

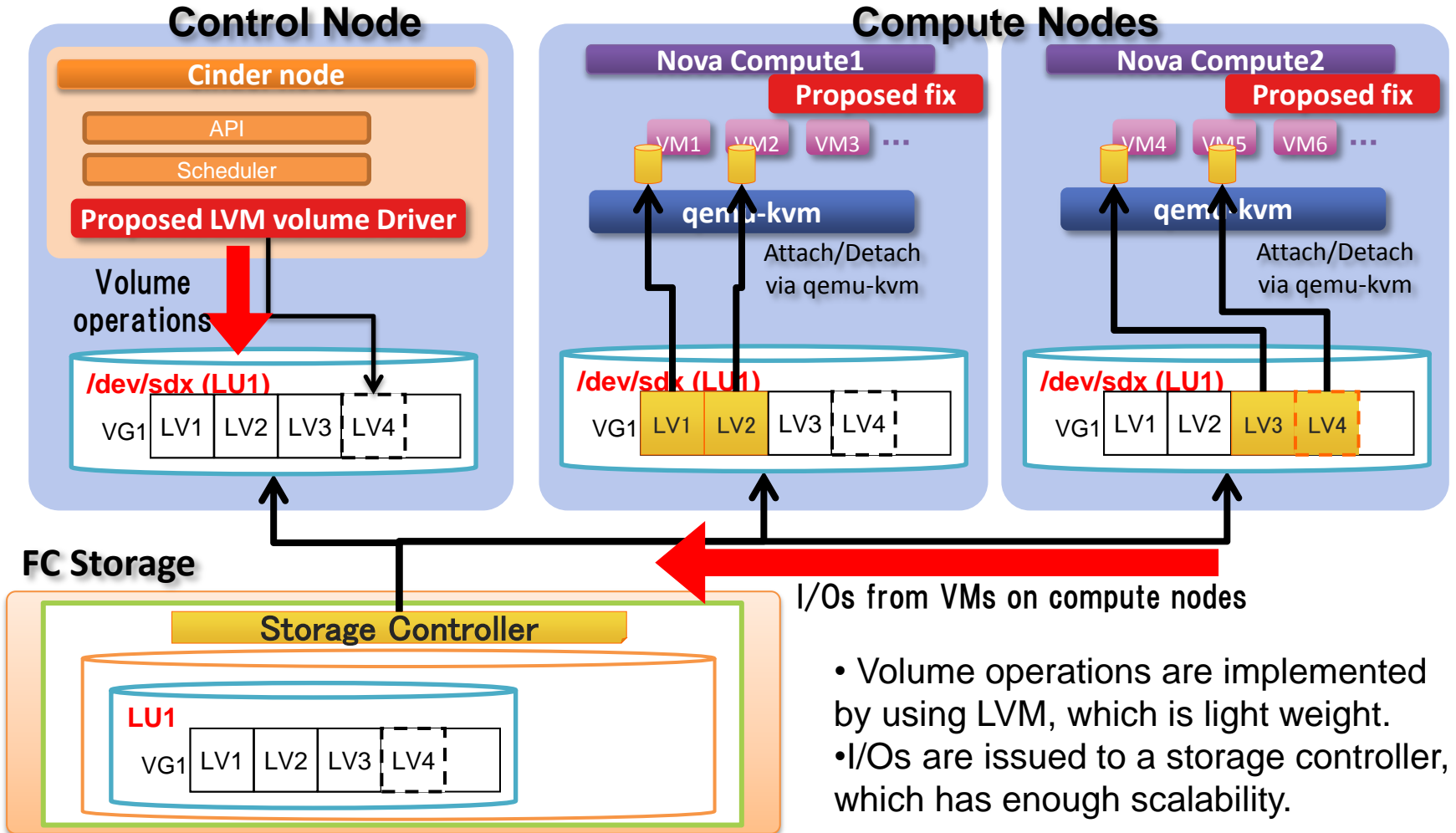
# **[Cinder] Support LVM on a shared LU**

