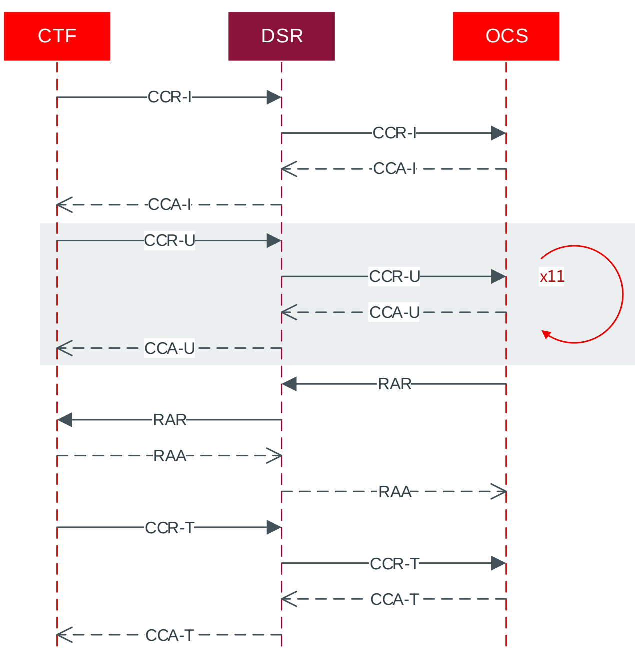
Oracle Communications Diameter Signaling Router Cloud Benchmarking Guide

Release 7.3.x

ORACLE COMMUNICATIONS | SEPTEMBER 2016

Table of Contents



CTF
DSR
OCS
CCR-I
CCR-I
CCA-I
CCA-I
CCR-U
CCR-U
x11
CCA-U
CCA-U
RAR
RAR
RAA
RAA
CCR-T
CCR-T
CCA-T
CCA-T

Table of Tables

Table of Figures

.

Acronyms

Table 1: Acronyms

Acronym	**Description**
COTS	Commercial Off The Shelf
CPU	Central Processing Unit
DA	Diameter Agent
DP	Database Processor
DSR	Diameter Signaling Routing
HDD	Hard Disk Drive
IDIH	Integrated Diameter Intelligence Hub
IPFE	IP Front End
KPI	Key Performance Indicator
MP	Message Processor
MPS	Messages Per Second
NIC	Network Interface Card
NOAM	Network Operations, Alarms, Measurements
NE	Network Element
OCDSR	Oracle Communications Diameter Signaling Router
P-DRA	Policy DIAMETER Routing Agent
RAM	Random Access Memory
SBR	Session Binding Repository
SBR(b)	SBR – subscriber binding database
SBR(s)	SBR – session database
SOAM	System (nodal) Operations, Alarms, Measurements
WAN	Wide Area Network

Terminology

Table 2: Terminology

Term	Description
1+1 Redundancy	For every 1, an additional 1 is needed to support redundant capacity. The specific redundancy scheme is not inferred (e.g. active-active, active-standby).
Geo-Diverse	Refers to DSR equipment located at geographically separated sites/datacenters
Geo-Redundant	A node at a geo-diverse location which can assume the processing load for another DSR signaling node(s)

Ingress Message Rate	A measure of the total Diameter messages per second ingressing the DSR. For this measure, a message is defined as any Diameter message that DSR reads from a Diameter peer connection independent of how the message is processed by the DSR.
Messages Per Second	A measure of the DSR Diameter message processing volume in messages per second. For this measure, a message is defined as: 1. DSR processing of an ingress Diameter message and either transmitting a single outgoing Diameter message or discarding the ingress message. The outgoing message may be a variant of, or a response to, the ingress message. 2. DSR transmission of any Diameter message, as required by DSR configuration, that is associated with incremental actions/events associated with #1 above. For example, the re-routing of a Request upon connection failure or the copying of a Request. Messages excluded from this measure are: • Diameter peer-to-peer messages: CER/CEA, DWR/DWA, and DPR/DPA • Ingress Diameter messages discarded by the DSR due to Overload controls • Answers received in response to Message Copy
N+K Redundancy	For every N, an additional K is needed to support redundant capacity. The specific redundancy scheme is not inferred (e.g. active-active, active-standby).
Node	A DSR node is a DSR signaling node (SOAM and subtending topology), an NOAM node or an SDS node. A node is synonymous with the network element (NE).
Site	A specific geographic location or datacenter where DSR application is installed.

References

1. Performance Best Practices for VMware vSphere® 5.5 (EN-000005-07)

2. DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN)

3. DSR Cloud Deployable Installation Guide - Available at Oracle.com on the Oracle Technology Network (OTN)

Introduction

The Oracle Communications Diameter Signaling Router (OCDSR or DSR) is deployable in the cloud as a Virtual Network Function (VNF). With DSR's added flexibility of being cloud deployable, operators must be able to manage the capacity and performance of the DSR in the cloud.

This document focuses on:

- How to benchmark DSR performance and capacity in a cloud deployed DSR
- Provides recommendations on performance tuning the DSR
- Provides benchmark data from our labs
- Provides information on the key metrics used to manage DSR performance and capacity
- Provides recommendations on how to use the data obtained from the metrics

About Cloud Deployable DSR

Oracle Communications Diameter Signaling Router (OCDSR or DSR) is deployed on a number of platforms. The DSR has a multiple deployment scenarios:

- » Bare-metal and hybrid (mixture of bare metal and virtual machines) - is the original deployment configuration of the DSR. It scales to very high performance and is widely deployed.
- » Fully virtualized - was introduced shortly after bare-metal. It provides virtualization of the DSR, but does not use a cloud manager, and does not co-reside with other applications. Provides a compact, cost-effective footprint and is widely deployed.
- » Cloud deployable –It provides full virtualization, assumes the DSR resources will be managed by a COTS cloud manager, and that the DSR can be one of many applications in the cloud. Cloud deployable DSR is the focus of this document.
- » Mix and match – DSR is a network of DSR signaling sites. The deployment infrastructure at each site can vary. E.g. bare-metal at one site, and then cloud deployed at another location.

What is a cloud deployable DSR?

A DSR that is ready and able to be deployed into a number of different cloud environments, including but not limited to:

- » A customer provided cloud infrastructure. The DSR is simply one of many applications.
- » A dedicated private cloud. The DSR may be the only application, or one of a small set of applications. Services and infrastructure may also be provided by Oracle and deployed at customer's sites. Often (but not necessarily) this is a deployment tuned specifically for the DSR.
- » A hosted cloud. The DSR is deployed in an Oracle or operator hosting cloud, and end-customers rent or lease the DSR application from the hosting provider.

Infrastructure matters

The DSR is capable of running on a huge variety of infrastructures, but not all infrastructures are the same and performance, capacity, and latency can vary dramatically based on the chosen infrastructure and how it is deployed. In general, the DSR works best in a high bandwidth, low-latency, high processing power environment (carrier grade cloud). Some considerations that impact DSR performance, capacity, latency:

- » Hardware – the CPUs, and NICs (network interface cards)
- » Hypervisor settings/configuration

» Uplinks, switches, WAN latency

DSR has excellent high availability and geo-diversity resiliency mechanisms that work in concert with cloud manager capabilities. Obviously the needed scale, availability, and resiliency of the deployment also impact the resource and infrastructure requirements.

Flexibility

DSR is flexibly deployed into many different clouds. It is unlikely that any two clouds are exactly the same and operators need to optimize for different reasons (e.g. power consumption may be critical for one operator, and WAN latency at another), varying sets of applications, and differing operational requirements. The performance and capacity of the DSR varies in each cloud, and the DSR application can no longer provide a guaranteed level of performance and capacity. However, the operator still needs to:

» Plan their networks – DSR's use resources, what impact will DSR have on their datacenters?

» Deploy DSR with predictable (if not exact) performance and capacity.

» Manage the capacity and performance of the DSR in their datacenters.

Methodology

There is a set of DSR specific tools, methods and documentation to assist in planning, deploying, and managing the capacity and performance of a cloud deployable DSR. This toolset is expected to be used in conjunction with information and tools provided by the infrastructure (hardware, cloud manager, hypervisor) vendors.

- » Planning for cloud deployable DSR
  - » Estimating required resources for a given DSR cloud deployment
    - Please contact your Oracle Sales Consultant. They have access to the DSR Cloud Dimensioning tool which estimates DSR cloud resources. This tool takes into account many factors not covered in this benchmarking guide, such as the overhead for optional DSR features not covered in the benchmarking guide, and recommended margins for redundancy.
  - » DSR Cloud Customer Documentation
    - Can be found with the DSR customer documentation at: http://docs.oracle.com/cd/E68457_01/index.htm
    - Look under the topic: "Cloud Planning, Installation, Upgrade, and Disaster Recovery"
- » Deploy DSR with predictable performance and capacity
  - » It is recommended that the DSR is run through a benchmark on the target cloud infrastructure to determine the likely capacity and performance in the target infrastructure. This information can then be used to adjust the initial deployment resources (if needed), and to help predict future resource requirements if and when the DSR grows.
  - » This document provides information on how to benchmark DSR performance and capacity. It also provides comprehensive benchmark results for a few select infrastructures. More benchmark data will be added to the document as it becomes available.
  - » This document also provides performance recommendations and observed differences for performance tuning decisions.
- » Manage the capacity and performance of the DSR
  - » The customer network is always changing- traffic patterns change, new applications are introduced. The infrastructure changes – new hardware, software/firmware updates. The operator needs to monitor and adjust the DSR resources for the changing conditions of the network and infrastructure.
  - » This document provides the key metrics and recommendations for monitoring the capacity and performance of a cloud deployed DSR.

Benchmarking Cloud Deployable DSR

This document is divided into several sections:

- Infrastructure Environment -This section provides details of the infrastructures used for the benchmark testing, including the hardware and software. It also describes key settings and attributes, and some recommendations on configuration.
- A benchmark section for each DSR server type - Each DSR server type is given independent treatment for its benchmark. Each section describes the traffic setup, and the observed results. It also provides metrics and guidelines for assessing performance on any infrastructure.

What to do with all this data?

This data is intended to provide guidance. Recommendations may need to be adapted to the conditions in a given operator's network. Each section below provides metrics that provide feedback on the running performance of the application.

When planning to deploy a DSR into any cloud environment, a few steps are recommended:

- Understand the initial deployment scenario for the DSR. Which features are planned, how much of what type of traffic? Of course, this may change once deployed, and the DSR can be grown or shrunk to meet the changing needs.
- Use the DSR cloud dimensioning tool to get an estimate of the types of DSR virtual servers needed, and an initial estimate of the quantity of the virtual machines and resources. Your Oracle Sales Consultant can run this tool for you based on your DSR requirements:
  - The tool allows for a very detailed model to be built of your DSR requirements including:
    - Required MPS by DIAMETER Application ID (S6a, Sd, Gx, Rx, etc.).
    - Required DSR applications such as Full Address Based Resolution (FABR) and Policy DRA (PDRA) and any required sizing information such as the number of subscribers supported for each application.
    - Any required DSR features such as Topology Hiding, Message Copy, IPSEC or Mediation that can affect performance.
    - Network-level redundancy requirements, such as mated pair DSR deployments where one DSR needs to support full traffic when one of the DSRs is unavailable.
    - Infrastructure information such as VMware vs. KVM, and Server parameters.
  - The tool will then generate a recommended number of VMs for each of the required VM types.
  - As noted below, these recommendations are just guidelines, since the actual performance of the DSR can vary significantly based on the details of the infrastructure.
- Based on the initial deployment scenario, determine if additional benchmarking is warranted:
  - For labs and trials, there is no need to benchmark performance and capacity if the goal of the lab is to test DSR functionality.
  - If the server hardware is different from the hardware used in this document then the performance differences can likely be estimated using industry standard metrics comparing single-threaded processor performance of the CPUs used in this document vs. the CPUs used in the customer's infrastructure. This approach will be most accurate for small differences in hardware (for instance different clock speeds for the same generation of Intel processors) and least accurate across processor generations where other architectural differences such as networking interfaces could also affect the comparison.
  - It is the operator's decision to determine if additional benchmarking in the operator's infrastructure is desired. Here's a few things to consider when deciding:
    - Benchmark infrastructure is similar to the operator's infrastructure, and the operator is satisfied with the benchmark data provided by Oracle.
    - Initial turn-up of the DSR is handling a relatively small amount of traffic and the operator prefers to measure and adjust once deployed.

