DSwitch: A Dual Mode Direct and Network Attached Disk

Quanlu Zhang, Yafei Dai, and Lintao Zhang
Peking University and Microsoft Research

Motivation

Power saving is a big concern in both data center and home/office scenarios.

- Idle state occupies about *50%* fraction of the time on average for PCs in homes and enterprises.
- Server utilization is only about *20-30%* in typical data center. For interactive services the utilization is even lower than *10%*.

DVFS☹ : difficult to make all the components of a computer power proportional (even for heterogeneous multi-core processor).

ACPI ☹ : storage is not available.

Others ☹ : apply special data placement policies.

Directly attached storage (DAS):

- High performance
  - -- Full bandwidth
  - -- Relatively predictable latency
- Low cost, less stress on the network, lower power during peak load
- When machine hibernates, data is unavailable ☹

Network attached storage (NAS):

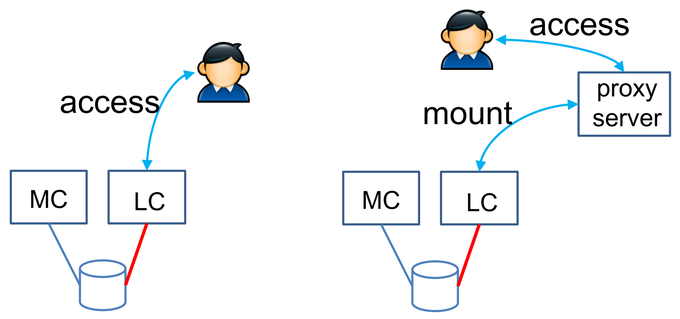
- Data is always available
- Low performance ☹
  - -- Bandwidth limited by network
  - -- Network congestion may cause latency unpredictability
- High cost, stress on the network, higher power during peak load ☹

☺ Combine the merits of DAS and NAS, flexibly couple and decouple storage and computation.

Use Cases

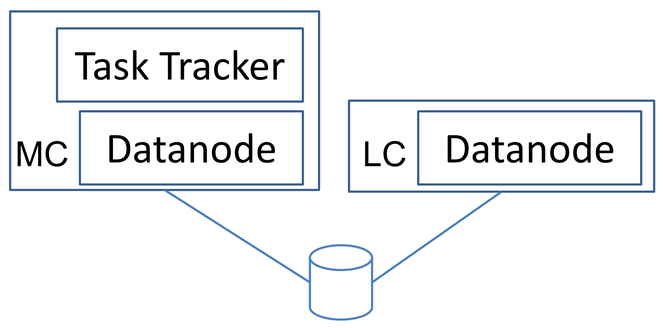
Office Scenario:

- Both service migration and VM migration can be deployed in office scenario.



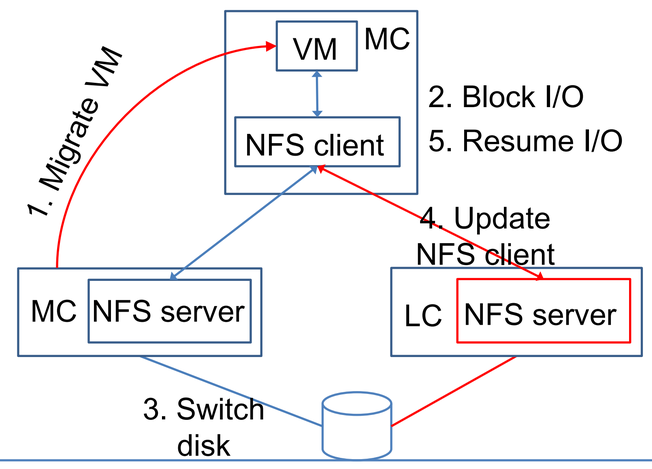
Hadoop:

- We keep Hadoop and its data placement policy unchanged.
- We use service migration, but only datanode is resumed on LC.



