

ONL/ONIE ON QorlQ AND QorlQ LS SERIES

FTF-DES-N1842

DR. XIN-XIN YANG SOFTWARE ENGINEERING MANAGER FTF-DES-N1842 MAY 17, 2016





### **AGENDA**

- What is ONL/ONIE
- Why ONL/ONIE in NXP
- Backport Kernel Rebase
- ONL/ONIE Development and Upstreaming
- Plan & Roadmap
- Summary



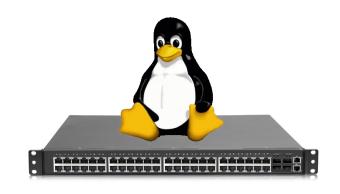
### WHAT IS ONL/ONIE



### ONL is a Linux Distribution for Bare Metal Switches

- A collection of software packages, utilities, drivers, and abstractions to run on OCP, bare metal, "brite box" hardware
  - i.e., a "NOS" that ONIE would install
  - Network forwarding devices built from commodity components
- ONL uses ONIE to install onto on-board flash memory
- Open Network Linux is a part of the Open Compute Project and is a component in a growing collection of open source and commercial projects

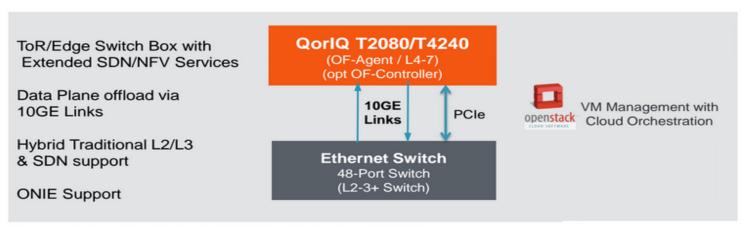
http://opennetlinux.org/





### Why Not Use an Existing Linux Distribution

- Does build on existing distribution Debian Wheezy
- Need to create ONIE installers for many platforms
- Need to manage switch-specific hardware (e.g.,SFPs)
- Switches are very similar to servers, but not quite



Network Service Switch (NSS) with L4/L7 SDN Support

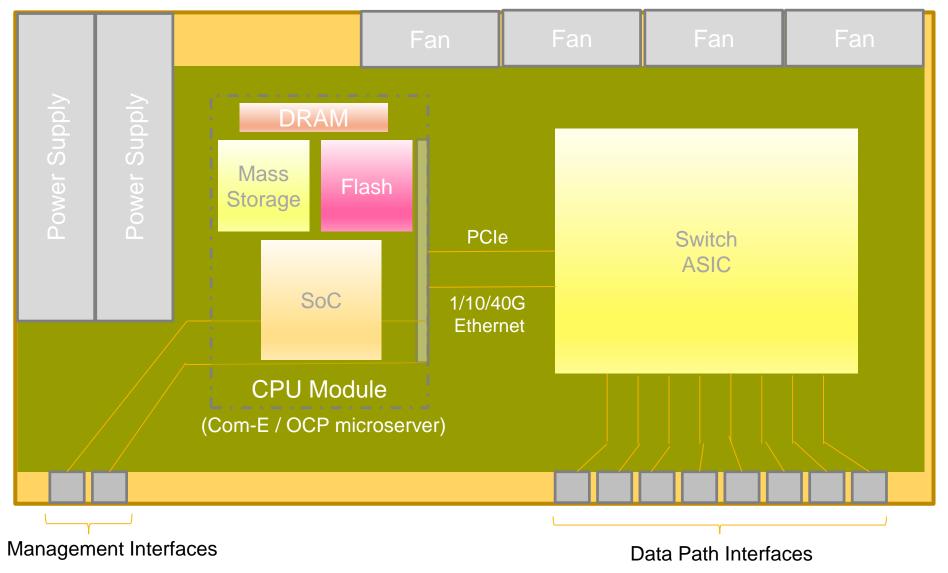


### Why Use ONL?

- Help ecosystem focus on innovation
  - Many annoying software details to run an OCP switch
  - Building platform drivers not high value asset; should be common
- Enables a reference NOS implementation
  - Hardware without software is useless
  - Package up details and best practices into one place
- Help bootstrap the Open ecosystem and OCP adoption
  - Allows commercial companies and DYI-folks to build OCP-based products faster



