



JUNIPER
NETWORKS

JUNIPER DAY

“Эволюция операционной
системы Junos”

16 октября 2018 | Москва

Ястребов Валерий
vyastrebov@juniper.net

LEGAL STATEMENT

This statement of direction sets forth Juniper Networks' current intention and is subject to change at any time without notice.

No purchases are contingent upon Juniper Networks delivering any feature or functionality depicted in this presentation.

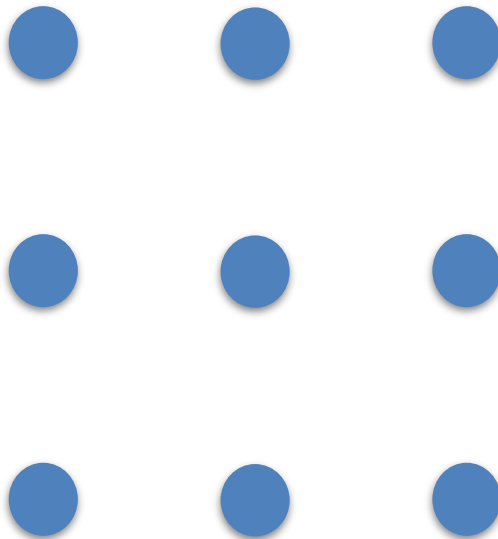
Содержание презентации

- Изменения в архитектуре JUNOS
- Архитектура JUNOS EVO
- Демонстрация возможностей
 - Визуализация JUNOS объектов в БД
 - In Service Software Upgrade (ISSU)
 - Распределенный узел

Разминка для ума

Условие:

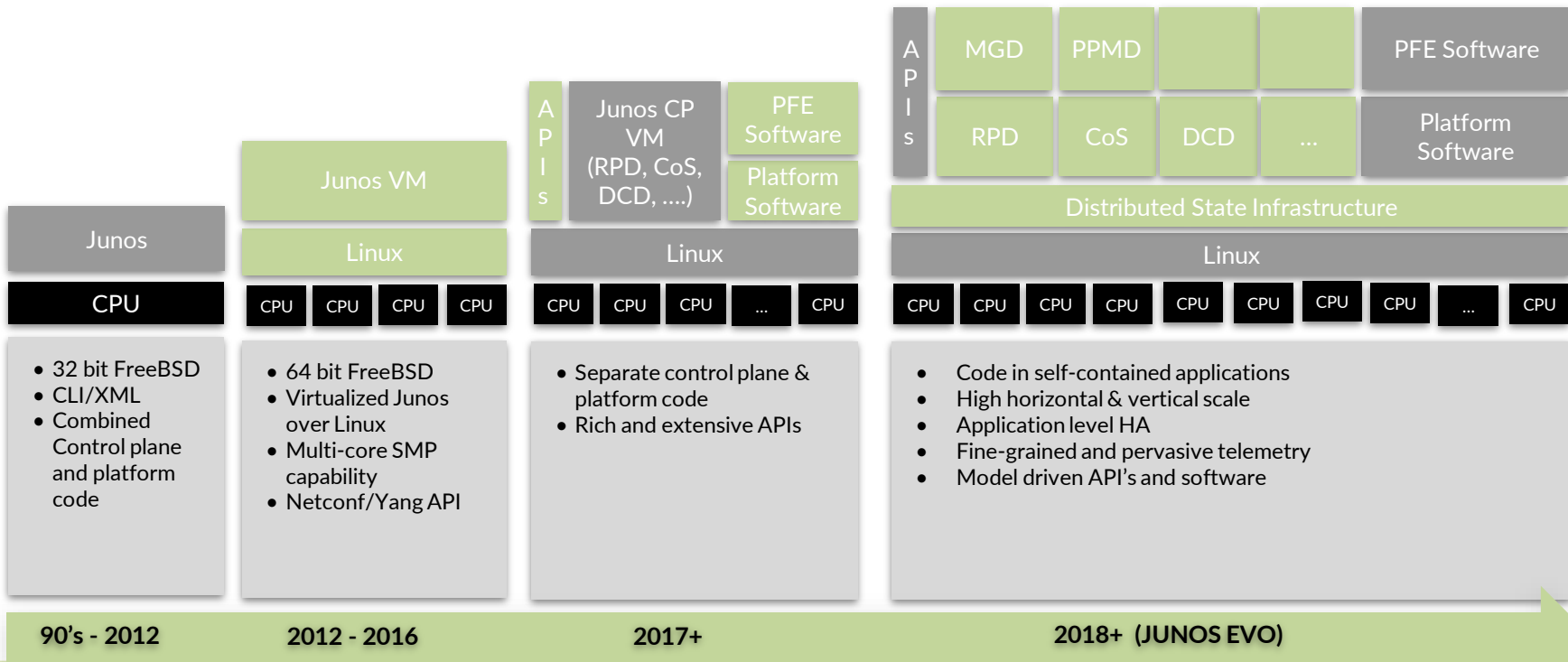
Нужно соединить нарисованные девять точек четырьмя прямыми линиями не отрывая ручки от листа бумаги



ИЗМЕНЕНИЯ В АРХИТЕКТУРЕ JUNOS И РЕЛИЗАХ

Эволюция Junos

Towards Open, Scalable and Flexible Software Architecture



Почему мы должны беспокоиться о архитектуре OS ?

Качество продукта

- Rich visibility ⇒ Лучшая изоляция
- Fast fix ⇒ Лучший MTTR
- Lean options ⇒ Исключение влияния не используемых опций
- Resilience by design ⇒ Лучший fault isolation

Быстрота развертывания продукта и сервисов

- Portability ⇒ Использование тех же самых инструментов в экосистеме
- Upgradeability ⇒ Включение нового функционала без downtime
- Testability ⇒ Предварительное тестирование перед включение в production
- Modularity ⇒ Постепенные интерактивные upgrades