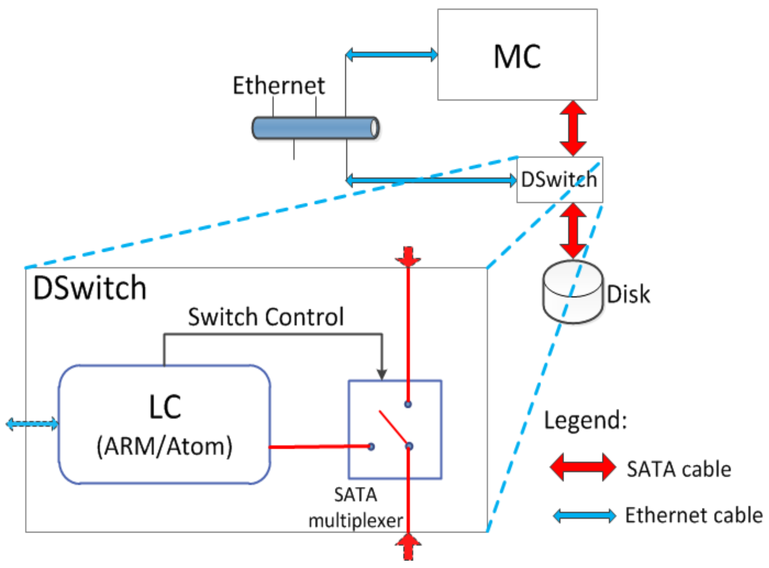
VM Consolidation (live migration):

- NFS is employed for data access
- NFS is modified a little, so as to cover the interruption of switching disks.

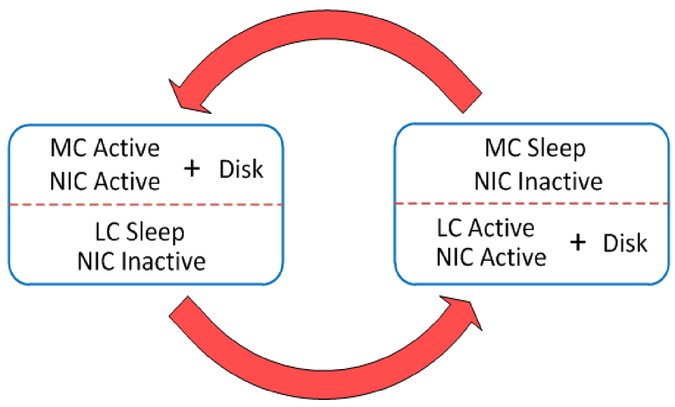


DSwitch Architecture

DSwitch Hardware Architecture:



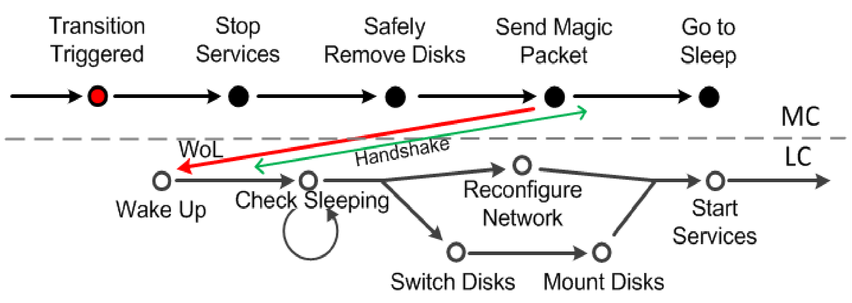
DSwitch works in two modes:



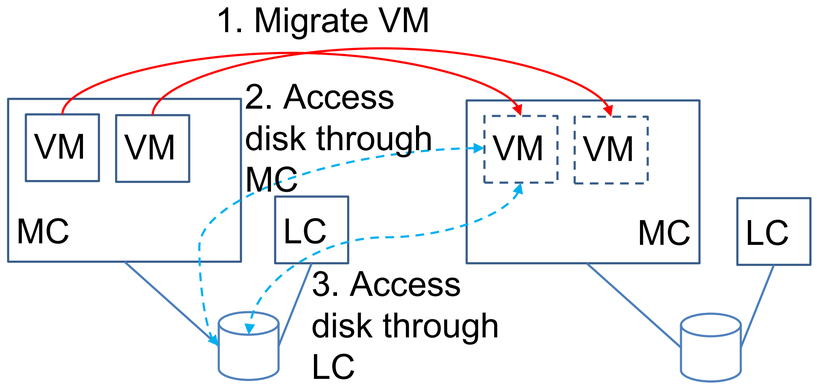
MC: the main computer
LC: the low power computer

