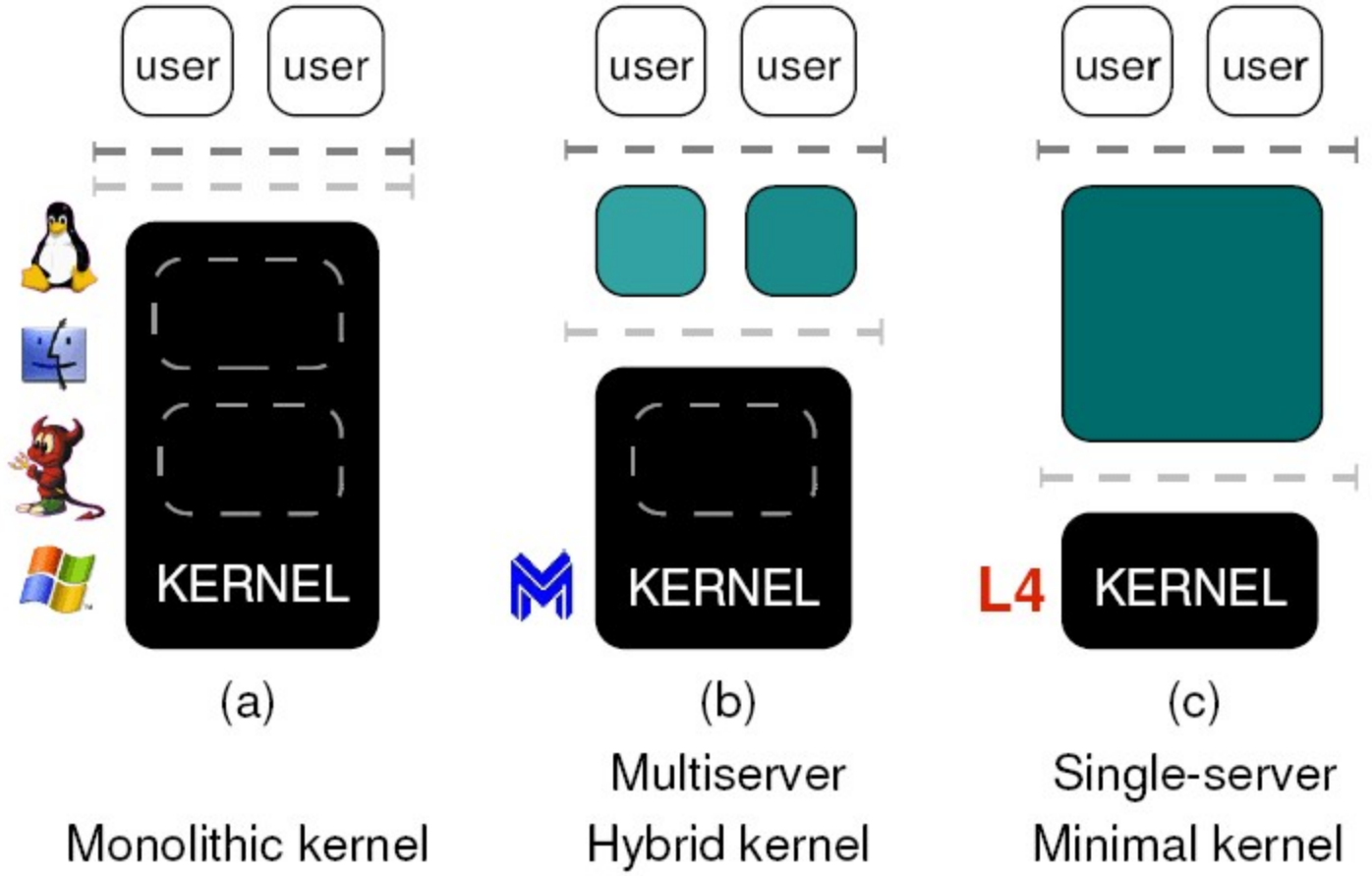
Operating System Reliability (MINIX 3)

