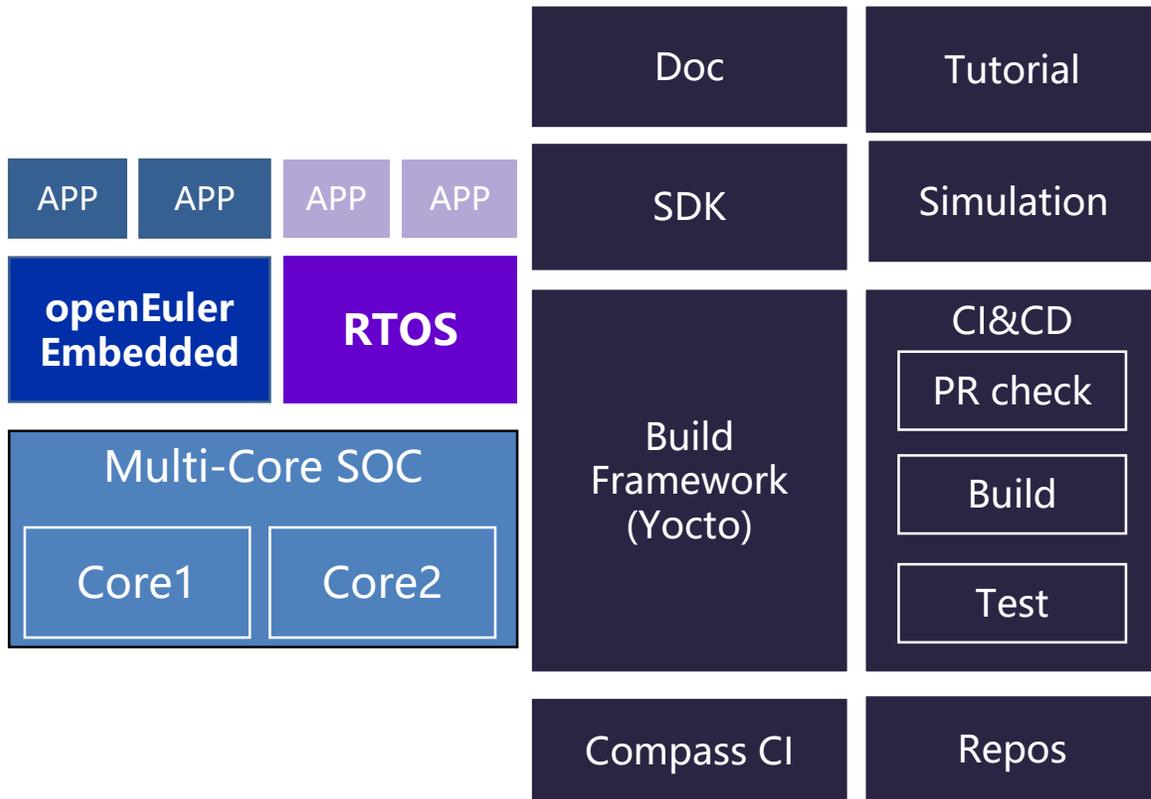
1. Простое создание патчей, как для ядра, так и для уровня пользователя.
2. Интерфейс для управления патчами, включая их установку, активацию, деактивация и удаление.

Используются следующие возможности:

1. Унифицированные патчи: SysCare скрывает разницу в деталях создания патча, предоставляя унифицированный инструмент для улучшения эффективности эксплуатации и обслуживания.
2. Применение патчей на лету для приложения уровня пользователя: SysCare поддерживает наложение патчей на многопоточные и много-процессные приложения в пространстве пользователя, что позволяет достичь эффекта без перезапуска процесса или потока.
3. Ленивый механизм: SysCare отлавливает все системные вызовы через ptrace и ждёт их завершения, что позволяет повысить эффективность успешного применения.

ОС для Embedded-устройств

17



- Возможность создания легковесных дистрибутивов для встраиваемых устройств
 - Поддержка фреймворка Yocto, создание дистрибутива ОС для любого типа устройств размером от 5 МВ
 - Наличие механизмов оптимизации дистрибутива и времени запуска
- Поддержка широкого спектра оборудования, включая SoC-решения на ARM, такие как Raspberry Pi и аналоги
- Наличие возможности использования ядра «реального времени» и RTOS (Real time Operating system)
- Наличие Distributed soft bus (DSoftBus) для взаимодействия встраиваемых и мобильных устройств между собой
- Наличие 80+ апробированных и готовых к использованию пакетов в специализированных дистрибутивах