---

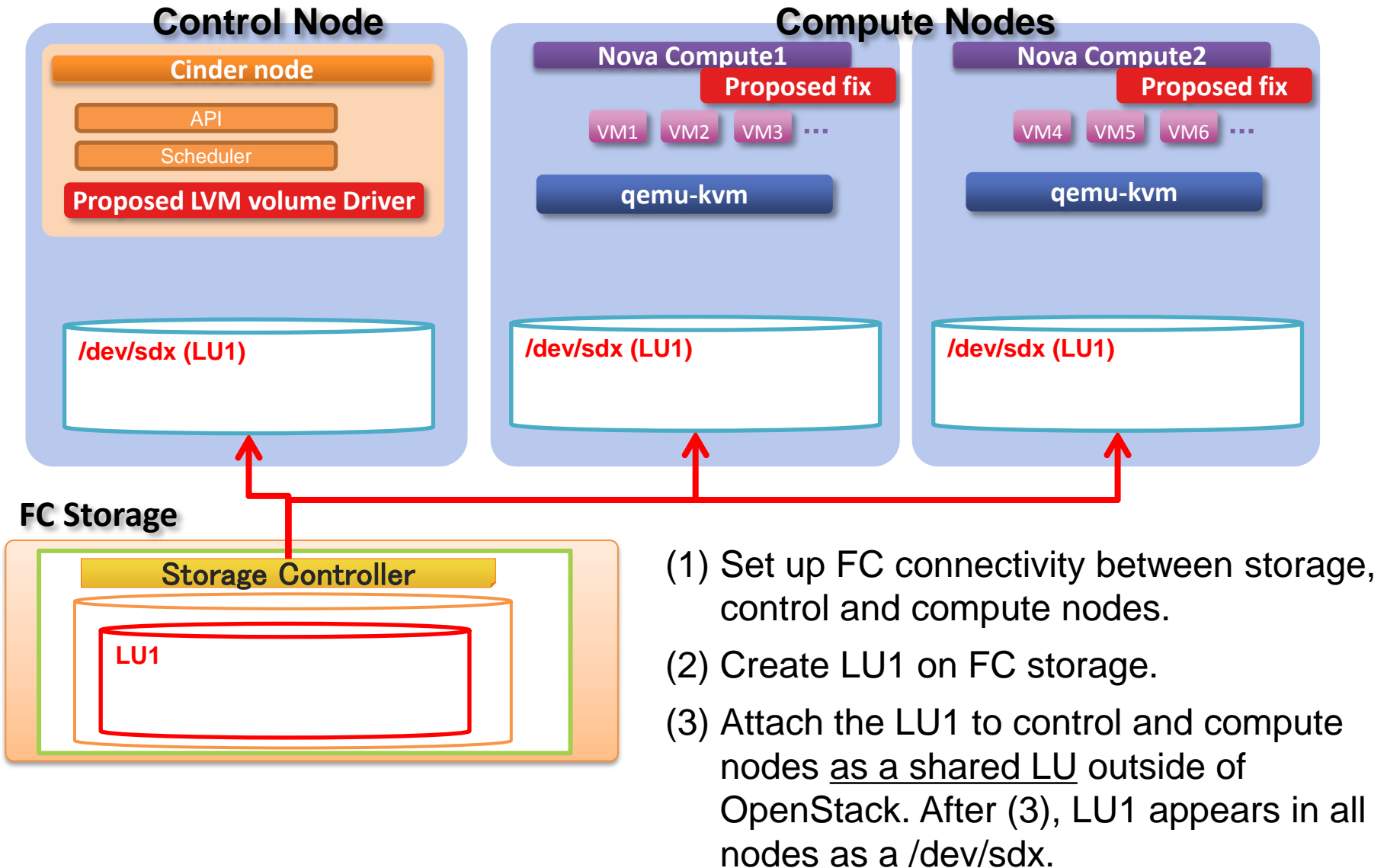
July/2/2014

Mitsuhiro Tanino <mitsuhiro.tanino@hds.com>

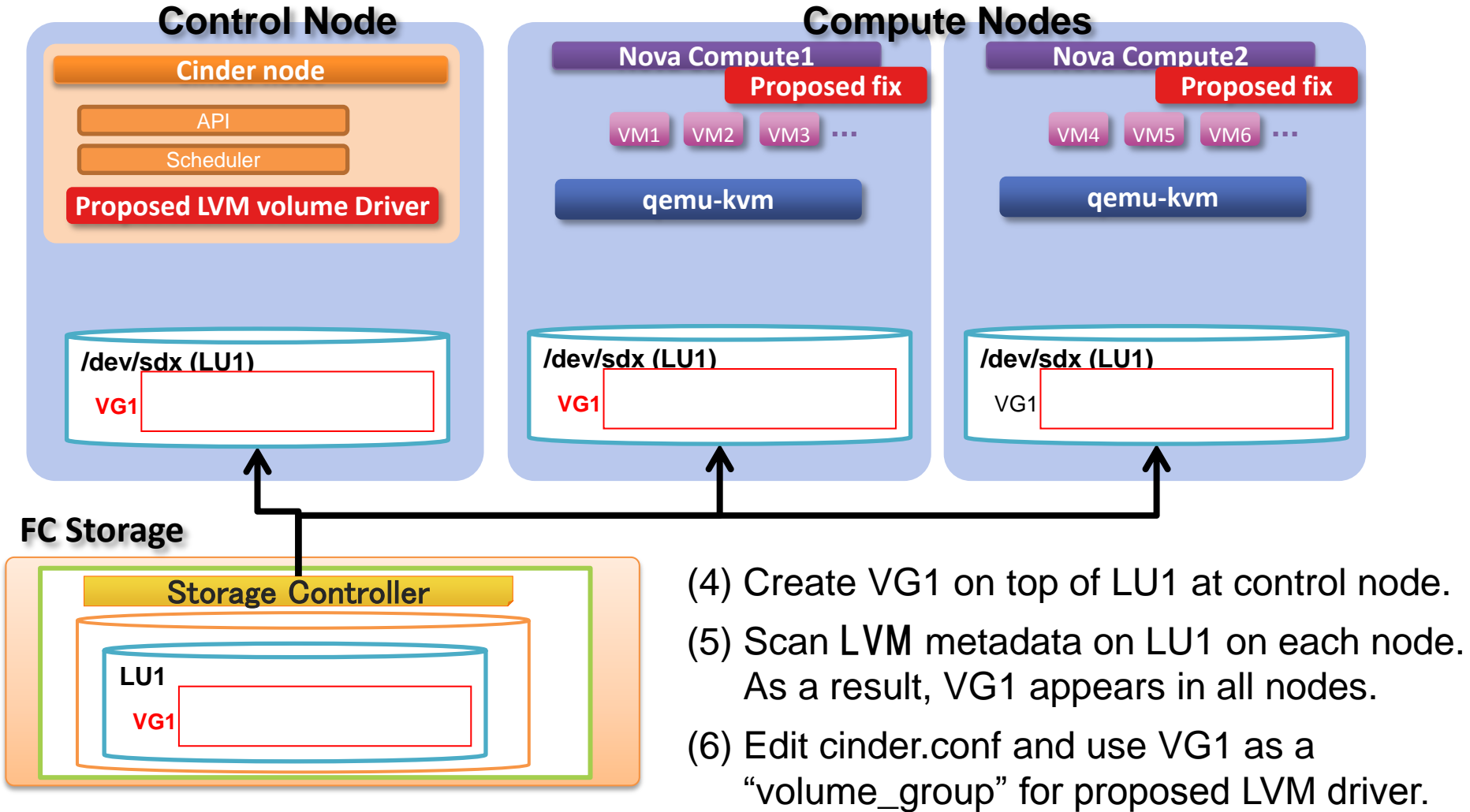
# 1-1. Big Picture



