

2013年 11月26-27日 遺伝研

**新学術「ゲノム支援」
平成25年度 情報解析講習会
Linuxサーバの構築**

東京工業大学 大学院生命理工学研究科
森 宙史

Hiroshi Mori

hmori@bio.titech.ac.jp

なぜ自分でサーバを構築する必要があるのか？

Q. サーバとパソコンは何が違うのか？

A. 通常、サーバは複数のコンピュータからネットワーク経由でアクセスされ、何らかのサービスを提供するコンピュータのことを言う。
24時間365日稼働を前提として信頼性の高い部品が使われていたりする

Q. 自分で構築しなくてもDDBJスーパーコンピュータ使えば良いのでは？

A. はい。待ち時間が気にならないのであれば。

Q. 色々な企業のクラウドサービスを使えば良いのでは？

A. ゲノム研究の情報解析で行う特殊な計算に対応できるサービスがあるのなら。

この1時間で扱う内容

ゲノム研究における情報解析を行えるサーバを構築するために、

[1] 目的に応じたハードウェアのスペックの目安

[2] OSの選択

[3] Linuxディストリビューションの選択

[4] 実際のLinux OSのインストール&初期設定手順

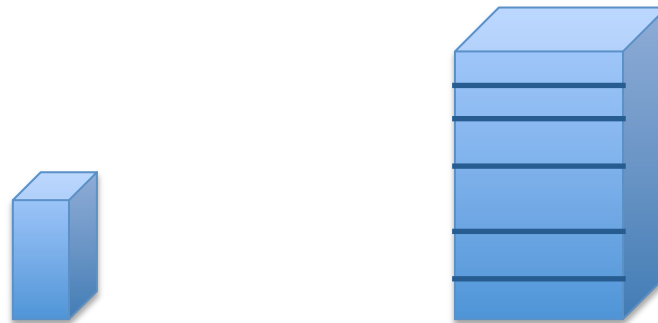
について解説する。

この1時間は、手持ちのPCを用いた実習はほとんど行いません

[1] 目的に応じたハードウェアのスペックの目安

(1) サーバの物理的なタイプ

- **タワー型** 通常のデスクトップ型PCの大きいもの。拡張性が高い。
- **ラック型** ラック内に縦にサーバを積み上げる。ある程度の数のサーバがあればタワー型より省スペース。縦に巨大(フルラックなら約2m)。
- **ブレード型** 電源やファンをサーバと分離して高密度にサーバを搭載可能。
- **その他**



[1] 目的に応じたハードウェアのスペックの目安

(2) サーバの機種の違い

サーバを選ぶ際には、

- ・ ディスクの種類と最大搭載数
- ・ メモリの種類と最大搭載数
- ・ CPUの種類と最大搭載数
- ・ 電源の容量

などに特に注意すべき。

しかし、お金は無限にある訳ではない...

[1] 目的に応じたハードウェアのスペックの目安

1. Webサーバ

研究プロジェクトのWebページを公開、論文のSupplementsを公開

メモリ数GB、1CPU数コア、ディスク数百GBで十分 (な場合が多い)

2. DBサーバ

研究で得られたデータを何らかのDB管理システムに格納してWebで公開、検索可能にする

メモリ数十GB(32GBくらいまで)、1CPU 数コア、ディスク数TBあれば大丈夫

[1] 目的に応じたハードウェアのスペックの目安

3. 原核・真核・メタゲノム・Transcriptomeアセンブル

特に重要なハードウェア: メモリ・CPU

- ・ 原核生物のゲノムアセンブル 数十GBメモリ (50GB あれば十分)、1CPU 数コア
- ・ Transcriptomeアセンブル 今のHiSeqのリード1億pairsなら、120GBあれば十分
(例えばTrinityなら、100万pairsにつき1GBメモリが目安)
(<http://trinityrnaseq.sourceforge.net/>)
- ・ メタゲノムのアセンブル 組成が単純な群集なら原核生物のアセンブルと同様
群集が多様な場合、大量のリードが必要であり、
数百GB-数TBメモリが必要な場合もある。 十数コア以上
- ・ 真核生物のゲノムアセンブル ゲノムサイズ、どれくらいheteroか、等に依存する
数十GB-数TBメモリ、十数コア以上

少なくとも100GB以上のメモリをのせないで、アセンブルは辛い

[1] 目的に応じたハードウェアのスペックの目安

4. 数千本以上の配列相同性検索

特に重要なハードウェア: CPU

- ・ ゲノムやメタゲノムの遺伝子アノテーション
- ・ 複数ゲノムの比較解析

入力配列が数千本以上になることが多く、計算を並列化して高速化する必要がある。

現在のnr相手のBLASTXやBLASTPは、並列化しても各プロセスで5GBほどメモリを使用するので、メモリもそれなりに必要。

CPU数、コア数が非常に重要、メモリもある程度必要

[1] 目的に応じたハードウェアのスペックの目安

5. マッピングツールでのReferenceゲノムへのマッピング

特に重要なハードウェア: ディスク

- Resequencing, RNA-Seq, ChIP-Seq解析におけるマッピング

リードのゲノムへのマッピングは高速でメモリ使用量も少ない。

結果のSAMファイルやBAMファイルが数十から数百GBになったりする。
samtools等でsortしたりすると、その規模のファイルが何個もできる。

マッピングを頻繁にするのなら、ディスクは少なくとも十数TBは必要

[1] 目的に応じたハードウェアのスペックの目安

3. 原核・真核・メタゲノム・Transcriptomeアセンブル

特に重要なハードウェア: メモリ・CPU

4. 数千本以上の配列相同性検索

特に重要なハードウェア: CPU

5. マッピングツールでのReferenceゲノムへのマッピング

特に重要なハードウェア: ディスク

[2] サーバのOSの選択

マイナスポイント

Windows Server

- ・ 様々なバイオインフォマティクスのソフトウェアが動かない
- ・ Virtual OSを使う手もあるが...
- ・ GUIでサーバの設定が可能

Mac OSX Server

- ・ バイオインフォマティクスのソフトウェアが結構対応している
- ・ Linuxコマンドが標準で使える
- ・ GUIでサーバの設定が可能
- ・ ユーザコミュニティが小さい

Linux

- ・ Linux専用のバイオインフォマティクスのソフトウェアも多数存在
- ・ GUIである程度サーバの設定が可能だが...
- ・ スーパーコンピュータのほとんどのOSもLinux

[3] Linuxディストリビューションの選択

Linuxディストリビューション

Linuxの本体であるカーネルと、独自のGUIや様々なアプリケーションをまとめて一つの塊(パッケージ)にし、容易に再配布出来るようにしたもの

- Debian系

Debian, Ubuntu, etc.

- Red Hat系

Red Hat Enterprise Linux, Fedora, CentOS, etc.

- Slackware系

openSUSE, etc.

- その他

Gentoo Linux, etc.

[4] CentOSのインストール & 初期設定

CentOS 6.4をマシンにインストールする際の手順

物理的に必要なもの

- CentOSをインストールする対象の、DVD-Rを読み込めるマシン
- DVD-R 1枚
- DVD-Rを書き込めるPC

1. CentOSをダウンロード

作業: Webブラウザで <https://www.centos.org/> にアクセス

artwork.) CentOS is free.

CentOS is developed by a small but growing team of core developers. In turn the core developers are supported by an active user community including system administrators, network administrators, enterprise users, managers, core Linux contributors and Linux enthusiasts from around the world.

CentOS has numerous advantages over some of the other clone projects including: an active and growing user community, quickly rebuilt, tested, and QA'ed errata packages, an extensive [mirror network](#), developers who are contactable and responsive, multiple free support avenues including [IRC Chat](#), [Mailing Lists](#), [Forums](#), a dynamic [FAQ](#).

This site makes use of Javascript. If your browser doesn't work with Javascript we recommend you use [Firefox](#).

CentOS 6 Releases

March 12th 2013

The CentOS team is pleased to [announce](#) the immediate availability of CentOS 6.4 for i386 and x86_64 architectures.

CentOS 6.4 is based on the upstream release EL 6.4 and includes packages from all variants. All upstream repositories have been combined into one, to make it easier for end users to work with.

There are some very important changes to this release compared with the previous versions of CentOS and we highly recommend reading this announcement along with the [Release Notes](#). Especially take a look at the "Known Issues" section.

There is also a minimal install CD that will get you a very small base install that you can add to.

And now: Have fun.

Release Notes: [CentOS](#)

Download: [i386](#) | [x86_64](#)

come together and talk about things that they care about most, and to share stories from their experiences working with CentOS in various scenarios.

Please see the [Wiki](#) for the latest information on the Austin Dojo. We would like to thank our Event Sponsors [Calxeda](#) and [Rackspace](#) for this event.

[Register Here](#)

CentOS Sponsor

The CentOS Project would like to thank the following product or service vendor for being a [CentOS sponsor](#):



(clicking the graphic opens a new window to the sponsor's site)

If you would like to help the CentOS project then details can be found on the [donation page](#)

Sponsors Links

The CentOS Project would like to thank the following sponsors of the project.

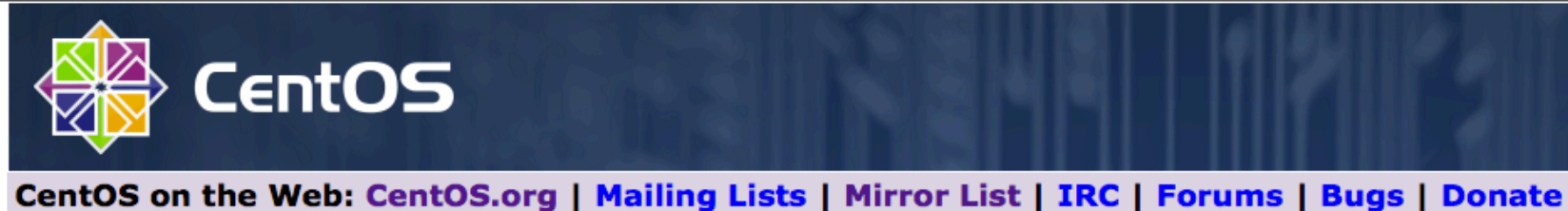
UK Dedicated Server Hosting

Website Backups by Comcure

If you would like to help the CentOS project by sponsoring a link then please contact us using the [web contact form](#)

1. CentOSをダウンロード

作業: ダウンロード元として、距離的に近いミラーサーバを選ぶ



In order to conserve the limited bandwidth available .iso images are not downloadable from mirror.centos.org

The following mirrors should have the ISO images available:

Actual Country -

http://ftp.jaist.ac.jp/pub/Linux/CentOS/6.4/isos/x86_64/
http://ftp.tsukuba.wide.ad.jp/Linux/centos/6.4/isos/x86_64/
http://ftp.ijj.ad.jp/pub/linux/centos/6.4/isos/x86_64/
http://www.ftp.ne.jp/Linux/packages/CentOS/6.4/isos/x86_64/
http://ftp.riken.jp/Linux/centos/6.4/isos/x86_64/
http://mirror.fairway.ne.jp/centos/6.4/isos/x86_64/
http://ftp.nara.wide.ad.jp/pub/Linux/centos/6.4/isos/x86_64/








Nearby Countries -

http://mirror.nus.edu.sg/centos/6.4/isos/x86_64/
http://mirrors.sin3.sg.voxel.net/centos/6.4/isos/x86_64/
http://mirror.bit.edu.cn/centos/6.4/isos/x86_64/
http://mirror.esocc.com/centos/6.4/isos/x86_64/
http://mirrors.163.com/centos/6.4/isos/x86_64/
http://mirrors.hust.edu.cn/centos/6.4/isos/x86_64/
http://mirrors.hustunique.com/centos/6.4/isos/x86_64/
http://centos.ustc.edu.cn/centos/6.4/isos/x86_64/
http://mirrors.btte.net/centos/6.4/isos/x86_64/
http://mirrors.stuhome.net/centos/6.4/isos/x86_64/

