



HVM PCI Pass-Through

XenSummit – April 2007

Guy Zana

- We are a Software Company
 - in stealth mode
- Not a Virtualization company per se
- Parts of what we do is Open Source

HVM requirements in a Desktop environment are different

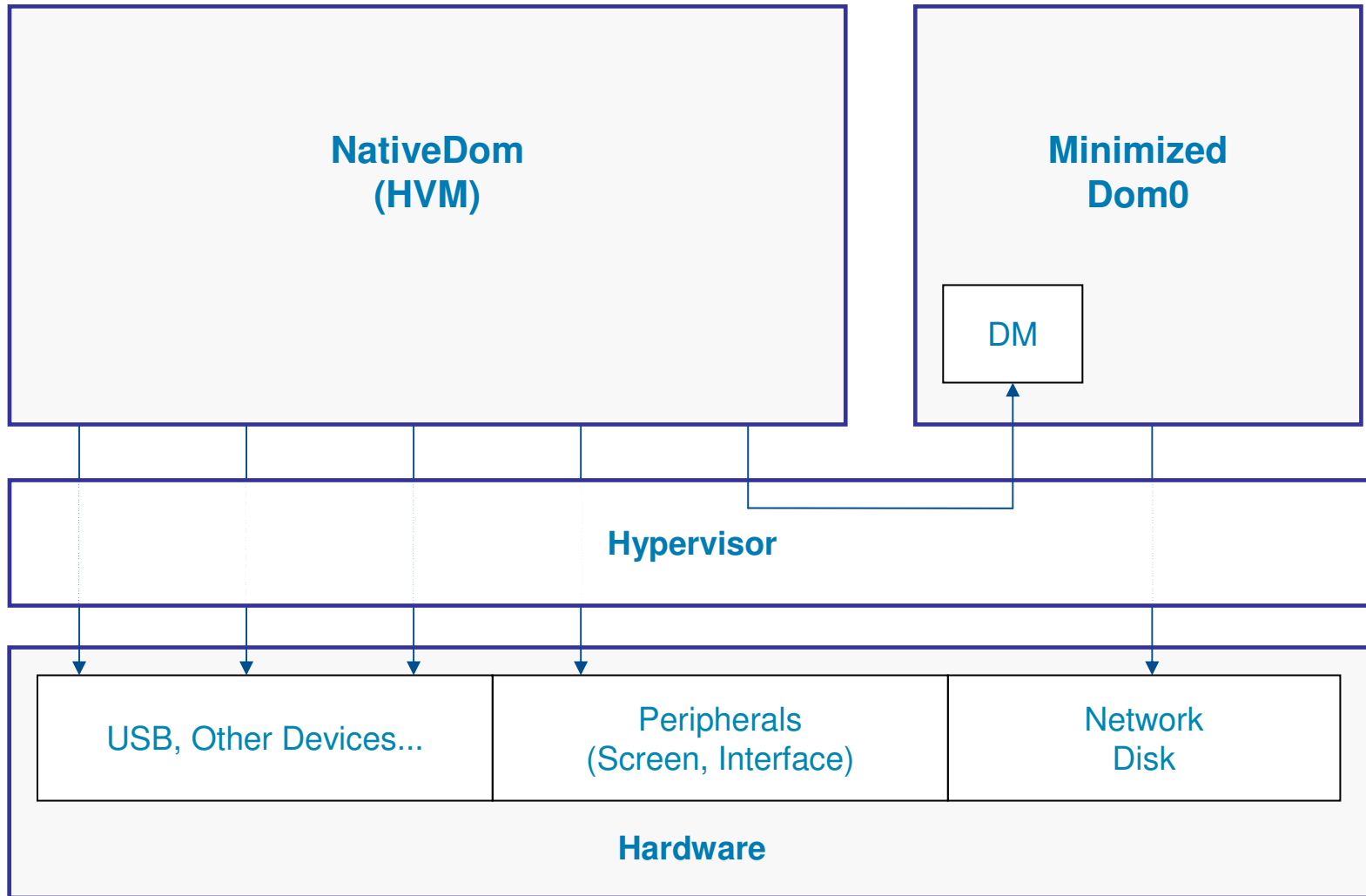
- There is an end-user
- Dynamic configuration (e.g. network)
- Unknown hardware characteristics
- P2V is a must

- Critical Success Factors
 - Performance
 - Security
 - Usability
 - Hardware Compatibility

- Address Translation
- Security
- Resource Management
- ACPI
- Integrated Devices (Chipsets)
- Hardware ROMs (VGABIOS, Etherboot...)

