

IoT時代のモヤモヤを自宅で解消

メカニズム丸見え! ラズパイ AIサーバを作る

ご購入はこちら

第7回 ロード・バランサでアクセスを分散する

土屋 健

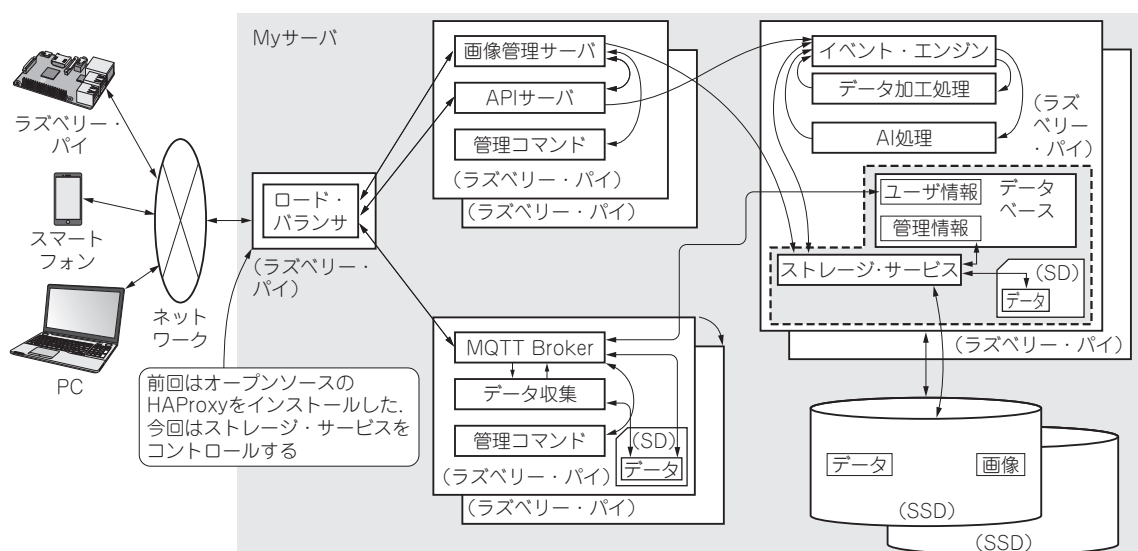


図1 前回作ったロード・バランサでストレージへのアクセスを振り分ける
連載ではAI搭載の本格サーバを自宅に構築する

ストレージ・サービスの構築は第5回(2017年8月号)までで行いました。前回(第6回, 2017年9月号)はオープンソースのロード・バランサをラズベリー・パイにインストールしました(図1)。

今回はロード・バランサ上でストレージ・サービスへのリクエストの振り分けが行えるような設定を追加します。

ロード・バランサの基本動作

ストレージ・サービスにロード・バランサを組み込む場合と組み込まない場合の動作について示します。冗長化済みストレージ・サービス(第5回までに作成した)が動作しているラズベリー・パイを2台、ロード・バランサ(前回作成)が動作しているラズベリー・パイを1台使う構成とします。

図2(a)はロード・バランサなしの場合の動作です。利用者はストレージ・サービスを構成するサーバを意識し、どちらかを選択してリクエストを送信します。

図2(b)はロード・バランサありの場合の動作です。ロード・バランサはストレージ・サービス1および2の状況に応じてリクエストを振り分けます。図2(b)はストレージ・サービス1が選択された場合です。利用者はロード・バランサにリクエストを送ることだけを考えればよく、その裏にあるサーバ構成については意識する必要がありません。

ストレージ・サービス1が故障した場合かつロード・バランサなし(これまでの)場合、利用者は自分でどちらかのサービスを選択してリクエストを投げます。応答がないとき、リクエストがエラーとなったことを自分で認識し、別のストレージ・サービスに再リクエストする必要があります。