1. CentOSをダウンロード

ここから先は、紙芝居: *x86_64-bin-DVD1.iso をダウンロードする

Index of /Linux/centos/6.4/isos/x86_64

Name	Last modified	Size	Description
 Parent Directory		-	
 0_README.txt	2013-04-05 20:09	1.4K	
 CentOS-6.4-x86_64-LiveCD.iso	2013-04-23 21:59	702M	
 CentOS-6.4-x86_64-LiveCD.torrent	2013-05-23 01:54	28K	
 CentOS-6.4-x86_64-LiveDVD.iso	2013-04-23 22:03	1.7G	
 CentOS-6.4-x86_64-LiveDVD.torrent	2013-05-23 01:54	68K	
 CentOS-6.4-x86_64-bin-DVD1.iso	2013-03-06 03:44	4.1G	
 CentOS-6.4-x86_64-bin-DVD1to2.torrent	2013-03-09 06:15	217K	
 CentOS-6.4-x86_64-bin-DVD2.iso	2013-03-06 03:44	1.4G	
 CentOS-6.4-x86_64-minimal.iso	2013-03-06 03:45	342M	
 CentOS-6.4-x86_64-netinstall.iso	2013-03-05 10:05	230M	
 md5sum.txt	2013-05-16 22:36	388	
 md5sum.txt.asc	2013-05-16 23:43	1.2K	
 sha1sum.txt	2013-05-16 22:39	436	
 sha1sum.txt.asc	2013-05-16 23:43	1.3K	
 sha256sum.txt	2013-05-16 22:44	580	
 sha256sum.txt.asc	2013-05-16 23:43	1.4K	

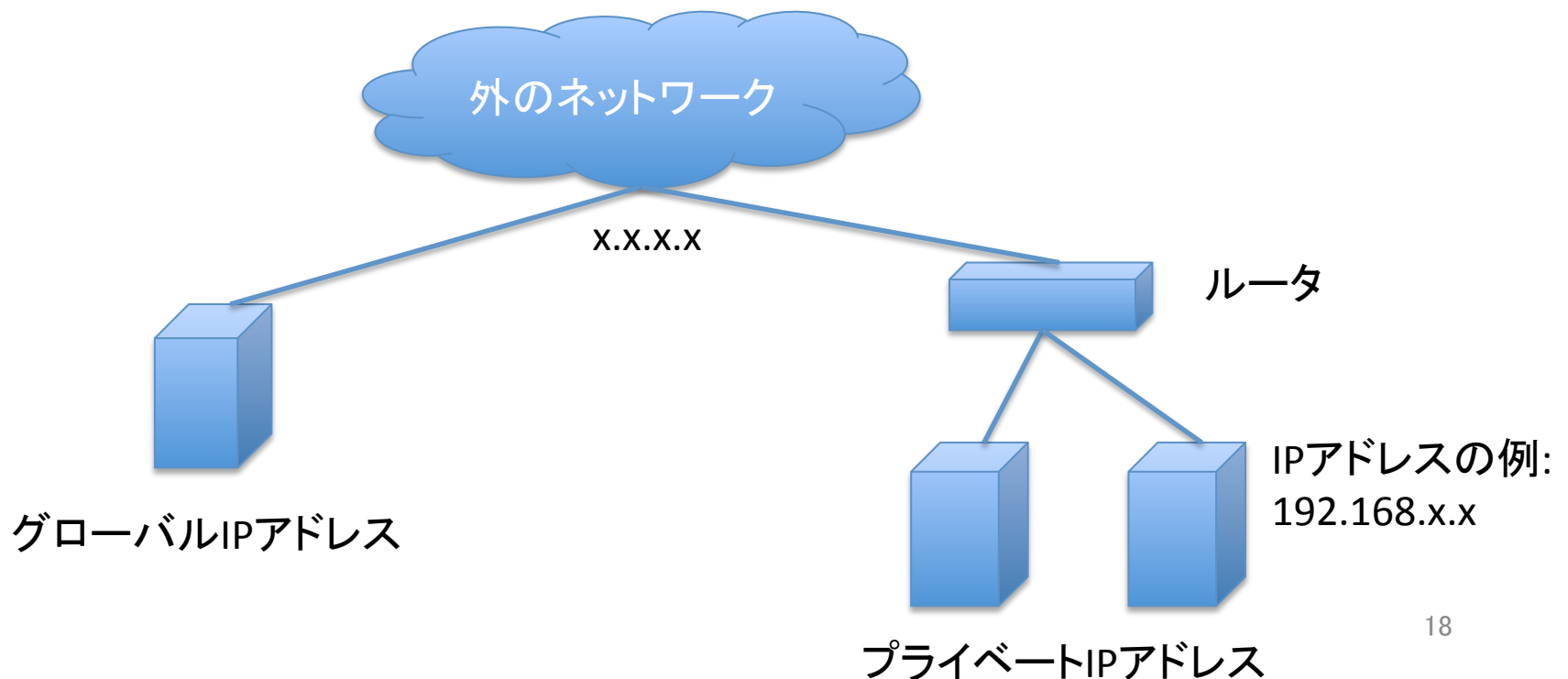
2. CentOS DVDを作成する

*x86_64-bin-DVD1.iso のダウンロードが終わったら、isoファイルと空のDVD-Rを基に、CentOSのインストール用DVDを作成する。

通常のDVD作成ソフトウェアを用いて作成でき、例えばMacならディスクユーティリティ等のソフトウェアで作成可能。

3. ネットワークの設計

どのようなネットワークにサーバを接続するかによって、事務手続きに時間がかかる場合もあり、セキュリティ設定も変わる。サーバにCentOSをインストールする前に、サーバをどのようなネットワークに接続するか、考えておく。具体的には、IPアドレス、ホスト名、どのようなプロトコルを使うのかなど

