

変える力を、ともに生み出す。

NTT DATAグループ



Live Migration support for OpenStack



2010年11月19日
株式会社NTTデータ

1. Live Migration

稼働中のVMインスタンスを停止させることなく他の物理マシンに移す手段
(Providing the scheme to migrate running VM instances from a physical machine to others)

1. **最小限のダウンタイム** (Almost no visible downtime)
2. **トランザクションロスなし** (No transaction loss)

2. Live Migration の必要性 (why Live Migration?)

1. メンテナンス(Maintenance)

ex. 物理マシンのアップグレードやパッチ

(Upgrade/installing the patches to hypervisors/BIOS.)

ex. RAID構成されたHDDの1ボリュームや、bondingされたNICの片系故障

(One of HDD volumes RAID / one of bonded NICs is out of order)

ex. 定期メンテナンス

(Regular period maintenance)

2. 負荷分散(Distributing high-load)

ex. 特定のVMインスタンスが過負荷になった場合、他VMを他物理マシンに逃がす

(a specific VM instance gets high load, other VM want to move)

3. 電力節約(Saving power)

物理マシンリソースが空き過ぎていた場合には一つにまとめる

(VM instances are too much scattered, move VM instances to a physical machine!)

3. 必要な機能 (inevitable functionalities)

1. **ライブマイグレーション**
(Live migration)
2. **特定の物理マシン上で稼働するVMインスタンス一覧の取得**
(Get a VM instances list running on the physical machine)
3. **特定の物理マシンの空きリソースの取得**
(Provide info which physical machine still has enough machine resource)

3. 必要な機能 (inevitable functionalities)

例 (Examples)

```
# nova-manage instance live_migration instance-1111 compute-node2  
instance-id destination
```

```
# nova-manage describe instances  
INSTANCE i-45610761 0.0.0.0 0.0.0.0 pending 0 c1.medium 2010-04-01T04:44:10.774Z  
cluster1 eki-30D00D36 eri-8FC50F4E compute-node1  
INSTANCE i-45610761 emi-18240C94 0.0.0.0 0.0.0.0 pending 0 c1.medium 2010-04-  
01T04:44:10.774Z cluster1 eki-30D00D36 eri-8FC50F4E compute-node2
```

誰か既に作ったらしく新規機能追加の必要なし(no need to develop)

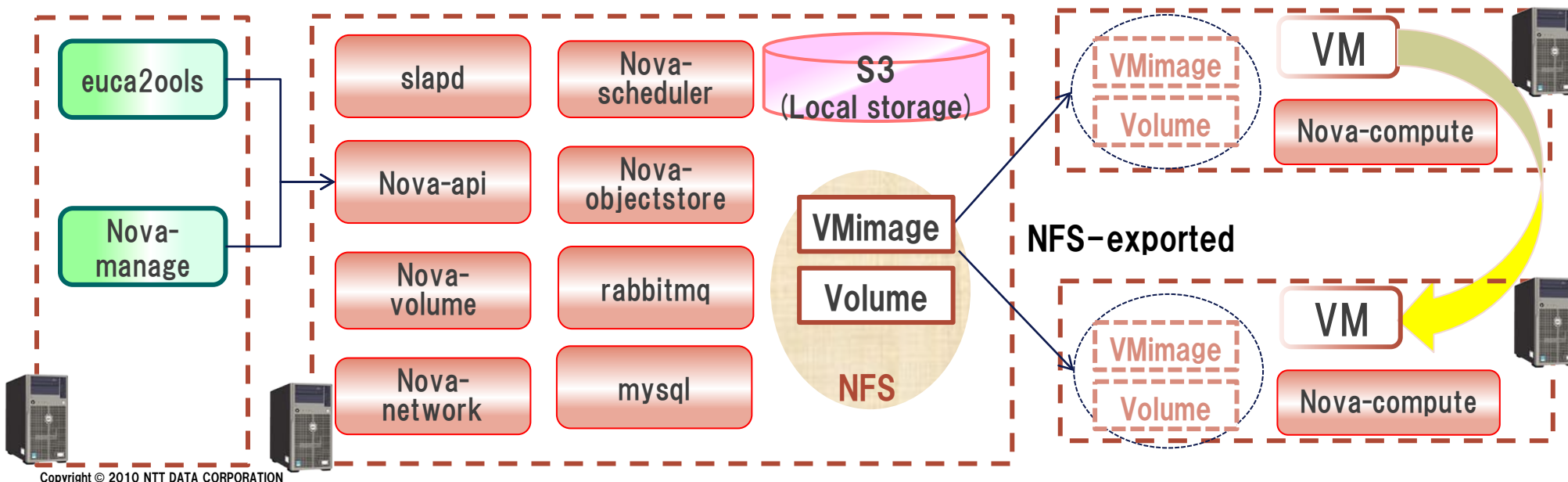
```
# nova-manage host show host1 ( no hostname given, getting info from all physical machine)  
TOTAL host1 10 20480 1000 (total resource per physical machine)  
NODE host1 project1 5 10240 100 (usage per project)  
NODE host1 project2 2 5012 30  
(nodename - projectname - total cpu usage - total memory usage - total HDD usage )
```

