Open Source Technology Center

ACPI in Linux

Len Brown

len.brown@intel.com

July 20, 2005

intel

Agenda

- ACPI Tutorial/Decoder-Ring
- What changed in the last year?
- ACPI 3.0
- What's next?

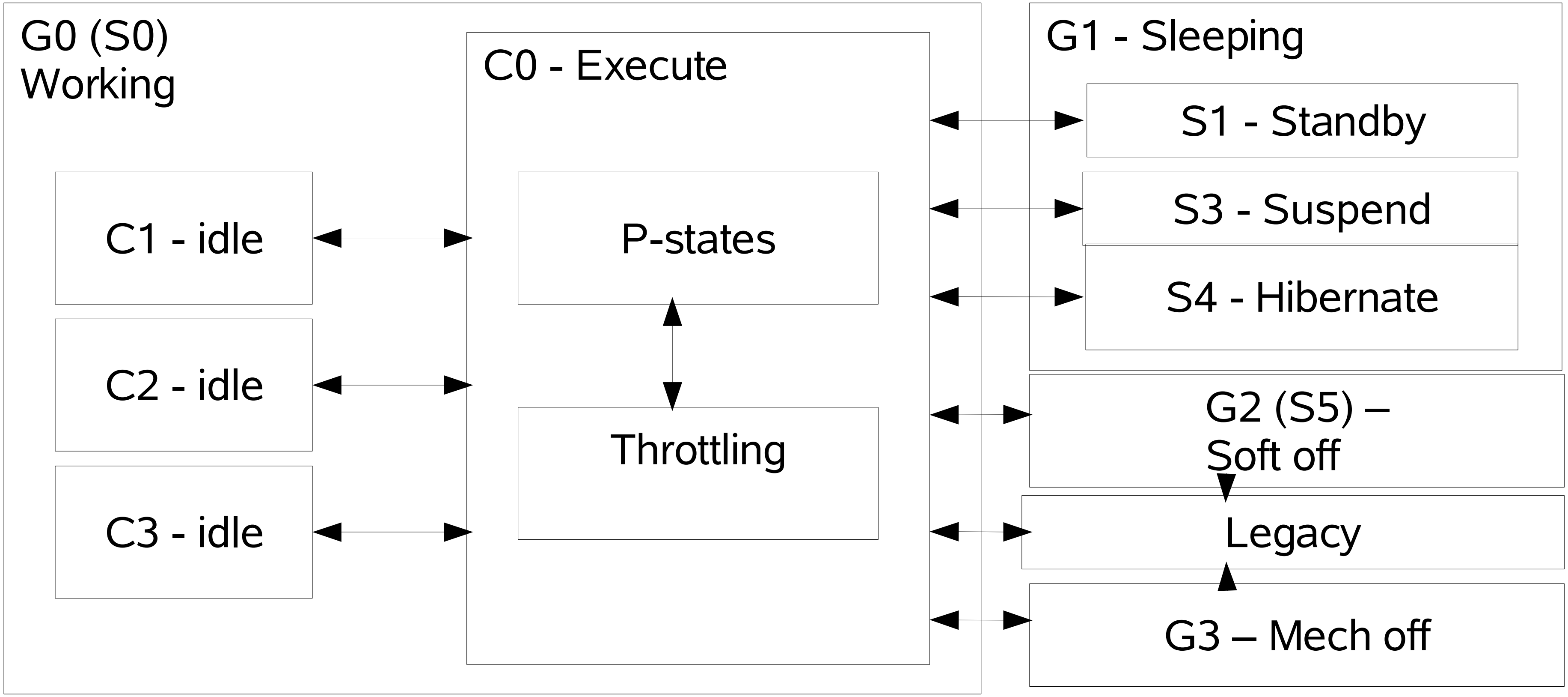
ACPI Tables

- RSDP – Root System Description Pointer
- RSDT – Root System Description Table
- FADT – Fixed ACPI Description Table
- DSDT – Differentiated System Description Table
- SSDT – Secondcary System Description Table
- SRAT – System Resource Affinity Table
- SLIT – System Locality Distance Inf. Table

