Telco VoIP Scalability Test Results

... for 10 Million Subscribers

Introduction

Scalability is one of the key objectives for Telecommunications Carriers that offer voice-over-IP (VoIP) to their customers. CommuniGate Systems, the leader in scalable carrier-class Internet Communications solutions, delivers telco-level performance with the CommuniGate Pro Dynamic Cluster SIP Farm for more than 10 million VoIP subscribers. As a leading IETF candidate to be the foundation of a global VoIP infrastructure, the skyrocketing adoption rate of Session Initiation Protocol (SIP) is forcing Service Providers to quickly grow their capacity to sustain massive amounts of signaling and media and ensure high quality of service and access.

CommuniGate Systems, working with Hewlett-Packard, Navtel Communications, and F5 Networks in an HP performance lab tested a real-world scenario for a VoIP Telco Provider with over 10-million consumer subscribers. The goals of this benchmark were to validate that with the CommuniGate Pro Dynamic Cluster SIP Farm running on HP Integrity servers, VoIP can truly be scaled to the levels required to meet Tier 1 service provider needs, while providing 99.999% reliability, and to provide configuration and optimization guidelines for the software and hardware required to deliver the solution.

The audience of this document is staff drawn from the following groups:

- Chief architects and network planners
- Technical specialists
- Systems testing staff
- Pre sales support staff

Overview

The benchmarking tests were run between January 23 and March 6, 2006, in the HP Atlanta Performance Center, Atlanta, GA.

Representatives from CommuniGate Systems, HP, F5 Networks, and Navtel Communications worked on site and remotely to architect and configure a test system based on the HP Integrity Superdome with an EVA 6000 storage system to test the performance of a large-scale consumer-profile SIP-based VoIP subscriber base. The test methodology and results were arranged to demonstrate and document the profile of the system, while providing Lessons Learned to continue to move forward with even better future results for large-scale Telco VoIP services.

Today, large-scale VoIP implementations may consist of 1- to 5-million-user subscriber bases, with one proprietary peer-to-peer implementation reported to be near 100 million users. Traditional circuit-based PSTN telephony operators regularly serve customer markets well over 50 million lines with high quality-of-service and proven historical uptime. The distributed Internet-wide global email user base exceeds 1 billion users. The target readers of this document are Communications Service Provider management and administration staff who must provide the same level of quality, service, and standardization with VoIP as is expected from the legacy circuit-switched telephone system. By providing these test results to the public, CommuniGate Systems hopes to encourage additional public testing and reporting by all major VoIP infrastructure vendors, in an effort to increase public awareness of VoIP scalability and redundancy, and to help move SIP towards a true, Internet-wide, standardized protocol for end-to-end communication.

v5

VoIP BENCHMARK

Test Methodology

The test methodology used in this benchmark focused on the rather straightforward testing of SIP and RTP generation to emulate a real-world environment consisting of both "inbound' and "outbound" calling.

For the purpose of this test, "inbound calling" is defined as a call to a SIP User Agent (UA) registered as a unique IP/port endpoint with the CommuniGate Pro SIP Farm cluster. Also for the definition of this test, a Caller is defined as a SIP User Agent which initiates a SIP call via a SIP INVITE, and a Callee is defined as a SIP User Agent which receives the initial SIP INVITE and completes the SIP conversation with a unique Caller. The CommuniGate Pro SIP Farm was configured as a flat domain of "example.lan" with 10-million unique accounts with a 1% "subscriber usage rate concurrency" using an Address-of-Record (AOR) of test[0-1250]-[0-7999].example.lan. In the tests described below, inbound calling was generated by the Navtel InterWatch system, which first registers all UA endpoints until all UAs are registered concurrently. Following the registration process, the Navtel device was configured to run a 1-hour load test at a set throughput rate, configured for the highest level tests to run 250 calls per second on each 1Gb port (where each port provides the registration of 32,000 UA endpoints, up to a maximum of 192,000 concurrent UAs for the total number of port cards (three port cards, with two 1Gb ports per port card – on each port card, one port performed as a Caller Group and one performed as a Callee Group) provided in the Navtel chassis used). Each call lasted a duration of 60 seconds.

"Outbound calling" was therefore defined as a call to a "remote" domain, where the status of registration of the SIP User Agents is considered unknown. All outbound calling was performed by the traffic generator "sipp" (http://sipp.sourceforge.net/). For the purposes of this test, a single remote domain ("remote.lan") with DNS SRV records pointing to two sipp systems acting as sipp User Agent Servers (UAS). Two additional identical systems acted as sipp User Agent Clients (UACs). All outbound calling was generated by sipp, with a call duration of 60 seconds and 250 calls per second per sipp UAC. Outbound calling could just as easily be performed using calls to digit-based Callees if using an ENUM (E164 Number Mapping) DNS-based system, though for this particular test just AORs were used.

Each test was run multiple times over a period of 1.5 months. When running combination inbound/outbound tests, the sipp generators would be started after the Navtel UAs had registered, and would shut down after the Navtel one-hour test had completed. The test was measured a successful test with zero unsuccessful calls and a SIP retransmission rate of less than 1% of all SIP packets using a 500ms T1 timer. A profiled test result with a total of 1,000 calls per second is provided below with detailed configuration, results, and supporting system statistics.

SIP Farm Architecture

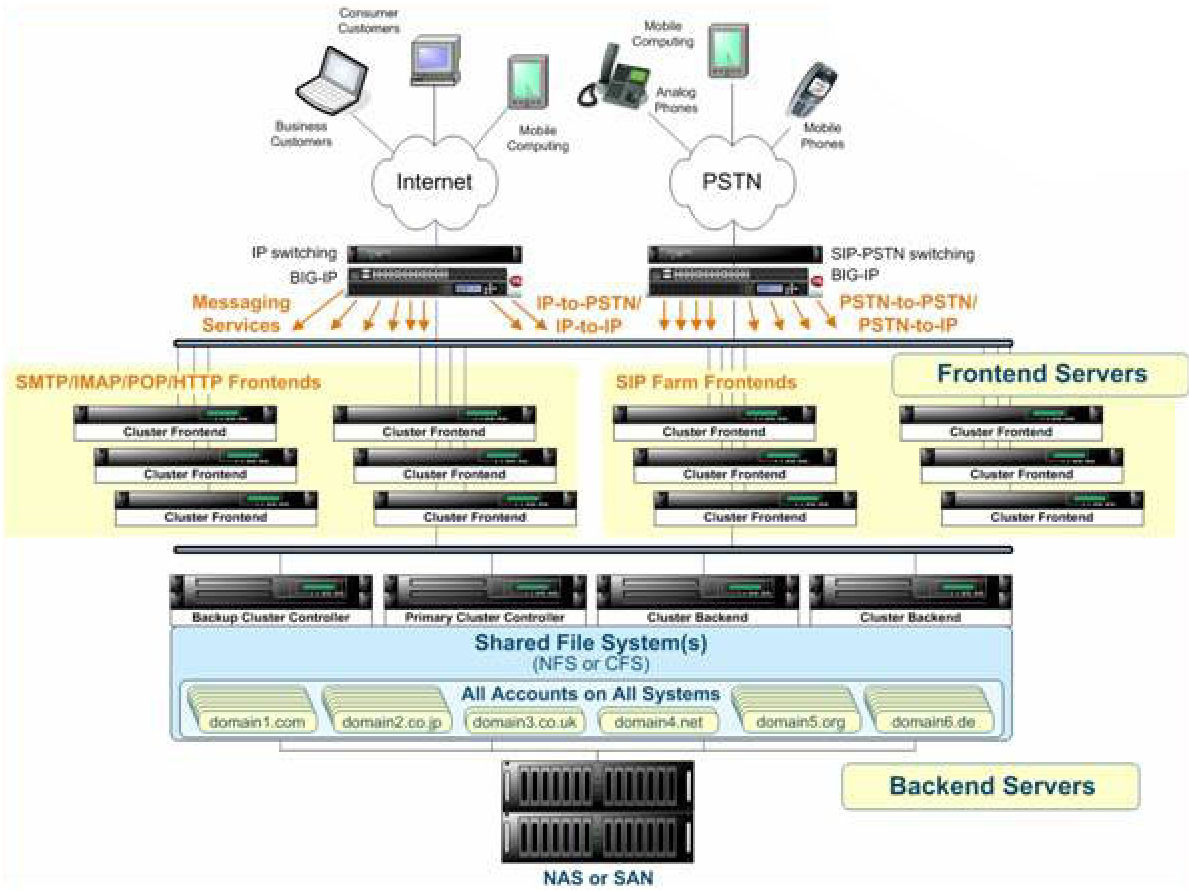
SIP Farm is CommuniGate Systems' latest technology for clustering voice-over-IP (VoIP) for 99.999% uptime, redundancy, and scalability. Both Dynamic Cluster and Super Cluster deployment implementations can be clustered with SIP Farm, and the members of a cluster allocated to the SIP Farm can be based on traffic or regional geographic node placement.

