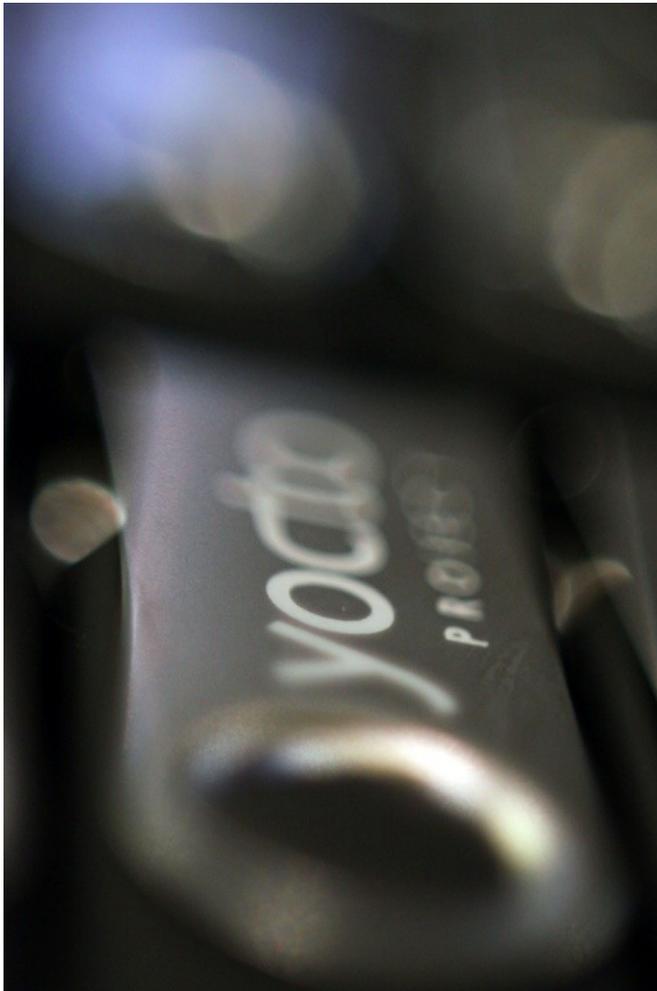


Tuning Embedded Linux

When Less is More

Darren Hart
Intel Corporation
October 17, 2011



Agenda

- Objectives, Motivation, and Target
- Current image type summary reports
- Require concepts and tools
- Iterate over configurations
 - Analyze the kernel and root fs for bloat
 - Identify configuration changes
 - Rebuild and compare reports
- Summary
- Next steps

Objectives

- Reduce raw image size
- Reduce static memory use
- Reduce dynamic memory use
- Minimize boot time

Motivation

- System-on-chip
 - On-die memory is expensive in terms of real-estate and power usage
- Mass market
 - Saving pennies on a smaller flash chip translates to real money
- Performance
 - Smaller images translate to more efficient cache use
- Power usage
 - Less memory means less power
- Smaller images reduce processing due to IO overhead
 - Fewer background services means longer idle states
- Boot time
 - Smaller images translate to less IO and decompression time
- Reduced development overhead
 - Smaller images contain less unnecessary code to build and validate

Real-World Examples

Thank you to the individuals who shared their experiences on the Yocto mailing list to generate these examples.

- Digital camera
 - 10 MB memory
 - Critical boot time
- Medical devices
 - 8 MB flash
 - 4 MB memory
- Network boot RAM FS
 - No flash on device
 - Entire FS in RAM
- Small headless systems
 - 8 MB SPI flash
 - MMC/SD for additional storage
- Partitioned flash
 - Smaller parallel NAND
 - Larger MMC/SD

Target

- Generate a Kernel + RootFS in under 4MB
- Boot in under 8MB
 - (4MB would be better)
- Boot to shell in under 2 seconds
- Maintain ipv4 functionality
- Avoid an initial RAM disk
 - (No cheating by building everything as modules)
- We'll use qemu86 for the purposes of this exercise

Sato: Size Report

- Contents

- Linux kernel
- Eglibc
- Udev
- Login
- X Server
- Sato Desktop and Applications

- Size Report

- BzImage: 4.0 MB
- RootFS: 118.0 MB
- Modules: ~~35.0 MB~~
- **Total: 122.0 MB**

- Memory Report

- RAM: 128 MB
- Early boot: 9.8 MB
- Login: 82.3 MB
- Kernel Freed: 444 KB

- Boot Time

- Kernel*: 4.26s
- Sato Desktop: 21.9s

* At "Freeing unused kernel memory"

Minimal: Size Report

- Contents

- Linux kernel
- Eglibc
- Udev
- Login

- Memory Report

- RAM: 32 MB
- Early boot: 8.6 MB
- Login: 15.8 MB
- Kernel Freed: 444 KB

- Size Report

- BzImage: 4.0 MB
- RootFS: 11.0 MB (-107.0 MB)
- Modules: ~~35.0 MB~~
- **Total: 15.0 MB**

