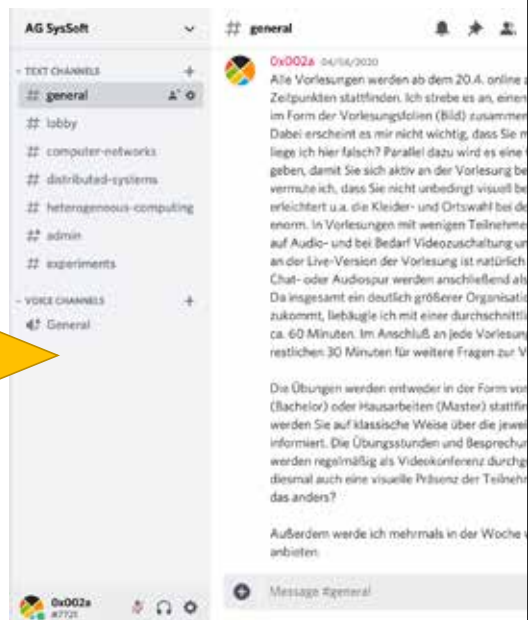


Kommunikationsformen

- Studip
 - Folienkopien
 - Übungsblätter
- Discord Server
 - Link in studip
- Email
 - sturm@uni-trier.de



2

<https://discord.gg/XEE4cAs5>

3

Übung (Portfolio)

- 3 Übungsblätter (jeweils maximal 10 Punkte)
 1. Architektur (Anfang November)
 2. Virtualisierung (Anfang Dezember)
 3. Concurrency (Anfang Januar)
- 1 Abschlußprojekt (maximal 20 Punkte)
 - Messen und Auswerten (Performanz, Latenz)
 - Wenige Seiten Ausarbeitung
 - Präsentation
- Note ergibt sich aus der erreichten Punktzahl

4



Was können wir?

- Grundlagen Betriebssysteme
- Virtueller Speicher
- Virtuelle Prozessoren (Threads)
- Locks, Mutex, Semaphore
- Dateisysteme
- C

6

Was wollen wir? Ich? 😊

- OS Architekturen
- Virtuelle Maschinen
- Container / Docker
- Many Cores / Concurrency / Parallelism
- Moderne Dateisysteme

7

1. Architektur

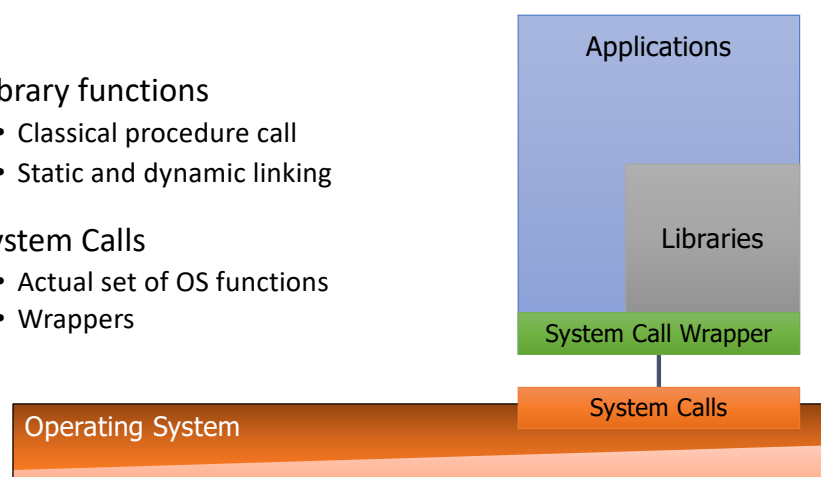
8

Von Monolithen und Mikrokernen

9

Ways to Access OS Functionality

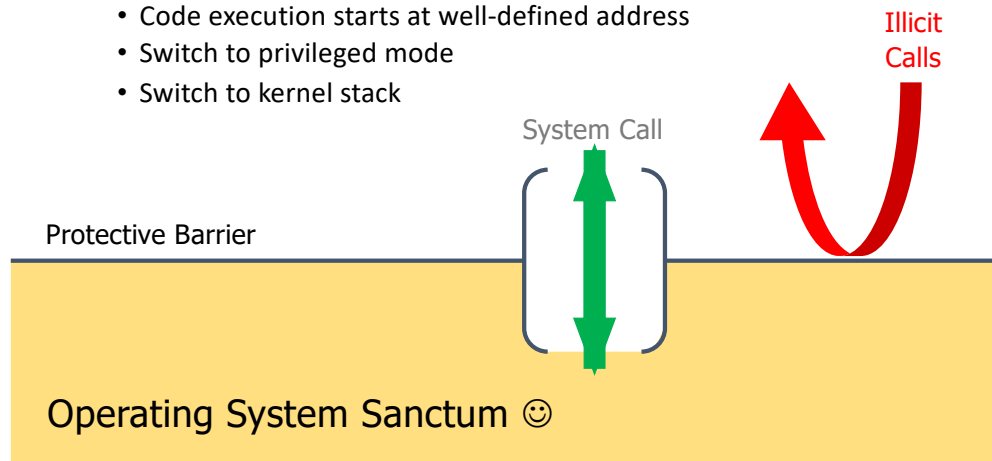
- Library functions
 - Classical procedure call
 - Static and dynamic linking
- System Calls
 - Actual set of OS functions
 - Wrappers



10

System Call

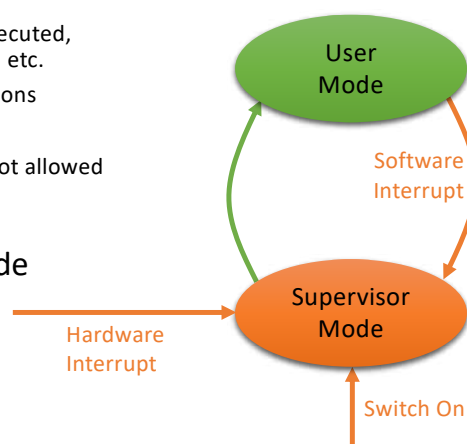
- The only legitimate access to the OS kernel
- System calls = software interrupt
 - Code execution starts at well-defined address
 - Switch to privileged mode
 - Switch to kernel stack



11

Supervisor Mode & User Mode

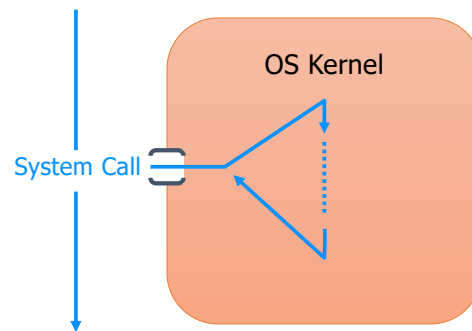
- Modern OS require CPU to differ between
 - Supervisor mode
 - Every instruction can be executed, including configuring MMU etc.
 - Access to all memory locations
 - User mode
 - Sensitive instructions are not allowed
 - Restricted memory access
- Only OS in supervisor mode