- Operator is satisfied with the high-availability and geo-diversity of the DSR, and is willing to risk initial overload conditions, and will adjust once the DSR is production.

» If desired, execute benchmarking testing on the target cloud infrastructure. Only benchmark those types of DSR servers needed for the deployment (e.g. if full address resolution is not planned, don't waste time benchmarking the SDS, SDS SOAM, or DPs).

  » Once that benchmarking is completed, take a look at the data for each server type, and compare it to the baseline used for the estimate (from the cloud dimensioning tool).

    - If the performance estimate for a given DSR function is X and the observed performance is Y, then adjust the performance for that DSR function to Y.
    - Recalculate the resources needed for deployment based on the updated values.

» Deploy the DSR

» Monitor the DSR performance and capacity as described later in the document. As the network changes additional resources may be required. Increase the DSR resources as described later in this document as needed.

Infrastructure Environment

This section describes the infrastructure that was used for benchmarking. In general, the defaults or recommendations for hypervisor settings are available from the infrastructure vendors (e.g. ESXi vendor recommendations and defaults found in Performance Best Practices for VMware vSphere® 5.5 (EN-000005-07)); whenever possible the DSR recommendations align with vendor defaults and recommendations. Benchmarking was performed with the settings described in this section. Operators may choose different values; better or worse performance compared to the benchmarks may be observed. When recommendations other than vendor defaults or recommendations are made, additional explanations are included in the applicable section.

There is a subsection included for each infrastructure environment used in benchmarking.

KVM (QEMU) / Oracle X5-2 – Infrastructure Environment

There are a number of settings that affect performance of the hosted virtual machines. A number of tests were performed to maximize the performance of the underlying virtual machines for the DSR application.

A Host Hardware

  » Oracle Server X5-2

    - CPU Model: Intel(R) Xeon(R) CPU E5-2699 v3 @ 2.30GHz
    - RAM: 128 GB
    - HDD: 2.3 TB of solid state drive (SSD) storage
    - NICs:
      - 4 x Intel Ethernet Controller 10-Gigabit x540-AT2

B Hypervisor

  » QEMU-KVM Version: QEMU 1.5.3, libvirt 1.2.8, API QEMU 1.2.8

CPU Technology

CPU topology indicates the number of Sockets / Cores / Threads configured for the virtual CPUs on a given Virtual Machine. For the best performance results it is recommended to have a single core and a single thread. Therefore, for 8vCPU virtual machine the recommended topology would be 8 Sockets, 1 Core, and 1 Thread.

Recommendation: *a single core and a single thread*

Network Settings

Network Adapters

VirtIO is a virtualizing standard for network and disk device drivers where just the guest's device driver "knows" it is running in a virtual environment, and cooperates with the hypervisor. This enables guests to get high performance network and disk operations, and gives most of the performance benefits of para-virtualization.

Vhost-net provides improved network performance over Virtio-net by totally bypassing QEMU as a fast path for interruptions. The vhost-net runs as a kernel thread and interrupts with less overhead providing near native performance. The advantages of using the vhost-net approach are reduced copy operations, lower latency, and lower CPU usage.

Recommendation: *Vhost-net driver is recommended.*

BIOS Power Settings

KVM Openstack allows provides three options for power settings:

- Power Supply Maximum: The maximum power the available PSUs can draw
- Allocated Power: The power is allocated for installed and hot pluggable components
- Peak Permitted: The maximum power the system is permitted to consume

Recommendation: *Set to Allocated Power*

Disk Provisioning

The two preferred disk image file formats available when deploying a KVM virtual machine:

  - **QCOW2**: Disk format supported by the QEMU emulator that can expand dynamically and supports Copy on Write.
  - **Raw Dump Representation**: Unstructured disk image format

QCOW2 provides a number of benefits over raw dump such as:

- Smaller file size, even on filesystems which don't support holes (i.e. sparse files)
- Copy-on-write support, where the image only represents changes made to an underlying disk image
- Snapshot support, where the image can contain multiple snapshots of the images history

Recommendation: *QCOW2 (Since DSR does not involve processes which are disk I/O intensive.)*

Container format being chosen is bare – no container or metadata envelope for the disk image. The container format string is not currently being used by OpenStack components.

Recommendation: *Bare*

Guest Caching Modes

The operating system maintains a page cache to improve the storage I/O performance. With the page cache, write operations to the storage system are considered completed after the data has been copied to the page cache. Read operations can be satisfied from the page cache if the data requested is in the cache. The page cache is copied to permanent storage using fsync. Direct I/O requests bypass the page cache. In the KVM environment, both the host and guest operating systems can maintain their own page caches, resulting in two copies of data in memory.

The following caching modes are supported for KVM guests:

- **Writethrough**: I/O from the guest is cached on the host but written through to the physical medium. This mode is slower and prone to scaling problems. Best used for a small number of guests with lower I/O requirements. Suggested for guests that do not support a writeback cache (such as Red Hat Enterprise Linux 5.5 and earlier), where migration is not needed.
- **Writeback**: With caching set to writeback mode, both the host page cache and the disk write cache are enabled for the guest. Because of this, the I/O performance for applications running in the guest is good, but the data is not protected in a power failure. As a result, this caching mode is recommended only for temporary data where potential data loss is not a concern.
- **None [Selected]**: With caching mode set to none, the host page cache is disabled, but the disk write cache is enabled for the guest. In this mode, the write performance in the guest is optimal because write operations bypass the host page cache and go directly to the disk write cache. If the disk write cache is battery-backed, or if the applications or storage stack in the guest transfer data properly (either through fsync operations or file system barriers), then data integrity can be ensured. However, because the host page cache is disabled, the read performance in the guest would not be as good as in the modes where the host page cache is enabled, such as write through mode.
- **Unsafe:** The host may cache all disk I/O, and sync requests from guest are ignored.

Caching mode None is recommended for remote NFS storage, because direct I/O operations (O_DIRECT) perform better than synchronous I/O operations (with O_SYNC). Caching mode None effectively turns all guest I/O operations into direct I/O operations on the host, which is the NFS client in this environment. Moreover, it is the only option to support migration.

Recommendation: *Caching Mode = None*

Memory Tuning Parameters

Swappiness

The swappiness parameter controls the tendency of the kernel to move processes out of physical memory and onto the swap disk. Because disks are much slower than RAM, this can lead to slower response times for system and applications if processes are too aggressively moved out of memory.

- vm.swappiness = 0: The kernel will swap only to avoid an out of memory condition.
- vm.swappiness = 1: Kernel version 3.5 and over, as well as kernel version 2.6.32-303 and over; Minimum amount of swapping without disabling it entirely.
- vm.swappiness = 10: This value is recommended to improve performance when sufficient memory exists in a system.
- vm.swappiness = 60: Default
- vm.swappiness = 100: The kernel will swap aggressively.

Recommendation: *vm.swappiness = 10*

Kernel Same Page Merging

Kernel Same-page Merging (KSM), used by the KVM hypervisor, allows KVM guests to share identical memory pages. These shared pages are usually common libraries or other identical, high-use data. KSM allows for greater guest density of identical or similar guest operating systems by avoiding memory duplication. KSM enables the kernel to examine two or more already running programs and compare their memory. If any memory regions or pages are identical, KSM reduces multiple identical memory pages to a single page. This page is then marked copy on write. If the contents of the page is modified by a guest virtual machine, a new page is created for that guest.

This is useful for virtualization with KVM. When a guest virtual machine is started, it only inherits the memory from the host qemu-kvm process. Once the guest is running, the contents of the guest operating system image can be shared when guests are running the same operating system or applications. KSM allows KVM to request that these identical guest memory regions be shared.

KSM provides enhanced memory speed and utilization. With KSM, common process data is stored in cache or in main memory. This reduces cache misses for the KVM guests, which can improve performance for some applications and operating systems. Secondly, sharing memory reduces the overall memory usage of guests, which allows for higher densities and greater utilization of resources.

The following 2 Services control KSM:

- **KSM Service:** When the ksm service is started, KSM will share up to half of the host system's main memory. Start the ksm service to enable KSM to share more memory.
- **KSM Tuning Service:** The ksmtuned service loops and adjusts KSM. The ksmtuned service is notified by libvirt when a guest virtual machine is created or destroyed.

Recommendation: *'ksm' service set to active and 'ksmtuned' service running on KVM hosts*

Zone Reclaim Mode

When an operating system allocates memory to a NUMA node, but the NUMA node is full, the operating system reclaims memory for the local NUMA node rather than immediately allocating the memory to a remote NUMA node. The performance benefit of allocating memory to the local node outweighs the performance drawback of reclaiming the memory. However, in some situations reclaiming memory decreases performance to the extent that the opposite is true. In other words, in these situations, allocating memory to a remote NUMA node generates better performance than reclaiming memory for the local node.

A guest operating system causes zone reclaim in the following situations:

- When you configure the guest operating system to use huge pages.
- When you use Kernel same-page merging (KSM) to share memory pages between guest operating systems.

Configuring huge pages and running KSM are both best practices for KVM environments. Therefore, to optimize performance in KVM environments, it is recommended to disable zone reclaim.

Recommendation*: Disable Zone Reclaim*

Transparent Huge Pages

Transparent huge pages (THP) automatically optimize system settings for performance. By allowing all free memory to be used as cache, performance is increased.

***Recommendation**: Enable THP*

VMware (ESXi) / HP Gen8 – Infrastructure Environment

There are a number of ESXi (VMware hypervisor) settings that affect performance of the hosted virtual machines. A number of tests were performed to maximize the performance of the underlying virtual machines for the DSR application.

- Host Hardware
  - » HP ProLiant BL460c Generation 8+ (40 Cores with Hyper threading enabled)
    - CPU Model: Intel(R) Xeon(R) CPU E5-2670 v2 @ 2.50GHz
    - RAM: 128 GB
    - HDD: 832G
    - NICs:
      - 2 x HP FlexFabric 10Gb 2-port 534M Adapter
      - 2 x HP Flex-10 10Gb 2-port 530M Adapter
- Hypervisor
  - » ESXi Version: VMware-VMvisor-Installer-5.5.0.update02-2068190.x86_64

Virtual Sockets vs. Cores per Virtual Socket

When defining a virtual machine the number of vCPUs must be assigned to a server. The user has options for setting the number of “Virtual Sockets” and the number of “Cores per Virtual Socket”. The product of these two parameters determines the number of vCPUs available to the virtual machine.

In following the *VMware* best practice Performance Best Practices for VMware vSphere® 5.5 (EN-000005-07), the default value of 1 core per socket was used. This configuration is referred to as “wide” and “flat.” This will enable vNUMA to select and present the best virtual NUMA topology to the guest operating system, which will be optimal on the underlying physical topology.

***Recommendation:** 1 core per socket, virtual socket set to the number of vCPUs required by the server role.*

Network Settings

Network Adapters

There is a number of networking adapter choices when deploying a virtual machine:

- » **E1000**: This adapter is an emulated version of Intel 82545EM Gigabit Ethernet Controller. VMXNET3 adapter is the next generation of Para virtualized NIC designed for performance.
- » **VMXNET3**: This adapter has less CPU overhead compared to e1000 or e1000e. Also, VMXNET3 is more stable than e1000 or e1000e. . VMXNET3 adapter is the next generation of Para virtualized NIC designed for performance. This is the vSphere default setting.
  - VMXNET family implements an idealized network interface that passes network traffic between the virtual machine and the physical network interface cards with minimal overhead.

***Recommendation:** VMXNET3. No observable differences were noticed between E1000 and VMXNET3 for DSR application testing.*

Virtual Network Interrupt Coalescing and SplitRx Mode

- **Virtual network Interrupt Coalescing**: This option reduces number of interrupts thus potentially decreasing CPU utilization. This may however increase network latency. By default this is enabled in ESX 5.5.
- **SplitRxMode**: This option uses multiple physical CPUs to process network packets received in single network queue. By default this is enabled in ESX 5.5 for VMXNET3 adapter type.