### **Bare Metal Switch**



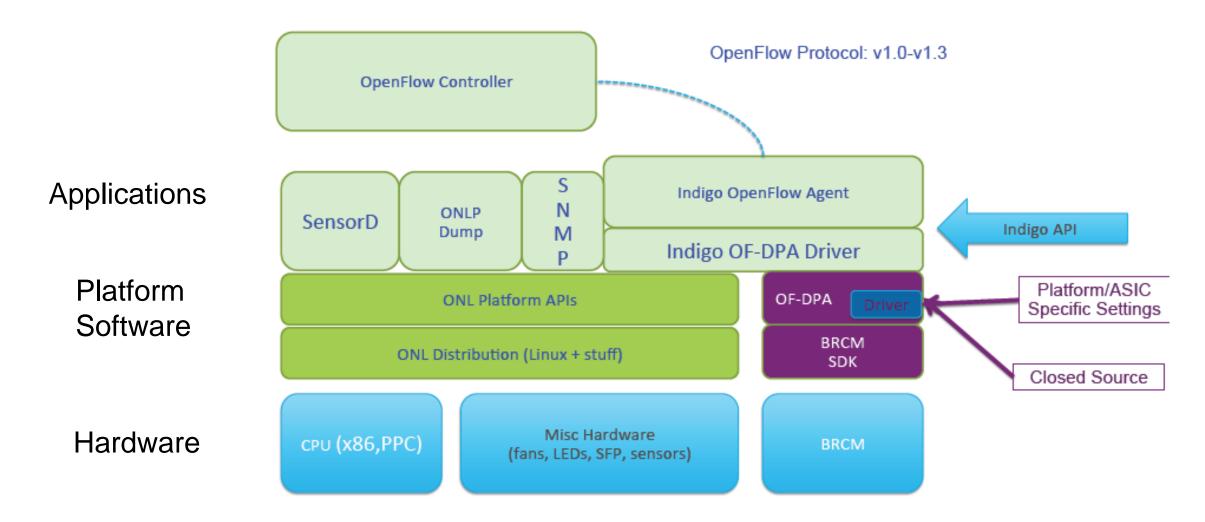


### **ONL** Architecture

**OpenNSL** OpenFlow Open Route **Applications** . . . Agent Cache Apps **Open Network Linux Platform** Platform Specific **Platform Abstraction Layer** Drivers Software **ONL** Kernel BRM or ASIC SDK BRM Switch Chips / Misc Hardware (Fans, DDR, USB, **CPU** SATA, PCIe, sensors, power supplies) ASIC Hardware **OCP Switch Hardware** 



### **ONL With OpenFlow Agent**





### **Supported Hardware - I**

### Quanta

Device	Ports	СРИ	Forwarding	ONL Certified	In Lab	ORC	OF-DPA	OpenNSL	SAI
QuantaMesh T1048-LB9	48x1G + 4x10G	FreeScale P2020	Broadcom BCM56534 (Firebolt3)	Yes	Yes	Yes	No	No	No
QuantaMesh T3048-LY2	48x10G + 4x40G	FreeScale P2020	Broadcom BCM56846 (Trident+)	Yes	Yes	Yes	Yes	No	No
QuantaMesh T3048-LY8	48x10G + 6x40G	Intel Rangely C2758 x86	Broadcom BCM56854 (Trident2)	Yes*	No	No	No	No	No
QuantaMesh T5032-LY6	32x40G	Intel Rangely C2758 x86	Broadcom BCM56850 (Trident2)	Yes*	No	No	No	No	No

### Accton/Edge-Core

Device	Ports	CPU	Forwarding	ONL Certified	In Lab	ORC	OF-DPA	OpenNSL	SAI
Accton AS4600-54T	48x1G + 4x10G	FreeScale P2020	Broadcom BCM56540 (Apollo2)	Yes	Yes	Yes	Yes***	Yes***	No
Accton AS5600-52X	48x10G + 4x40G	FreeScale P2020	Broadcom BCM56846 (Trident+)	Yes	Yes	Yes	No	No	No
Accton AS5610-52X	48x10G + 4x40G	FreeScale P2020	Broadcom BCM56846 (Trident+)	Yes	Yes	Yes	No	No	No
Accton AS5710-54X	48x10G + 6x40G	FreeScale P2041	Broadcom BCM56854 (Trident2)	Yes	Yes	Yes	Yes***	Yes***	No
Accton AS6700-32X	32x40G	FreeScale P2041	Broadcom BCM56850 (Trident2)	Yes	Yes	Yes	No	No	No
Accton AS5712-54X	48x10G + 6x40G	Intel Rangely C2538 x86	Broadcom BCM56854 (Trident2)	Yes	Yes	Yes	Yes***	Yes***	No
Accton AS6712-32X	32x40G	Intel Rangely C2538 x86	Broadcom BCM56850 (Trident2)	Yes	Yes	Yes	Yes***	Yes***	No
Accton AS5812-54T	48x10G + 6x40G	Intel Rangely C2538 x86	Broadcom BCM56864 (Trident2+)	Yes	Yes	No	No	No	No
Accton AS5812-54X	48x10G + 6x40G	Intel Rangely C2538 x86	Broadcom BCM56864 (Trident2+)	Yes	Yes	No	No	No	No
Accton AS6812-32X	32x40G	Intel Rangely C2538 x86	Broadcom BCM56864 (Trident2+)	Yes	Yes	No	No	No	No
Accton AS7712-32X	32x100G	Intel Rangely C2538 x86	Broadcom BCM56960 (Tomahawk)	Yes	Yes	Yes	Yes***	Yes***	No
Accton Wedge-16X	16x40G	Intel Rangely C2550 x86	Broadcom BCM56864 (Trident2+)	Work In Progress**	Yes	No	No	Yes	No



### **Supported Hardware - II**

### DNI/Agema

Device	Ports	CPU	Forwarding	ONL Certified	In Lab	ORC	OF-DPA	OpenNSL	SAI
AG-7448CU	48x10G + 4x40G	FreeScale P2020	Broadcom BCM56845 (Trident)	Yes	Yes	Yes	No	No	No

### Dell

Device	Ports	СРИ	Forwarding	ONL Certified	In Lab	ORC	OF-DPA	OpenNSL	SAI
S4810-ON	48x10G + 4x40G	FreeScale P2020	Broadcom BCM56845 (Trident)	Yes	Yes	Yes	No	No	No
S4048-ON	48x10G + 6x40G	Intel Atom C2338	Broadcom BCM56854 (Trident2)	Yes	Yes	Yes	No	No	No
S6000-ON	32x40G	Intel Atom S1220	Broadcom BCM56850 (Trident2)	Yes	Yes	Yes	No	No	No
Z9100-ON	32x100G	Intel Atom C2538	Broadcom BCM56960 (Tomahawk)	Yes	Yes	No	No	No	No

