Caching and High Availability

in the AWS environment

presentation by Patrick Neville

Intro

- Who am I ?
- Why use Caching?
- Why High Availability?
- Always think scalability!

Caching Options

Multiple Options Available

- Varnish Cache
- Nginx Cache-Control (Nginx-Core)
- Squid-Cache
- LiteSpeed Page Cache

Which Cache To Integrate?

This answer depends on two things.

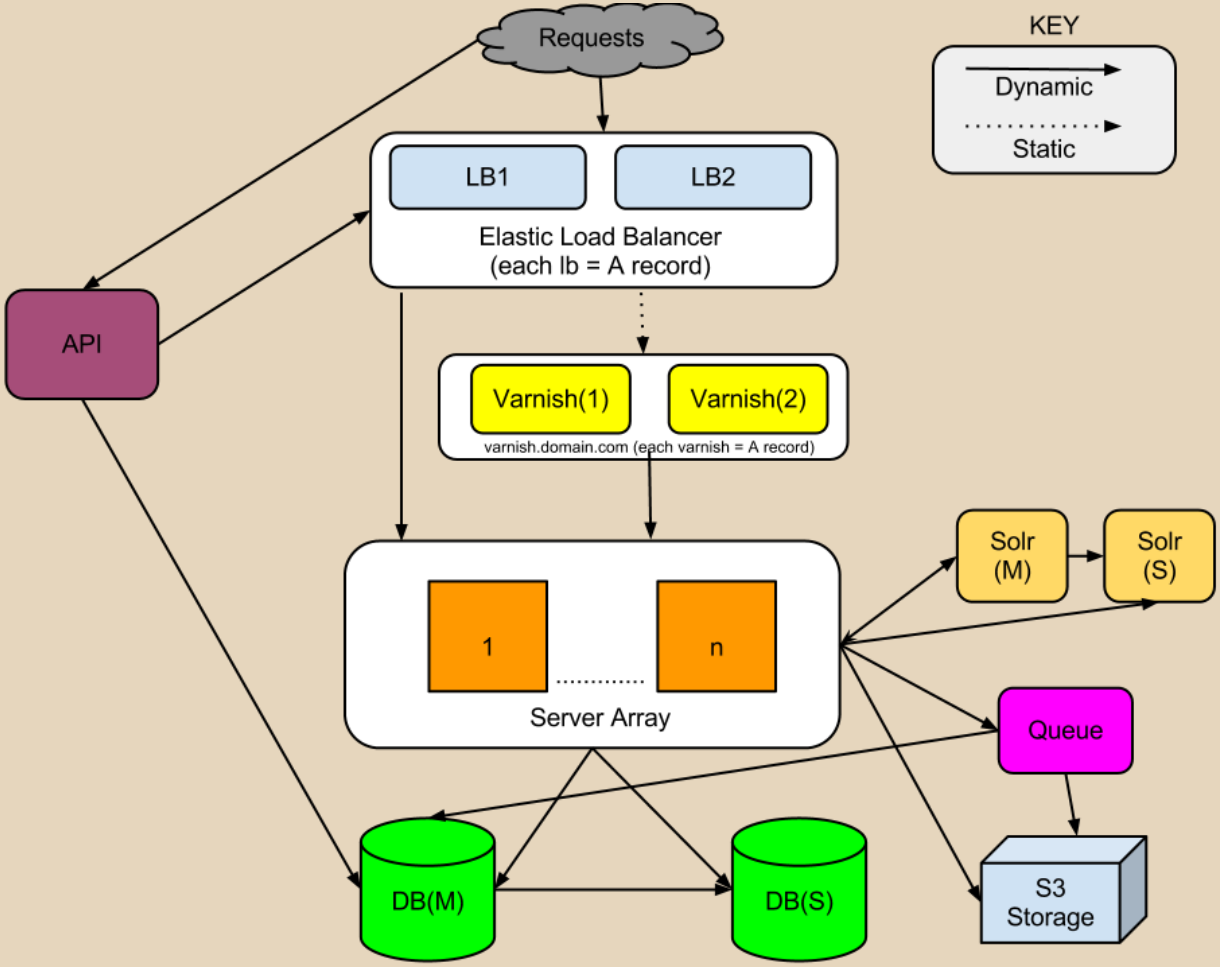
Application Environment:

- Server Farm (Cloud and Physical)
- Single Application (Cloud and Physical)

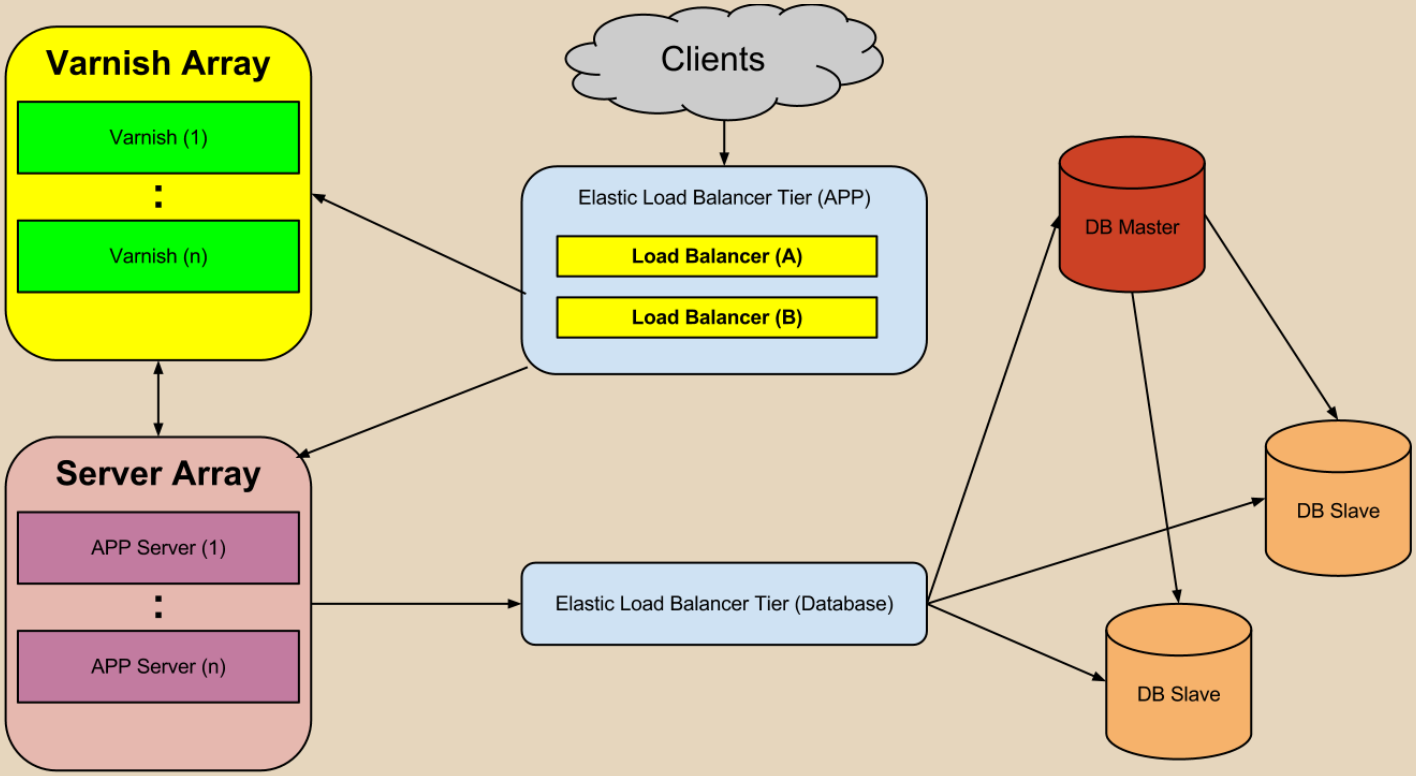
Application Need:

- Static and Pseudo Static Cache
- Web Acceleration

Integration of Varnish (Single App)



Integration of Varnish (Server Farm)



Outage-Proofing Practices

Resource Placement

Place in more than one zone:

- Load Balancers
- Cache Tier
- App Servers
- Databases

Maintain capacity to absorb zone/region fail

Replication and Failover

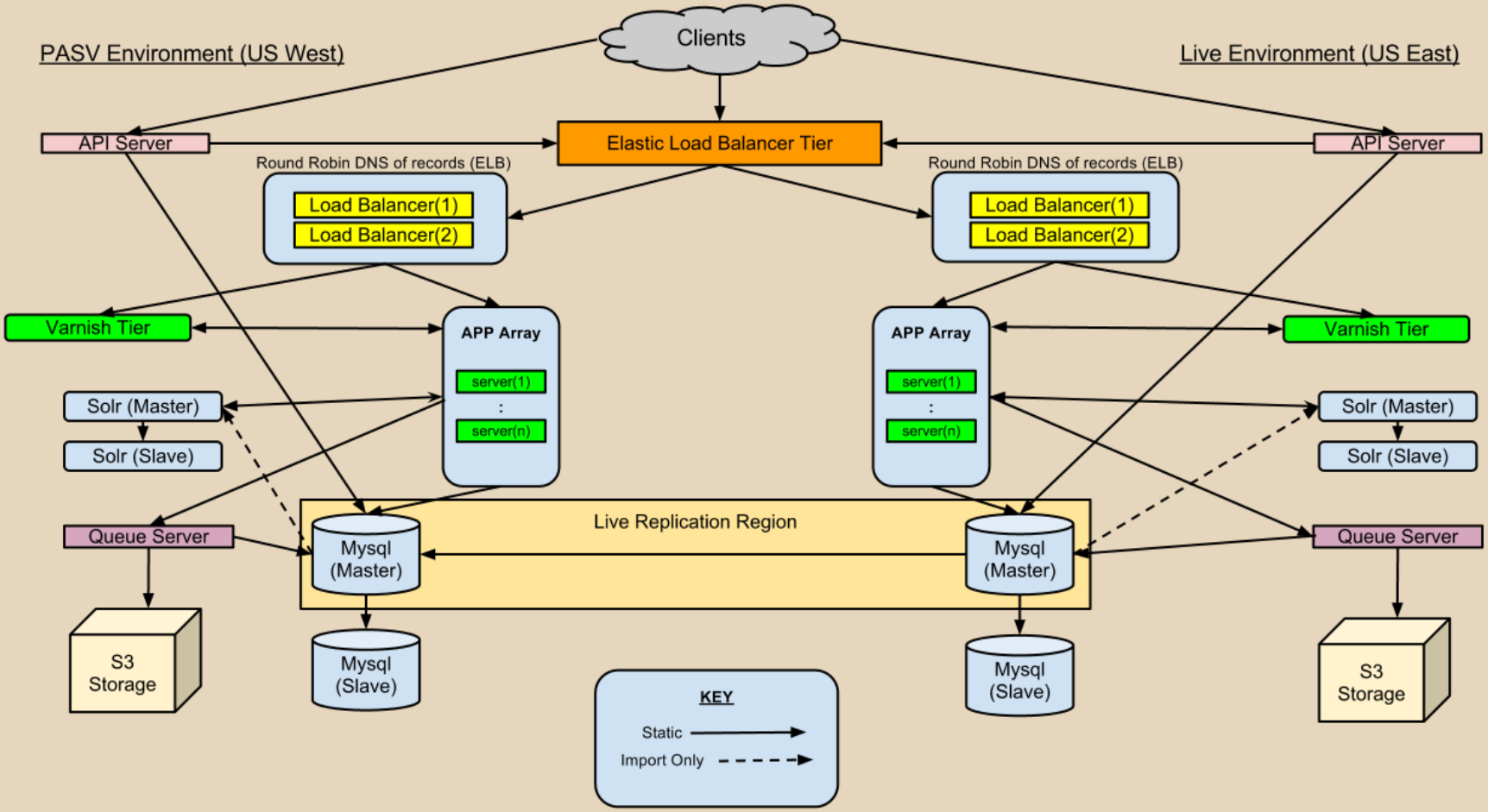
- Replicate data across zones
- Backup across regions
- Monitoring/Alerts

