]pexip[

<customer> Solution Design (Infinity Only)

Prepared by: <sa> – <title>

<date>

Solution Design Summary

The purpose of the solution design document is to complement the Pexip Scope of Work (SOW) and provide <customer> insight into deployment model, use cases, requirements, and user acceptance testing for a successful implementation. Examples are provided to assist <customer> with both configurations and user acceptance testing. Links to Pexip's robust documentation library have been provided, which are relevant to the solution design, further enhancement, and growth. Additional information on Pexip Infinity, integrations, and more can be found at: https://docs.pexip.com

Infinity Design Introduction

Pexip Infinity is an award-winning, self-hosted, virtualized, and distributed multipoint conferencing platform. It can be deployed in an organization's own data center, or public cloud such as Microsoft Azure, Amazon Web Services (AWS), or Google Cloud Platform (GCP), or in a hybrid combination of on-premises and cloud. It enables scaling of video, voice, and data collaboration across organizations, enabling everyone to engage in high-definition video, web, and audio conferencing. This document provides a high level of the services and deliverables that Pexip shall provide to the customer on the solution design. It also defines the customer's responsibilities in completing the project. This document is designed to be a general outline with more specific tasks, deliverables, and timelines to be discussed and agreed upon during the initial project kick-off meeting and other pre-installation meetings.

Executive Overview

<customer> will use Pexip Infinity to provide video collaboration between participants and standard-based endpoints. Pexip allows <customer> to embed a best-in-class video solution into their existing workflows and provide endpoint registration. By providing a robust suite of open APIs and SDKs, Pexip enables <customer> to create custom video workflows for their staff and customers alike. Pexip's infrastructure enables customers like <customer> to quickly scale and provide services during high-demand times. Pexip provides an easy-to-join process for users, leveraging a browser webapp or joining with standard-based endpoints. <customExecSection>

Solution Overview

<customer> provided background: Pexip Infinity will be installed on <platform> and provide <customSO>

- CPU models available:
  - <cpu_model>
- Management Node Specs: <mgt>

]pexip[

- Primary Location: <TBD>. Calculations are performed by: (vCPU*cpuGHz) / 1.4 GHz = HD calls (720p). We will use 720p as a baseline for the calculations below
    - Primary Location: <location1>
    - <location1>
        - Conference Node Small Specs: <cnsm1>
        - Conference Node Medium Specs: <cnmed1>
        - Conference Node Large Spec: <cnlg1>
        - Proxy Edge Node Small Specs: <edgesm1>
        - Proxy Edge Node Medium Specs: <edgemed1>
    - <location2>
        - Conference Node Small Specs: <cnsm2>
        - Conference Node Medium Specs: <cnmed2>
        - Conference Node Large Specs: <cnlg2>
        - Proxy Edge Node Small Specs: <edgesm2>
        - Proxy Edge Node Medium Specs: <edgemed2>
- Background
    - <hypervisor>
    - Registration for up to <endpoints>
        - OTJ for up to <otj> endpoints
    - (<cap>) concurrent capacity
    - Integrations with Pexip Infinity:
        - Cisco UCM: <ucm>
        - Cisco VCS: <vcs>
        - Poly: <poly>
        - Genesys: <genesys>
        - PSTN: <pstn>

Pexip's capacity planning takes into consideration a variety of factors such as:

- Server capacity and hardware (cpu type)
- Type of call: Full HD, HD, SD, or audio only
- Number of unique VMRs
- Type of gateway call: Inbound/Outbound legs on same transcoding node or not
- Compared to a single HD 720p call:
    - Full HD 1080p call uses twice the resource, SD uses half, and audio-only use 1/16
- Backplane reservation: Each conference instance on each transcoding node reserves a backplane connection at a resource level corresponding to the conference Maximum call quality setting, to allow the conference to become geographically distributed if required.