The CommuniGate Pro Dynamic Cluster maintains the information about all servers enabled for SIP Farm. Incoming SIP UDP packets and TCP connections are distributed to those servers using F5 BIG-IP Local Traffic Management devices configured to perform basic load balancing (for this test, BIG-IP SIP features were not used).

The receiving server detects if the received packet must be processed on a certain SIP Farm server - it checks if the packet is a response or an ACK packet for an existing transaction or if the packet is directed to a task created on a certain server. In this case the packet is relayed from one SIP Farm cluster member to the target member.

Packets not directed to a particular SIP Farm server are distributed to all SIP Farm members based on the CommuniGate Pro cluster algorithms and the currently available set of the active SIP Farm cluster members. In the case of the addition or loss of a SIP Farm member (such as a hardware failure), the traffic is redistributed to other SIP Farm members to maintain consistent signalling.

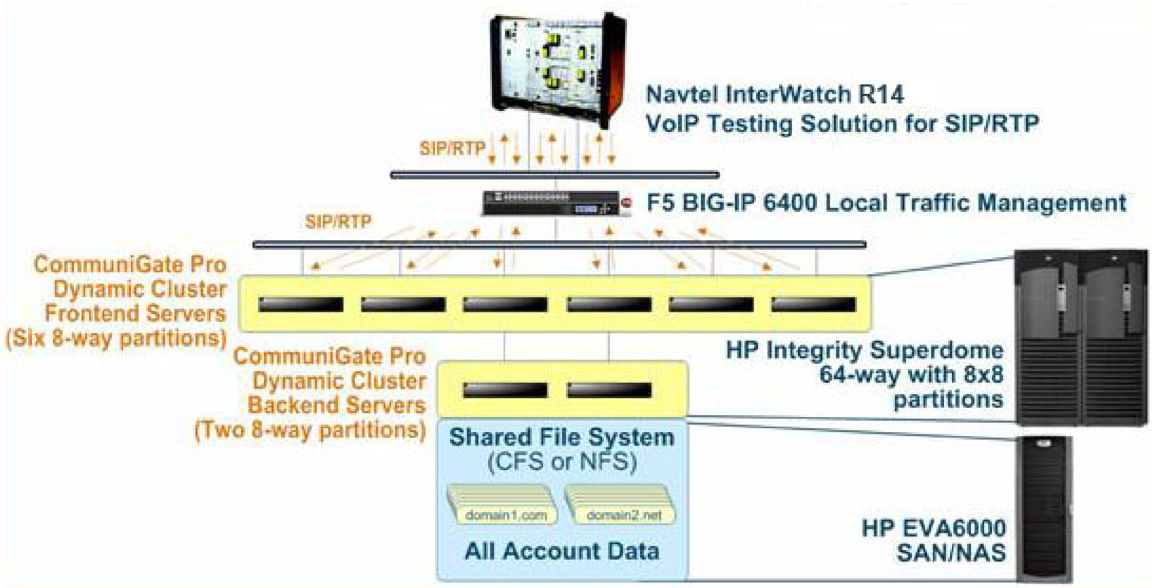
The following diagram on the next page demonstrates a "12x4" Dynamic Cluster (12 Frontends and 4 Backends) using optional "SIP Farm Specialization", which allows for a subset of all Dynamic Cluster members to be allocated as part of the SIP Farm. This technique can be used to protect the Quality of Service (QOS) of voice and real-time requirements by separating the email traffic from the SIP/RTP traffic, while continuing to maintain the single-system image of the Dynamic Cluster with consolidated identity management.



CommuniGate Pro Dynamic Cluster SIP Farm

Real-time communications (voice/video/multimedia over IP, instant messaging) demand a very low latency, synchronous application infrastructure to ensure high quality of service and reliability. In the Dynamic Cluster SIP Farm, CommuniGate Frontend servers service all accounts and handle all connections to or from the external network. Proxied SIP and RTP traffic only touches the Frontend servers, which also behave as the application servers running all voice and conference applications. CommuniGate Pro Backend servers provide registration, authentication, cluster management, and delivery/retrieval functions for all voicemail and Multimedia in the Inbox (MITI). In a CommuniGate Pro VoIP environment, Frontend servers should generally be optimized for high CPU and network usage, while Backend servers should be optimized for disk I/O capacity.

CommuniGate Pro 5.0 SIP Farm allows providers to sustain a million busy-hour call attempts and call rates of over 300 calls per second on each cluster Frontend member. The SIP Farm provides a large cluster capable of call rates of tens of millions of Busy Hour Call Attempts (BHCAs), while providing innate resilience in case of system or even site failure. SIP Farm brings to VoIP what the Dynamic Cluster brings to asynchronous communications (such as email) - unsurpassed scalability and capacity, as evidenced by CommuniGate Pro's world-record setting SPECmail benchmark records at spec.org [http://www.spec.org/mail2001/results/mail2001.html] (SPEC® and the benchmark name SPECmail® are registered trademarks of the Standard Performance Evaluation Corporation).



CommuniGate Pro Dynamic Cluster with SIP Farm
Demonstration for Telco VoIP 10-Million-Account Subscriber Base

Benchmark Architecture

The test platform is an HP Integrity Superdome with 64 Intel Itanium2 CPUs partitioned into eight separate partitions of 8 CPUs each, with the HP Super-Scalable Processor Chipset sx1000, running HP-UX 11iv2. The system was configured as a CommuniGate Pro "6x2" Dynamic Cluster (six Frontend servers and two Backend servers), with the six Frontend servers configured as a multi-node SIP Farm.

The test used the Navtel InterWatch R14 VoIP Testing Solution to generate SIP and RTP traffic. The InterWatch can emulate up to 256,000 unique SIP endpoints, generate 128,000 simultaneous RTP streams and up to 10 Million BHCAs on a single chassis.

The traffic load is distributed across the CommuniGate Pro Frontend servers using an F5 BIG-IP 6400 running BIG-IP Local Traffic Management version 9.

System Specifications

HP Integrity Superdome with sx1000 chipset (one)

- 64 x 1.5 GHz Itanium 2 processors
- 256 GB memory
- Divided into eight 8-CPU hard partitions
- OS: HP-UX 11i v2 (three total partitions, two as CommuniGate Pro Dynamic Cluster Back end Servers and one as a CommuniGate Pro Frontend servers)
- OS: RedHat ES 4.0 Linux for Integrity (five total partitions, all used as CommuniGatePro Dynamnic Cluster Frontend Servers/SIP Farm members)
- Application: CommuniGate Pro v5.0

HP EVA 6000 Storage Array (one)

- 2C8D enclosure with 2048 controller, 146G 15krpm drives arranged in two logical volumes(one 1.5 TB volume, one 0.5 TB volume), dualcontrollers with 256MB cache each, 2 x 2Gbps fibre host ports, VRAID-5 configuration.