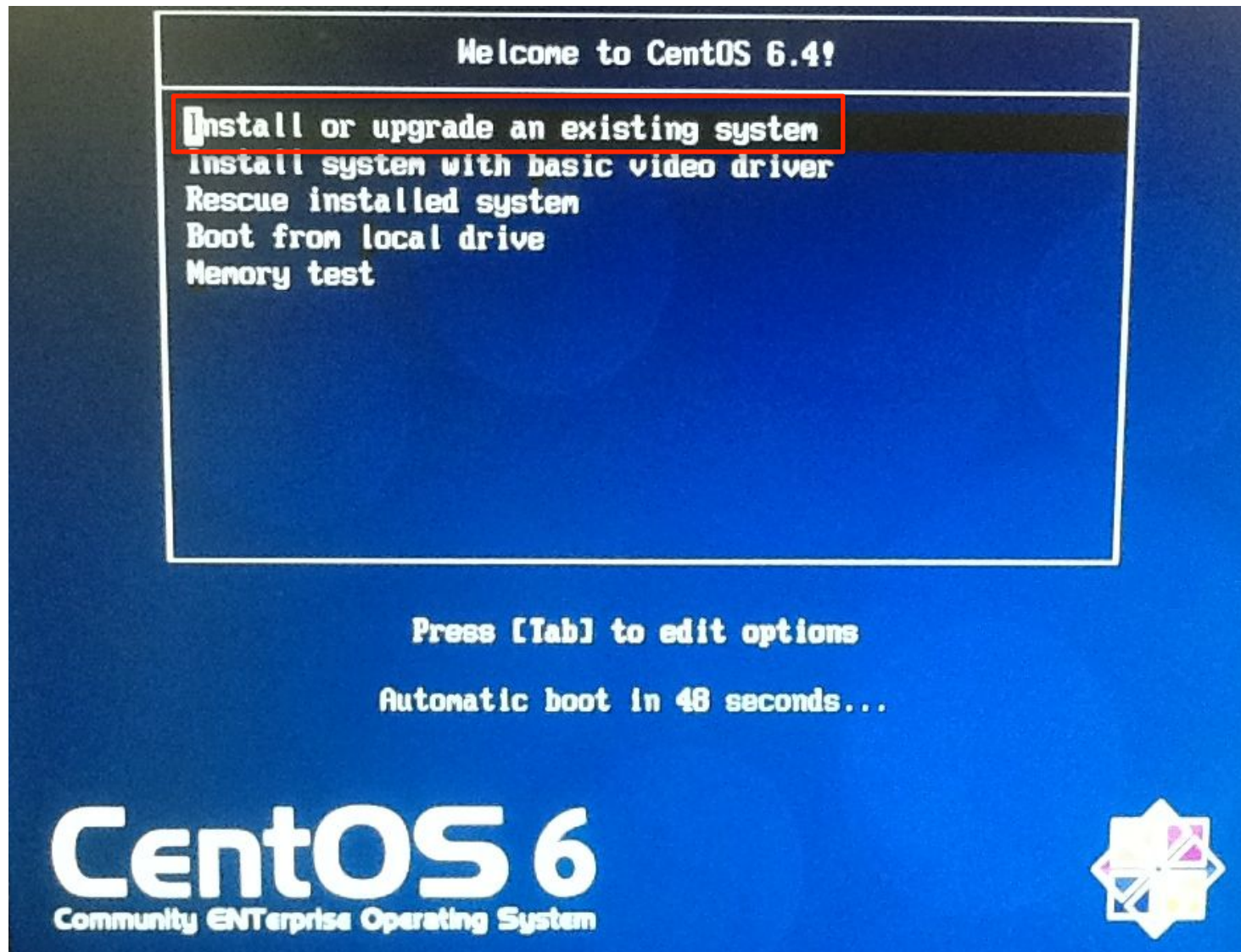


4. BIOS設定でDVDからOSを起動するように設定

この設定方法はサーバのメーカー・機種によってバラバラ。

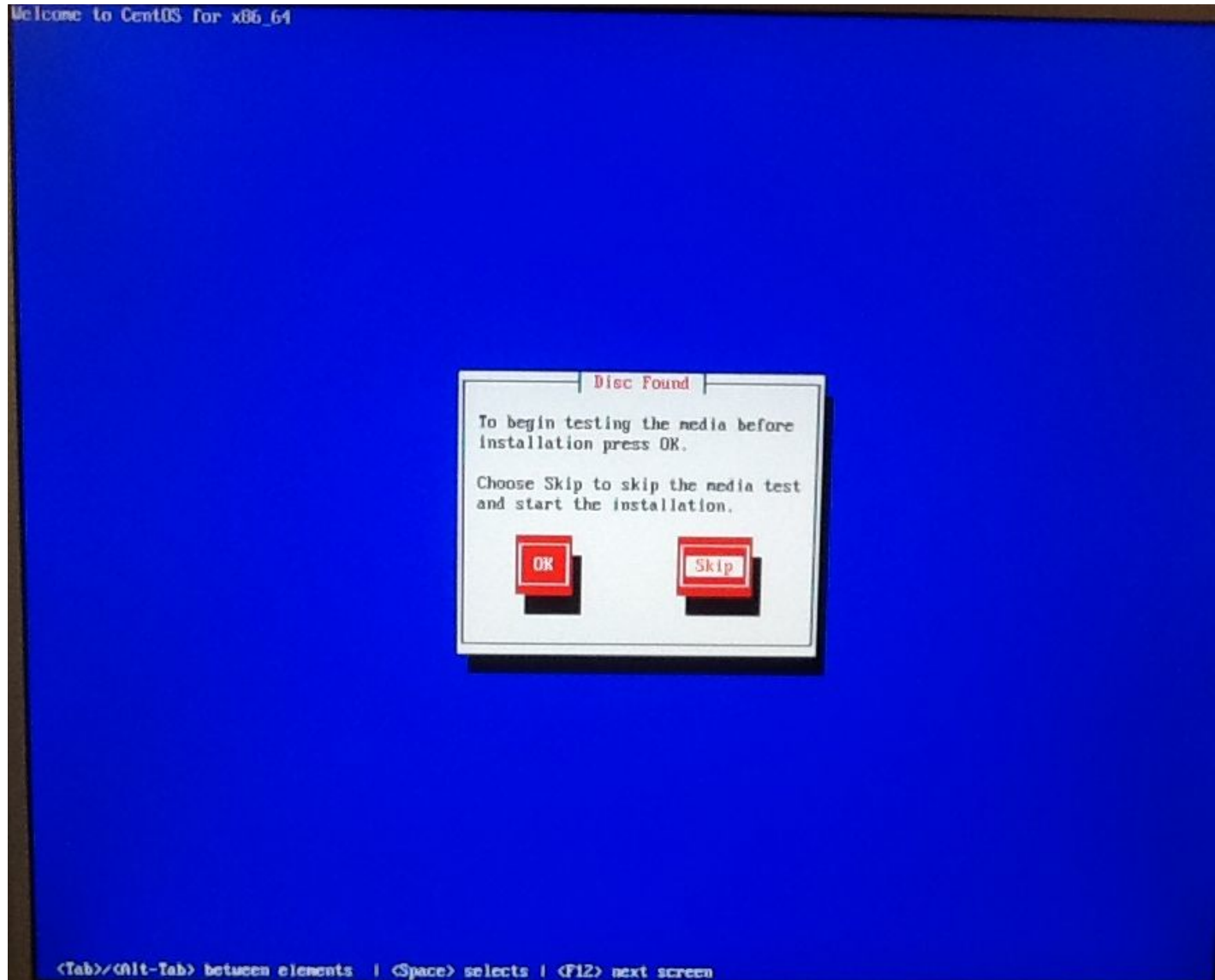
5. Welcome to CentOS画面

インストールDVDの作成に失敗していると、そもそもこの画面が出ない



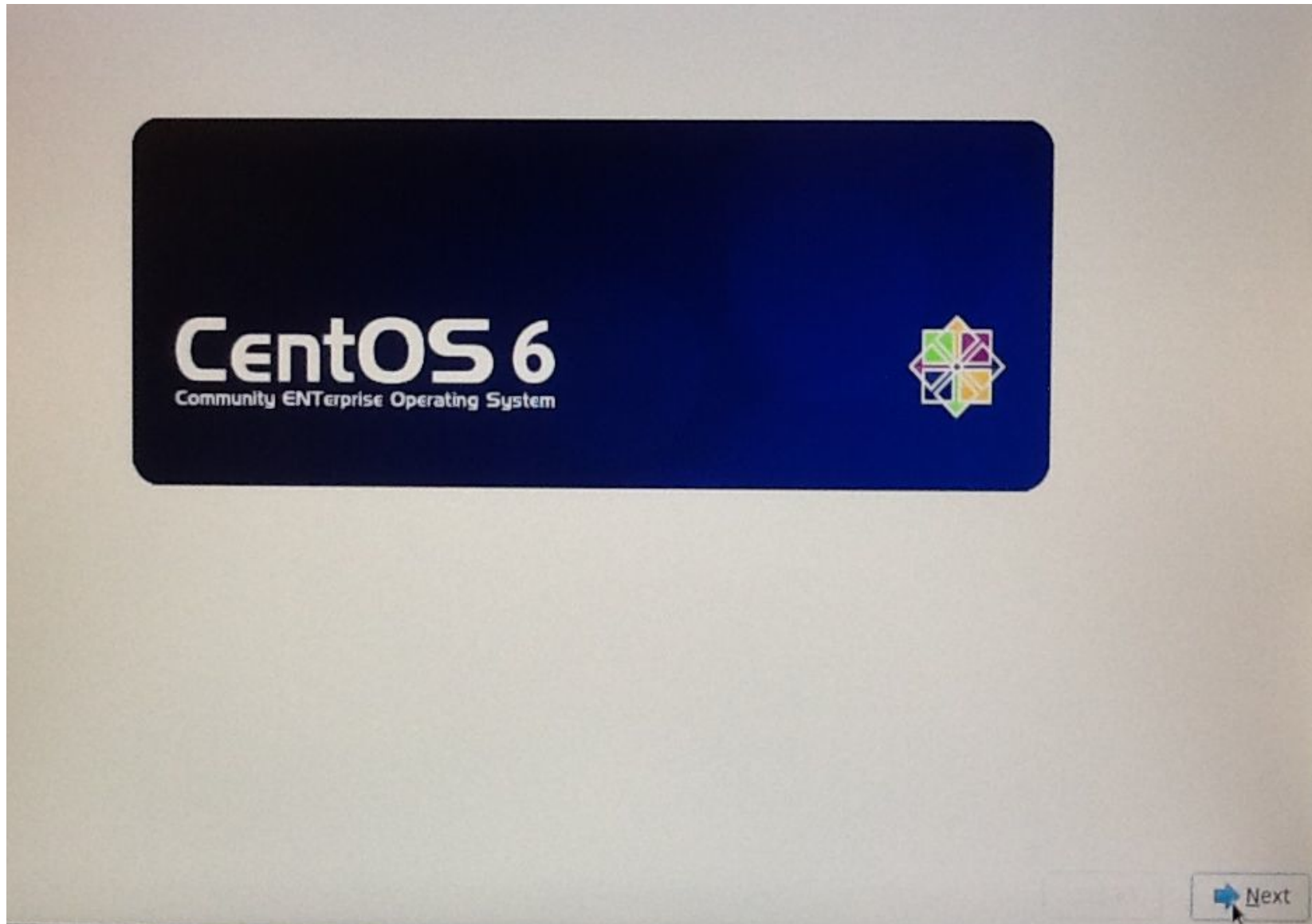
6. インストールDVDのチェック

インストールDVDが問題無いかチェックできるが、時間かかるのでskip



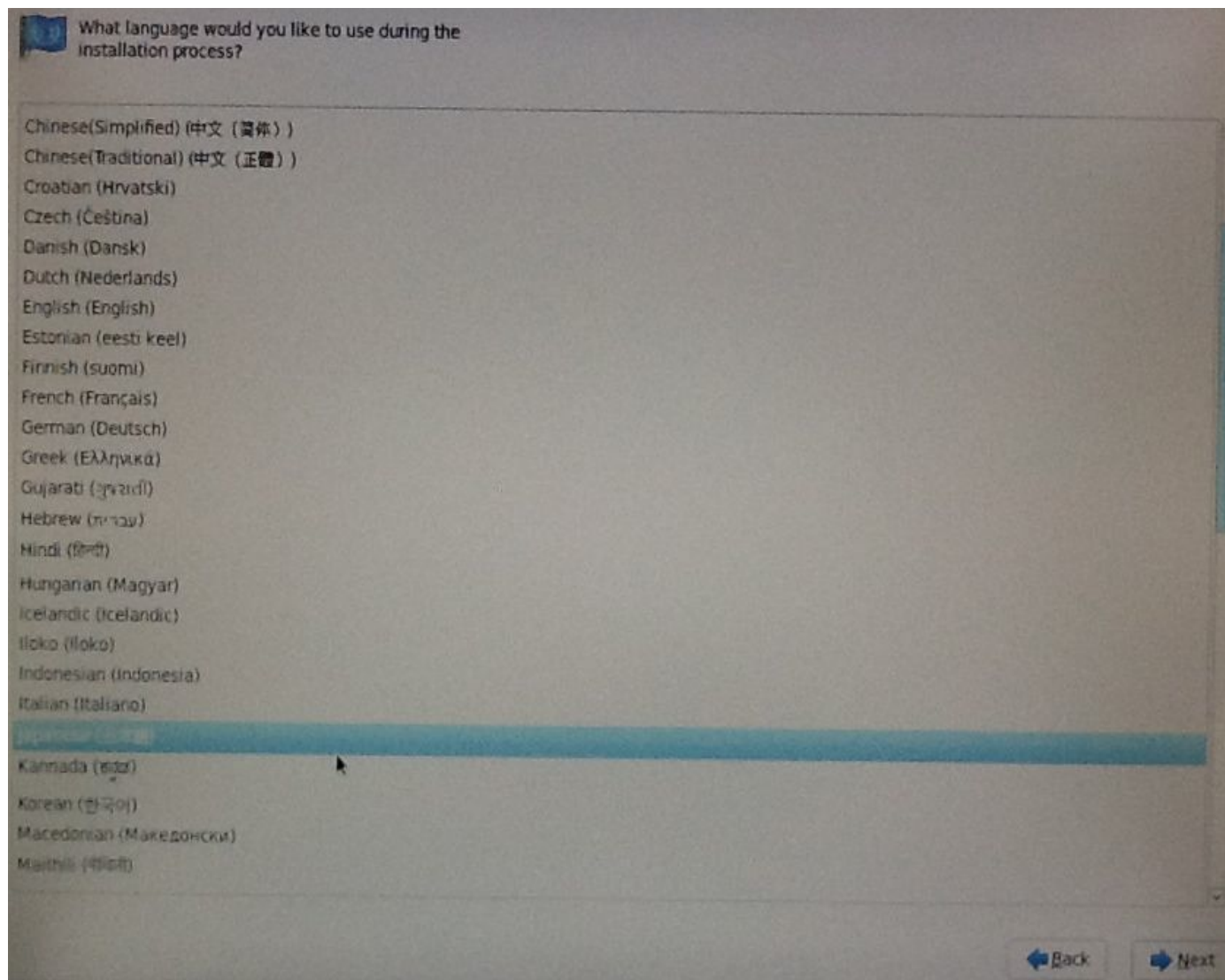
7. CentOSインストールの開始

Nextを押す



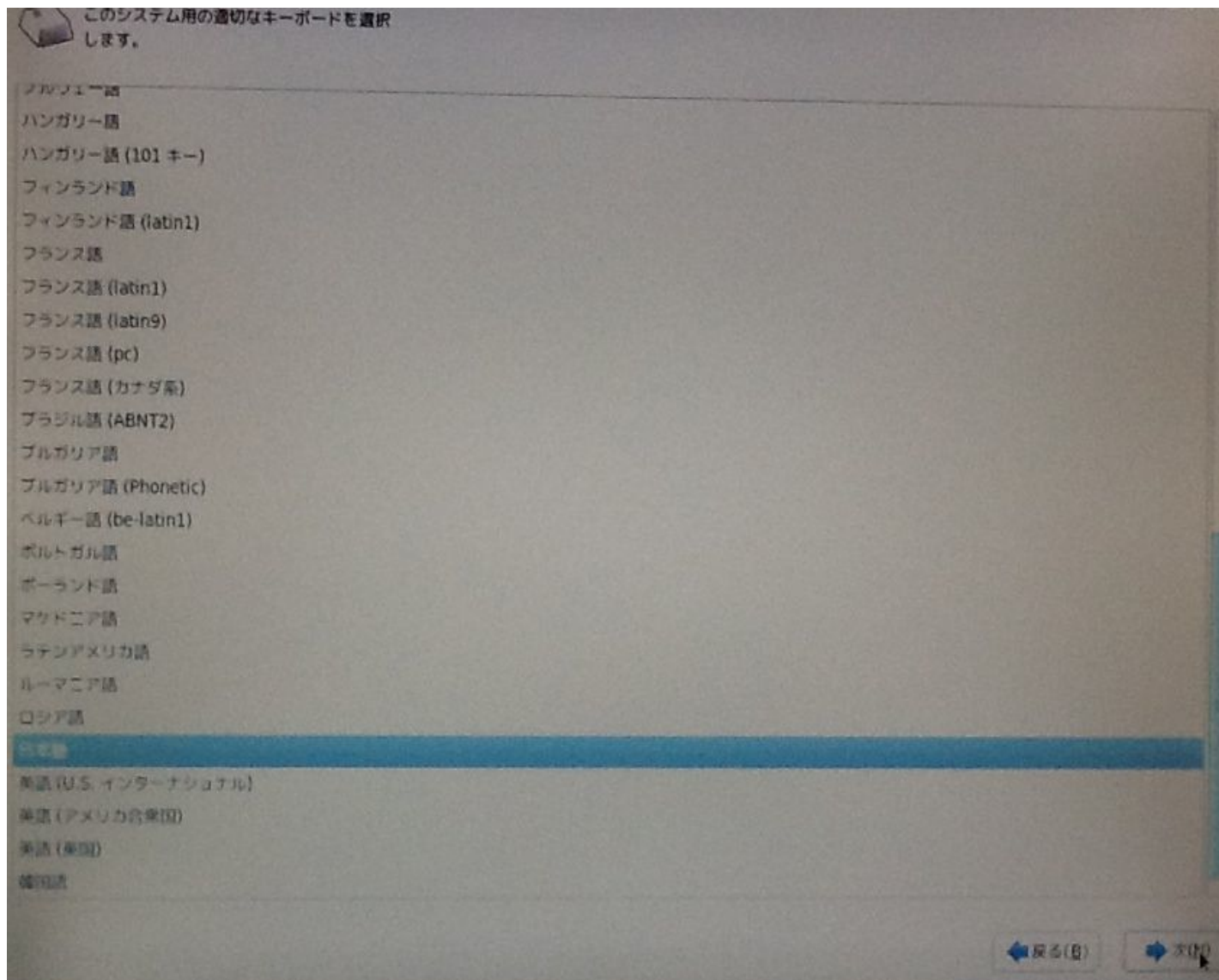
8. インストールに使用する言語を選択

インストール作業で使用する言語として日本語を選択