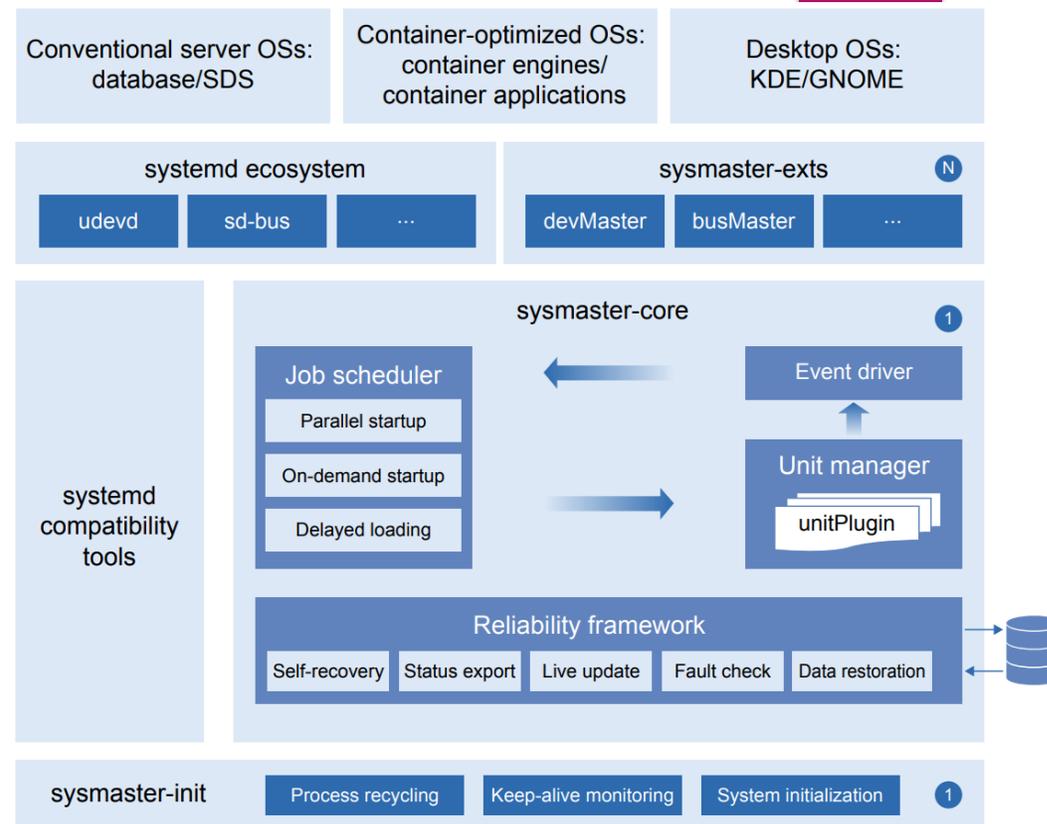
SysMaster: минималистичная система инициализации для встраиваемых устройств

18

Минималистичная, ультра-лёгкая система инициализации, призванная составить конкуренцию systemd на embedded устройствах, в облачных вычислениях и даже серверах.

sysMaster состоит из трех компонентов:

- **sysmaster-init**, новая реализация PID 1, применима к встраиваемым системам с такими функциями, как инициализация системы, мониторинг поддержания работоспособности сервисов.
- **sysmaster-core** выполняет основные функции управления сервисами и включает в себя надёжную платформу, обеспечивающую оперативное обновление и быстрое самовосстановление в случае сбоев, обеспечивая доступность сервиса в режиме 24/7.
- **sysmaster-exts** предлагает набор компонентов (таких как **devMaster** для управления устройствами и **busMaster** для обмена данными по шине), которые обеспечивают ключевые системные функции. Вы можете выбрать компоненты для использования по мере необходимости.



DevMaster: диспетчер устройств (аналог system-udev)

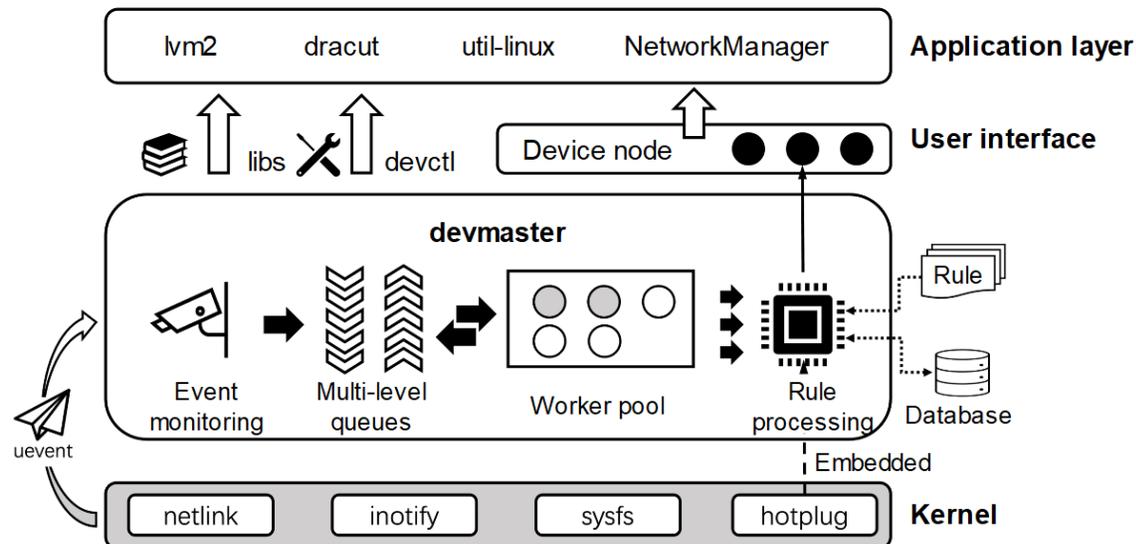
19

Диспетчер устройств является связующим звеном между программным обеспечением работающем в пространстве пользователя и базовыми физическими устройствами.

Как программное обеспечение, devmaster состоит из демона, клиентского инструмента и динамической библиотеки.

Демон devmaster использует механизмы ядра, такие как netlink, notify и sysfs, для мониторинга событий устройства и запуска задач обработки правил.

Клиентский инструмент devctl предоставляет набор команд CLI и общедоступных интерфейсов для отладки правил, управления демонами и запроса состояния устройства.



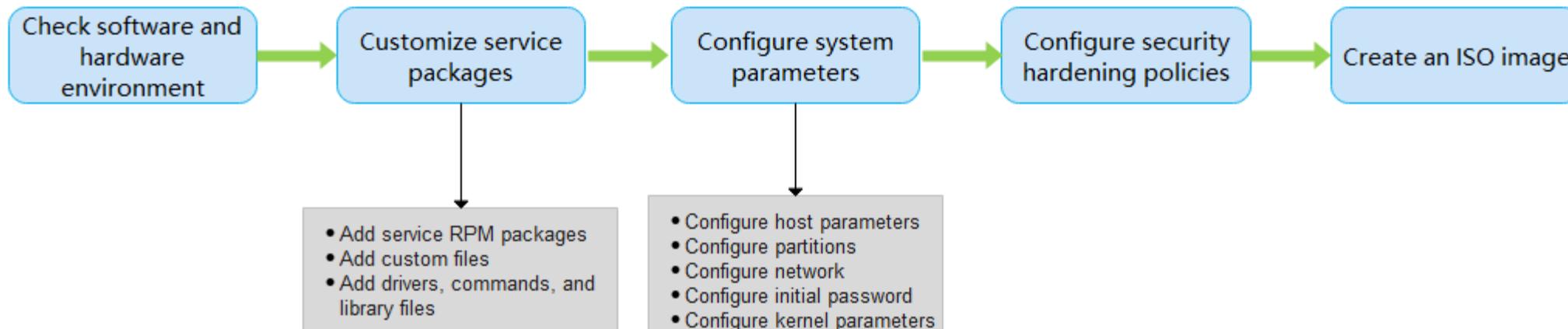
ImageTailor: Изменяем установочный образ