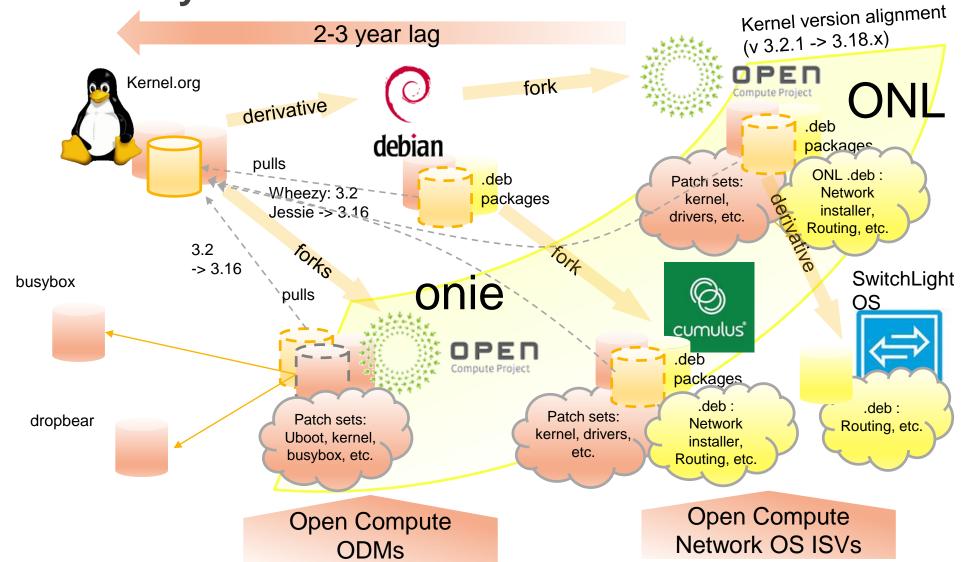
### Interface Masters Technologies, Inc.

Device	Ports	CPU	Forwarding	ONL Certified	In Lab	ORC	OF-DPA	OpenNSL	SAI
Niagara 2948X12XLm	48x10G + 12x40G	Intel/AMD x86	Broadcom BCM56850 (Trident2)	Work In Progress**	No	No	Yes***	Yes***	No
Niagara 2960X6XLm	60x10G + 6x40G	Intel/AMD x86	Broadcom BCM56850 (Trident2)	Work In Progress**	No	No	Yes***	Yes***	No
Niagara 2972Xm	72x10G	Intel/AMD x86	Broadcom BCM56850 (Trident2)	Work In Progress**	Yes	No	Yes***	Yes***	No
Niagara 2932XL	32x40G	Intel/AMD x86	Broadcom BCM56850 (Trident2)	Work In Progress**	No	No	Yes***	Yes***	No
Niagara 2948X6XL	48x10G + 6x40G	Intel/AMD x86	Broadcom BCM56850 (Trident2)	Work In Progress**	No	No	Yes***	Yes	No

http://opennetlinux.org/hcl



### **ONL/ONIE Ecosystem**





# WHY ONL/ONIE IN NXP



### **Changing Requirements for Linux Distributions**

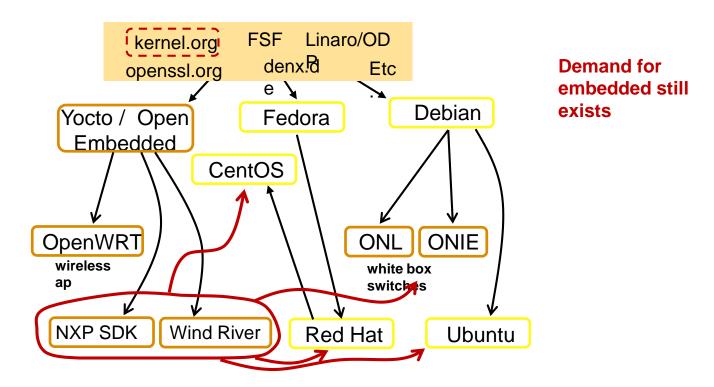
- Requirements are changing due to technology shifts and convergences
  - -NFV
  - Appliance / Server convergence
- Suggested NXP responses
  - Rebalance team goals more to enabling ecosystem distributions via basic enablement "upstreaming"
  - Resource allocation for special-purpose distributions.
  - Eventually, budget for commercial Enterprise distributions.
- Note: Discussion is what Linux distribution customers demand, not valueadd software on top of Linux



### **Linux Distribution Requirements Shift**

- Demand <u>broadens</u> away from embedded (and NXP SDK) towards enterprise and some enterprise-derived special purpose distributions.
- Biggest reason is convergence of network appliances & servers but also
  - Server ecosystem dominance in ARMv8
  - More powerful SoCs
  - Intel encroachment (WB switches)

