

# Replugging the Modern Desktop

Kay Sievers <[kay.sievers@suse.de](mailto:kay.sievers@suse.de)>  
David Zeuthen <[davidz@redhat.com](mailto:davidz@redhat.com)>

Linux Plumbers Conference  
Portland, OR, Sept 2009

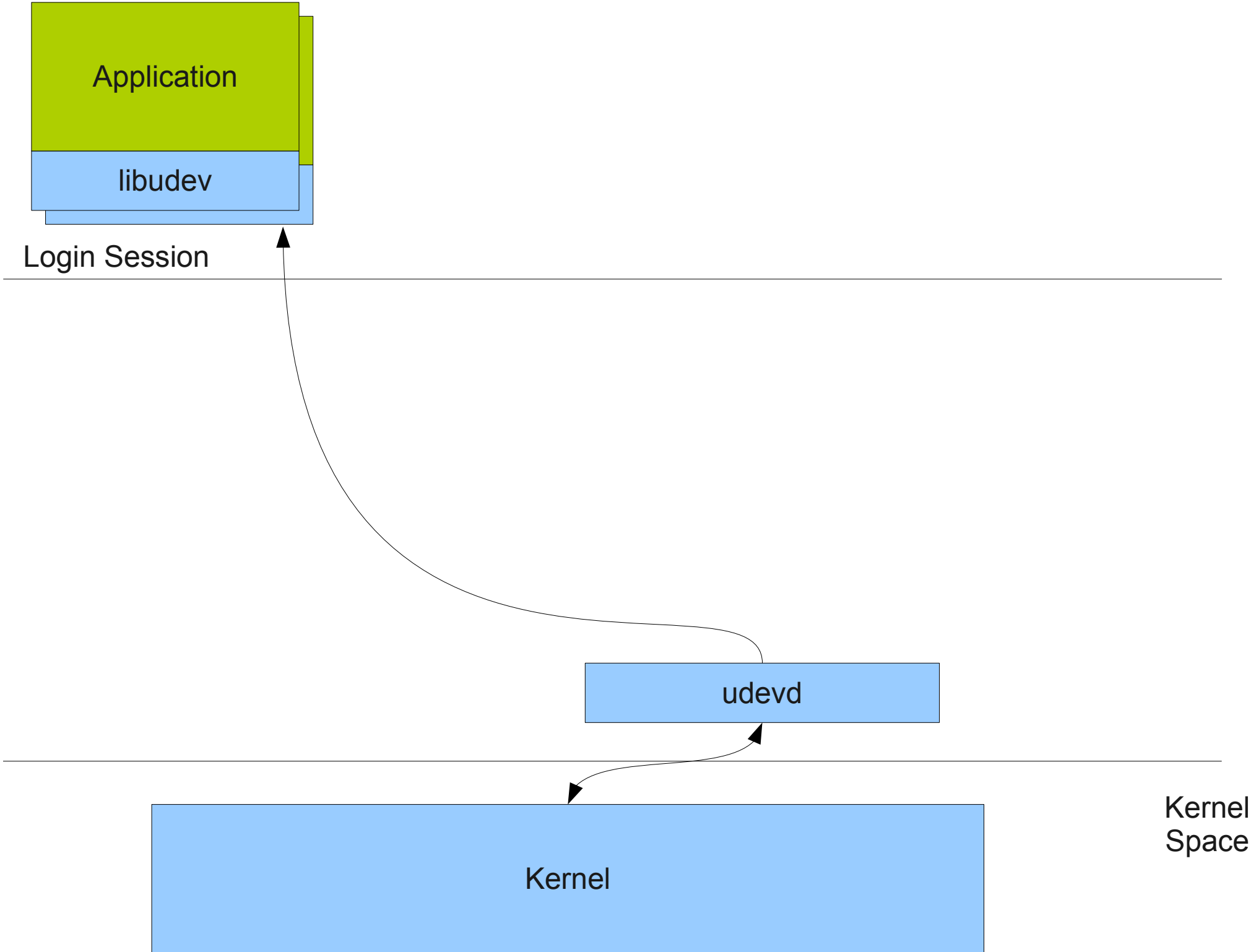
# History

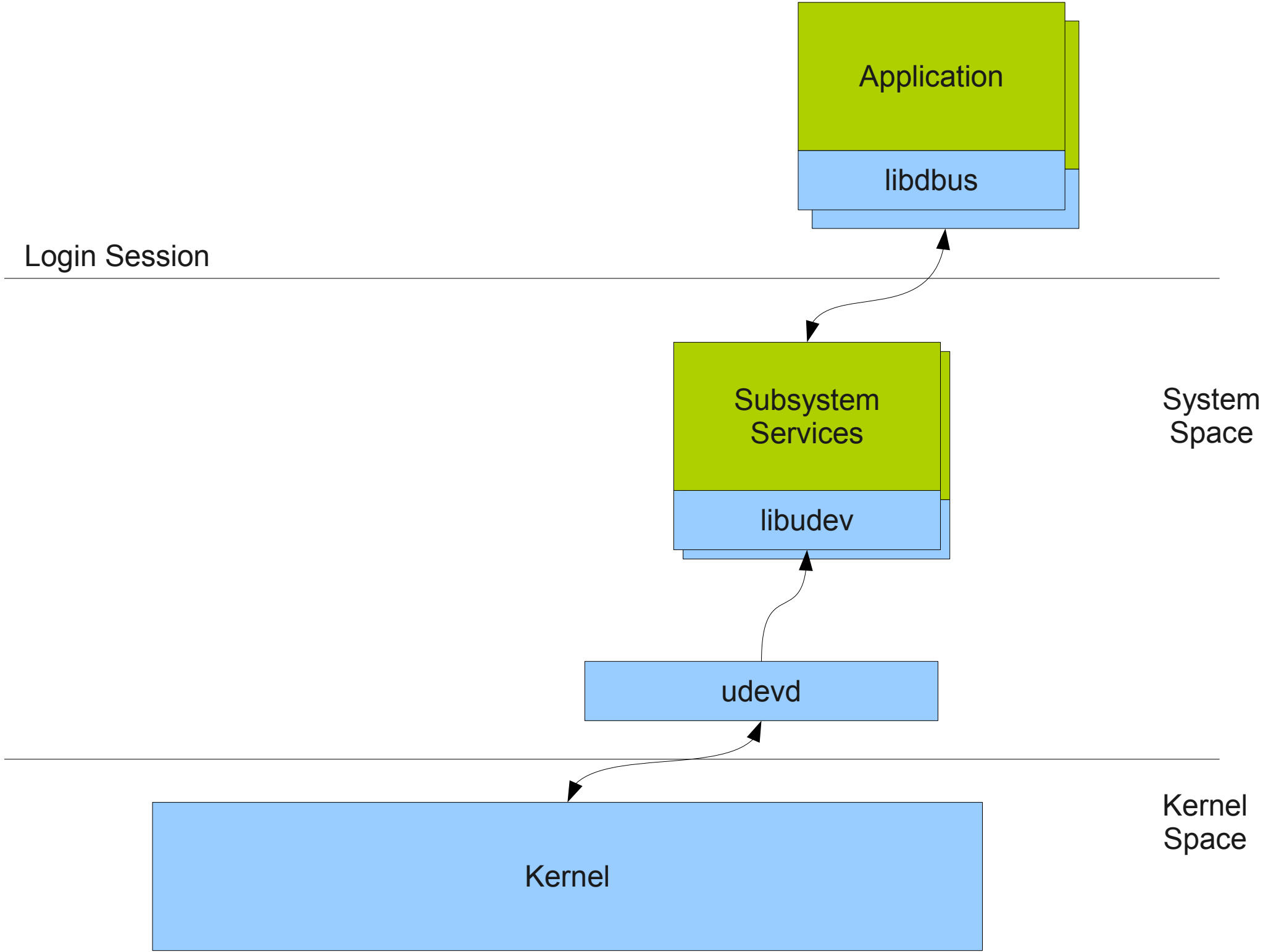
- Back in the day
  - /sbin/hotplug, scan entire /dev, /proc/scsi/scsi, /proc/partitions
  - magicdev, supermount, subfs
  - User conf / passwords stored in /etc or hard-coded
  - Millions of LOC running as uid 0