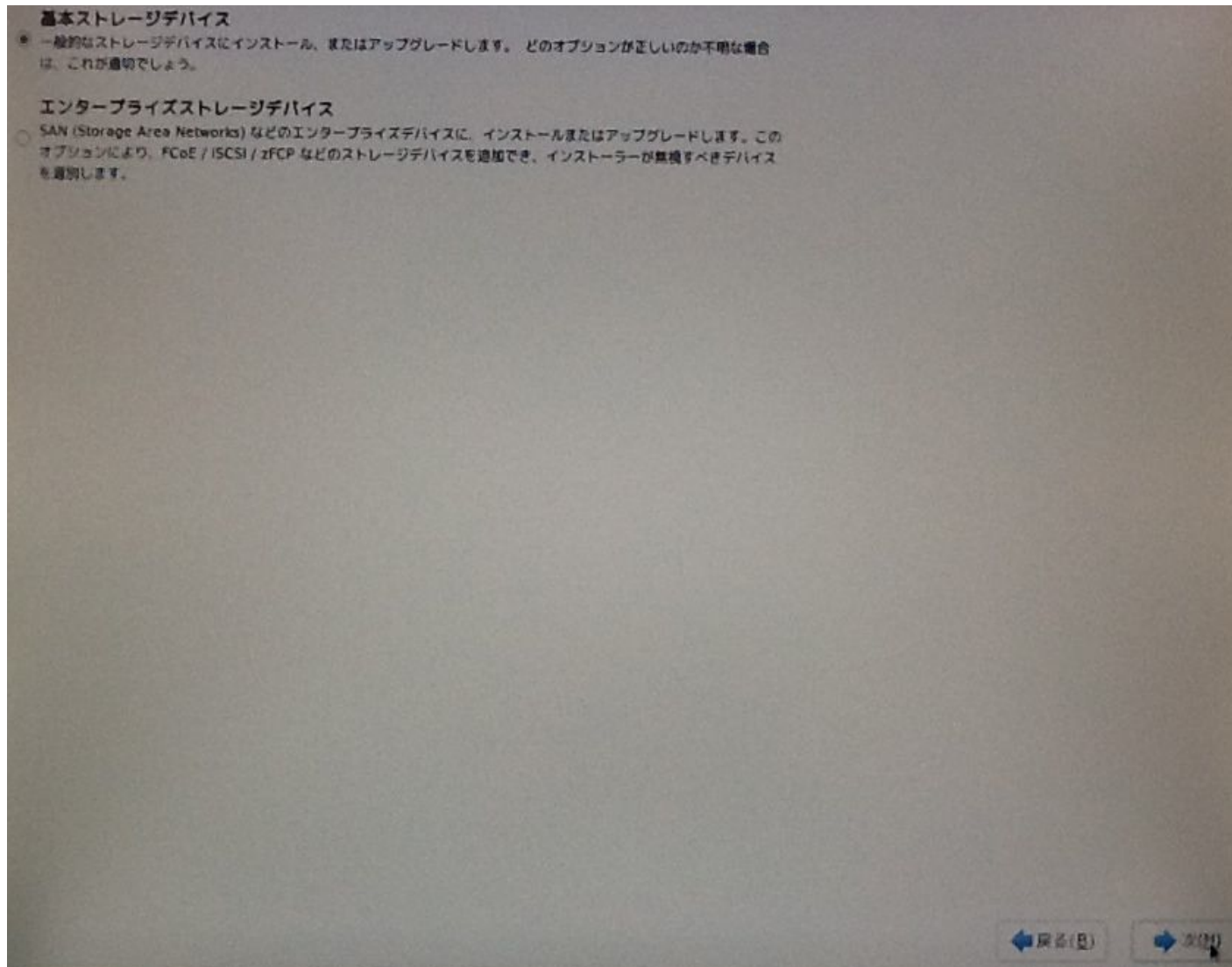
9. キーボードの言語を選択

日本語キーボードを使用しているなら、日本語を選択



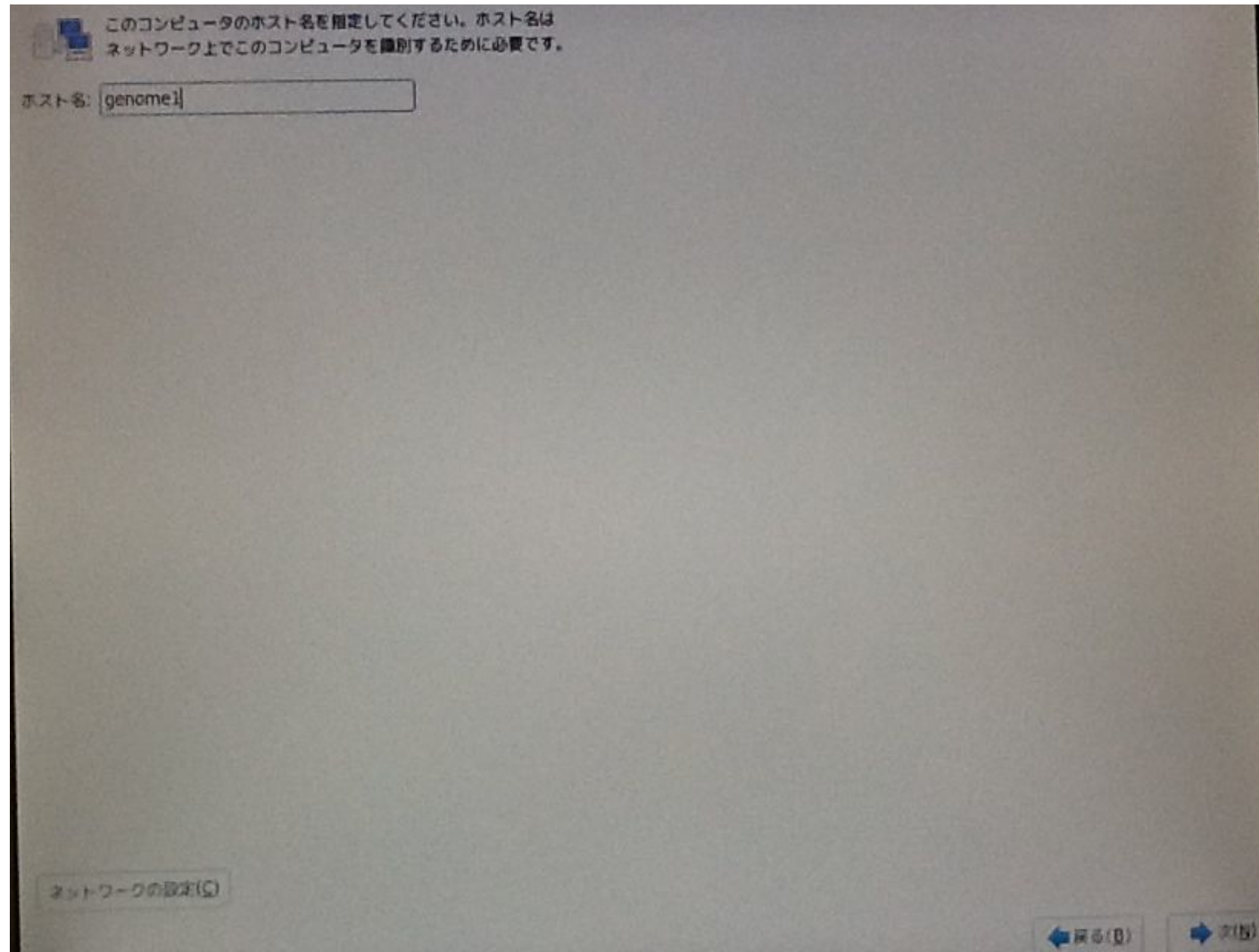
10. CentOSをインストールするストレージを選択

外付けのディスク等にCentOSをインストールする場合以外は、基本ストレージデバイスを選択



11. サーバのホスト名を設定

例: hoge.ac.jp、genome1など。グローバルIPアドレスを取得した場合は、そのホスト名は世界中でユニークである必要がある



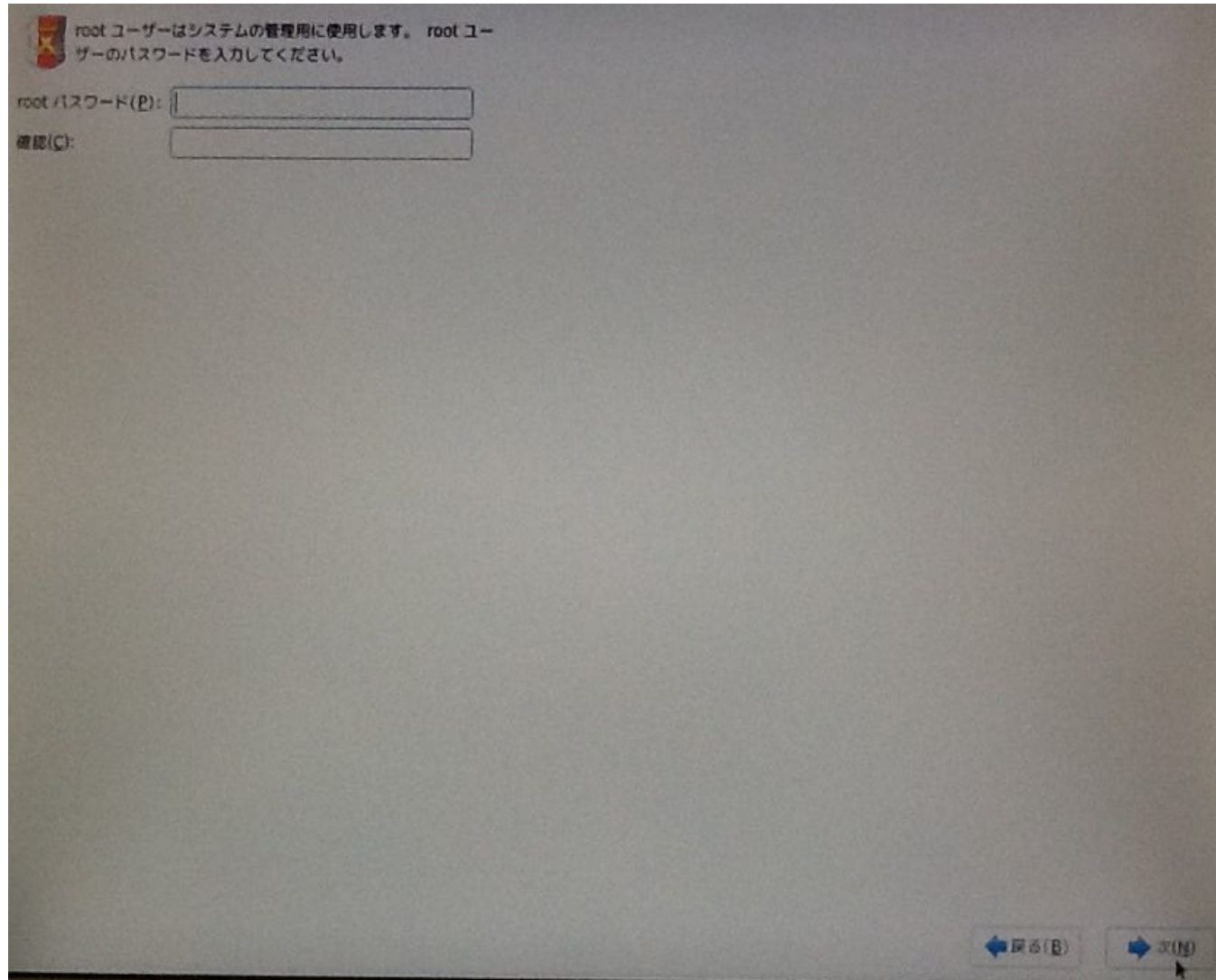
12. タイムゾーンの設定

アジア/東京 を選択。左下の、システムクロックでUTCを使用する、のチェックを外す



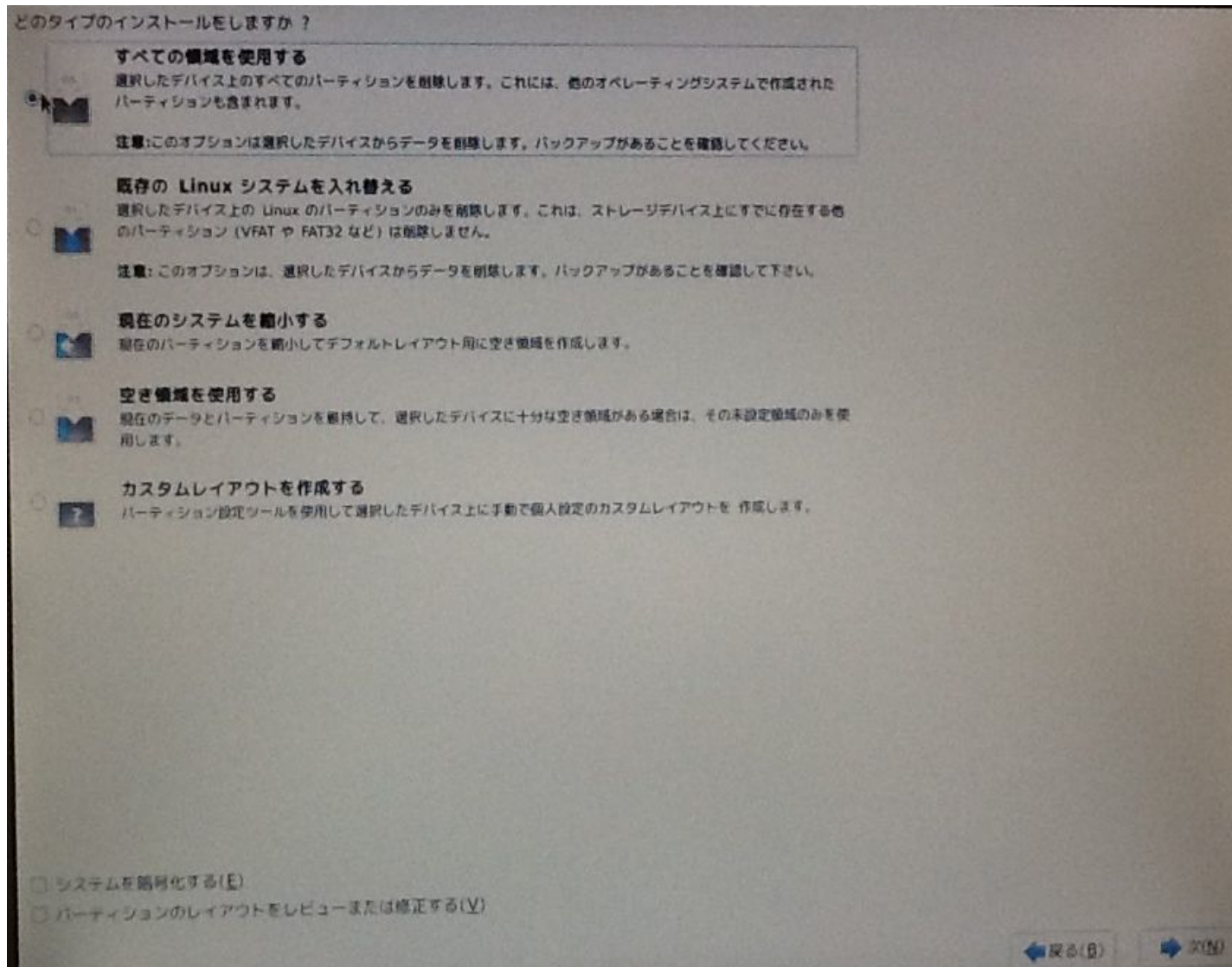
