

1. How much bandwidth do you need to handle calls simultaneously?

Number of users of the system: 10 million.

Assume 25% users make a phone call everyday = 2.5 million calls.

Assume each call to last 1 minute on average. That's 2.5 million minutes.

Number of minutes in a day = $24 * 60 \sim 25 * 60 = 100 * 15 = 1500$.

Number of calls at any point in time = $\text{total_talk_time} / \text{minutes_in_a_day}$

= $2.5\text{M}/1500$

= 2500 calls.

For PSTN calls, we need dedicated channels.

Each call will take about 64kbps. (Modems have 56 kbps)

Bandwidth has to be at least $64 \text{ kbps} * 2500 \text{ calls} = 64 * 2.5 \text{ Mbps} = 160 \text{ Mbps}$.

During peak hours, the load will be significantly higher. Assume $160 * 3 \sim \mathbf{500 \text{ Mbps}}$.

The bandwidth rate of PSTN is *quite high*. This will make for the bulk of the cost.

For VOIP calls, we can reuse channels.

Each concurrent call will take about 100kbps.

Bandwidth has to be at least $100 \text{ kbps} * 2500 \text{ calls} = 250 \text{ Mbps}$

During peak hours, the load will be significantly higher. To reduce network congestion, you may want to bump up the bandwidth.

Assume $250 * 4 \sim \mathbf{1 \text{ Gbps}}$.

Despite this being twice the requirement of the PSTN bandwidth, the VOIP lines are cheaper.

Total bandwidth requirement will be a mix of VOIP and PSTN. Depending on the user requirements, you can purchase the respective bandwidth.

2. How much RAM memory do you need for caching call states? What if you there is a national emergency which triggers a huge calling spike?

We mentioned having 2500 calls at any given point in time. Each call can be defined by it's ID, and some parameters like caller, callee, current state, start time, etc...

Let's assume we need to store 20 fields per call. Each field is a string of 20 characters.
~400 characters per call.

$2500 \text{ calls} * 400 \text{ characters per call} = 10^6 \text{ characters} = 1 \text{ MB}$

During a crisis, call volume can shoot up to 20-50x of the normal volume. In terms of RAM memory, that's just 50 MB. The requirement here is negligible.

3. Can you guess which part of the system will be most affected if call volumes shoot up?

A dramatic call volume increase most likely to affect all bottlenecks of the system. The call state manager will need to scale up, but that seems manageable. The bigger issue is the switch service.

The switch has to dynamically scale up to the huge spikey demand. We must have rate limiting and scaling measures in place for this scenario. Dynamically acquiring PSTN lines will be hard. This is the part that is most likely to fail.

4. What is the total amount of storage required for the entire system?

This is a hard estimate, since the internal services may be complex and we have not accounted for services such as log aggregators, rate limiters, analytics, etc...

The call state manager will need state transitions of each call recorded. Permanent.

The invoice manager will have the summary of each call. Permanent.

The switch will have detailed logs of state changes. Temporary.

We may need to record calls from suspicious parties for regulatory/compliance purposes.

The routing service will need service-provider data and call states to make decisions.

It's best to clarify what the interview wants. If they insist on the total storage, elaborate on each system and sum the components.

5. What is the total storage required to record calls?

2.5 million calls are made everyday. Of these, about 0.01% will be suspicious enough to track.

Hence we have $2.5M * 10^{-4} = 250$ calls to be intercepted.

Each call lasts a minute on average = 250 minutes of storage a day.

Assume a 1 minute wav file takes 10 MB storage.

Total storage to record calls per day = 10 MB per minute * 250 minutes = 2.5 GB everyday

If we assume 3 copies for fault tolerance and a life of 10 years for a recorded call, we get

$2.5 \text{ GB} * 3 * 365 * 10$

$\sim 2.5 \text{ GB} * 1000 * 10$

$\sim 30 \text{ TB}$

The total storage requirement for recorded calls is 30 TB.