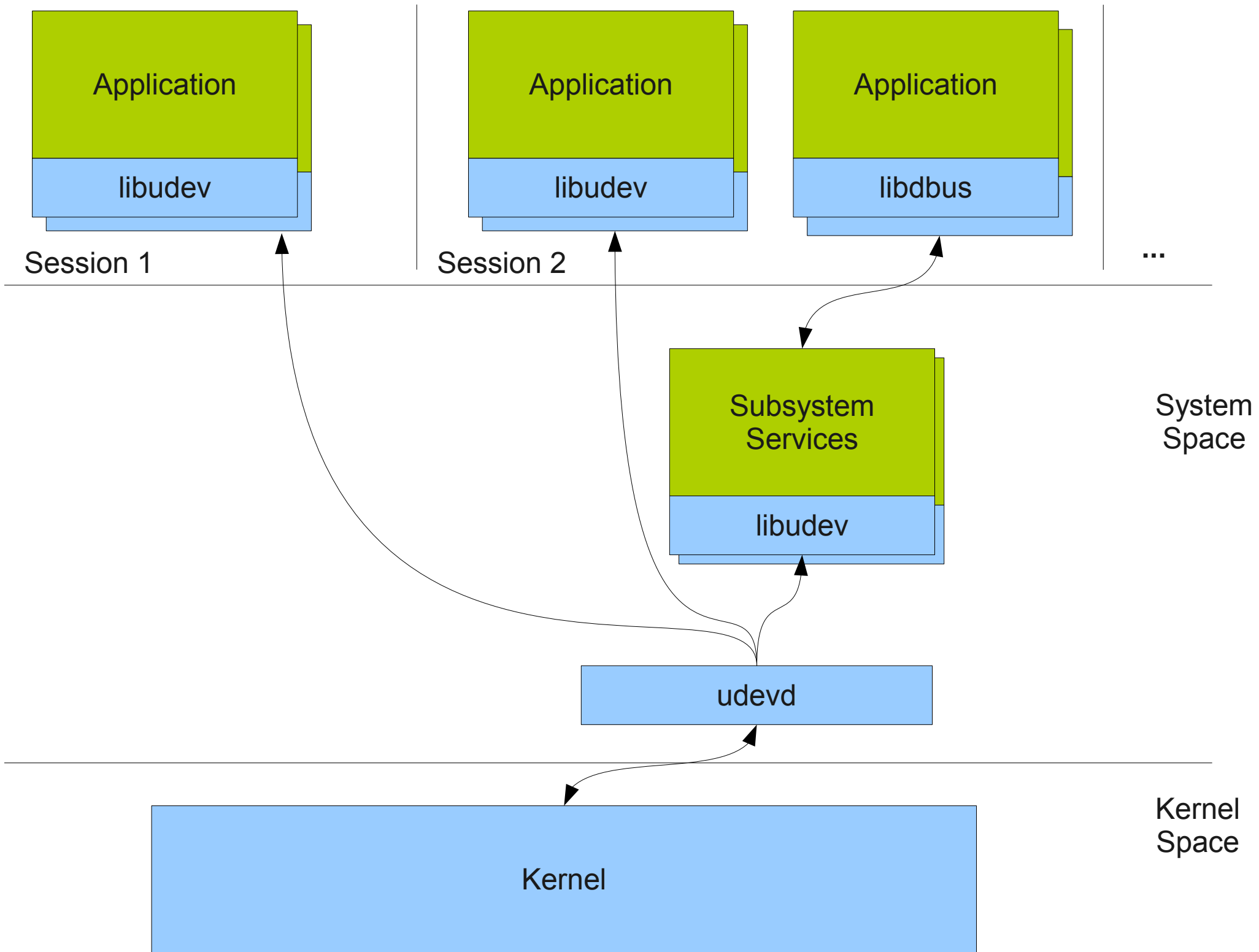
# History

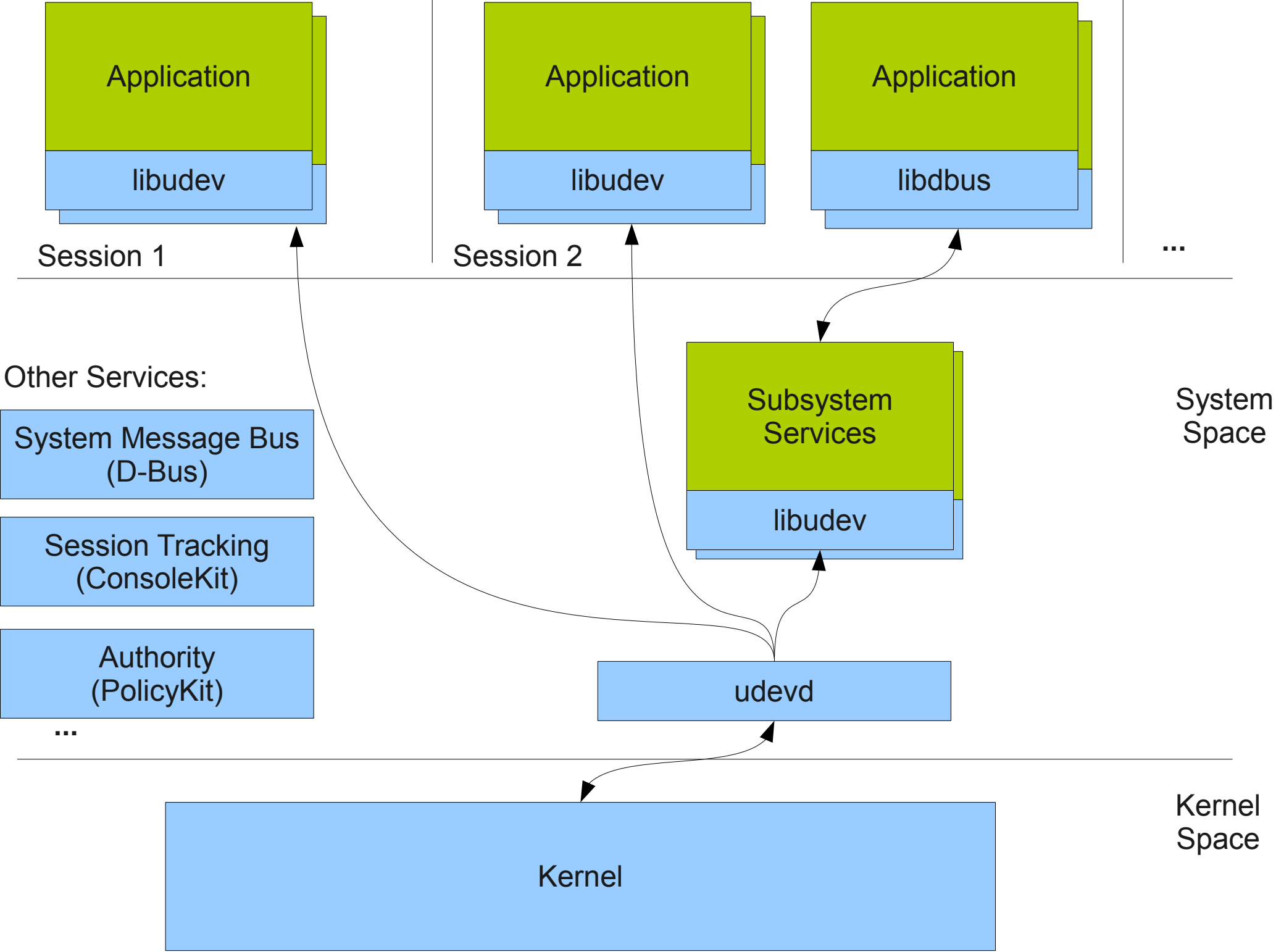
- Back in the day
  - /sbin/hotplug, scan entire /dev, /proc/scsi/scsi, /proc/partitions
  - magicdev, supermount, subfs
  - User conf / passwords stored in /etc or hard-coded
  - Millions of LOC running as uid 0
- Early Desktop Integration
  - HAL, D-Bus, PolicyKit
  - Separate Mechanism and Policy
  - But... Implementation too complex, not scalable, not focused, too many abstractions

- Cutting the same cake in a different way
- 1<sup>st</sup> piece: Move device discovery/enumeration, classification, quirks, probing, event propagation to udev
- 2<sup>nd</sup> piece: Write libudev
- 3<sup>rd</sup> piece: Dedicated system services for major subsystems
  - DeviceKit-disks, DeviceKit-power, NetworkManager, PulseAudio, Bluez, Gypsy, ...
- 4<sup>th</sup> piece: Port the world to subsystem services
  - Apps using simple subsystems use libudev (Cheese)









## kernel devices show up in a device tree in /sys

```
/sys/devices
|-- pci0000:00
...
|   |-- 0000:00:1f.2
|   |   |-- driver -> ../../../../bus/pci/drivers/ahci
...
|   |   |-- host0
|   |       |-- subsystem -> ../../../../../../bus/scsi
|   |       |-- target0:0:0
|   |           |-- subsystem -> ../../../../../../bus/scsi
...
|   |   |-- 0:0:0:0
|   |       |-- subsystem -> ../../../../../../bus/scsi
...
|   |   |-- block
|   |       |-- sda
|   |           |-- subsystem -> ../../../../../../../../../../class/block