12

Everything Is Inside The Kernel

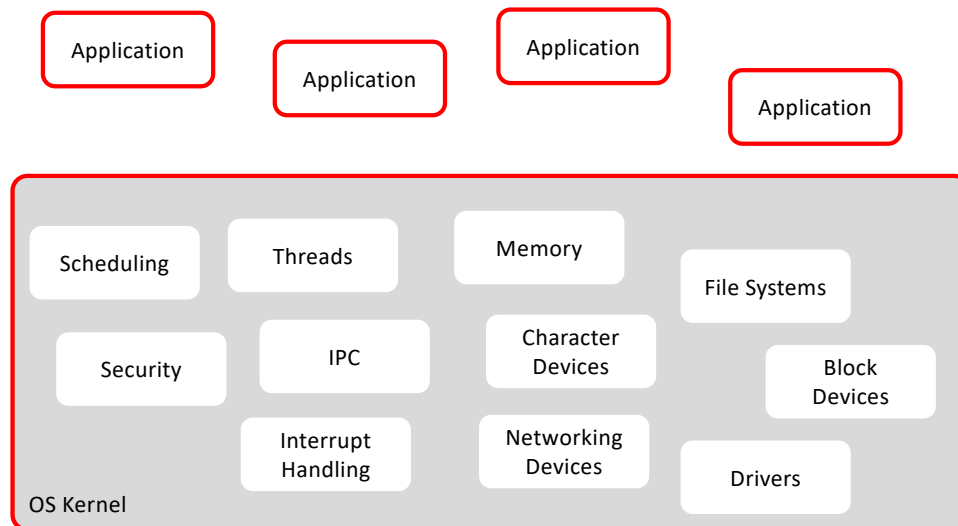
- Thread issuing system call „stays“ inside kernel until call is completed
- All required functionality must be inside kernel
- Context-switch in case of blocking
- **Monolith**



13



Monolith

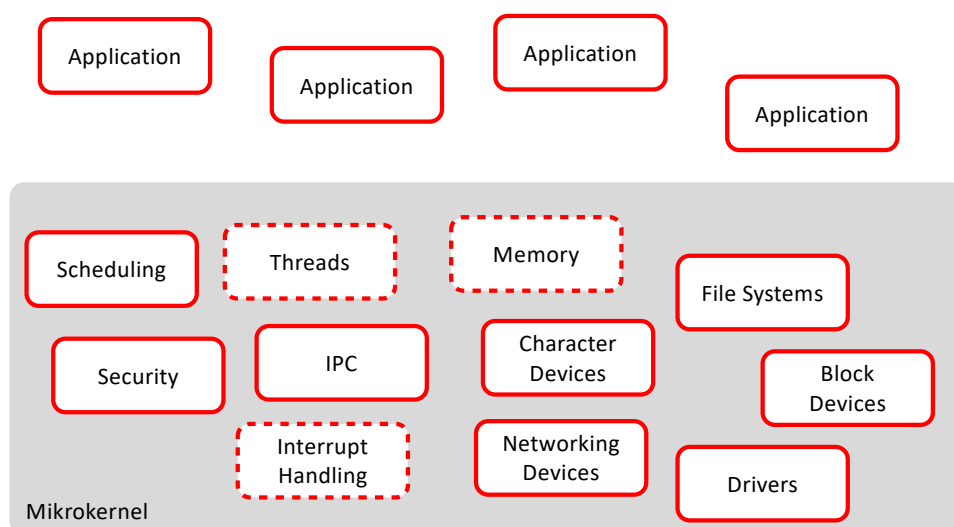


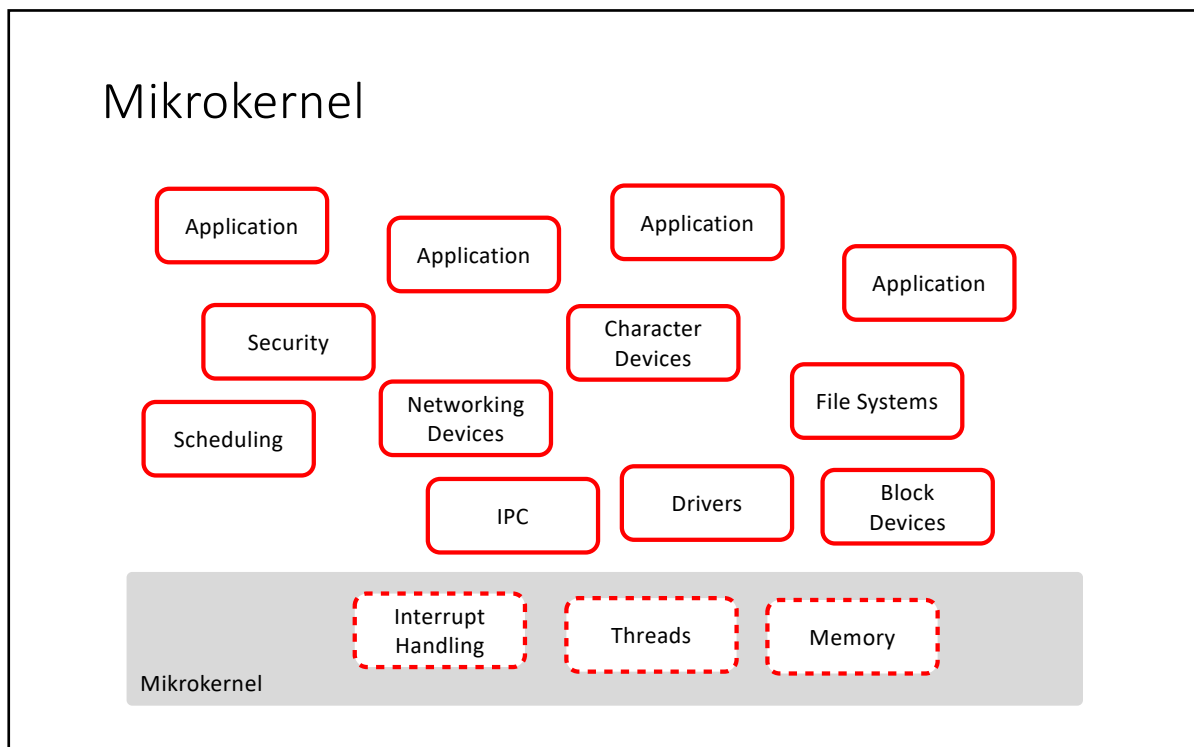
Cathedral



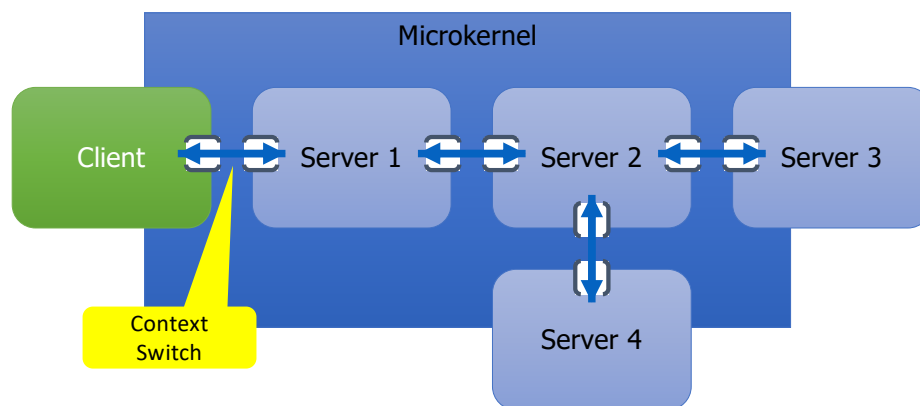


Mikrokernel





Next to Nothing Inside Kernel



21

Microkernel

- Implements essential abstractions
 - Protection
 - Communication primitives to cross address spaces
 - Interrupt preprocessing and delivery
- Goal: No single application can monopolize resources
- Dislocating all remaining functionality into user mode server
 - File server, Communication server, Graphics server, ...
- From a scientific point of view the most attractive solution