- · Ease of use in some cases
- Standardization





### **Details of the Changes**

- Server ODMs, OEMs and operators require a single stable image for consolidating all of their server equipment
  - One unified asset to deploy to all of their equipment (of the same type)
  - Control OPEX related to validation of platforms and management of equipment
- Embedded Solutions (Yocto, Linaro, Enea) insufficient (for this usage)
  - Server users (e.g. carrier operators) use automated provisioning to perform a one-time install of a certified Enterprise image to white box & NFV servers
  - Requires inclusion of QorlQ platform support and drivers (esp. net driver)
- Rely on commercial distributions
  - Linux kernel and suite of server applications certified to work against it
  - Kitchen sink approach: distribution contains all the platform software the operator may conceivably need, pre-built (i.e. in binary form) and pre-tested
  - Long-term support provided against a stable (i.e. well-tested in field) configuration
- Limited set of vendors:
  - Red Hat (primary)
  - Canonical (new entrant largest platform vendor for OpenStack cloud): Ubuntu
  - SUSE (predominantly Europe)

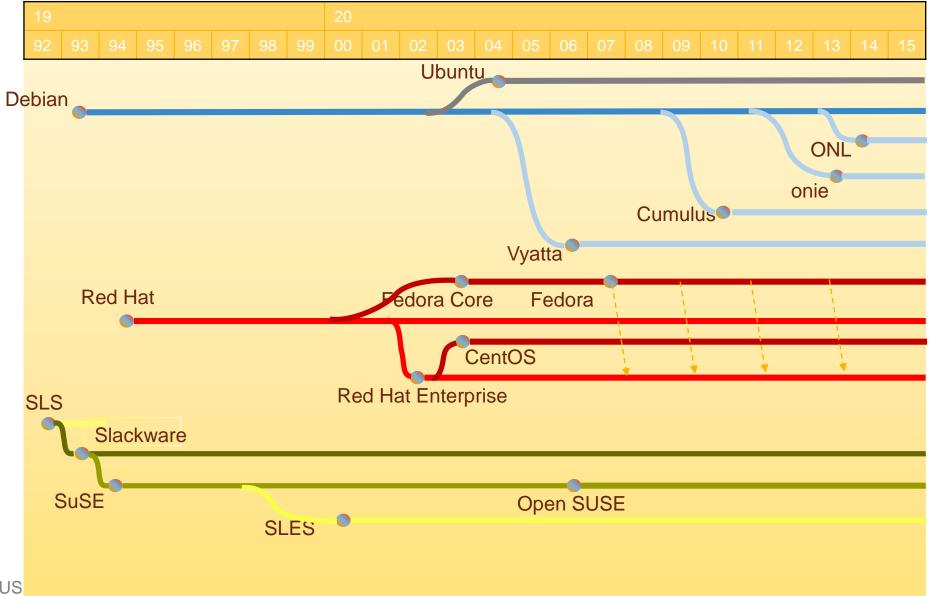


### **Distributions**

- Two key components
  - Linux kernel standardized to a stable configuration / revision level
  - Root file system containing user space applications and dynamically loadable kernel modules for standard drivers
- Commercial distributions usually rely on kernels and user space packages derived from an upstream community-driven distribution (feeder)
  - E.g. Debian, Slackware, Gentoo
  - Generally share build tools, package management system, etc. with progenitor
  - Frequent cross-pollination between feeder and derivative (i.e. not strictly a \*fork\*); e.g. derivative rebases off new feeder releases, bug fixes, enhancements submitted upstream
  - Some commercial distros sponsor community distros; e.g. Red Hat → Fedora, CentOS
- All derive from a release branch of the mainline kernel.org linux kernel development tree
  - Often distinguished by how closely they track to kernel.org releases
  - Community distributions typically released more frequently and closer to kernel.org releases
  - Enterprise distros focus on stability with less frequent releases based on long-term support "branches"

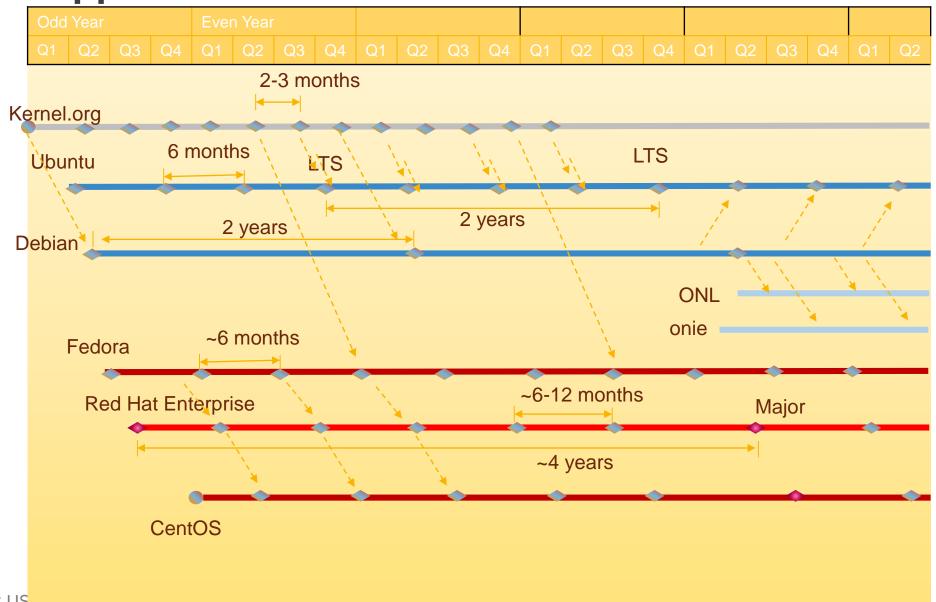


### Family Tree of Key Enterprise / Network Distributions





### **Release Approaches**





### **ARMv8 Support Status**

Distribution	ARMv8 Status	Release	Date
Fedora	Supported† (secondary architecture)	21	December 2014
Debian	Supported	"Jessie"	June 2015
Ubuntu	Supported	14.10‡ † 15.04	Oct 2014
CentOS	Alpha	7.1	June 2015
RHEL	Preview	7.1	Upcoming (2015)