|   |           |-- alignment_offset
|   |           |-- capability
|   |           |-- dev
```

## **new devices and changes are announced over netlink with uevents:**

```
recvmsg(3,  
    {msg_name(12)={sa_family=AF_NETLINK, pid=0, groups=00000001},  
      msg_iov(1)=[{"add@/devices/pci0000:00/0000:00:1f.2/host0/target0:0:0/...\\0  
ACTION=add\\0  
DEVPATH=/devices/pci0000:00/0000:00:1f.2/host0/target0:0:0/0:0:0:0/...\\0  
SUBSYSTEM=block\\0  
MAJOR=8\\0  
MINOR=0\\0  
DEVNAME=sda\\0  
DEVTYPE=disk\\0  
SEQNUM=1584\\0"}, ...],  
    ...
```

## udev rules to:

- add properties to store in database
- create meaningful symlinks
- run programs to configure/setup the device

```
SUBSYSTEM=="block", KERNEL=="sd*", ENV{DEVTYPE}=="disk", \
  IMPORT{program}="ata_id --export $tempnode"
```

```
/lib/udev/ata_id --export /dev/sda
ID_TYPE=disk
ID_BUS=ata
ID_MODEL=SAMSUNG_MMCQE28G8MUP-0VA
ID_MODEL_ENC=SAMSUNG\x20MMCQE28G8MUP-0VA
ID_REVISION=VAM08L1Q
ID_SERIAL=SAMSUNG_MMCQE28G8MUP-0VA_SE837A4759
ID_SERIAL_SHORT=SE837A4759
```

```
SUBSYSTEM=="block", KERNEL=="sd*", ENV{DEVTYPE}=="disk", \
  ENV{ID_SERIAL}=="?*", SYMLINK+="disk/by-id/$env{ID_BUS}-$env{ID_SERIAL}"
```



## receive event with libudev:

```
struct udev_monitor *monitor;  
struct udev_device *device;  
  
monitor = udev_monitor_new_from_netlink(udev, "udev");  
udev_monitor_enable_receiving(monitor);  
udev_monitor_filter_add_match_subsystem_devtype(monitor, "block", "disk");  
device = udev_monitor_receive_device(monitor);
```