Инновации

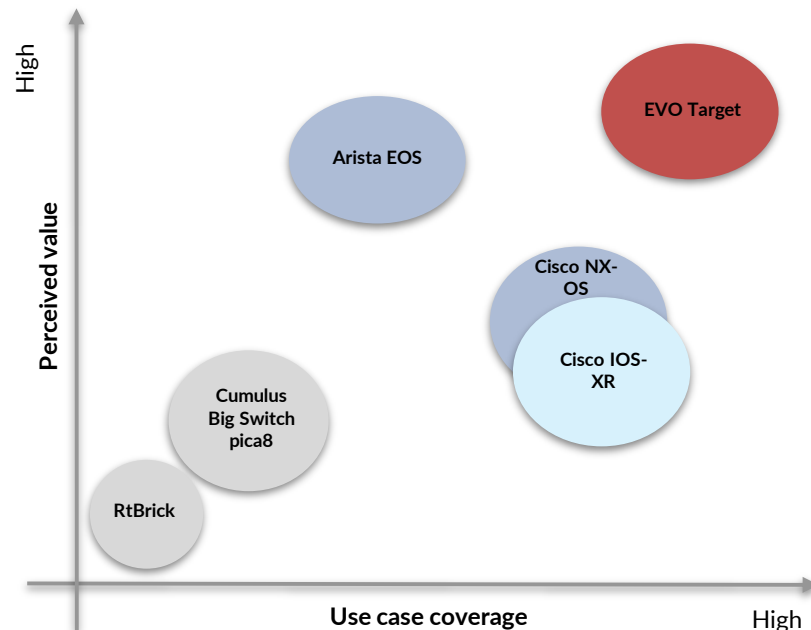
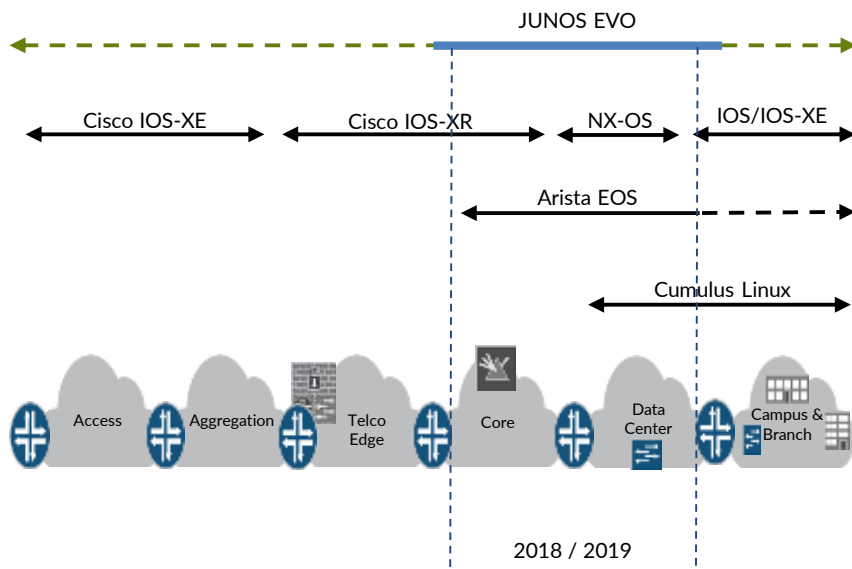
- Open system ⇒ Добавление собственных (3x) приложений
- Extensibility ⇒ В будущем использование новых возможностей
- Developer Documentation ⇒ Документация

Универсальность

- Homogeneity ⇒ Однородность приложений и средств
- Versatility ⇒ Универсальность платформы для разных use-cases

JUNOS EVO сценарии использования

Junos EVO: Единая операционная система – множество сценариев



JUNOS EVO 2018/19 Планы

в Q4-2018 – 2019 Junos Evo в основном сфокусируется на Cloud Data Center market.
QFX5200 – первая платформа с данной операционной системой (Ноябрь 2018)

	3Q'18	4Q'18-1Q'19	1Q'19	2Q'19	2H'19
Платформа	QFX5200 (TH)	+ QFX10003 (Jun)	+ QFX5220 (TH3)	+ QFX5120 (TD3)	+ Платформы на BT chipset
Сценарии использования	L3 Spine, L2/L3 TOR	+ LSR, IP Core, Peering, DCI	+ Private Cloud, Cable Aggregation	+ FSI, Telco Cloud	
SW Функционал	DC-OS capabilities, BGP/ISIS/OSPF, LAG, ZTP (LLDP, DHCP options)	MPLS (RSVP, LDP, SR), IP Multicast, EPE, BGP-LS, PCEP, Sflow, 128-ECMP, OCST, JET	L2 (Ucast + Mcast), L3VPN, EVPN (VXLAN + MPLS), Resilient Hashing, Flowspec, uRPF, CEM, OVSDB, jFlow, DDoS, MACsec, ISSU	MC-LAG, QinQ, Timing, SCU/DCU	Feature parity with prev releases on BT

JUNOS EVO Изменения в название имиджа

`{package-prefix}[-${media}]-${platform}-${arch}-${abi}-${release}[-${edition}].[tgz|img|iso]`

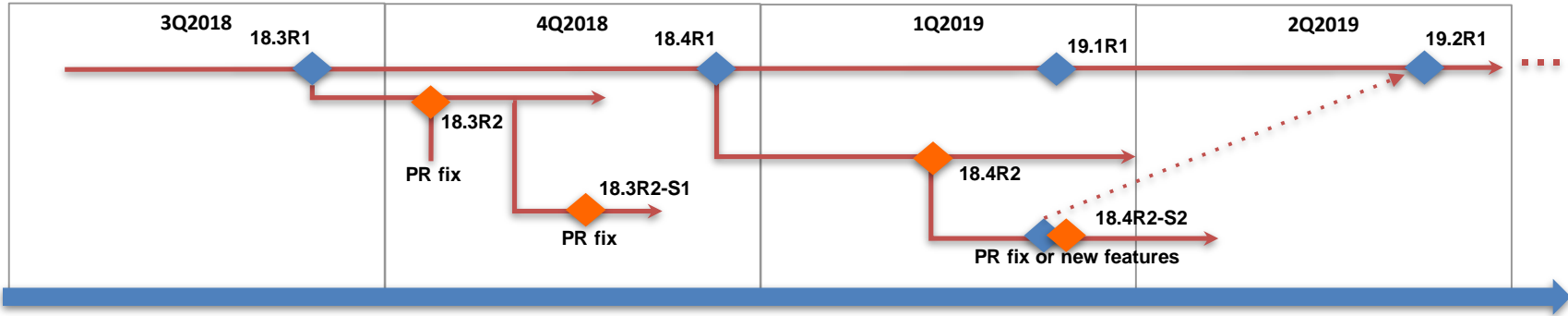
Occam on 5K

- `{install-media-host}-{usb}-{qfx-5e}-{x86}-{64}-{18.2R1.9}.img`
- `{jinstall-host}-{qfx-5e}-{x86}-{64}-{18.2R1.9}-{secure-signed}.tgz`
- `{install-media-host}-{usb}-{qfx-5e-flex}-{x86}-{64}-{18.2R1.9}-{secure}.img`