High Availability

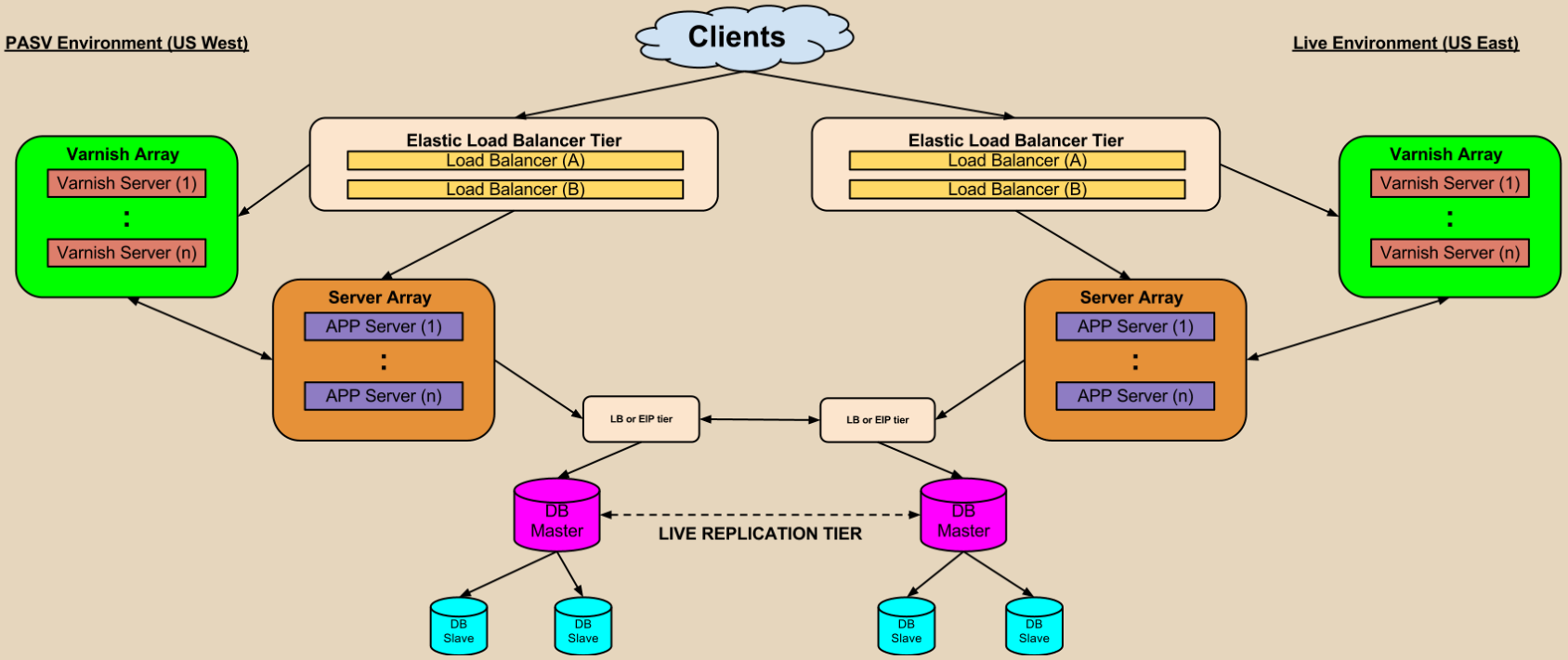
Different types of availability depend on the needs of your application.

- Failover Availability
- Geographic Availability
- Disaster Recovery
- ALL THREE !