Table 3: Virtual Network Interrupt Coalescing and SplitRX Mode

Network Setting:	Default	Virtual network Interrupt Coalescing: Disabled	SplitRxMode: Disabled
DSR.Cpu (Avg / Max)	~40.7% / ~44.5%	~42% / ~45.5%	~38.8% / ~40.6%
System.CPU_UtilPct (Avg / Max)	~44.4% / ~53%	~44.4% / ~55.5%	~41.8% / ~53.3%
Latency	Observed as same in DSR application benchmarking		

Recommendation: *Virtual network interrupt coalescing: Enabled, SplitRxMode: Enabled.*

Power Settings

VMware ESXi allows assignment of power management profiles. These profiles allow the user to configure the host to save power while balancing performance. The power management profiles use the host's processor ACPI power setting. Many host manufacturer's bios overrides the *ESXi* settings.

Table 4: Power Management Profiles

ESXi Power Mode:	High Performance	Balanced Performance
System.CPU UtilPct (Avg / Max)	~40% / ~60%	~38% / ~55%
Dsr.Cpu (Avg / Max)	~38% / ~48%	~36% / ~44%
Used % / Run % / System %	~472 / ~388 / ~49%	~462 / ~376 / ~49%
Wait % / Idle %	~407% / ~1013	~419% / ~1023

The data in the table above is collected from a DA MP, but similar trends are observed on the other DSR virtual server types. A small but significant difference was observed between balanced and high performance power settings. However, the data did not indicate a large enough deviation to vary from the hardware vendor's guidelines. DSR benchmark testing was performed with balanced performance settings.

Recommendation: *Refer to host hardware vendor power management guidelines for virtualization.*

Hardware Assisted Virtualization

VMware ESXi automatically determines if a virtual machine can use hardware support for virtualization based on processor type. Several settings were selected for assessing performance:

A. Automatic
B. Use software for instruction set and MMU virtualization [i.e.
C. Use Intel® VT-x/AMD-V™ for instruction set virtualization and software for MMU virtualization
D. Use Intel® VT-x/AMD-V™ for instruction set virtualization and Intel® EPT/AMD RVI for MMU virtualization

- Also testing with “Node Interleaving” setting Enabled [i.e. NUMA disabled], with no noticeable changes in performance.

Table 5: Virtualization Performance by Processor

MMU Virtualization Setting:	A.	B.	C.	D.
System CPU UtilPct (Max / Avg)	57.5/38%	71.5/43%	71.5/43%	53/38%
Dsr.Cpu (Max / Avg)	43.5/36.3%	50/38.6%	50/38.5%	43/36.3%

The data in the table above is provided from a DA MP. Trends for other servers are similar. The automatic (default) settings provide performance better than options B and C above, and fairly equivalent to option D.

Recommendation: *Refer to host hardware vendor guidelines for virtualization. Defaults recommended.*

Virtual Machine Storage Configuration

Storage Type Adapter

Testing was performed with the default “LSI Logic Parallel” option. No testing was performed against recent virtual SCSI adapters (LSI Logic SAS and VMware para-virtualized (PVSCSI.) At the time of testing the default was considered as the most stable and compatible.

Recommendation: *Default “LSI Logic Parallel”*

Disk Provisioning

The following disk provisioning options are available when deploying a virtual machine:

- **Thick Provision Lazy Zeroed**: All space needed for the VM is allocated during creation. Data on the host disk is zeroed out at a later time on first write from the virtual machine.
- **Thick Provision Eager Zeroed**: All space needed for the VM is allocated during creation. The data on the host disk will be zeroed out during creation. Time to deploy the virtual machine will be increased with this option. This option will support fault tolerant features provided by the infrastructure.
- **Thin Provision**: This option uses only the amount needed by the virtual machine disk. The image will grow as needed until the allocated capacity is reached.

With the high availability of the DSR, storage should be allocated at the time the VM is created, so thin provisioned is not recommended. When instantiating a fairly typical DSR VM with 60G of storage, the lazy zeroed disk was created almost instantaneously. Whereas the eager zeroed disk took about 7 minutes to initialize. Lazy zeroed is recommended.

Recommendation: *“Thick Provisioned Lazy Zeroed”*

Large Memory Pages

VMware ESXi Large-page support enables server applications to establish large-page memory regions. The use of large pages can potentially increase TLB access efficiency and thus improve program performance. By default Large page support is enabled on *VMware ESXi Server* although the full benefit of large pages comes only when guest OS and applications use them as well. By default large page support is not enabled in the DSR product.

The following settings were evaluated:

- Default settings [i.e. Large memory pages support enabled on host and large pages configured as 0 on guest]
- Large memory pages support enabled on host and 1024 large pages configured on guest
- Large memory pages support disabled on host

Recommendation: *Default settings. No visible advantage was observed when comparing iterative memory stats as observed through /proc/meminfo. No visible advantage could be observed in using large pages.*

Benchmark Testing

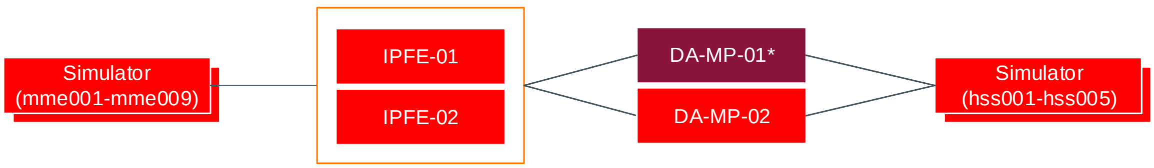
The way the testing was performed and the benchmark test set-up is the same for each benchmark infrastructure. Each section below describes the common set-up and procedures used to benchmark, and then the specific results for the benchmarks are provided for each benchmark infrastructure.

DA MP Relay Benchmark

Overview

This benchmarking case illustrates conditions for an overload of a DSR DA MP. Simulator message rate is increased until the DA-MP Overload mechanisms are triggered causing messages to be discarded. For this benchmarking scenario 9 MME clients and 5 HSS servers were created for s6a Diameter message pattern.

Topology



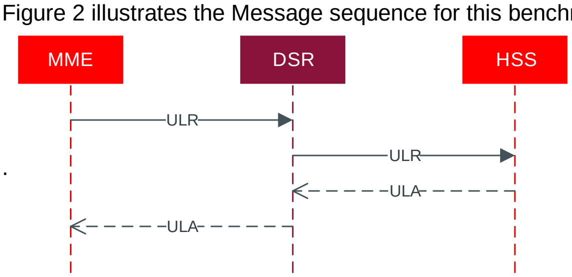
*System under test

Figure 1 – DA-MP Testing Topology

Figure 1 illustrates the topology used for this testing. For this testing all traffic traverses DA-MP-01. For DSR release 7.3 the maximum number of messages per second for each DA MP is an engineered configured value provided by the DA MP profile. With the hardware used for this benchmarking effort the engineered maximum (10k msg/s) was reached before the DA MP server became overloaded. As a workaround the number of configured vCPUs was decreased on the DA-MP-01 to “2”. This reduces the amount of available CPU and provides the means to illustrate the condition of a DA MP CPU overload.

Message Flow

Figure 2 illustrates the Message sequence for this benchmark case.



.

Figure 2 – DA-MP Message Sequence

Table 6: DA-MP Test Set-up

Messages	Traffic Details

Message	Distribution
ULR, ULA	100%

Detail	Distribution
Average Message Size	2.2K
Cross DSR Routing	75%

Observations

Figure 3 – DA-MP CPU Occupancy details the utilization of CPU of the affected DA MP. As the message rate increases the amount of CPU necessary to process these incoming messages increases. The DA MP will be marked as congested (around 60% CPU occupancy) and start dropping messages when this threshold is reached the DA MP may enter congestion level 1. Looking at the figure below, it should be observed that the average and maximum CPU occupancy using the KVM/ X5-2 configuration is much better than the VMware/Gen8 numbers due primarily to the faster processors in the newer X5-2.

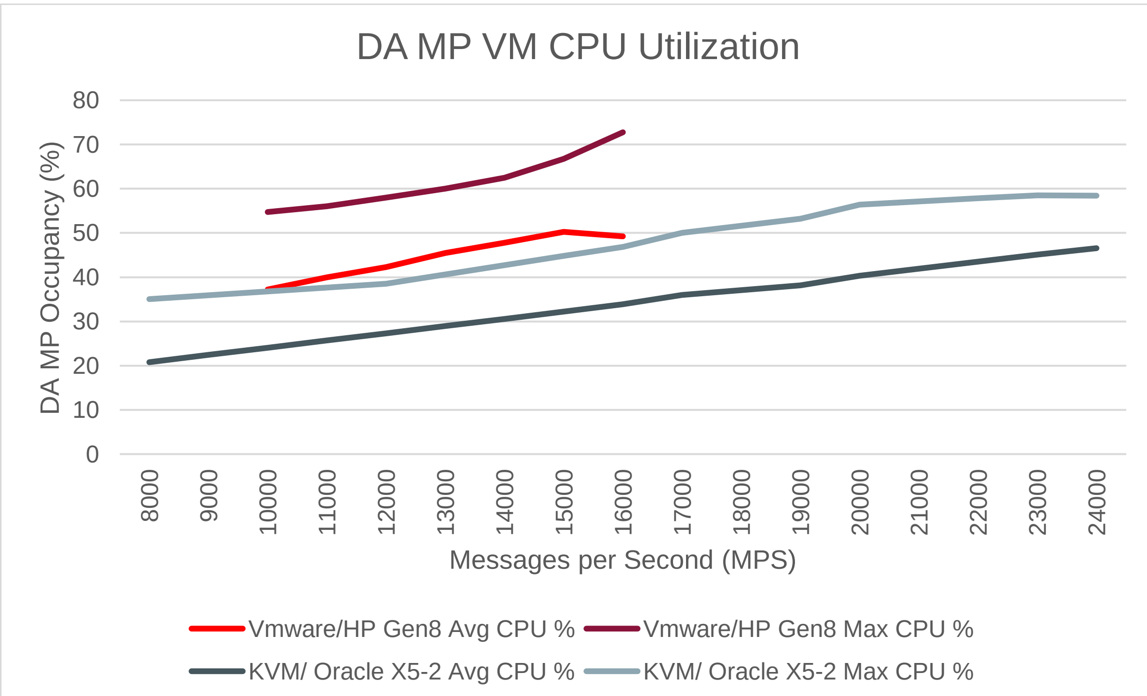


Figure 3 – DA-MP CPU Occupancy

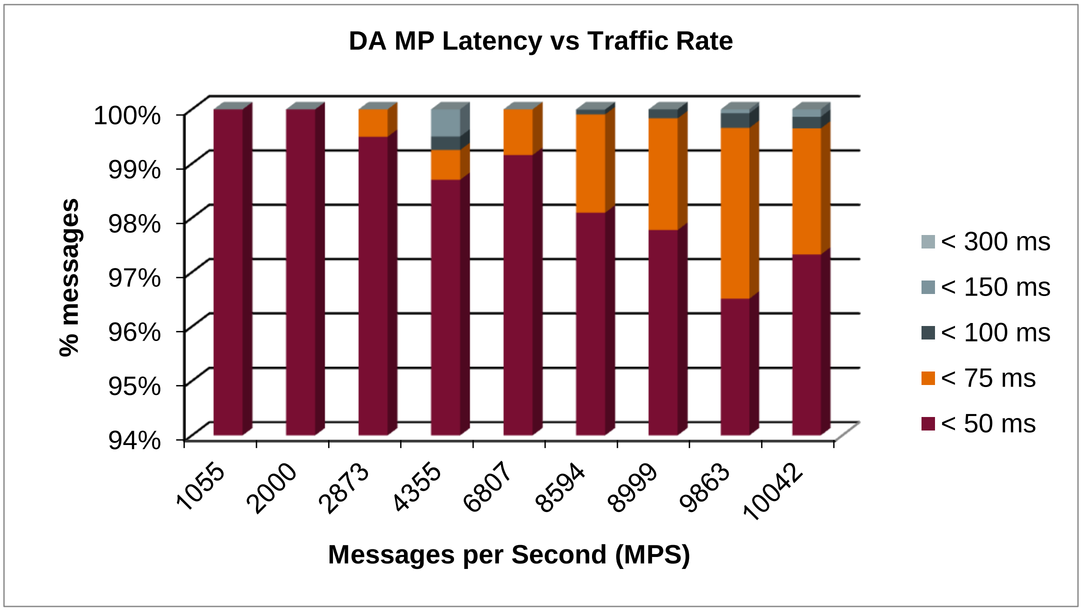


Figure 4 – DA-MP Round-Trip Latency for VMware/Gen8

It was observed as the number of incoming messages increased the round trip message latency increases. Figure 4 illustrates the increase in latency as the message rate increases. For example: at 2000 MPS 100% of the messages were observed to <50 ms of cross DSR latency. For this benchmark, CPU occupancy and latency stayed within acceptable levels up through the entire tested range.