22

Comparison

Monolith

- Pro
 - Fast
- Contra
 - Keeping a good architecture and structure requires a lot of self-discipline
 - No isolation between different parts of kernel
 - Hard to extend
 - Complex to adapt

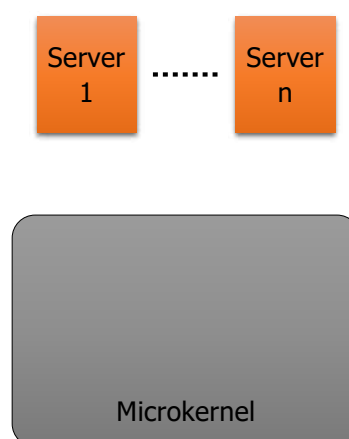
Microkernel and Server

- Pro
 - Good architecture and structure possible
 - Easier to extend
 - Easier to adapt
 - Debugging of servers with user mode tools
- Contra
 - Slow (too many context switches)

23

Accelerating Microkernels

- Sensible and/or time-critical servers return into kernel
 - Microkernel becomes „still structured“ monolith
- Realization
 - Additional linking step replaces IPC with functions calls
 - Of course, no protection anymore
- Applied in most modern an microkernel-based operating systems
 - MACH
 - NT and subsequent systems



24

Improving Monoliths

- Example Linux
 - Introduction of a module concept
 - Parts of operating systems (especially drivers) will be loaded on demand
 - Most modules can be unloaded again
- Execution, testing, and debugging not in user mode
- Still no protection



25

LINUX

26

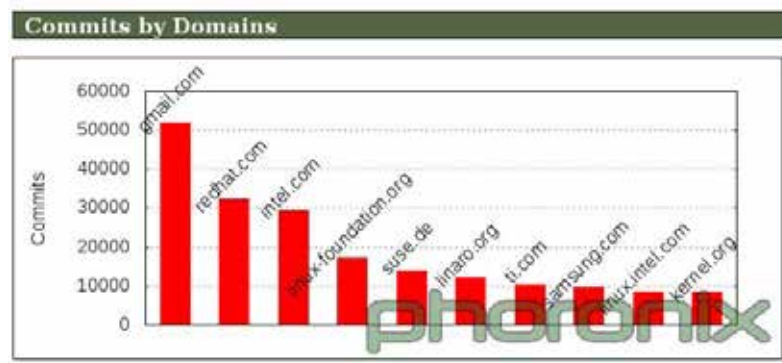
Linux 5.3.7 (21.10.2019)

- 108.6 MB .tar.xz
- 968 MB mit 65.216 Dateien
- Insgesamt
 - Assembler 376.252
 - C Source (*.c) 18.743.746
 - C Header (*.h) 5.585.729
- Davon in arch, drivers
 - Assembler 370.400+2.459
 - C (*.ch) 2.216.470+16.136.460
- Also reiner Kernel
 - 3393 Assembler
 - 5.976.545 C (*.ch)