Junos Evo on 5K

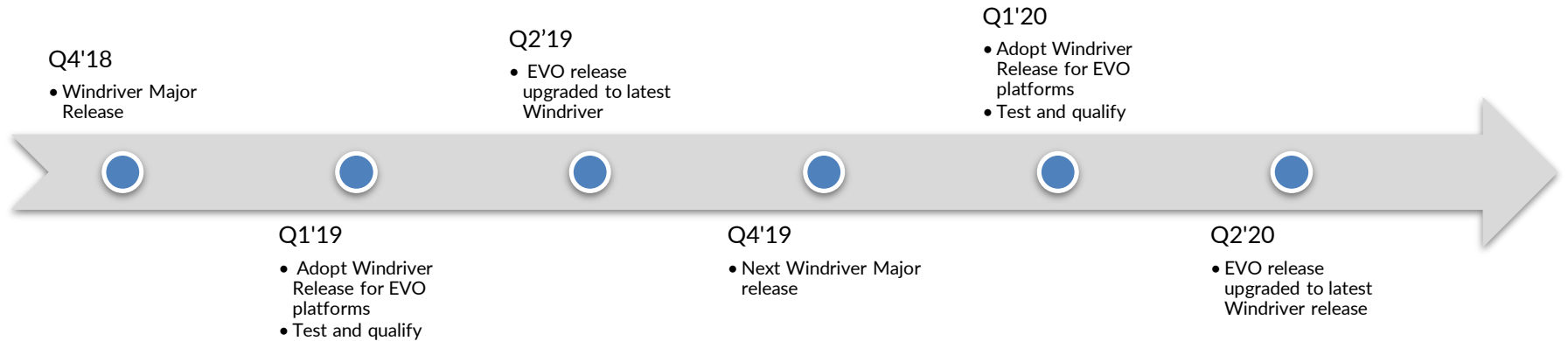
- `{junos-install-media}-{usb}-{qfx-5e}-{x86}-{linux-64}-{18.2R1.9}.iso`
- `{junos-install}-{qfx-5e}-{x86}-{linux-64}-{18.2R1.9}.iso`
- `{junos-install}-{qfx-5e}-{x86}-{linux-64}-{18.2R1.9-ip}.iso`
- `{junos-install}-{qfx-5e}-{x86}-{linux-64}-{18.2R1.9-nvo}.iso`

JUNOS EVO Планы выпуска релизов



- JUNOS EVO даты выпуска выровняют с mainline releases(-ами) и нумерацией
- Релизы R1/MR полновесные релизы
- SRs – инкрементальные релизы. Могут быть с новым функционалом или с Bug fix
 - SRs – так же поставляются как полновесные релизы, но изменена логика upgrade(a) для SR
 - Install tool вычисляются и производят апгрейд только дельта компонент
 - Включена опция «Dry run» для анализа изменений в имидже (выведет лист компонентов для апгрейда)
 - В Post-FRS будет включен инструмент offline для анализа последствий обновления (Post FRS)
 - Включена опция «-force full upgrade»

JUNOS EVO – Обновление host операционной системы Windriver Linux



EVO АРХИТЕКТУРА И ТЕРМИНОЛОГИЯ

JUNOS EVO – Какие цели продукта ? (1)

Host OS - Linux

- Обширная база разработчиков
- Использование open source tools где возможно
- Использование стандартных средств Linux tools

Убрать kernel состояния (states)

- Больше нет kernel состояний и propagation
- Минимизировать изменения в kernel
- Поддержка распределенных узлов и масштабирование системы

Модульность

- Формализованные интерфейсы между компонентами
- Изоляция между компонентами
- Изоляция на уровне сбоев
- Использование в различных use-cases
- Поддержка ISSU per app

JUNOS EVO – Какие цели продукта ? (2)

Богатая инфраструктуру под приложения (linux app)

- Упрощение разработки новых приложений
- Использования современных языков программирования

Симметричная инфраструктура

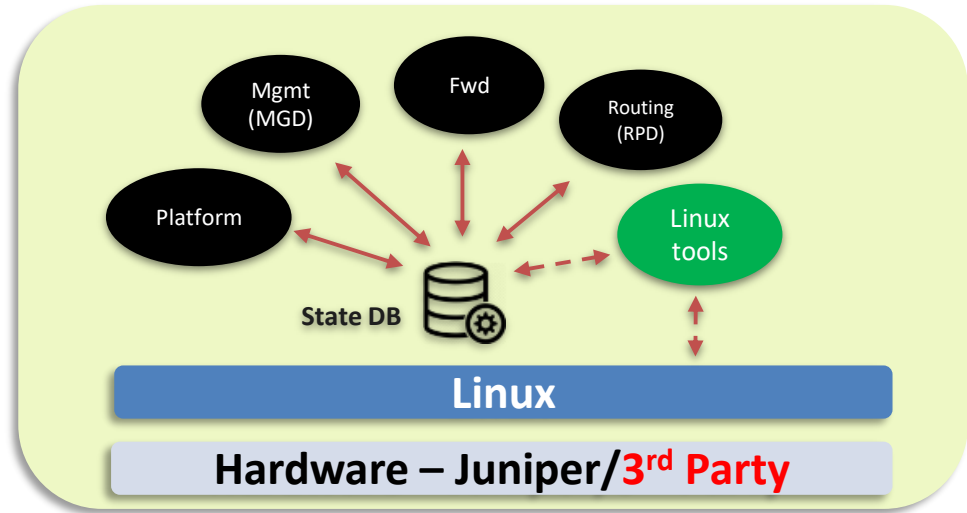
- Приложения могут быть распределенные и находится в нескольких локациях
- Автоматическая система определения распределенных приложений и узлов
- Поддержка произвольных топологий – single node, chassis, VC, VM cluster, etc.

Современная среда для разработки приложений

- Надежная система модульного тестирования
- Pre-commit / Dry Run тестирование и использование виртуального Junos EVO.

