

# DSASのこれから

～もっとダイナミックなシステムへ～

**サーバ/インフラ Tech Meeting**



**KLab**

2008年8月8日

KLab 株式会社  
Kラボラトリー  
安井 真伸

# 自己紹介

- 安井 真伸 (やすい まさのぶ)
- KLab株式会社に勤務
  - 「Kラボラトリー」という部門でDSASをいじってます
- 最近の出来事
  - 複数サーバ向けのファイル転送ツールを作ってみたり
  - C言語で書いた関数をアセンブラで書き直してみたり
- 好きな事
  - 人前でしゃべること

# はじめに

- インフラの担当者って
  - 障害発生時には圧倒的な存在感を発揮します
  - システムが安定していると存在を忘れ去られます
  - 仕事の内容を一言で説明するのはとても難しいです
  - 「普段なにしてるの？」と聞かれると少し凹みます

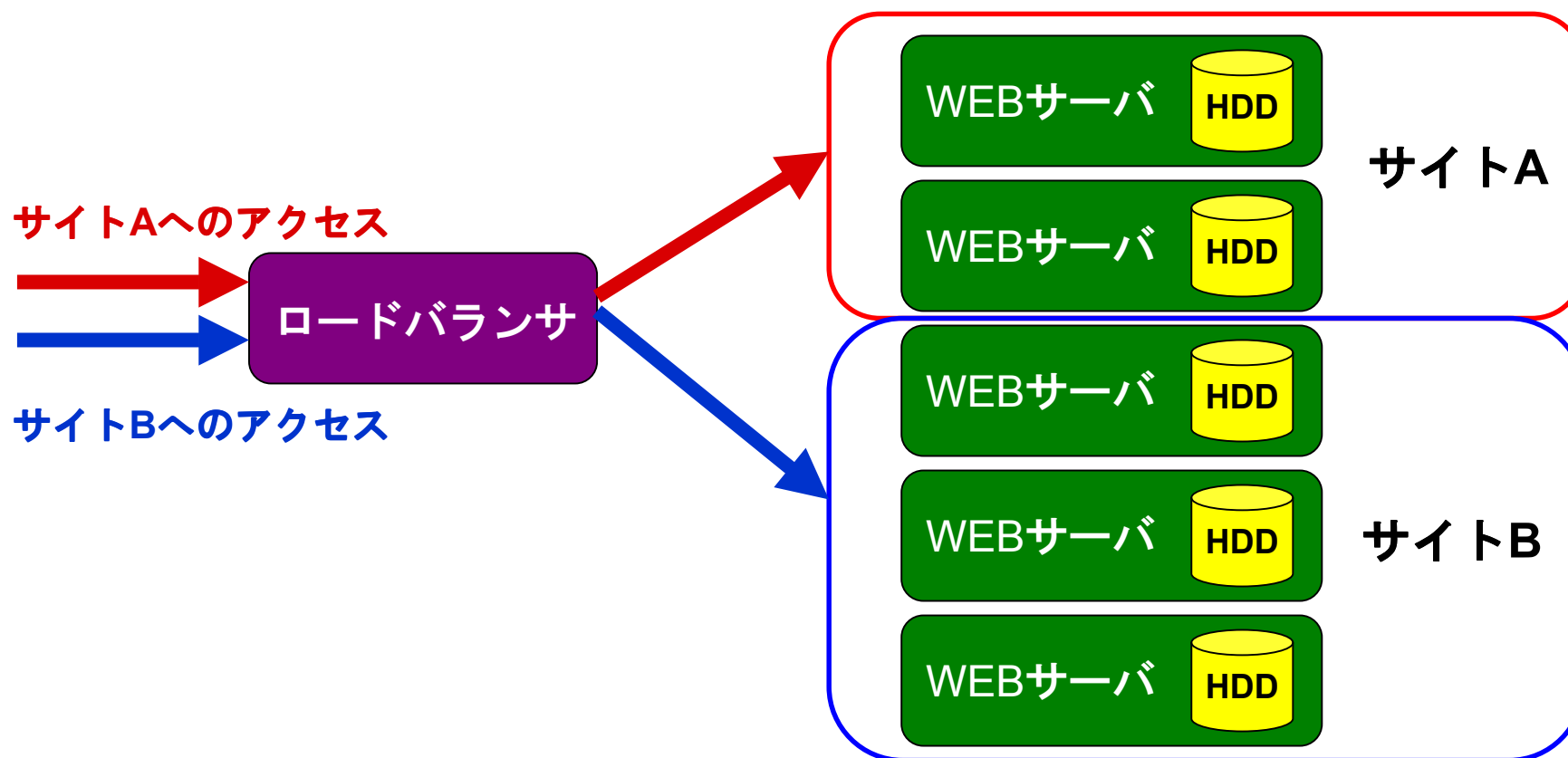
**DSASの日常の様子をお話したいと思います**

# DSASってなに？

- **KLabで運用してるサーバ・インフラの名前です**
  - 「ダイナミック・サーバ・アサイン・システム」の略です
  - 名前が付いているとなにかと便利です
  - KLabの商品ですが、近所のスーパーでは売ってません
- **全部で300台から400台くらいのサーバがあります**
  - 東京に2カ所、福岡に1カ所のデータセンターがあります
- **オープンソースベースのシステムです**
  - Linux, Apache, PHP, Tomcat, その他もろもろ