IPA Herfstdagen on Security
November 24, 2005

Jorrit Herder
<jnherder@cs.vu.nl>
Vrije Universiteit Amsterdam

Typical OS Structures



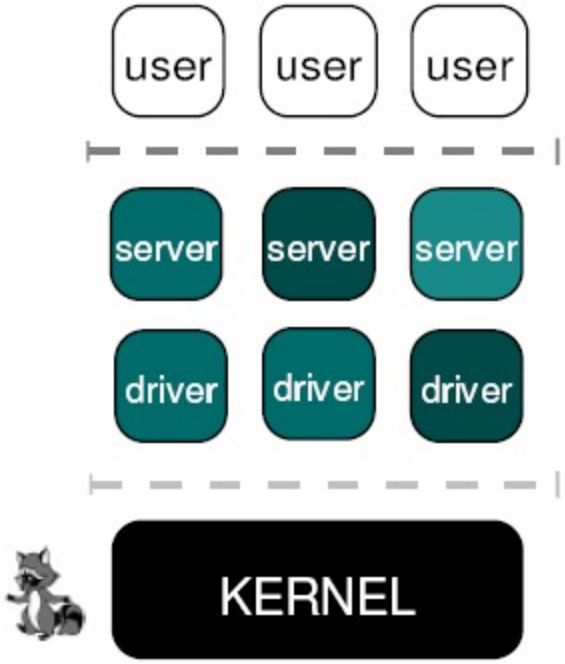
Monolithic Systems

- **Fundamental design flaws**
  - All code at highest privilege level
  - No proper fault isolation
  - Huge amount of code in the kernel
  - Untrusted, 3rd-party code in the kernel
  - Hard to maintain and configure

What's Missing?

- **Principle of Least Authority**
  - Compartmentalize the systems
  - Isolate faults within the components
  - Allow components to communicate
  - Assign only the required privileges
- **Fault Tolerance**
  - Expect components to crash
  - Withstand failures and recover

Structure of MINIX 3



(d)
Multiserver
Minimal kernel

- **Design Principles**
  - Simplicity
  - Least Authority
  - Fault Tolerance
- **Facts**
  - Kernel < 4000 LOC
  - User-space OS servers
  - Disk I/O ~9% overhead
  - Full-speed ethernet

MINIX 3: Structural Measures

- **Minimal kernel (< 3800 LOC) reduces # fatal bugs**
- **OS is collection of isolated, user-mode processes**
- **Reliable IPC: small, fixed-size message passing**
- **Deadlock avoidance and deadlock detection**
- **Buffer overruns in OS prevented and damage limited**
- **Bad pointers in OS are caught with MMU hardware**
- **Scheduler detects and tames infinite loops in OS**
- **Monitor and restart malfunctioning OS services**

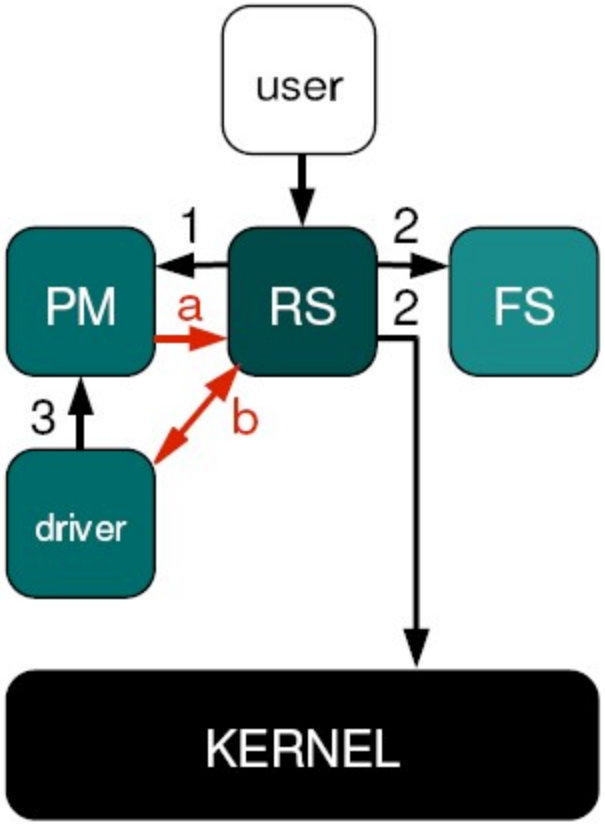
- *Compare this to a monolithic system …*

MINIX 3: Per-Process Policies

- **IPC only possible if type and target are allowed**
- **Fine-grained control over request types allowed**
- **Access to individual I/O ports can be restricted**
- **Access to remote memory, e.g., video RAM**
- **Scheduling priority and quantum size**

- *More to come …*

MINIX 3: Reincarnation Server



Start in controlled way

1. Encapsulate in new process
2. Assign minimal privileges
3. Execute binary

Monitor system services

a) Immediate crash detection

b) Periodic status checks

Fix problems

- Kill and restart fresh copy

Summary & Conclusion

- **Different OS structures and properties**
  - Fundamental problems with monolithic systems
  - Inherent benefits of modular systems
- **OS reliability *is* possible: MINIX 3**
  - Multiserver OS with minimal kernel (< 4000 LOC)
  - Improvements over other operating systems
    - We reduce the number of fatal bugs
    - We limit the damage bugs can do
    - We can recover from common failures
- **Download and more information: www.minix3.org**