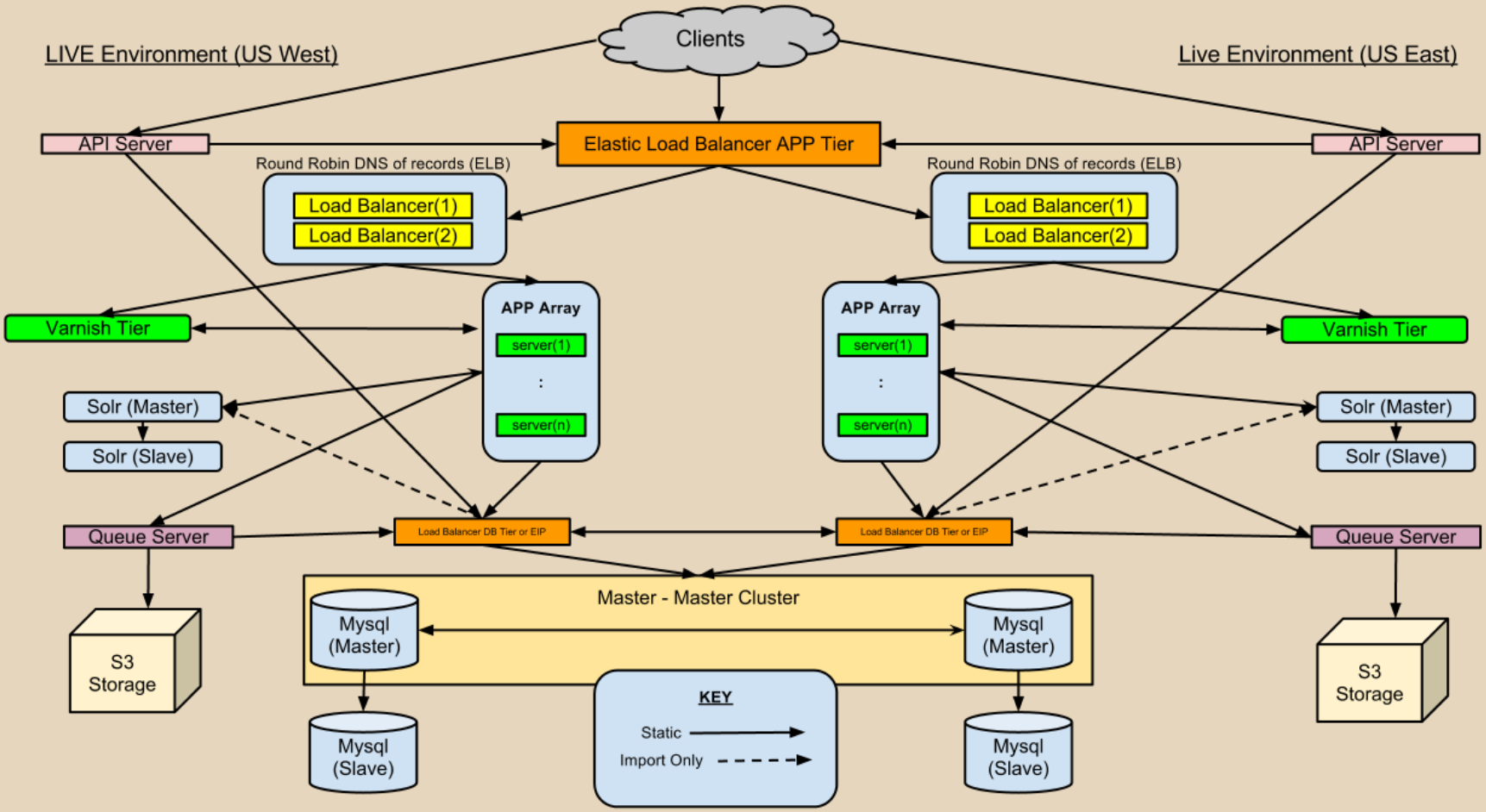
High Availability LIVE - PASV (Single)