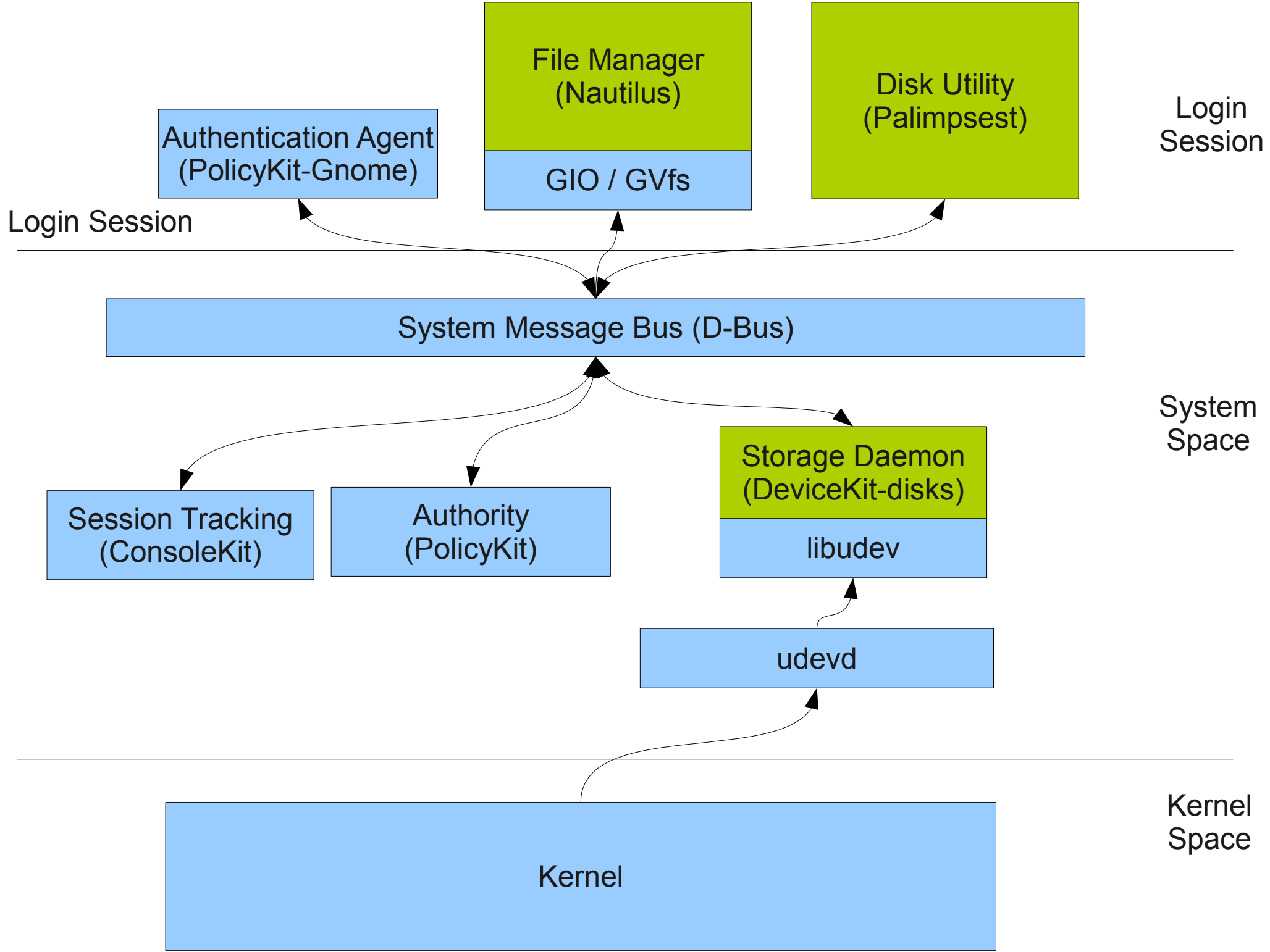
message multiplexing in the kernel

messages filtered inside the kernel with berkeley packet filter

libgudev javascript example

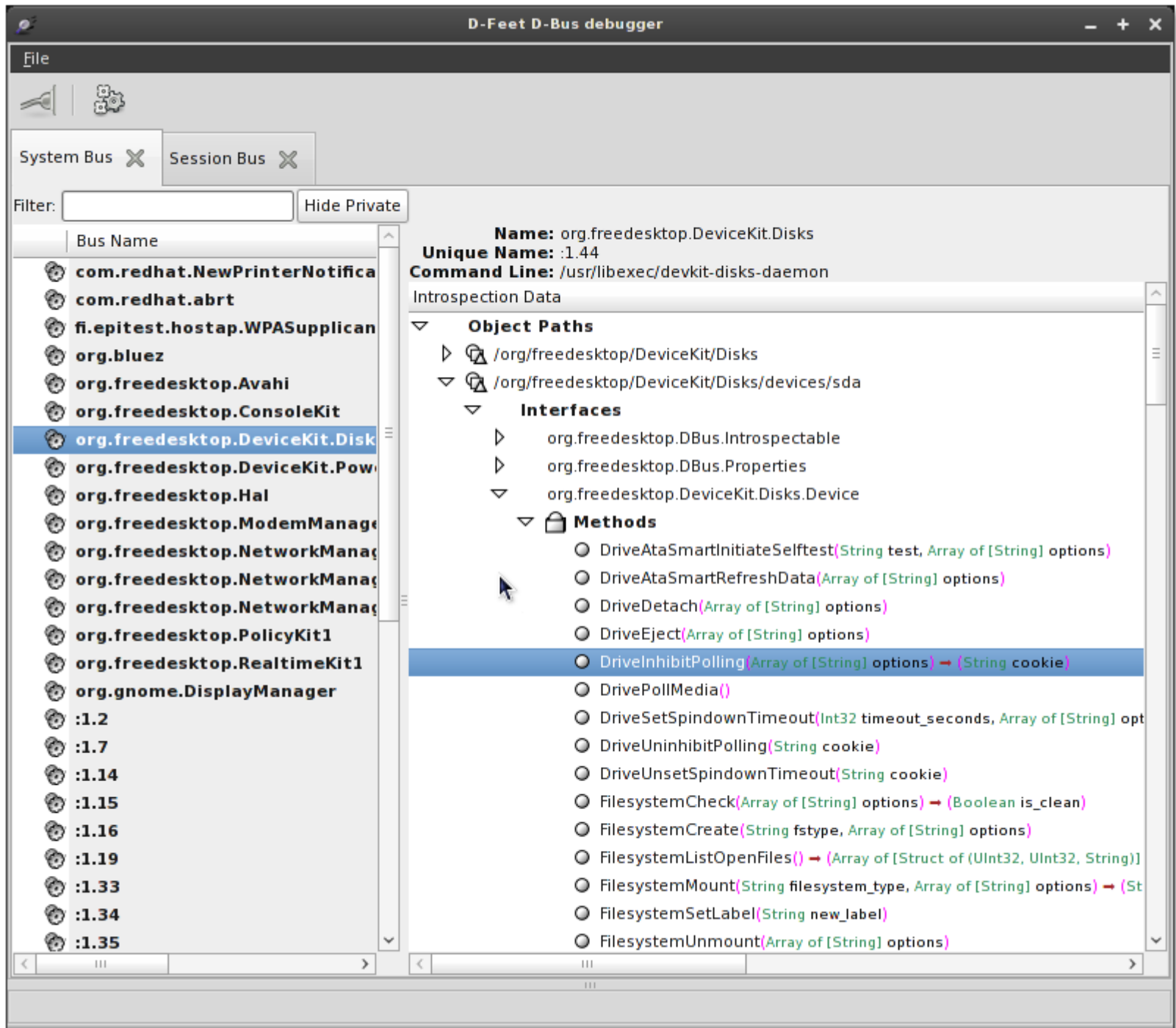
# Storage Subsystem Daemon

- DeviceKit-disks
- Consumes udev information
- Started on demand
- High-level API w/ progress reporting
  - Mount, Unmount, Eject, Poll, Fsck
  - Partitioning, Formatting, FS Label
  - ATA SMART monitoring
  - MD-RAID (Create, Start/Stop, Check, ...)
  - Drive spindown



