

負荷量に応じて動的にサーバ台数と ルールを変更する 拡張ロードバランサと補助システム

20G477 畑智裕
最所研究室

用語解説

- キャッシュサーバ

- 外部に提供しているサービスやデータのキャッシュ(複製)を保存し, 提供するサーバ

- オリジンサーバ

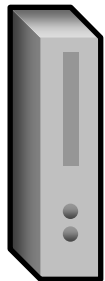
- 外部に提供しているサービスやデータのオリジナルデータを保持しているサーバ

用語解説

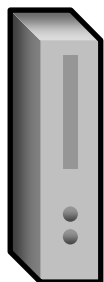
- スループット
 - 一定時間あたりのデータ処理能力
- スケールアウト
 - システムを構成する仮想マシン(サーバ)の台数を増やすこと
- スケールイン
 - システムを構成する仮想マシン(サーバ)の台数を減らすこと
- オートスケール
 - サーバ負荷に応じて、自動的にクラウドサーバの台数を増減させる機能のこと

ホストマシン3

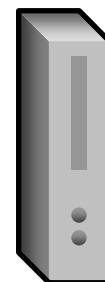
オリジンサーバ



仮想キャッシュサーバ



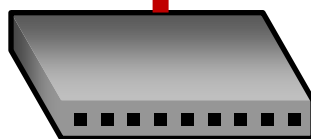
仮想キャッシュサーバ



...