<sup>†</sup> Does not support ARMv8 processors with GIC500 in legacy mode (key interoperability issue)



<sup>‡</sup> Long-term support release – next official LTS not for 2 years

### **QorlQ Power Support Status (QorlQ non-DPAA Only)**

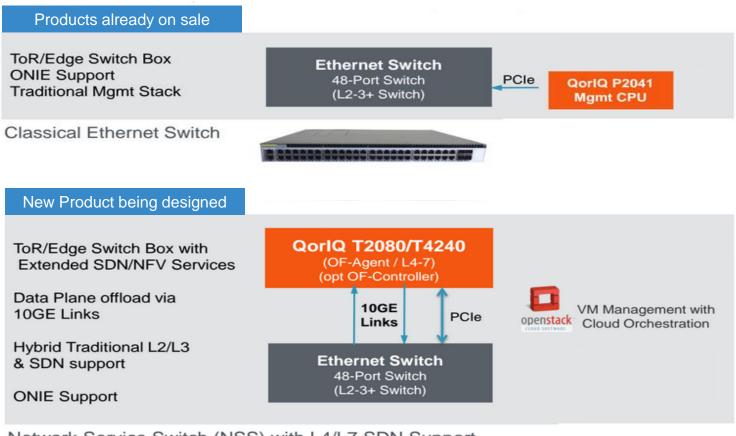
Distribution	PPC Status	Release	Date
Fedora	32-bit support EOL (F17) Power Architecture-based 64-bit	Fedora Core 4 Fedora 8?	June 2005 2007?
Debian	Supported	2.2† "Jessie"‡	August 2000 June 2015
Ubuntu	EOL (community supported)	n/a	n/a
CentOS	EOL (CentOS 5)	n/a	n/a
RHEL OpenPower (thru IBM)	EOL (RHEL 6) † Supported ‡	n/a RHEL 5.11	n/a 2014

<sup>† 32-</sup>bit Power Architecture- based



<sup>‡</sup> Power Architecture-based 64-bit platform support; with user space compatibility issues

### White Box Switch Reference Offerings with NXP Processor



Network Service Switch (NSS) with L4/L7 SDN Support

Accton will launch a 32X100G Switch based on LS2080A in the mid-year.



# BACKPORT – KERNEL REBASE



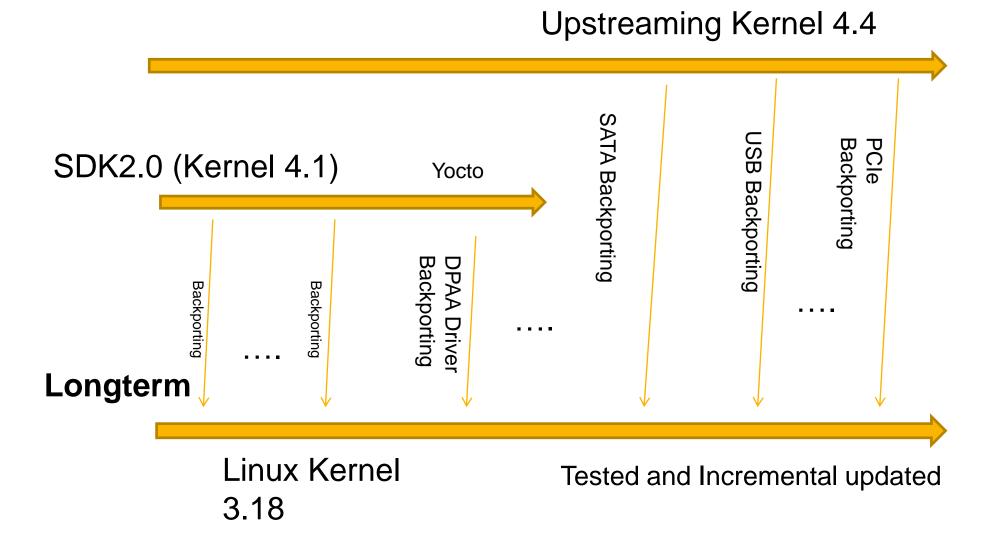
### **Kernel Version Gap**

- Kernel version in latest ONL: 3.18
- Kernel version in latest ONIE: 4.1
- Latest long term kernel version in community: 4.4.
- Latest kernel version in NXP Linux SDK
  - SDK1.9: 3.12
  - SDK2.0: 4.1

Backport is needed



### **Kernel Backport**





### **Platform and Core Support**

- Take T2080 as example, it contains 4 dual-thread e6500 Power Architecture processors organized in one cluster and it can only run 64-bit kernel.
- For this type backporting, we use both module analysis method and bisection method. Module analysis method can help us finish the early stage of backporting and bisection method can help to find out the bug as quickly as possible.
- The key patches as followed. (focused on e6500, 64-bit kernel, dual-thread and platform support)

```
From a2From 8815from segmentification states and secondary from the secondary from the secondary from the secondary support to the secondary support to assume signed-off-by: Signed-off-by: Rumar Gala <a href="mailto:square-off-by: laid support from the support from the support from the support from the support in the support from the support support from the support from the support support from the support support from the support from the support support from the support support from the support support support support from the support support support support support support from the support suppo
```



