Challenges Revisited in Supporting Virt CPU Hotplug on architectures that don’t Support CPU Hotplug (like ARM64)

OR: Supporting hotplug on architectures that don’t.

Salil Mehta (Huawei)
James Morse (Arm)
What is it good for?

Cloud orchestration:

- Pre-deploy a VM, then scale it by adding CPUs when the workload is deployed.
History: what are we talking about?

+ ‘Hotplug’ in the arm world typically means CPU online/offline.

+ Physical CPU hotplug involves moving the CPU package between machines.
  - Making CPUs that were not present at boot, present.

Return Value:

An Integer containing a device status bitmap:

- Bit [0] - Set if the device is present.
- Bit [1] - Set if the device is enabled and decoding its resources.
- Bit [2] - Set if the device should be shown in the UI.
- Bit [3] - Set if the device is functioning properly (cleared if device failed its diagnostics).
- Bit [4] - Set if the battery is present.
- Bits [31:5] - Reserved (must be cleared).

ACPI 6.5, 6.3.7 _STA (Device Status)
The only way to win, is not to play...

- CPUs are really a slice of the system.
  - Each CPU has a GIC redistributor, a chunk of cache, a PMU, RAS ERR nodes, MPAM MSCs ...
  - A group of CPUs may come with an ITS, an IOMMU a PCIe root-complex, memory controllers, miscellaneous non-discoverable devices.

- There is no hardware that does anything of this.

- No appetite for updating numerous specifications to describe hotplug support.

- Whatever is defined for virtual machines needs to work on physical hardware too – the OS doesn’t know its in a VM.
Do you think you can tell?

- For a virtual machine, this is just a CPU online/offline policy.
  - Nothing about the ‘SoC topology’ changes.
- The OS must know that all resources remain present the whole time.
- It must look and smell like physical CPU hotplug.
CPUs can check out any time they like …. 

- ACPI MADT advertises CPUs as disabled, but online-capable.
- Firmware policy prevents a CPU from being brought online.
  - For a VM, the VMM plays the part of firmware.
  - Enforced by PSCI CPU_ON returning DENIED.

- ACPI notifications are used to toggle the _STA enabled bit.
- CPUs are registered (allowing them to be brought online) when they are enabled.
Once during boot
Runtime updates

ACPI:device_check → ACPI:_STA → cpu_register() → sysfs

Add to cpu_online_mask

Remove from cpu_online_mask

ACPI:eject_request