- Boot Time

- Kernel: 3.84s
- Login: 9.5s

Components

- Root filesystem
 - Packages
 - Boot
 - Libraries
 - Applications
 - Package configuration
 - Filesystem
- Linux kernel
 - Policy
 - Subsystems
 - Architecture
 - Drivers

Guiding Principles

- Prepare a budget
 - Linux Kernel: 1 MB
 - Root FS: 3 MB
- Don't sweat the small stuff (90% rule)
- Avoid difficult to maintain hacks
 - At first anyway...
- Leverage device specific options
- Develop in a separate layer

Concepts: Storage

- ELF Sections
 - text: the code itself
 - data: initialized data
 - bss: uninitialized data
- Image Size
 - Includes text and data sections only, not bss.
- Measure size in blocks with df (not in bytes with du)

```
$ df mnt-stage1/  
Filesystem 1K-blocks  Used  Available Use% Mounted on  
/dev/loop1 8059    5407    2243    71% mnt-stage1
```

Concepts: Memory

- Static Memory
 - The text, data, and bss sections.
- Dynamic Memory
 - Memory allocated at runtime
 - Stacks
 - Hashtables
 - Allocators
 - Page Cache
 - Reservations
- Temporary Memory
 - Decompression
 - `__init__`

Tools

- Identify, quantify, and record your changes

- ksize.py
- dirsize.py
- merge_config.sh

```
$ bitbake -u depexp -g core-image-*
```

- Scripts available here until merged upstream

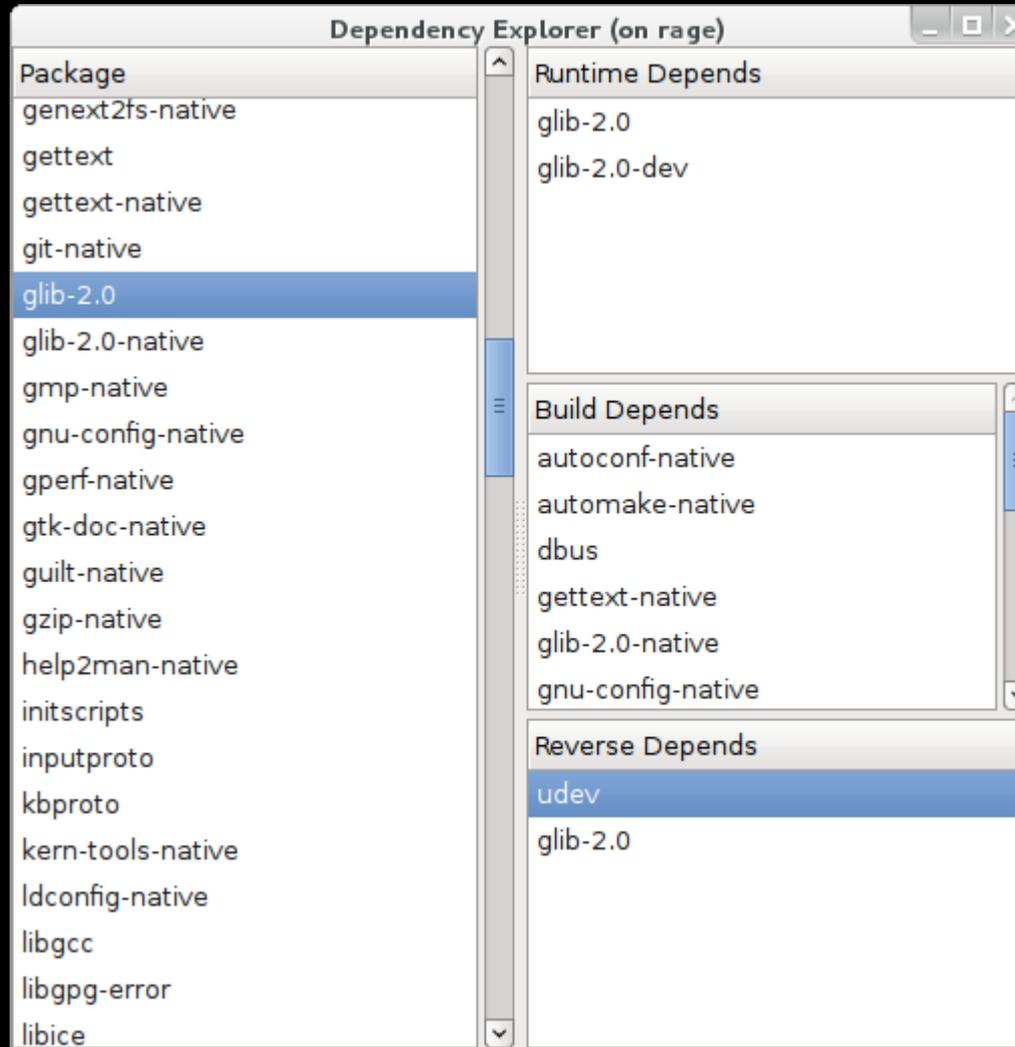
- <http://dvhart.com/darren/yocto/tiny/>