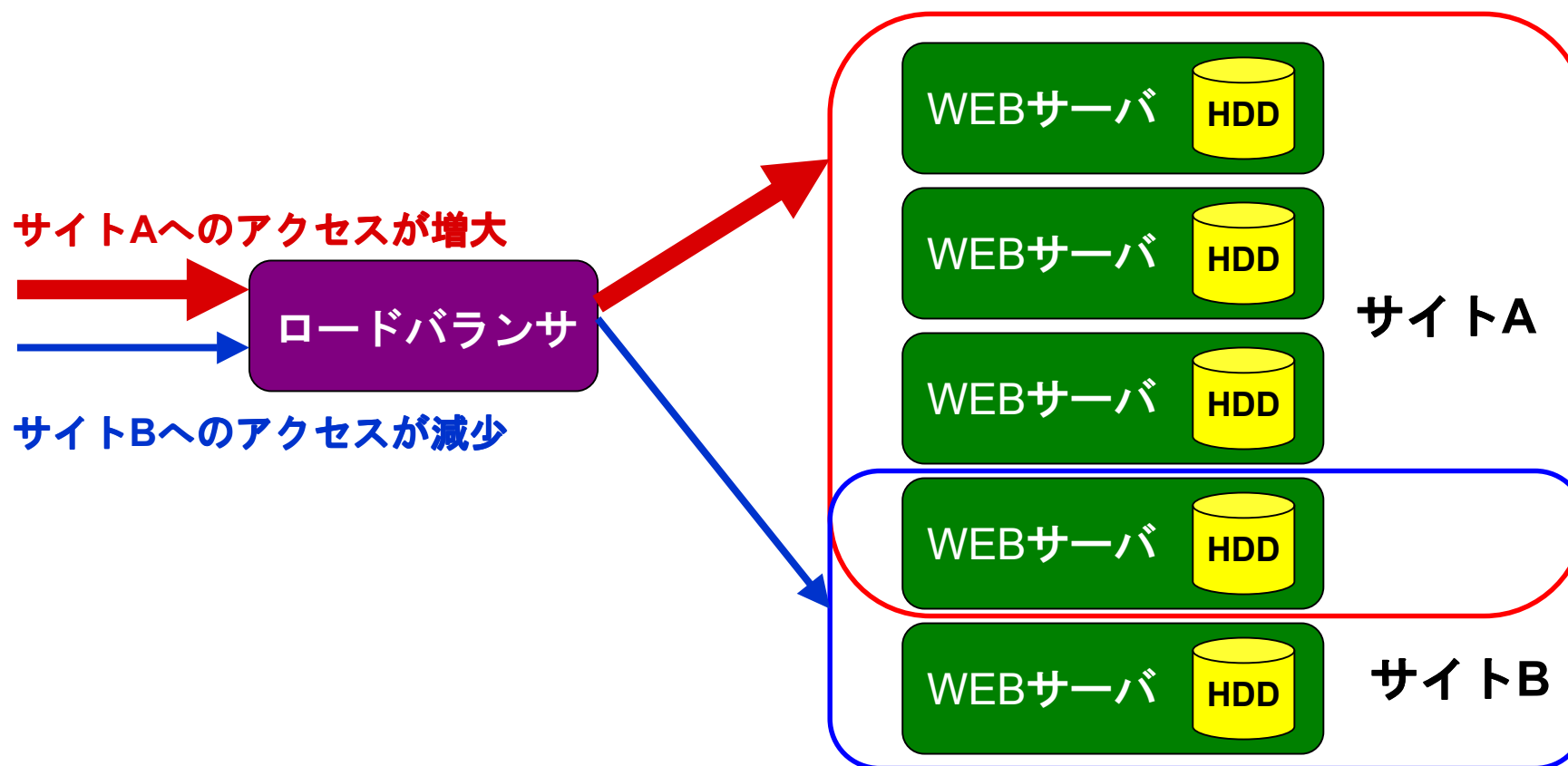
# なにがダイナミックなの？

サイト毎のサーバ台数を動的に変更できるという特徴に由来



# なにがダイナミックなの？

サイト毎のサーバ台数を動的に変更できるという特徴に由来



でも、それってさ . . .

ロードバランサの設定を  
変えてるだけじゃね？



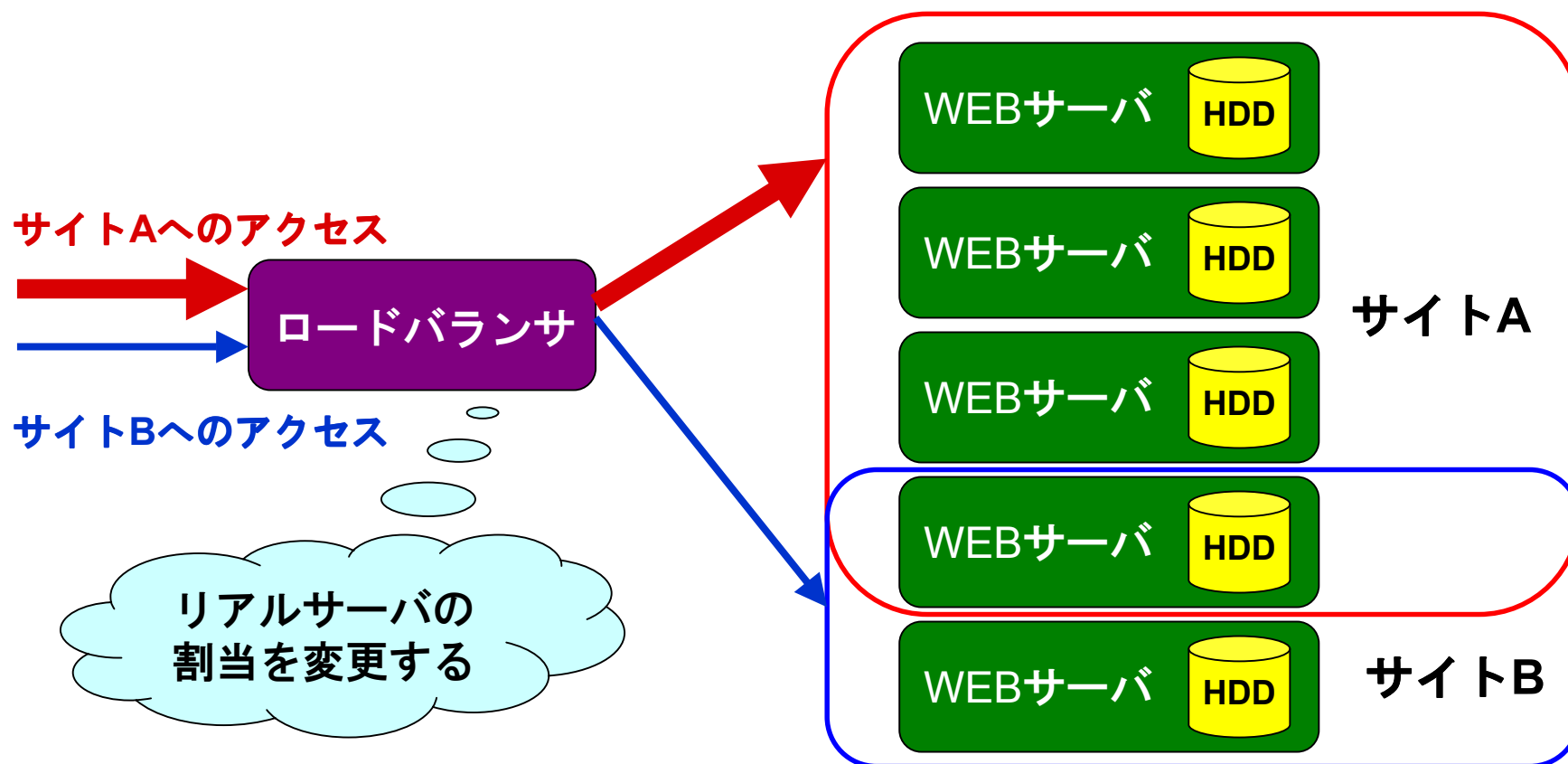
そのとおりです (><)