HP Integrity 4640 (five, used for sipp load generation and DNS named)

- 4 x 1.6 GHz mx-2 Itanium CPUs (each)
- 8GB RAM (each)
- OS: 64-bit RedHat ES 4.0 Linux for Integrity
- Application: sipp traffic load generation tool

Navtel InterWatch R14 VoIP Testing Solution Chassis (one)

- Three 2x1GB port cards

F5 Networks BIG-IP 6400 Application Switch (one)

- BIG-IP Local Traffic Management version 9
- 4GB memory

HP Procurve 5308xl Switch (one)

- Gigabit network

System and Application Tuning

The tuning options used for these systems includes the following:

HP-UX System Tuning:

```
kctune maxfiles_lim=32768
kctune nfile=65536
kctune maxfiles=16384
kctune ninode=16384
kctune maxdsiz=2147481600
ndd -set /dev/sockets socket_buf_max 1048576
ndd -set /dev/sockets socket_qlimit_max 8192
```

Linux:

```
sysctl -w net.ipv4.neigh.default.gc_
thresh3=262144
# max open files
echo 262144 > /proc/sys/fs/file-max
# kernel threads
echo 131072 > /proc/sys/kernel/threads-max
# socket buffers
echo 111616 > /proc/sys/net/core/wmem_default
echo 4194304 > /proc/sys/net/core/wmem_max
echo 111616 > /proc/sys/net/core/rmem_default
echo 4194304 > /proc/sys/net/core/rmem_max
# netdev backlog
echo 4096 > /proc/sys/net/core/netdev_max_back-
log
# socket buckets
echo 131072 > /proc/sys/net/ipv4/tcp_max_tw_
buckets
# port range
echo '16384 65535' > /proc/sys/net/ipv4/ip_lo-
cal_port_range
# disabling forwarding
echo 0 >> /proc/sys/net/ipv4/ip_forward
```

CommuniGate Pro:

/var/CommuniGate/Startup.sh:

```
### Startup.sh begin
SUPPLPARAMS="--DefaultStackSize 131072 --use-
NonBlockingSockets --closeStuckSockets --Crea-
teTempFilesDirectly 10 --ClusterFrontend --si-
pUDPReceiveBuffer 4M"
ulimit -n 32768
### Startup.sh end
```

/etc/init.d/CommuniGate:

Used pthreads option for Linux (commented out LD_ASSUME_KERNEL)

```
    # LD_ASSUME_KERNEL=2.4.1
    # export LD_ASSUME_KERNEL
```

CGP Configuration Options:

Cluster-wide Options:

Cluster Domain: example.lan

```
Account Storage:
    Foldering Method: Hashed 2 Levels
    Generate Index: Yes
```

Per-System Settings (Frontends):

SIP Settings:

```
Sending:
    Client Transactions Limit: 300000
    Processors (threads): 10
Receiving:
    Enqueuers: 15
    Server Transaction Limit: 300000
    Processors (threads): 10
Real-Time Settings:
    Signal Transaction Limit: 300000
    Signal Processors (threads): 10
    Nodes Transaction Limit: 300000
    Nodes Processors (threads): 10
```

Navtel InterWatch Configuration Options Overview:

2-3 Caller Groups, each with 32,000 registered UA endpoints
2-3 Callee Groups, each with 32,000 registered UA endpoints
Default settings for T1 and other timeouts
Registration interval (per Group) of 20ms
Single Proxy Address (all Groups): 10.10.0.110 (the F5 Load Balancer)
IP address space for UA clients: 10.10.1.1-10.10.200.254
Domain for UA clients: remote.lan

F5 Configuration Overview:

BIG-IP virtual server represents a pool of six CommuniGate Pro Frontend Servers: listening on 10.10.0.110
Inactive timeout: 5 minutes
Tests with two different NAT methods were performed with identical results. Method 1 should be used if preserving the individual server address is required (which may be necessary when the server is also performing media proxying or termination functions).

```
Method 1) Outbound NAT for each Frontend:
    Origin Address -> NAT Address
    192.168.0.113 -> 10.10.0.113
    192.168.0.114 -> 10.10.0.114
    192.168.0.115 -> 10.10.0.115
    192.168.0.116 -> 10.10.0.116
    192.168.0.117 -> 10.10.0.117
    192.168.0.118 -> 10.10.0.118
```

```
Method 2) Single Outbound NAT Address, with RTP
port ranges defined for each CommuniGate Pro
Frontend System:
    Origin Address -> NAT Address
    192.168.0.113 -> 10.10.0.110
    192.168.0.114 -> 10.10.0.110
    192.168.0.115 -> 10.10.0.110
    192.168.0.116 -> 10.10.0.110
    192.168.0.117 -> 10.10.0.110
    192.168.0.118 -> 10.10.0.110
```

sipp Configuration

URL : http://sipp.sourceforge.net/

source code: sipp.cumulus.2006-01-24

sipp Compile Options:

```
# gunzip sipp-xxx.tar.gz
# tar -xvf sipp-xxx.tar
# cd sipp
# make ossl
```

sipp Scenarios:

UAS: sipp-uas-hpvoip4.xml

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE scenario SYSTEM "sipp.dtd">

<!-- This program is free software; you can redistribute it and/or      -->
<!-- modify it under the terms of the GNU General Public License as      -->
<!-- published by the Free Software Foundation; either version 2 of the -->
<!-- License, or (at your option) any later version.                     -->
<!--                                                                     -->
<!-- This program is distributed in the hope that it will be useful,     -->
<!-- but WITHOUT ANY WARRANTY; without even the implied warranty of      -->
<!-- MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the       -->
<!-- GNU General Public License for more details.                        -->
<!--                                                                     -->
<!-- You should have received a copy of the GNU General Public License  -->
<!-- along with this program; if not, write to the                       -->
<!-- Free Software Foundation, Inc.,                                     -->
<!-- 59 Temple Place, Suite 330, Boston, MA  02111-1307 USA              -->
<!--                                                                     -->
<scenario name="Basic UAS responder">
  <!-- By adding rrs="true" (Record Route Sets), the route sets          -->
  <!-- are saved and used for following messages sent. Useful to test   -->
  <!-- against stateful SIP proxies/B2BUAs.                              -->
  <recv request="INVITE" crlf="true">
  </recv>
  <!-- The '[last_*]' keyword is replaced automatically by the           -->
  <!-- specified header if it was present in the last message received -->
  <!-- (except if it was a retransmission). If the header was not       -->
  <!-- present or if no message has been received, the '[last_*]'       -->
  <!-- keyword is discarded, and all bytes until the end of the line    -->
  <!-- are also discarded.                                               -->
  <!--                                                                   -->
  <!-- If the specified header was present several times in the        -->
  <!-- message, all occurences are concatenated (CRLF seperated)        -->
  <!-- to be used in place of the '[last_*]' keyword.                    -->
  <send>
    <![CDATA[

      SIP/2.0 180 Ringing
      [last_Via:]
      [last_From:]
      [last_To:];tag=[call_number]
      [last_Call-ID:]
      [last_CSeq:]
      Contact: <sip:[local_ip]:[local_port];transport=[transport]>
      Content-Length: 0

    ]]>
  </send>
  <pause milliseconds="500"/>
  <send retrans="500">
```

```
    <![CDATA[

      SIP/2.0 200 OK
      [last_Via:]
      [last_From:]
      [last_To:];tag=[call_number]
      [last_Call-ID:]
      [last_CSeq:]
      Contact: <sip:[local_ip]:[local_port];transport=[transport]>
      Content-Type: application/sdp
      Content-Length: [len]

      v=0
      o=user1 53655765 2353687637 IN IP[local_ip_type] [local_ip]
      s=-
      c=IN IP[media_ip_type] [media_ip]
      t=0 0
      m=audio [media_port] RTP/AVP 0
      a=rtpmap:0 PCMU/8000

    ]]>
  </send>
  <recv request="ACK"
        optional="true"
        rtd="true"
        crlf="true">
  </recv>
  <recv request="BYE">
  </recv>
  <send>
    <![CDATA[

      SIP/2.0 200 OK
      [last_Via:]
      [last_From:]
      [last_To:]
      [last_Call-ID:]
      [last_CSeq:]
      Contact: <sip:[local_ip]:[local_port];transport=[transport]>
      Content-Length: 0

