Voice Verify scalable vs. non-scalable version

Voice Verify can be deployed in two variants:

1. Non-scalable
2. Scalable

The main differences are summarized in the following table:

–	Non-scalable	Scalable
# of servers/VMs needed	1	10+
# of parallel calls Voice Verify can handle	1 – 350*	unlimited
vertical scaling	✓	✓
horizontal scaling	×	✓
public domain needed**	×	✓
secured communication***	×	✓
accepts SIP calls	✓	×
accepts HTTP streaming	✓	✓
accepts WebSocket streaming	✓	✓
batch import of voice recordings	✓	✓
WebHook support	✓	✓

*in case a server/virtual machine with more than 50 CPUs is used – see Hardware requirements.
**public domain is needed only for secured communication in case no SSL certificate is owned
***secured communication is optional.

Typically, the scalable version is suitable for huge deployments with more than 350 calls processed in parallel.

Voice Verify deployment

1) Non-scalable version

Voice Verify is delivered as a single virtual machine. Both on-premise and cloud deployments are supported. We are able to provide the package for the following hypervisors:

1. VMware
2. Amazon Web Services
3. Microsoft Azure

Hardware requirements

Phonexia Voice Verify as a virtual machine needs only specific hardware for successful operation. The HW specification is as follows.

CPU

Phonexia technologies are optimized for INTEL CPUs. Recommended series are

- INTEL Xeon E5 generation 3 or 4
- INTEL Xeon Gold
- INTEL Xeon Platinum

The specific model selection depends on the expected traffic. To cover the peaks in the estimated load on Phonexia Voice Verify, the system needs enough dedicated CPU cores. A rough formula to calculate CPU sizing is that **1 CPU core can handle 7 concurrent calls**.

CPU cores are a narrow point for scaling. Other components are perceived by Phonexia as not that crucial or costly.

RAM

RAM required for the smooth operation of Phonexia Voice Verify also depends on the expected traffic.

1GB of RAM for 7 concurrent streams, plus 8GB for the whole system is a sufficient estimation.

Disk storage

There are two virtual disks required by Phonexia Voice Verify – a system disk and data disk.

- System disk
  - Requires 10 GB
  - Contains:
    - Voice Verify
    - License file
    - Audio Source Profile (= calibration profile)
- Data disk
  - Capacity is defined by the number of calls processed in parallel
  - Crucial from a Disaster Recovery (DR) perspective + updates
  - Contains
    - Voiceprints
    - Logs
    - PBX instances database

Logs are created during various activities by Phonexia Voice Verify (mainly API requests) and are deleted after some time (90 days). The amount of logs depends on the traffic.

The basic formula for the estimation of required disk capacity is dependent on the amount of audio processed by Phonexia Voice Verify. This formula is: **1 minute of 1 audio stream with usual usage** (2 verification queries per second) **creates 100kB of logs**.

As an example, **one stream running 24 hours a day straight** generates **15GB logs during 90 days**. This disk capacity is then required to keep all the necessary logs for this stream.

Networking requirements

The Voice Verify virtual machine needs to meet the following requirements:

- static IP address (typically an IP reservation on DHCP server)
- has to be reachable for API requests
- needs to be able to connect to a PBX (in case SIP calls are used)
- allowed ports
  - TCP 22 – SSH connection
  - TCP 80 – WebSockets, Kibana, Grafana
  - TCP 5060 – SIP
  - TCP 8000 – Voice Verify
  - UDP 20000-20350 – RTP
- domain requirements

- the customer chooses a domain name on which he wants to run Voice Verify
- The required DNS configuration for domain „mydomain.com“ is in the next table:

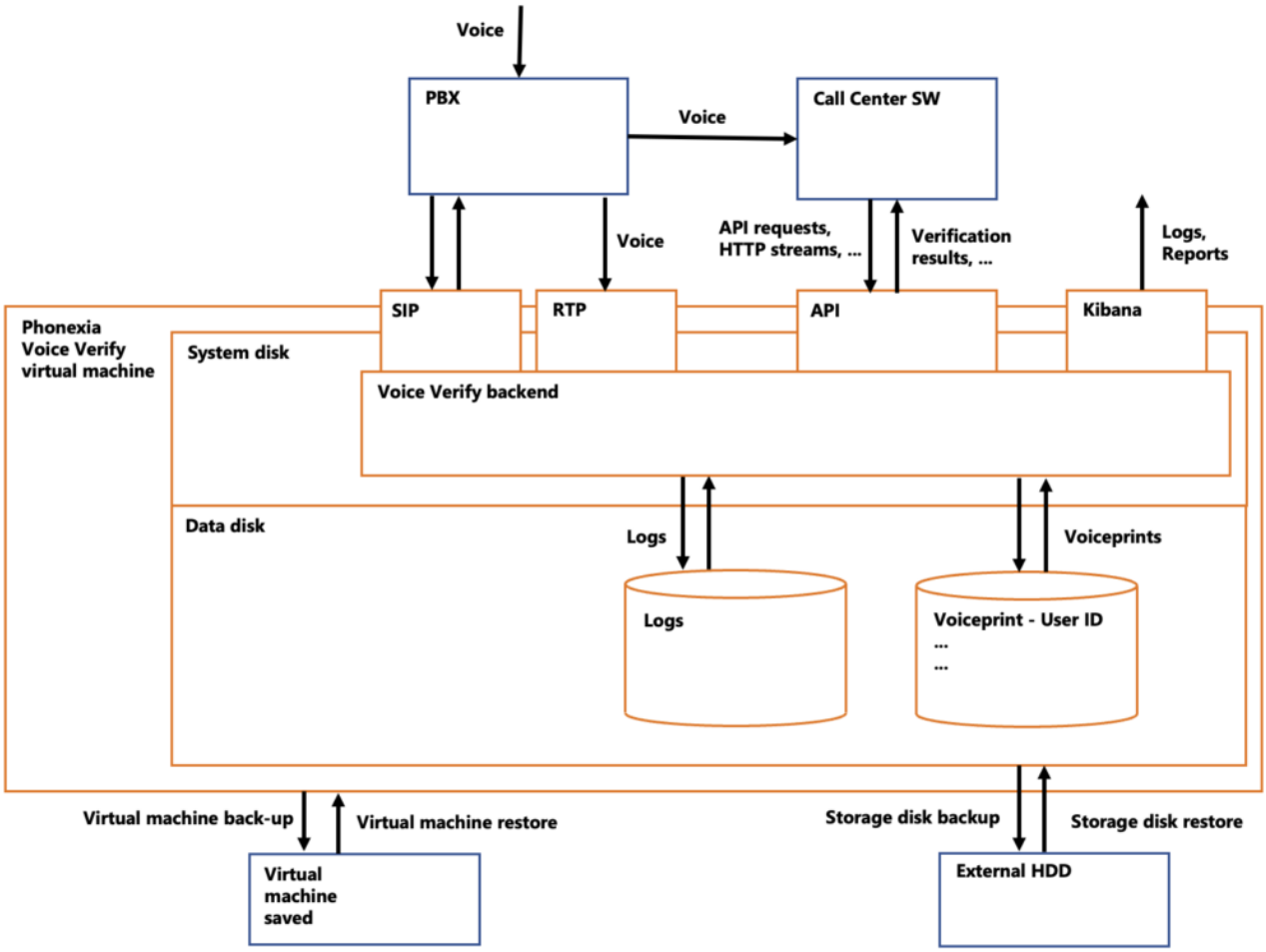
Record	Type	Value
mydomain.com	A	PUBLIC IP ADDRESS
*.mydomain.com	CNAME	mydomain.com

The domain does not necessarily need to exist, it is possible to include the domain name and corresponding URLs into `hosts` file
(Windows `C:\Windows\System32\drivers\etc\hosts` or Linux/Mac `/etc/hosts`).
Example:

```
<IP_ADDRESS> mydomain.com
<IP_ADDRESS> websocket.mydomain.com
<IP_ADDRESS> voiceverify.mydomain.com
<IP_ADDRESS> elasticsearch.mydomain.com
<IP_ADDRESS> kibana.mydomain.com
<IP_ADDRESS> grafana.mydomain.com
```

where `<IP_ADDRESS>` is the IP address of the virtual machine.

Architecture



2) Scalable version

The Voice Verify scalable version needs at least 10 virtual machines (or physical servers) in order to run. The solution consists of several components designed for high availability (for most of the components) and for scalability. Both on-premise and cloud deployments are possible. The deployment process is semi-automatic and it requires cooperation between the customer's and Phonexia's DevOps Engineers, who perform deployment to the customer's environment.

Hardware requirements

The customer defines the maximum expected load for Voice Verify. Phonexia's DevOps Engineers calculate the server requirements, including the count of virtual machines/servers and specifications for CPU, RAM and HDD.

CPU

Phonexia technologies are optimized for INTEL CPUs. Recommended series are

- INTEL Xeon E5 generation 3 or 4
- INTEL Xeon Gold
- INTEL Xeon Platinum

Example

Configuration for 500 parallel calls:

Virtual machine/server count	CPU	RAM	HDD
4x	4	8	30 GB
18x	8	16	30 GB
3x	4	4	30 GB

In total, 25 virtual machines/servers with 172 CPU cores, 332 GB RAM and 750 GB HDD.

Networking requirements

Network communication is secure, all endpoints use certificates on customer defined subdomain. Most services use common HTTP on TCP 80.