ACPI Configuration

- **Replaces Legacy configuration standards and proprietary methods**
- **Enumerates Motherboard devices:**
  - Processor
  - Memory
  - I/O devices
  - Interrupts
  - Hot Plug
- **ACPI comes with the BIOS – it does not know anything about add-in devices.**
- **ACPI augments, but does not replace PCI config.**

ACPI States


G0 (S0)
Working
C0 - Execute
C1 - idle
C2 - idle
C3 - idle
P-states
Throttling
G1 - Sleeping
S1 - Standby
S3 - Suspend
S4 - Hibernate
G2 (S5) –
Soft off
Legacy
G3 – Mech off

System Sleep S-States

- S0 = Executing
- S1 = Standby
-
- S3 = Suspend to RAM
- S4 = Suspend to Disk
- S5 = Soft Power-off

C-States – CPU Idle Power States

- Only C0 executes instructions
- C1 – Cn do no execute instructions
- Deeper C-states save more power cost more latency
- /proc/acpi/processor/*/power

Processor Performance P-States

- P0 = full MHz (and voltage)
- P1 … Pn = progressively lower MHZ and Volts
-
- Different platforms offer different states
- /sys/devices/system/cpu/cpu0/cpufreq
- speedstep-centrino, acpi-cpufreq, cpufreq-stats

Throttling T-states

- ~Lineary power savings w/ Mhz
- Older system had only throttling to manage processor performance.
- P-states used for modern hardware, with T-states used only for thermal emergencies.
-
- cpufreq sub-system can use only 1 method at a time. eg. p4-clockmod vs cpufreq-acpi
- ACPI in addition to hardware guards TM1/TM2

Processor Power Saving Example

C-State	P-State	Mhz	Volt	Watt
C0	P0	1600	1.5	24.5
	P1	1300	1.4	22.0
	P2	1100	1.2	12.0
	P3	600	1.0	6.0
C1,C2	from P0	0	1.5	7.3
	from P3	0	1.0	1.8
C3	from P0	0	1.5	5.0
	from P3	0	1.0	1.0
C4	(any)	0	0.8	0.5