ホストマシン2

リクエスト振分・負荷監視・キャッシュサーバ管理



ロードバランサ

ホストマシン1

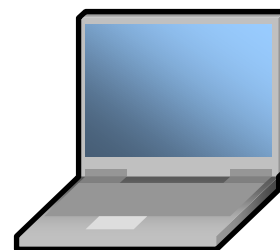
リクエスト



クライアント

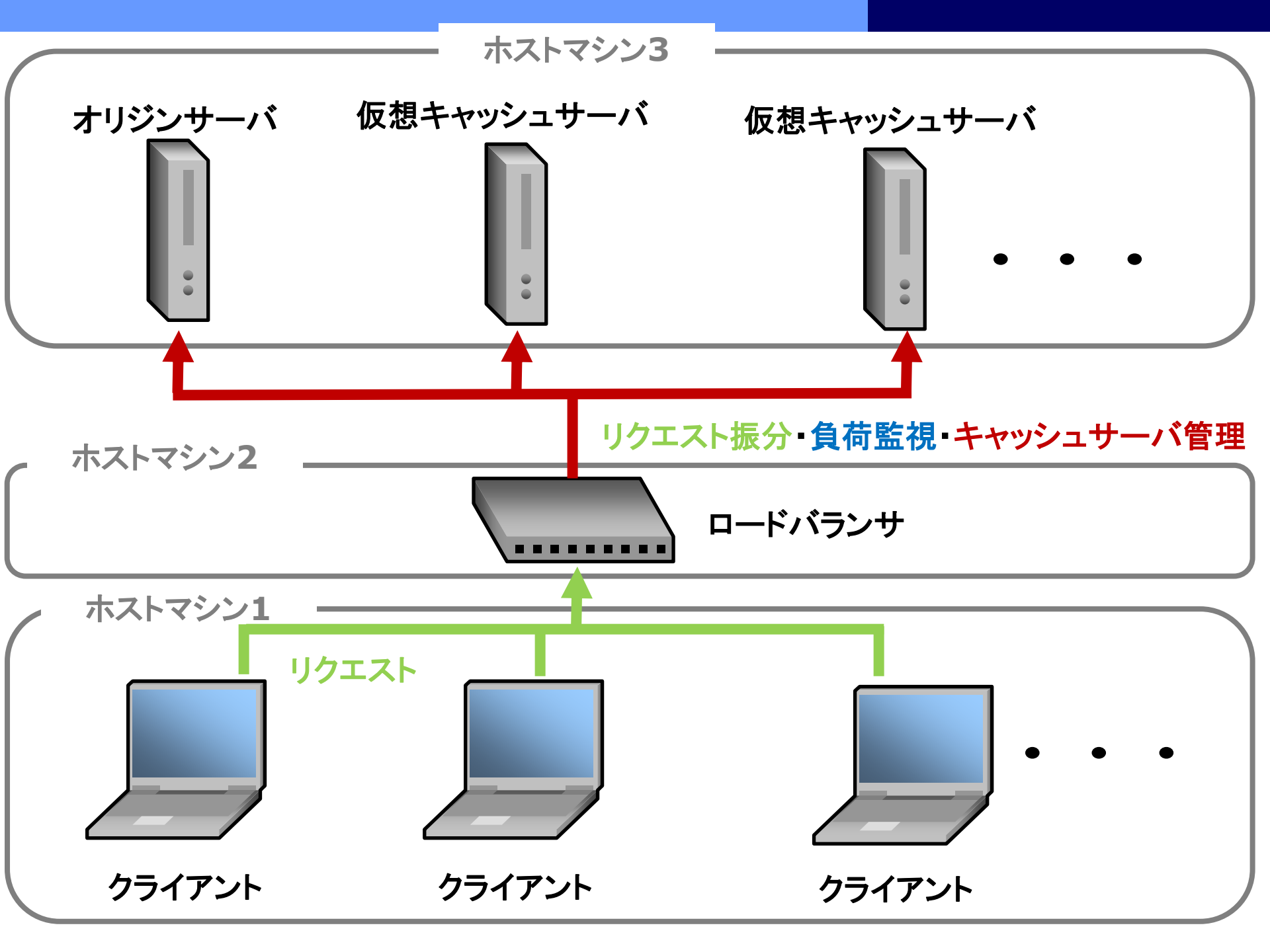


クライアント



クライアント

...



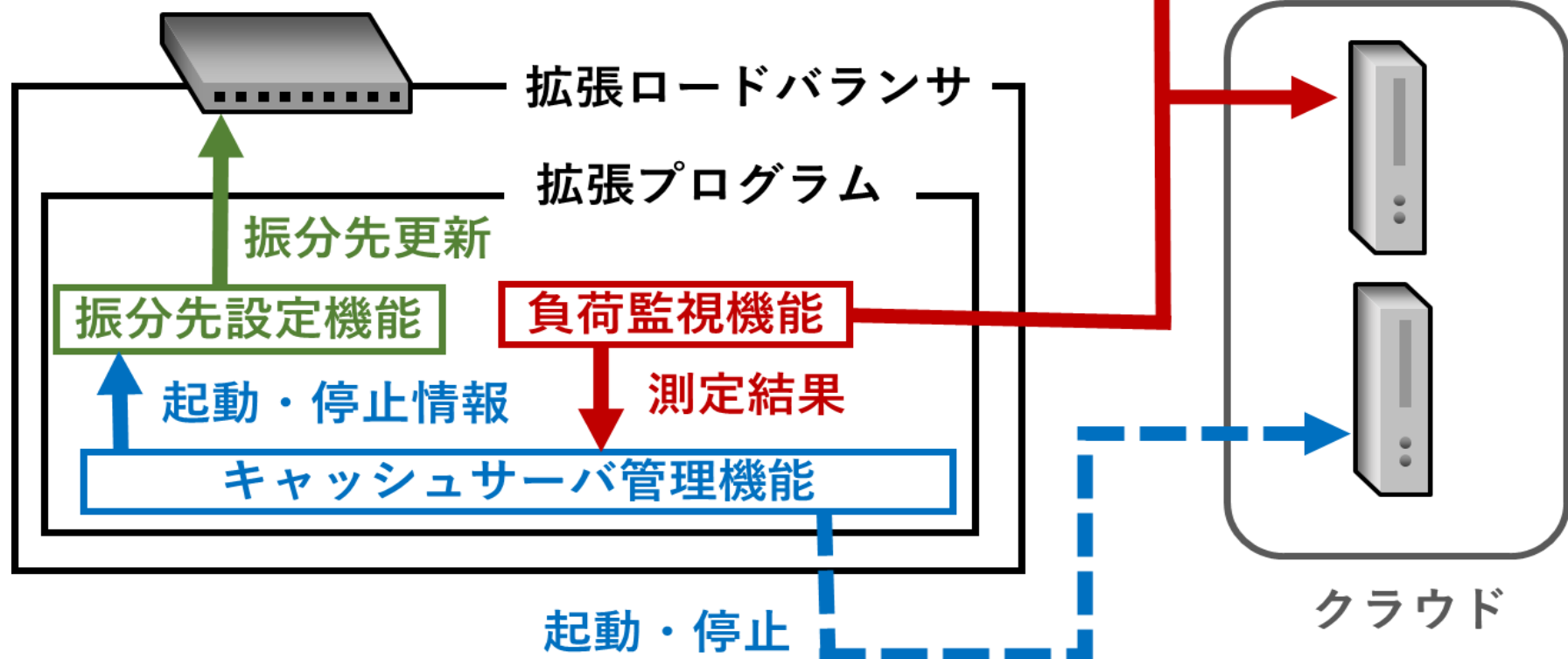
負荷監視機能: 各サーバの負荷量を測定

キャッシュサーバ管理機能:

負荷状況に合わせた起動・停止

振分先更新機能:

ロードバランサの振分先情報を更新



オリジンサーバ

負荷量測定

キャッシュサーバ

拡張ロードバランサ

拡張プログラム

振分先更新

振分先設定機能

起動・停止情報

負荷監視機能

測定結果

キャッシュサーバ管理機能

起動・停止

クラウド