Pexip NOTED: *Based on information provided by <customer> above, <pexipNoted>*

]pexip[

Pexip recommended host server specifications: We recommend 3rd or 4th generation Intel Xeon Series processors (Ice Lake / Sapphire Rapids). Earlier Intel Xeon Scalable Series processors and Intel Xeon E5/E7 -v3 and -v4 series processors are also supported where newer hardware is not available. We recommend prioritizing the newer hardware for transcoding nodes. Additionally, if Hyper-Threading is in use, we recommend 2.6GHz (or faster) base clock speed on 3rd gen Intel Xeon Scalable Series (Ice Lake) or newer. A 2.8GHz+ for older Intel Xeon processors where Hyper-Threading is in use and 2.3GHz where Hyper-Threading is **not** in use. When Hyper-Threading is in use, we recommend that conferencing nodes are NUMA pinned to their sockets to avoid memory access bottlenecks. 1GB RAM for each vCPU that is allocated to the Pexip nodes. CPU and RAM must be dedicated and the population of all memory channels. For VMware, we support ESXi 6.7, 7.0. and 8.0; KVM. No oversubscription if running in EVC mode. AVX2 or AVX512 instruction sets for conferencing nodes. Although the conferencing nodes will normally not use more than 1-2 Mbps per video call, we recommend 1Gbps network interface cards or switches to ensure free flow of traffic between Pexip Infinity nodes in the same data center. We do **not** recommend 100 Mbps NIC. Management node and conference node disks should be thick provisioned. RAID 1 mirroring for disk redundancy and the RAID controller must have enabled cache. Additional information may be found at:
https://docs.pexip.com/server_design/server_design_intro.htm

Quantities based on solution design:

* *Quantities assume the availability of free space on the host server. Proxy Edge node servers should be placed in a DMZ. We have recommended an optional failover node for xCode and Proxy Edge, these will serve as a failover for full redundancy, in which, in the case of any one node failure, the remaining nodes can serve as full-capacity*

Image 1: Sample Compute Layout

Pexip components descriptions

- 
- Management Node: Required (1)
  - Administrative interface
  - Does not handle signaling or media
  - Platform and conference Status
  - Robust API suite (config, status, history, command,…)

- 
- Conference Node (Transcoding role): Required (1:Many per deployment)
  - Signaling and media processing
  - Distributed and scalable
  - Robust API suite

- 

]pexip[

- Conference Node (Proxy edge role): Required (1:Many per deployment)
    - Signaling
    - Forwards media to transcoding
    - Distributed and scalable
    - Robust API suite

High-Level Solution Overview Section

Image 2: VMWare. **NOTE:** DMZ Proxy edge can support single NIC with IPv4 NAT or dual NIC setups

Image 3: Clients to Conference Node – **Required for connections**

Image 4: Administrator to Management Node – **Required for access/provisioning**

Image 5: Optional Authentication Methods to Management Node

Image 6: Optional ServicesA

Image 7: Optional ServicesB

- In general, Pexip Conferencing Nodes should be considered compute intensive, and Management Nodes reflect a more general-purpose workload

Required: DNS

DNS will allow video clients to find conference nodes upon lookup. This section will show examples of A-Records and SRV records, but there may be additional requirements depending on the final scope of the project. It is highly recommended to review the DNS guidelines KB article on docs.pexip.com. **NOTE for Poly device registration:** Poly requires that NAPTR records be created to allow for registration to multiple nodes for redundancy, which differs from the priority method Poly had with Clarti infrastructure

The following types of video clients will require DNS to be set up for the system to function.

- SIP and H.323 endpoint registration messages to be routed to Pexip
- Calls from SIP and H.323 endpoints to be routed to Pexip Priv
- Pexip Connect clients to make calls to Pexip

The system will need A Records set up for the initial infrastructure.

Example A-Records

DNS A-Records	Public DNS Records for each public-facing node in the solution
Conf-01	*conf-01.foo.mycompany.com → 10.22.4.15*
Edge-01	*Edge-01.foo.mycompany.com → Public IP or 1:1 NAT IP*
Pex-Mgr	*pex-mgr.foo.mycompany.com → 10.22.4.4*

NAPTR (Optional)	Used for Poly device registration
vc.example.com	*Points to a SRV domain (vc.example.com) for services (SIP or SIPs)*

If your deployment will include communication from standards-based endpoints such as SIP or H.323, Pexip will also require SRV records. This is to allow signaling and media services to transmit on specific ports. As a general rule, every Edge Node and Conference Node should be configured with SRV records, although this may vary depending on the situation. Weighted DNS records can also play into an HA/DR strategy – Please ask your Pexip Engineer for best practices on this.