    ]]>
  </send>
  <!-- Keep the call open for a while in case the 200 is lost to be     -->
  <!-- able to retransmit it if we receive the BYE again.               -->
  <pause/>
  <!-- definition of the response time repartition table (unit is ms)   -->
  <ResponseTimeRepartition value="10, 20, 30, 40, 50, 100, 150, 200"/>
  <!-- definition of the call length repartition table (unit is ms)     -->
  <CallLengthRepartition value="10, 50, 100, 500, 1000, 5000, 10000"/>
</scenario>
```

UAC: sipp-uac-hpvoip4.xml

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE scenario SYSTEM "sipp.dtd">

<!-- This program is free software; you can redistribute it and/or      -->
<!-- modify it under the terms of the GNU General Public License as      -->
<!-- published by the Free Software Foundation; either version 2 of the -->
<!-- License, or (at your option) any later version.                     -->
<!--                                                                     -->
<!-- This program is distributed in the hope that it will be useful,    -->
<!-- but WITHOUT ANY WARRANTY; without even the implied warranty of      -->
```

```
<!-- MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the      -->
<!-- GNU General Public License for more details.                       -->
<!--                                                                    -->
<!-- You should have received a copy of the GNU General Public License  -->
<!-- along with this program; if not, write to the                      -->
<!-- Free Software Foundation, Inc.,                                    -->
<!-- 59 Temple Place, Suite 330, Boston, MA  02111-1307 USA             -->
<!--                                                                    -->
<scenario name="Basic Sipstone UAC">
  <!-- In client mode (sipp placing calls), the Call-ID MUST be         -->
  <!-- generated by sipp. To do so, use [call_id] keyword.                -->
  <send retrans="500">
    <![CDATA[

      INVITE sip:[service]@[remote_ip]:[remote_port] SIP/2.0
      Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
      From: sipp <sip:sipp@[local_ip]:[local_port]>;tag=[call_number]
      To: sut <sip:[service]@[remote_ip]:[remote_port]>
      Call-ID: [call_id]
      CSeq: 1 INVITE
      Contact: sip:sipp@[local_ip]:[local_port]
      Max-Forwards: 70
      Subject: Performance Test
      Content-Type: application/sdp
      Content-Length: [len]

      v=0
      o=user1 53655765 2353687637 IN IP[local_ip_type] [local_ip]
      s=-
      c=IN IP[media_ip_type] [media_ip]
      t=0 0
      m=audio [media_port] RTP/AVP 0
      a=rtpmap:0 PCMU/8000

    ]]>
  </send>
  <recv response="100" optional="true">
  </recv>
  <recv response="180" optional="true">
  </recv>
  <!-- By adding rrs="true" (Record Route Sets), the route sets         -->
  <!-- are saved and used for following messages sent. Useful to test  -->
  <!-- against stateful SIP proxies/B2BUAs.                             -->
  <recv response="200" rtd="true">
  </recv>
  <!-- Packet lost can be simulated in any send/recv message by         -->
  <!-- by adding the 'lost = "10"'. Value can be [1-100] percent.       -->
  <send>
    <![CDATA[

      ACK sip:[service]@[remote_ip]:[remote_port] SIP/2.0
      Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
      From: sipp <sip:sipp@[local_ip]:[local_port]>;tag=[call_number]
      To: sut <sip:[service]@[remote_ip]:[remote_port]>[peer_tag_param]
      Call-ID: [call_id]
      CSeq: 1 ACK
      Contact: sip:sipp@[local_ip]:[local_port]
      Max-Forwards: 70
      Subject: Performance Test
      Content-Length: 0

    ]]>
```

```
    </send>
    <!-- This delay can be customized by the -d command-line option       -->
    <!-- or by adding a 'milliseconds = "value"' option here.             -->
    <pause/>
    <!-- The 'crlf' option inserts a blank line in the statistics report. -->
    <send retrans="500">
      <![CDATA[

        BYE sip:[service]@[remote_ip]:[remote_port] SIP/2.0
        Via: SIP/2.0/[transport] [local_ip]:[local_port];branch=[branch]
        From: sipp <sip:sipp@[local_ip]:[local_port]>;tag=[call_number]
        To: sut <sip:[service]@[remote_ip]:[remote_port]>[peer_tag_param]
        Call-ID: [call_id]
        CSeq: 2 BYE
        Contact: sip:sipp@[local_ip]:[local_port]
        Max-Forwards: 70
        Subject: Performance Test
        Content-Length: 0

      ]]>
    </send>
    <recv response="100" optional="true">
    </recv>
    <recv response="200" crlf="true">
    </recv>
    <!-- definition of the response time repartition table (unit is ms)   -->
    <ResponseTimeRepartition value="10, 20, 30, 40, 50, 100, 150, 200"/>
    <!-- definition of the call length repartition table (unit is ms)     -->
    <CallLengthRepartition value="10, 50, 100, 500, 1000, 5000, 10000"/>
  </scenario>
```

sipp Startup Options:

UAS (2):

```
host164# sipp -sf sipp-uas-hpvoip4.xml -d 0 -i 10.10.0.164 -p 5060 -trace_err
host165# sipp -sf sipp-uas-hpvoip4.xml -d 0 -i 10.10.0.165 -p 5060 -trace_err
```

UAC (2):

```
host166# sipp -sf sipp-uac-hpvoip4.xml -r 200 -rp 1000 -l 50000 -d 60000 -i
10.10.0.166 -p 5060 -trace_err -rsa 10.10.0.110:5060 10.10.0.165:5060
host167# sipp -sf sipp-uac-hpvoip4.xml -r 200 -rp 1000 -l 50000 -d 60000 -i
10.10.0.167 -p 5060 -trace_err -rsa 10.10.0.110:5060 10.10.0.164:5060
```

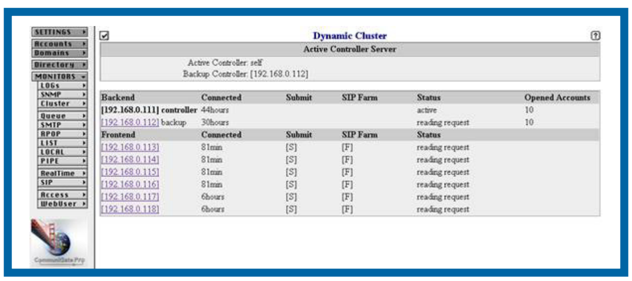
CommuniGate Pro SIP Farm Monitoring

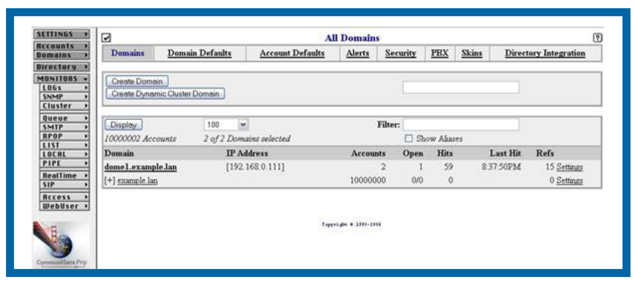
The following screenshots pages display the CommuniGate Pro cluster monitoring for the 6x2 cluster, and with a 10-million account subscriber base. The data storage required to create the 10-million accounts enabled for all services (SIP, PBX, IMAP, POP, SMTP, etc.) with password and related metadata as well as an INBOX mailbox for voicemail/email was approximately 31GB, on a single logical volume made available to all CommuniGate Pro Dynamic Cluster Backend Servers using NFS.

6x2 Dynamic Cluster with SIP Farm and a 10-million subscriber base

Cluster Setup
Test Domain: example.lan
Accounts: 10,000,000

Hostname	Cluster Function	Unique "Public" IP Address	Virtual (Load-balanced) Public IP Address/ Pool	Private IP Address	Operating System
be1.example.lan	Backend	10.10.0.111	n/a	192.168.0.111	HP-UX 11iv2
be2.example.lan	Backend	10.10.0.112	n/a	192.168.0.112	HP-UX 11iv2
fe1.example.lan	Frontend/ SIP Farm	10.10.0.113	10.10.0.110	192.168.0.113	RHES 4.0 Linux
fe2.example.lan	Frontend/ SIP Farm	10.10.0.114	10.10.0.110	192.168.0.114	RHES 4.0 Linux
fe3.example.lan	Frontend/ SIP Farm	10.10.0.115	10.10.0.110	192.168.0.115	HP-UX 11iv2
fe4.example.lan	Frontend/ SIP Farm	10.10.0.116	10.10.0.110	192.168.0.116	RHES 4.0 Linux
fe5.example.lan	Frontend/ SIP Farm	10.10.0.117	10.10.0.110	192.168.0.117	RHES 4.0 Linux
fe6.example.lan	Frontend/ SIP Farm	10.10.0.118	10.10.0.110	192.168.0.118	RHES 4.0 Linux





Benchmark Results

6x2 CommuniGate Pro SIP Farm Dynamic Cluster

Registration Test: Registered 192,000 User-Agent Endpoints (simultaneously)

The top screenshot on the opposing page demonstrates the Navtel InterWatch SIP UA Functional & Performance Test Tool application with active registrations of 192,000 UA endpoints. A fully stocked Navtel R14 InterWatch chassis can register up to 256,000 UA endpoints; however for this benchmark, only 3 of 4 available port card slots were in use.