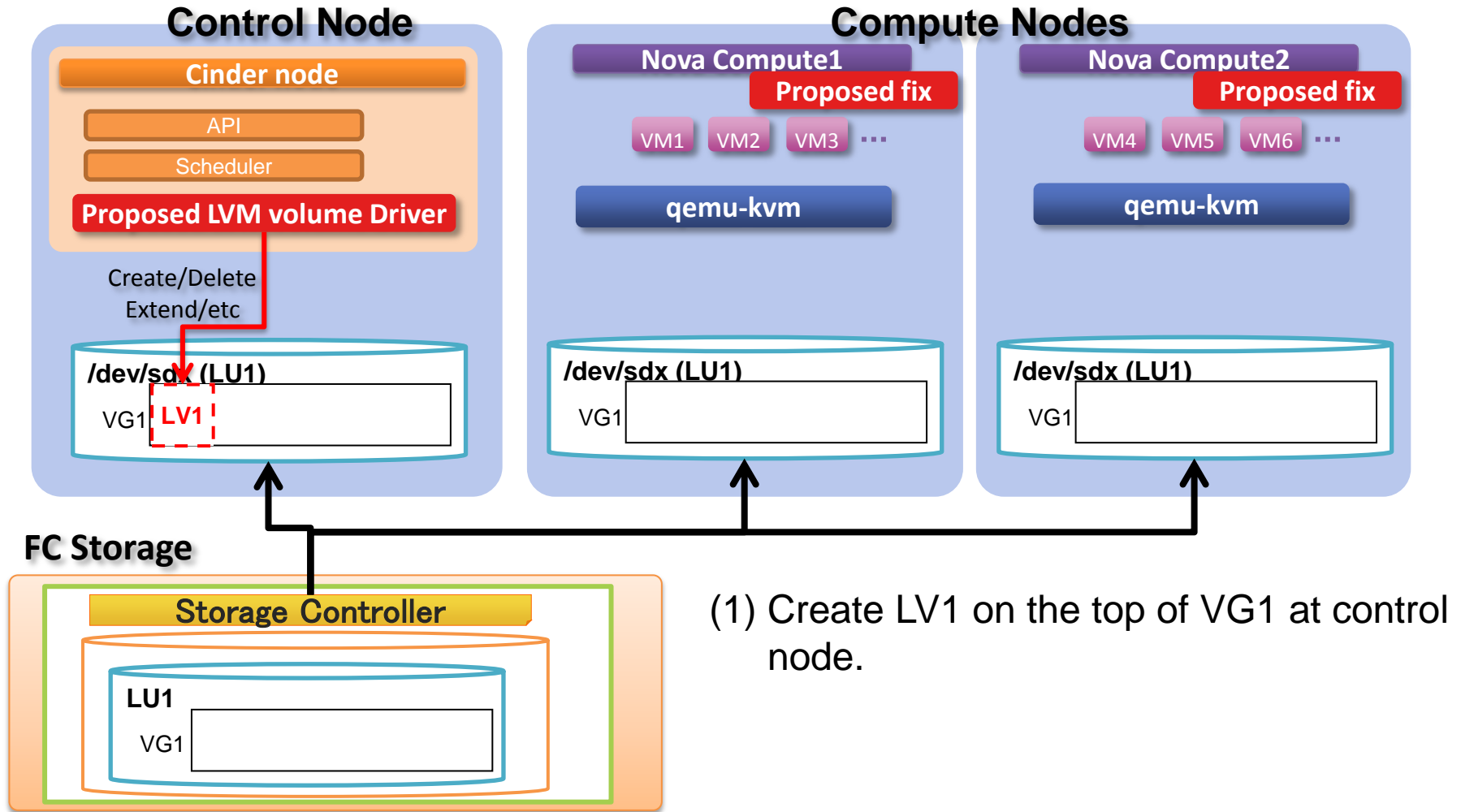
## 1-2. Preparation by administrator