しかも！

人が手でやっています  
＼ (^ ^ ) ／

# なにがダイナミックなの？

サービス毎のサーバ台数を動的に変更できるという特徴に由来

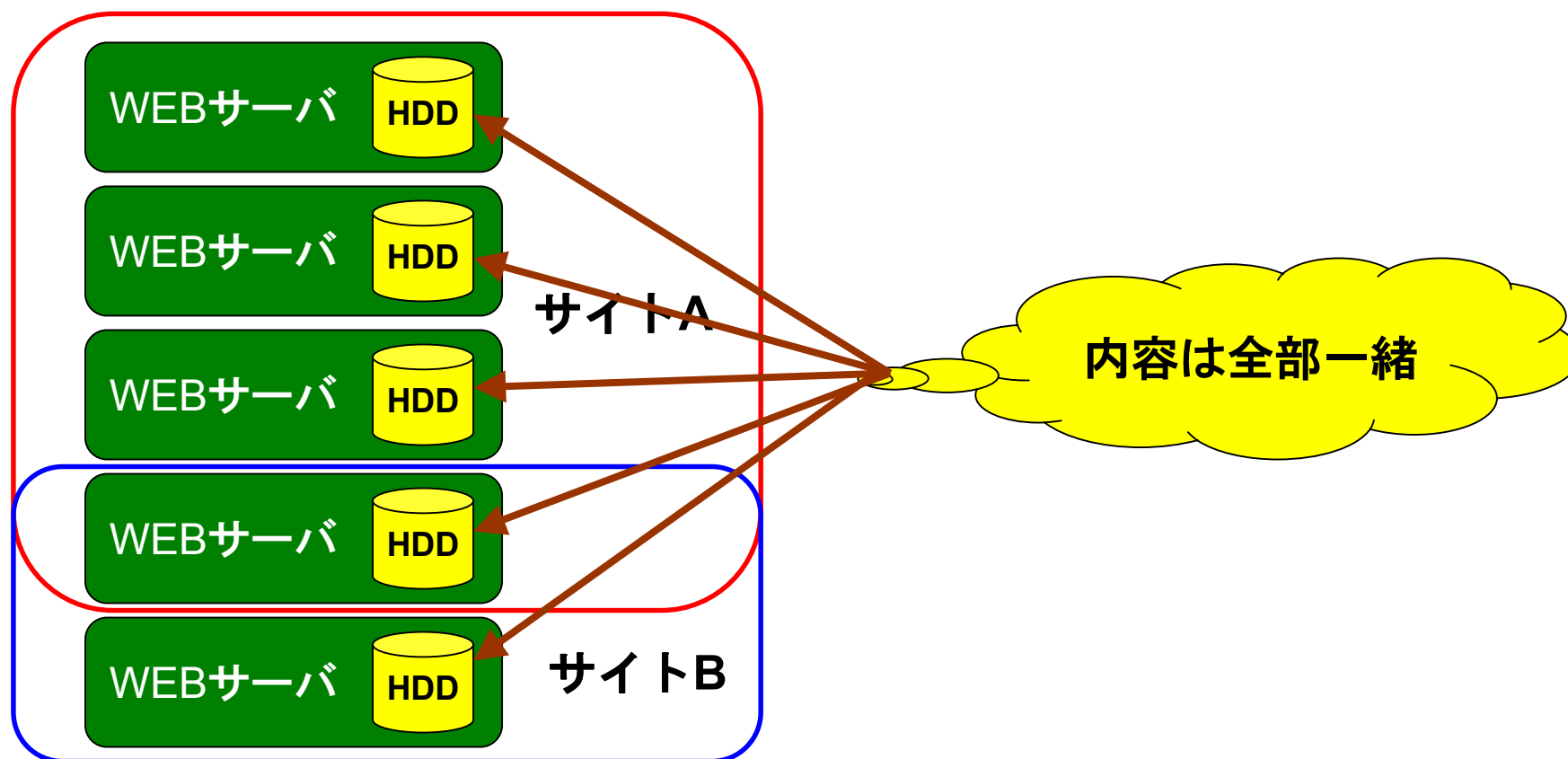


とはいえ

いろいろ工夫しています

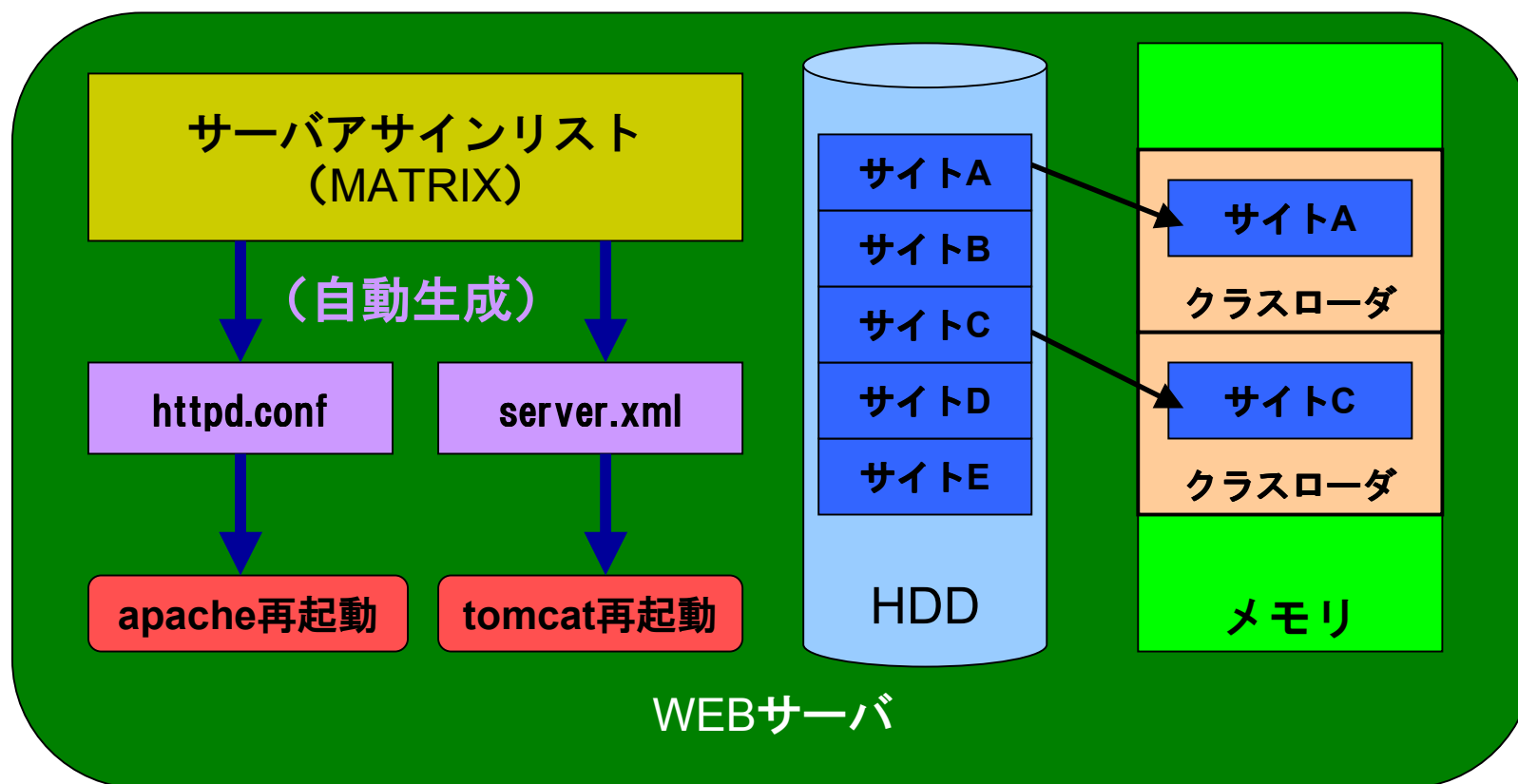
# いろいろ工夫しています(1)

全サーバのディスクの内容を同じに保つ