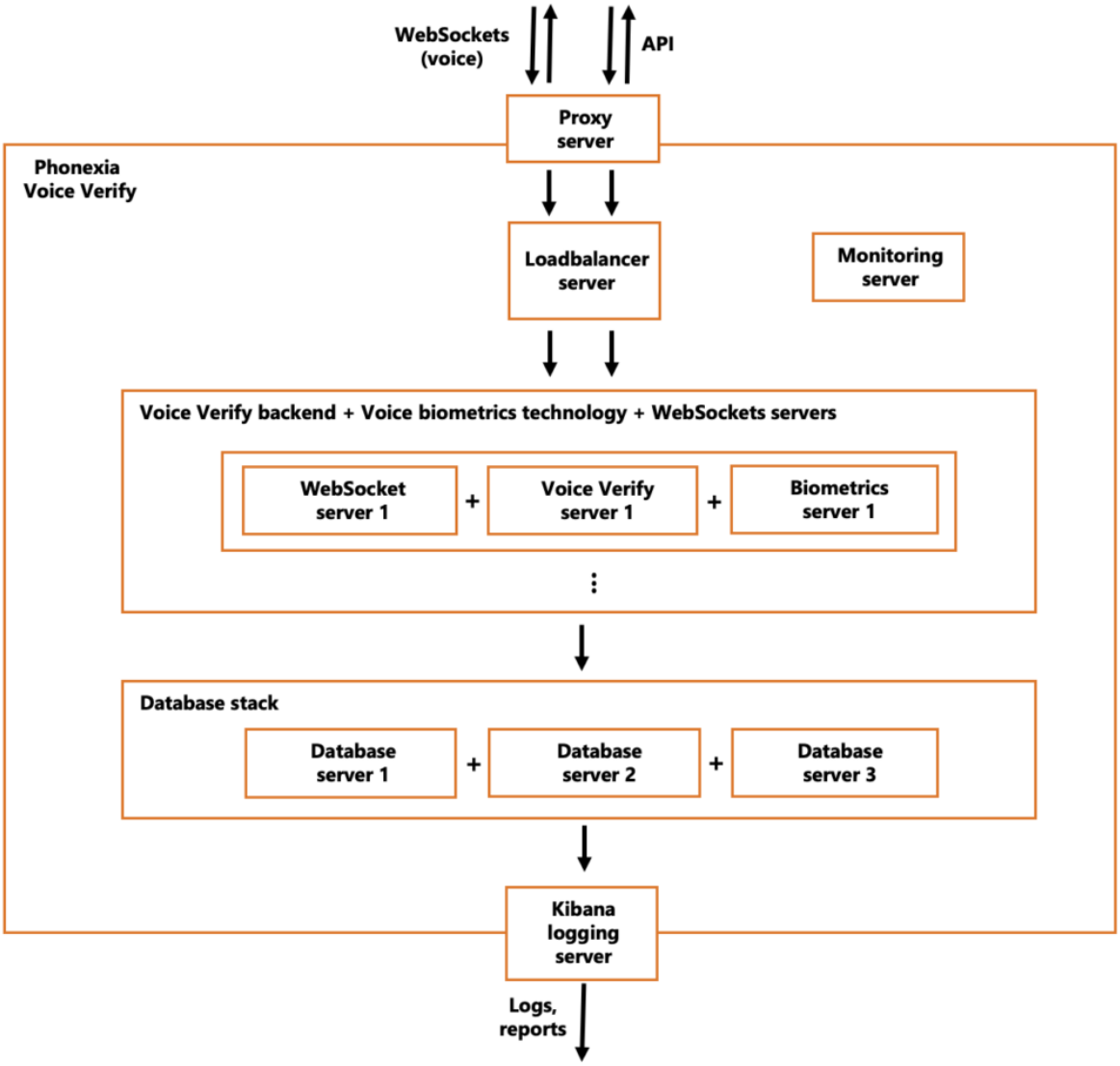
- all servers must have an internal static IP address and they have to be able to communicate with each other
- one of them must be publicly accessible with a static public IP address
    - the customer chooses a domain name on which he wants to run Voice Verify
    - this domain must be redirected to the static public IP address mentioned above
      The required DNS configuration for domain „mydomain.com“ is in the following table:

Record	Type	Value
mydomain.com	A	PUBLIC IP ADDRESS
*.mydomain.com	CNAME	mydomain.com

- SSH access to virtual machines/servers (for deployment and updates only)
- all servers must be deployed to the same subnet
- allowed ports
    - inside the subnet
        - TCP port 2377

        - TCP and UDP port 7946
        - UDP port 4789
    - incoming communication from the public – only to the public IP address defined above (server/virtual machine hosting proxy server)
        - TCP 80 – HTTP
        - custom
            - TCP <custom_port_1> – can be set for WebSocket connection
            - TCP <custom_port_2> – can be set for Voice Verify API
            - by default, both of these run on TCP 80

Architecture



To obtain more detailed information about the scalable Voice Verify infrastructure, please contact Phonexia’s consulting team.