5. 前提条件と想定条件 (pre-requisites and assumptions)

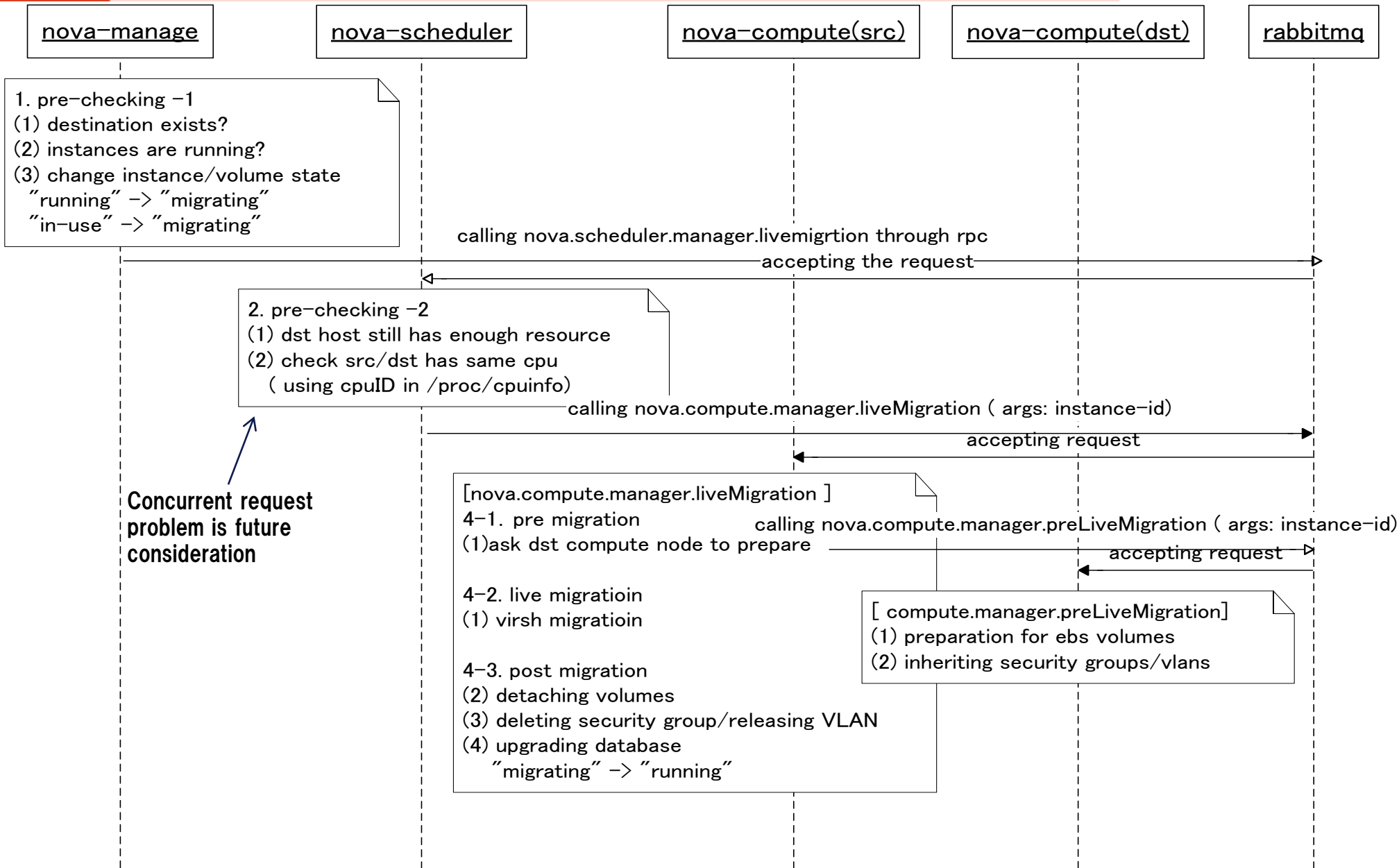
	Explanation
OS	Ubuntu Lucid 10.04 (both physical machine and VM)
Hypervisor	KVM (other hypervisors support may be discussed later)
Storages	共有ストレージを前提とする (The directory which instances are running and creating volumes must be a part of shared storage (using NFS))
Authentication	管理者のみ実行可能(nova-manageの拡張とし、API組込は後リリースで) Admin only (meaning nova-manage extension. Extending to APIs should be discussed later)
Network connectivity	ライブマイグレーション対象の二つの物理マシンは同セグメント Source /dest physical machines must belong to same availability zone (segment) .
EBS volume	ライブマイグレーション前後でEBSボリュームをマウントし続ける VM instances must continue mounting volumes.
Floating/ Fixed_ip	ライブマイグレーション前後でFloating/Fixed は変わらない VM instances must use same fix/floating ips.
VLAN	VM instances must belong same VLAN.
Security group	VM instances must use same security group including filtering rules.