# いろいろ工夫しています(2)

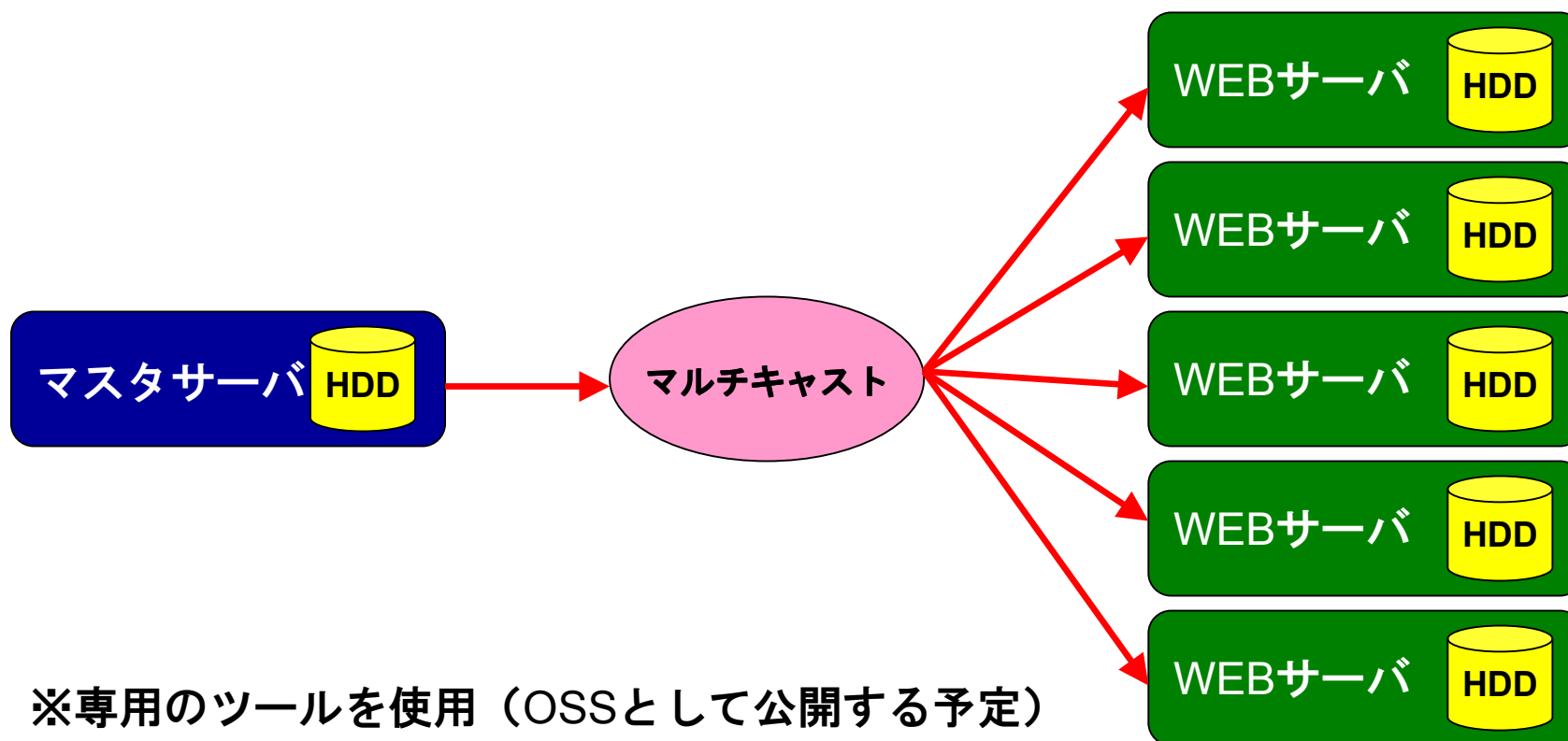
## 必要なクラスローダのみを起動する





# いろいろ工夫しています(3)

## ファイルの更新はマルチキャストで一括転送



※専用のツールを使用 (OSSとして公開する予定)

## その他にも

- ネットワーク構成をあんな感じにしたり
- ストレージサーバでこんなことしたり
- DBサーバをそんな感じで使ってみたり
- セッションサーバをごによごによしたり