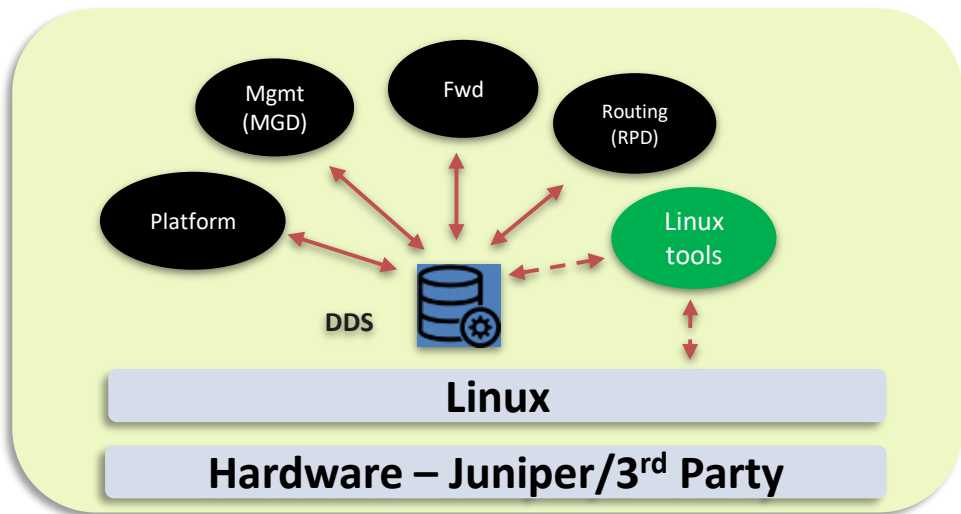
Нативная операционная система Linux OS

- Минимальные изменения в Ядро
- ПО Juniper как LKMs или приложения
- Поддержка vendor or COTS silicon
- Поддержка стандартных Linux tools



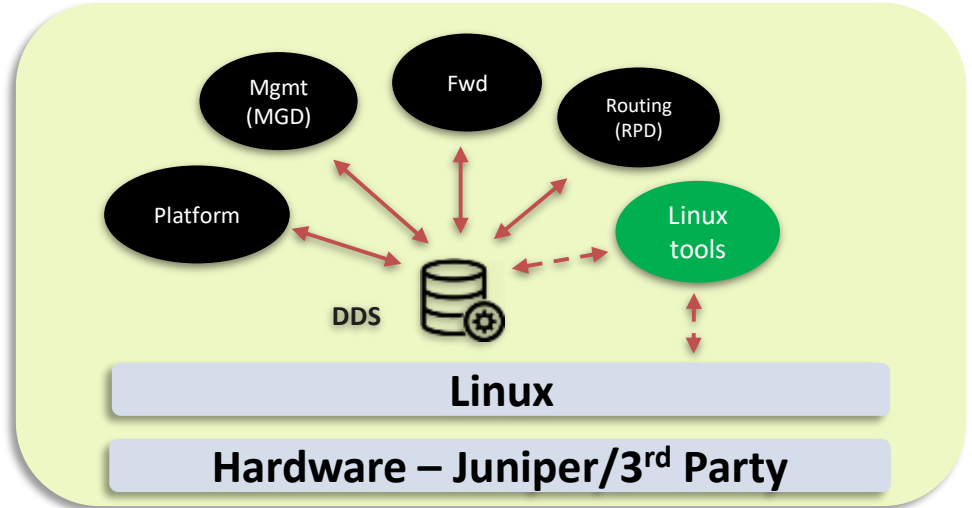
Состояния теперь хранятся в Common Object Database (DDS)

- База DDS теперь хранит как операционные так и конфигурационные состояния – committed configuration, состояния интерфейсов, маршруты, состояния аппаратных компонентов и т.д.
- Приложения – являются потребителями состояний в БД DDS
- Задача БД DDS's – хранение состояний и обмен состояниями с потребителями (app) – больше ничего
- DDS – никаких интерпретация состояний
- Отделение логики интерпретации от логики хранения

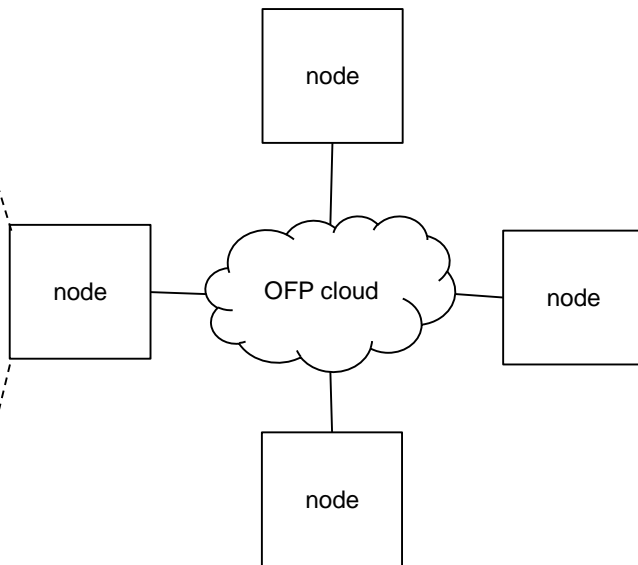
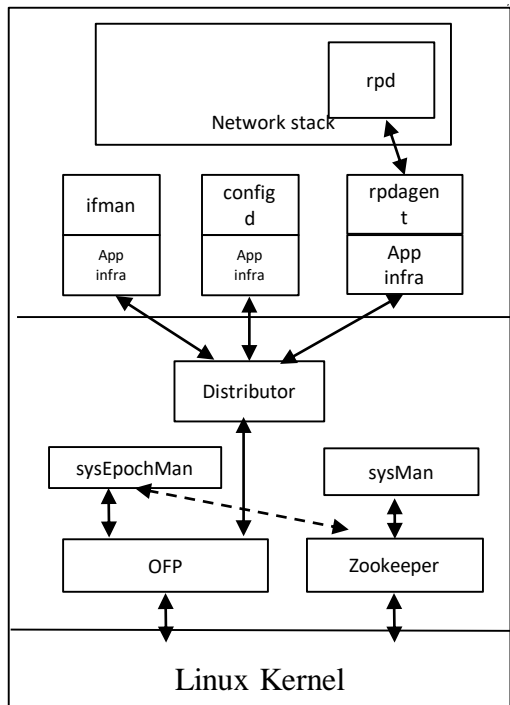


База данных хранения объектов (DDS)

- Количество транзакций не влияет на интерпретацию приложением – only the final state is important
- Производители и потребители состояний разделены во времени
- Eventually consistent – медленные потребители не влияют на скоростных потребителей
- Приложения согласовывают свои состояния из DDS при рестарте (own up old state)



Немного глубже



OFP – В распределенной системе является одним из основных элементов. Протокол коммуникаций между Node

ZooKeeper – служба, предназначенная для хранения конфигурации, а так же выполняет распределенную синхронизацию процессов.