Device D-States

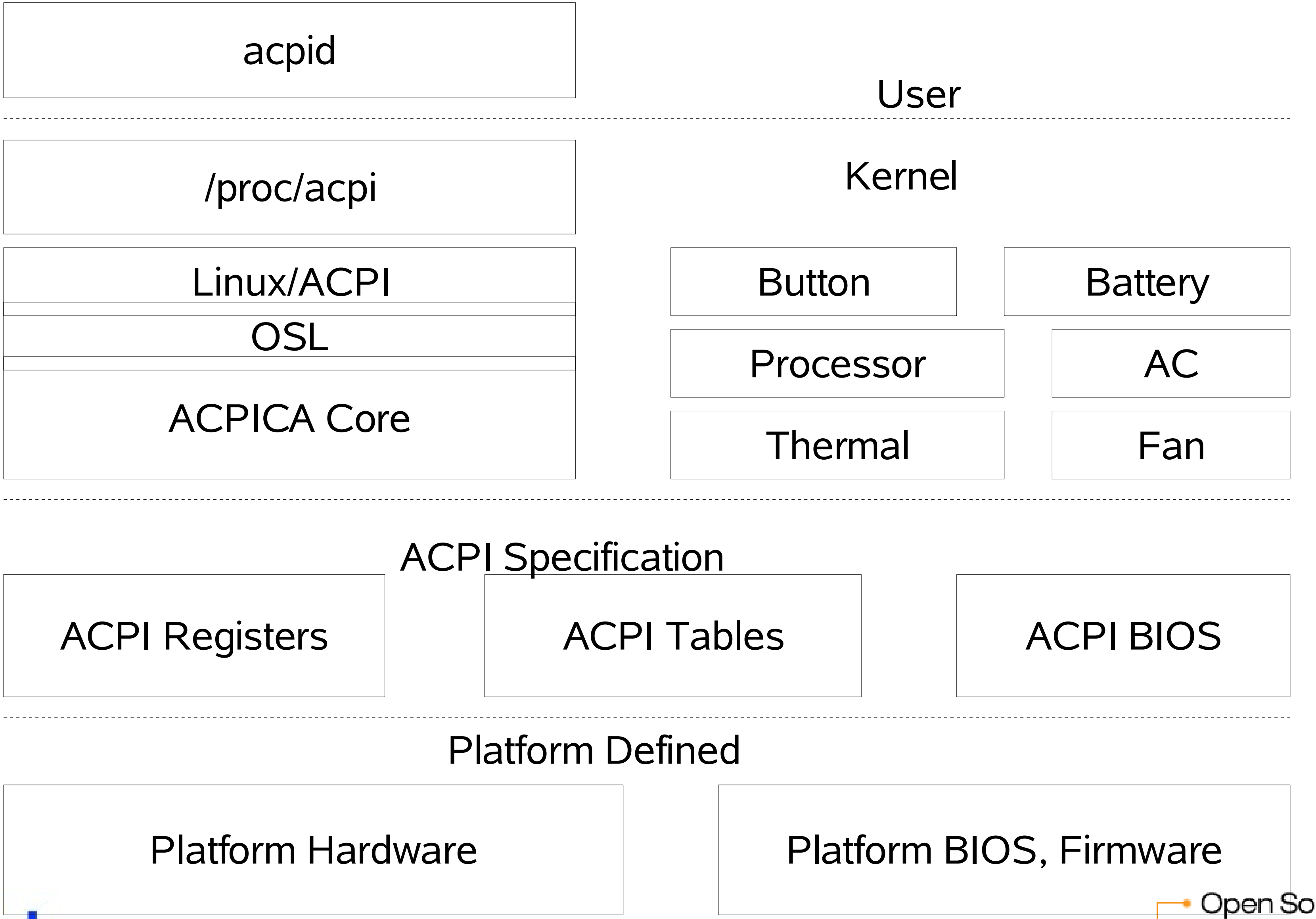
- D0 – full on
- D3 – full off
- vs. PCI Power Management
- Semanntics are device-class specific
-
- Linux just starting to use ACPI D-states

ACPI Events

- **System Control Interrupt (SCI)**
  - Fixed Feature Events (FF)
  - General Purpose Events (GPE)
  - /proc/acpi/event

Linux/ACPI Deployment

- This year the major Linux distributors are all shipping with ACPI built-in and enabled-by-default on all ACPI capable architectures (i386, x86_64, ia64)