# この辺の詳細は.....



## ■目次

- 1章 サーバ/インフラ構築入門 .....冗長化/負荷分散の基本
- 2章 ワンランク上のサーバ/インフラの構築 .....冗長化、負荷分散、高性能の追求
- 3章 止まらないインフラを目指すさらなる工夫 .....DNSサーバ、ストレージサーバ、ネットワーク
- 4章 性能向上、チューニング .....Linux単一ホスト、Apache、
- 5章 省力運用 .....安定したサービスへ向けて
- 6章 あのサービスの舞台裏 .....自律的なインフラへ、ダイナミックなシステムへ

少し前置きが長くなりましたが、、

**DSASの概要はこんな感じです**

ところで

「ダイナミック」ってどういう意味？

# 国語辞典で調べてみました

- **ダイナミック**
  - 力強く、生き生きとしているさま。躍動的。力動的。
- **サーバ**
  - (1883-1957) イタリアの詩人。伝統的詩法を駆使し、平明な口語体で物語性あふれる音楽的な抒情詩を綴った。詩集「カンツォニエーレ」など。
- **アサイン**
  - 割り当てること。割り振ること。代入すること。
- **システム**
  - ある作業をコンピューターで行なっている場合に中核となっている部分をさす語



というわけで、「DSAS」とは・・・

**「生き生きとしたサバを  
割り当てるシステム(?)」**

お寿司屋さんへ売ってそうですね！  
＼ (^ ^ ) ／

**この手の単語の意味を  
国語辞典で調べるのは  
おすすりめしりません**

冗談はこれくらいにして・・・

話をもどします

# 普段やってること

基本的に何をするかは自由です。

人にとってテーマや内容は様々です。

- Linuxのカーネルソース読んでみたり
- IPVSの拡張モジュール書いてみたり
- カーネルの内部情報を参照する方法を調べたり

なぜそんなことやってるの？



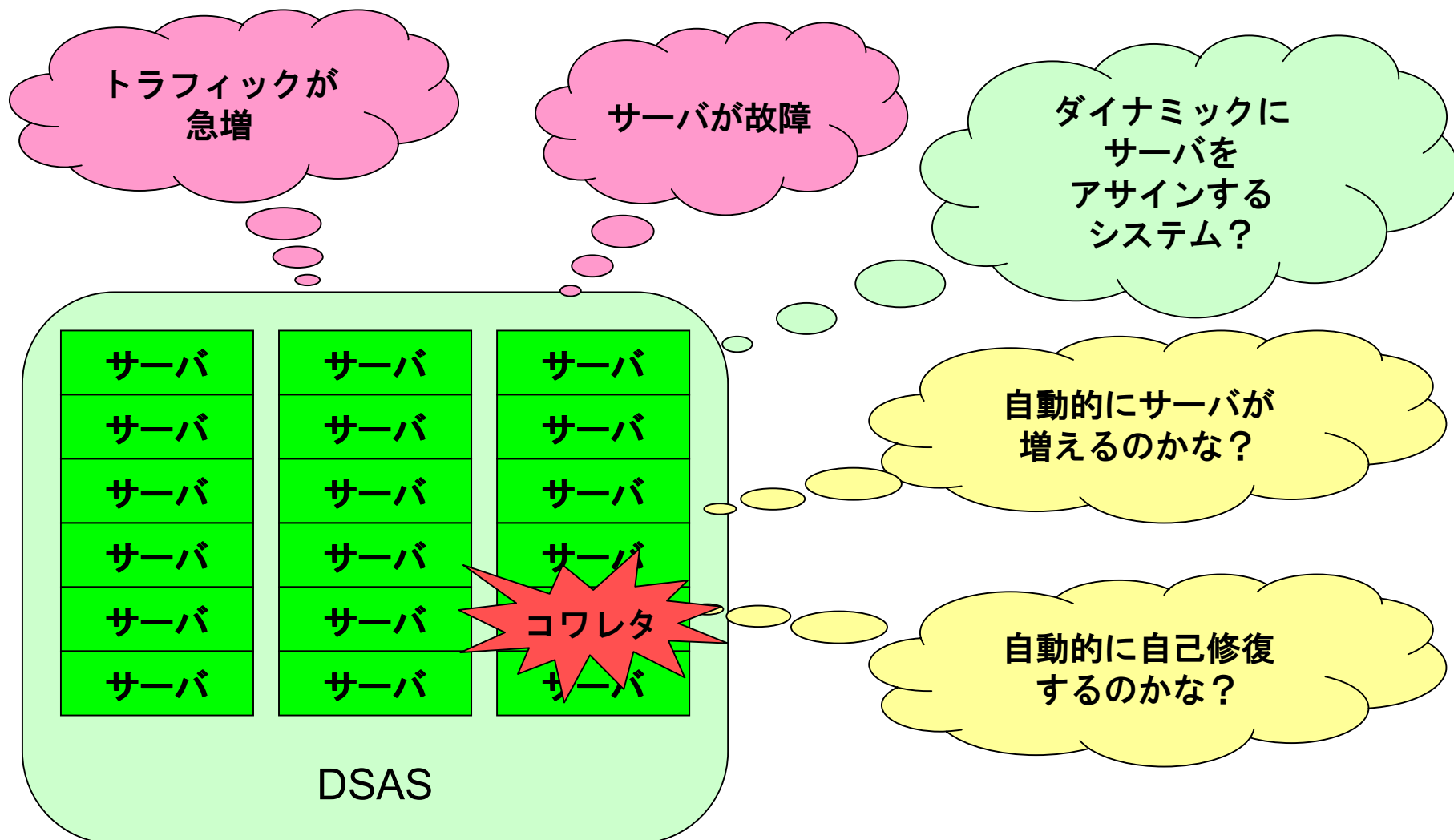
これから順を追ってお話します

# 「ダイナミック」がつく単語

- ダイナミックDNS
- ダイナミックルーティング
- ダイナミックリンクライブラリ
- DHCP(Dynamic Host Control Protocol)