13. サーバの管理者 (root) のパスワードを設定

特にグローバルIPアドレスを取得してサーバを運用する場合は、rootパスワードは辞書に載っている単語等にはしてはいけない



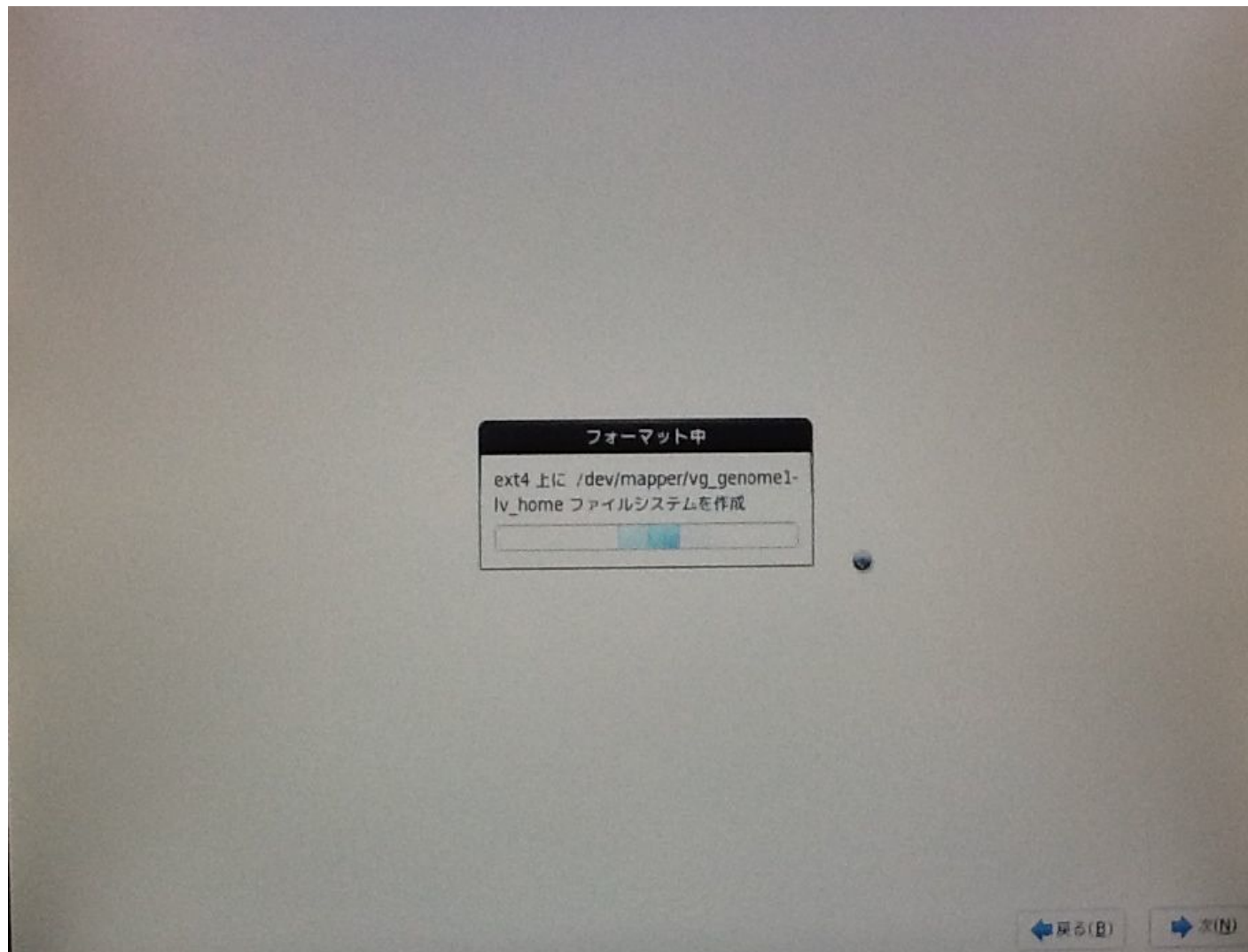
14. インストールタイプの設定

ディスクのパーティション設定を変えたり、ソフトウェアRAIDを構築する必要が無い場合は、すべての領域を使用する、を選択する



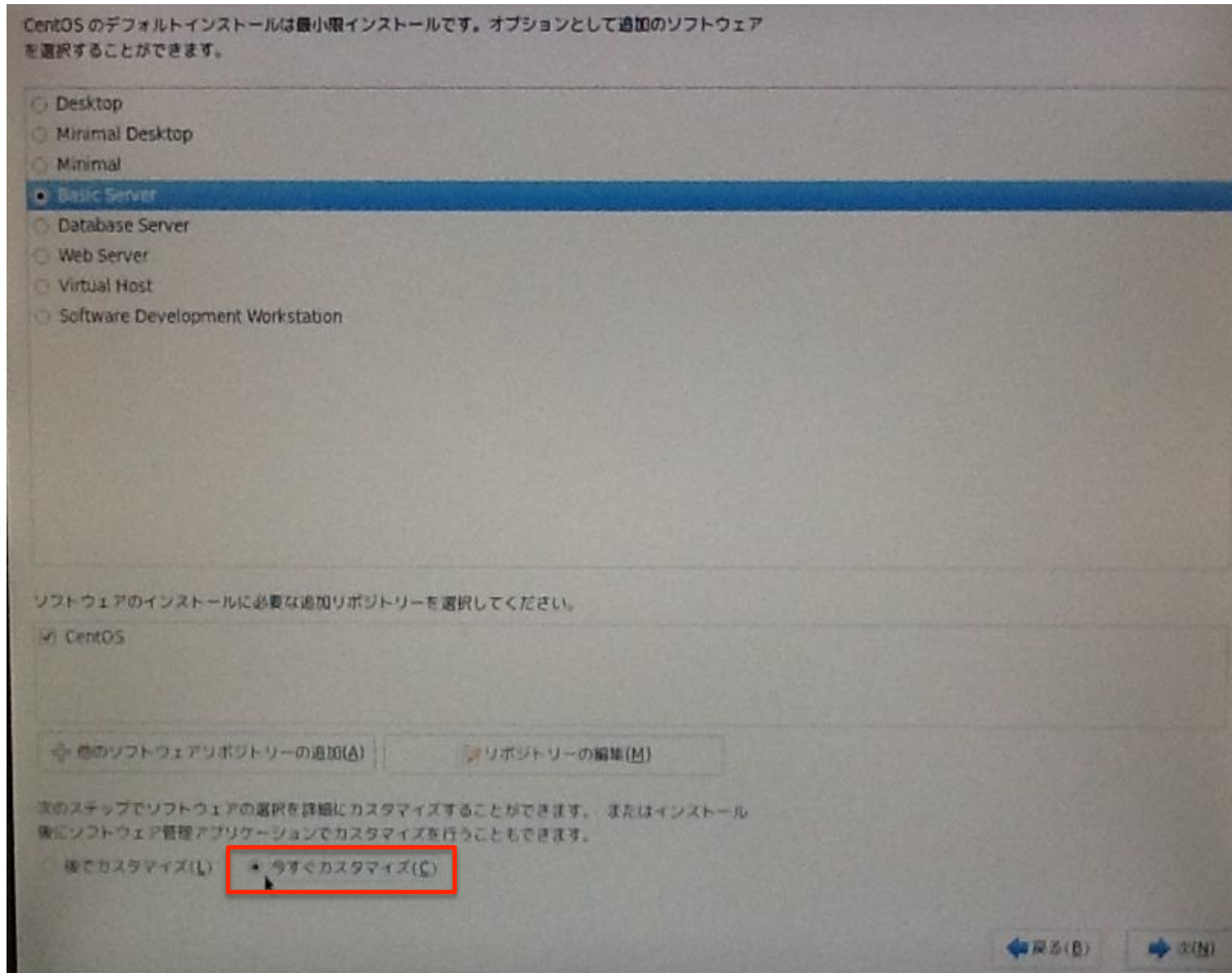
14. インストールタイプの設定

すべての領域を使用する、を選択したら、ディスクの初期化が始まる



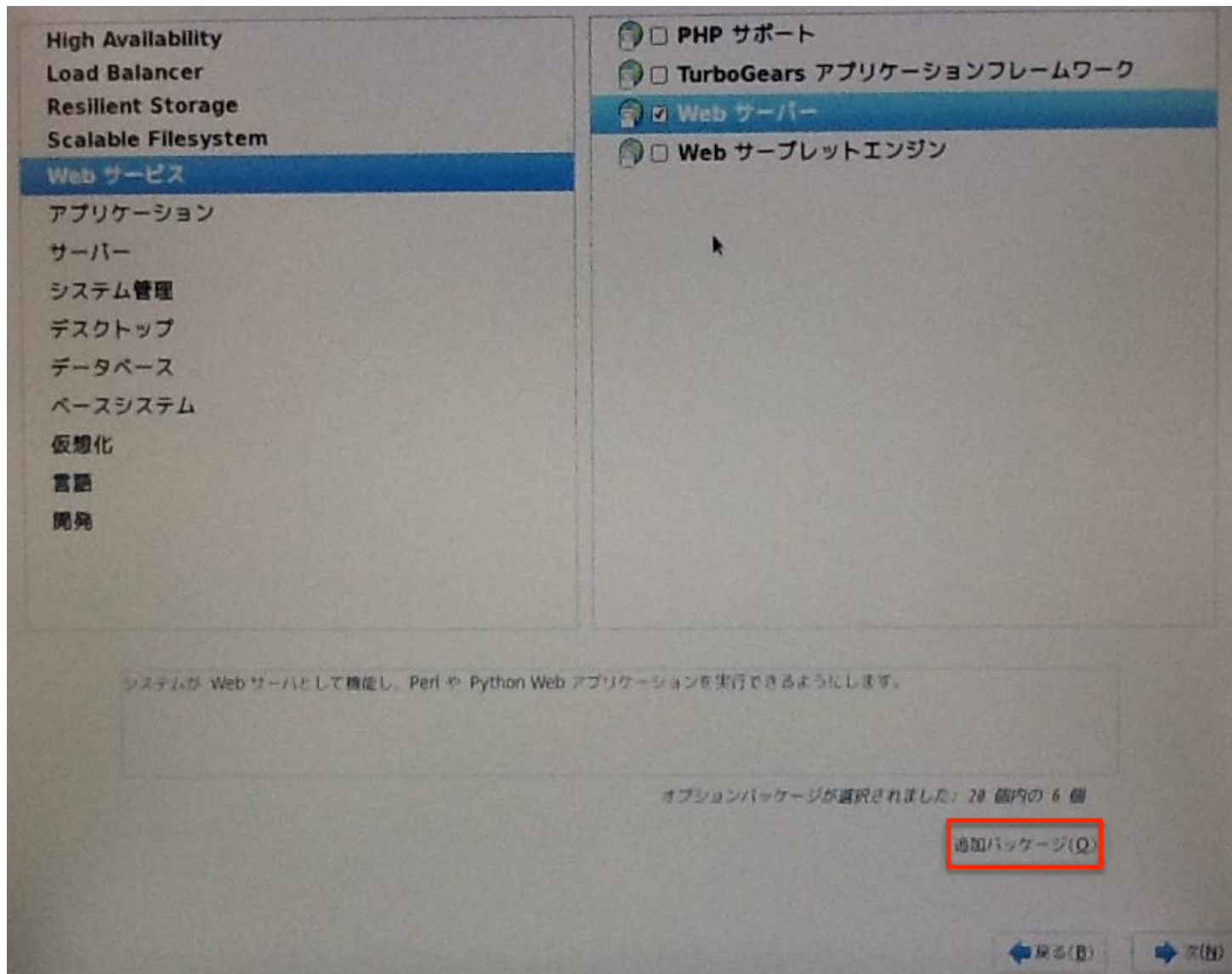
15. OSのインストール時に同時に入れるソフトウェアを選択

Basic Server、左下の、今すぐカスタマイズ、を選択



16. インストールするソフトウェアを追加で選択

情報解析を行うサーバとして必要なソフトウェアを追加で選択