### **DPAA** Driver Backport

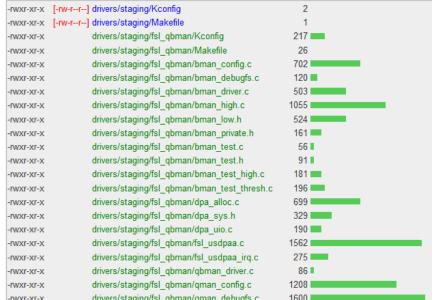
- DPAA driver is on the way of upstreaming.
- Need to backport from NXP SDK2.0.
- In order to make the DPAA related networking driver work, we need follow modules at least: QBMAN, FMAN, PAMU, PME and etc.
- Simply replacing the driver folder cannot make the driver work.
  - The real difficulty lies in connecting DPAA driver and other components of kernel, such as SKB, IOMMU and etc.

#### fsl\_qbman: Add drivers for the Freescale DPAA Q/BMan

extract from FreeScale QorIQ SDK

```
Signed-off-by: Geoff Thorpe <Geoff.Thorpe@freescale.com>
Signed-off-by: Bharat Bhushan <Bharat.Bhushan@freescale.com>
Signed-off-by: Bogdan Hamciuc <bogdan.hamciuc@freescale.com>
Signed-off-by: Hai-Ying Wang <Haiying.Wang@freescale.com>
Signed-off-by: Jeffrey Ladouceur <jeffrey.ladouceur@freescale.com>
Signed-off-by: Jia-Fei Pan <Jiafei.Pan@freescale.com>
Signed-off-by: Kumar Gala <galak@kernel.crashing.org>
Signed-off-by: Priyanka Jain <Priyanka.Jain@freescale.com>
Signed-off-by: Vakul Garg <vakul@freescale.com>
Signed-off-by: Emil Medve <Emilian.Medve@Freescale.com>
Integrated-by: Li Jingyu <b41874@freescale.com>
```

#### Diffstat





### Driver Support - PCI/PCIe, SATA, USB etc.

- Different from DPAA related driver, the patches for PCI/PCIe, SATA, USB etc have been upstreamed already.
- They should be backported from the upstreaming kernel.
- There are numerous patches for this type of driver because of the large gap between kernel versions. Simply applying all the related patches between the two versions is inefficient. Need to pick up the key patches individually and apply them accordingly.



### **Backport Procedure**

The key point in backport procedure - The specific platform and driver should be upgraded, but the kernel shouldn't! mit cc50fad2fee92520ee279e47e3e3222f6536d4d6

diff --git a/drivers/net/fman/Makefile b/drivers/net/fman/Makefile

Author: Li JingYu <b41874@freescale.com>

Tue Mar 4 13:20:15 2014 +0800

fmd: backport some interfaces to old version

- Code Architecture
- **Function Interface**

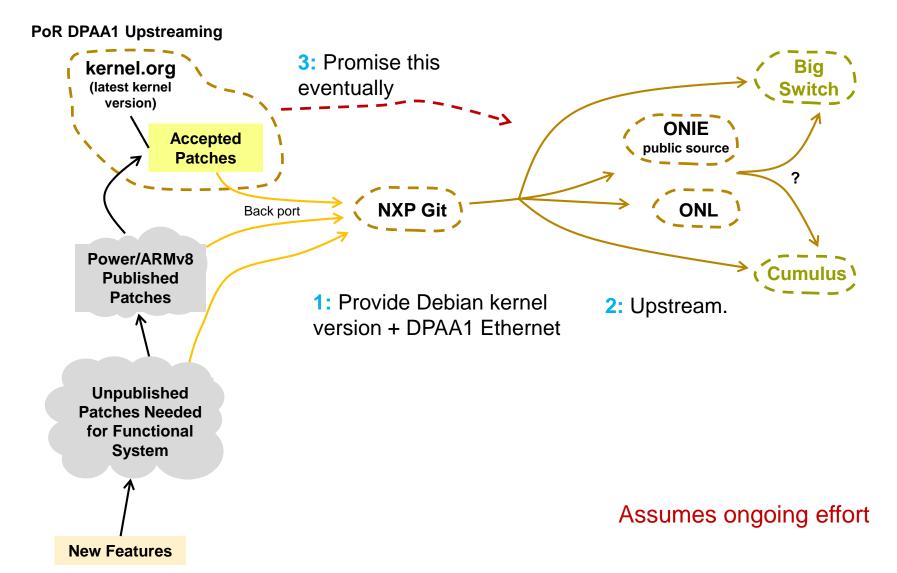
```
n device with of device
                                  index 7115296..07edaf0 100755
Driver Model
                                                                                                             rq save nort with local irq save
                                   -- a/drivers/net/fman/Makefile
                                                                                                             rg restore nort with local irg restore
                                  +++ b/drivers/net/fman/Makefile
                                                                                                             drivers/net
                                  @@ -4.7 +4.7 @@
diff --git a/drivers/net/fman/src/xx/xx linux.c b/drivers/;
                                                               static int fm port remove(struct of device *of dev)
index 57d2a48..84e9c10 100755
 --- a/drivers/net/fman/src/xx/xx linux.c
                                                                     t LnxWrpFmPortDev *p LnxWrpFmPortDev;
                                                                     t LnxWrpFmDev *p LnxWrpFmDev;
+++ b/drivers/net/fman/src/xx/xx linux.c
                                                              @@ -1342,13 +1341,10 @@ static const struct of device id fm port match[] = {
@@ -354,14 +354,14 @@ uint32 t XX DisableAllIntr(void)
                                                               MODULE DEVICE TABLE (of, fm port match);
                                                              #endif /* !MODULE */
     unsigned long flags;
     local irq save(flags);
     return (uint32 t)flags;
                                                               static struct of platform driver fm port driver = {
 void XX RestoreAllIntr(uint32 t flags)
                                                                      .name = "fsl-fman-port",
                                                                      .match table = fm port match,
                                                                      .owner = THIS MODULE,
     local irq restore((unsigned long)flags);
                                                                     .probe = fm port probe,
                                                                     .remove = fm port remove
```