20

Специально для того чтобы пользователи могли самостоятельно видоизменять установочный образ в дистрибутиве OpenScaler представлен инструмент именуемый imageTailor.

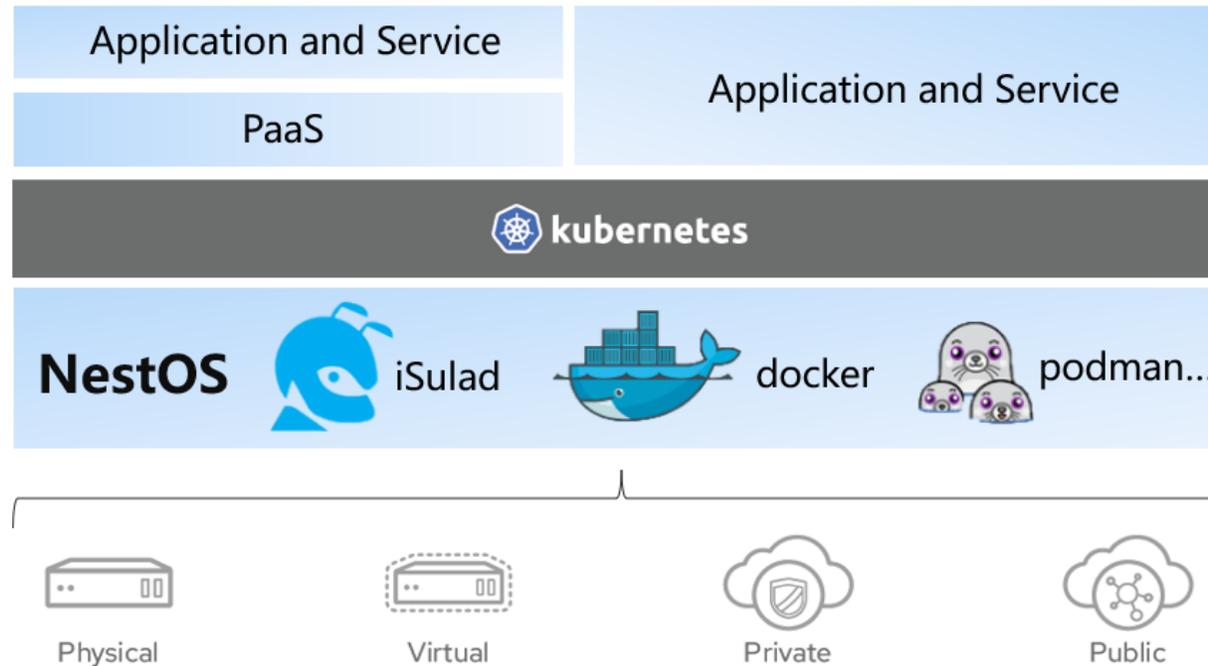
В числе его ключевых функций стоит отметить следующее:

- Настройка системных пакетов: настройка системных команд, библиотек и драйверов на основе списка устанавливаемых пакетов RPM.
- Изменение конфигурации системы: настраивает имя хоста, службы запуска, часовой пояс, сеть, разделы, загружаемые драйверы и версию ядра.
- Добавление пакета программного обеспечения: добавляет в систему пользовательские пакеты RPM или файлы.