# Palimpsest Demo

(d-feet, fs labels, mkfs, ATA SMART from USB, new-ui)



```
davidz@x61:~/Hacking
Every 2.0s: tree /dev/disk/
Wed Sep 23 16:22:13 2009
/dev/disk/
|-- by-id
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN -> ../../sda
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part1 -> ../../sda1
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part2 -> ../../sda2
|   |-- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN -> ../../sda
|   |-- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN-part1 -> ../../sda1
|   |-- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN-part2 -> ../../sda2
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0 -> ../../sdb
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0-part1 -> ../../sdb1
|-- by-label
|   |-- FIRSTNAME -> ../../sdb1
|   |-- FedoraRawhide -> ../../sda1
|-- by-path
```

```
root@x61:~
[root@x61 ~]# udevadm monitor --udev
monitor will print the received events for:
UDEV - the event which udev sends out after rule processing
```

Palimpsest Disk Utility

File Edit Help

80 GB Solid-State Disk  
ATA INTEL SSDSA2MH080G1GC  
MBR Partition Table

- Fedora (Rawhide)  
32 GB Linux Ext3 (version 1.0)
- 2.0 GB Unrecognized  
Unknown or Unused
- 46 GB Free  
Unallocated Space
- Generic STORAGE DEVICE  
Generic STORAGE DEVICE  
4.1 GB Media, MBR Partition Table
- FIRST NAME**  
4.1 GB FAT (32-bit version)

**FIRST NAME**  
4.1 GB / 3.8 GiB / 4,071,000  
FAT (32-bit version) File System  
Partition 1 (W95 FAT32)

**Partition**  
The attributes of the partition can be edited... can also be deleted to make room for other...

Partition Label:

Type:

Bootable

**Mountable Filesystem**  
The volume contains a mountable filesystem

Label:

```
davidz@x61:~/Hacking
[davidz@x61 Hacking]$ devkit-disks --monitor
Monitoring activity from the disks daemon. Press Ctrl+C to cancel.
```

```
davidz@x61:~/Hacking
Every 2.0s: tree /dev/disk/ Wed Sep 23 16:22:39 2009
/dev/disk/
-- by-id
|  -- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN -> ../../sda
|  -- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part1 -> ../../sda1
|  -- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part2 -> ../../sda2
|  -- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN -> ../../sda
|  -- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN-part1 -> ../../sda1
|  -- scsi-SATA_INTEL_SSDSA2MH0C842101HD080DGN-part2 -> ../../sda2
|  -- usb-Generic_STORAGE_DEVICE_000000009407-0-0 -> ../../sdb
|  -- usb-Generic_STORAGE_DEVICE_000000009407-0-0-part1 -> ../../sdb1
-- by-label
|  -- Fedora\x20\x28Rawhide\x29 -> ../../sda1
|  -- NEW\x20NAME -> ../../sdb1
-- by-path
|  -- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0 -> ../../sdb
|  -- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0-part1 -> ../../sdb1
|  -- pci-0000:00:1f.2-scsi-0:0:0:0 -> ../../sda
|  -- pci-0000:00:1f.2-scsi-0:0:0:0-part1 -> ../../sda1
|  -- pci-0000:00:1f.2-scsi-0:0:0:0-part2 -> ../../sda2
-- by-uuid
|  -- 7456-7B63 -> ../../sdb1
|  -- 786263c4-5e28-4cdc-97b8-1ab6e221c344 -> ../../sda1

4 directories, 17 files
```

```
root@x61:~
[root@x61 ~]# udevadm monitor --udev
monitor will print the received events for:
UDEV - the event which udev sends out after rule processing

UDEV [1253748149.094123] change /devices/pci0000:00/0000:00:1d.7/usb2/2-1/2-1:1.0/host6/target6:0:0/6:0:0:0/block/sdb/sdb1 (block)
UDEV [1253748149.189025] change /devices/pci0000:00/0000:00:1d.7/usb2/2-1/2-1:1.0/host6/target6:0:0/6:0:0:0/block/sdb/sdb1 (block)
```

Unknown or Unused

46 GB Free  
Unallocated Space

**Generic STORAGE DEVICE**  
Generic STORAGE DEVICE  
4.1 GB Media, MBR Partition Table

**NEW NAME**  
4.1 GB FAT (32-bit version)

**Partition**  
The attributes of the partition can be edited... can also be deleted to make room for other...

Partition Label:

Type:

Bootable

**Mountable Filesystem**  
The volume contains a mountable filesystem

Label:

```
davidz@x61:~/Hacking
[davidz@x61 Hacking]$ devkit-disks --monitor
Monitoring activity from the disks daemon. Press Ctrl+C to cancel.
job-changed: /org/freedesktop/DeviceKit/Disks/devices/sdb1
changed: /org/freedesktop/DeviceKit/Disks/devices/sdb1
job-changed: /org/freedesktop/DeviceKit/Disks/devices/sdb1
changed: /org/freedesktop/DeviceKit/Disks/devices/sdb1
changed: /org/freedesktop/DeviceKit/Disks/devices/sdb1
```

```
davidz@x61:~/Hacking
Every 2.0s: tree /dev/disk/ Wed Sep 23 16:26:21 2009
/dev/disk/
|-- by-id
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN -> ../../sda
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part1 -> ../../sda1
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part2 -> ../../sda2
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN -> ../../sda
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN-part1 -> ../../sda1
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN-part2 -> ../../sda2
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0 -> ../../sdb
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0-part1 -> ../../sdb1
|-- by-label
|   |-- Fedora\x20\x28Rawhide\x29 -> ../../sda1
|   |-- NEW\x20NAME\x202 -> ../../sdb1
|-- by-path
|   |-- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0 -> ../../sdb
|   |-- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0-part1 -> ../../sdb1
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0 -> ../../sda
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0-part1 -> ../../sda1
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0-part2 -> ../../sda2
|-- by-uuid
|   |-- 786263c4-5e28-4cdc-97b8-1ab6e221c344 -> ../../sda1
|   |-- a1560eaf-215f-4352-97b8-6f6e3d451b73 -> ../../sdb1
4 directories, 17 files
```

```
root@x61:~
[root@x61 ~]# udevadm monitor --udev
monitor will print the received events for:
UDEV - the event which udev sends out after rule processing

UDEV [1253748368.965545] change /devices/pci0000:00/0000:00:1d.7/usb2/2-1/2-1:1.0/host6/target6:0:0/6:0:0/block/sdb/sdb1 (block)
```

Unknown or Unused  
46 GB Free  
Unallocated Space

**Generic STORAGE DEVICE**  
Generic STORAGE DEVICE  
4.1 GB Media, MBR Partition Table

- NEW NAME 2  
4.1 GB Linux Ext2 (version 1.0)

**Partition**  
The attributes of the partition can be edited. The partition can also be deleted to make room for other partitions.