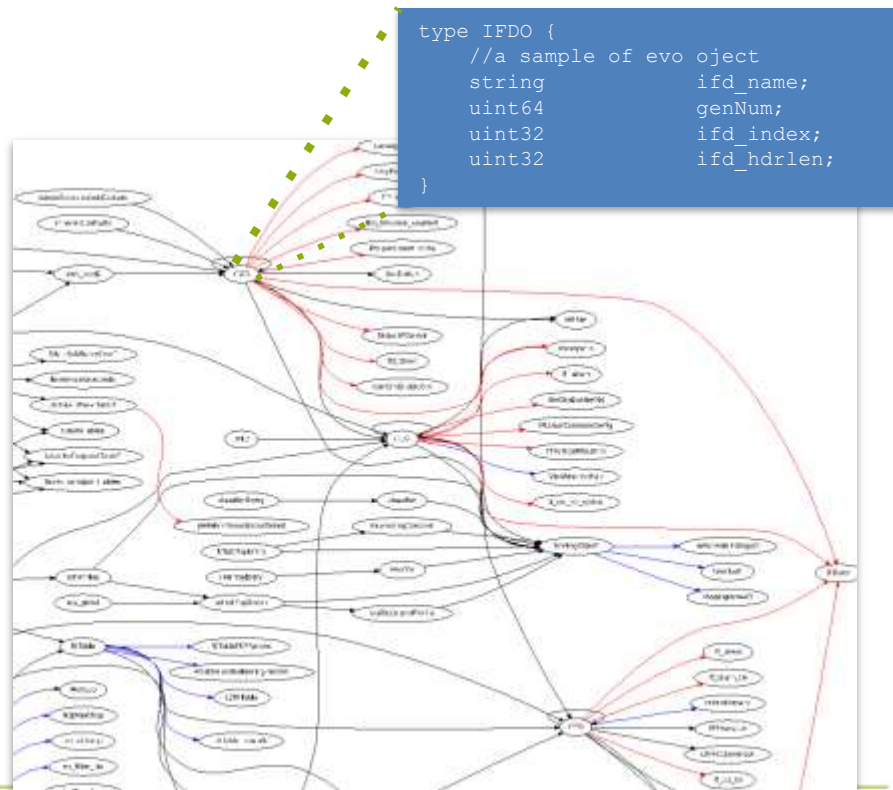
SysEpochMaster – election master, мониторинг Node в распределенной системе

Distributor – отвечает за сохранение системных состояний в базе данных DDS, координирует доступ к базе данных от приложений

SysMan – выбирается с помощью Zookeeper. На каждой ноде присутствует Active/Standby. Отвечает за регистрацию приложений в качестве Active приложения или Standby приложения

Все состояния в базе данных смоделированы как объекты

- Все состояния в DDS смоделированы как объекты
- Все объекты описаны в едином формате data modelling language (EVL)
- Объекты организованы в графы
- При необходимости автоматическое преобразование иерархической схеме в другие языки моделирования – такие как YANG
- Python/C++ APIs для работы с DDS состояниями





Currently available sessions

Name	Type	Source	Created
vBracka	Live Roamc	10.102.182.84	Mon Oct 08 2018 18:58:10 GMT-0700 (PDT)
teal	Live Roamc	10.102.226.39	Tue Oct 09 2018 06:41:16 GMT-0700 (PDT)
rkvelaga	Live Roamc	10.216.129.212	Tue Oct 09 2018 22:35:34 GMT-0700 (PDT)
rank	Live Roamc	10.82.252.190	Thu Oct 11 2018 15:17:24 GMT-0700 (PDT)
WebConfig-Te sz2	Live Roamc	10.102.227.118	Thu Oct 11 2018 19:50:40 GMT-0700 (PDT)
dhw-bron	Live Roamc	brj-ae-duck-p1-1	Thu Oct 11 2018 22:33:44 GMT-0700 (PDT)
10.102.229.10 3	Live Roamc	10.102.229.103	Fri Oct 12 2018 03:15:23 GMT-0700 (PDT)
vyastrebov	Live Roamc	10.49.172.122	Fri Oct 12 2018 03:58:00 GMT-0700 (PDT)