The particular test profiled above was a Navtel-only (i.e., "inbound calling") only test, demonstrating 2,160,000 successfully completed 60-second SIP calls (no RTP) over a one-hour period, for a call rate of 600 calls per second (200 calls per second per Calling Group port), with 0 unsuccessful calls.

Profiled Test: 500 calls per second from Navtel + 500 calls per second using sipp (simultaneously)
February 26, 2006, from approximately 1700-1900 EST.

The following Profiled Test demonstrates the results of a 1,000 calls-per-second test, with 500 SIP calls per second of "inbound calling" (no RTP) with the Navtel InterWatch SIP UA Functional & Performance Test Tool application, and simultaneously 500 calls per second generated by the sipp tool on four HP Integrity 4640s, representing "outbound calling".

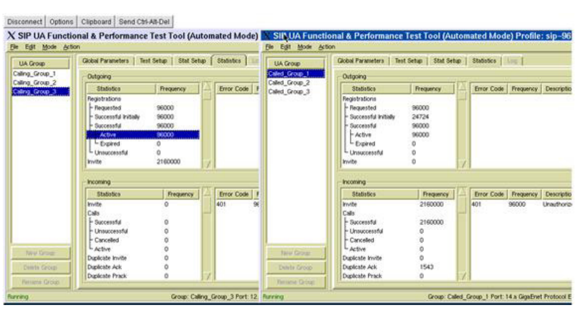
Registration Test: Total Unique User Agents Registered: 192,000 Total Throughput = 2,160,000 Completed SIP Calls per hour (600 calls per second, 60s call duration, zero unsuccessful calls)			
Test Group	**Call Function**	**Unique Registered User-Agent End Points**	**Completed Calls Per Second (60s call duration)**
Calling Group 1	Caller (inbound)	32,000	200
Calling Group 2	Caller (inbound)	32,000	200
Calling Group 3	Caller (inbound)	32,000	200
Called Group 1	Callee (inbound)	32,000	n/a
Called Group 2	Callee (inbound)	32,000	n/a
Called Group 3	Callee (inbound)	32,000	n/a
	Total	**192,000**	**600**

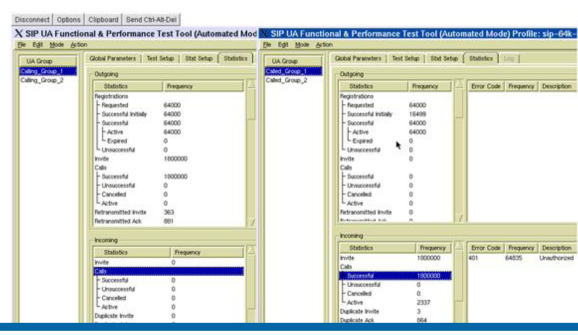
Navtel InterWatch Results

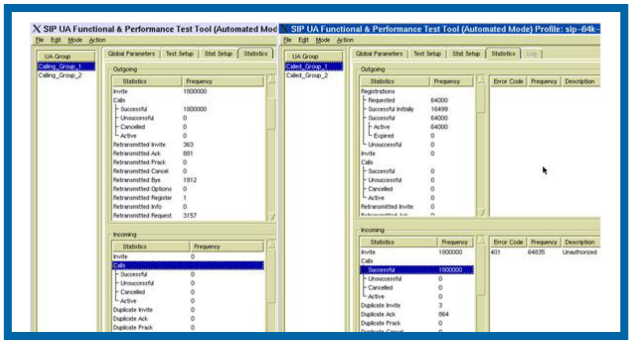
The Navtel InterWatch was configured with two Calling Groups to total 64,000 registered Callers, and two Called Groups totaling 64,000 registered Callees, for a total of 128,000 registered accounts for this test. The call rate was set to 250 calls per second per Calling Group, for a total of 500 calls per second from the Navtel chassis.

The following two screenshots are from the same results screen, but with the one on the following page adjusted to display the total number of retransmissions.

Profiled Test: Total Unique User Agents Registered: 128,000 Total Throughput = 3,600,000 Completed SIP Calls per hour (1,000 calls per second, 60s call duration, zero unsuccessful calls)			
Test Group	**Call Function**	**Unique Registered User-Agent End Points**	**Completed Calls Per Second (60s call duration)**
Navtel InterWatch Results			
Calling Group 1	Caller (inbound)	32,000	250
Calling Group 2	Caller (inbound)	32,000	250
Called Group 1	Callee (inbound)	32,000	n/a
Called Group 2	Callee (inbound)	32,000	n/a
Sipp Results			
Sipp system “sut164”	Callee (outbound)	n/a	n/a
Sipp system “sut165”	Callee (outbound)	n/a	n/a
Sipp system “sut166”	Caller (outbound)	n/a	250
Sipp system “sut167”	Caller (outbound)	n/a	250
	Total	**128,000**	**1,000**







During this test, the Navtel device reported 0 unsuccessful calls and a one-hour total of 1,800,000 successful 60-second calls at a rate of 500 calls per second, and a retransmission rate of significantly less than 1%.

Sipp Results

Similarly, the sipp UAS and UAC components demonstrated a sustained call rate total of 500 calls per second with far less than 1% retransmission.

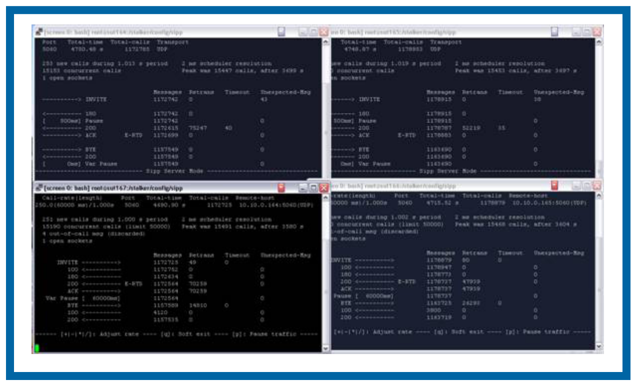
System Statistics During the Profiled Test

The 6x2 cluster system nodes (6 CommuniGate Pro Frontend Servers – one HP-UX and five Linux, and 2 CommuniGate Pro Backend Servers – both HP-UX), demonstrated a large amount of available CPU and memory resources, although with a large amount of network traffic.

CPU Usage:

The CommuniGate Pro SIP Farm cluster on Superdome demonstrated a large amount of available CPU availability during the tests. The following table lists the maximum CPU usage (as measured by the HP-UX and Linux operating systems, using the tools "sar" or "top") during the 1,000 calls-per-second test:

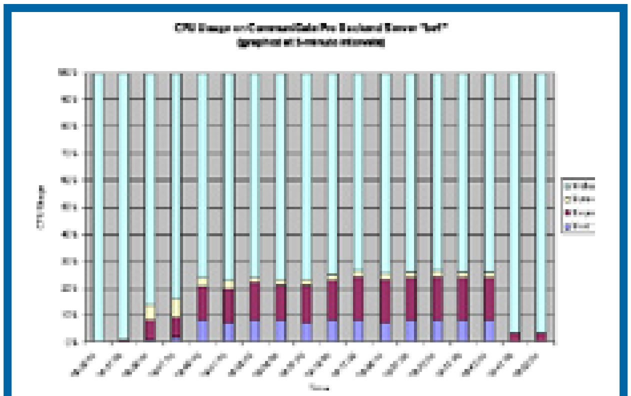
Real-time traffic (e.g., SIP, XMPP, RTP) consumes primarily CPU resources and requires very low latency packet transfer, and for these tests was expected to be the primary bottleneck in the system. However, it was discovered that for these particular tests, the actual limiting factor was primarily