Indicative Alarms / Events

During benchmark testing the following alarms/events were observed as it crossed into congestion.

Table 7: DA-MP Alarms/Events

Number	Severity	Server	Name	Description
5005	Minor	IPFE	IPFE Backend in Stasis	A backend server not accepting new connections but continuing to process existing connections
22200	Major	DA-MP	Communication Agent Routed Service Congested	The Diameter process is approaching or exceeding its engineered traffic handling capacity
22215	Major	DA-MP	Ingress Message Discarded: DA MP Overload Control	Ingress message discarded due to DA MP (danger of) CPU congestion.

Measurements

Measuring DA-MP Utilization

In this section, only the key recommended metrics for planning expansions of the DA-MP are discussed. There are many more measurements available on the DA-MP, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN).

The key metrics for managing the DA-MP are:

Table 8: DA-MP Utilization Metrics

Measure-ment ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
10204	EvDiameterProcessAvg	MP Performance	Server Group	Average percent Diameter Process CPU utilization (0-100%) on a MP server.	When running in normal operation with a mate in normal operation, and this measurement exceeds 30% of the rated maximum capacity, OR exceeds 60% of the rated capacity when running without an active mate.	If additional growth in the system is anticipated, then consider adding an additional DA-MP. It's possible that the traffic mix is different than originally dimensioned (e.g. 40% IPSEC instead of the originally dimensioning 5%). In these cases, re-assess the dimensioning with the actual traffic/application mix and add additional DA-MPs as needed.
10205	TmMpCongestion	MP Performance	Server Group	Total time (in milliseconds) spent in local MP congestion state	Any number greater than 0.	After eliminating any configuration, anomalous traffic spikes or major failure conditions, then is a late indication that additional DA MPs are needed. It is highly desirable that planning for additional DA-MPs happens before this condition occurs.
10133	RxMsgSizeAvg	Diameter Performance	Server Group	The average ingress message size in Diameter payload octets.	Average message size > 2000 bytes	DA-MP dimensioning assumes 2K average message size. This information is used to dimension IPFEs and DIH/IDIH. No action required if there are no alarms associated with the PDU message pool (available memory for messages). If PDU message pool is exhausting, contact Oracle.
31056	RAM_UtilPct_Average	System	System	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle.
31052	CPU_UtilPct_Average	System	System	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When running in normal operation with a mate in normal operation, and this measurements exceeds 30% of the rated maximum capacity, OR exceeds 60% of the rated capacity when running without an	If additional growth in the system is anticipated, then consider adding an additional DA MP. It's possible that the traffic mix is different than originally dimensioned (e.g. 40% IPSEC instead of the originally dimensioning 5%). In these cases, re-assess the dimensioning with the actual traffic and application mix and add additional DA-

					active mate.	MPs blades as needed.

Measuring DA-MP Connection Utilization

In this section, only the key recommended metrics for planning expansions of the DA-MP connections are discussed. There are many more measurements available on the DA-MP connections, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN).

The key metrics for managing the DA-MP connections are:

Table 9: DA-MP Connection Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
10500	RxConnAvgMPS	Connection Performance	connection	Average Ingress Message Rate (messages per second) utilization on a connection.	Minor alarm is set by default at 50%, major at 80%. Ingress message rate per connection is customer configurable with a max per connection of 10,000	Configure additional connections

Suggested Resolution

If congestion alarms shown in Table 7: DA-MP Alarms/Events then add additional DA-MPs to avoid CPU congestion. However, if the connection alarm shown in Table 9: DA-MP Connection Metrics is seen, than adding additional connections for that peer will help distribute the load and alleviate the connection alarm.

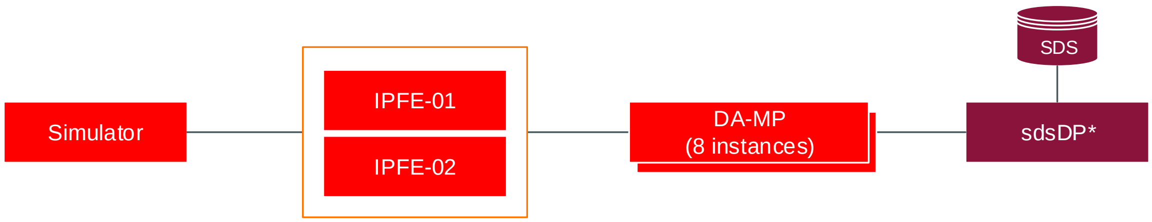
In general, the growth mechanism for DA MPs is via horizontal scaling. That is by adding additional DA MPs. The current maximum number of the DA MPs per DSR signaling NE is 16. If the number of DA MPs at given DSR Signaling NE is nearing its limits, the consider adding an additional DSR signaling NE. A DSR network supports up to 32 DSR signaling NEs.

SDS DP

Overview

This benchmarking case details a DP server type when overloaded. The role of the DP server is to process incoming messages from the DA MP and perform database lookups with a user defined key (IMSI, MSISDN, or Account ID and MSISDN or IMSI.) If the key is contained in the database, the DP will return the realm and FQDN associated with that key. The returned realm and FQDN can be used by the DSR Routing layer to route the connection to the desired endpoint.

Topology



*System under test

Figure 5 - SDS DP Testing Topology

To accomplish overload of the DP servers a number of configuration changes were made to the test setup.

» The Diameter request message size was reduced to 500 bytes. This allows the greatest number of queries to the DP without overloading the DA-MP.
» The IMSIs [in User-Name AVP] were selected to ensure they were not contained in the SDS Database. This ensures the greatest load on the DP database.
» DP Responses indicating NO MATCH are handled in FABR to send Diameter SUCCESS back to seagull. This reduces the egress routing overhead and CPU on the DA-MP.
» CPU thresholds were disabled on DA-MPs by setting engineered threshold to 86 / 88 / 90 / 92 for DOC / Minor / Major / Critical thresholds.

SDS DB Details

The SDS database was first populated with subscribers. This population simulates real-world scenarios likely encountered in a production environment and ensure the database is of substantial size to be queried against.

» SDS DB Size: 15 M MSISDNs / 15 M IMSIs / 15 M Subscriber Profiles
» AVP Decoded: User-Name for IMSI

Message Flow

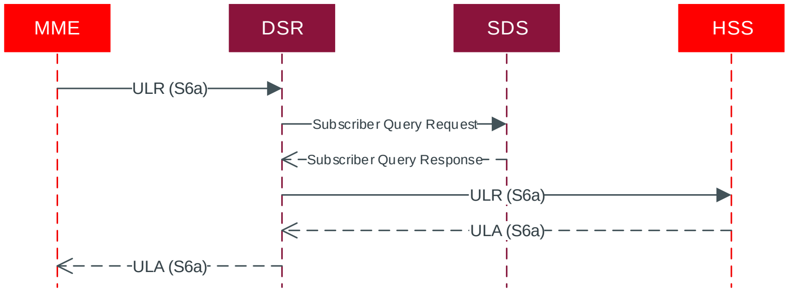


Figure 6 - SDS DP Message Sequence

Table 10: SDS DP Message Details

Messages

Message	Distribution
ULR, ULA	100%

Traffic Details

Detail	Distribution
Average Message Size	2.2K
Cross DSR Routing	75%

Observations

As the number of DP Lookups increases so does the CPU occupancy of the Database Processor (DP). In general, for processing functions, the recommendation is to keep the CPU occupancy at or below 60%. So in this case, the DP running on the benchmark infrastructure with the DP profile from DSR Cloud Deployable Installation Guide - Available at Oracle.com on the Oracle Technology Network (OTN), the expected capacity of the DP is ~28K lookups/s for VMware/HP Gen8. Typical only 50% of applicable traffic needs to do a lookup (i.e. responses don't need to lookup), and depending of the message flow, it could be less. But in this case, typically a single DP could support 28K x 2 = 56K MPS of address resolution traffic (e.g. S6a).

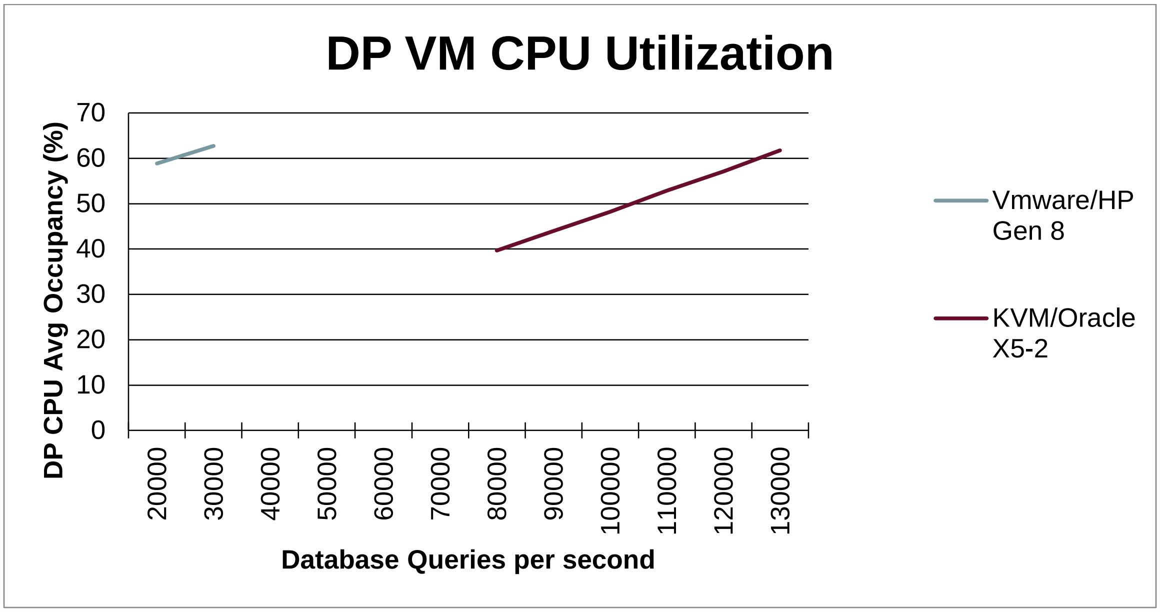


Figure 7 - DP CPU Util vs Traffic Rate (DP Queries)

Indicative Alarms / Events

Table 11: SDS DP Alarms/Events

Number	Severity	Server	Name	Description
19900	Minor	sdsDP	Process CPU Utilization	The Process, which is responsible for handling all traffic, is approaching or exceeding its engineered traffic handling capacity.
19822	Major	DA-MP	Communication Agent Routed Service Congested	Communication Agent Routed Service Congested
19825	Major / Critical	DA-MP	Communication Agent Transaction Failure Rate	The number of failed transactions during the sampling period has exceeded configured thresholds.
19826	Major	sdsDP	Communication Agent Connection Congested	Communication Agent Connection Congested
19831	Info	DA-MP	Communication Agent Service Operational State Changed	Communication Agent Service Operational State Changed, Instance DPService
19816	Info	DA-MP	Communication Agent Connection state Changed	Configuration Mode = Configured, Admin State = Enabled, Connect Mode = Server, Operational State = Degraded, Congestion Level = 1, Overload Level = 1, Transport Congestion Level = 0

Measurements

Measuring DP Utilization

In this section, only the key recommended metrics for managing the performance of the DP are discussed. There are many more measurements available on the DP, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN).