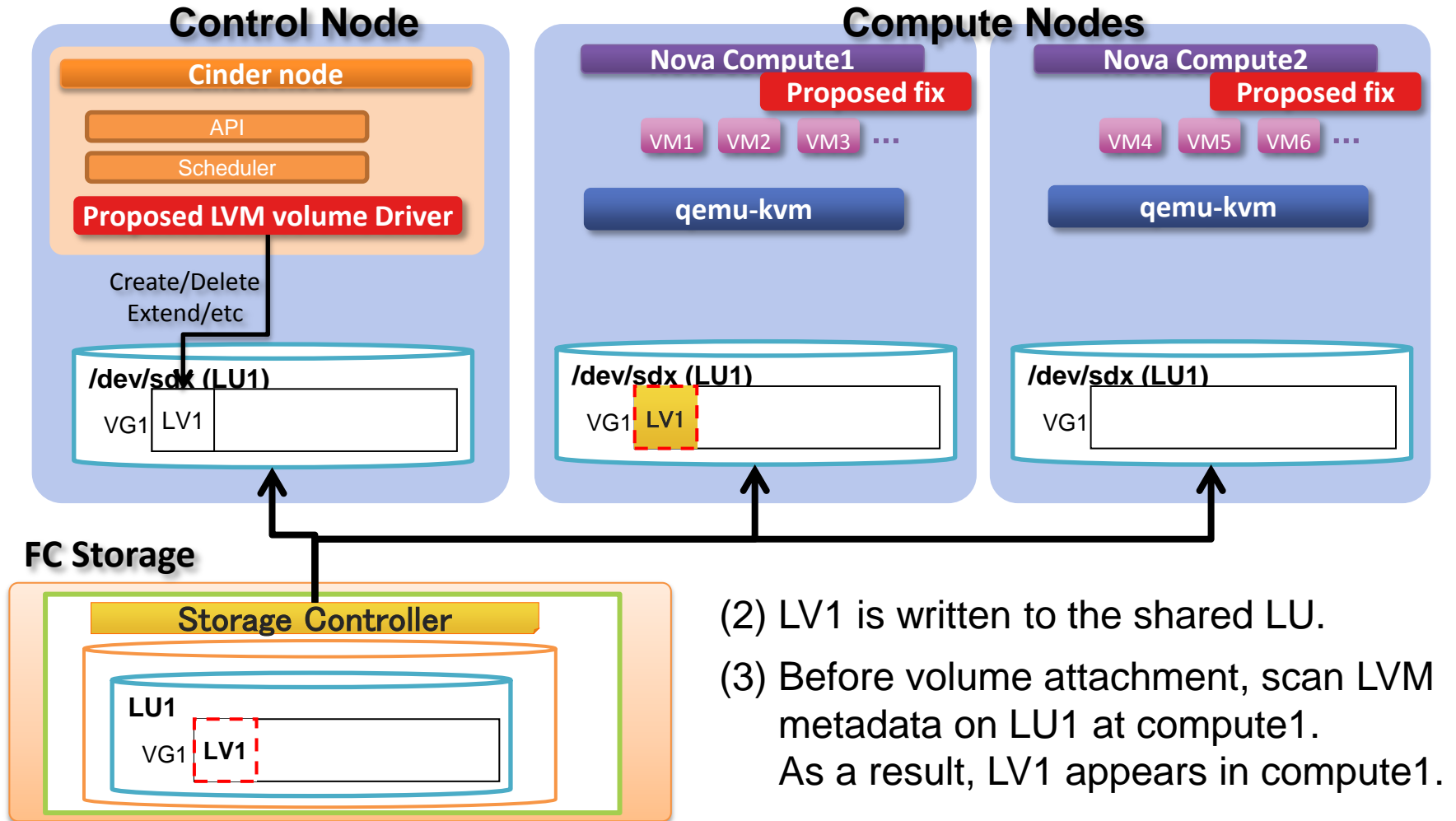
## 1-2. Preparation by administrator