Partition Label:

Type: W95 FAT32 (0x0b)

Bootable

**Mountable Filesystem**  
The volume contains a mountable filesystem.

Label:

```
root@x61:~
[root@x61 ~]# e2label /dev/sdb1 "NEW NAME 2"
[root@x61 ~]#
```

```
davidz@x61:~/Hacking
Every 2.0s: tree /dev/disk/ Wed Sep 23 16:29:38 2009
/dev/disk/
|-- by-id
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN -> ../../sda
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part1 -> ../../sda1
|   |-- ata-INTEL_SSDSA2MH080G1GC_CVEM842101HD080DGN-part2 -> ../../sda2
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN -> ../../sda
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN-part1 -> ../../sda1
|   |-- scsi-SATA_INTEL_SSDSA2MH0C8VEM842101HD080DGN-part2 -> ../../sda2
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0 -> ../../sdb
|   |-- usb-Generic_STORAGE_DEVICE_000000009407-0-0-part1 -> ../../sdb1
|-- by-label
|   |-- Fedora\x20Rawhide\x29 -> ../../sda1
|   |-- My\x20Super\x20Pe -> ../../sdb1
|-- by-path
|   |-- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0 -> ../../sdb
|   |-- pci-0000:00:1d.7-usb-0:1:1.0-scsi-0:0:0:0-part1 -> ../../sdb1
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0 -> ../../sda
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0-part1 -> ../../sda1
|   |-- pci-0000:00:1f.2-scsi-0:0:0:0-part2 -> ../../sda2
|-- by-uuid
|   |-- 786263c4-5e28-4cdc-97b8-1ab6e221c344 -> ../../sda1
|   |-- F588-BE80 -> ../../sdb1
4 directories, 17 files
```

```
root@x61:~# udevadm monitor --udev
monitor will print the received events for:
UDEV - the event which udev sends out after rule processing

UDEV [1253748368.965545] change /devices/pci0000:00/0000:00:1d.7/usb2/2-1/2-1:1.0/host6/target6:0:0/6:0:0:0/block/sdb/sdb1 (block)

UDEV [1253748570.550025] change /devices/pci0000:00/0000:00:1d.7/usb2/2-1/2-1:1.0/host6/target6:0:0/6:0:0:0/block/sdb/sdb1 (block)
```

Unknown or Unused  
46 GB Free  
Unallocated Space

**Generic STORAGE DEVICE**  
Generic STORAGE DEVICE  
4.1 GB Media, MBR Partition Table

**My Super Pe**  
4.1 GB FAT (32-bit version)

**Partition**  
The attributes of the partition can be edited. The partition can also be deleted to make room for other partitions.

Partition Label:

Type:


Bootable

**Mountable Filesystem**  
The volume contains a mountable filesystem.

Label:

```
root@x61:~# mkfs.vfat -n "My Super Penguin" /dev/sdb1
mkfs.vfat 3.0.1 (23 Nov 2008)
[root@x61 ~]#
```

Drive:

 **320 GB Hard Disk - ST932042 1ASG**  
Disk is healthy
**Status**

Updated: Less than a minute ago - [Update Now](#)

Self-tests: Last self-test completed OK - [Run self-test](#)

Model: ST9320421ASG

Firmware Version: SD13

Serial Number: 5TJ08ZSC

Powered On: 81.2 days

Temperature: 32° C / 90° F

Bad Sectors: None

Self Assessment: Passed

Overall Assessment: ● Disk is healthy

 Don't warn me if the disk is failing
**Attributes**

ID	Attribute	Assessment	Value
	Time needed to spin up the disk		Threshold: 0 Value: N/A
4	<b>Start/Stop Count</b> Number of spindle start/stop cycles	<span style="color: green;">●</span> Good	Normalized: 100 Worst: 100 Threshold: 20 Value: 306
5	<b>Reallocated Sector Count</b> Count of remapped sectors. When the hard drive finds a read/write/verification error, it mark the sector as "reallocated" and transfers data to a special reserved area (spare area)	<span style="color: green;">●</span> Good	Normalized: 100 Worst: 100 Threshold: 36 Value: 0 sectors
7	<b>Seek Error Rate</b> Frequency of errors while positioning	<span style="color: green;">●</span> Good	Normalized: 65 Worst: 60 Threshold: 30 Value: 21494915521
			Normalized: 98

```

davidz@x61:~/Hacking
Every 2.0s: tree /dev/disk/
Wed Sep 23 16:49:59 2009
/dev/disk/
|-- by
|   |-- File
|   |-- Help
|   |-- SMART Data
|   |-- 320 GB Hard Disk - ST932042 1ASG
|   |-- DISK FAILURE IS IMMINENT
|   |-- Status
|   |-- Updated: Less than a minute ago - Update Now
|   |-- Self-tests: Last self-test completed OK - Run self-test
|   |-- Model: Maxtor 96147H8
|   |-- Firmware Version: BAC51KJ0
|   |-- Serial Number: N80BR8EC
|   |-- Powered On: 94.3 days
|   |-- Temperature: Unknown
|   |-- Bad Sectors: 71 bad sectors
|   |-- Self Assessment: FAILING
|   |-- Overall Assessment: DISK FAILURE IS IMMINENT
|   |-- Backup all data and replace the disk
|   |-- Don't warn me if the disk is failing
|   |-- Attributes
|   |-- ID | Attribute | Assessment |
|   |-- 4 | Start/Stop Count | N/A
|   |-- 5 | Reallocated Sector Count | Warning
|   |-- 6 | Read Channel Margin | Good