There are two key components of the subscriber database within a DSR Signaling node: the Database Processors (DPs), and OAM component which runs on the System OAM blades. The key metrics for managing the DPs are:

Table 12: SDS DP Utilization Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
4170	DpQueriesReceived	DP	System (per DP)	The total number of queries received per second.	When running in normal operation with a mate in normal operation, and this measurement exceeds 30% of the benchmarked maximum capacity, OR exceeds 60% of the benchmarked capacity when running without an active mate.	The operator should determine if additional growth in the number traffic requiring subscriber database look-ups is continuing to grow. If so, an estimate of the additional rate of database lookups should be calculated and additional DPs should be planned for.
31056	RAM_UtilPct_Average	System	System (per DP)	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
31052	CPU_UtilPct_Average	System	System (per DP)	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When running in normal operation with a mate in normal operation, and this measurements exceeds 30% of the rated maximum capacity, OR exceeds 60% of the benchmarked capacity when running without an active mate.	Oracle considers this measurement of lesser importance to the DpQueriesReceived. However, this measurement in conjunction with DpQueriesReceived can be used to indicate the need to add additional DPs.

While memory is a consideration for the DPs, the SDS provides the centralized provisioning for the entire DSR network.

The OAM application related to the DPs (DP SOAM) runs at each DSR Signaling NE requiring the Full Address Resolution feature. Currently these are fixed sized VMs with no horizontal or vertical scaling recommended as no need for scaling these VMs has been observed. The following two metrics should be monitored,

Table 13: DP SOAM Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
31056	RAM_UtilPct_Average	System	System (per DP SOAM)	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
31052	CPU_UtilPct_Average	System	System	The average CPU	If the average CPU	Contact Oracle

			(per DP SOAM)	usage from 0 to 100% (100% indicates that all cores are completely busy).	utilization exceeds 80% utilization	

Suggested Resolution

Add additional DP servers to accommodate database lookups.

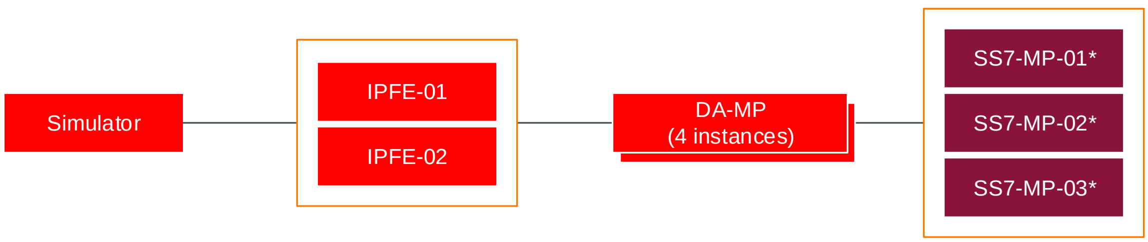
In general, the growth mechanism for DPs is via horizontal scaling. That is by adding additional DPs. The current maximum number of the DPs per DSR signaling NE is 10. This amount of scaling currently well exceeds capacities of the DA MPs driving queries to the DPs.

SS7 MP

Overview

This benchmarking case attempts to overload the SS7-MP server type. The SS7-MP server type is responsible for transforming messages between SS7 and Diameter protocols. Both Diameter and MAP messages were sent from the simulator to the DSR.

Topology



*System under test

Figure 8 - SS7 MP Testing Topology

» Additional DA-MPs were added to ensure DA-MPs were not the point of congestion.

Message Flow

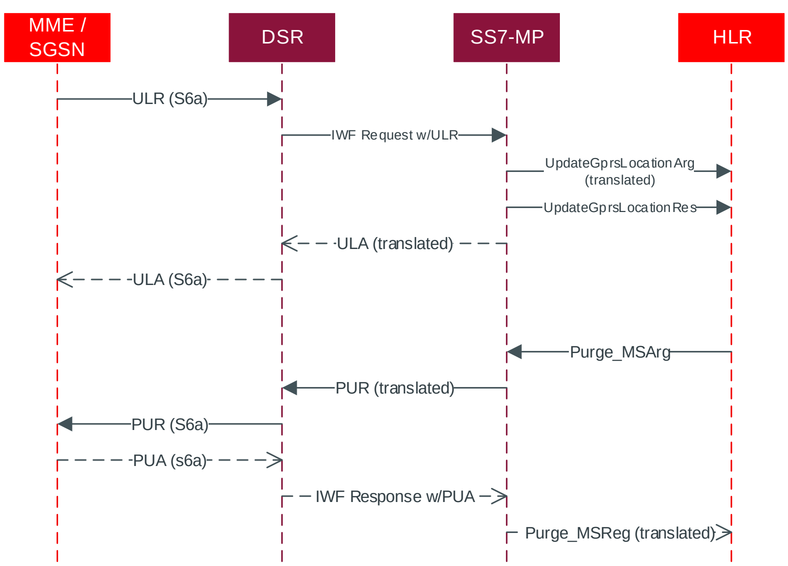


Figure 9 - SS7 MP Message Flow

Table 14: SS7 MP Message Detail

Detail	Distribution

Diameter originated	50%
MAP originated	50%

Observations

As SS7 ingress traffic increases the amount of CPU needed to process the incoming messages increases. For VMware/HP Gen8, it was observed SS7 process CPU Utilization alarms (19250) occurred when the SS7 MPs reached a rate of between 12K and 13K ingress messages per second. At 60% CPU occupancy, the SS7 MP is running 10-12K MPS. The equivalent data is shown for KVM/Oracle X5-2 in the figure below.

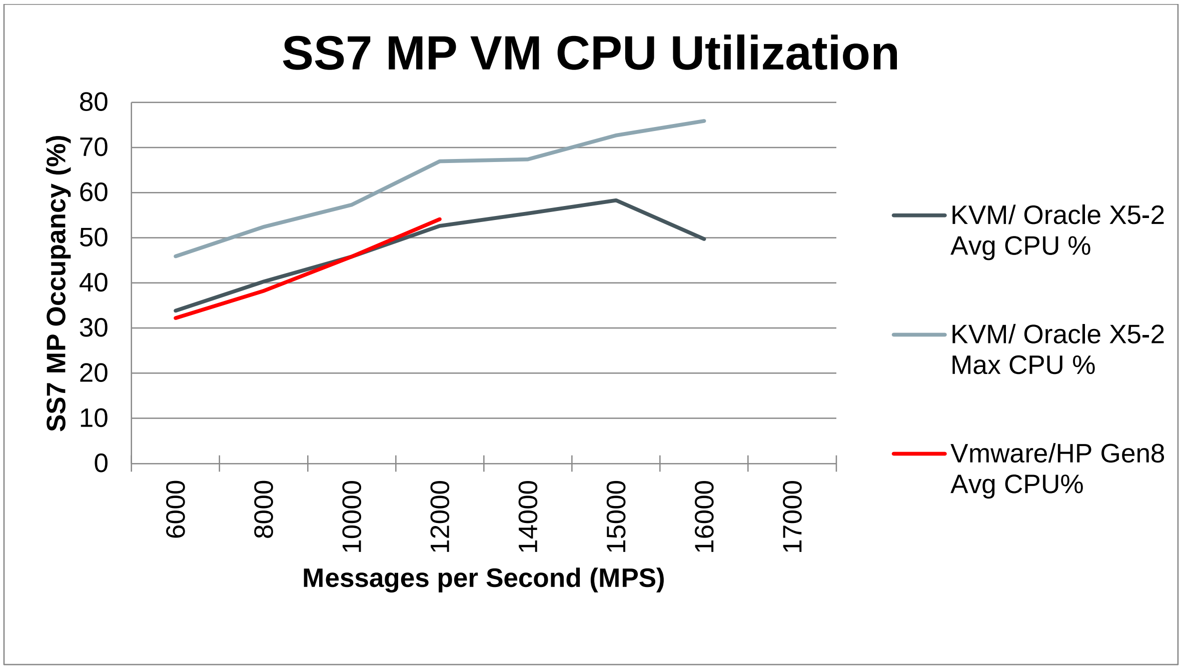


Figure 10 - SS7 MP CPU Occupancy(%) vs. MPS

Indicative Alarms / Events

Table 15: SS7 MP Alarms/Events

Number	Severity	Server	Name	Description
19250	Minor	DA-SS7MP	SS7 Process CPU Utilization	The SS7 Process, which is responsible for handling all SS7 traffic, is approaching or exceeding its engineered traffic handling capacity.

Measurements

Measuring SS7 MP Utilization

In this section, only the key recommended metrics for planning expansions of the SS7 MP are discussed. There are many more measurements available on the SS7 MP, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN).

The key metrics for managing the SS7 MP and associated IWF DA-MP are:

Table 16: SS7 MP Metrics

Measurement ID (or KPI ID)	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
11054	MAP Ingress Message Rate	MAP Diameter Interworking function (MD-IWF)	SS7 MP	Average number of MAP messages (both requests and responses) received per second from SS7 network	When running in normal operation with a mate in normal operation, and this measurement exceeds 30% of benchmarked maximum, OR exceeds 60% of the benchmarked capacity when running without an active mate.	If additional growth in MAP traffic is expected, an estimate of the traffic should be calculated and additional SS7 MPs should be planned for. This condition could also be an anomalous spike in traffic, and the operator may choose to ignore the occurrence. Above 40% in normal operation indicates an immediate need for additional SS7 MPs
31056	RAM_UtilPct_Average	System (SS7 MP)	System	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
31052	CPU_UtilPct_Average	System (SS7 MP)	System	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When running in normal operation with a mate in normal operation, and this measurements exceeds 30%, OR exceeds 60% when running without an active mate.	If additional growth in the system is anticipated, then consider adding an additional SS7 MP.

Suggested Resolution

Add additional SS7 MPs to accommodate additional MAP traffic.

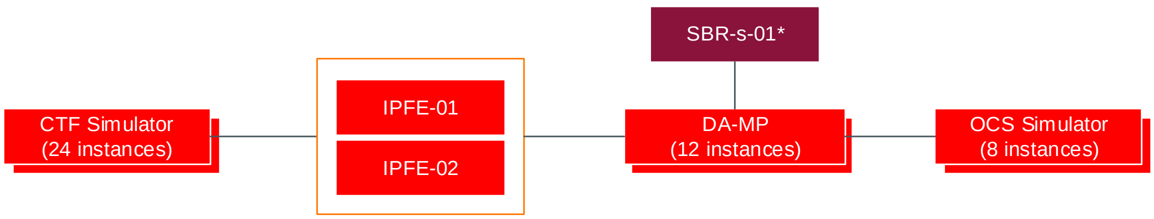
In general, the growth mechanism for SS7 MPs is via horizontal scaling. That is by adding additional SS7 MPs.

OCDRA Session SBR

Overview

The SBR is benchmarked running both as session and subscriber SBR roles. The session SBR information is shown first. The SBR was tested with 50 million sessions on the single SBR session server group.

Topology



*System under test

Figure 11 - SBR Testing Topology

» The number of DA-MPs was increased from 8 to 12 to generate enough traffic to cause congestion on SBR(s). Further increase in numbers of DAs is not possible due to hardware constraints.

» 12 DA-MPs were not able to generate enough traffic to cause congestion of the SBR(s). To accomplish overload the SBR(s) resource profile was reduced from 12 vCPUs to 8 vCPUs.

» Topology Hiding was enabled for this benchmark case to create additional overhead on the SBR(s).

» 24 PCEF, 32 AF clients and 6 PCRF servers were used for traffic generation, though the system contained a large number of connections and routing configuration (as part of a larger testing configuration).

» Executed on a Geo-diverse setup where site 2 had spare SBR(s). Only DA-MPs on site1 were used to generate traffic.