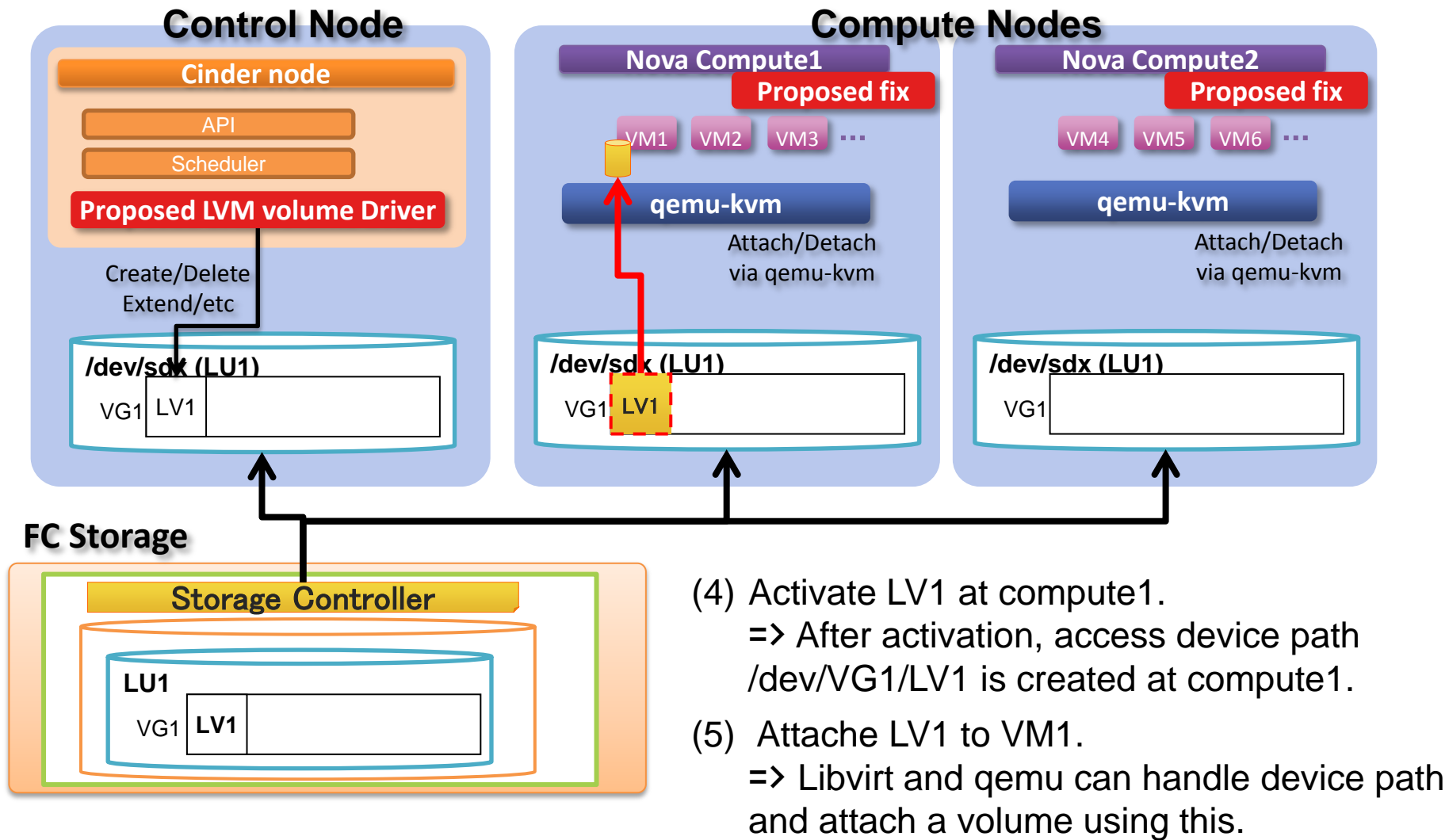
# 1-3. Work flow of volume creation and attachment