```

```

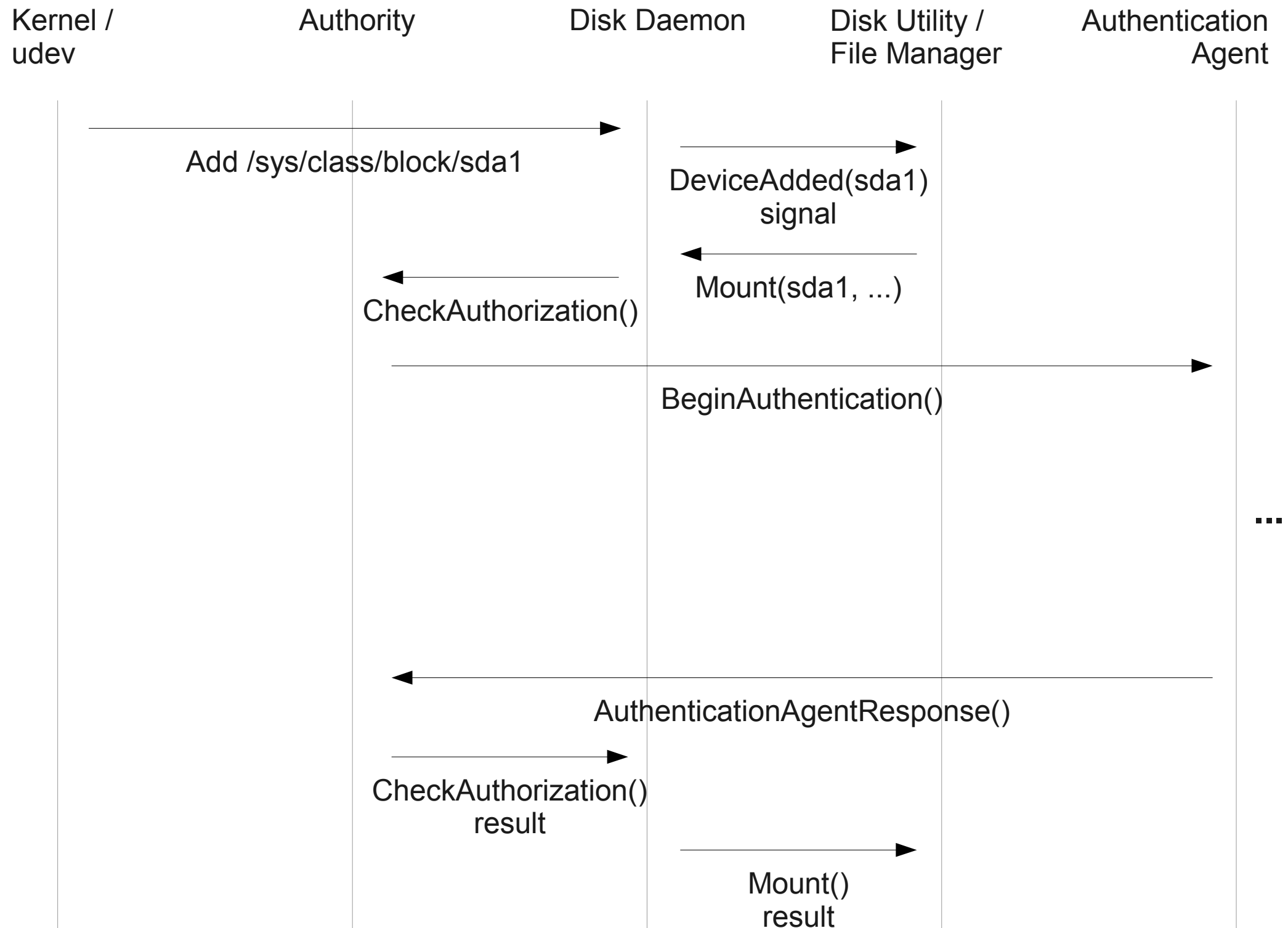
root@x61:~
partition table:
scheme: mbr
count: 1
drive:
vendor: ST932042
model: 1ASG
revision: 0100
serial: 508C88888888
detachable: 1
can spindown: 1
rotational media: 1
ejectable: 0
media:
compat:
interface: usb
if speed: 480000000 bits/s
ATA SMART: Updated at Wed 23 Sep 2009 04:49:57 PM PDT
overall assessment: Disk failure is imminent

Attribute      Current|Worst|Threshold  Status  Value      Type      Updates
-----
raw-read-error-rate      253|252| 0    n/a      441745     Old-age  Online
spin-up-time             187|183| 63    good     54 msec    Pre-fail Online
start-stop-count         252|252| 0    n/a      3210       Old-age  Online
reallocated-sector-count 226|226| 63    good     69 sectors Pre-fail Online
read-channel-margin      253|253|100   good     0          Pre-fail Offline
seek-error-rate          253|252| 0    n/a      62100      Old-age  Online
seek-time-performance    253|239|187   good     0          Pre-fail Online
power-on-minutes         247|247| 0    n/a      94.3 days  Old-age  Online
spin-retry-count         212|210|223   FAIL     176093659235 Pre-fail Online
calibration-retry-count  253|252|223   good     73         Pre-fail Online
power-cycle-count        249|249| 0    n/a      1810       Old-age  Online
reallocated-event-count  253|253| 0    n/a      0          Old-age  Offline
current-pending-sector   251|251| 0    n/a      2 sectors  Old-age  Offline
offline-uncorrectable    253|253| 0    n/a      0 sectors  Old-age  Offline
udma-crc-error-count     199|180| 0    n/a      29259      Old-age  Offline
multi-zone-error-rate    253|252| 0    n/a      339689     Old-age  Online
soft-read-error-rate     253|252| 0    n/a      648540503441 Old-age  Online
ta-increase-count        253|252| 0    n/a      441745     Old-age  Online
run-out-cancel           253|252|180   good     0          Pre-fail Online
shock-count-write-open   253|252| 0    n/a      441745     Old-age  Online
shock-rate-write-open    253|252| 0    n/a      441745     Old-age  Online
spin-high-current        230|229| 0    n/a      0          Old-age  Online
spin-buzz                 242|241| 0    n/a      0          Old-age  Online
offline-seek-performance 253|253| 0    n/a      0          Old-age  Offline
attribute-96             253|253| 0    n/a      0          Old-age  Offline
attribute-97             253|253| 0    n/a      0          Old-age  Offline
attribute-98             253|253| 0    n/a      0          Old-age  Offline
attribute-99             253|253| 0    n/a      0          Old-age  Offline
attribute-100            253|253| 0    n/a      0          Old-age  Offline
attribute-101            253|253| 0    n/a      0          Old-age  Offline
[root@x61 ~]#