NestOS: Минималистичная ОС для узлов контейнерной виртуализации

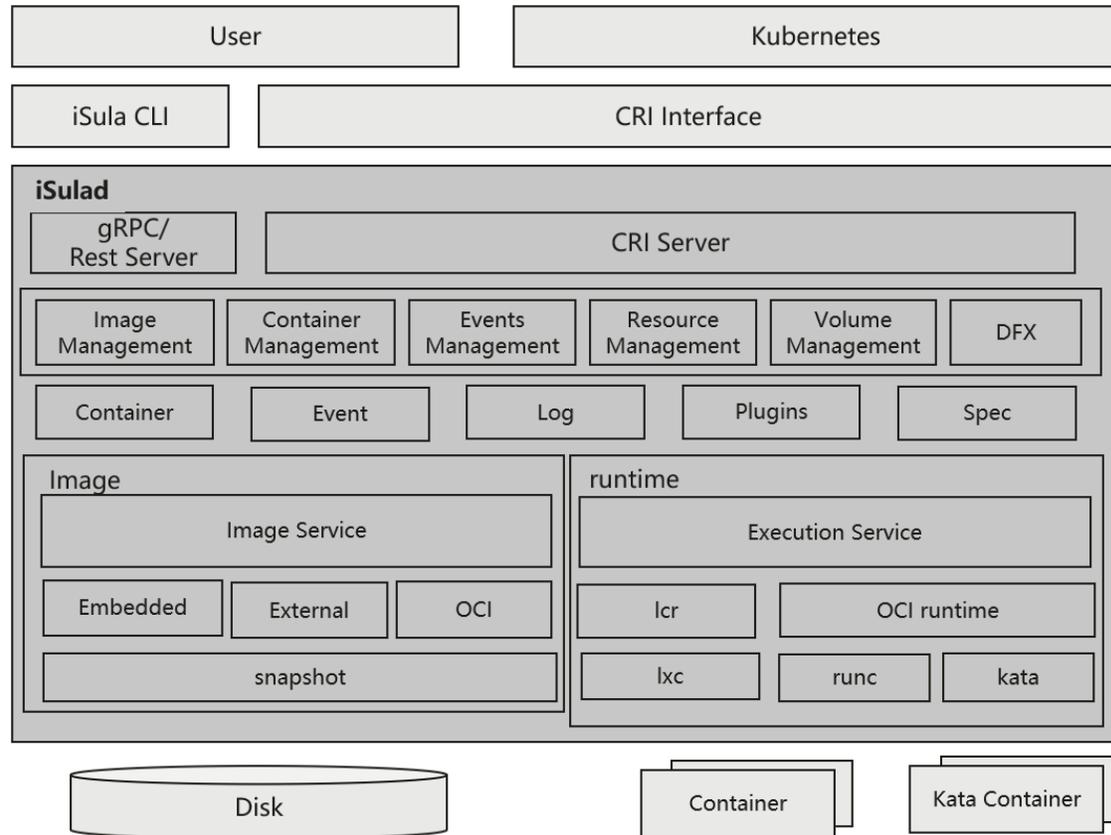
21



- Содержит минимально необходимый набор компонентов
- Базовая ОС для запуска кластера k8s
- Для записи доступны только разделы /etc, /var. Остальные разделы доступны только на чтение.
- Удобная система обновления (используются образы системы, обновляемые при помощи rpm-ostree).
- Наличие инструментов для создания своих сборок NestOS, содержащих сразу необходимый набор софта или конфигурационных файлов
- Поддержка "из коробки" среды выполнения контейнеров iSulad

iSula: легковесные контейнеры

22



iSula разработан на C/C++ и использует меньше вычислительных ресурсов для оптимизации производительности

- Обеспечивает совместимость со стандартами kubernetes CRI и сетевыми плагинами CNI
- Поддерживает изоляцию namespace и Cgroup v1/v2
- Поддерживает application container, system container и secure container
- Поддержка Seccomp, технологии SELinux Security-Enhanced
- Соответствует спецификации OCI (на базе среды выполнения lxc)
- Поддержка встроенных и внешних форматов образов, а также спецификаций образов OCI и образов докеров.

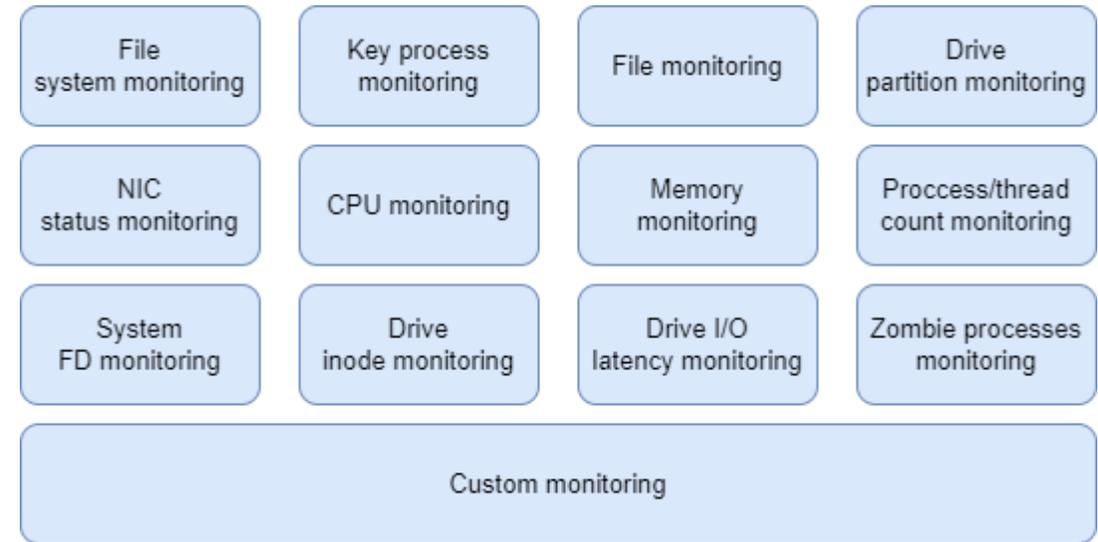
Sysmonitor — облегчённая система мониторинга сервера

Данный проект демонстрирует минималистский подход, по сути сводя всю систему мониторинга к запущенному процессу sysmonitor и сообщениям в файле лога /var/log/sysmonitor.log

23

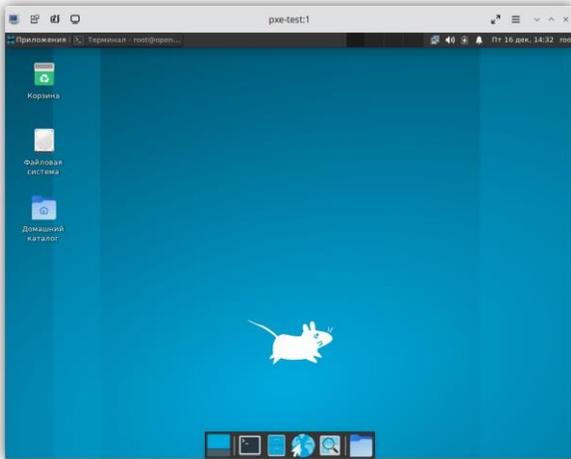
возможности sysmonitor покрывают следующие виды мониторинга:

- ошибки при работе файловых систем;
- ключевых процессов (если потребуется, у sysmonitor есть возможность выполнить команды для восстановления процесса);
- слежения за важными файлами и фиксации в лог событий их удаления или изменения;
- использования разделов диска;
- статусов сетевых адаптеров;
- использования CPU (при этом можно следить за конкретным ядром);
- использования памяти;
- количества процессов/потоков;
- файловых дескрипторов;
- количества inode смонтированных файловых систем;
- задержки i/o дисковой подсистемы;
- количества зомби процессов;
- Также sysmonitor позволяет создавать собственные варианты мониторинга, по параметрам, которые выберет сам системный администратор.