5. 前提条件 (pre-requisites and assumptions)

開発環境(Development environment)



6. デザイン- Live Migration処理フロー (Design for execution of Live Migration)



- APIへの組み込み
- 他ハイパーバイザのサポート(Other hypervisors support)
- VPN 対応(Using VPN considerations)
- Block Migration (共有ストレージなしのライブマイグレーション)
(a live migration that shared storage is not necessary.)

詳細はBexarリリース終了後の
design summitで議論
(Detail discussion should be
started the design summit after
Bexar release)

Bexar Design Summitにおける議論

8. Design summitでの議論 (discussions at bexar design summit)

1. 誰が実行すべき機能か(who should execute Live Migration?)

さまざまなユースケースを想定した場合のロールの割当て (role assign considering to various use cases)

 **Bexar Releaseはnova-manageへ実装
(implement to nova-manage at Bexer Release)**

2. 誰が物理ノードの空リソースを把握するか？(who knows vacant resource?)

 **Scheduler !**

3. Concurrency request

管理者が空リソースを確認し、ライブマイグレーションを実行する迄に、空リソースが埋まる状況の想定 (Considerations requests between admin checking vacant resource and executing live migration)

 **Bexerでは直前で再チェックはするが、抜本的な対策は後リリースで検討
(re-checking just before executing live migration, then discuss later)**

And many people started to discuss “what this feature should be?” …… .. 10

9. Design summitでの結論

(conclusion of discussions at bexar design summit)

1. ライブマイグレーションが必要であることは全員賛成

(Everyone agreed with implementing live migration feature to OpenStack)

2. Bexerリリースまでは基本機能をシンプルに実装すること

(“**Implementing the basic functionality at the simplest way**” is our target for Bexer release.)

10. その後のステータス



OpenStack Compute (Nova)

Overview Code Bugs **Blueprints** Translations Answers

Blueprints for OpenStack Compute (nova)

OpenStack Compute (nova) » Blueprints »

Show only blueprints containing: Search blueprints

1 → 75 of 107 results

First • Previous • **Next** ▶ • Last

Priority	Blueprint	Design	Delivery	Assignee	Series
Essential	xenapi-raw-disk-images	Approved	Not started	Ewan Mellor	bexar
High	audit-logging ⓘ	Approved	Beta Available	Todd Willey	bexar
High	austin-microsoft-hyper-v-support ⓘ	Approved	Beta Available	Chiradeep Vittal	bexar
High	bexar-shared-ip-groups	Approved	Not started	Eric Day	bexar
High	ipv6-support	Approved	Unknown	NTT PF Lab.	bexar
High	rescue-mode ⓘ	Approved	Beta Available	Anso Labs	bexar
High	resize-servers	Approved	Unknown	Anso Labs	bexar
High	snapshot-instance	Approved	Unknown	Anso Labs	bexar
High	web-based-serial-console ⓘ	Approved	Started	Anso Labs	bexar
High	bexar-distributed-scheduler	Pending Approval	Unknown	Eric Day	bexar
High	bexar-openstack-api-floating-ips	Pending Approval	Not started		bexar
Medium	bexar-migration-live	Approved	Started	NTT DATA	bexar
Medium	guest-agent ⓘ	Approved	Started	Anso Labs	bexar

Blueprint status is changed to “approved”!

年明けすぐのfeature freezeに向けて奮闘中！
Our engineers are now struggling to develop targeting to next feature freeze!

変える力を、ともに生み出す。

NTT DATAグループ