なんとなく、「自動的」とか、「自立的」みたいなイメージってありません？

# 「ダイナミックなシステム」から連想されるもの



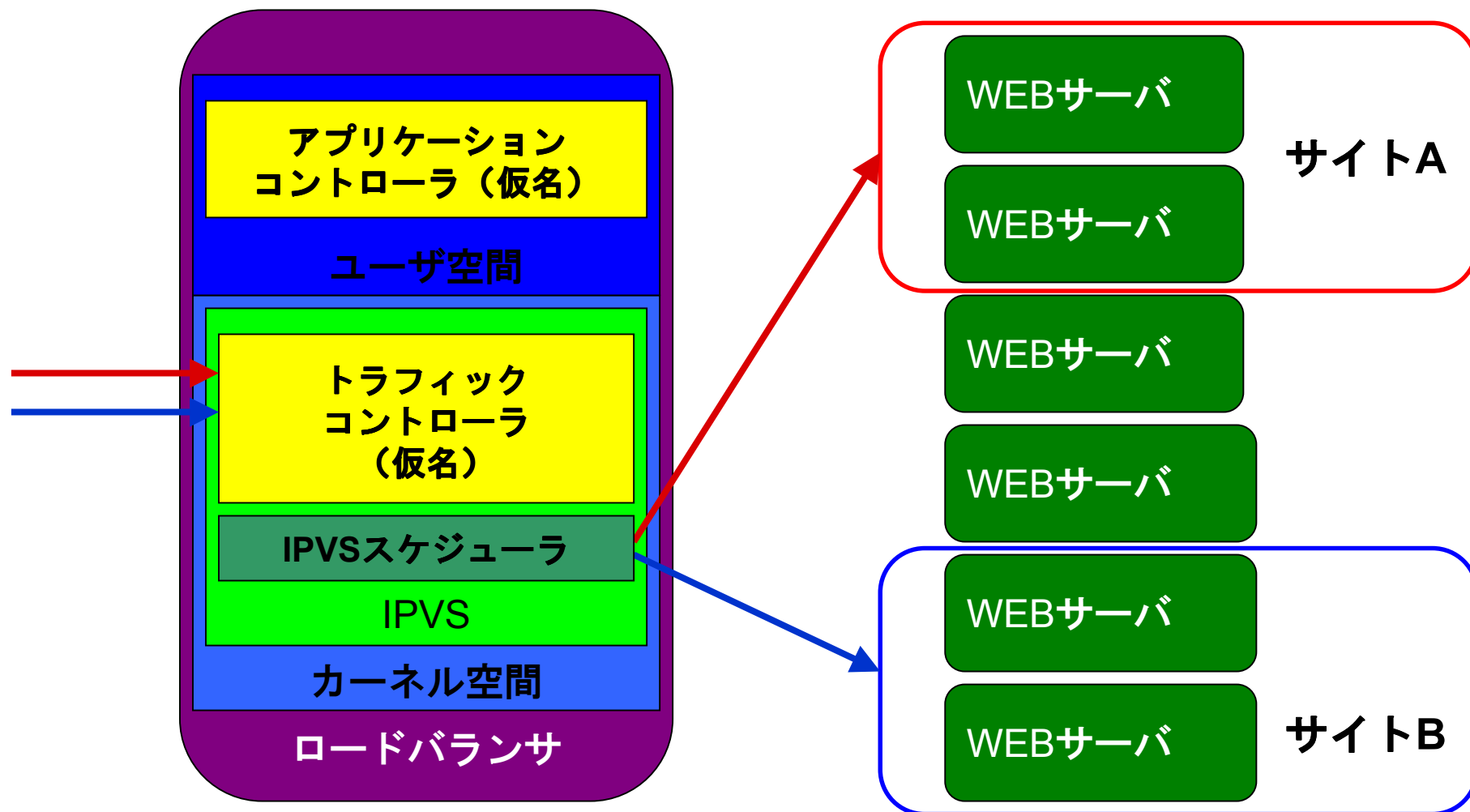
名前負けしてる気がします(^^;