Message Flow

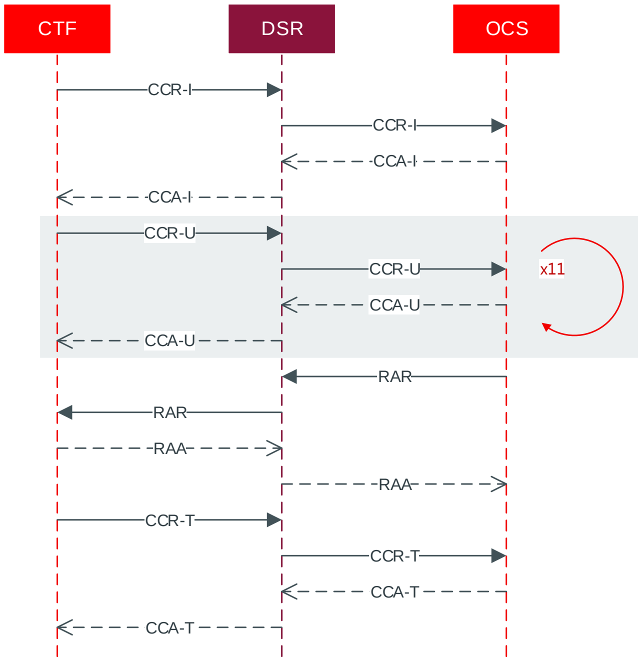


Figure 12 - SBR Gy Message Sequence

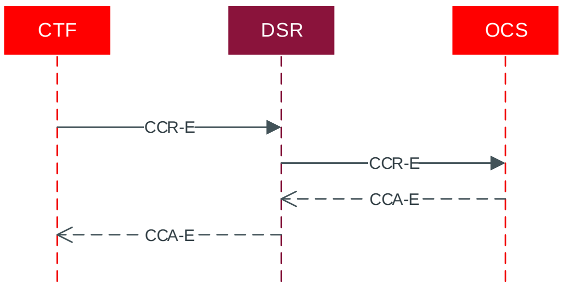


Figure 13 - SBR Gy Message Sequence for SMS

Table 17: SBR Test Set-up

Message	Distribution
CCR-I, CCA-I	6.5%
CCR-U, CCA-U	70.5%
CCR-T, CCA-T	6.5%

RAR, RAA[1]	6.5%
CCR-E, CCA-E[2]	10%
STR, STA	12.8%
Rx RAR, RAA	6.4%

Observations

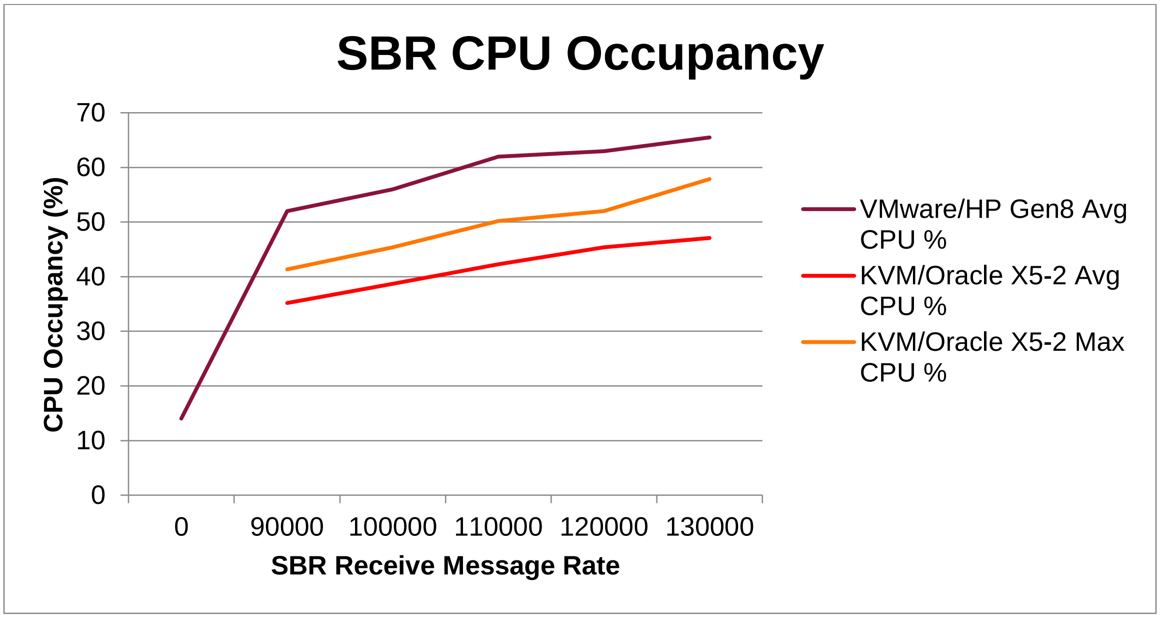


Figure 14 - SBR CPU Occupancy vs. Messages per Second

Like with the other messaging intensive components of the DSR, this processor too is recommended to stay at or below 60% CPU occupancy for planning purposes. From Figure 14, with 5 Million sessions, the SBR runs about ~100K receive messages per second at 60% CPU occupancy on VMware/HP Gen8. Notice how much lower the average occupancy is for KVM/Oracle X5-2.

Indicative Alarms / Events

Table 18: SBR Alarms/Events

Number	Severity	Server	Name	Description
19825	Minor / Major / Critical	DA-MP	Communication Agent Transaction Failure Rate	The number of failed transactions during the sampling period has exceeded configured thresholds.
22106	Major	DA-MP	Ingress Message Discarded: DA MP Ingress Message Rate Control	Ingress message discarded due to connection (or DA-MP) ingress message rate exceeding connection (or DA-MP) maximum ingress MPS.

[1] Assumed to be 1 transaction per OCS Session
[2] CCR-E / CCA-E usage is associated with sending and receiving SMSs and is assumed to be 10% of total traffic

22328	Minor / Major	DA-MP	Ingress MPS Rate	The diameter connection specified in the alarm instance is processing higher than normal ingress messaging rate.
22726	Minor / Major / Critical	SBR(s)	SBR Queue Utilization Threshold Exceeded	SBR server stack event queue utilization threshold has been exceeded.

Measurements

Key metrics for managing the Session SBR(s) are:

Table 19: SBR Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
31052	CPU_UtilPct_Average	System	System (SBR(s))	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When this measurement exceeds 60% of the benchmarked capacity.	Contact Oracle
31056	RAM_UtilPct_Average	System	SBR(s)	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
11372	SbrPolicySessionRecsAvg	SBR Session Performance	Server Group	The number of policy sessions in progress	If P-DRA function is enabled and OC-DRA is not enabled and average exceeds benchmarked capacity. If both P-DRA and OC-DRA are enabled this average must be combined with the SbrOcSessionRecsAvg and the combined average exceeds benchmarked capacity.	Contact Oracle
11441	SbrOcSessionRecsAvg	SBR Session Performance	Server Group	The number of online Charging sessions in progress	If OC-DRA function is enabled and P-DRA is not enabled and average exceeds benchmarked capacity. If both P-DRA and OC-DRA are enabled this average must be combined with the SbrPolicySessionRecsAvg and the combined average exceeds benchmarked capacity.	Contact Oracle

Suggested Resolution

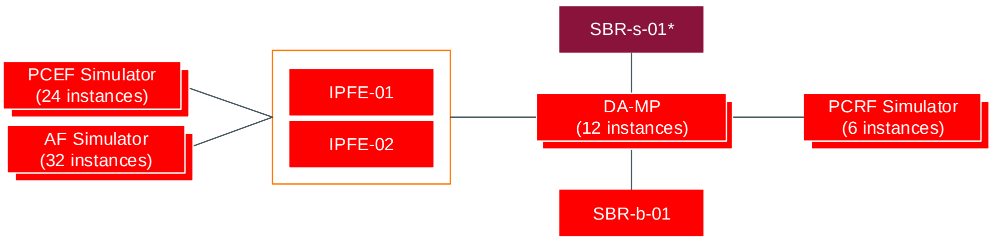
If either additional Sessions or MPS capacity is required is required then additional Server Groups may be added to an existing SBR(s) using the SBR reconfiguration feature. There can be up to 8 Server Groups in the SBR(s).

PDRA Subscriber Binding SBR (SBR(b))

Overview

The SBR (b) is only used with a PDRA application. The SBR (b) is a network scoped resource. Deployments with many DSR signaling nodes and PDRA may only have the SBR (b) at a subset of the sites. The SBR (b) is benchmarked as described below.

Topology



*System under test

Figure 15 - SBR (b) Testing Topology

» 12 DA-MPs were used to generate enough traffic which can cause congestion on the SBR(s). Further increase in numbers of DA-MPs was not possible due to hardware constraints.

» 12 DA-MPs were not able to generate enough traffic to cause congestion on the SBR(s). To accomplish congestion the resource profile on the SBR(s) was changed from the required 12 vCPU to 8 vCPU.

» 24 CTF clients and 7 OCS servers were created on the simulator, though system has a large number of configured connections and routing configuration (inherited from previous testing).

» This benchmark case was executed on a Geo-diverse setup where site 2 had spare SBR(s). Only DA-MPs on site1 were used to generate traffic.

Message Flow

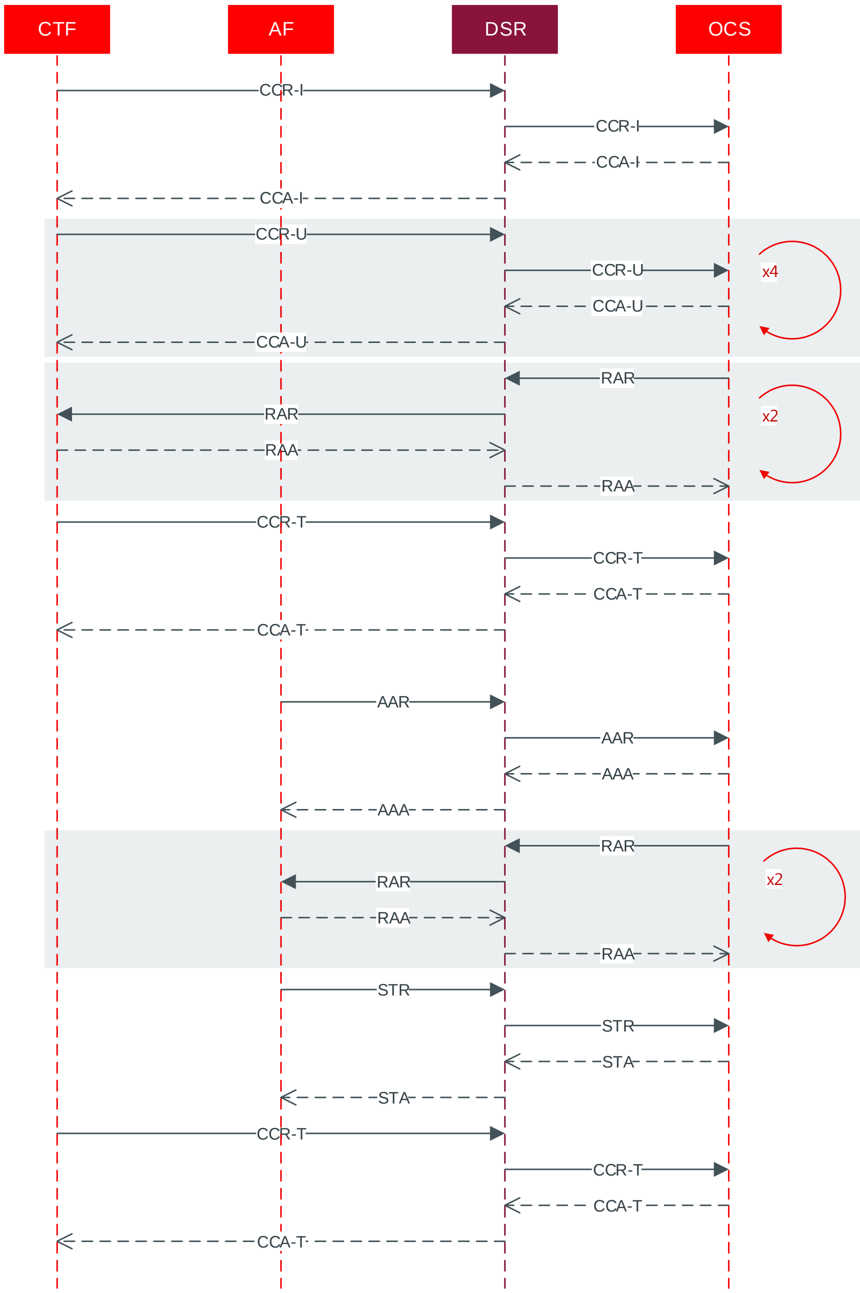


Figure 16 - SBR (b) Message Sequence

Table 20: SBR (b) Test Set-up

Messages

Message	Distribution
CCR-I, CCA-I	8.5%
CCR-U, CCA-U	25.5%
CCR-T, CCA-T	8.5%
Gx RAR, RAA	25.5%
AAR, AAA Initial	12.8%
STR, STA	12.8%
Rx RAR, RAA	6.4%