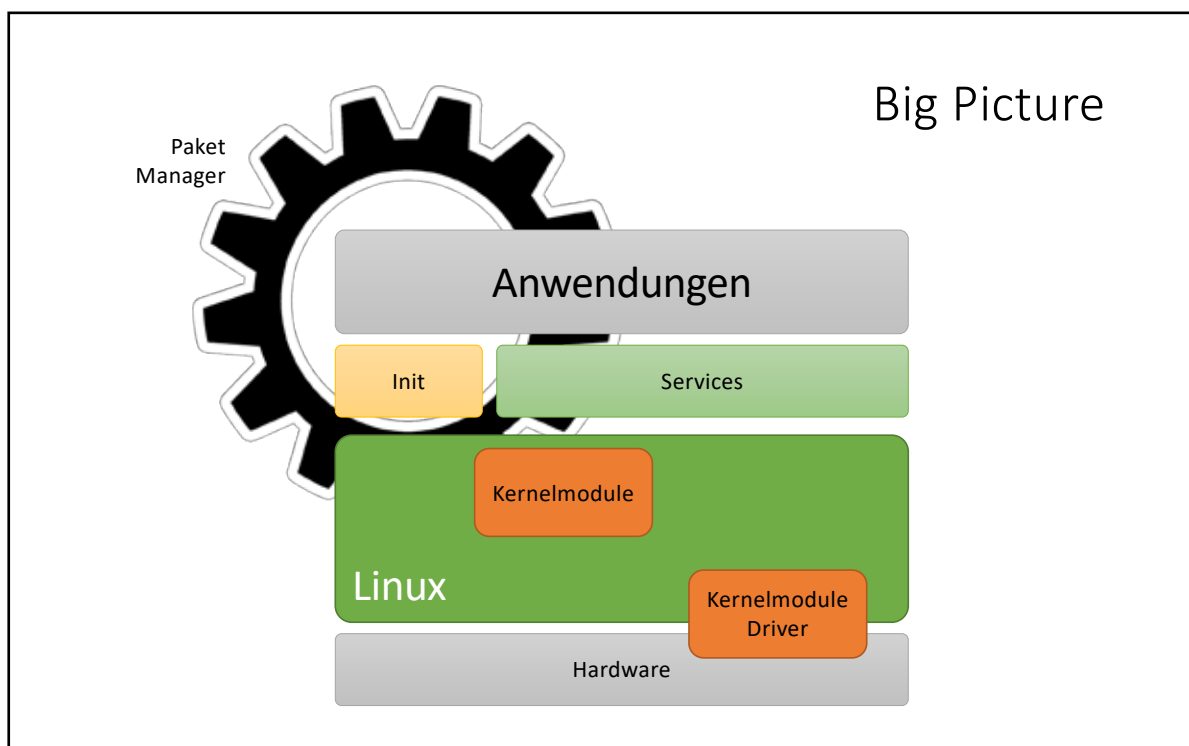
GitStats

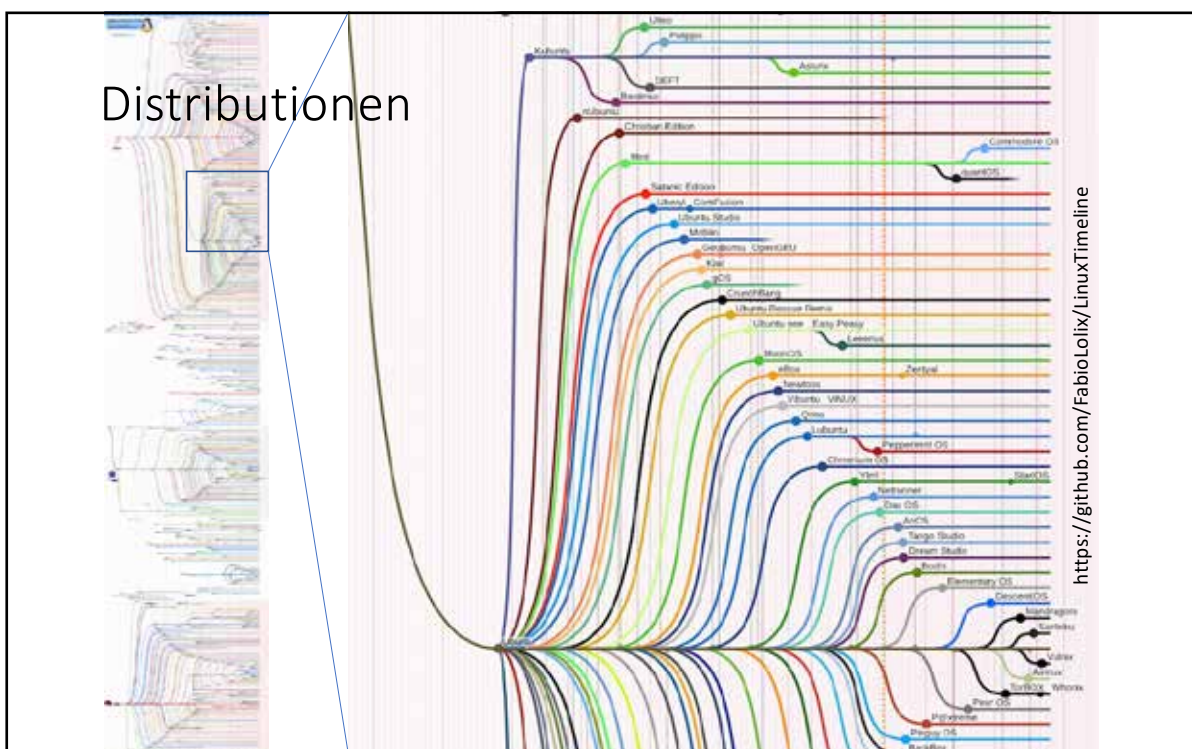
- 548273 Commits
- 14335 Autoren



Linux 4.3/arch

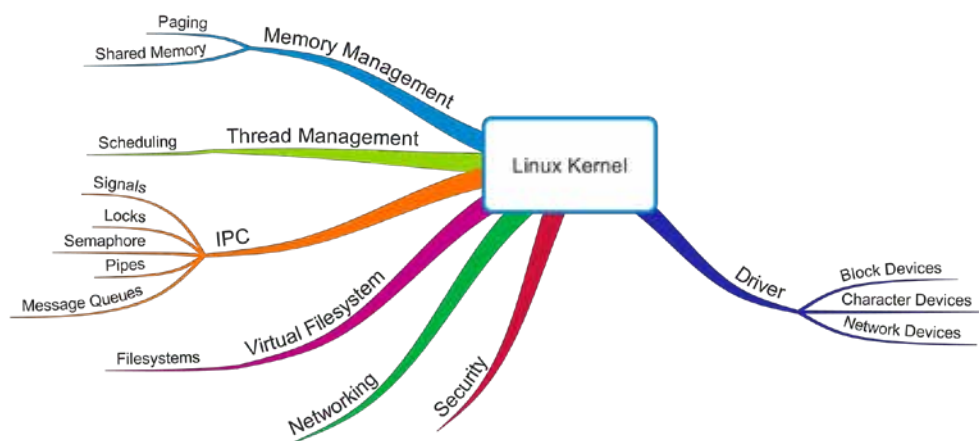
alpha arc arm arm64 avr32 blackfin c6x cris frv h8300
hexagon ia64 m32r m68k metag microblaze mips
mn10300 nios2 openrisc parisc powerpc s390 score sh
sparc tile um unicore32 x86 xtensa



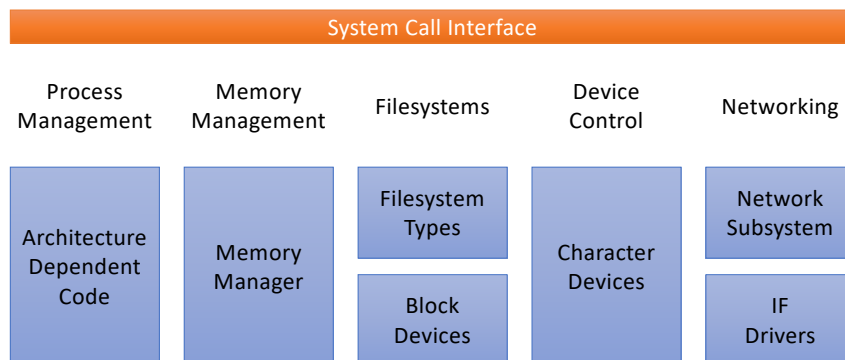


Distributionen

- distrowatch.com
- 262 aktive Distributionen
- TOP 5 (November 2022)
 - MX Linux
 - EndeavourOS
 - Mint
 - Manjaro
 - Ubuntu



Linux



37

History

- Archetype MINIX
 - Minimal-UNIX for teaching purposes
 - A. Tanenbaum, Amsterdam
- Linus Torvalds implements his own system
- First appearance: comp.os.minix, 29.3.1991

„Hello everybody,

I've had minix for a week now, and have upgraded to 386-minix (nice), and duly downloaded gcc for minix. Yes, it works – but ... optimizing isn't working, giving error message [...]. Is this normal?"



38

A minix clone was born

„Hello everybody out there using minix –

I’m doing a (free) operating system (just a hobby, won’t be big and professional like gnu) for 386(486) AT clones. This has been brewing since April, and is starting to get ready. I’d like any feedback on things people like/dislike in minx, as my OS resembles it somewhat (same physcal layout of the file-system (due to practical reasons) among other things).

... I’ll get something practical within a few months, and i’d like to know what features most people would want. Any suggestions are welcome, but i won’t promise I’ll implement them :-)”

39

Chronology

- Version 0.01, September 1991
- Version 0.11, December 1991
 - 0.11+VM over christmas
- Version 0.12, January 1992
 - Increasing number of interested people
 - Quarrel on OS architecture with A. Tanenbaum (Linux is obsolete)
- Version 0.95, March 1992
- We are approaching version 1
 - 0.95a, 0.95a.1, 0.95c, 0.95c+, pre0.96, 0.96a (Mai 1992), 0.96b, 0.96b.2, 0.98, 0.98.2, 0.99, 0.99.13, 0.99.13k,
- Version 1.0, 13. March 1994

40

Linux is obsolete (1)

- Tanenbaum

„I was in the U.S. for a couple of weeks, so i haven't commented much on LINUX (not that i would have said much had i been around), but for what it is worth, i have a couple of comments now.

As most of you know, for me MINIX is a hobby, something that i do in the evening when i get bored writing books and there are no major wars, revolutions, and senate hearings being televised live on CNN. My real job is a professor and researcher in the area of operating systems.

As a result of my occupation, I think I know a bit about where operating systems are going in the next decade or so. Two aspects stand out [kernel architecture and portability].