# 1-3. Work flow of volume creation and attachment

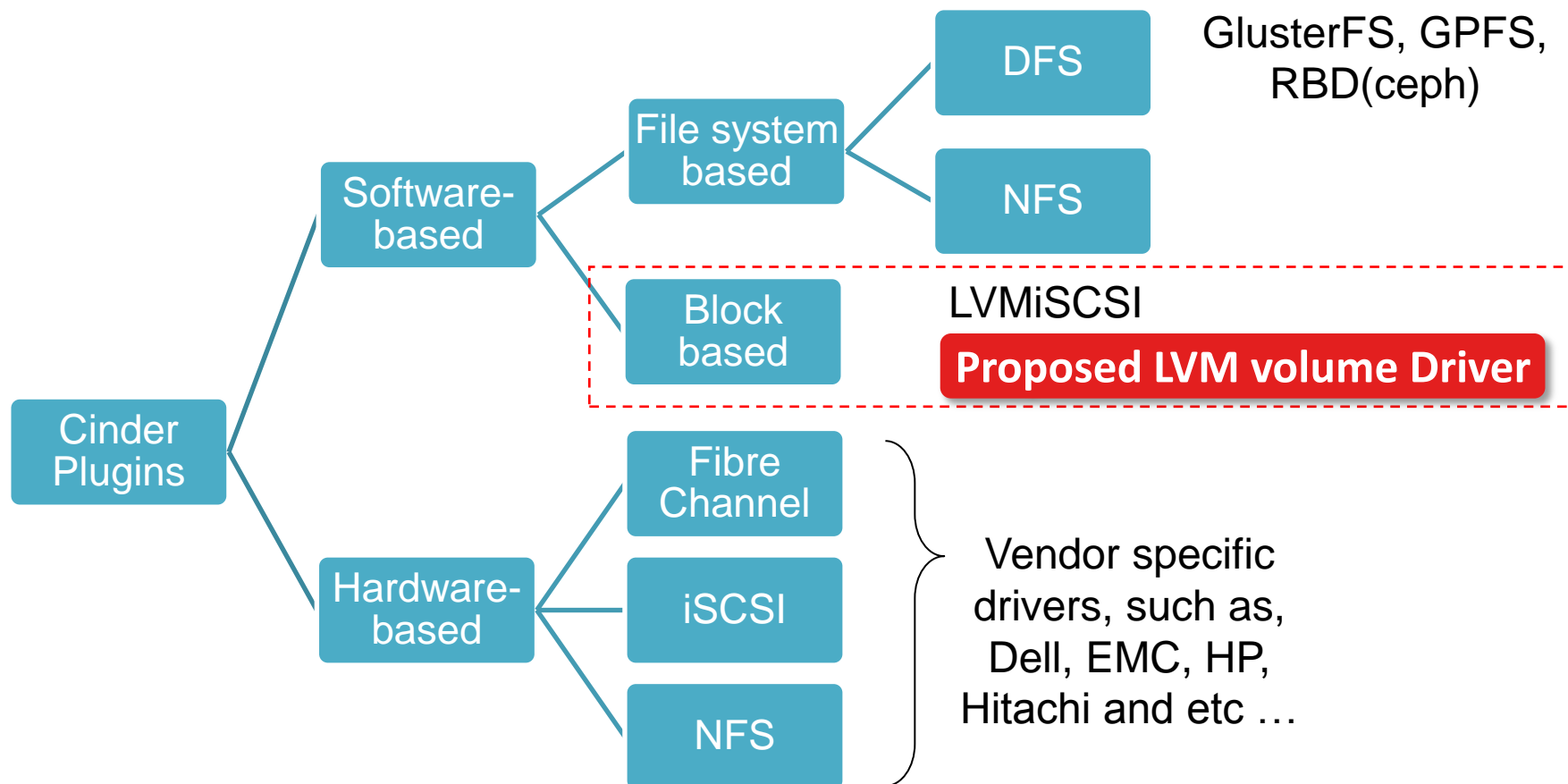


# 1-3. Work flow of volume creation and attachment



## 2. Target of Proposed LVM volume driver

There are many kinds of Cinder plugins available.





### 3. Benefit of Shared LVM driver

- Reduce hardware based storage workload by offloading the workload to software based volume operation.
  - ✓ *In order to use expensive storage more efficiently, I think it is better to reduce hardware based storage workload by offloading the workload to software based volume operation on a case by case basis.*
- Enable cinder to any kinds of shared storage volumes without specific cinder storage vendor driver.
  - ✓ *User can apply this driver to a storage which does not have specific cinder driver or does not have features such as efficient snapshot, backup.*

### 3. Benefit of Shared LVM driver

- Provide quicker volume creation and snapshot creation without storage workloads.
  - ✓ *LVM has lightweight volume creation. And also lightweight snapshot using copy on write.*
- Better I/O performance using direct volume access via Fibre channel.
  - ✓ *This driver can issue I/O without scsi target, therefore performance is better compared to LVMiSCSI driver.*
  - ✓ *I/O bandwidth and latency are almost same as raw FC volume. (See test result at P13 and P14)*

