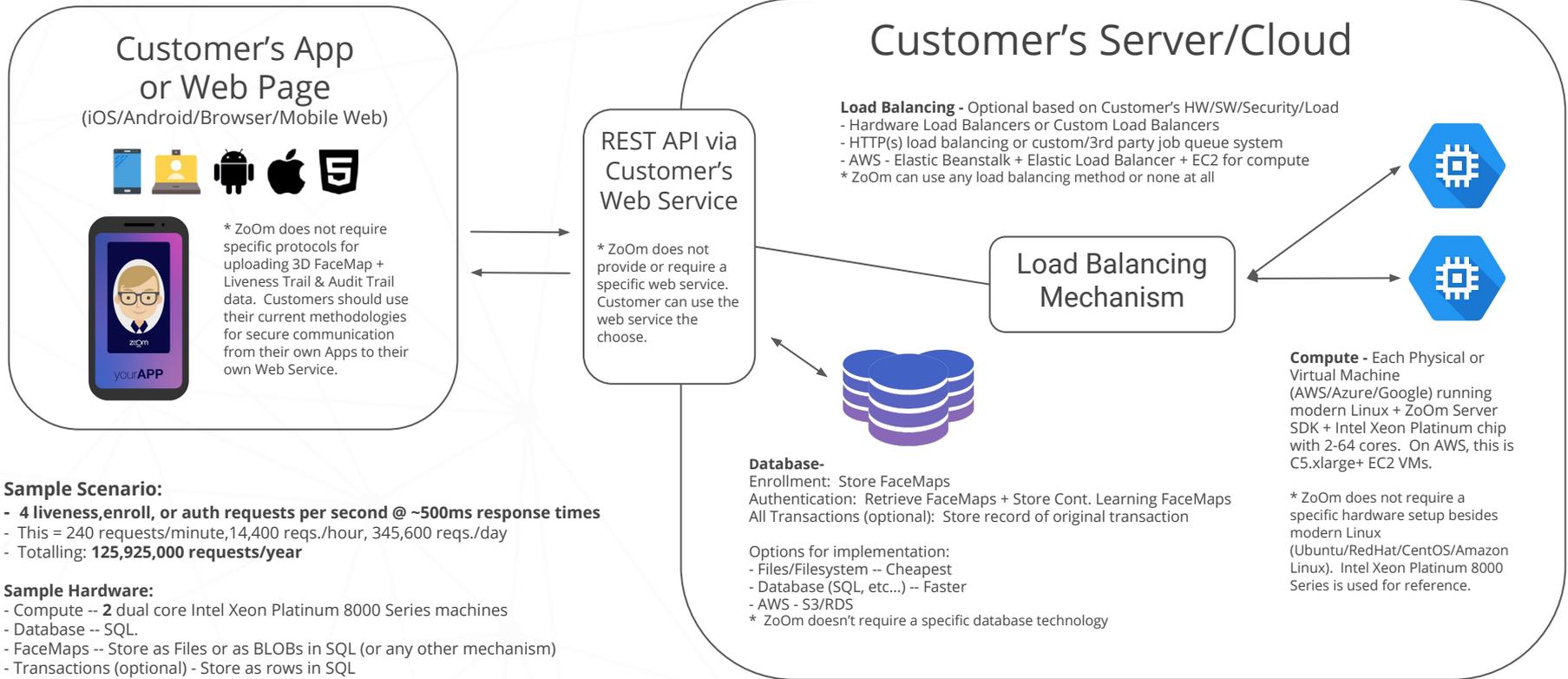




# Dedicated Server / Cloud Architecture

& Hardware Infrastructure Requirements





**Sample Scenario:**

- 4 liveness, enroll, or auth requests per second @ ~500ms response times
- This = 240 requests/minute, 14,400 reqs./hour, 345,600 reqs./day
- Totalling: **125,925,000 requests/year**

**Sample Hardware:**

- Compute -- 2 dual core Intel Xeon Platinum 8000 Series machines
- Database -- SQL.
- FaceMaps -- Store as Files or as BLOBs in SQL (or any other mechanism)
- Transactions (optional) - Store as rows in SQL
- Hardware Load Balancer (if necessary)
- **Cost with AWS = ~\$1000/year** (compute), but CPU, GPU compute costs drop annually.

**Calculating Your Infrastructure Requirements:**

- To double the requests/sec OR the geographic regions multiply everything by 2.
- To ensure resiliency to load spikes and prevent catastrophic failures, it is standard industry practice to multiply all hardware by 3
- To handle more than 1.5x peak load, more testing and analysis is required to determine additional hardware requirements
- GPU machines are more expensive, but can dramatically increase performance if response times of under 500ms under heavy load are required