[...]“



41

Linux is obsolete (2)

- Tanenbaum (contd.)

„While I could go into a long story here about the relative merits of the two designs, suffice to say that among the people who actually design operating systems, the debate is essentially over. Microkernels have won.

LINUX is a monolithic style system. This is a giant step back into the 1970s. That is like taking an existing, working C program and rewriting it in BASIC.

I think it is a gross error to design an OS for any specific architecture, since that is not going to be around all that long. MINIX was designed to be reasonably portable, and has been ported from Intel line to the 680x0 (Atari, Amiga, Macintosh), SPARC, and NS32016. LINUX is tied fairly closely to the 80x86. Not the way to go.“

42

Linux is obsolete (3)

- Torvalds reply

„Well, with a subject like this, I'm afraid I'll have to reply. Apologies to minix-users who have heard enough about linux anyway. I'd like to be able to just 'ignore the bait,' but ... Time for some serious flamewar!

[...] Look at who makes money off minix, and who gives linux out for free. Then talk about hobbies. Make minix freely available, and one of my biggest gripes with it will disappear.

[...] (Bezugnehmend auf Tanenbaums Beruf) That's one hell of a good excuse for some of the brain damages of minix. I can only hope (and assume) that Amoeba doesn't suck like minix does.“

43

Linux is obsolete (4)

- Torvalds reply (contd)

„PS. I apologize for sometimes sounding too harsh: minix is nice enough if you have nothing else. Amoeba might be nice if you have 5-10 spare 386's lying around, but I certainly don't. I don't usually get into flames, but I'm touchy when it comes to linux :)“

44

Linux is obsolete (5)

- Torvalds email next day

„I wrote:

Well, with a subject like this, I'm afraid I'll have to reply.

And reply I did, with complete abandon, and no thought for good taste and netiquette. Apologies to [Andrew Tanenbaum], and thanks to John Nall for a friendly "that's not how it's done" letter. I overreacted, and am now composing a (much less acerbic) personal letter to [Andrew Tanenbaum]. Hope nobody was turned away from linux due to it being (a) possibly obsolete (I still think that's not the case, although some of the criticisms are valid) and (b) written by a hothead :-)

Linus "my first, and hopefully last flamefest" Torvalds"

45

Linux is obsolete (6)

- Wise words by somebody else

„Many if not most of the software we use is probably obsolete according to the latest design criteria. Most users could probably care less if the internals of the operating system they use is obsolete. They are rightly more interested in its performance and capabilities at the user level.

I would generally agree the microkernels are probably the wave of the future. However, it is in my opinion easier to implement a monolithic kernel. It is also easier for it to turn into a mess in a hurry as it is modified.

Regards,

Ken"

- Ken Thompson, „Inventor“ of UNIX

46

Linux is obsolete (7)

- The fight is over

- Tanenbaum

„I still maintain the point that designing a monolithic kernel in 1991 is a fundamental error. Be thankful your are not my student. You would not get a high grade for such a design :-)

[...] Writing a new OS only for the 386 in 1991 get you your second 'F' for this term. But if you do real well on the final exam, you can still pass the course.”