# ONL/ONIE DEVELOPMENT AND UPSTREAMING

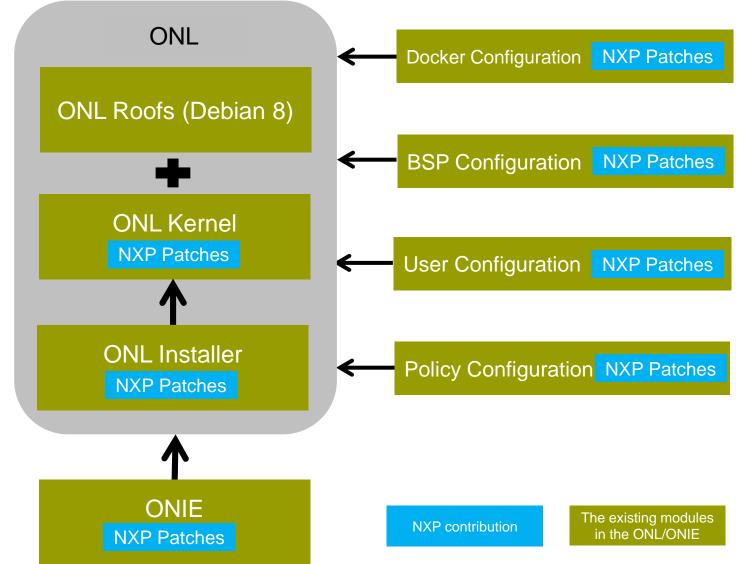


### **General Process**





**Key Modification Needed** 



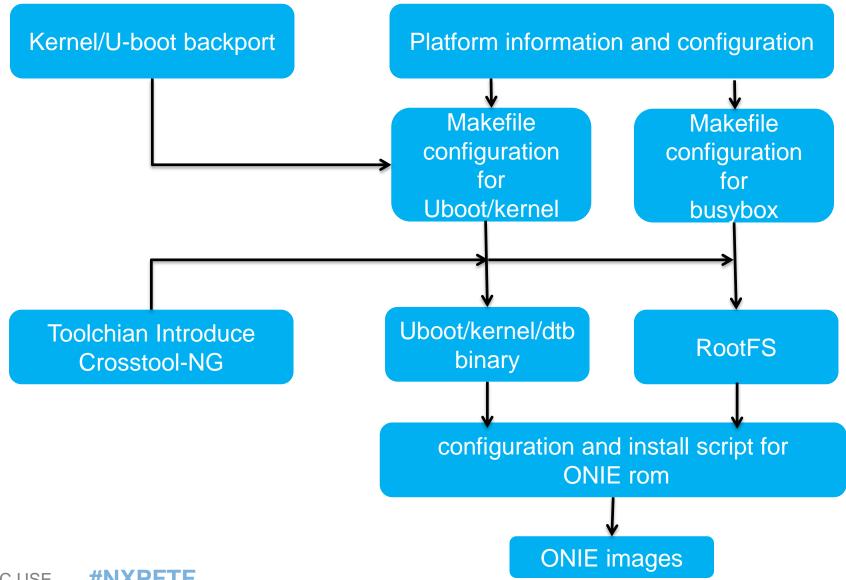


### **Feature List**

- Platform support
- DDR Memory
- Serial Console support
- 1G/10G network port.
- USB host.
- PCIe Host bus.
- IFC NOR.
- IFC NAND.
- SD/MMC
- GPIO
- 12C