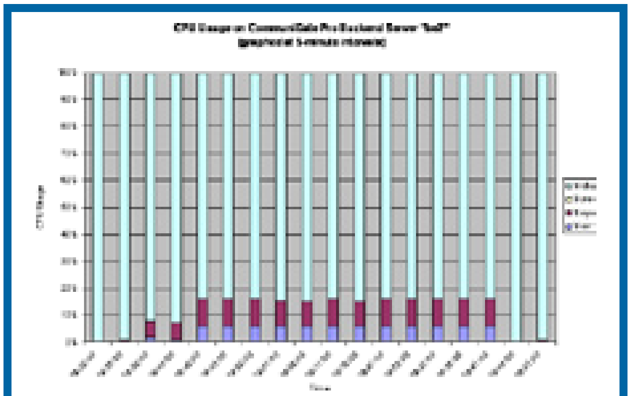
System Hostname	Operating System	Maximum Measured CPU Usage (%) (1-minute intervals) User (%)	System (%)	Wait (%)	Minimum Measured Idle CPU availability (%)
Backend Systems (partitions on the Superdome)					
be1 (Backend) (*also NFS server)	HP -UX 11iv2	8%	17%	7%	73%
be2	HP -UX 11iv2	6%	10%	0%	83%
Frontend Systems (partitions on the Superdome)					
fe1 (Frontend)	RedHat ES 4	6.2%	2.3%	0.8%	87.8%
fe2	RedHat ES 4	7.8%	3.7%	0.2%	85.3%
fe3	HP -UX 11iv2	5%	10%	1%	86%
fe4	RedHat ES 4	6.2%	2.5%	0.5%	88.1%
fe5	RedHat ES 4	7.0%	2.3%	0.1%	87.7%
fe6	RedHat ES 4	12.6%	3.7%	0.2%	85.5%

the ability of the operating system UDP stack and physical network interfaces to successfully receive very high rates of UDP packets (average UDP incoming packet rate of over 2200 UDP packets/second per Frontend system). For sizing purposes, it is highly recommend to increase the number and throughput specifications of physical network interfaces, as well as to use horizontal scaling (more, smaller Frontends rather than vertical scaling (fewer, bigger Frontends). A future similar benchmark will benefit from these Lessons Learned.

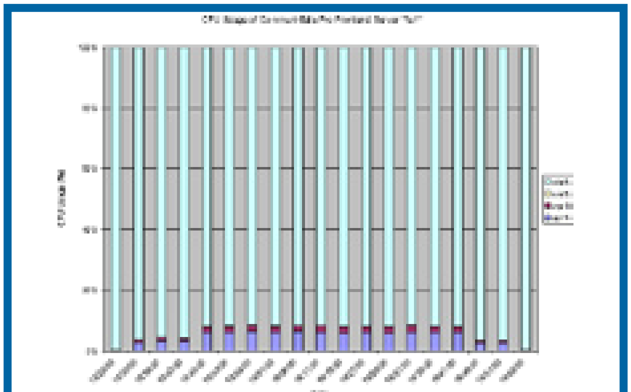
The following graphs visually demonstrate the CPU usage of the systems during the profiled test. [Note: the 1-minute gap between the 1800 and 1900 hours was caused by the 60-minute data gathering interval on each server, where the first measurement (1900) contained no valid data.]



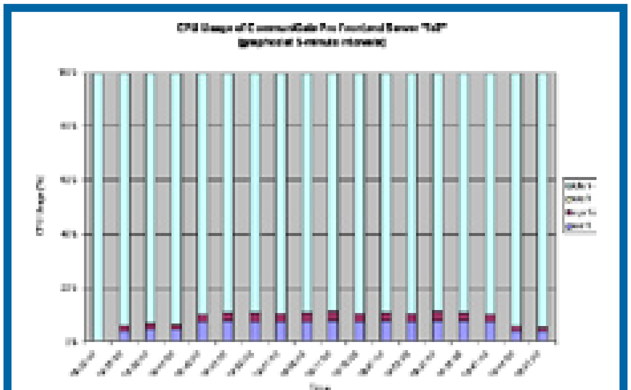
Backend Server Partition "be1" (HP-UX 11iv2 IA64)
(also acting as the NFS server for be2)



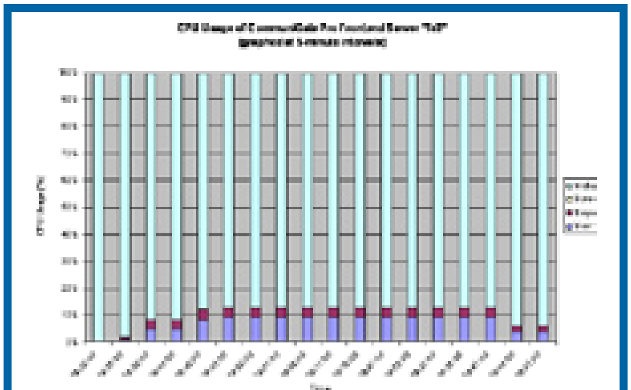
Backend Server Partition "be2" (HP-UX 11iv2 IA64)



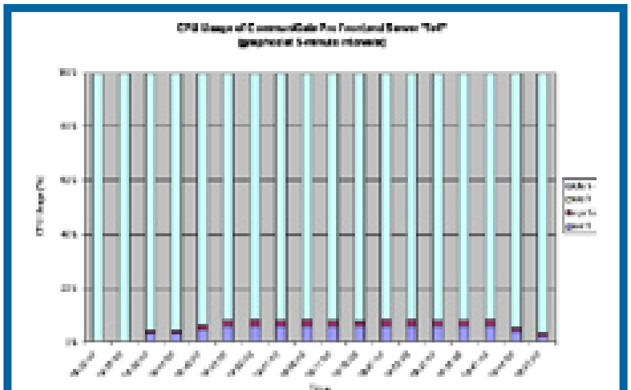
Frontend Server Partition "fe1" (RedHat Linux ES 4.0 IA64)



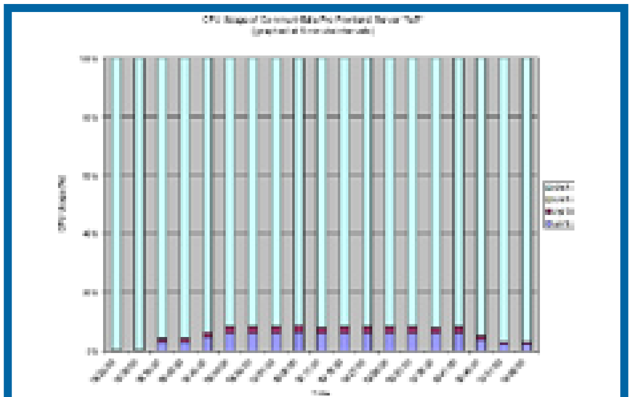
Frontend Server Partition "fe2" (RedHat Linux ES 4.0 IA64)



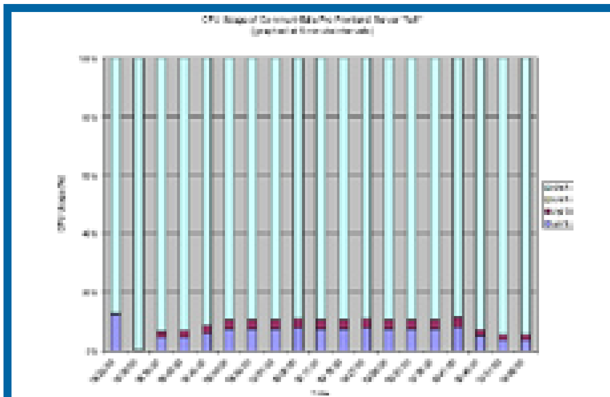
Frontend Server Partition "fe3" (HP-UX 11iv2 IA64)



Frontend Server Partition "fe4" (RedHat Linux ES 4.0 IA64)



Frontend Server Partition "fe5" (RedHat Linux ES 4.0 IA64)