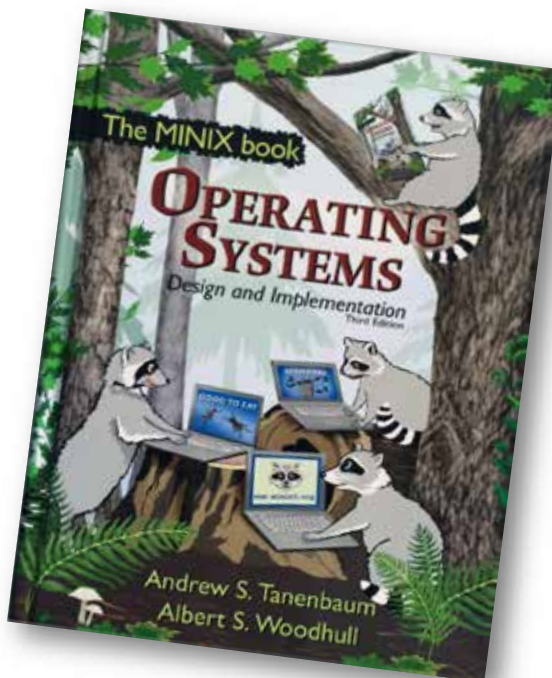
- Torvald

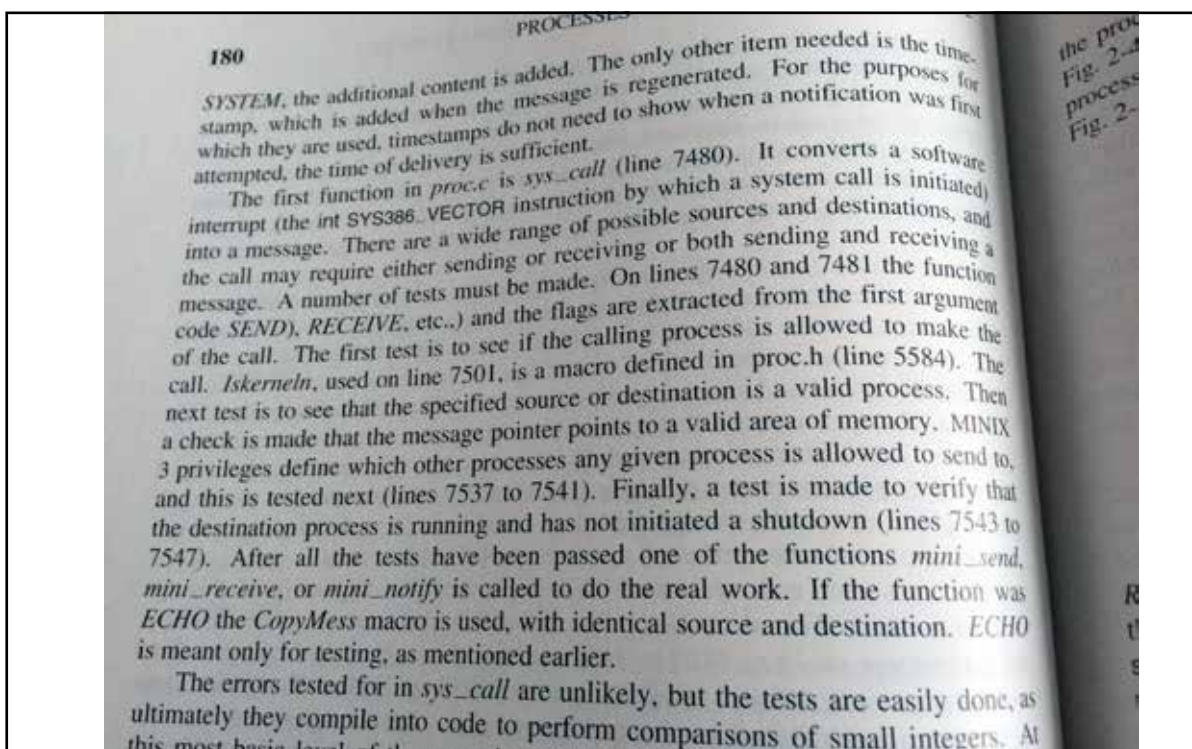
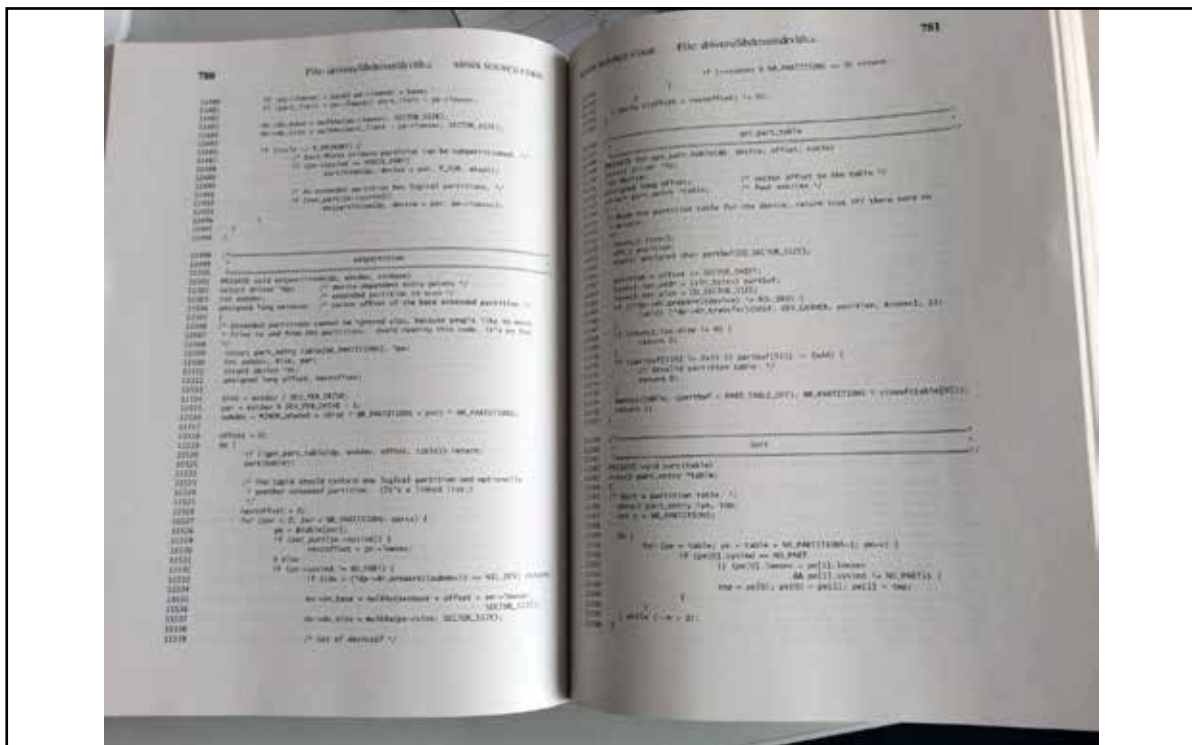
„Well, I probably won't get too good grades even without you: I had an argument (completely unrelated – not even pertaining to OS's) with the person here at the university that teaches OS design. I wonder when I'll learn :)

47

Minix

- 1987, A. Tanenbaum
- Für Lehrzwecke
 - Mikrokernell





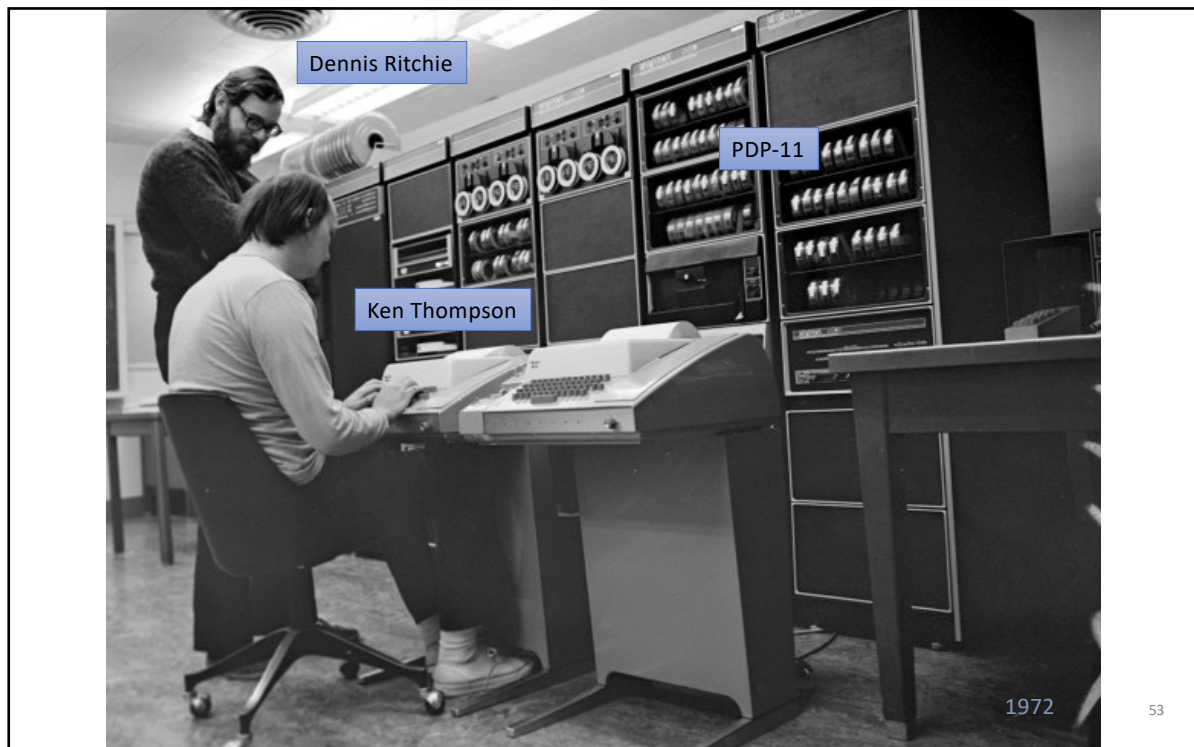
The UNIX roots

51

In the beginning ...

- 1969: First UNIX version (Bell Labs)
 - „Private“ research project by Ken Thompson „to use an otherwise idle PDP-7“
 - Dennis Richie: C programming language
 - Sources available to universities from the very beginning
- Predecessors
 - MULTICS: Ambitious goal at Bell Labs was archetype for many of the central abstractions such as file system, shell, etc.
 - CTSS (MIT)
 - GENIE (Berkeley): `fork()`

52



53

UNIX Roots

- UNIX was programmed with high-level programming language
 - Triumphant procession for programming language C
 - About 3% assembler only
- UNIX was distributed as sources
 - AT&T (Bell Labs) had no commercial interest (at first)
- UNIX featured properties only much larger and more expensive systems offered at this time
- UNIX provides many elementary functions that can be combined in very different ways (synergy)
 - fork and pipes

54



The UNIX Way

- Systemprogrammiersprache C
- Bausteinprinzip
- stdin / stdout
- Pipes
- Quellen für alle 😊

Linux 5.3.7 (21.10.2019)

- 108.6 MB .tar.xz
- 968 MB mit 65.216 Dateien
- Insgesamt
 - Assembler 376.252
 - C Source (*.c) 18.743.746
 - C Header (*.h) 5.585.729
- Davon in arch, drivers
 - Assembler 370.400+2.459
 - C (*.ch) 2.216.470+16.136.460
- Also reiner Kernel
 - 3393 Assembler
 - 5.976.545 C (*.ch)



Linux ? (?)