Create new session

Router File

Address

10.49.172.122

Username

root

Password

Embed/Implic

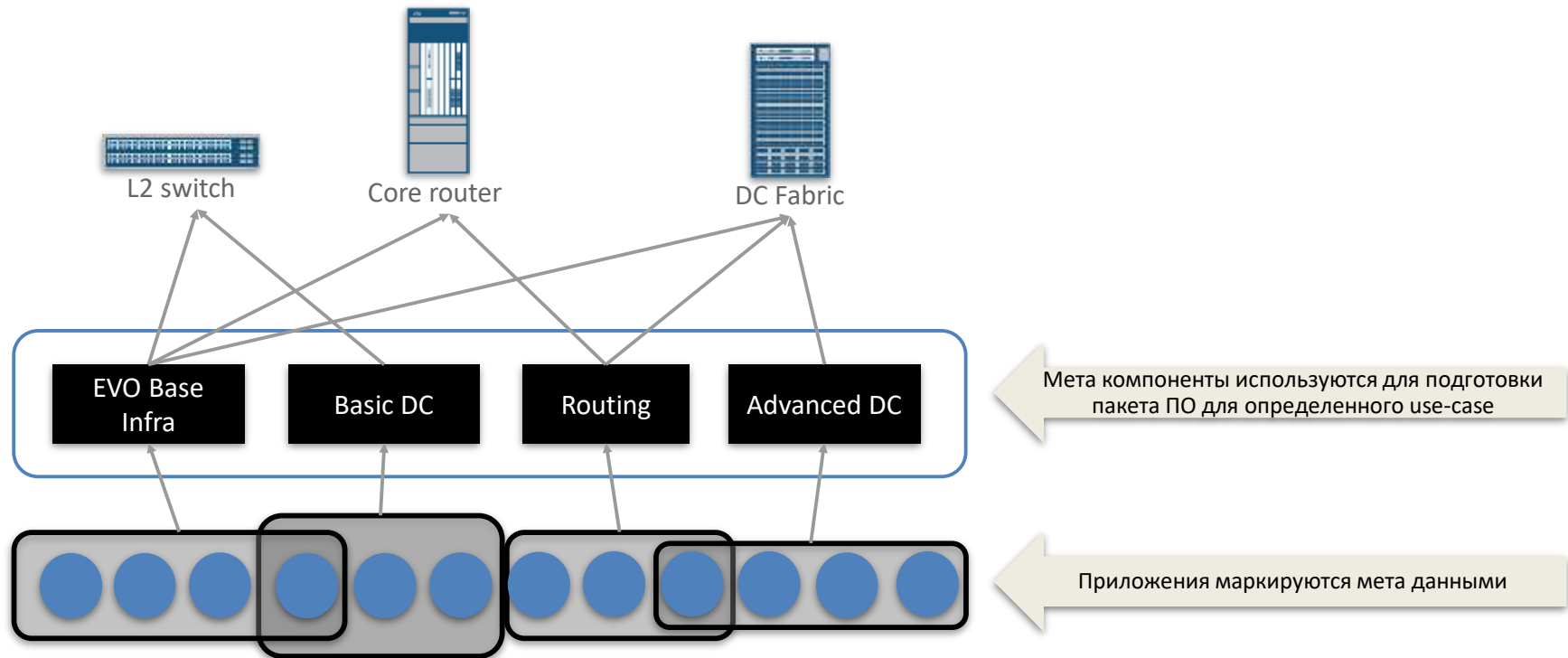
vyastrebov

Login

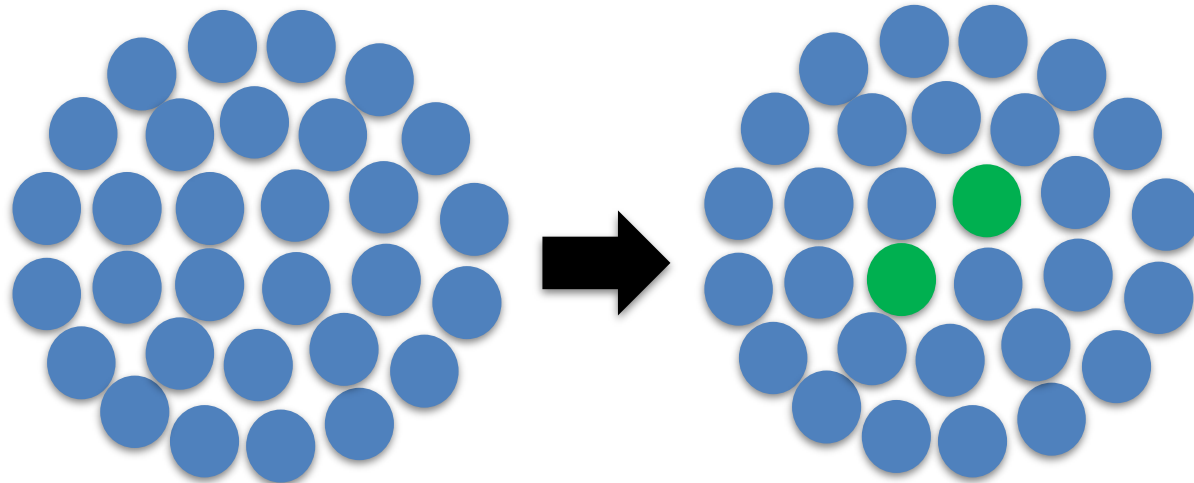
Previously used

10.49.172.122

JUNOS EVO гибкий подход к формированию install-package



JUNOS EVO процедура апгрейда



JUNOS EVO приложения с SW
version 1

Junos EVO приложения с SW
version 2

Процедура ISSU (In-service software upgrade):

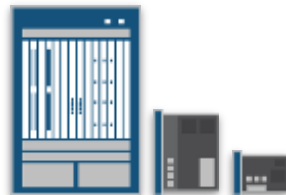
1. Используется процедура diff для оценки изменений существующего ПО и нового
2. Сохраняются состояния устройства
3. Обновляются и рестартуют только модули требующие изменения

EVO инкрементальные upgrade(ы)



hotfix

Осуществляется обновление только отдельных приложений (app)



JAM

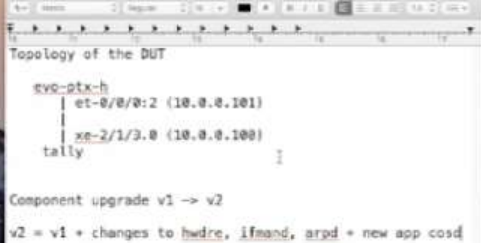
Обновление аппаратных драйверов. Версия ПО остается той же



s/w JAM

Добавление программного функционала

ДЕМОНСТРАЦИЯ



```

64 bytes from 10.0.0.100: icmp_seq=469 ttl=64 time=1.148 ms
64 bytes from 10.0.0.100: icmp_seq=470 ttl=64 time=1.152 ms
64 bytes from 10.0.0.100: icmp_seq=471 ttl=64 time=55.387 ms
64 bytes from 10.0.0.100: icmp_seq=472 ttl=64 time=2.477 ms
64 bytes from 10.0.0.100: icmp_seq=473 ttl=64 time=1.329 ms
64 bytes from 10.0.0.100: icmp_seq=474 ttl=64 time=1.298 ms
64 bytes from 10.0.0.100: icmp_seq=475 ttl=64 time=3.951 ms
64 bytes from 10.0.0.100: icmp_seq=476 ttl=64 time=3.070 ms
64 bytes from 10.0.0.100: icmp_seq=477 ttl=64 time=1.183 ms
64 bytes from 10.0.0.100: icmp_seq=478 ttl=64 time=68.048 ms
64 bytes from 10.0.0.100: icmp_seq=479 ttl=64 time=1.180 ms
64 bytes from 10.0.0.100: icmp_seq=480 ttl=64 time=1.065 ms
64 bytes from 10.0.0.100: icmp_seq=481 ttl=64 time=2.019 ms
64 bytes from 10.0.0.100: icmp_seq=482 ttl=64 time=3.955 ms
64 bytes from 10.0.0.100: icmp_seq=483 ttl=64 time=1.441 ms

```

```

root@evo-ptx-h-re0:~# cli
root@evo-ptx-h-re0:~# exit

root@evo-ptx-h-re0:~# cli
root@evo-ptx-h-re0:~# request system software add /var/tmp/ptx828-2.iso restart | no-...