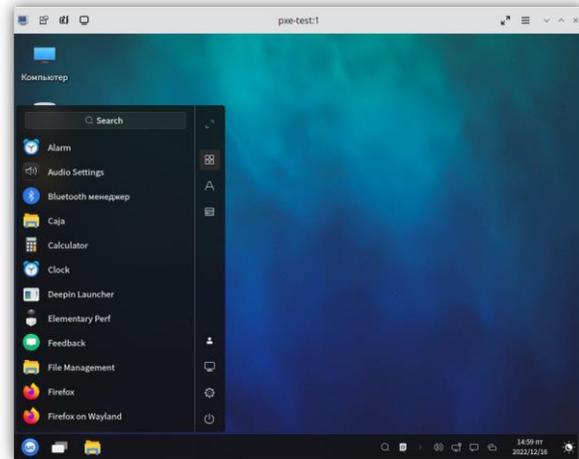


Пять популярных графических окружений

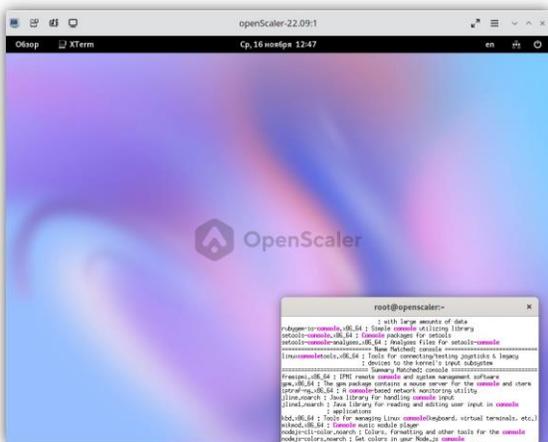
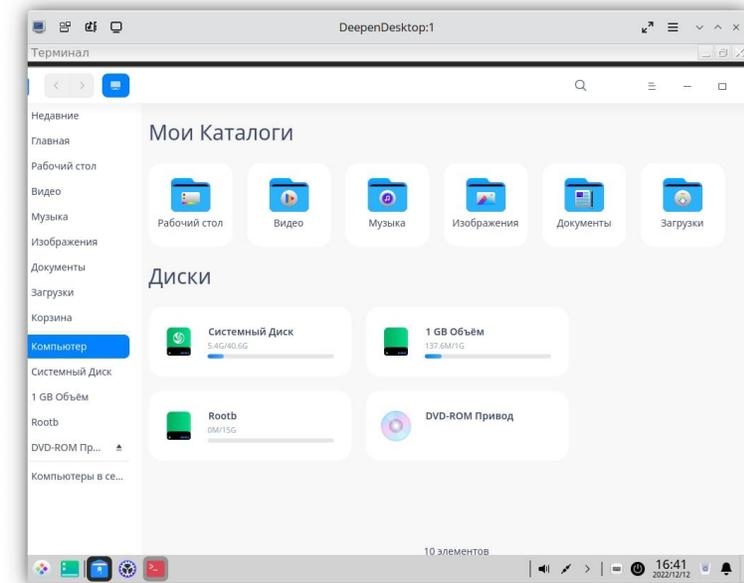
24



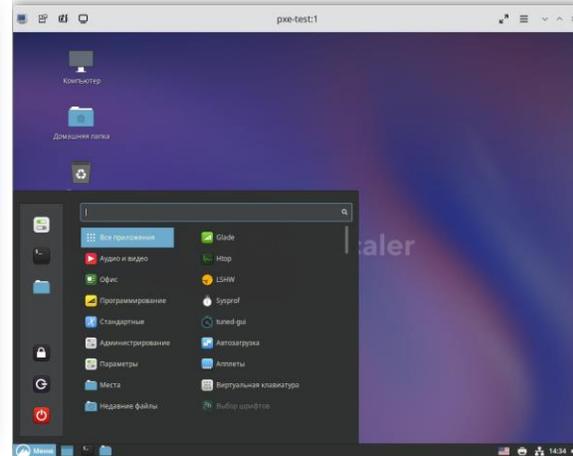
XFCE



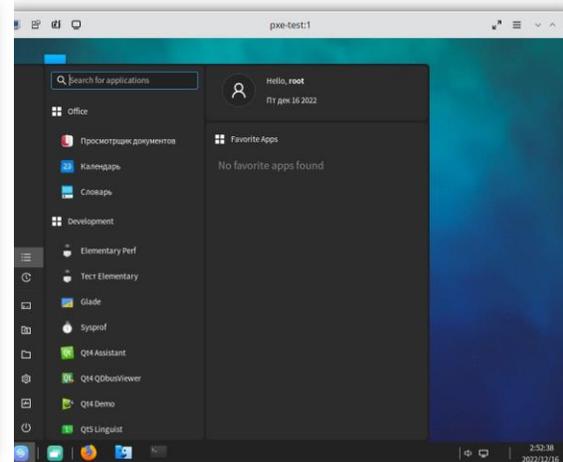
UKUI-Desktop



GNOME



Cinnamon (RU)



Kirian-Desktop (RU)

OpenScaler – в Yandex Cloud

Yandex  Cloud

Поиск

[Сервисы](#) [Решения](#) [Почему Yandex Cloud](#) [Ресурсы](#) [Тарифы](#) [Документация](#) [Блог](#)

← [Marketplace](#)

ОпенСкейлер

Обновлено 23 ноября 2023 г.