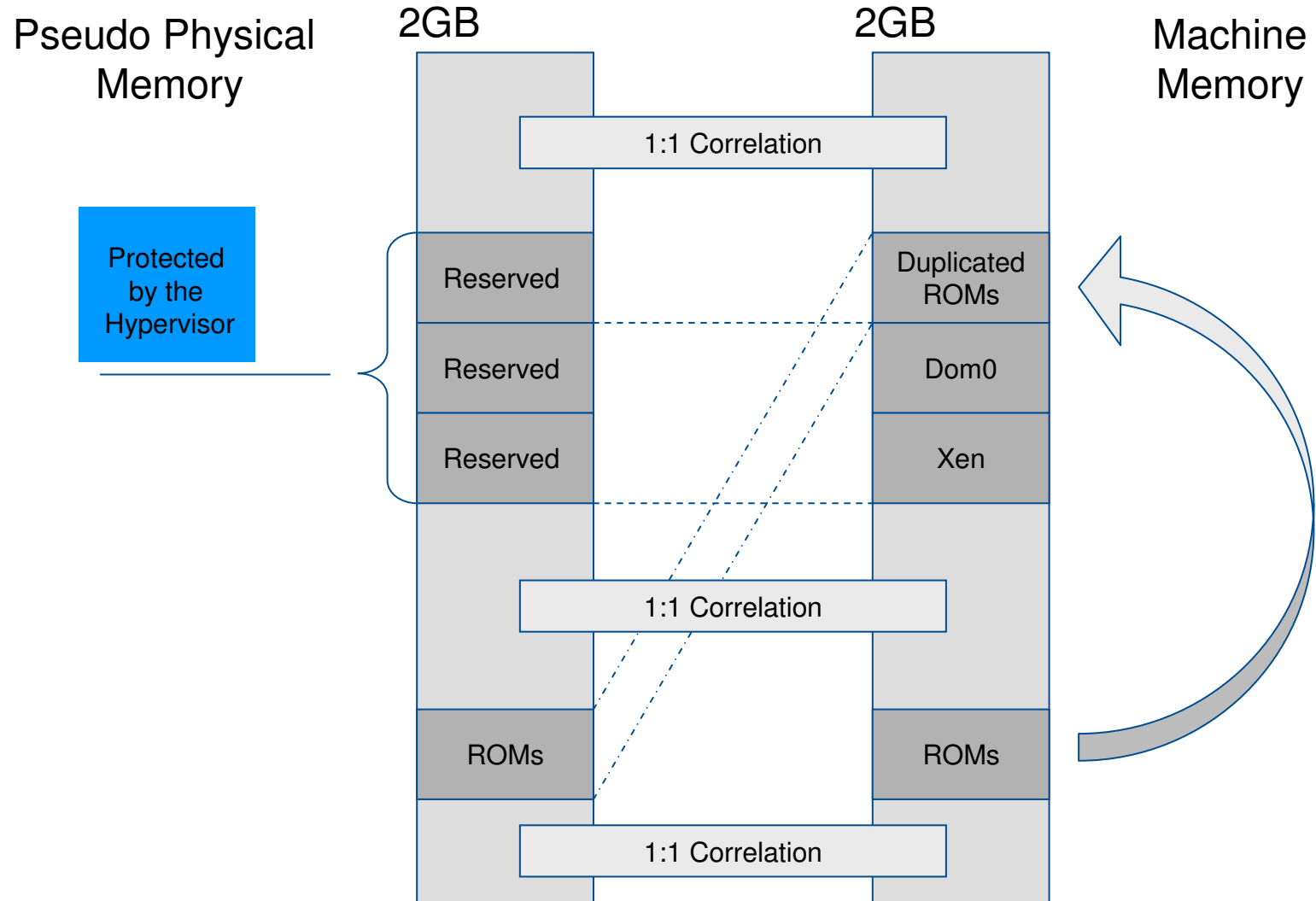
Outline

- Basic Concepts
- Memory Layout
- PCI Configuration
- ACPI et al.
- HW Resource Routing
- Development Status



- Hardware Devices Categorization
 - Real vs. Emulated
 - Dedicated vs. Dynamically Allocated
 - Shared vs. Non-Shared
- Supporting IOMMU
- Supporting Non-IOMMU Configurations
 - This is what we have today
 - 100s of millions of computers that should be supported

- Target: Non-IOMMU Systems
- Solution: 1:1 Memory Layout (NativeDom)
 - pfn == mfn
 - Motivation
 - DMA
 - Implementation
 - E820 Refactoring
 - Protecting Memory



- E820 Map

```
(XEN) NEO: Initializing...
(XEN) NEO: Machine Available RAM: 2120036352
(XEN) NEO: RAM Available for Xen: 254406656
(XEN) NEO: Dom0 E820 After manipulations:
(XEN) 0000000000000000 - 000000000008f000 (usable)
(XEN) 000000000008f000 - 00000000000a0000 (reserved)
(XEN) 00000000000e0000 - 00000000000100000 (reserved)
(XEN) 00000000000100000 - 00000000007400000 (usable)
(XEN) 00000000007400000 - 00000000006f300000 (1:1)
...
(XEN) 000000007e64f000 - 000000007e6a5000 (ACPI NVS)
(XEN) 000000007e6a5000 - 000000007e6aa000 (ACPI data)
...
```

- 0xCF8 / 0xCFC Trapping
 - Allowing HVM access to the real PCI config space
 - Exposing the PCI configuration space as a composition of real and emulated devices
- Hiding Certain Real Devices
- Resource Balancing
 - BAR emulation – detecting changes
- PCI-PCI Bridges
 - Fake devices

- ACPI – Provide configuration for the system
- PT-Devices might need ACPI info.
- Providing a “smart” ACPI world view for the HVM
- Applying logic when parsing
 - Expose
 - Hide
 - Manipulate
 - Add new tables/routines etc...