```

```

64 bytes from 10.0.0.101: icmp_seq=468 ttl=64 time=1.434 ms
64 bytes from 10.0.0.101: icmp_seq=469 ttl=64 time=1.499 ms
64 bytes from 10.0.0.101: icmp_seq=470 ttl=64 time=71.982 ms
64 bytes from 10.0.0.101: icmp_seq=471 ttl=64 time=1.472 ms
64 bytes from 10.0.0.101: icmp_seq=472 ttl=64 time=1.389 ms
64 bytes from 10.0.0.101: icmp_seq=473 ttl=64 time=1.377 ms
64 bytes from 10.0.0.101: icmp_seq=474 ttl=64 time=1.441 ms
64 bytes from 10.0.0.101: icmp_seq=475 ttl=64 time=1.845 ms
64 bytes from 10.0.0.101: icmp_seq=476 ttl=64 time=1.543 ms
64 bytes from 10.0.0.101: icmp_seq=477 ttl=64 time=34.648 ms
64 bytes from 10.0.0.101: icmp_seq=478 ttl=64 time=1.489 ms
64 bytes from 10.0.0.101: icmp_seq=479 ttl=64 time=1.381 ms
64 bytes from 10.0.0.101: icmp_seq=480 ttl=64 time=1.467 ms
64 bytes from 10.0.0.101: icmp_seq=481 ttl=64 time=1.461 ms
64 bytes from 10.0.0.101: icmp_seq=482 ttl=64 time=1.419 ms

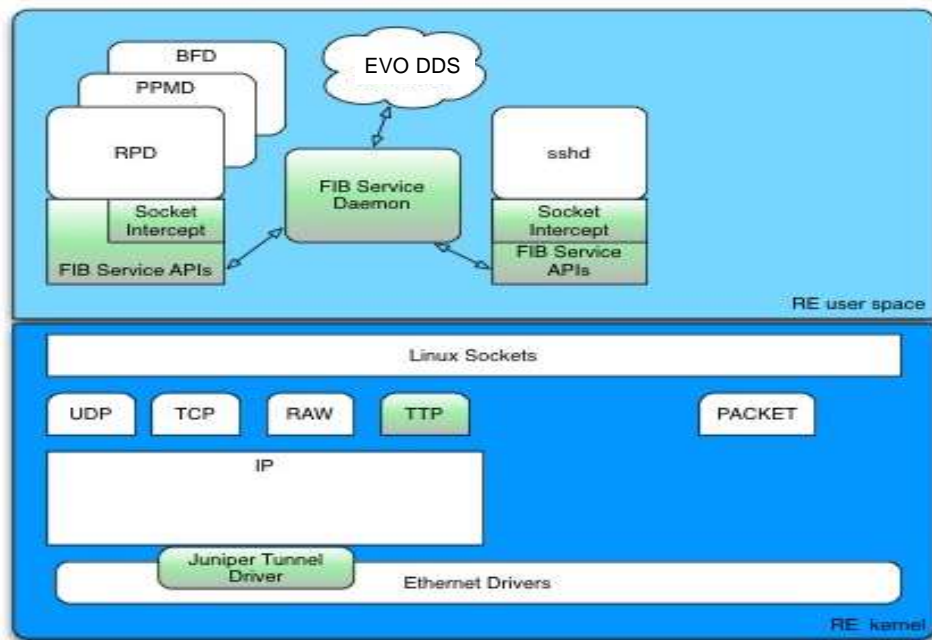
```

```

root@evo-ptx-h-re0:~# cli
root@evo-ptx-h-re0:~#

```

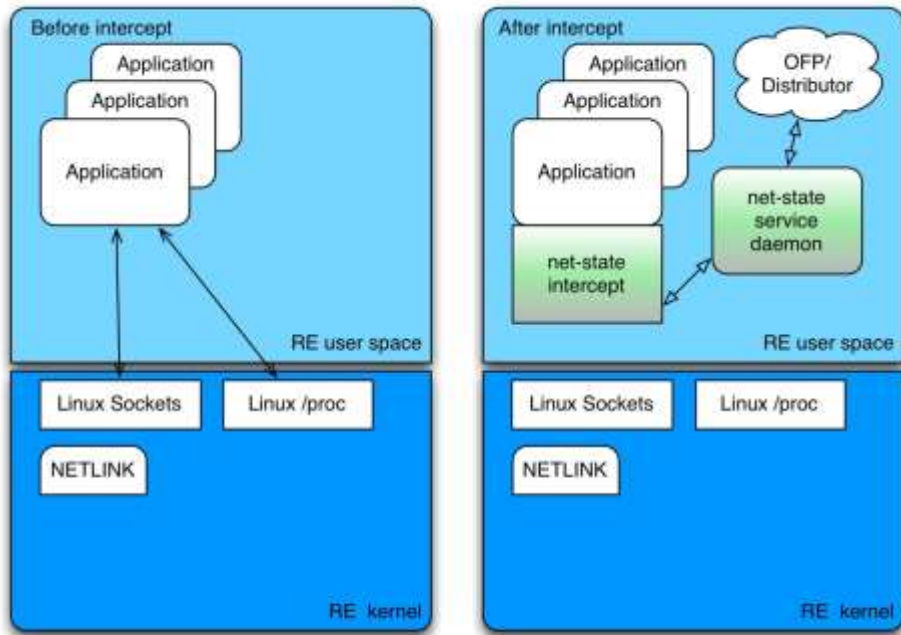
JUNOS EVO: сетевой стек



- Без VM, EVO Linux – единая операционная система и сетевой стек.
- RPD/PPMD нативно работают поверх LINUX как EVO приложения
- Сетевой стек Junos Evo прозрачен для предкомпилированных LINUX приложений, socket interceptor осуществляют перехват вызовов в сторону LINUX socket API (Socket Intercept “hijacked” socket APIs)
- Контрольные пакеты перехватываются Juniper Tunnel driver (pseudo tunnel device)
- Новый FIB Service Daemon и service APIs обеспечивают routing в Linux user space оперируя напрямую с DDS