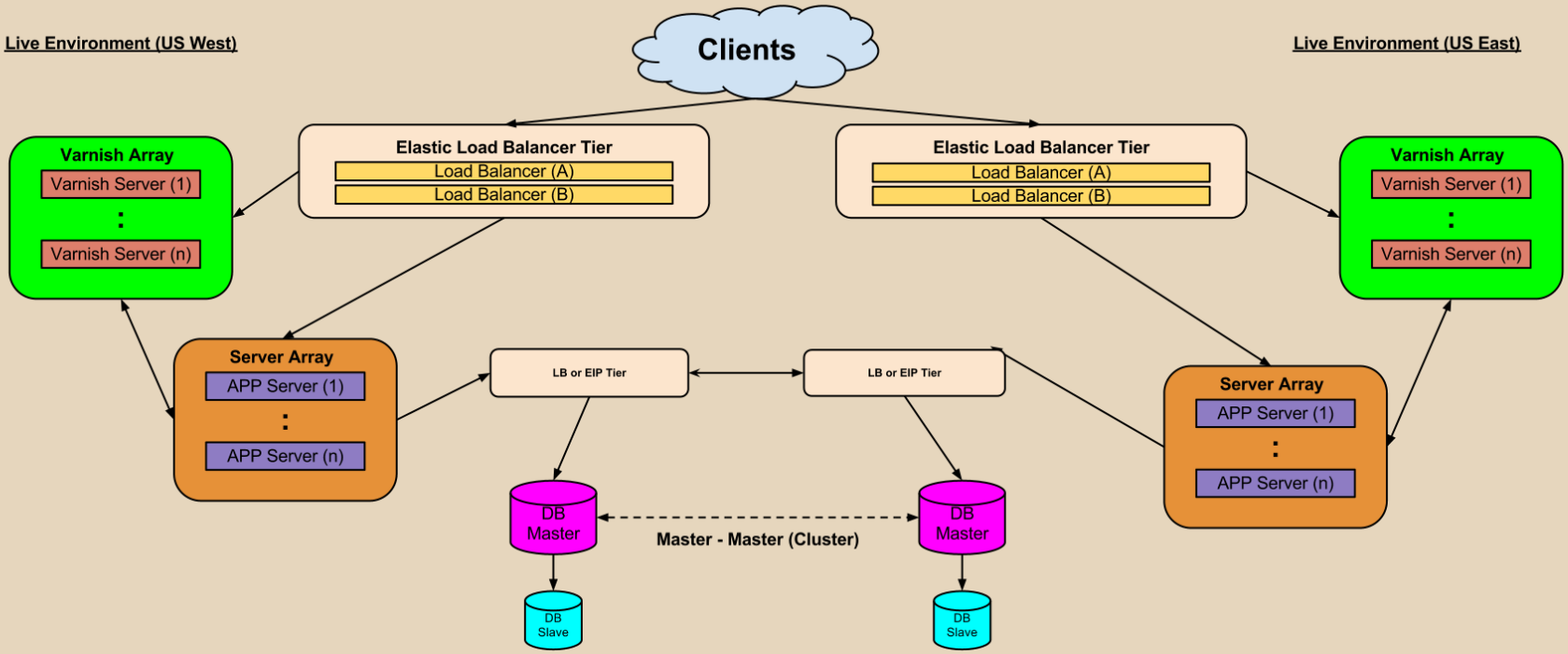
High Availability LIVE - PASV (Server Farm)



High Availability LIVE - LIVE (Single)



High Availability LIVE - LIVE (Server Farm)



High Availability Disaster Recovery

Options are based on data and needs of your application:

- Secondary Cloud Provider Options
- Data Retention and Storage Options
- Encryption and Accessibility Needs

Environment Choice

Which AWS option would suite scalability for the needs of the environment.

1. EC2 environment (Pros & Cons)
2. VPC environment (Pros & Cons)

Questions and Answers

Can't think of anything? Feel free to contact:

pat@theresumator.com