Дистрибутив OpenScaler включает следующие передовые программные наработки:

- A-Tune — систему для автоматической оптимизации настроек с помощью механизма машинного обучения. С помощью технологий искусственного интеллекта подбираются оптимальные параметры конфигурации операционной системы для повышения общей эффективности работы системы в соответствии с рабочей нагрузкой;
- собственный стек легковесной виртуализации (StratoVirt) и контейнеризации (iSulad);
- инструмент secPaver для настройки политик информационной безопасности SELinux;
- Kernel Live update — обновление ядра ОС без перезапуска виртуальной машины или физического сервера
- NFS MutiPath — Предлагается установить несколько каналов связи между клиентом и сервером в рамках одной точки монтирования для поддержки передачи данных ввода-вывода по нескольким каналам, повышая производительность одной точки монтирования. Кроме того, периодически проверяется состояние соединения, чтобы обеспечить быструю отработку отказа ввода-вывода при сбое соединения.
- SysCare — технология SysCare позволяет произвести установку патча для процесса пользователя без его перезапуска.

25

Первая коммерческая версия "Скейлер" в реестре Минкомсвязи РФ

26

Скейлер

Сведения обновлены 27.11.2023

Реестровая запись №20010 от 27.11.2023

Произведена на основании поручения Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации от 27.11.2023 по протоколу заседания экспертного совета от 13.11.2023 №749пр

Правообладатели программного обеспечения

Полное наименование (коммерческая организация без преобладающего иностранного участия)

Идентификационный номер (ИНН)

Государство регистрации в качестве юридического лица

9717098243

Россия

ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "СКАЛА-R"

Общие сведения

История изменений

Описание программного обеспечения

Класс программного обеспечения по классификатору, утвержденному приказом от 22.09.2020 № 486

Основной

02.09 Операционные системы общего назначения

Коды продукции

Коды продукции в соответствии с Общероссийским классификатором продукции по видам экономической деятельности

58.29.11 Системы операционные на электронном носителе

Информация об уязвимостях

Данная информация содержится в Банке данных угроз безопасности информации ФСТЭК России



Уязвимости отсутствуют

OpenScaler – Учебник по администрированию

27

 ЛИЧИ
ТЕХНОЛОГИИ

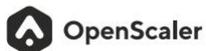


КУРС ОБУЧЕНИЯ СИСТЕМНОГО АДМИНИСТРАТОРА

дистрибутива

OpenScaler Linux

Справочная книга
Версия 1.0



Компанией Личи Технологии с содружестве с сообществом разработчиков OpenScaler был выпущен первый учебный курс по работе с системой на уровне системного администратора

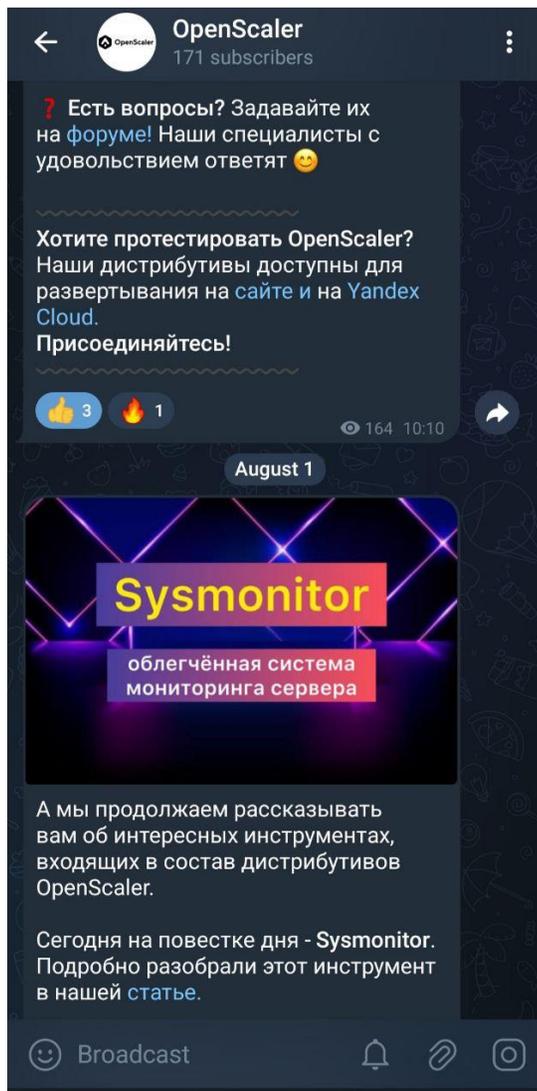
Курс состоит из двух свободно доступных книг:

- Теоретическая часть в виде справочной книги
- Практический сборник заданий

На базе платформы проведения онлайн-тестирований ИТ-Планета была запущена официальная экзаменационная сертификация на основе данного курса

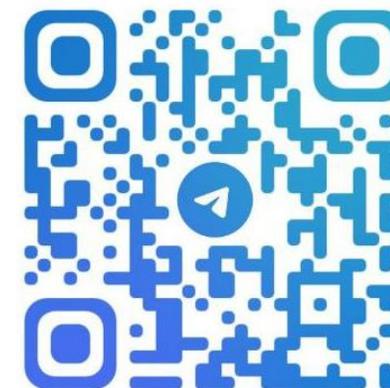
 ЛИЧИ
ТЕХНОЛОГИИ

Много полезного вы найдете в нашем Телеграмм канале



- Технические статьи
- Мероприятия
- Новости
- Интересные материалы

Подписывайтесь!



@OPENSCALER

Спасибо!