DSwitch Design

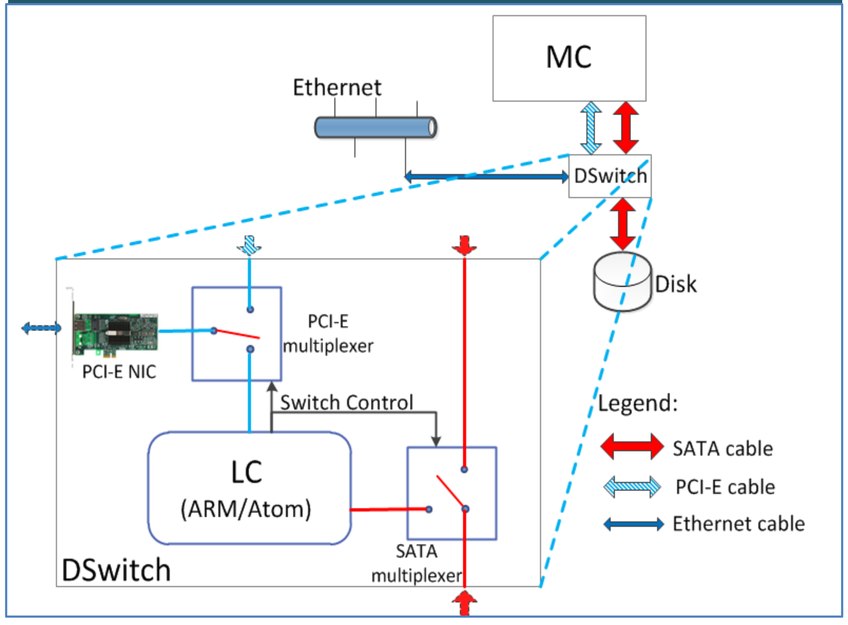
Switch with Service Migration



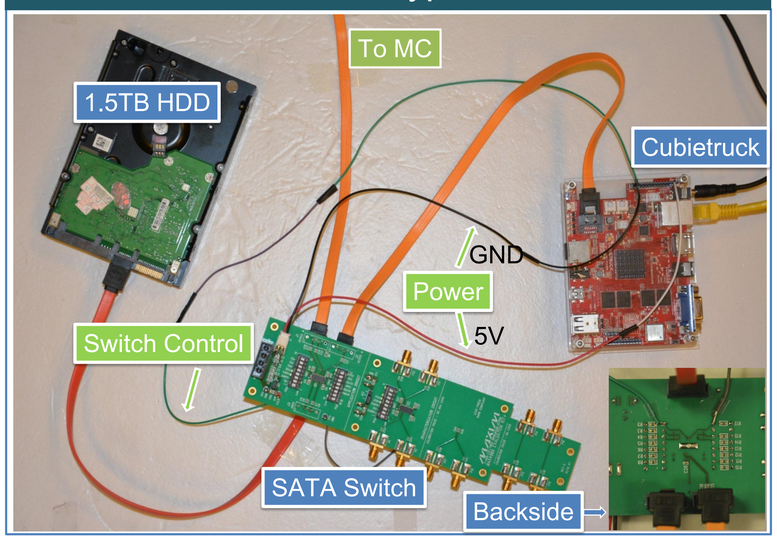
Switch with VM Migration



Alternative Design



Prototype



Performance

Power Consumption

- SATA multiplexer consumes around 0.45W.
- DSwitch working in the low power network attached mode can achieve a power saving of 91.9%-97.5%.

Power consumption of different machines

Computers (W)	Lightweight I/O	Idle	S3 (suspend to RAM)	Save
Workstation	218.7	211.2	3.3	97.5%
Game machine	116.3	107.5	2.4	95.9%
OptiPlex 9010	56.0	41.5	1.3	91.9%
Cubietruck	2.8	1.6	--	--

Reachability of Service Migration

Disruption of network and services

Events (second)	Min	Avg	Max
Enter S3	2.6	2.8	3.2
Resumes from S3	4.0	4.9	5.1
Network MC to LC	3.0	3.5	4.0
Network LC to MC	7.0	13.9	22.0
Service MC to LC	9.2	12.7	15.3
Service LC to MC	11.0	17.0	22.8

Interruption of VM Migration

From MC to LC: 5.36s
From LC to MC: 4.01s
The running of VMs is almost not interrupted.