acpid

User

/proc/acpi

Kernel

Linux/ACPI

OSL

ACPICA Core

Button

Battery

Processor

AC

Thermal

Fan

ACPI Specification

ACPI Registers

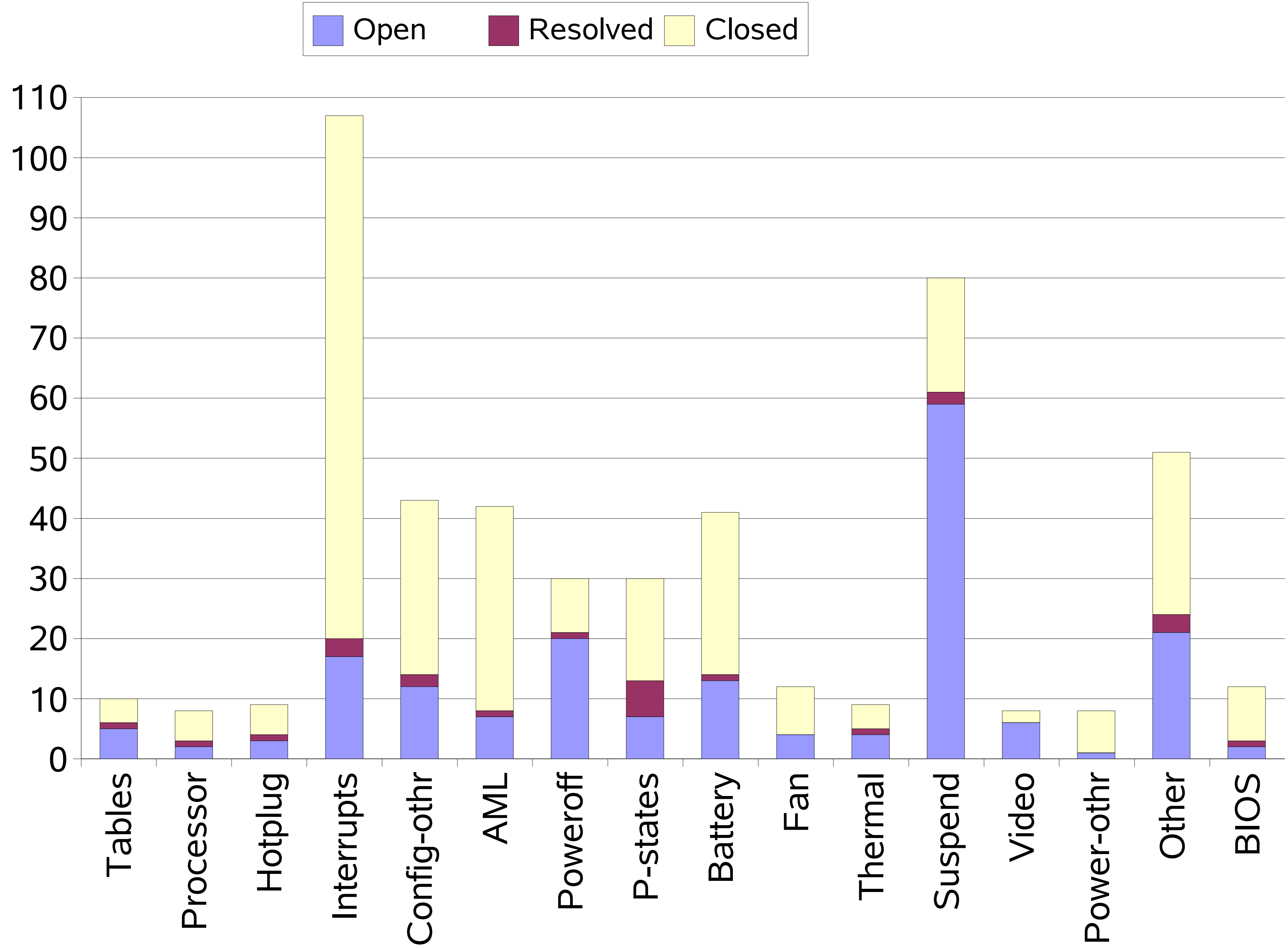
ACPI Tables

ACPI BIOS

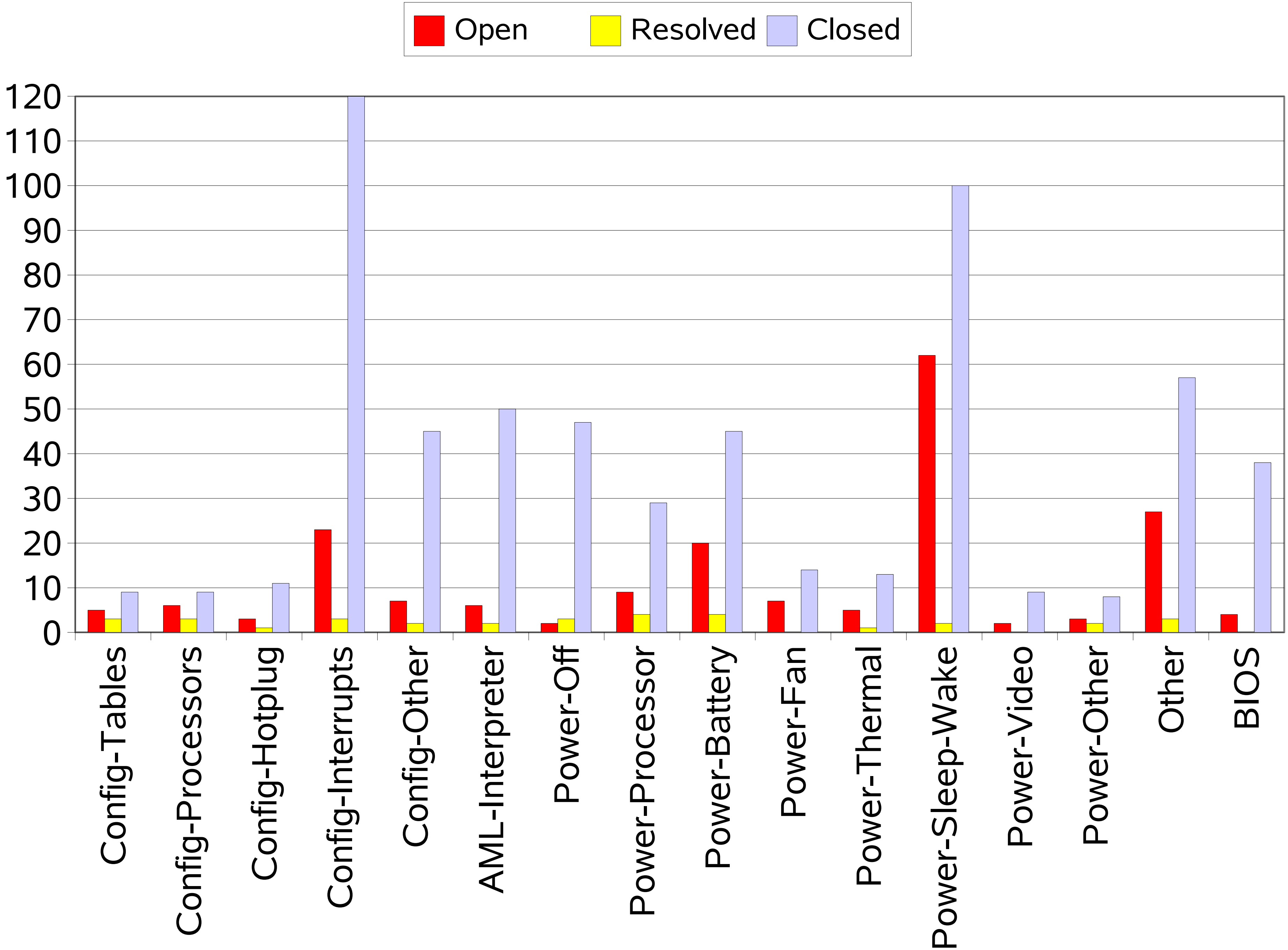
Platform Defined

Platform Hardware

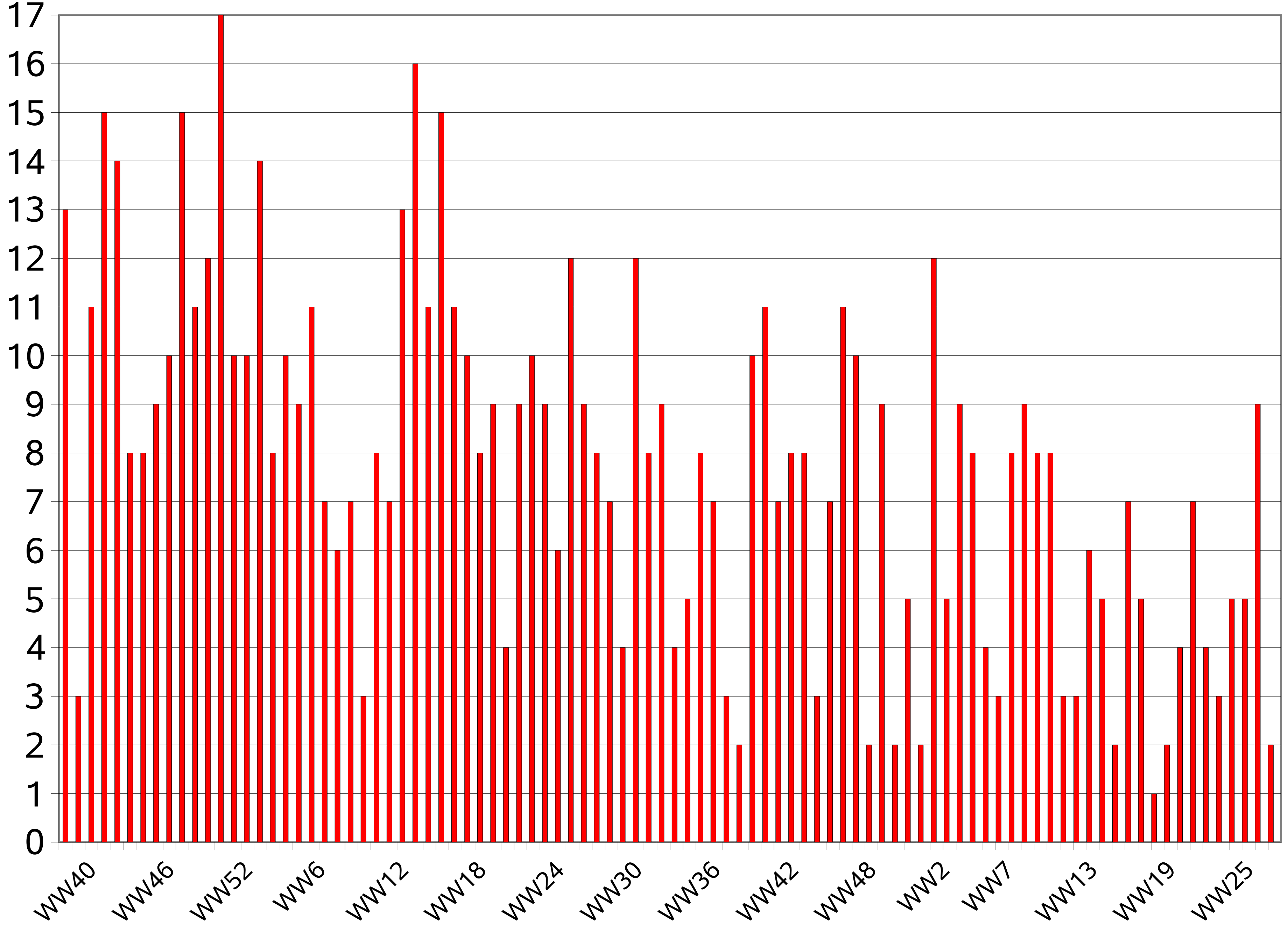
Platform BIOS, Firmware



Open
Resolved
Closed
110
100
90
80
70
60
50
40
30
20
10
0
Tables
Processor
Hotplug
Interrupts
Config-othr
AML
Poweroff
P-states
Battery
Fan
Thermal
Suspend
Video
Power-othr
Other
BIOS



Open
Resolved
Closed
120
110
100
90
80
70
60
50
40
30
20
10
0
Config-Tables
Config-Processors
Config-Hotplug