Minimal: Root FS

```
$ cat dirsize-100k.log
9850251 .
3878968 ./lib
1457504 ./lib/libc-2.13.so
173908 ./lib/libm-2.13.so
158617 ./lib/libacl.so.1.1.0
127228 ./lib/ld-2.13.so
...
696977 ./lib/modules/3.0.4-yocto-standard+/kernel/drivers/video
645004 ./lib/udev
2907574 ./usr
2516900 ./usr/lib
1047940 ./usr/lib/libgio-2.0.so.0.2800.8
1036944 ./usr/lib/libglib-2.0.so.0.2800.8
249756 ./usr/lib/libgobject-2.0.so.0.2800.8
299502 ./usr/share
170680 ./usr/share/pci.ids.gz
124206 ./usr/share/usb.ids.gz
1263456 ./sbin
691588 ./sbin/ldconfig
137012 ./sbin/udevadm
133132 ./sbin/udev
115932 ./sbin/v86d
1138391 ./etc
1044480 ./etc/dev.tar
659740 ./bin
602752 ./bin/busybox
Displayed 7968656/9850251 bytes (80.90%)
```

Glib?

```
$ bitbake -u depexp -g core-image-minimal
```



Minimal → Stage 1

- Reduce size with minimal impact on features
- We can get by with devtmpfs and mdev
- We don't **need** a VGA display, we have serial
- Drop udev and v86d

```
$ cat conf/local.conf | head -n 7
##### MINIMAL STAGE 1 Mods #####
# Drop udev (and glib) and use mdev
# Save 4MB from minimal image rootfs!
VIRTUAL-RUNTIME_dev_manager = ""
# Drop v86d from qemu86 required packages
MACHINE_ESSENTIAL_EXTRA_RDEPENDS_qemu86 = ""
```

Filesystem Options

- Minimal builds ext3 by default
- ext3 requires a 1k block journal
 - 1 MB with 1024 byte blocks (instead of 4096)
- If we don't **need** the journal, we can save 1 MB by using ext2
 - 5.3 MB ext3
 - 4.0 MB ext2
- For a small image, you are most likely going to use JFFS2 or UBIFS anyway

Stage 1: Size Report

- Contents

- Linux kernel
- Eglibc
- Login

- Size Report

- bzImage: 4.0 MB (minimal)
- rootfs: 4.0 MB (-7.0 MB)
- ~~modules: 35.0 MB~~
- **Total: 8.0 MB (-7.0 MB)**

- Memory Report

- RAM: 32 MB
- Early boot: 8.6 MB
- Login: 15.7 MB
- Kernel Freed: 444 KB

- Boot Time

- Kernel: 3.54s
- Login: 7.19s

Stage 1: Root FS

```
$ cat dirsize-30k.log
3878774 .
2242550 ./lib
1457504 ./lib/libc-2.13.so
173908 ./lib/libm-2.13.so
127228 ./lib/ld-2.13.so
96624 ./lib/libpthread-2.13.so
91956 ./lib/libnsl-2.13.so
79620 ./lib/libresolv-2.13.so
46672 ./lib/libnss_files-2.13.so
35956 ./lib/libcrypt-2.13.so
34588 ./lib/libnss_compat-2.13.so
30624 ./lib/librt-2.13.so
807168 ./sbin
691588 ./sbin/ldconfig
34300 ./sbin/init.sysvinit
659740 ./bin
602752 ./bin/busybox
50308 ./bin/tinylogin
87565 ./usr
50168 ./usr/bin
80786 ./etc
34406 ./etc/init.d
Displayed 3553628/3878774 bytes (91.62%)
```

Stage 1: Kernel

```
$ ls -s bzImage
4064 bzImage-qemu86.bin
```

```
$ cat ksize.log
```

```
Linux Kernel
```

	total	text	data	bss
-----	-----	-----	-----	-----
vmlinux	9657412	7538548	529616	1589248
-----	-----	-----	-----	-----
drivers/built-in.o	2549250	2385650	133508	30092
net/built-in.o	1194464	1137786	29358	27320
kernel/built-in.o	1033129	723329	45832	263968
fs/built-in.o	948917	926681	18564	3672
sound/built-in.o	699821	684877	9624	5320
arch/x86/built-in.o	459019	277038	87265	94716
mm/built-in.o	345158	294330	23816	27012
block/built-in.o	126489	119272	5741	1476
crypto/built-in.o	84412	82364	2028	20
lib/built-in.o	52607	52561	38	8
security/built-in.o	46993	44778	1879	336
ipc/built-in.o	36996	35880	1100	16
init/built-in.o	31256	20186	10921	149
firmware/built-in.o	15375	15375	0	0
usr/built-in.o	516	516	0	0
-----	-----	-----	-----	-----
sum	7624402	6800623	369674	454105
delta	2033010	737925	159942	1135143

Stage 1 → Stage 2

- 91.62% of the Root FS is composed of:
 - Eglibc
 - Busybox
- 66.53% of the Kernel image is composed of:
 - Drivers
 - Networking
 - Core kernel
 - Filesystems
 - Sound
- Bound to be more fluff in the kernel image