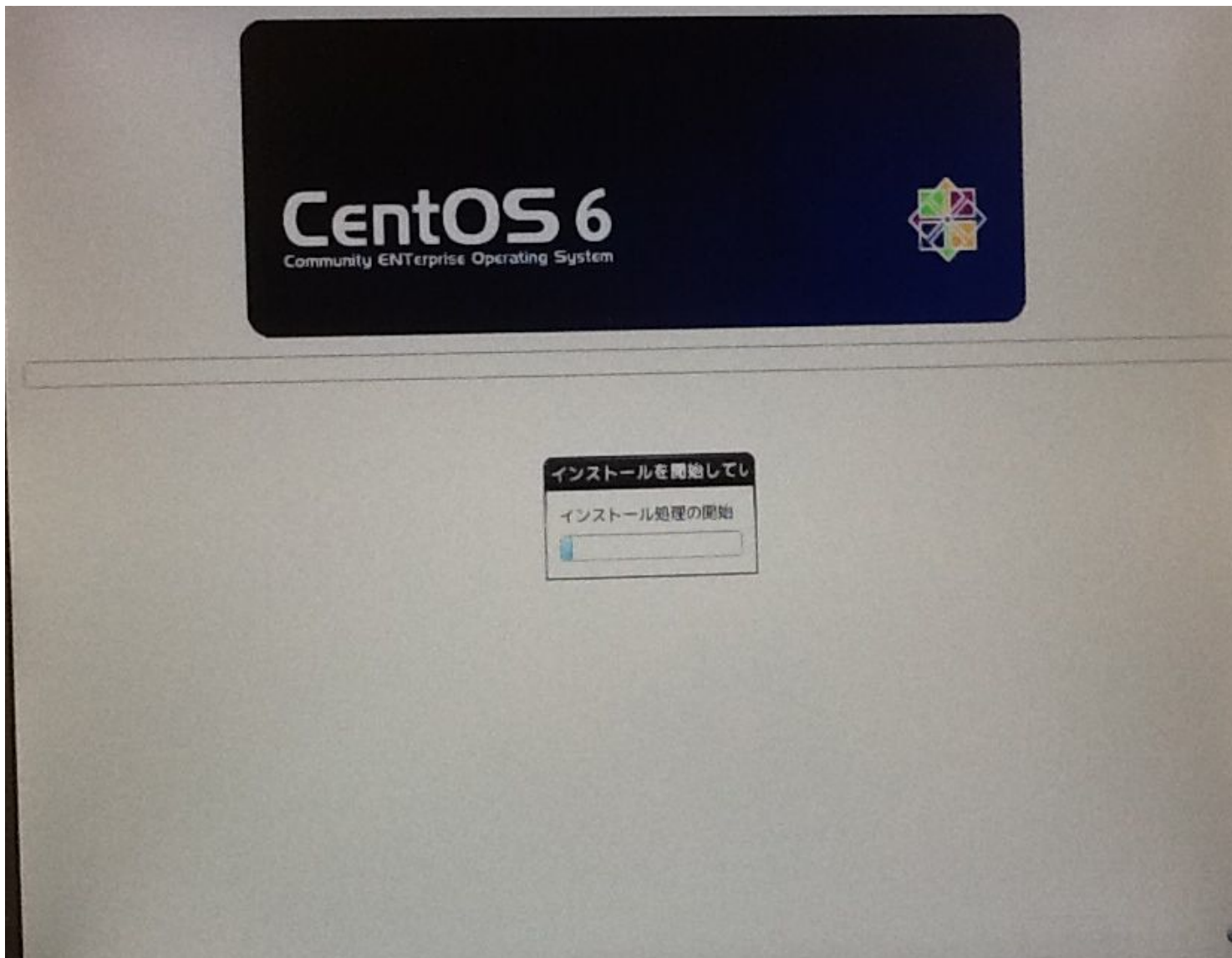
16. インストールするソフトウェアを追加で選択

以下は追加選択の例。あくまで私の個人的な好みです

- Webサービス --> Webサーバー
- アプリケーション --> Emacs
- アプリケーション --> インターネットブラウザ
- サーバー --> システム管理ツール + (追加パッケージ) screen
- デスクトップ --> X Window System
- デスクトップ --> グラフィカル管理ツール
- デスクトップ --> デスクトップ
- デスクトップ --> フォント
- デスクトップ --> 汎用デスクトップ
- ベースシステム --> Rubyサポート
- ベースシステム --> ベース + zsh
- ベースシステム --> 数学/科学系および並列計算 + scipy, numpy
- 開発 --> 開発ツール
- 開発 --> その他の開発