**このままでは悔しいので、  
どうすれば実現できるか考えます**

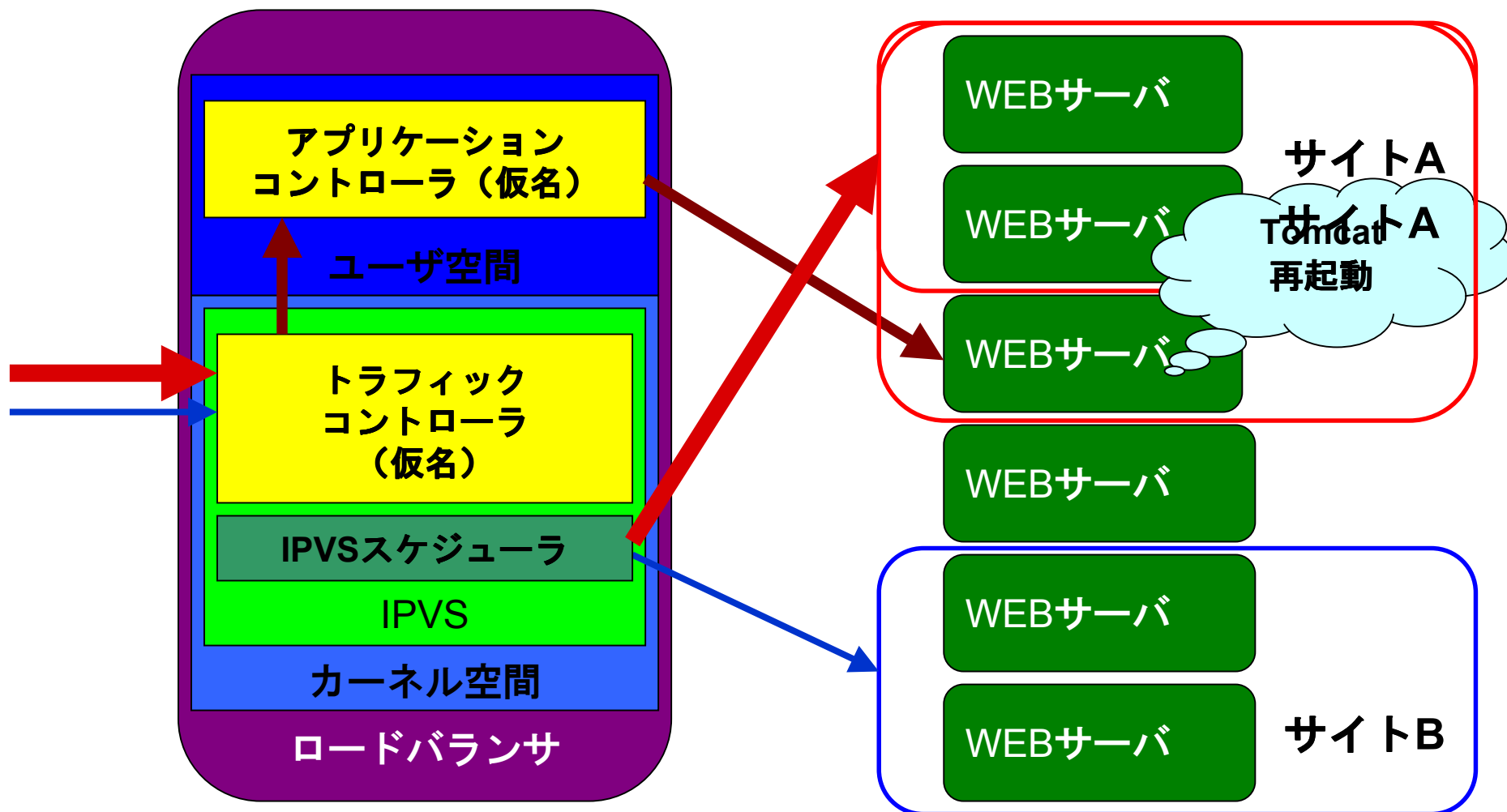
## 自動的にサーバを追加するには

- サイト毎のトラフィックの推移を記録しなきゃ
  - ロードバランサの機能を拡張するとできるかな
  - トラフィック測定はIPVSのモジュールでやるとよさそう
- サーバを増や前にTomcatを再起動しなきゃ
  - ロードバランサからTomcatを再起動するには
    - sshしてコマンド実行する？
    - httpなどで再起動リクエストを発行する？
  - どちらにせよカーネル空間でやるべき処理ではない
  - ユーザ空間で常駐するプロセス(デーモン)が必要

# 自動的にサーバを追加するには



# 自動的にサーバを追加するには





やればできそうな気がします

## 自動的にサーバを追加するには

しかし、懸念点や疑問もいっぱいできてきま

す

- IPVSの機能を拡張するにはどうすればいいの？

- Linuxのカーネルモジュールってどうやって作るの？
- カーネルモジュールとユーザプロセスの通信方法は？
- 人がいないところで勝手に構成が変わるのは怖い？
- すべてのサイトのトラフィックが同時に上がったら？

## 実現に向けてやるべきこと

- IPVSのソースを読んで構造を理解する
- とりあえずカーネルモジュールを作ってみる
- カーネルモジュールと通信するデーモンを作ってみる

## 普段やってること(再掲)

基本的に何をするかは自由です。

人にとってテーマや内容は様々です。

- Linuxのカーネルソース読んでみたり
- IPVSの拡張モジュール書いてみたり
- カーネルの内部情報を参照する方法を調べたり

などなど

この話はひとつの例ですが

普段やってる事はこんな感じです

最後にもうひとつ



## 実際に役に立った事例を紹介

# 少しでも過去を振り返ってみます

(WindowsNTでファイルサーバやDBサーバを運用してました)

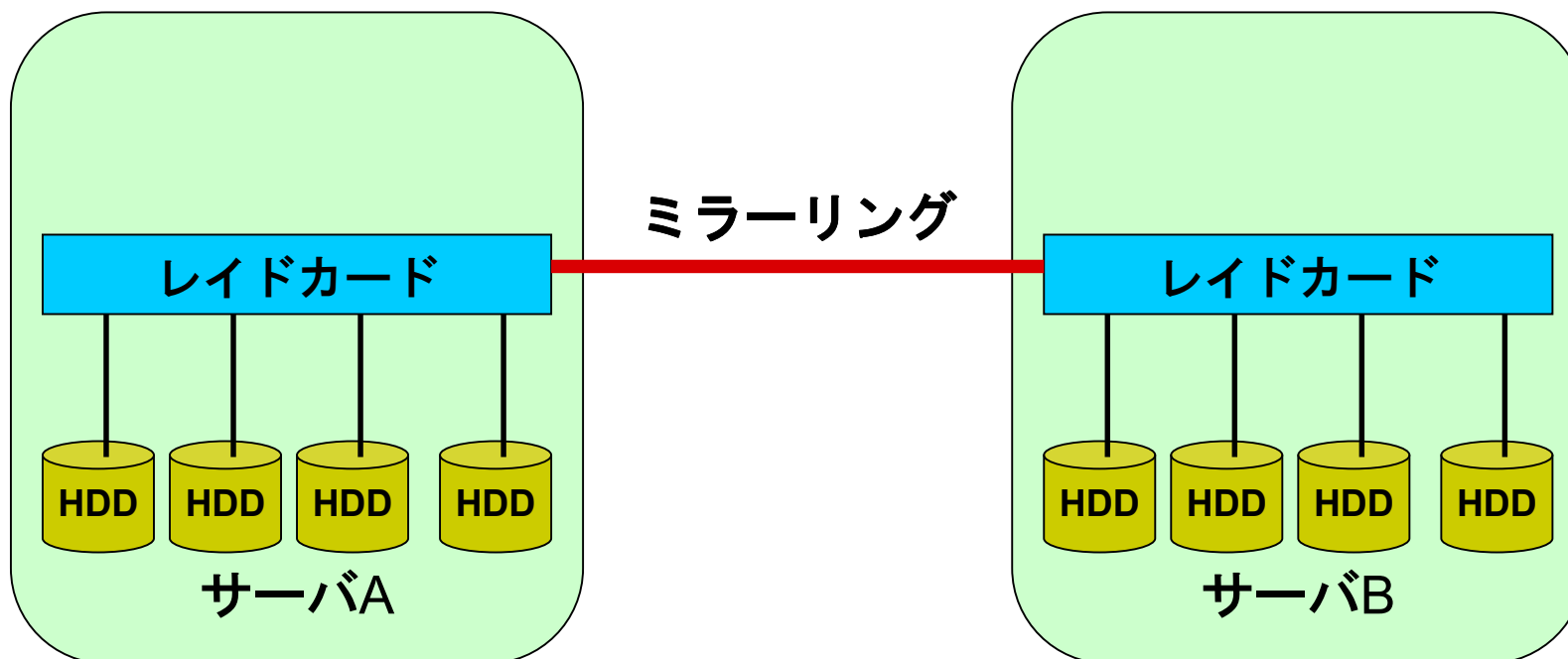
- **毎朝テープデバイスにバックアップをとっていた**
  - 当時はDATやDLTなどにバックアップをとるのが当たり前
  - テープの交換が超面倒！
  - デバイスエラーで固まることもしばしば！
  - 本当にリカバリできるのか不安で仕方がない！
- **サーバを買ってオンラインバックアップすることを考えた**
  - しかし、当時はサーバ1台買うだけでも結構大変
  - とても「バックアップ用にサーバ買って☆」とは言えなかった
  - バックアップ用途のためだけに保守費用が倍になるのは割に合わない