- Letting the native driver drive the real hardware
- Resources: PIO/MMIO/Interrupts
- PIO and MMIO is trivial
- Interrupts
 - Assert when an interrupt occur
 - De-Assert when the corresponding bit in the IRR is low.
- Update Mapping Event
 - Updating access functions



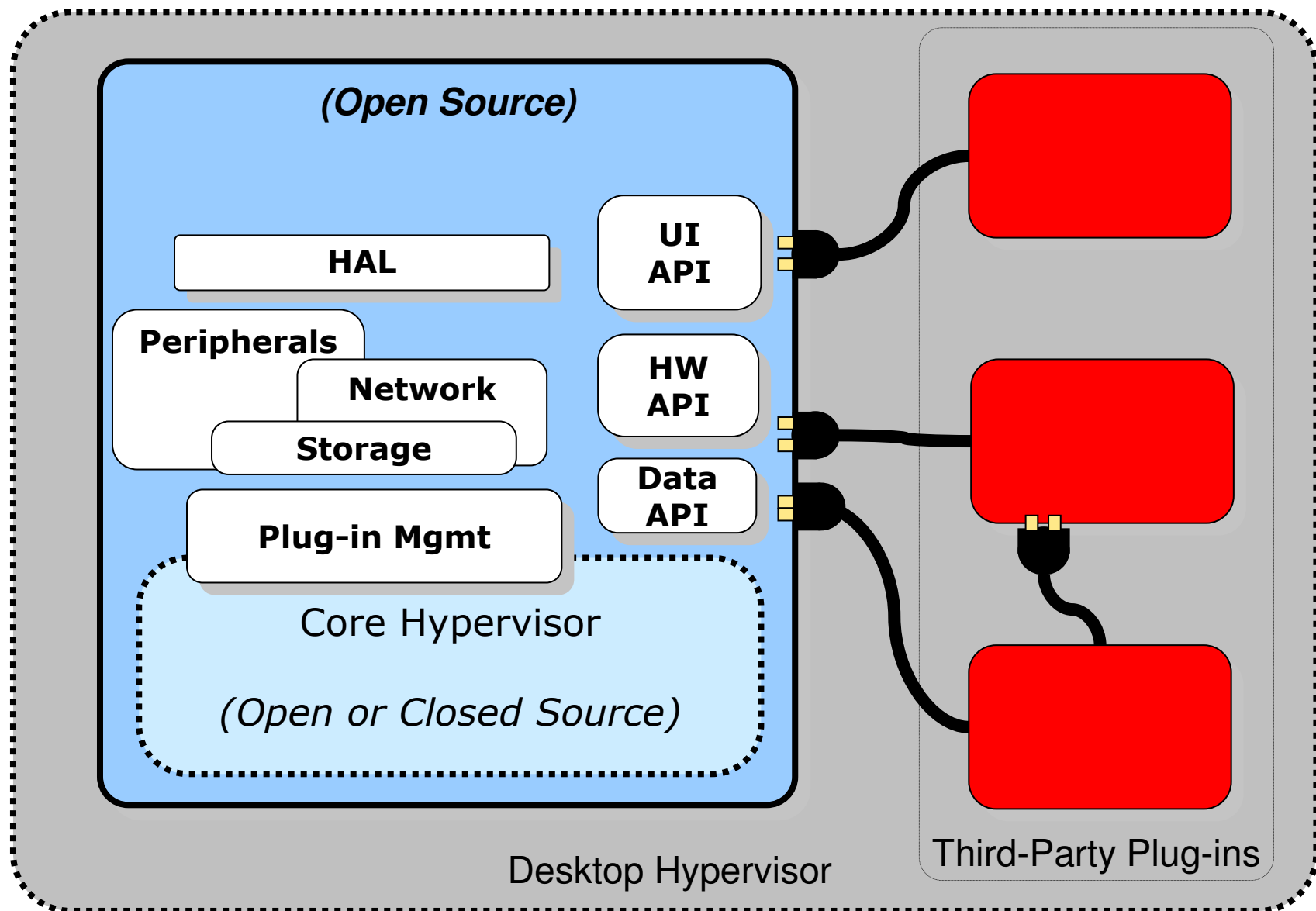
The Big Picture

Etay Bogner

- Why?
 - Well, only one Hypervisor can be installed...
- Requires a Plug-in framework for 3rd party ISVs
 - Define APIs
- Our offering: build such a framework for existing Hypervisors
 - As an Open Source project

Enabling 3rd party ISVs (such as ourselves), to build desktop SoftAppliances

Desktop Hypervisor Schematics



- Looking for Cooperation/Collaboration
 - 3rd Party ISVs like Neocleus
- Please contact us at “DHF at neocleus.com”
 - “Guy at neocleus.com”
 - “Etay at neocleus.com”



Thank You