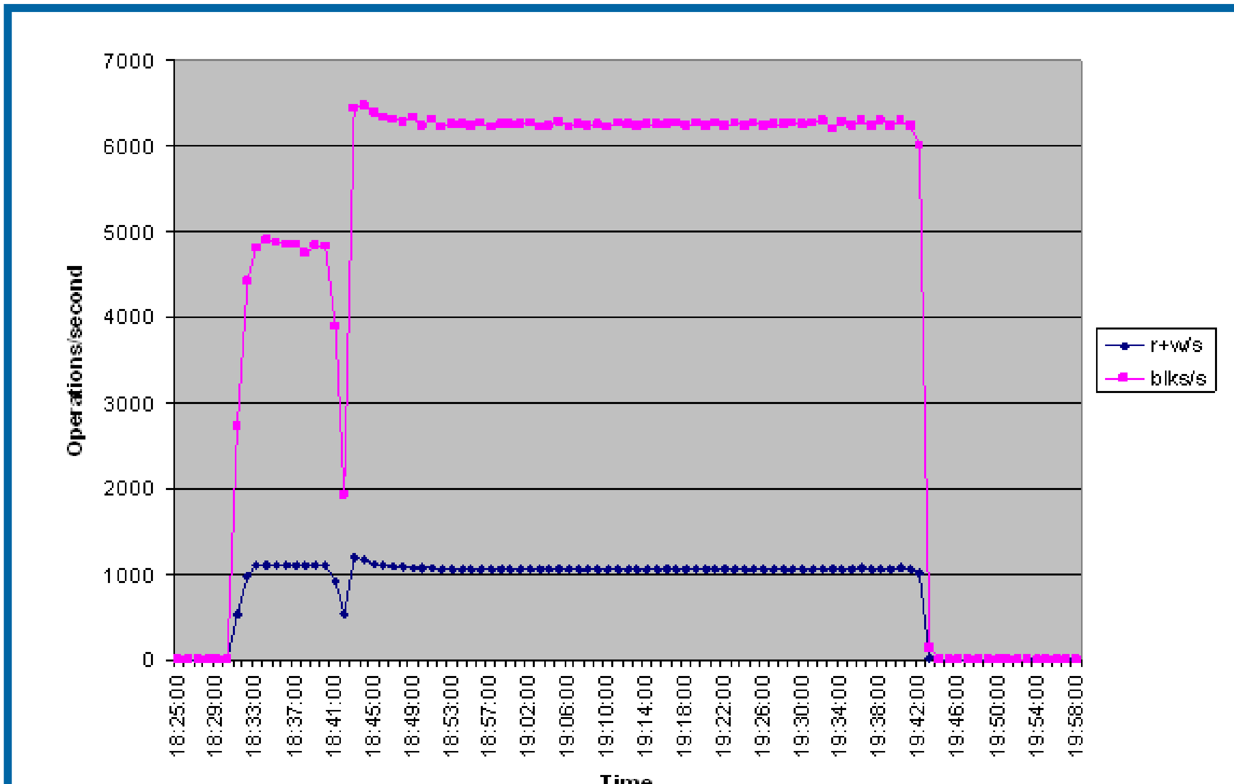


Frontend Server Partition "fe6" (RedHat Linux ES 4.0 IA64)

Disk I/O Usage:

All disk I/O (other than logging) for the SIP Farm architecture was concentrated on one system, "be1", for the profiled test. In a more traditional Dynamic Cluster SIP Farm architecture, the Shared File System accessed by all Backend systems would reside on an NFS file system being served by dedicated NAS server, or using a SAN-based Cluster File System. Due to the constraints of time and hardware for this particular benchmark, be1 was providing not only "Cluster Controller" duties but also acted as the NFS server, with the data volume also being mounted by the second Backend system, "be2". While this architecture is not thought to have been a limiting factor in the test – for best performance and redundancy, a dedicated storage controller is strongly recommended.

The following diagram graphs the total number of reads and writes per second, as well as blocks per second, measured on be1 during the profiled test. In this diagram, near 100% of the disk I/O in use would be either SIP registration (authentication) or a query of SIP registration data (e.g., "what is the IP address and port of the registration for user test100?")

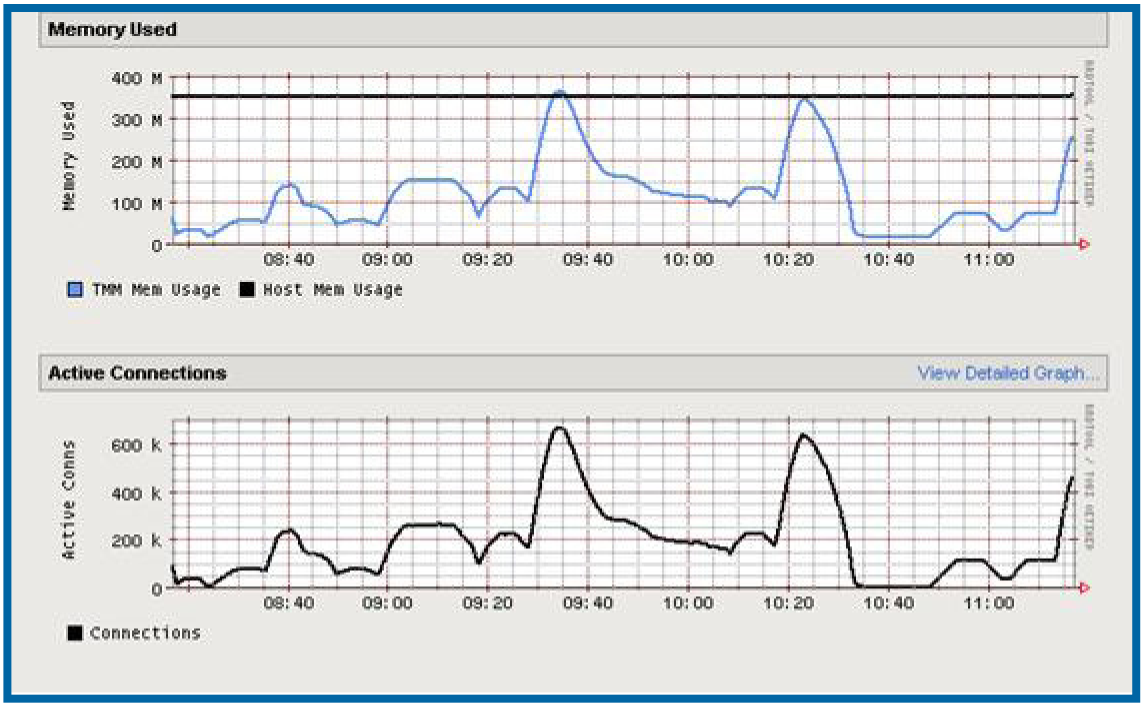


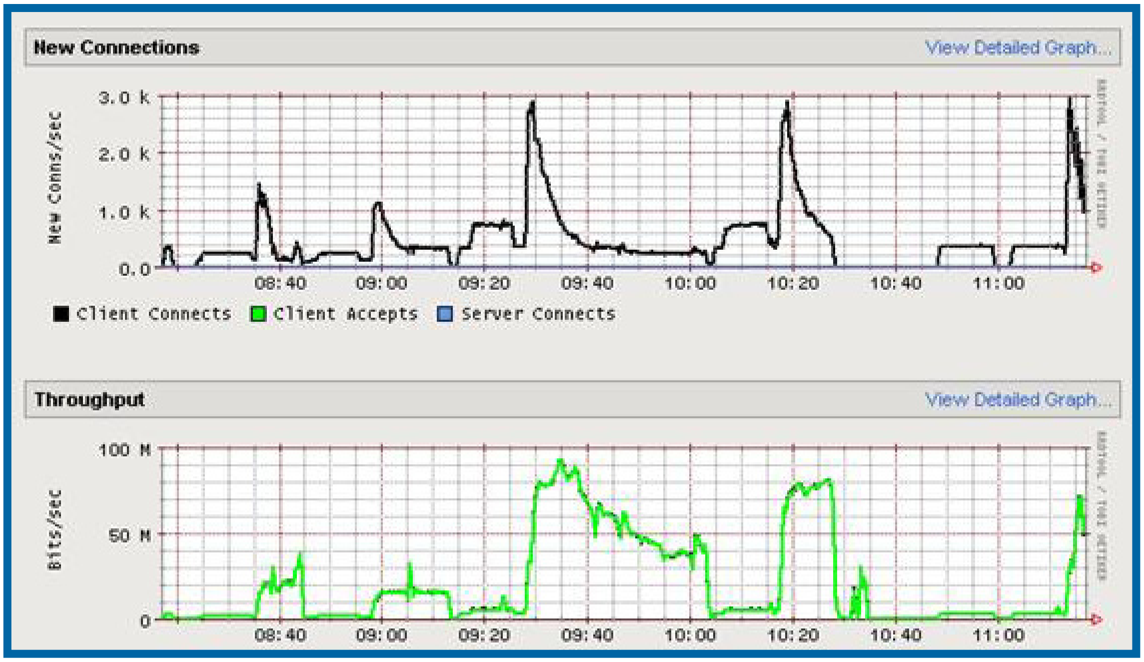
Raw Disk Usage of CommuniGate Pro Backend Server "be1"

F5 BIG-IP Performance

The F5 Networks BIG-IP 6400 Application Switch with BIG-IP Local Traffic Management version 9 and 4GB of memory demonstrated a large capacity for graceful traffic handling under the significant loads generated in these tests.