Drivers

drivers	total	text	data	bss
drivers/built-in.o	2549250	2385650	133508	30092
drivers/net/built-in.o	499378	488591	10339	448
drivers/usb/built-in.o	256540	226215	27697	2628
drivers/md/built-in.o	245896	240667	4017	1212
drivers/acpi/built-in.o	245894	218314	25752	1828
drivers/ata/built-in.o	198861	183896	10761	4204
drivers/tty/built-in.o	196733	165026	26755	4952
drivers/scsi/built-in.o	123556	117492	5516	548
drivers/input/built-in.o	115474	112337	2709	428
drivers/pci/built-in.o	105975	101094	2733	2148
drivers/ide/built-in.o	104091	102287	1540	264
drivers/video/built-in.o	95058	86002	1180	7876
drivers/hid/built-in.o	78498	74450	4012	36
drivers/base/built-in.o	62975	61402	1481	92
drivers/pnp/built-in.o	34517	33268	1233	16
drivers/cdrom/built-in.o	28387	26847	484	1056
drivers/rtc/built-in.o	21447	20851	452	144
drivers/i2c/built-in.o	19640	18999	612	29
drivers/char/built-in.o	13472	11644	824	1004
drivers/thermal/built-in.o	9002	8206	760	36
drivers/gpu/built-in.o	7977	7869	92	16
drivers/firmware/built-in.o	7534	6730	580	224
drivers/cpuidle/built-in.o	7176	6548	604	24
drivers/power/built-in.o	5199	4251	740	208
drivers/leds/built-in.o	4125	3997	124	4
drivers/connector/built-in.o	4060	4000	24	36
drivers/block/built-in.o	3344	3276	56	12
drivers/clocksource/built-in.o	1956	1656	292	8
drivers/hwmon/built-in.o	818	790	8	20
sum	2497583	2336705	131377	29501
delta	51667	48945	2131	591

Networking

net	total	text	data	bss
net/built-in.o	1194464	1137786	29358	27320
net/ipv4/built-in.o	364644	346523	13037	5084
net/core/built-in.o	196473	188607	4781	3085
net/sunrpc/built-in.o	178398	158816	3102	16480
net/mac80211/built-in.o	152576	152020	444	112
net/wireless/built-in.o	131551	128631	2664	256
net/xfrm/built-in.o	52381	50921	1076	384
net/sched/built-in.o	22183	21023	1148	12
net/netlink/built-in.o	21614	20934	520	160
net/unix/built-in.o	19811	18423	348	1040
net/*.o	16690	16282	392	16
net/packet/built-in.o	16356	16092	264	0
net/netfilter/built-in.o	9509	7637	1268	604
net/ipv6/built-in.o	4865	4865	0	0
net/dns_resolver/built-in.o	3525	3457	60	8
net/ethernet/built-in.o	1887	1875	12	0
net/8021q/built-in.o	1386	1386	0	0
sum	1193849	1137492	29116	27241
delta	615	294	242	79

Core Kernel

kernel	total	text	data	bss
kernel/built-in.o	1033129	723329	45832	263968
kernel/*.o	535934	466134	24338	45462
kernel/trace/built-in.o	305798	142282	14860	148656
kernel/time/built-in.o	94008	40975	3065	49968
kernel/events/built-in.o	40549	39613	808	128
kernel/debug/built-in.o	29591	10074	190	19327
kernel/irq/built-in.o	20706	18754	1924	28
kernel/power/built-in.o	4442	4278	148	16
sum	1031028	722110	45333	263585
delta	2101	1219	499	383

Filesystems

fs	total	text	data	bss
fs/built-in.o	948917	926681	18564	3672
fs/*.o	319243	312988	4435	1820
fs/nfs/built-in.o	230495	222498	7765	232
fs/ext3/built-in.o	104159	104087	60	12
fs/proc/built-in.o	68568	68244	236	88
fs/lockd/built-in.o	56621	51349	4144	1128
fs/ext2/built-in.o	50828	50728	92	8
fs/jbd/built-in.o	37086	37038	28	20
fs/quota/built-in.o	22937	22225	588	124
fs/sysfs/built-in.o	19958	19526	396	36
fs/notify/built-in.o	16864	16552	264	48
fs/debugfs/built-in.o	9259	9195	48	16
fs/partitions/built-in.o	7571	7311	260	0
fs/nls/built-in.o	4636	4572	64	0
fs/devpts/built-in.o	2335	2263	68	4
fs/ramfs/built-in.o	2304	1976	328	0
sum	952864	930552	18776	3536
delta	-3947	-3871	-212	136

Sound

sound	total	text	data	bss
sound/built-in.o	699821	684877	9624	5320
sound/pci/built-in.o	482464	474748	6972	744
sound/core/built-in.o	212882	205834	2596	4452
sound/*.o	9256	8620	444	192
sum	704602	689202	10012	5388
delta	-4781	-4325	-388	-68

Linux Kernel Config Fragments

- Entire defconfigs make it difficult to quantify cost of individual options
- Better to assemble config fragments
- Avoid modules and the initial RAM disk
- Start with allnoconfig
- Merge fragments with `merge_config.pl`
 - Generates a `.config`
 - Warns on overrides
 - Warns on missing `CONFIG_` options
(possibly due to missing dependencies)