29



openEuler

<https://openeuler.org>



OpenScaler

<https://openscaler.ru/>



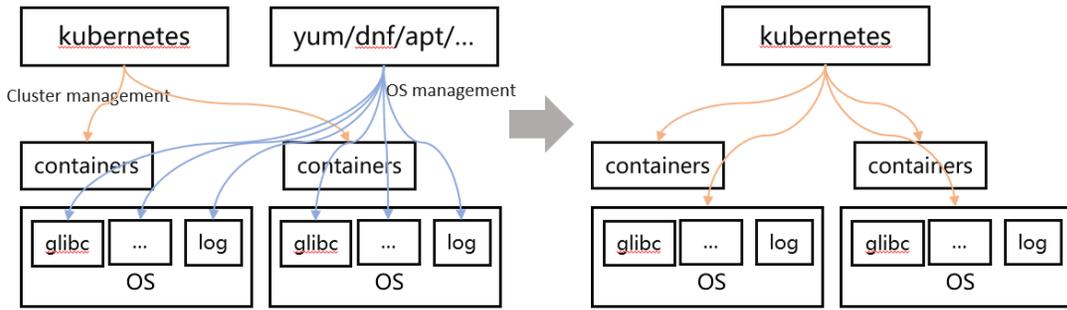
LICI
TECHNOLOGIES

<https://www.lichi-tech.ru/>

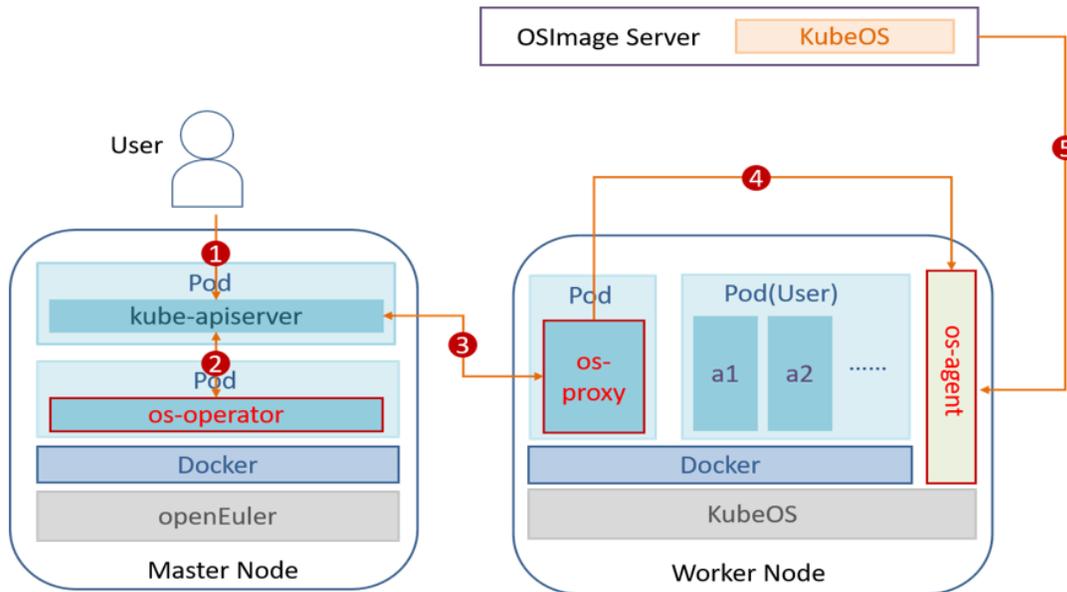
<https://t.me/openscaler>

KubeOS: упрощение управления обновлениями ОС

30



- Используйте API Kubernetes для управления ОС узлов кластера.
- Подключение ОС к Kubernetes через CRD + Operator
- Управление ОС через утилиты управления Kubernetes
- Упрощение обновления и отката ОС



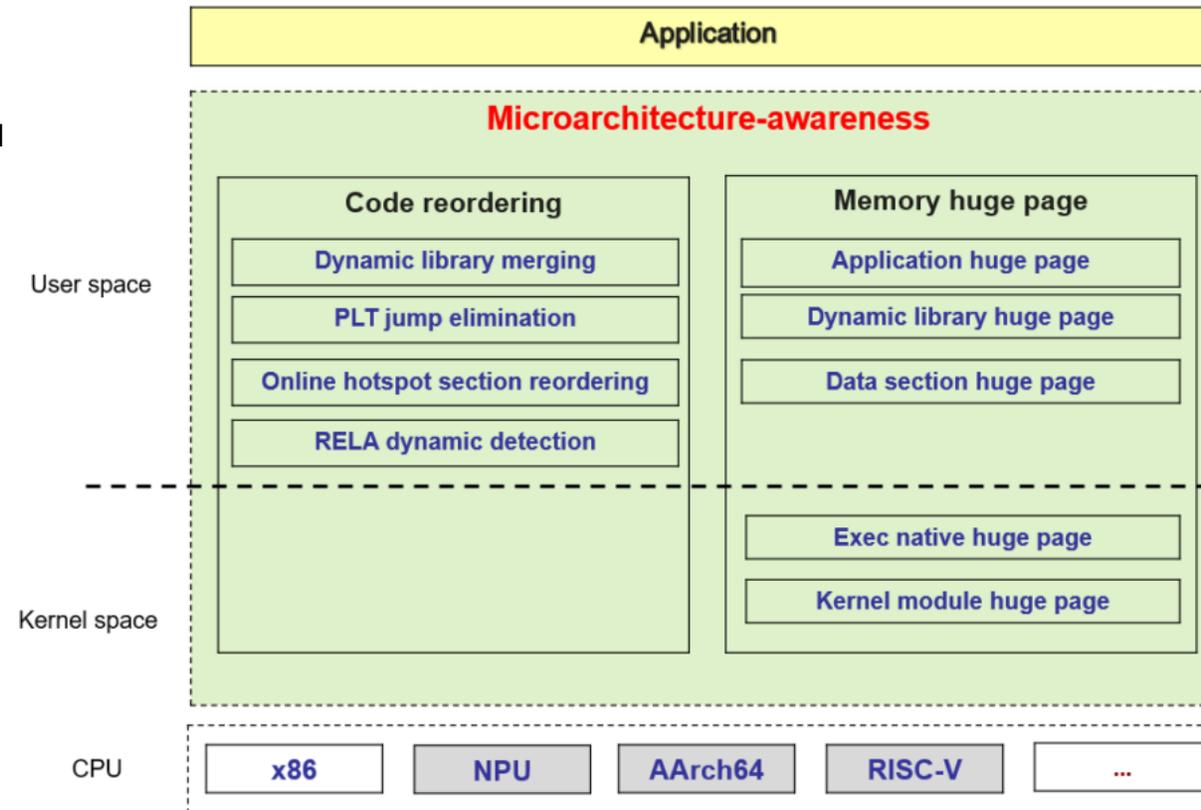
sysBoosts: ускоряем процесс запуска программ

31

sysBoost производит переупорядочивание исполняемого кода динамических библиотек, для повышения производительности программы, путем уменьшения количества сегментов кода и количества кеш-промахов.