Config-Interrupts
Config-Other
AML-Interpreter
Power-Off
Power-Processor
Power-Battery
Power-Fan
Power-Thermal
Power-Sleep-Wake
Power-Video
Power-Other
Other
BIOS



New bugs/week

Linux/ACPI changes in past year

- Linux PNP = PNPACPI, PNPBIOS, PNPISA
- Hot Plug CPU, Memory
- SMP Deep C-states
- SMP S3/S4 progress
- video control progress
- hot-key progress
- ACPICA Interpreter progress

ACPI Specification history

- 1996 ACPI 1.0
- 2000 ACPI 2.0
- 2004 ACPI 3.0

Device Drivers vs. Power Management

- Devices need to recognize when idle and take power saving action.
- Device drivers should not have to care if power management is handled by PCI-PM or ACPI.
- Today, no connection between (conventional) motherboard device drivers and ACPI for resources or power management.

Linux need Power Policy Mgr

- User specifies policy
- manager communicates it to devices and sub-systems

Configuration Todo

- **Fix device tree**
- **Abstract Device Enumeration:**
  - PCI/ACPI/Legacy resources
- **Docking Station support -**

Power Management Todo

- Suspend/Resume reliability++
- Abstract Device power states
  - PCI/ACPI D-states
- System Power policy manager
- C-states: tick timer stop in idle
- Hot Keys ++

References

- http://www.acpi.info
- http://acpi.sourceforge.net