```

**A hard disk may be failing**

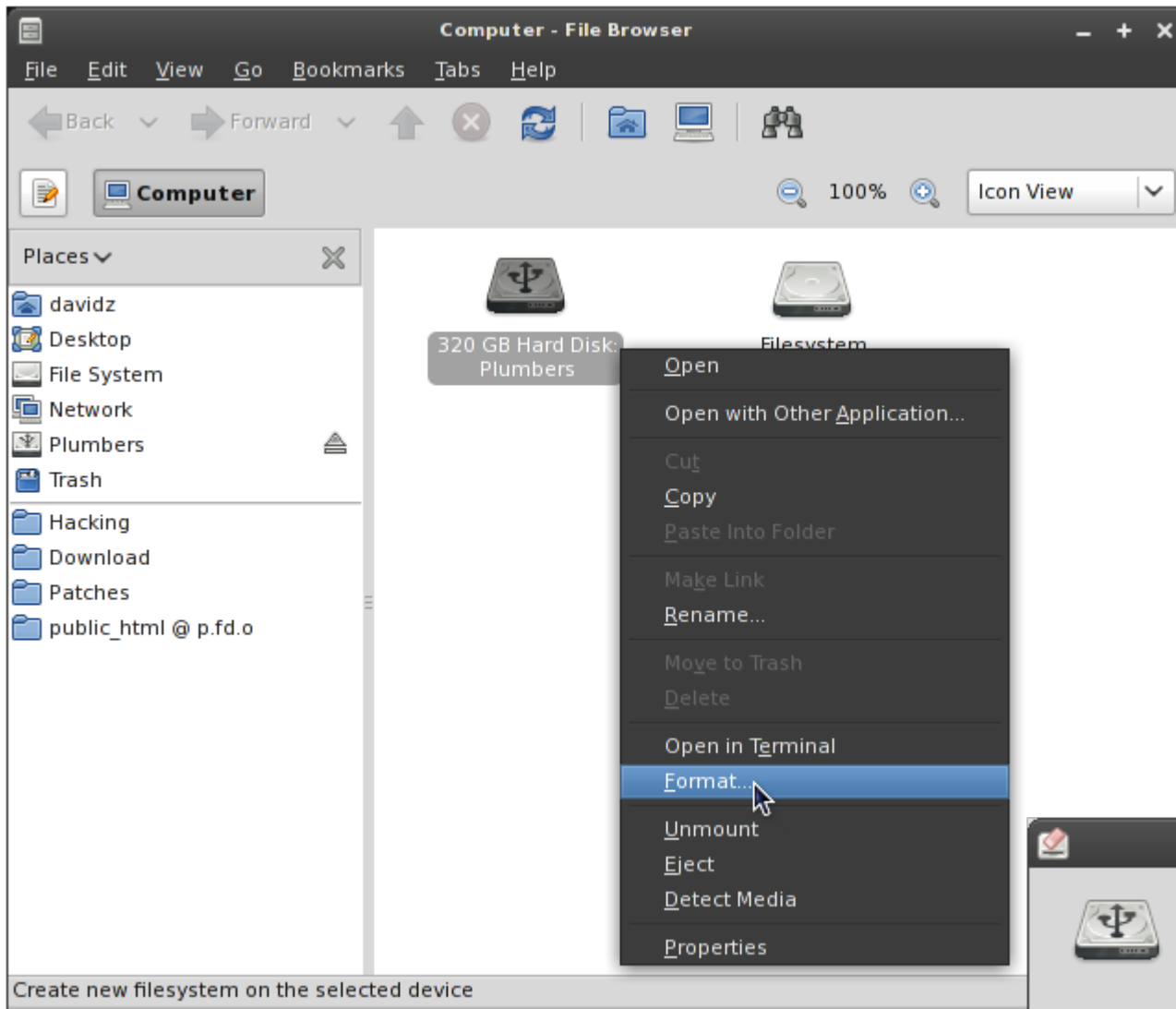
One or more hard disks report health problems. Click the icon to get more information.



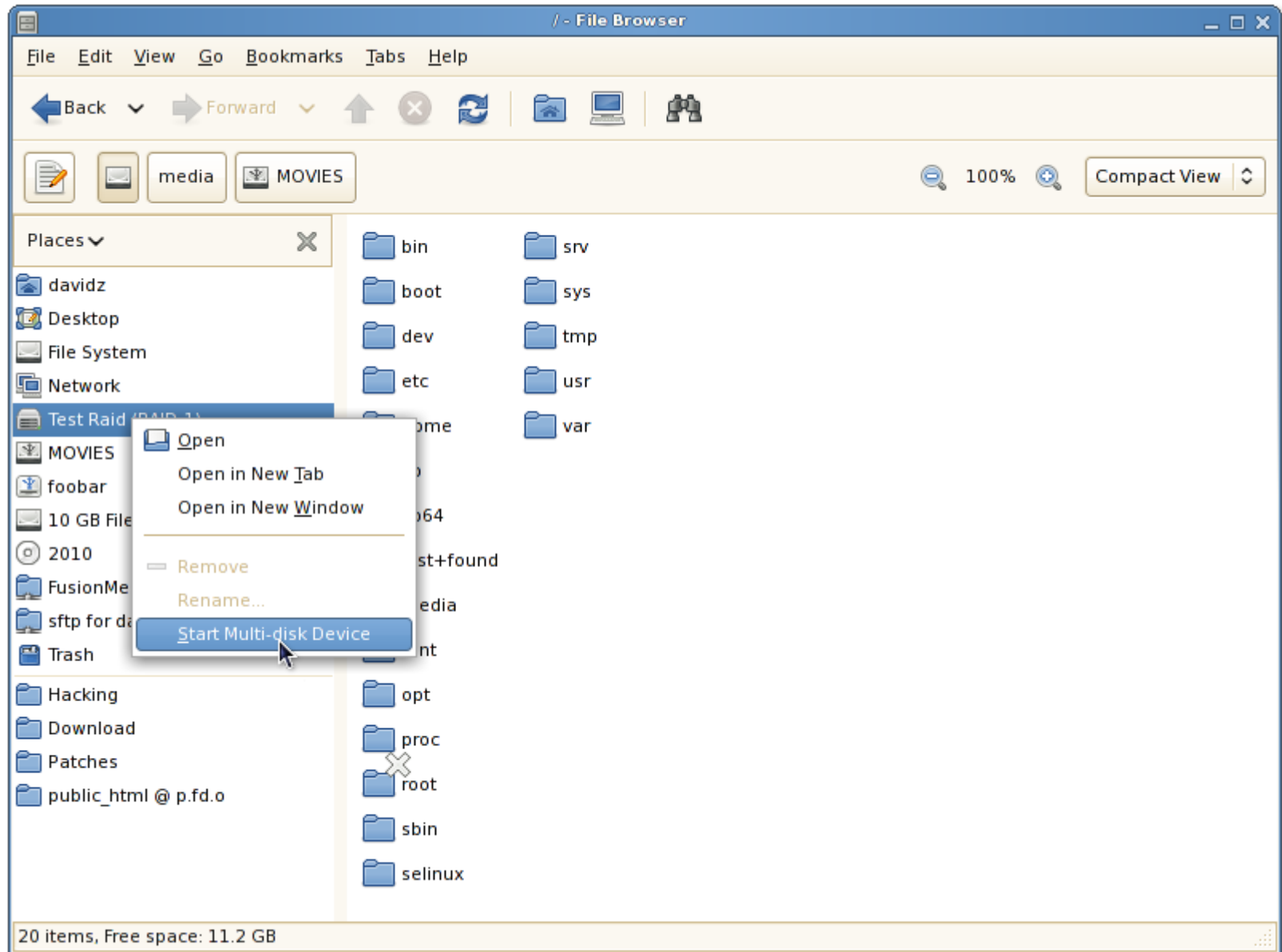
# Desktop Integration

- GVfs volume monitor
- GNOME Power Manager

# Formatting



# MD RAID





# Power Management Preferences



On AC Power

On Battery Power

General

## Actions

Put computer to sleep when inactive for:

Never



When laptop lid is closed:

Suspend



Spin down hard disks when possible

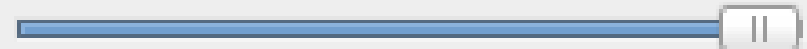
## Display

Put display to sleep when inactive for:

30 minutes



Set display brightness to:



100%

Dim display when idle

Help

Make Default

Close

Questions?

# Docs / References

- <http://www.kernel.org/pub/linux/utils/kernel/hotplug/libudev/>
- <http://www.kernel.org/pub/linux/utils/kernel/hotplug/gudev/>
- <http://hal.freedesktop.org/docs/DeviceKit-disks/>
- <http://hal.freedesktop.org/docs/DeviceKit-power/>
- <http://hal.freedesktop.org/docs/polkit/>