В основе работы sysBoosts используются следующие технологии:

- **Динамическое слияние библиотек:** разрозненные сегменты кода и сегменты данных объединяются, когда динамический загрузчик загружает динамические библиотеки. Huge page memory используется для улучшения коэффициента попадания iTLB.
- **Устранение перехода PLT:** когда код приложения вызывает функцию динамической библиотеки, выполнение перенаправляется в PLT, а затем в фактическую функцию. Устранение перехода PLT может улучшить IPC.
- **Онлайн-переупорядочивание сегментов кода горячих точек:** по умолчанию код упорядочивается динамической библиотекой. Технология онлайн-переупорядочивания может переупорядочивать код горячих точек по сегментам.
- **Механизм exec Huge Page:** Механизм Huge Page пользовательского режима требует определенной конфигурации приложения и перекомпиляции. Механизм exec Huge Page напрямую использует Huge page memory, когда ядро загружает файл ELF, без необходимости изменения приложений.



Gazelle: повышаем пропускную способность сетевого ввода-вывода приложений

32

Gazelle — это высокопроизводительный стек протоколов для пользовательского режима.

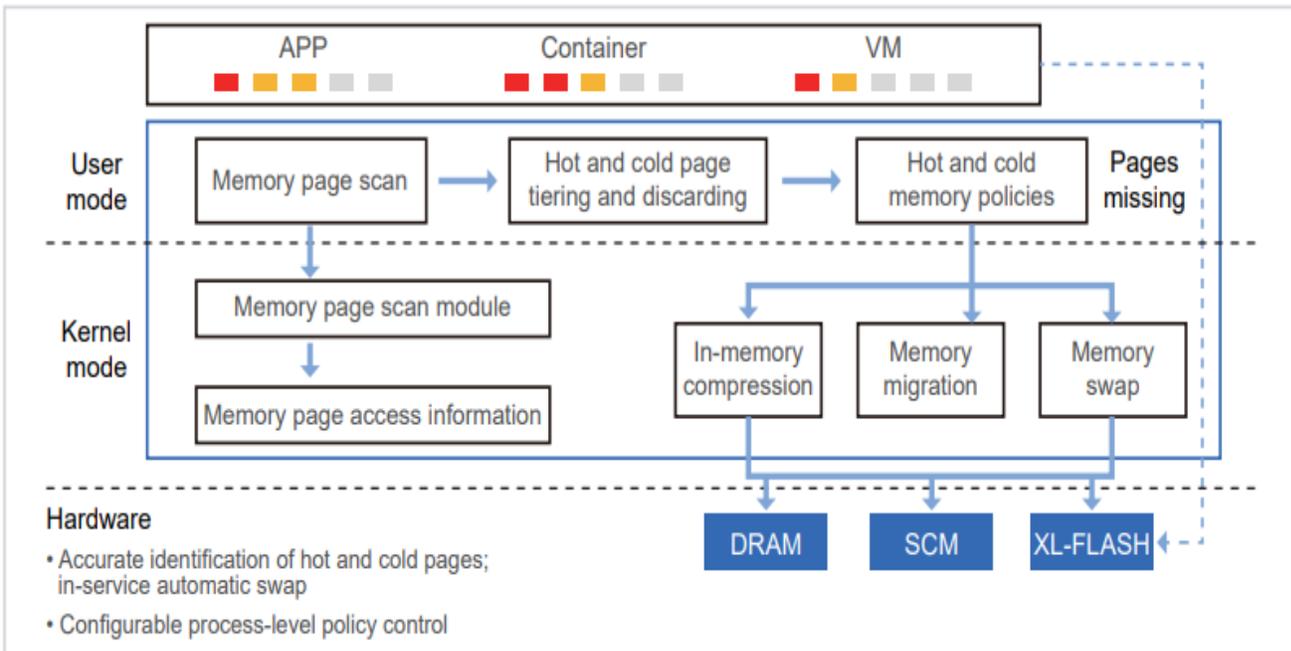
Он напрямую считывает и записывает пакеты сетевого интерфейса NIC в пользовательском режиме на основе DPDK (Data Plane Development Kit, набор для разработки на уровне данных, состоящий из библиотек для ускорения рабочих нагрузок по обработке пакетов) и передает пакеты через память HugePages memory и использует стек протоколов LwIP (Lightweight TCP/IP, небольшая независимая реализация пакета протоколов TCP/IP).

Gazelle значительно повышает пропускную способность сетевого ввода-вывода приложений и ускоряет работу сети с базами данных, такими как MySQL и Redis. Его ключевые особенности и преимущества:

- Высокопроизводительные пакеты с нулевым копированием и без блокировки, которые можно гибко масштабировать и адаптивно планировать.
- Универсальность, совместимая с POSIX без каких-либо изменений и применимая к различным типам приложений.

eMem: новый взгляд на раздел подкачки

33



- DRAM составляет более 40% стоимости серверов центров обработки данных
- Разрыв в производительности ЦП и DRAM увеличивается: за 10 лет количество ядер ЦП увеличилось в 7 раз, а объем памяти увеличился только в 4 раза
- eMem хорошо подходит для приложений, использующих большой объем ОЗУ, но при ограниченных обращениях к ней к примеру, таких как MySQL, Redis, или Nginx
- Особенности:
 - Механизм контроля: за счет наличия возможности настройки по сравнению со стандартным kswap
 - Механизм сканирования: сканирует процессы в пользовательском и привилегированном режиме.
 - Вытеснение "горячих"/"холодных" данных.
 - Политика вытеснения: гибкие политики вытеснения данных.