]pexip[

Example Internal SRV-Records

Name	Service	Protocol	Priority	Weight	Port	Target host
vc.example.com.	h323cs	tcp	10	10	1720	*conf-01.foo.mycompany.com*
vc.example.com.	h323cs	tcp	10	10	1720	*conf-02.foo.mycompany.com*
vc.example.com.	h323ls	udp	10	10	1719	*conf-01.foo.mycompany.com*
vc.example.com.	h323ls	udp	10	10	1719	*conf-02.foo.mycompany.com*
vc.example.com.	h323rs	udp	10	10	1719	*conf-01.foo.mycompany.com*
vc.example.com.	h323rs	udp	10	10	1719	*conf-02.foo.mycompany.com*
vc.example.com.	sip	tcp	10	10	5060	*conf-01.foo.mycompany.com*
vc.example.com.	sip	tcp	10	10	5060	*conf-02.foo.mycompany.com*
vc.example.com.	sips	tcp	10	10	5061	*conf-01.foo.mycompany.com*
vc.example.com.	sips	tcp	10	10	5061	*conf-02.foo.mycompany.com*
vc.example.com.	sip	udp *	10	10	5060	*conf-01.foo.mycompany.com*
vc.example.com.	sip	udp *	10	10	5060	*conf-02.foo.mycompany.com*

Example External SRV-Records

Name	Service	Protocol	Priority	Weight	Port	Target host
vc.example.com.	h323cs	tcp	10	10	1720	*edge-01.foo.mycompany.com*
vc.example.com.	h323cs	tcp	10	10	1720	*edge-02.foo.mycompany.com*
vc.example.com.	h323ls	udp	10	10	1719	*edge-01.foo.mycompany.com*
vc.example.com.	h323ls	udp	10	10	1719	*edge-02.foo.mycompany.com*
vc.example.com.	h323rs	udp	10	10	1719	*edge-01.foo.mycompany.com*
vc.example.com.	h323rs	udp	10	10	1719	*edge-02.foo.mycompany.com*
vc.example.com.	sip	tcp	10	10	5060	*edge-01.foo.mycompany.com*
vc.example.com.	sip	tcp	10	10	5060	*edge-02.foo.mycompany.com*
vc.example.com.	sips	tcp	10	10	5061	*edge-01.foo.mycompany.com*
vc.example.com.	sips	tcp	10	10	5061	*edge-02.foo.mycompany.com*
vc.example.com.	sip	udp *	10	10	5060	*edge-01.foo.mycompany.com*
vc.example.com.	sip	udp *	10	10	5060	*edge-02.foo.mycompany.com*

Required: Certificates

TLS certificates are used by the Management Node and each Conferencing Node to verify their identity to clients connecting to them over HTTPS (web) or SIP TLS. Examples of various clients could include:

- video endpoints
- web browsers (including the Pexip Connect web app)
- Pexip Connect mobile clients (certificates are mandatory for these clients)
- third-party video network infrastructure devices
- Outlook clients (if the VMR Scheduling for Exchange service is enabled).
- API calls

You can use Pexip's inbuilt Certificate Signing Request (CSR) generator to assist in acquiring a server certificate from a Certificate Authority.

Please review the entire KB article focused on Certificate Management for more information.

Required: Firewall Rules:

Pexip's Comprehensive list of firewall ports can be found at: https://docs.pexip.com/admin/port_usage.htm

The deployment checklist may be found at: https://docs.pexip.com/admin/deployment_checklist.htm#deployment

User acceptance testing example:

]pexip[

Step	Step Detail	Expected Result	Actual Result	(Pass/Fail)	Exceptions / Notes
1					
2					
3					
4					
5					
6					
7					

NOTE: Please consult your Pexip Engineering team for Pexip Private Cloud as the list of ports will vary depending on deployment

Please refer to Pexip's Scope of Work for an outline of the project, deliverables, and acceptance

Date	Resource	Version	Notes
<date>	<sa>	1.0	Initial Solution Design