Minimal Linux Kernel Config

- Start with the bare minimal for an x86-32 machine:
 - defconfig (x86_32_allnoconfig)
 - core.cfg
 - smp.cfg
 - rtc-pc.cfg
- Some basic policy:
 - serial.cfg
 - devtmpfs.cfg
 - sysfs.cfg
 - ext2.cfg
 - ext3.cfg
 - net.cfg
 - vt.cfg
 - fb.cfg
 - debug.cfg
- QEMU “hardware” support
 - ata.cfg
 - e1000.cfg
 - floppy.cfg
 - usb.cfg
 - vga.cfg
 - intel-hda.cfg

Stage 2: Size Report

- Contents

- Linux kernel
- Eglibc
- Login

- Size Report

- BzImage: 1.8 MB (-2.2 MB)
- RootFS: 4.0 MB (stage 1)
- **Total: 5.8 MB (-2.2 MB)**

- Memory Report

- RAM: 32 MB
- Early boot: 4.49 MB
- Login: 9.37 MB
- Kernel Freed: 240 KB

- Boot Time

- Kernel: 0.90s
- Login: 3.38s

Stage 2: Kernel

```
$ ls -s bzImage
4064 bzImage-qemux86.bin
```

```
$ cat ksize.log
```

```
Linux Kernel
```

	total	text	data	bss
-----	-----	-----	-----	-----
vmlinux	5214442	3569634	276744	1368064
-----	-----	-----	-----	-----
drivers/built-in.o	1285171	1175622	78161	31388
sound/built-in.o	559278	548606	8456	2216
kernel/built-in.o	538539	322032	77555	138952
net/built-in.o	475916	451509	17507	6900
fs/built-in.o	456887	451541	3370	1976
arch/x86/built-in.o	289285	219562	44515	25208
mm/built-in.o	231360	189117	16543	25700
block/built-in.o	77877	74707	1722	1448
lib/built-in.o	33087	32999	80	8
ipc/built-in.o	22097	21365	724	8
init/built-in.o	13549	8215	5221	113
security/built-in.o	3738	3722	8	8
-----	-----	-----	-----	-----
sum	3986784	3498997	253862	233925
delta	1227658	70637	22882	1134139

Stage 2 → Stage 3

- 91.62% of the Root FS is composed of:
 - eglibc
 - busybox
- 44.13% of the Kernel image is composed of:
 - drivers
 - sound
 - Filesystems
- Let's see what we can shave off from each

Kernel: Only the Essentials

- Drop everything but the essentials for boot, serial console, and networking
- Drop from policy
 - vt.cfg
 - ext3.cfg
 - fb.cfg
- Drop from Qemux86 “Hardware” support
 - floppy.cfg
 - usb.cfg
 - vga.cfg
 - intel-hda.cfg

Root FS: Busybox

- Drop all the vt services from busybox, this needs a simple patch to avoid opening tty devices
- Drop ipv6 and all the Linux module utilities
- Use a busybox bbappend recipe and a new defconfig

Root FS: eglibc

```
# Reconfigure eglibc for a smaller installation
# Comment out any of the lines below to disable them in the eglibc build
DISTRO_FEATURES_LIBC_TINY = "libc-libm libc-crypt"
DISTRO_FEATURES_LIBC_REGEX = "libc-posix-regex"
DISTRO_FEATURES_LIBC_NET = "libc-inet libc-nis"
DISTRO_FEATURES_LIBC_MINIMAL = "libc-utmp libc-getlogin"

DISTRO_FEATURES_LIBC = "${DISTRO_FEATURES_LIBC_TINY} \
    ${DISTRO_FEATURES_LIBC_MINIMAL} \
    ${DISTRO_FEATURES_LIBC_REGEX} \
    ${DISTRO_FEATURES_LIBC_NET}"

# Comment out any of the lines below to disable them in the build
DISTRO_FEATURES_TINY = "ext2 pci"
DISTRO_FEATURES_NET = "ipv4 nfs"

DISTRO_FEATURES = "${DISTRO_FEATURES_TINY} \
    ${DISTRO_FEATURES_NET} \
    ${DISTRO_FEATURES_LIBC}"
```

- Dropping 'who' and tools like 'grep' and 'sed' allow the removal of libc-posix-regex, libc-utmp, and libc-getlogin, but start to limit functionality

Root FS: System Services

- Drop tinylogin, modutils-initscripts, and netbase
- Define a new image type, core-image-tiny which is built using a new task-core-tiny task

```
RDEPENDS_task-core-tiny = "base-files base-passwd \  
                          busybox initscripts"  
  
# task-core-tiny RDEPENDS on a subset of what task-core-boot does:  
#RDEPENDS_task-core-boot = "base-files base-passwd \  
#                          busybox initscripts \  
#                          ${@base_contains("MACHINE_FEATURES", "keyboard", "keymaps", "", d)} \  
#                          modutils-initscripts netbase \  
#                          ${VIRTUAL-RUNTIME_login_manager} \  
#                          ${VIRTUAL-RUNTIME_init_manager} \  
#                          ${VIRTUAL-RUNTIME_dev_manager} \  
#                          ${VIRTUAL-RUNTIME_update-alternatives} \  
#                          ${MACHINE_ESSENTIAL_EXTRA_RDEPENDS}"
```

Stage 3: Size Report

- Contents

- Linux kernel
- Eglibc
- Busybox shell

- Size Report

- BzImage: 1.2 MB (-0.6 MB)
- RootFS: 3.2 MB (-0.8 MB)
- **Total: 4.4 MB (-1.4 MB)**

- Memory Report

- RAM: 32 MB
- Early boot: 3.42 MB
- Login: 6.66 MB
- Kernel Freed: 220 KB

- Boot Time

- Kernel: 0.60s
- Shell: 2.13s

Now What?