 New Modules

JUNOS EVO: сетевой стек -2

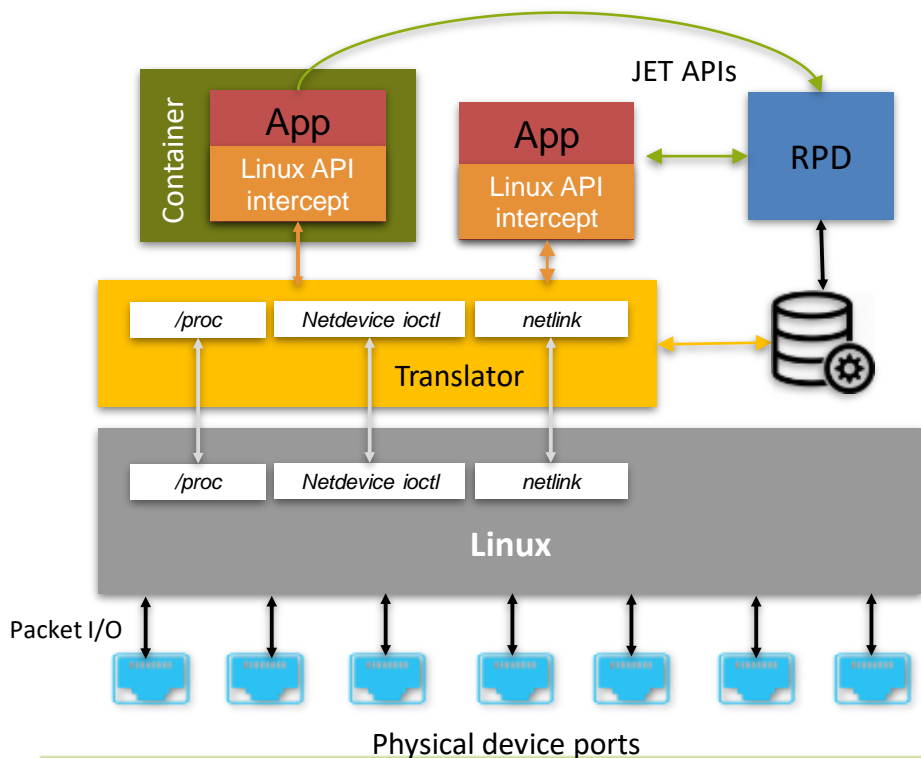


- С использованием интерсептера LD_PRELOAD транспарентная поддержка Linux APIs
- Packet IO, netlink, /proc/net, netdevice

```
root@EVOvPTX1_RE0-re0:~# uname -a
Linux EVOvPTX1_RE0-re0 4.4.66-juniper-02729-g1d01e24 #1 SMP PREEMPT Wed Nov 15 10:56:58 PST 2017 x86_64 GNU/Linux
root@EVOvPTX1_RE0-re0:~# cli show route
```

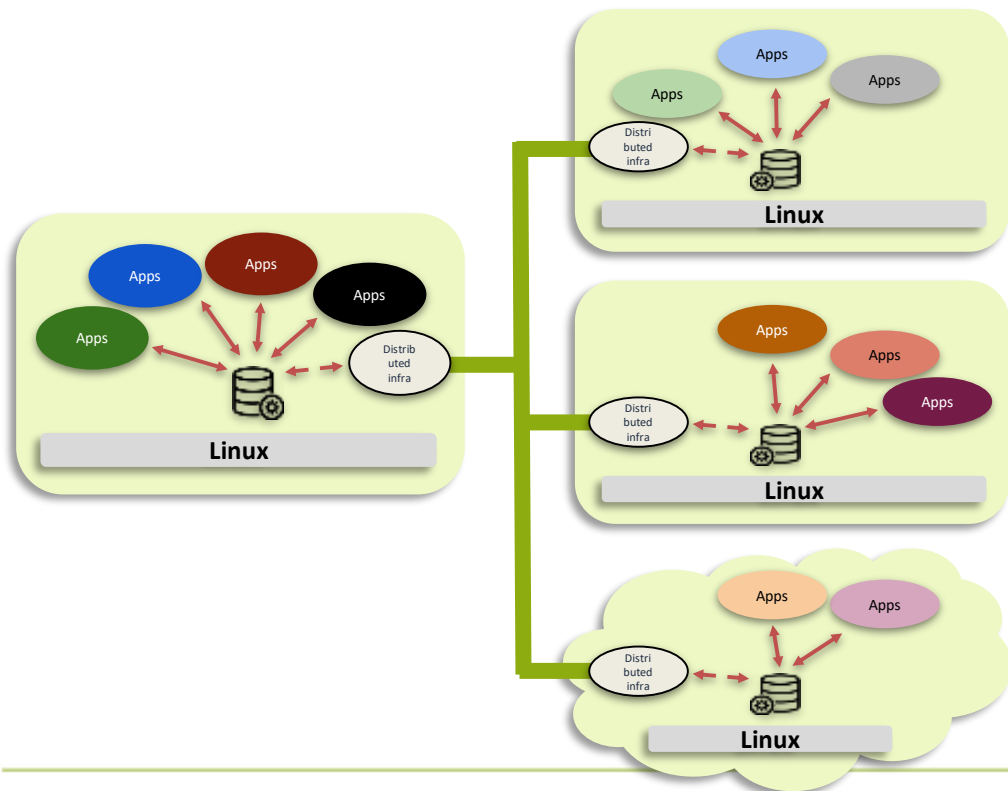
```
iper-02729-g1d01e24 #1 SMP PREEMPT Wed Nov 15 10:56:58 PST 2
```

3rd Party Support On EVO



- 3rd party приложения могут работать нативно или в контейнере
- Контейнер в свою очередь может быть самостоятельным Linux дистрибутивом со своим пакетным менеджером
- Junos toolchain для разработки собственных приложений
- Транспарентная поддержка Linux APIs с использованием LD_PRELOAD (интерцептер)
- Может быть использован JET APIs для доступа к Junos EVO/DDS
- Оболочки Python/C++ используются для чтения из базы данных DDS
- DevOps инструменты для управления устройством

JUNOS Evo распределенная операционная система по задумке

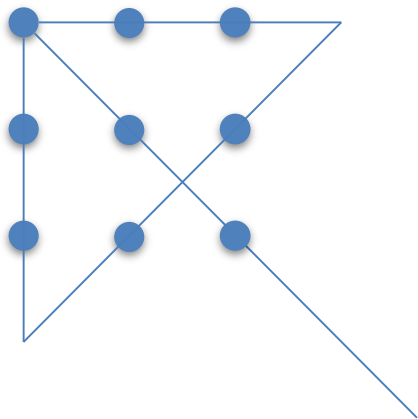


- Приложения являются location agnostic (протокол OFP)
- Ноды (physical or virtual) используют вшитый механизм discover для поиска друг друга и используют минимальные настройки
- Поддержка произвольных топологий
- Режимы резервирования приложений и Node
- Приложения: active/standby
- Масштабирования при добавлении ресурсов

Заключение

- Junos Evo – это новый шаг в эволюции Junos
- Экосистема новой архитектуры позволяет добиться лучшего качества, более быстрого развертывания, инновационных решений пользователей
- Конфигурация и все данные хранятся в БД DDS. Не требуется поить отдельно приложения для получения состояний, все запросы делаются в сторону БД
- Единая архитектура приложений и драйверов (например: RPD, MGD и т.д.)
- Первый релиз ПО для QFX5200 (Q3`18) и далее QFX10003 (Q4`18)
- Расширение use-case данной архитектуры на другие домены, в том числе магистраль
- Junos Evo будет выпускаться под единым знакомым брендом Junos

Решение:



Для решения задачи нужно выйти за границы!

Спасибо!