Traffic Details

Detail	Distribution
Gx w/ IPv6 Alternate Key	100%
Gx w/ IPv4 Alternate Key	0%
Gx with MSISDN Alternative Key	100%
Gx Topology Hiding	0%
Rx Topology Hiding	0%

Observations

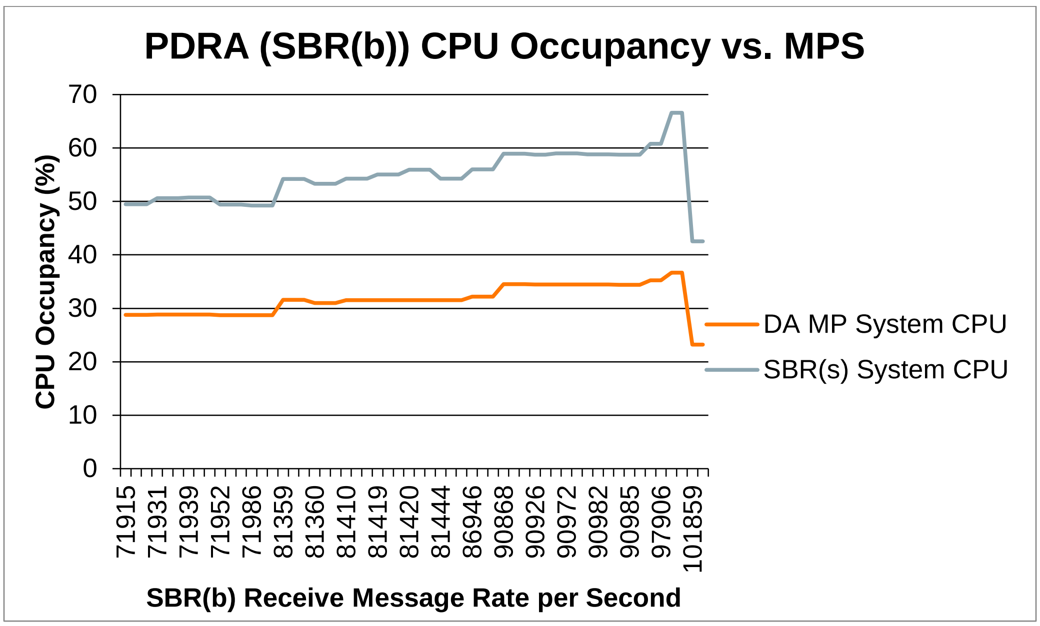


Figure 17: SBR (b) CPU Occupancy vs. Message Rate

Like with the other messaging intensive components of the DSR, this processor too is recommended to stay at or below 60% CPU occupancy for planning purposes. From Figure 17 with 5 Million sessions, the SBR (b) runs about ~100K (VMware/HP Gen8) receive messages per second at 60% CPU occupancy. Data was not collected for KVM/Oracle X5-2, however it was observed that the SBR (b) is well below 60% CPU occupancy at 125K MPS.

Indicative Alarms / Events

Table 21: SBR (b) Alarms/Events

Number	Severity	Server	Name	Description
19825	Minor / Major / Critical	DA-MP	Communication Agent Transaction Failure Rate	The number of failed transactions during the sampling period has exceeded configured thresholds.

19826	Major	DA-MP, SBR(s)	Communication Agent Connection Congested	Communication Agent Connection Congested
19846	Major	DA-MP, SBR(s)	Communication Agent Resource Degraded	Communication Agent Resource Degraded
22051	Critical	SOAM	Peer Unavailable	Unable to access the Diameter Peer because all of the diameter connections are Down.
22101	Major	SOAM	Connection Unavailable	Connection is unavailable for Diameter Request/Answer exchange with peer.
22715	Minor	SBR(s)	SBR Audit Suspended	SBR audit is suspended due to congestion.
22725	Minor / Major	SBR(s)	SBR Server In Congestion	SBR server operating in congestion.
22732	Minor / Major	SBR(s)	SBR Process CPU Utilization Threshold Exceeded	SBR process CPU utilization threshold has been exceeded.

Measurements

Key metrics for managing the Session SBR (b) blades are:

Table 22: Session SBR (b) Blades Metrics

Measurement ID	**Name**	**Group**	**Scope**	**Description**	**Recommended Usage**	
					Condition	Actions
31052	CPU_UtilPct_Average	System	System (SBR)	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When this measurement exceeds 60% utilization	Contact Oracle
31056	RAM_UtilPct_Average	System	SBR	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
11372	SbrPolicySessionRecsAvg	SBR Session Performance	Server Group	The number of policy sessions in progress	If P-DRA function is enabled and OC-DRA is not enabled and average exceeds benchmarked capacity. If both P-DRA and OC-DRA are enabled this average must be combined with the SbrOcSessionRecsAvg and the combined average exceeds benchmarked capacity.	Contact Oracle
11441	SbrOcSessionRecsAvg	SBR Session Performance	Server Group	The number of online Charging sessions in progress	If OC-DRA function is enabled and P-DRA is not enabled and average exceeds benchmarked capacity. If both P-DRA and OC-DRA are enabled this average must be combined with the SbrPolicySessionRecsAvg and the combined average exceeds benchmarked	Contact Oracle

					capacity.	

Key metrics for managing the Binding SBR (b) servers are:

Table 23: Binding SBR (b) Server Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
31052	CPU_UtilPct_Average	System	System (blade)	The average CPU usage from 0 to 100% (100% indicates that all cores are completely busy).	When this measurement exceeds 60% occupancy.	Contact Oracle
31056	RAM_UtilPct_Average	System	Blade	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle
11374	SbrPolicyBindingRecsAvg	SBR Binding Performance	Server Group	Average number of active SBR Policy bindings.	When this average exceeds benchmarked capacity.	Contact Oracle

Suggested Resolution

If either additional Bindings or MPS capacity is required is required then additional Server Groups may be added to an existing SBR(b) using the SBR reconfiguration feature. There can be up to 8 Server Groups in the SBR(b).

NOAM

Overview

Specific benchmark data for the DSR NOAM is not provided in this release as the DSR Cloud deployable footprint is modest and system testing of the DSR indicates that NOAM growth in not currently needed.

Indicative Alarms / Events

The DSR Network OAM is potentially a RAM intensive function. The Network OAM is designed not to exceed the available memory; however RAM is the most likely resource constraint.

Measurements

Measuring Network OAM Utilization

In this section, only the key recommended metrics for managing the performance of the Network OAM are discussed. There are many more measurements available, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN)[2].

The key metric for managing the Network OAM Servers are:

Table 24: Network OAM Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
31056	RAM_UtilPct_Average	System	System (server)	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle

Suggested Resolution

The NOAM can be vertically scaled; however this action is not anticipated to be necessary with the DSR 7.1.1 cloud deployable footprint. Please contact Oracle support for additional guidance as needed.

SOAM

Overview

Specific benchmark data for the DSR NOAM is not provided in this release as the DSR Cloud deployable footprint is modest and system testing of the DSR indicates that NOAM growth in not currently needed.

Indicative Alarms / Events

A key metric for managing the System OAM blades is:

Table 25: System OAM Metrics

Measurement ID	Name	Group	Scope	Description	Recommended Usage	
					Condition	Actions
31056	RAM_UtilPct_Average	System	Blade	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle

Suggested Resolution

Vertical and horizontal scaling of the DSR is not supported or indicated in this release. Please contact Oracle support for additional guidance as needed.

IPFE

Overview

The IPFE was exercised in both VMware and KVM environments. Table 26 shows the measurement capacity of the IPFE. Note that there are three main factors that determine the throughput limits:

- The number of TSAs (one or more) on the IPFE
- Whether there are more than 2,000 connections
- Whether the average message size is less than the MTU size.

Under most conditions the throughput of the IPFE is 1Gbit/sec. However under the worst case of all three of the above conditions the throughput of the IPFE drops to 800 Mbits/Sec.

Note that when monitoring IPFE capacity that since much of the IPFE work is done at the kernel level, the CPU utilization numbers returned by the IPFE application level don't fully reflect all of the IPFE overhead on the system.

	Single TSA on IPFE Pair		2 or more TSAs on IPFE Pair	
	Avg Msg Size < 1 MTU	**Avg Msg Size >= 1 MTU**	**Avg Msg Size < 1 MTU**	**Avg Msg Size >= 1 MTU**
2,000 Connections or less	1 Gbit/sec	1 Gbit/sec	1 Gbit/sec	1 Gbit/sec
More than 2,000 Connections	1 Gbit/sec	800 Mbits/sec	1 Gbit/sec	1 Gbit/sec

Table 26: IPFE Throughput

Indicative Alarms / Events

In this section, only the key recommended metrics for managing the performance of the IPFE are discussed. There are many more measurements available on the IPFE, and these can be found in DSR Alarms, KPIs, and Measurements – Available at Oracle.com on the Oracle Technology Network (OTN).

Measurements

The key metrics for managing the IPFE blades are:

Table 27: IPFE Metrics

Measurement ID	**Name**	**Group**	**Scope**	**Description**	**Recommended Usage**	
					Condition	Actions
5203	RxIpfeBytes	IPFE Performance	Server Group	Bytes received by the IPFE	If the number of (bytes * 8 bits/byte)/(time interval in s) is > benchmarked capacity (Gbps)	If the traffic is expected to grow then, consider adding an additional IPFE pair
31052	CPU_UtilPct_Average	System	System	The average CPU usage from 0 to	When running in normal operation with a mate in	Contact Oracle

			(IPFE)	100% (100% indicates that all cores are completely busy).	normal operation, and this measurements exceeds 30% occupancy, OR exceeds 60% occupancy when running without an active mate.	
31056	RAM_UtilPct_Average	System	System (IPFE)	The average committed RAM usage as a percentage of the total physical RAM.	If the average Ram utilization exceeds 80% utilization	Contact Oracle

Suggested Resolution

Horizontal scaling by adding an additional pair of IPFEs per DSR Signaling NE as indicated.

IDIH

Overview

The IDIH (IDIH application, IDIH mediation, IDIH Database VMs) are considered a best effort trouble shooting tool for the DSR. Benchmarking data is not currently provided for the IDIH VMs.

Suggested Resolution

Contact Oracle support.

Appendix A: Summary of Benchmark Data

The information shown below is a summary of the benchmark data described throughout the document. This data is intended to provide guidance. Recommendations may need to be adapted to the conditions in a given operator's network.

The data below summarizes the observed results based on the test setups described throughout this document.