- ? MB .tar.xz
- ? GB mit ? Dateien
- Insgesamt
 - Assembler ?
 - C Source (*.c) ?
 - C Header (*.h) ?
- Davon in arch, drivers
 - Assembler ?
 - C (*.ch) ?
- Also reiner Kernel
 - ? Assembler
 - ? C (*.ch)



Linux 6.0.2 (15.10.2022)

- 133.8 MB .tar.xz
- 1.4 GB mit 77.955 Dateien
- Insgesamt
 - Assembler 369.769
 - C Source (*.c) 22.270.978
 - C Header (*.h) 8.877.910
- Davon in arch, drivers
 - Assembler 361.287 + 2.303
 - C (*.ch) 2.187.201 + 21.635.367
- Also reiner Kernel
 - 6179 Assembler
 - 7.326.320 C (*.ch)



POSIX

- Portable Operating System Interface
 - API-Standard für UNIX-artige Systeme
- IEEE, The Open Group, Austin Group
- **POSIX:2008** (formally: IEEE Std 1003.1-2008)
 - XBD (Base Definitions, 84 C Header files)
 - XCU (Shell and Utilities)
 - XSH (System Interfaces, 1123 system calls & library f's)
 - XRAT (Rational, Explanation)
- 3700 Seiten

Bleibende Werte



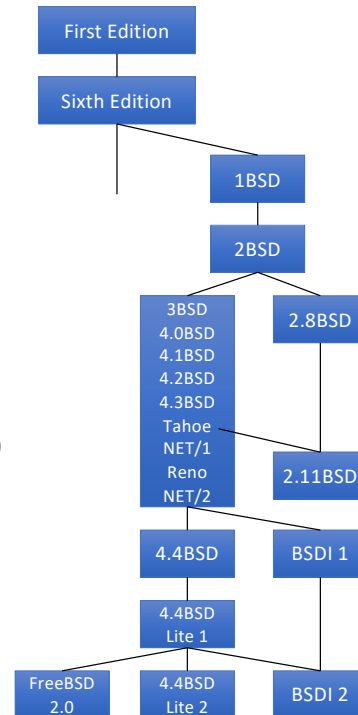
Important UNIX families

- First Edition up to Tenth Edition, Plan 9
 - The official Bell Labs branch
 - Plan 9 (4th edition available since 2003)
- BSD branch
 - Advancements in an academic environment
- System V
 - Advancements in a commercial environment
- SunOS / Solaris
- Linux

62

BSD family

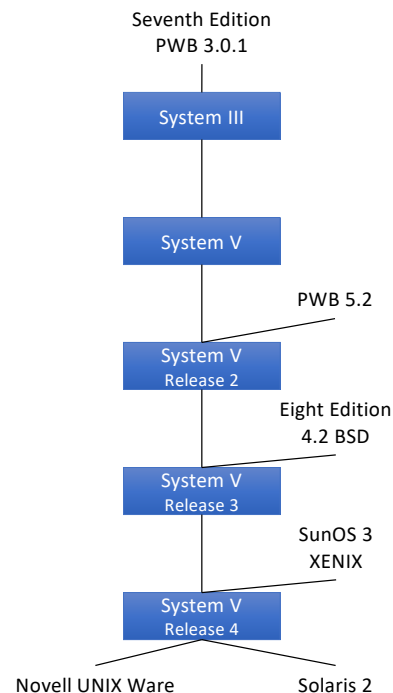
- Early branch in academia
 - University of Berkeley, San Diego
- Integration of TCP/IP protocol stack
 - ARPANET
 - Socket API
 - TCP variants (Reno, Tahoe, Vegas, ...)
- License difficulty
 - Several approaches to provide license/free version (Lite x)
- FreeBSD



63

System V

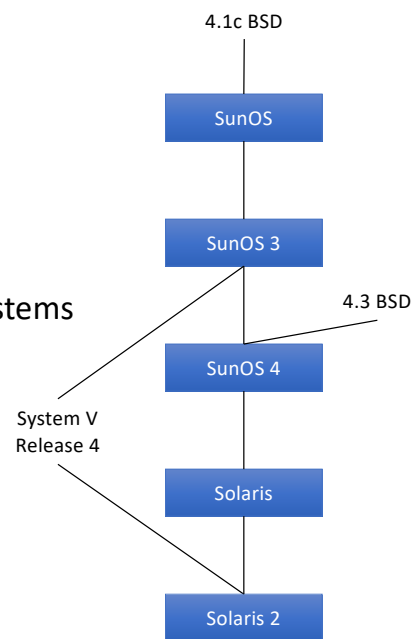
- Commercial version
 - Demerger of AT&T (1985) inhibits successful marketing
- Important improvement
 - Copy-On-Write
- „System V IPC“
 - Shared Memory
 - UNIX Streams
 - Semaphores, etc.



64

SunOS / Solaris

- UNIX branch of SUN Microsystems
- Consolidation of BSD and System V branches
- Also available for PC



65

MACH Microkernel



1985, CMU
u.a. B. Bershad und
R. Rashid (seit 1991 bei Microsoft,
Begründet MSR; NT Kernel)

WENN.com

History

- Famous microkernel approach starting 1984
 - Richard Rashid, CMU, previously working on the Accent network operating system kernel, now head of Microsoft Research
- Microkernel to cope with the ever increasing complexity of the UNIX operating system
 - Reduce the number of features in the kernel to make it less complex
- Mach kernel implements processor and memory management
 - File system, networking, I/O in a user-level Mach task

67

Mach Abstractions

- Task
 - Container for all the resource of one or more threads
 - Includes virtual memory, ports, processors, ...
- Thread
 - Basic unit of execution
- Port
 - In-kernel message queue with capabilities
- Message
 - Collection of data sent between threads in different tasks using ports
- Memory Object
 - Container of data mapped into a task's address space

68

Mach Monoliths or Microkernels

- Monoliths due to performance issues (Mach 2 and 2.5)
 - BSD, OSF/1, NEXTSTEP, OPENSTEP
- True Microkernel with version 3
 - Started at CMU, continued by OSF
 - Kernel preemption and RT scheduling
 - Low-level devices represented as ports
 - System-Call redirection (handling in user space)
 - Continuations
 - threads may block by specifying a function to be called upon completion

