

Ensuring Allocation for
Required BARs

And

Evenly Allocating
Optional BARs

Defining Required and Optional BARs

Example, NVMe:

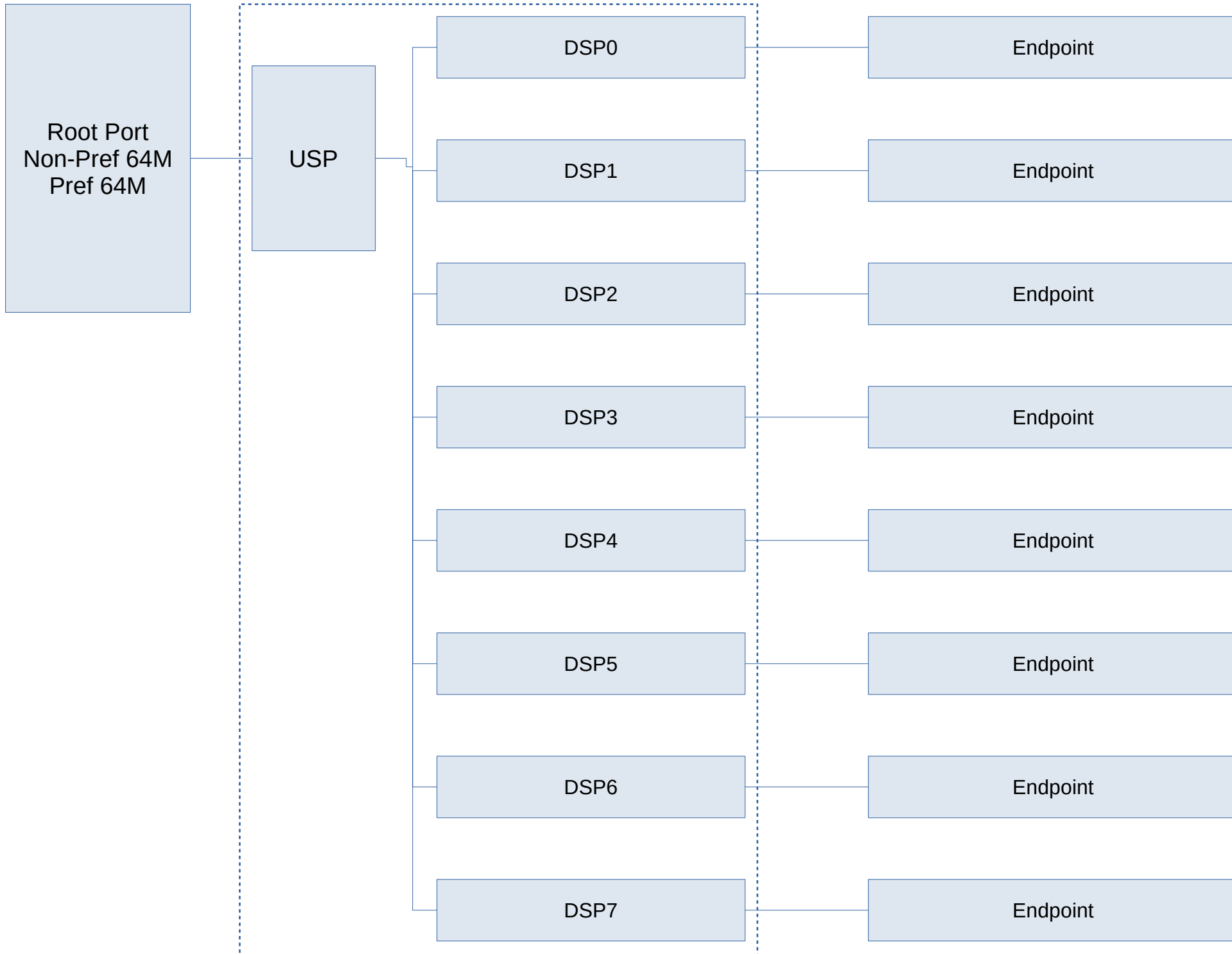
BAR 0 (M) – Registers

BAR N (O) - CMB/PMR

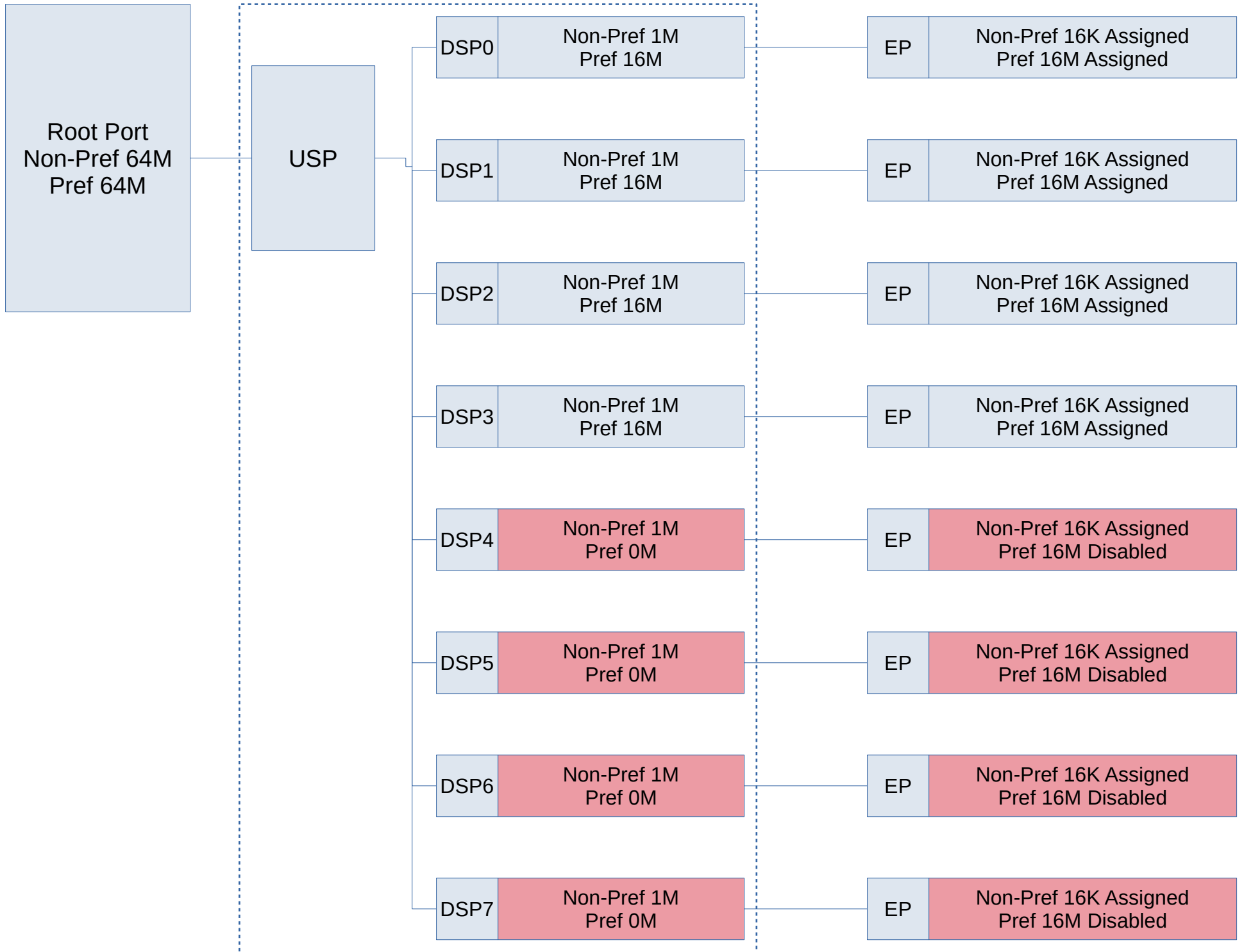
Problem statement:

1. Limited resources on PCI domains often doesn't allow for the use of optional BARs
2. Max-first resource assignment can cause endpoints to not have required BARs assigned
3. Limited resources and Max-first assignment can cause uneven resource assignments between different endpoints in the same hierarchy

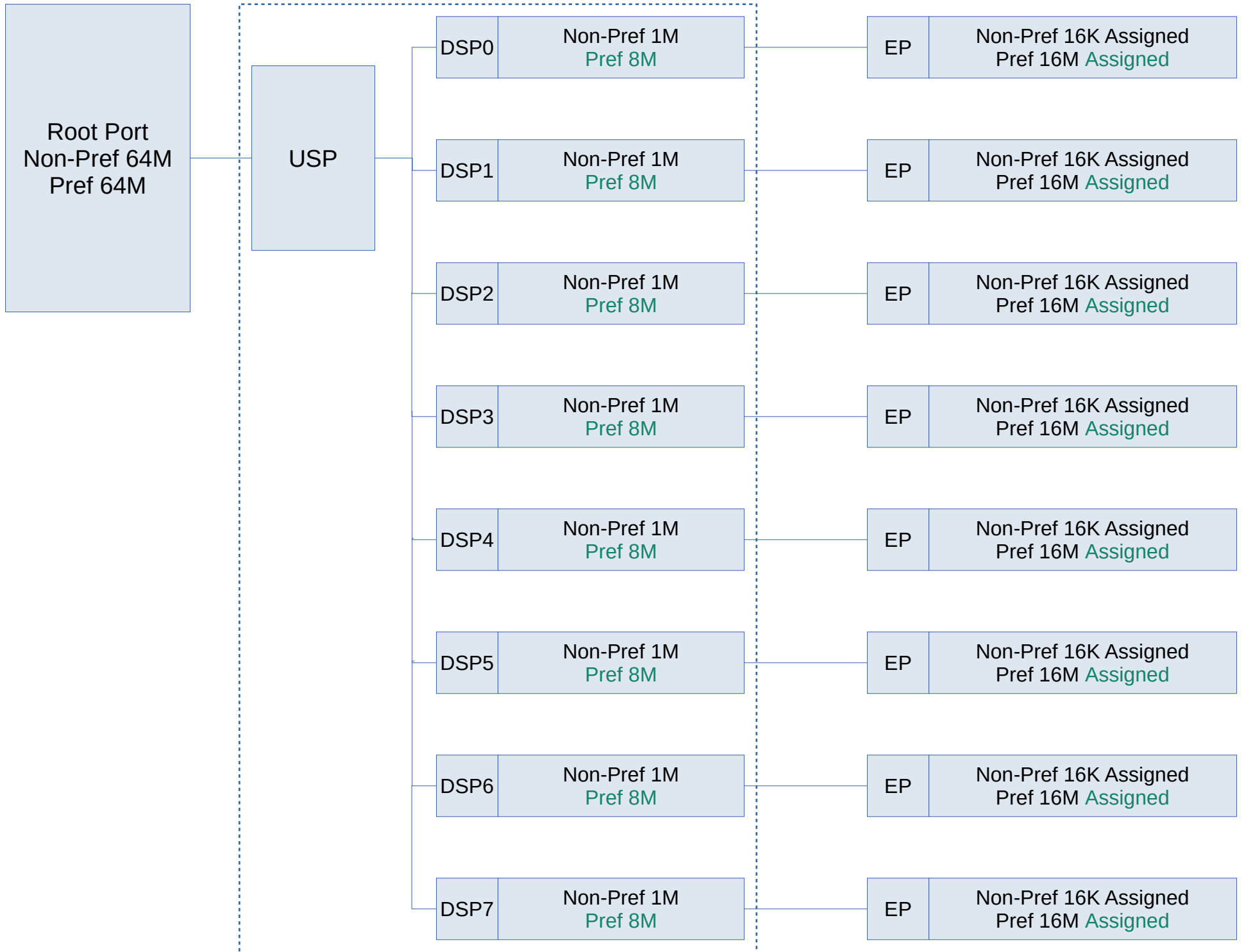
Example Hierarchy



Less Ideal



More Fair

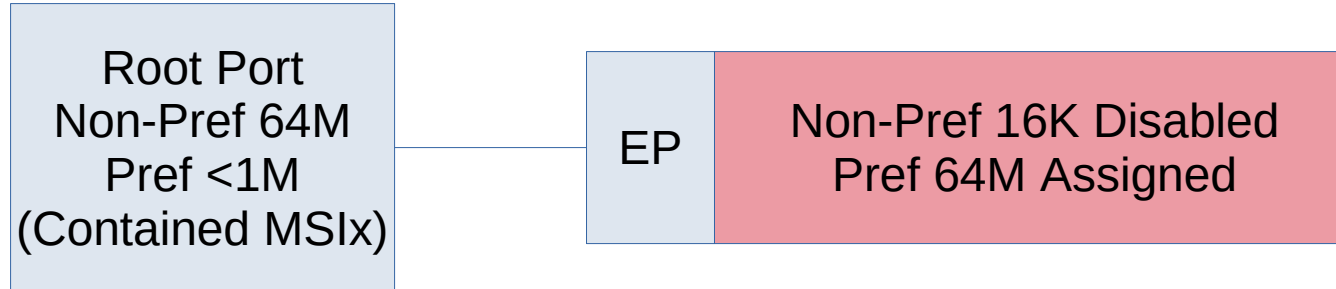


02:00.0 Non-Volatile memory controller:

...

Memory at <ignored> (64-bit, non-prefetchable)

Memory at a0000000 (64-bit, prefetchable) [size=64M]