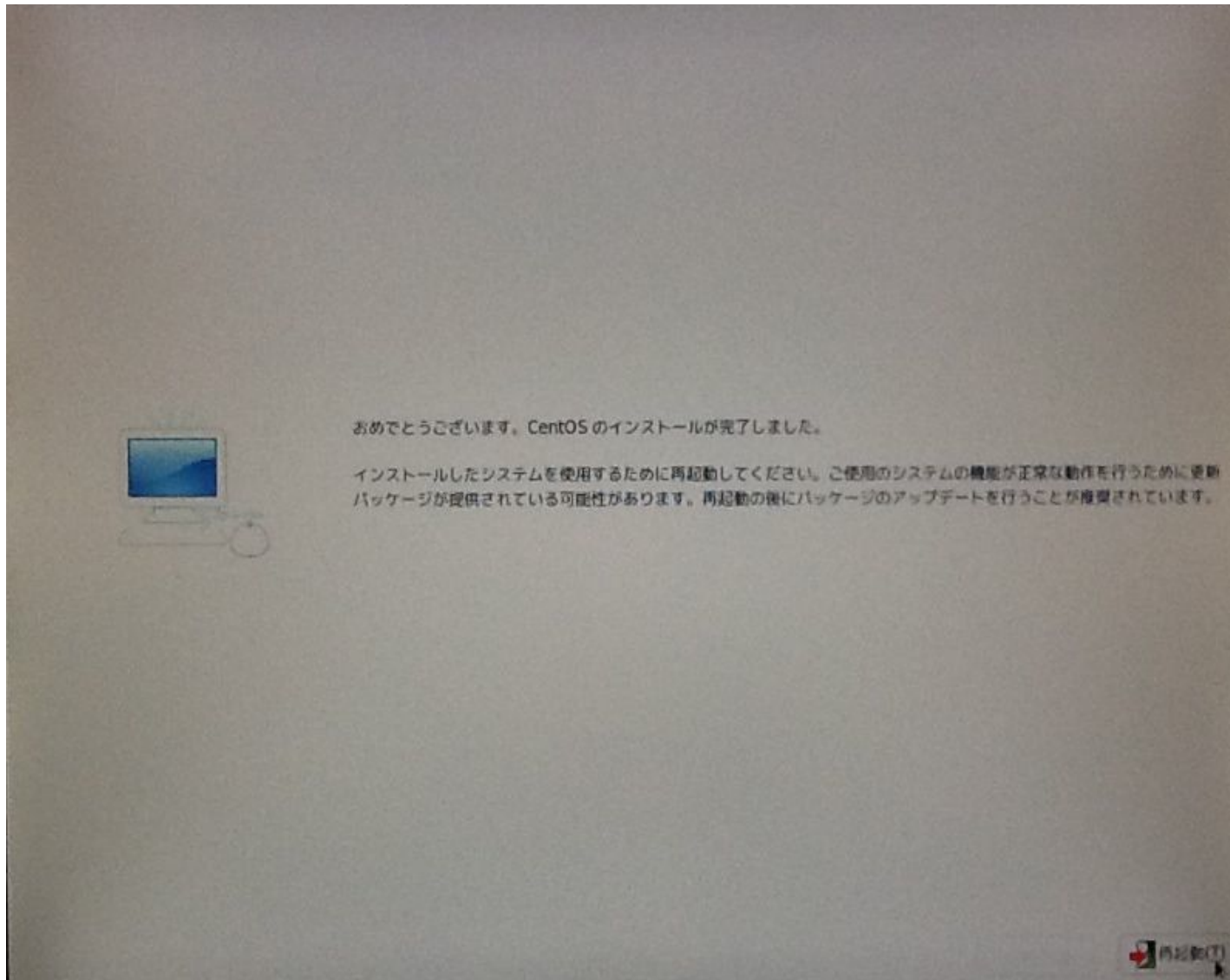
17. インストールの開始

10分以上かかることが多い



18. サーバの再起動

一度電源が切れた後にDVDを取り出す。再起動中にBIOS設定を行い、サーバのハードディスクからOSを起動するように設定する必要がある

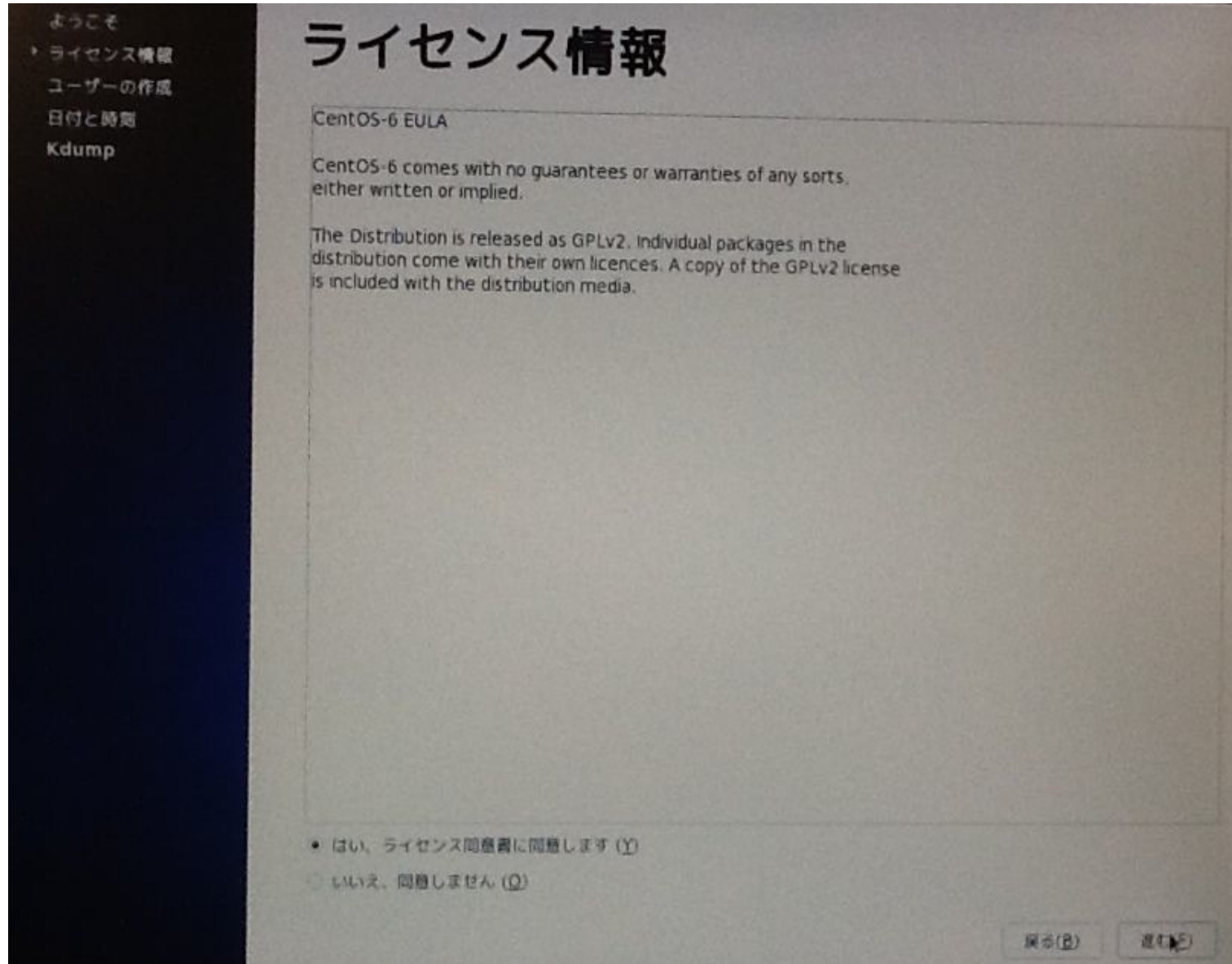


19. ようこそ画面



20. CentOSのライセンスに同意するか否か

ライセンス情報を読み、同意可能なら、同意しますを選択



21. ユーザーの作成

root以外に少なくとも1人分ユーザーアカウントを作成する必要がある

ようこそ
ライセンス情報
ユーザーの作成
日付と時刻
Kdump

ユーザーの作成

システムでの(管理用途ではない)普通の作業のために、'ユーザー'を作成することを推奨します。以下の情報を入力し、システムに'ユーザー'を作成します。

ユーザー名 (U):

フルネーム (E):

パスワード (P):

パスワードの確認 (C):

もしも Kerberos や NIS のようなネットワーク認証が必要な場合、「ネットワークログインを使用する」をクリックしてください。

ネットワークログインを使用する (L)

ユーザーを作成している時点でもっと制御（ホームディレクトリや UID の指定）が必要な場合は、「高度な設定 (Advanced)」ボタンをクリックして下さい。

高度な設定 (A)...

戻る(B) 進む(F)

22. 日付と時刻を設定

ネットワーク上で日付と時刻を同期化しますというチェックボックスはこの段階では特に選択する必要は無い

ようこそ
ライセンス情報
ユーザーの作成
▶ 日付と時刻
Kdump

日付と時刻

システム用に日付と時刻を設定してください。

日時 (T)

現在の日時: 平成25年11月24日 01時51分05秒

ネットワーク上で日付と時刻を同期化します (Y)

手操作であなたのシステムの日時を設定する:

日付 (D) 時間

< 2013 > < 11月 >

日	月	火	水	木	金	土	日
					1	2	
3	4	5	6	7	8	9	
10	11	12	13	14	15	16	
17	18	19	20	21	22	23	
24	25	26	27	28	29	30	

時 (H): 01

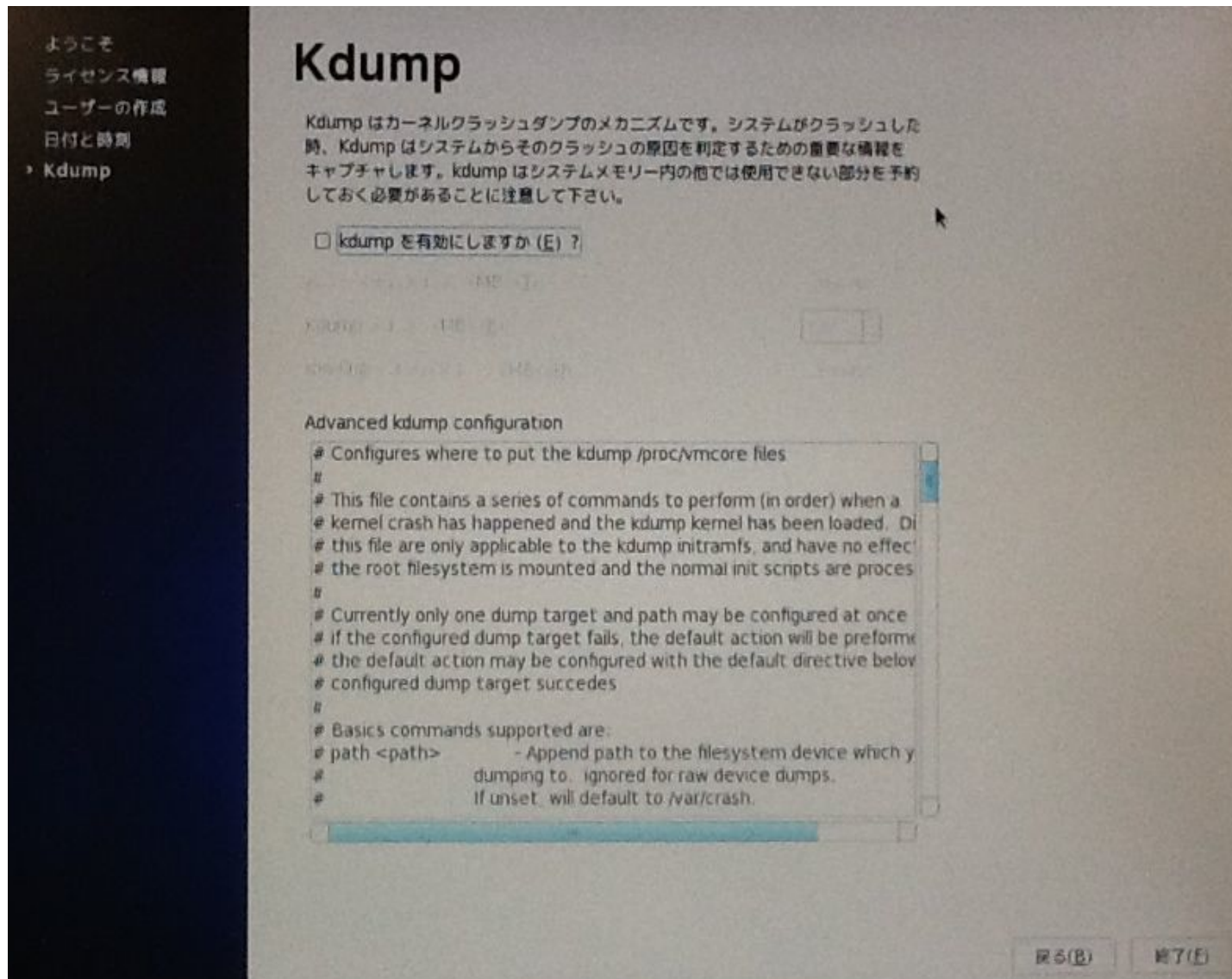
分 (M): 48

秒 (S): 14

戻る (B) 進む (F)