69

Mac OS X

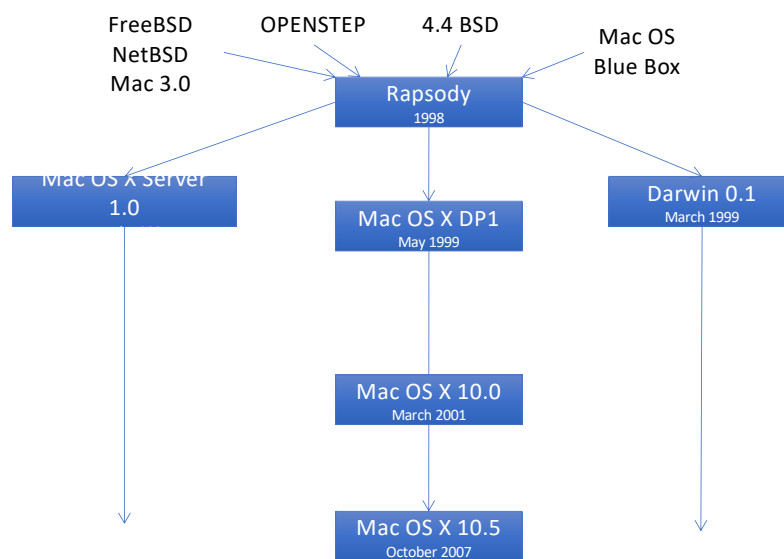
70

Roots

- Quite complex history (see Singh for details)
- OS Kernel
 - NuKernel first, replaced by Mach kernel later
- OS Personality
 - Various approaches including TalOS, Copland, Gershwin, BeOS,
 - NEXTSTEP and OPENSTEP (with Sun Microsystems) and Objective-C
- Final products
 - Mac OS 8 and 9
 - Rhapsody (1997)

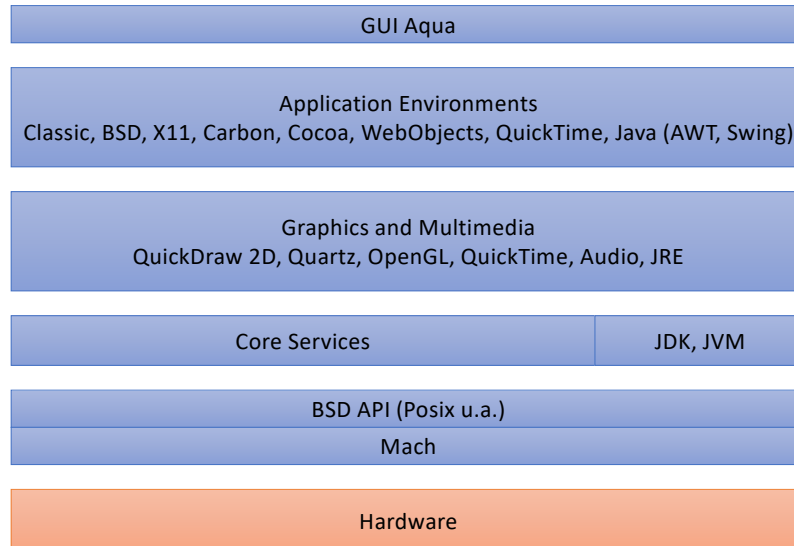
71

Family Tree



72

Architecture



73

Microsoft Windows

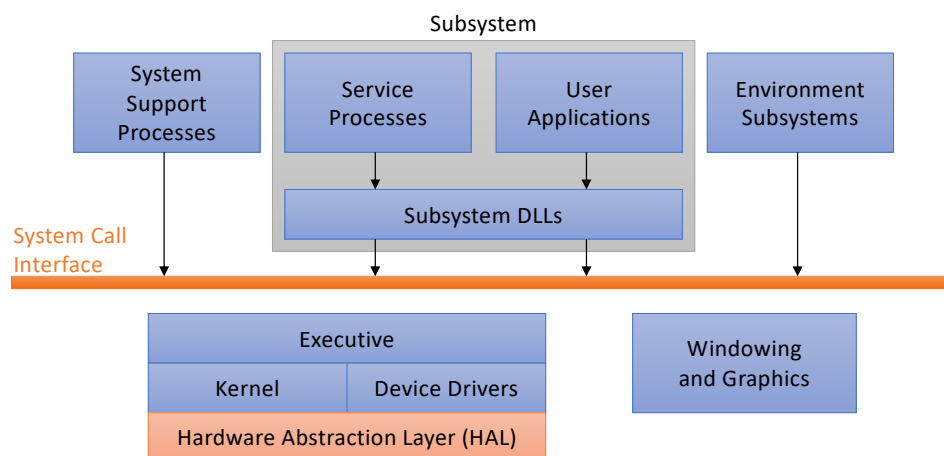
74

Overview

- Since Windows NT back in 1989 ...
 - 32 bit and 64 bit operating systems
 - Preemptive and reentrant scheduling
 - Full virtual memory support
 - Support for multiprocessor systems
- Mikrokernel with time-critical services “re-integrated”
- OS personality = subsystems
 - Windows subsystem common to all (provides Win32 API)
 - POSIX subsystem
 - There was a OS/2 subsystem once

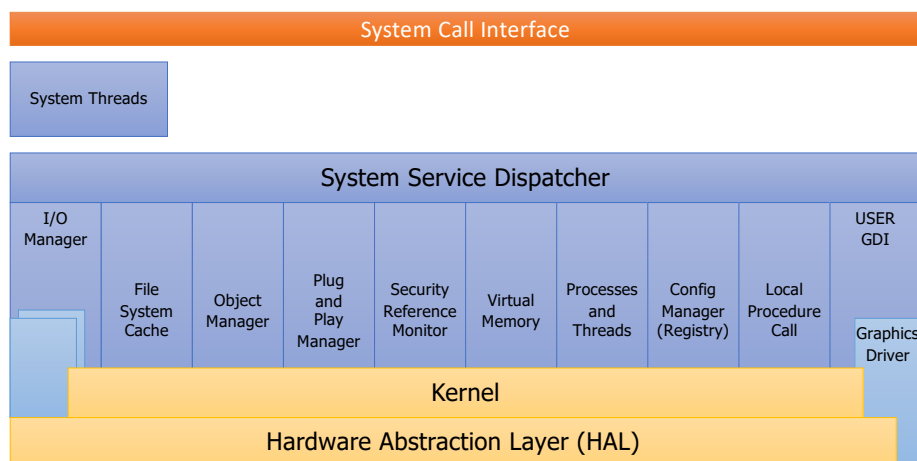
75

Overall Architecture



76

Windows Executive



77

Literature

- D. Bovet, M. Cesati, Understanding the Linux Kernel, 3rd Edition, O'Reilly, 2006
- Eric Levenez, Various chronologies on operating systems, <http://www.levenez.com>
- J. Mauro, R. McDougall, Solaris Internals, Sun Microsystems Press, 2001
- M.K. McKusik, K. Bostic, M.J. Karels, J.S. Quarterman, The Design and Implementation of the FreeBSD Operating System, Addison-Wesley, 2014
- G. Moody, rebel code – Linux and the Open Source Revolution, The Penguin Press, 2001
- M.E. Russinovich, D.A. Solomon, Microsoft Windows Internals, 4th Edition, Microsoft Press, 2005
- Amit Singh, Mac OS X Internals – A Systems Approach, Addison-Wesley, 2007
- A.S. Tanenbaum, A.S. Woodhull, The Minix Book – Operating Systems, 3rd Edition, Prentice-Hall, 2006
- L. Torvalds, just for fun, Harper Business, 2001

78