## 4. Comparison of Proposed LVM volume driver

	LVMiSCSI	Proposed LVM volume Driver	FC (Vendor Driver)	
Implementation of volume	LV (managed by LVM)	LV (managed by LVM)	LU (managed by storage)	
Volume Operation	By software (LVM)	By software (LVM)	By hardware (Storage)	Less volume operations to storage.
Work Load	Server side	Server side	Storage side	
Supported Storage	Any storage (Storage independent)	Any storage (Storage independent)	Specific storage (Requires specific plugin)	Better support coverage.
Volume Access Path	Via software iSCSI target	Direct from fibre channel	Direct from fibre channel	Better I/O performance.
Performance		Better I/O performance	Better I/O performance	
HA	Active/Passive: ○ Active/Active : —	Active/Passive: ○ Active/Active : —	Active/Passive: ○ Active/Active : ○	A/A is not supported.


## 4. Comparison of Proposed LVM volume driver

	LVMiSCSI	Proposed LVM volume Driver	FC (Vendor Driver)
Admin work - Initial setup	5 steps  (a) Setup network connectivity for all nodes. (b) OS and OpenStack setup (c) Create a storage volume and attach it to cinder node. (d) Create volume group on top of storage volume (e) Configure cinder.conf	5 steps  (a) Setup network connectivity for all nodes. (b) OS and OpenStack setup (c) Create a storage volume and attach it to cinder node. (d) Create volume group on top of storage volume (e) Configure cinder.conf	4 steps  (a) Set up FC connectivity between storage, cinder node and compute nodes. (b) OS and OpenStack setup (c) Create storage pool in the storage. (d) Configure cinder.conf

Volume group creation step is required.

## 4. Comparison of Proposed LVM volume driver

	LVMiSCSI	Proposed LVM volume Driver	FC (Vendor Driver)
Admin work - add new server	2 steps  (a) Setup network connectivity for new node. (b) OS and OpenStack setup	3 steps  (a) Set up FC connectivity between storage and new compute node (b) Set up an access control to a Shared LU from new node. (c) OS and OpenStack setup	2 steps  (a) Set up FC connectivity between storage and new compute node. (b) OS and OpenStack setup



Access control(FCZone or  
Shared LU) is required

## 5. Comparison of I/O performance

- Following results are I/O performance of issuing direct I/O from single instance to single volume.

Bandwidth of I/O [O\_DIRECT]

Unit: [KB/s]

	I/O size	LVMiSCSI	SharedLVM	Raw FC volume
Read	1M	49788.0	85146.0	88203.0
	512K	62382.0	160517.0	157810.0
	4K	4026.4	8630.8	8865.2
Write	1M	76691.0	141315.0	144173.0
	512K	59200.0	142408.0	144006.0
	4K	3870.1	7738.9	7867.1
Randread	1M	16152.0	17665.0	17105.0
	512K	13677.0	20397.0	19971.0
	4K	417.3	480.6	476.8
Randwrite	1M	15606.0	17067.0	16526.0
	512K	13666.0	20381.0	19955.0
	4K	417.1	4803.3	476.5

I/O performance of SharedLVM and Raw FC volume are almost same.

\*Larger is good.

## 5. Comparison of I/O performance

Average latency of I/O [O\_DIRECT]

Unit: [msec]

	I/O size	LVMiSCSI	SharedLVM	Raw FC volume
Read	1M	20.56	12.02	11.60
	512K	8.20	3.18	3.24
	4K	0.99	0.46	0.45
Write	1M	13.35	7.24	7.10
	512K	8.64	3.59	3.55
	4K	1.03	0.51	0.50
Randread	1M	50.47	52.31	54.22
	512K	28.75	21.58	22.13
	4K	8.38	7.76	7.84
Randwrite	1M	13.37	5.84	5.83
	512K	8.68	3.51	3.50
	4K	1.42	0.75	0.73

Average latency of SharedLVM and Raw FC volume are almost same.

\*Smaller is good.