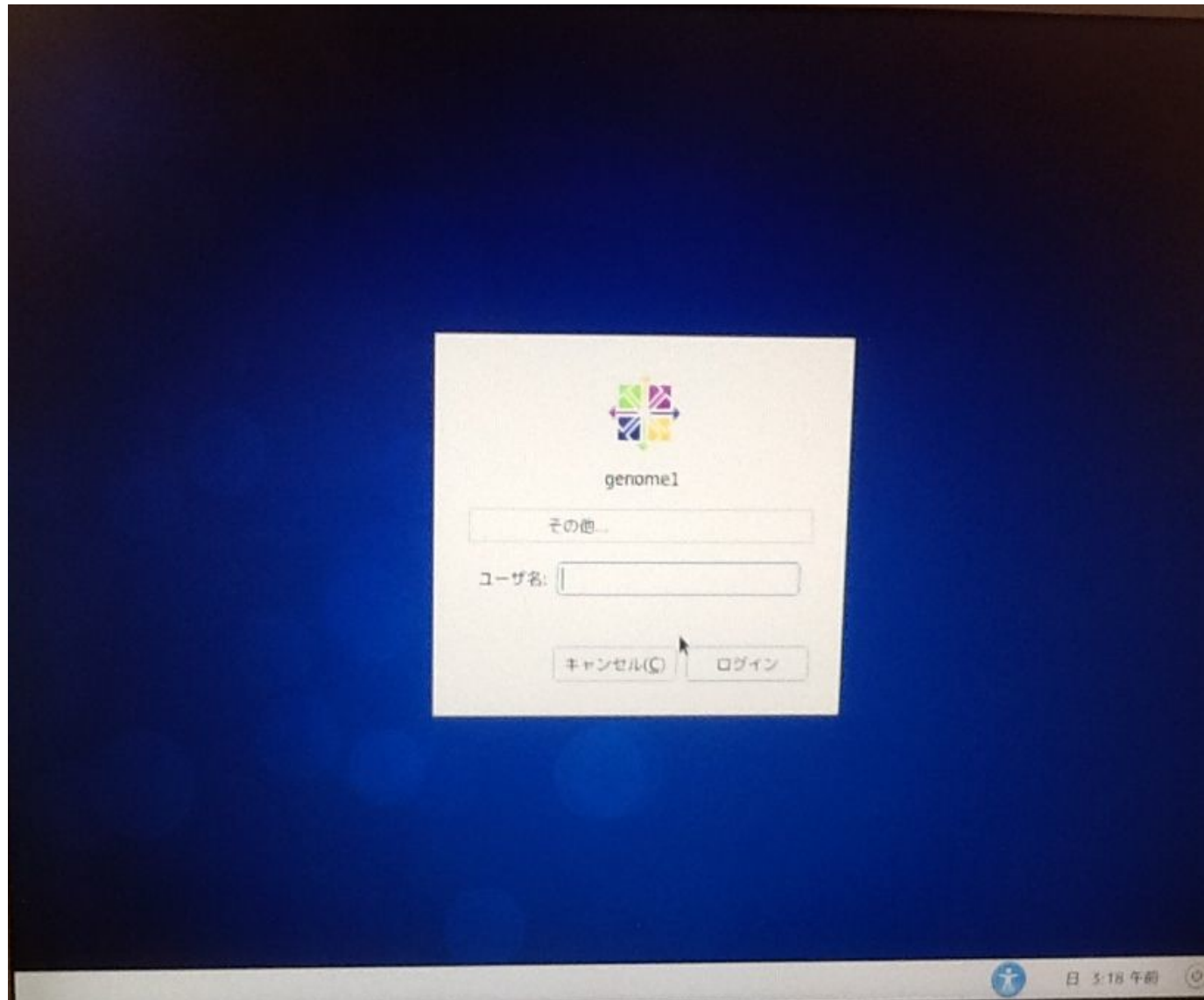
23. Kdumpの設定

OSがクラッシュした場合に様々な情報を保存するKdumpを有効にするか否かの設定。通常は無効で問題無い



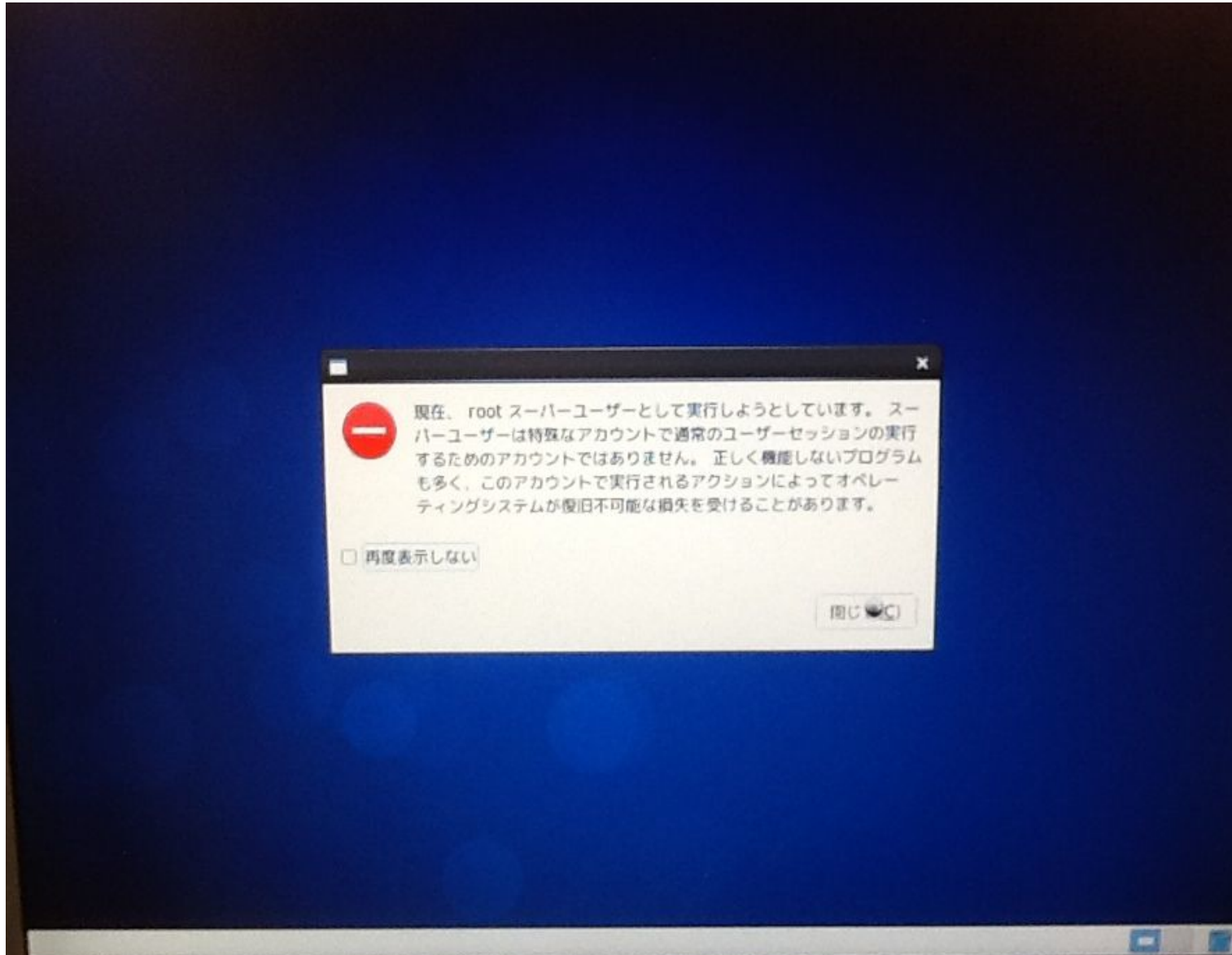
24. rootでログイン

インストールプロセスに問題が無ければ、ログイン画面が出るので、rootで先ほど設定したパスワードでログインする



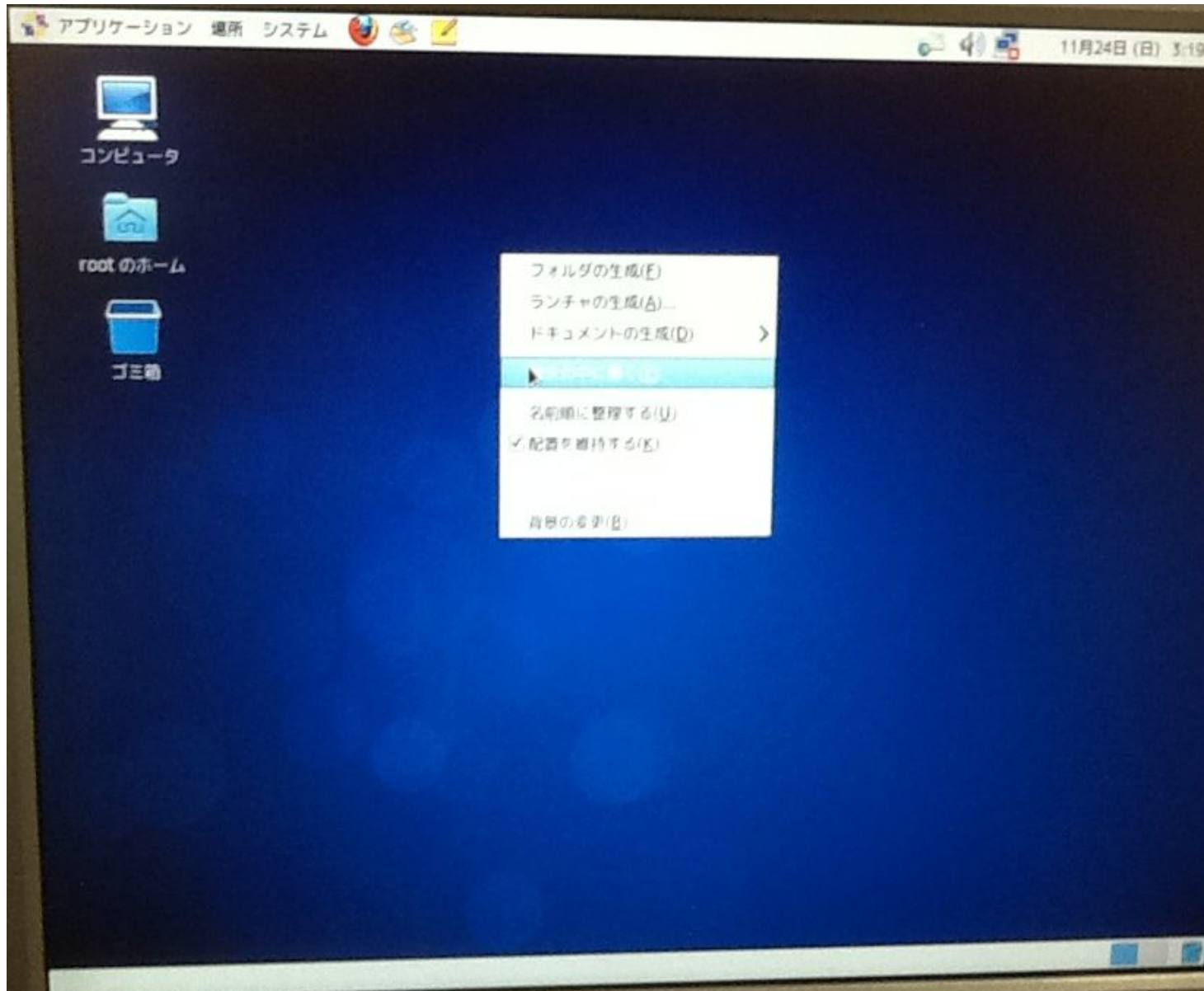
24. rootでログイン

Rootログインはやめた方がよい旨の警告メッセージが出るが、無視



25. ログイン後、端末を立ち上げる

ログインできたら、マウスを右クリックし、端末の中に開く、を選択



25. ログイン後、端末を立ち上げる

コマンドライン入力可能な、端末が立ち上がる