Table 28: Benchmark Data Summary

Benchmark characteristics:

Benchmark Run	Benchmark A: VMware/HP Gen8	Benchmark B: KVM/ Oracle X5-2
Application Software	DSR 7.0.1 (running Oracle Linux)	DSR 7.1.1 (running Oracle Linux)
Host VM	VMware (VMware-VMvisor-Installer-5.5.0.update02-2068190.x86_64)	KVM (QEMU 1.5.3)
HW	HP Gen8 blades	Oracle Server X5-2
CPU Clock Speed	2.5 GHz	2.30GHz
VM Profiles/Flavors	Per reference Performance Best Practices for VMware vSphere® 5.5 (EN-000005-07)	Per DSR Cloud Installation Guide

VM Name	VM Purpose	Benchmarked Performance		
		Unit	Quantity	
			Benchmark A: VMware/HP Gen8	Benchmark B: KVM/ Oracle X5-2
DSR NOAM	Network Operation, Administration, Maintenance (and Provisioning)	VM	1+1	1+1
DSR SOAM	Site (node/Network Element) Operation, Administration, Maintenance (and Provisioning)	VM	1+1	1+1
DA MP	Diameter Agent Message Processor	MPS	10,000	24,000
DA MP w/ IWF	Diameter Agent Message Processor with IWF	MPS	10,000	24,000
IPFE	IP Front End	megabits/s	800	800
SS7 MP	SS7 Message Processor for MAP Diameter interworking function	MPS	10,000	12,000
SBR(s)	Subscriber Binding Repository (session) for Policy DRA	Diameter sessions	5,000,000	5,000,000
		receive msg/s	100,000	125,000[1]
SBR(b)	Subscriber Binding Repository (binding) for Policy DRA	session binds	5,000,000	5,000,000
		receive msg/s	100,000	>125,000[1]
DP SOAM	Database Processor Site (node) Operation, Administration, Maintenance for address	VM	1+1	1+1

	resolution and subscriber location functions			
DP	Database Processor for address resolution and subscriber location functions	MPS requiring DP lookups (usually 50% of FABR traffic)	28,000	80,000
SDS	Subscriber Database Processor for address resolution and subscriber location functions	subscriber entities	30,000,000	30,000,000
Query Server	Allows customers to query FABR subscriber data via a MySQL interface	N/A	N/A	N/A
iDIH Application	Integrated Diameter Intelligence Hub for troubleshooting	VM	1	1
iDIH Mediation	Integrated Diameter Intelligence Hub for troubleshooting	VM	1	1
iDIH Database	Integrated Diameter Intelligence Hub for troubleshooting	VM	1	1

[1]Note that other functions such as Suspect Binding Audits need to have capacity reserved for them. A better limit for a production system would be 50,000 MPS of actual traffic.

Appendix B: Detailed Infrastructure Settings

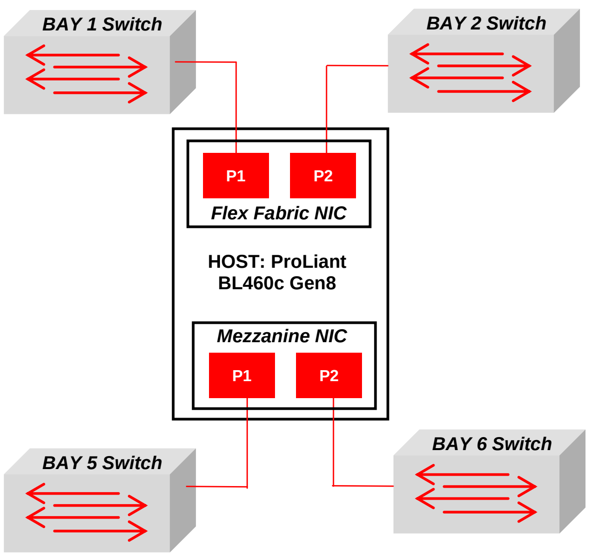
Table 29: Detailed Infrastructure Settings

Attribute	Particulars – VMware/HP Gen8	Particulars – KVM/Oracle X5-2
Model	ProLiant BL460c Gen8	Oracle Server X5-2
Processor Type	Intel(R) Xeon(R) CPU E5-2670 v2 @ 2.50GHz	Intel(R) Xeon(R) CPU E5-2699 v3 @ 2.30GHz
CPU Cores	40 [2 CPU Sockets [10 x 2 Cores, each with Hyper threading Active]	72 [2 CPU Sockets [18 x 2 Cores, each with Hyper threading Active]
RAM	131072 MB [DDR3-1866 Hz]	128 G [DDR3-2133 Hz]
CPU Cache Memory	25MB (Intel® Smart Cache) 5MB (1x5MB) Level 3 cache	L1d cache: 32K L1i cache: 32K L2 cache: 256K L3 cache: 46080K
Memory	131072 MB [DDR3 1866MHz RDIMMs at 1.5V]	131723288 kB [8 [out of 24] DIMM slots installed with 16384 MB each] Speed: 2133 MHz Type: DDR4
Number and Type of NICs	2 x HP FlexFabric 10Gb 2-port 534M Adapter 2 x HP Flex-10 10Gb 2-port 530M Adapter [Mezzanine Slot(s)]	4 [Intel Corporation Ethernet Controller 10-Gigabit X540-AT2]
BIOS Power Settings	HP Static High Performance Mode	Power Supply Maximum: Maximum power the available PSUs can draw Allocated Power: Power allocated for installed and hot pluggable components Peak Permitted: Maximum power the system is permitted to consume (set to Allocated Power)
HDD	HP Smart Array P220i Controller Disk Drive Interface *6Gb/s SAS (Serial Attached SCSI)* Cache Memory *512MB flash backed write cache (FBWC) cache standard* RAID Support *RAID 1 (mirroring)*	2.3 TB of solid state drive (SSD) storage
Hypervisor Information	VMware-VMvisor-Installer-5.5.0.update02-2068190.x86_64 (ESXi)	KVM - QEMU 1.5.3
Cloud Manager Information	vSphere Client Version 5.5.0 VMware vCenter Server Version 5.5.0	OpenStack release Kilo [Release: 2015.1.1, Release date: July 30, 2015]
DSR Version	DSR 7.0.1	DSR 7.1.1

Appendix C: Networking Configuration For Tests

Physical Networking

BAY 1 / 2 Switch	
Type	HP 6125G
Bandwidth with Aggregation switch	4 x 1Gb
Bandwidth with Host NICs	1 Gb



BAY 5 / 6 Switch	
Type	HP 6125G
Bandwidth with Aggregation switch	10 Gb
Bandwidth with Host NICs	1 Gb

Figure 18 - Physical Networking for HP Gen8

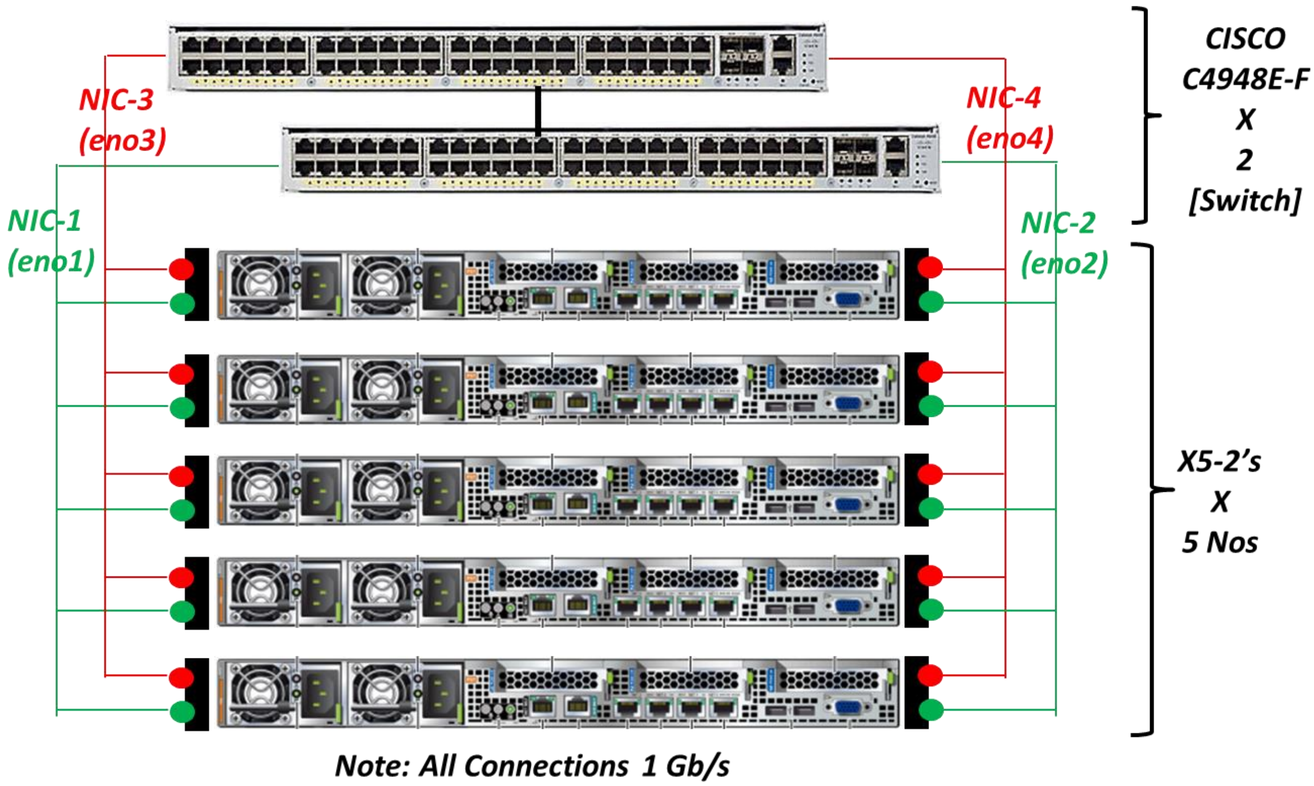
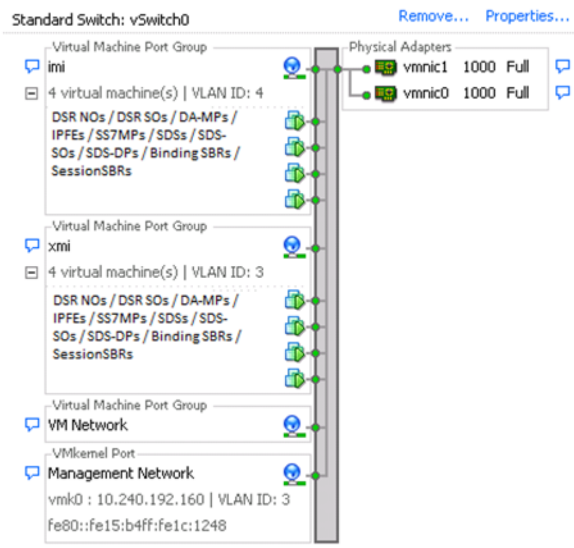


Figure 19: Networking for X5-2

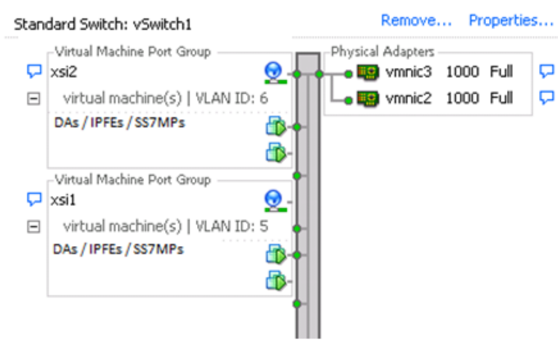
Note:

- eno1 / eno2 / eno3 / eno44 are the Ethernet interface names on X5-2's. NICx indicate the links from switch to the server(s).
- All links are 1 Gb/s

Host Networking



vSwitch0 Configuration	
Nos of Ports	120
MTU Size	1500
Promiscuous Mode	Reject
MAC Address Changes	Accept
Forged Transmits	Accept
Load Balancing	Failover Only
Network Failure Detection	Link Status Only
Notify Switches	Yes
Failback	Yes
Active / Standby Adapters	Vmnic0 / vmnic1



vSwitch1 Configuration	
Nos of Ports	120
MTU Size	1500
Promiscuous Mode	Reject
MAC Address Changes	Accept
Forged Transmits	Accept
Load Balancing	Failover Only
Network Failure Detection	Link Status Only
Notify Switches	Yes
Failback	Yes
Active / Standby Adapters	Vmnic2 / vmnic3

Figure 20 - Host Networking

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

White Paper Oracle Communications Diameter Signaling Router Cloud Benchmarking Guide
September 2016