The following screenshots were taken from the BIG-IP management interface during the profiled 1,000-calls-per-second test, demonstrating a maximum memory usage of near 350 MB while nearing 700,000 active connections (all UDP). Even with new connections arriving at a rate of 3,000 new connections per second and 100Mb traffic volume, the BIG-IP continued distributing traffic to the SIP Farm pool without performance degradation or failure, demonstrating a significant amount of available headroom for burst peak rates or additional growth:





Benchmark Results Conclusion

HP Superdome with Intel Itanium 2

This benchmark test demonstrated to satisfaction the performance and ease-of-administration of a CommuniGate Pro Dynamic Cluster with SIP Farm on the Superdome architecture. The Itanium2 processors with SX10000 chipset provided excellent cost-performance for real-time communications, providing substantial CPU resources with very small latency for hundreds of thousands of simultaneous voice callers.

Navtel Communications

This test emulated a VoIP infrastructure based solely on UDP packet transfer, which mirrors real-world VoIP expectations in the immediate market (moving forward, we would expect to see an increasing growth in the use of SIP-TLS signalling for secure session initiation, and a future benchmark should be investigated focusing on the use of SIP-TLS as well as higher-load RTP generation). Navtel Communications provides the industry's highest scalability and performance test solutions for testing NGN and IMS converged networks. Navtel's products are used by network equipment manufacturers and operators around the globe to measure the performance and scalability of converged networks and network elements.

F5 Networks

The F5 Networks BIG-IP demonstrated a significant capacity to handle any number of UDP packets sent to it – these tests did not stress the BIG-IP in memory usage or number of connections handled, peaking at only about 800,000 simultaneous connections while the BIG-IP 6400 is specified at 2,000,000 connections maximum. For a truly redundant VoIP infrastructure, BIG-IP devices should be implemented as a redundant pair (active/standby).

BIG-IP includes many advanced functions which may be employed to more intelligently manage traffic. (e.g. deep packet inspection, iRules, SIP monitor, SIP persistence, connection aggregation, LAN and WAN acceleration, application optimization, etc.), however, these were beyond the scope of this benchmark.

F5 Networks is the global leader in Application Delivery Networking. F5 provides solutions that make applications secure, fast, and available for everyone, helping organizations get the most out of their investment. By adding intelligence and manageability into the network to offload applications, F5 optimizes applications and allows them to work faster and consume fewer resources. F5's extensible architecture intelligently integrates application optimization, protects the application and the network, and delivers application reliability - all on one universal platform. Over 10,000 organizations and service providers worldwide trust F5 to keep their applications running. The company is headquartered in Seattle, Washington with offices worldwide. For more information, go to www.f5.com.

CommuniGate Systems

The CommuniGate Pro Dynamic Cluster with SIP Farm demonstrated a very high performance profile, with similar results when running either a 2x2 Cluster (2 Frontends, 2 Backends) or a larger 6x2 Cluster (6 Frontends, 2 Backends). As CommuniGate Pro runs as a single C++-based process on each cluster member, these processes ran at a maximum usage of only 1GB memory and peaked at about 40% CPU usage on the Superdome partitions (and an average CPU usage of less than 20%), providing a large amount of headroom for scalable growth or burst peak loads. SIP Farm cluster members could be removed and added even during tests, using the "MAKE NOT READY" graceful shutdown mode as well as sudden process termination, demonstrating the ability to move active sessions to alternate SIP Farm cluster members automatically. The BIG-IP in this environment was not configured to be "cluster aware", so all SIP traffic was simply distributed using a round-robin based mechanism, with equal weighting to all available SIP Farm cluster members (by checking the SIP listener ports every few seconds).

Lessons Learned

- While the EVA 6000 provided vast amounts of I/O capacity, an alternate storage layout using a high-performance cluster file system or NAS head to the EVA 6000 would have removed the potential of storage access bottlenecks. The test method used in this benchmark - with one of the CommuniGate Pro Backend Servers also acting as the NFS file server - caused an imbalance in I/O distribution for registered accounts (both authentication and registration-to-IP mapping). Account metadata queries for authentication and "inbound' calling can cause significant storage read and write access requirements.

- Using a vast number of independent IP addresses (up to 192,000) but keeping them on a flat network space (single subnet) caused some networking inefficiencies for systems and load balancers, which had to maintain huge ARP tables. Adding an "external" gateway to the test would have separated the client systems from the SIP infrastructure in a way that more accurately emulates the real-world.

• Due to the very sizable amount of largely UDP-based traffic, increasing the number of physical NIC interfaces can improve overall performance. Using multiple bonded interfaces for the Internet-facing routes in addition to separate (and potentially also bonded) interfaces to the intra-cluster and storage network adds significant capacity for SIP and RTP traffic.

• The SIP Farm architecture with software-based SIP servers must be sized to take into account for Operating System performance characteristics. In general, it is recommended to scale the SIP Farm horizontally with additional Frontends in order to grow, rather than simply increasing server size by increasing CPUs.

• Tuning the systems and applications to maximize UDP buffer sizes and ultimately throughput is critical to maintaining stability under heavy UDP packet volume and peak load bursts. Stocking the Navtel InterWatch chassis with additional port cards (up to 4 can be used) can greatly increase both the maximum number of concurrent endpoints as well as simultaneous registrations and calls.

• When using HP-UX for highly-threaded applications such as CommuniGate, it is critical to ensure that the appropriate patches are applied. Applying these patches to the HP-UX systems demonstrated a 3x increase in performance – by greatly improving the ability of the application to quickly accept and process very large amounts of UDP packets. The following patches were applied:

- PHCO_33675 - s700_800 11.23 pthread library cumulative patch
- PHKL_34032 - s700_800 11.23 ksleep cumulative patch

Summary

This benchmark was deemed a success by all partner companies, completing the goals set out before the benchmark. A 10-million account subscriber base, emulating a real-world Consumer environment for SIP-based voice-over-IP, including active call rates of 1% of the subscriber base, as well as very low latency and transmission, can be successfully deployed on the HP Integrity Superdome server with EVA storage while utilizing F5 BIG-IP for local traffic management. As subscriber bases continue to grow and as subscriber profiles for VoIP continue to demonstrate increased customer confidence and usage - and therefore increased demands on the VoIP infrastructure, the Superdome and BIG-IP provide a platform for long-term performance and scalability, and the CommuniGate Pro Dynamic Cluster with SIP Farm provides the software platform for growth-over-time and identity management for millions of users, enabling Telcos and Service Providers to easily administer their subscriber base and scale the infrastructure as needed with growth. Future benchmarks and customer case studies will continue this demonstration.

Contacts

CommuniGate Systems
Contact: Thom O'Connor, Director, Product Architecture
655 Redwood Hwy, Suite 275
Mill Valley, CA 94941 USA
T: 415 383 7164
Toll Free: 800 262 4722
F: 415 383 7461
http://www.communigate.com

Hewlett-Packard Company
Contact: Jamie Chui, Partner Alliance
19091 Pruneridge Avenue
Cupertino, CA 95014-0691 USA
T: 650 351 6596
http://www.hp.com

Navtel Communications, Inc.
Contact: Alvin C. Francis, Director of Marketing
55 Renfrew Drive
Markham, Ontario L3R 8H3
T: 905 948 7535
Toll Free Canada/US: 800 465 9400
F: 905 475 6524
http://www.navtelcom.com

F5 Networks
Contact: Craig Hovey, Business Development Technology Manager
401 Elliot Avenue West
Seattle, WA 98119
T: 206 272 5555
Toll Free: 888-888BIGIP
http://www.f5.com