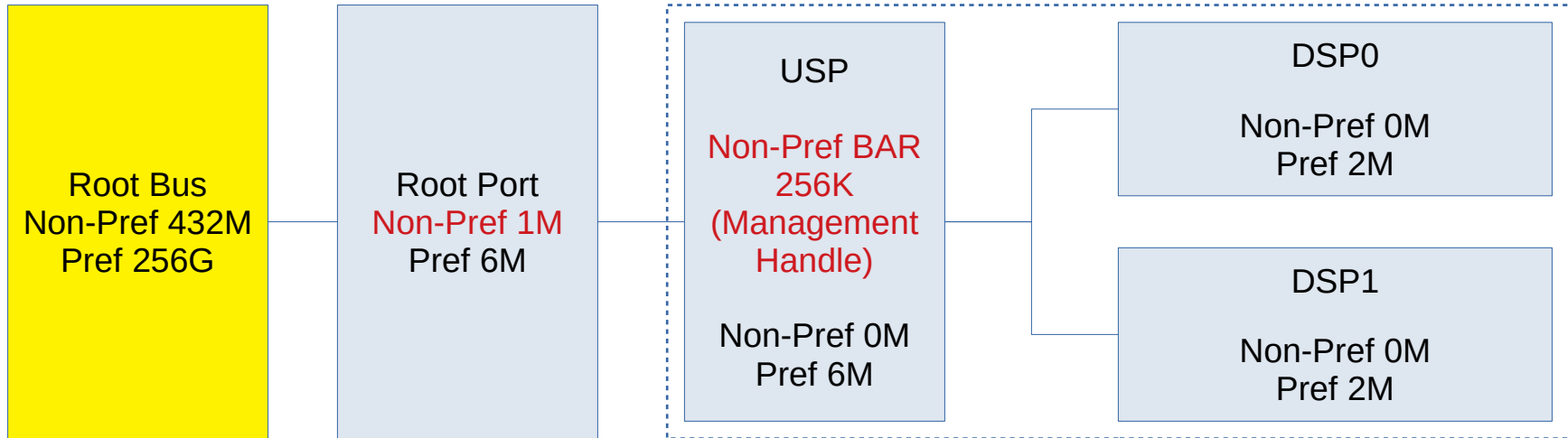


```
static int __pci_assign_resource(struct pci_bus *bus, struct pci_dev *dev,
                               int resno, resource_size_t size, resource_size_t align)
{
    ...
    /*
     * If we didn't find a better match, we can put any memory resource
     * in a non-prefetchable window. If this resource is 32 bits and
     * non-prefetchable, the first call already tried the only possibility
     * so we don't need to try again.
     */
    if (res->flags & (IORESOURCE_PREFETCH | IORESOURCE_MEM_64))
        ret = pci_bus_alloc_resource(bus, res, size, align, min, 0,
                                     pcibios_align_resource, dev);

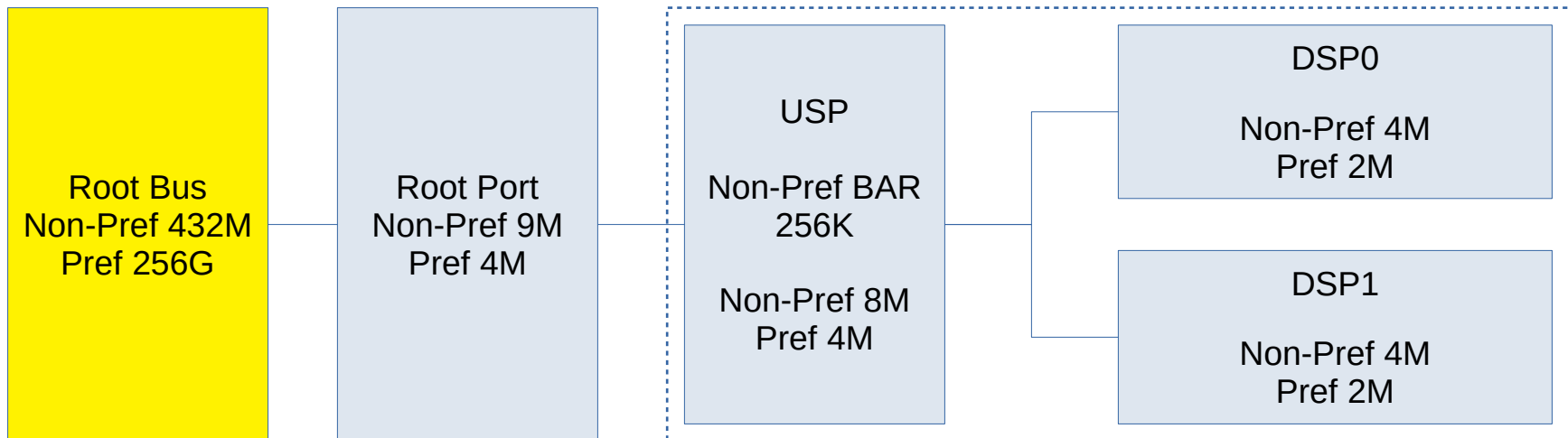
    return ret;
}
```

Bjorn/next as of 5.3-rc6

BIOS Non-Pref Assignment



pci=realloc



Proposal: Two-stage sort and assignment

1. Determine which BARs are required and which are optional
 - a. Class type lookup
 - b. Optional BARs give minimum size, ex. 1MB
2. Sort required BAR resource requirements
3. Assign windows to required BARs for all devices in hierarchy
4. Repeat for Optional BARs

Allowing truncated optional BARs:

1. Determine hierarchy's max remaining size after required BARs are assigned
2. Spread evenly or round-robin (needs to be fair to devices with differing size optional BARs)
3. Apply limit at each bridge to fit portion of optional BARs, down to minimum size given by lookup

Concerns: Device misunderstanding own address space (P2P misrouting)