- Kernel
 - networking
 - SMP
 - ACPI
 - SysV IPC, Futexes
 - Printk
- Eglibc
 - networking
 - regular expressions
- To get below 4.0 MB, we should look at uclibc

Stage 3 → Stage 4

- Switch to uclibc

```
DISTRO_FEATURES_NET = "ipv4 nfs"

DISTRO_FEATURES = "${DISTRO_FEATURES_TINY} \
                  ${DISTRO_FEATURES_NET} \
                  ${DISTRO_FEATURES_LIBC}"

TCLIBC = "uclibc"
```

Stage 4: Size Report

- Contents

- Linux kernel
- Uclibc
- Busybox shell

- Size Report

- BzImage: 1.2 MB (stage 3)
- RootFS: 1.5 MB (-1.7 MB)
- **Total: 2.7 MB (-1.7 MB)**

- Memory Report

- RAM: 32 MB
- Early boot: 3.42 MB
- Login: 5.84 MB
- Kernel Freed: 220 KB

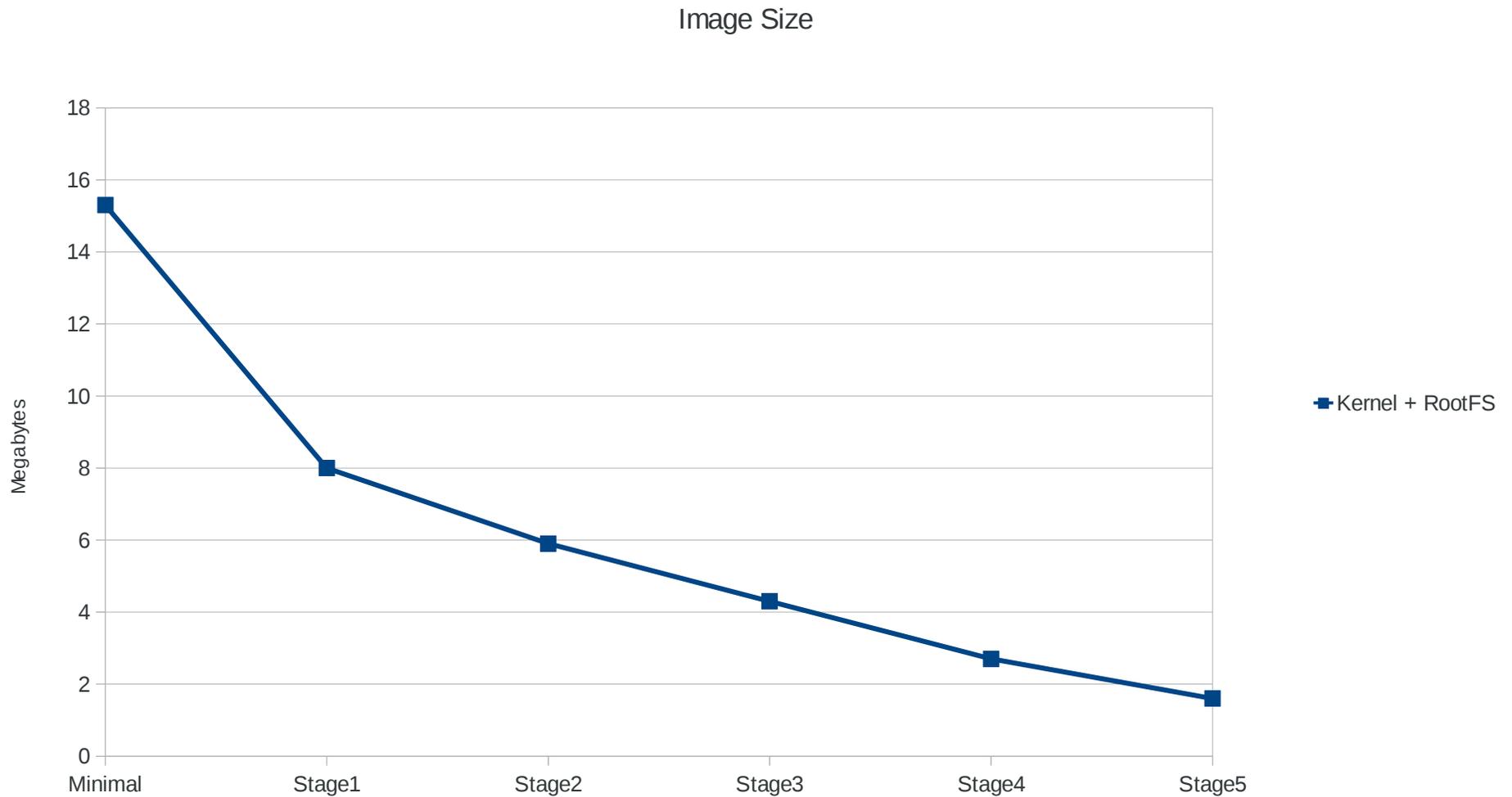
- Boot Time

- Kernel: 0.61s
- Shell: 2.07s

Stupid Small

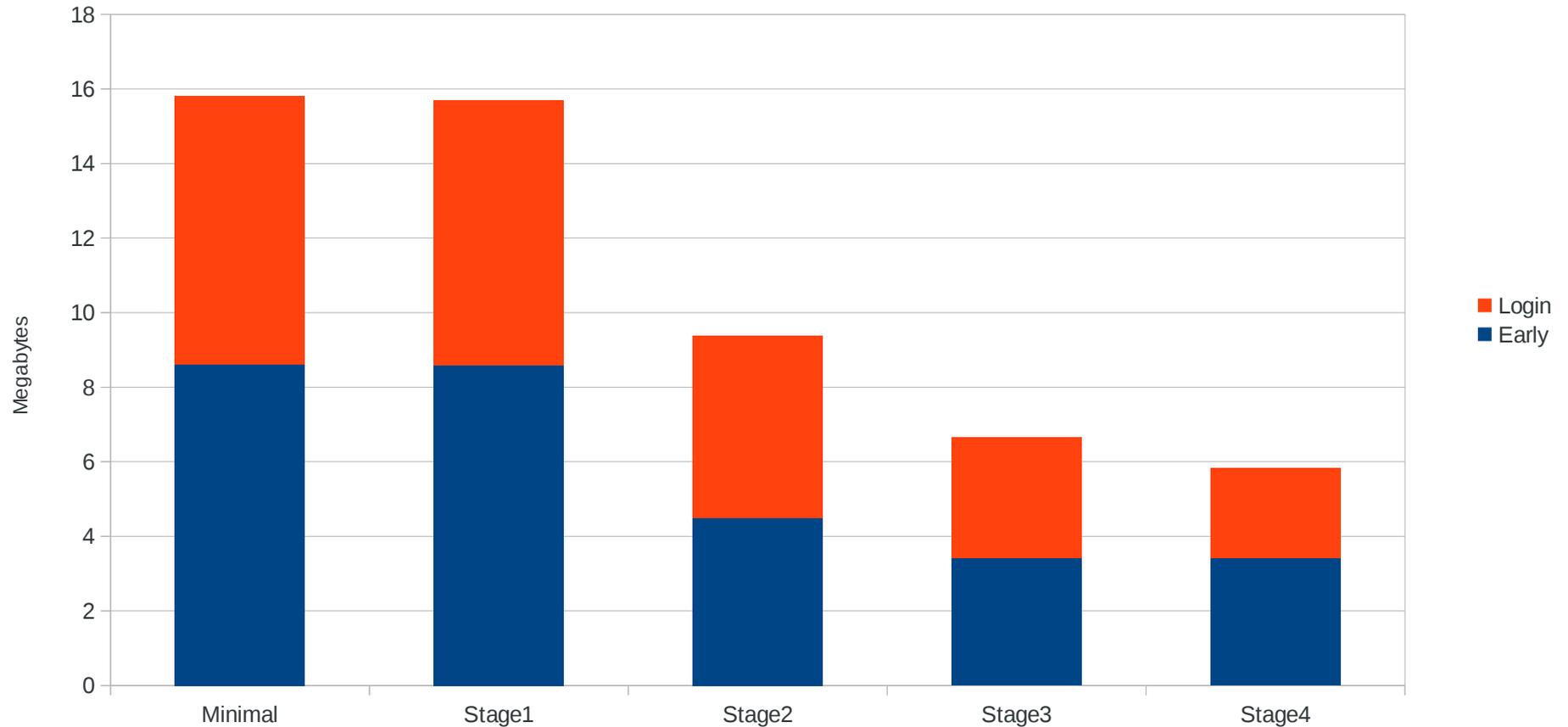
- You can go further still if you want
 - Drop networking support (uclibc and kernel)
 - Cripple Busybox (grep, network tools, etc)
 - Cripple Linux kernel (acpi, smp, ipc, futex, printk)
- Size Report
 - bzImage: 585K (-0.7 MB)
 - rootfs: 1.1MB (-0.4 MB)
 - **Total: 1.6 MB (-1.1 MB)**
- Memory Report
 - We removed printk and proc!
- Boot Time
 - Shell: 1.28s
- You have lost a lot of functionality to get here