# 少しでも過去を振り返ってみます

- **定時バックアップには問題がある**
  - バックアップをとってる最中はDBサーバを止めなければいけない
  - オープン中のファイルはバックアップできない
  - もし夜中に障害が起こると、丸一日分のデータが消失する
  - この問題は誰もが認識していた
- **2台のサーバでリアルタイムにバックアップできれば、、、**
  - 片方が壊れてもすぐに切り替えることができるかも
  - 毎朝のバックアップが不要になるはず
  - バックアップのためにプロセスを止める必要もなくなるかな
  - これができれば、サーバを増設するメリットは十分にあるかな

## 実現する方法を模索(妄想)

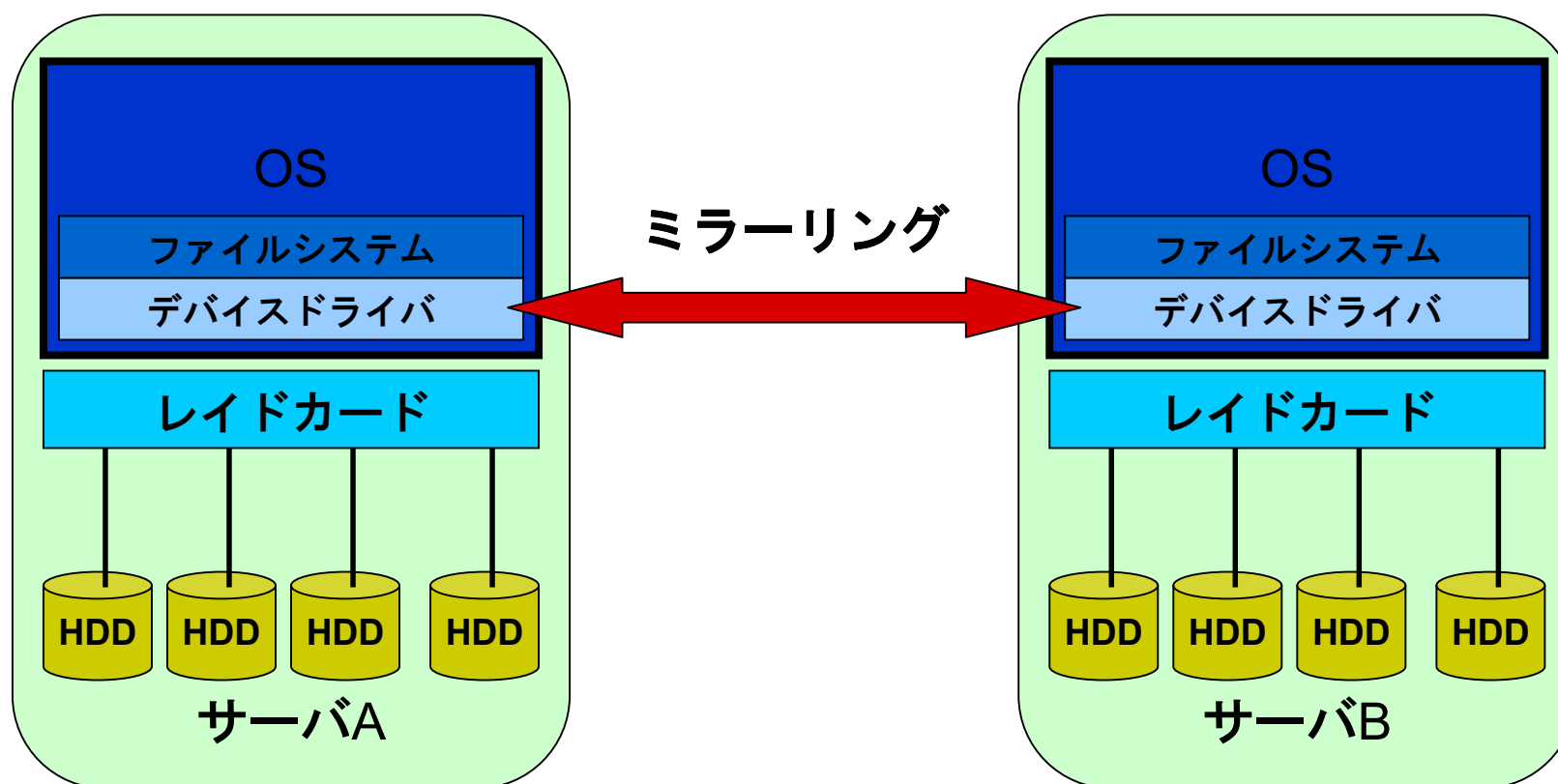
- ハードウェアを使った構成を考えてみました



**ニーズを満たす製品は見あたらず**

# 実現する方法を模索(妄想)

ソフトウェアでどうにかできないか考えてみました



**すでに同じ事を考えていた人が！**

# DRBDに遭遇

- **遭遇した当初は導入不可**
  - なにせWindowsサーバだったもので……
  - 自分のマシンで軽く動かして喜んでいました
- **Linuxでファイルサーバを構築する必要に迫られる**
  - DRBDの存在を知っていたので
  - システムの設計段階から導入を検討
  - 気合いをいれて検証を開始

**現在はDSASのネットワークストレージとして  
元気に稼働しています**



# まとめ

## ● 運用は楽しく

- 夢を見るのは素敵なことです(ただし睡眠時間が必要です)
- 実現する方法を模索(妄想)するのは楽しいです
- 日々の運用経験から、新しい仕組みが生まれます
- 単純作業の繰り返しだけでは飽きてしまいます

## ● 気が向いたら自分で実装してみるのもいいかも

- たとえだめもとでも、作ろうとすることで得るものがあります
- 将来実現できる環境が整ったときにその経験は役に立ちます
- もしかすると全く関係ない場面で役に立つかもしれません

**一見地味だけどクリエイティブなお仕事だと思いまっす↑**

結局言いたかったことは

「インフラって楽しいよ！」

今回の話はここまで

**ご静聴ありがとうございました**