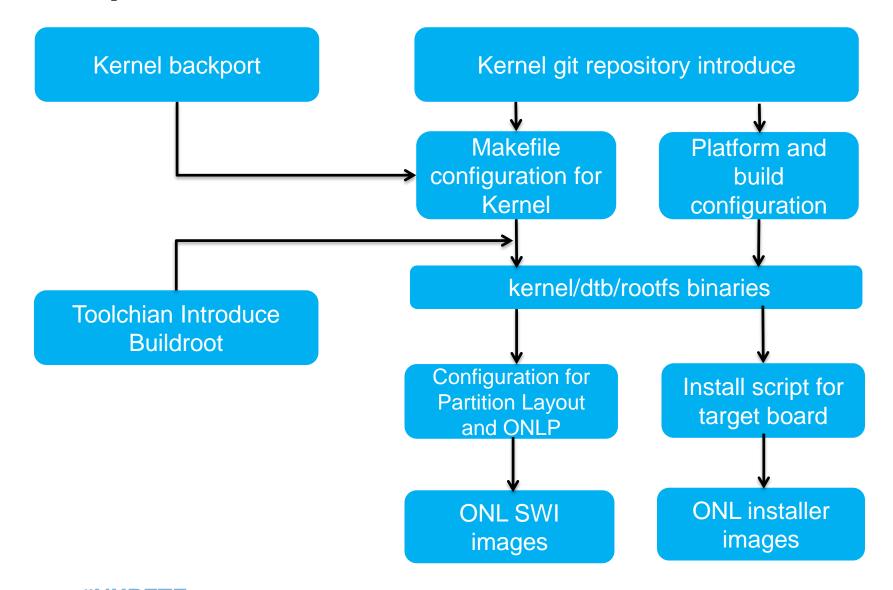


### **ONIE Development Process for New Platform**





### **ONL Development Process for New Platform**





### PLAN & ROADMAP



### Plan for T2080

ID	Task Name	Duration	Start_Date	Finish_Date	Predecessors
1	New ONL warmup and build environment setup	10 days	14-Mar-16	25-Mar-16	
2	T2080 backport to 3.18	10 days	14-Mar-16	25-Mar-16	
3	Backport testing and issue fixing	5 days	28-Mar-16	1-April-16	2
4	ONL configuration development for T2080	5 days	28-Mar-16	1-April-16	1
5	ONL Installer development for T2080	5 days	5-April-16	11-April-16	4
6	Test case design and automation scripts development	10 days	12-April-16	25-April-16	
7	ONL development for T2080	15 days	12-April-16	30-April-16	5
8	ONL testing and issue fixing	5 days	4-May-16	10-May-16	7
9	Legal reviewing and patch submission	3 days	11-May-16	13-May-16	8

Key Milestones



### Plan for LS2080A

ID	Task Name	Duration	Start_Date	Finish_Date	Predecessors
1	ONIE development for LS2080A	10 days	5-April-16	18-April-16	
2	ONIE testing and patch submission	5 days	19-April-16	25-April-16	1
3	LS2080A backport to 3.18	20 days	28-April-16	24-May-16	
4	Backport testing and bug fixing	5 days	25-May-16	31-May-16	3
5	ONL configuration development for LS2080A	10 days	16-May-16	27-May-16	
6	ONL Installer development for LS2080A	10 days	30-May-16	3-June-16	5
7	Test case design and automation script development	5 days	6-June-16	10-June-16	
8	ONL development for LS2080A	20 days	6-June-16	4-July-16	6
9	ONL testing and bug fixing	5 days	5-July-16	11-July-16	8
10	Legal review and ONL patch submission	3 days	12-July-16	14-July-16	9

Key Milestones



### Plan for Upstreaming Maintenance

ID	Task Name	Duration	Start_Date	Finish_Date	Predecessors
1	Track the status on ONL community for Comments handle and patch maintaining	40 days	15-July-16	9-Sep-16	



### Plan for LS1043A ONIE

ID	Task Name	Duration	Start_Date	Finish_Date	Predecessors
1	LS1043A backport to U-boot -2015.10	5 days	28-March-16	1-April-16	
2	LS1043A backport to Kernel-4.1	10 days	5-April-16	18-April-16	1
3	ARMv8 Toolchain support in ONIE	10 days	28-March-16	11-April-16	
4	ONIE configuration and development for LS1043A	10 days	12-April-16	25-April-16	2, 3
5	ONIE testing and bug fixing	5 days	26-April-16	30-April-16	4
6	Legal review and patch submission to ONIE community.	3 days	4-May-16	6-May-16	5
7	Comments handle and push patches be accepted by ONIE community	15 days	9-May-16	27-May-16	6

**Key Milestones** 



## SUMMARY





# SECURE CONNECTIONS FOR A SMARTER WORLD

### ATTRIBUTION STATEMENT

NXP, the NXP logo, NXP SECURE CONNECTIONS FOR A SMARTER WORLD, CoolFlux, EMBRACE, GREENCHIP, HITAG, I2C BUS, ICODE, JCOP, LIFE VIBES, MIFARE, MIFARE, MIFARE Classic, MIFARE DESFire, MIFARE Plus, MIFARE Plus, MIFARE Flex, MANTIS, MIFARE ULTRALIGHT, MIFARE4MOBILE, MIGLO, NTAG, ROADLINK, SMARTLX, SMARTMX, STARPLUG, TOPFET, TrenchMOS, UCODE, Freescale, the Freescale logo, AltiVec, C 5, CodeTEST, CodeWarrior, ColdFire, ColdFire+, C Ware, the Energy Efficient Solutions logo, Kinetis, Layerscape, MagniV, mobileGT, PEG, PowerQUICC, Processor Expert, QorlQ, QorlQ Qonverge, Ready Play, SafeAssure, the SafeAssure logo, StarCore, Symphony, VortiQa, Vybrid, Airfast, BeeKit, BeeStack, CoreNet, Flexis, MXC, Platform in a Package, QUICC Engine, SMARTMOS, Tower, TurboLink, and UMEMS are trademarks of NXP B.V. All other product or service names are the property of their respective owners. ARM, AMBA, ARM Powered, Artisan, Cortex, Jazelle, Keil, SecurCore, Thumb, TrustZone, and µVision are registered trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. ARM7, ARM9, ARM11, big.LITTLE, CoreLink, CoreSight, DesignStart, Mali, mbed, NEON, POP, Sensinode, Socrates, ULINK and Versatile are trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. All rights reserved. Oracle and Java are registered trademarks of Oracle and/or its affiliates. The Power Architecture and Power.org word marks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org. © 2015–2016 NXP B.V.