Image Size by Stage

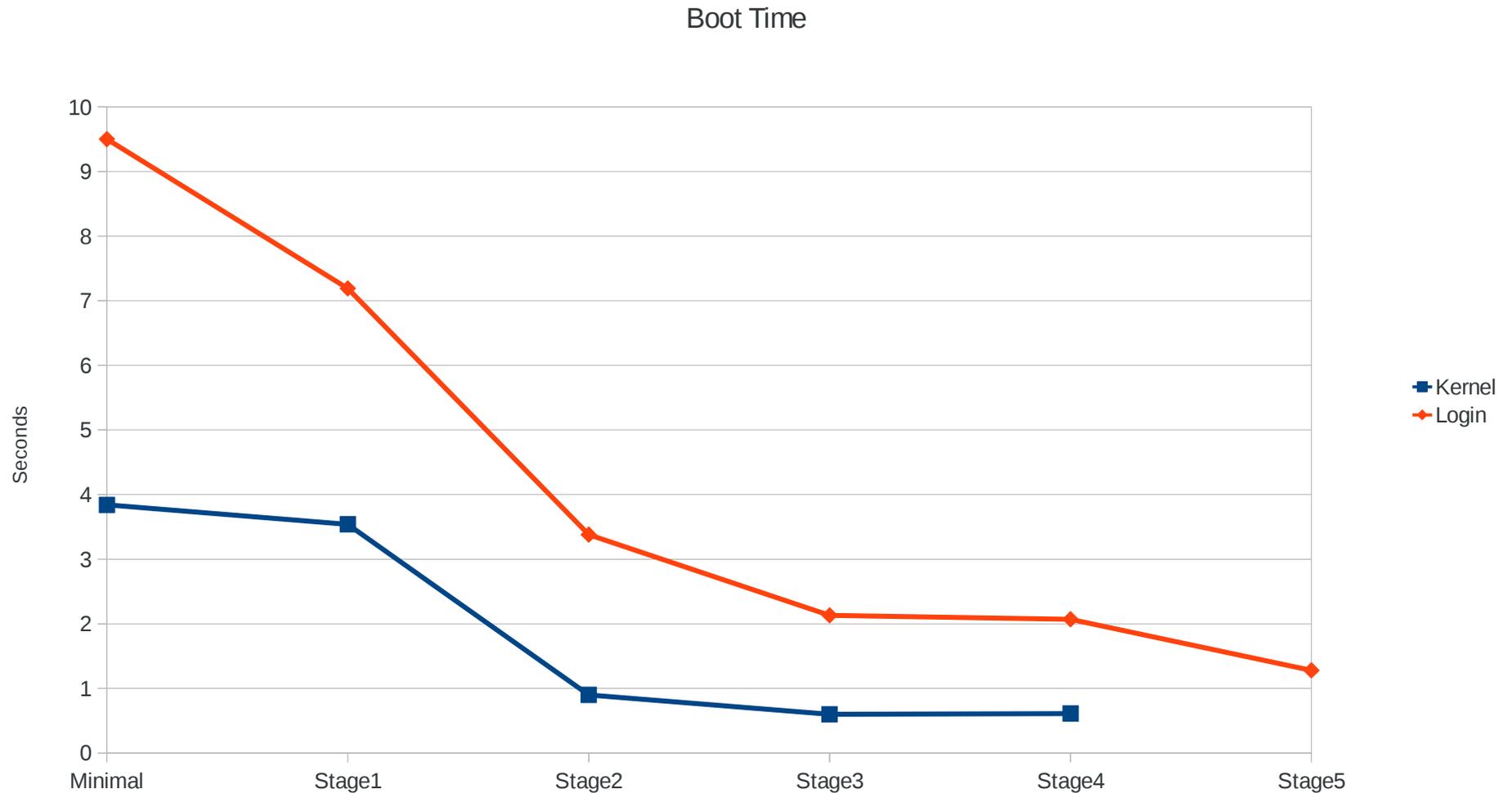


Memory Usage Summary

Memory Usage



Boot Time Summary



Next Steps

- Bitbake config fragments
 - Incorporate config fragment management from the Yocto Project kernel tools into the Bitbake recipe
- Distribution package feature mechanism
 - Prepare a distro package feature configuration mechanism for fine tuning recipe configs, such as bitbake and linux-yocto.
 - Eglibc and uclibc have similar mechanisms, but may need to be modified for a consistent implementation across recipes.
- Define one or more poky-tiny distributions and images
 - Your input is needed here
 - Do we define a no-network image?
 - Do we define a smaller graphical image?
 - Perhaps something with directfb instead of X

Resources

- Yocto Project and Meta-Tiny
 - <http://www.yoctoproject.org>
 - <http://git.yoctoproject.org/cgi/user-contrib/dvhart/meta-tiny>
- ELCE 2010 Videos
 - The Right Approach to Minimal Boot Time by Andrew Murray
 - <http://free-electrons.com/blog/elce-2010-videos/>
- Andi Kleen's Memory Usage Papers
 - <http://halobates.de/memorywaste.pdf>
 - <http://halobates.de/memory.pdf>
- Phil Blundell's meta-micro layer
 - <http://cgит.openembedded.org/meta-micro/>

yocto .

PROJECT

Legal

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT.

Intel may make changes to specifications, product descriptions, and plans at any time, without notice.

All dates provided are subject to change without notice.

Intel is a trademark of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2011, Intel Corporation. All rights are protected.