Technical Resources

[Configuring QoS for WAN]

Reference: Simware Technical Library

Date: April-2016

Version : 1

1 INTRODUCTION

Simware provides a QoS file (OSPL.xml) optimized for its use for real time simulation running on a LAN. Sometimes you would need to change some of this parameters in order to optimize the performance in your network. To do this optimization you would need to change to Qos parameters in ospl file. This document describes and explains the changes made to the default LAN QoS configuration used in Simware in order to configure WAN communications over public networks or virtual private networks (VPN)

QoS configuration file is in XML format and has several tags for different configuration subjects. Some of them have global scope and others have service scope.

The tags with global scope are:

- Domain
- Tracing

The tags with service scope are:

- Networking: OSPL protocol communication config.
- DDSI: DDSI standard protocol config. (Deprecated)
- DDSI2: DDSI standard protocol config. (optimized)
- Durability

2 DOMAIN

The changes under Domain tag.

2.1 ROLE

This (optional) element specifies the role of the instantiated domain. For dynamic discovery, this role will be used to define the communication scope of instantiated domains on other nodes in the system. The purpose of specifying roles within the system is to 'overlay' the underlying physical network with a node's scope-of-interest that allows to "bound" topology discovery effort and related overhead in large scale (WAN) systems. The effect is that only nodes with the Role matching Discovery Scope expression discover each other.

3 NETWORKING

The Networking service is responsible for forwarding data to the network and for receiving data from the network. It can be configured to distinguish multiple communication channels with different QoS policies assigned to be able to schedule sending and receival of specific messages to provide optimal performance for a specific application domain.

3.1 GLOBALPARTITION ADDRESS

This element specifies the global or default networking partition. This global networking partition transports data that is either meant to be global, like discovery heartbeats, or that is not mapped onto any other networking partition. It is also used for **resending not acked reliable packets**.

Although OpenSplice documentation states that the GlobalPartition address is a list of one or more unicast, multicast or broadcast addresses. If more than one address is specified, then the different addresses should be separated by commas (,) semicolons (;) or spaces ( ). Samples for the global partition will be sent to all addresses that are specified in this list of addresses. **Only the first address is taken into account for the global partition.**

If this address is **multicast or broadcast** then it is used as a "default partition" for heartbeats and resent packets. If the first address is a unicast address then a "default partition" with a broadcast address is created.

The rest of the address in the list is only used for sending unmapped data.

3.2 NetworkPartition

Every NetworkPartition has a name, an address and a connected flag. Enabling the compression attribute doesn't seem to have the desired effect (checked with wireshark).

Using NetworkPartition, IgnoredPartitions and Partitionsmappings elements and Partition QoS we can separate among internal node data, LAN Data and WAN data, thus avoiding internal node data traffic going through LAN, and local data traffic collapsing the WAN connection.

3.3 IgnoredPartitions

This element can be used to create a "Local Partition" that is only available on the node on which it is specified, and therefore won't generate network-load. Any DCPS partition-topic combination specified in this element will not be distributed by the Networking service.

3.4 PartitionMappings

This element specifies a mapping between a network partition and a partition-topic combination.

The Networking Service will match any DCPS messages to the DCPSPartitionTopic expression and determine if it matches. The PartitionExpression and TopicExpression are allowed to contain a '*' wild card, meaning that anything matches. An exact match is considered better than a wild card match. For every DCPS message, the best matching partition is determined and the data is sent over the corresponding networking partition as specified by the matching NetworkPartition element.

3.5 Channels

The set of channels defines the behaviour of the 'network' concerning aspects as priority, reliability and latency budget. By configuring a set of channels, the Networking Service is able to function as a 'scheduler' for the network bandwidth. It achieves this by using the application-defined DDS QoS policies of the data to select the most appropriate channel to send the data. Channels are bidirectional.

3.5.1 Priority

Messages sent to the network have a transport_priority quality of service value. Selection of a networking channel is based on the priority requested by the message and the priority offered by the channel. The priority settings of the different channels divide the priority range into intervals. Within a channel, messages are sorted in order of their transport priority QoS.

Giving higher priority to the reliable channel ensures that reliable messages, which should be used for control, are sent with zero waiting time, even if the bandwidth is full with data messages.

3.5.2 PortNr

This element specifies the port number used by the Channel. Messages for the channel are sent to the given port number. Each channel needs its own unique port number. Please note that 'reliable' channels use a second port, which is the specified **PortNr + 1**.

3.5.3 Resolution

The resolution indicates the number of milliseconds that this channel sleeps between two consecutive resends or packing actions. The minimum value is 1.

3.5.4 FragmentSize

The networking module will fragment large message into smaller fragments with size FragmentSize. These fragments are sent as datagrams to the UDP stack. Operating system settings determine the maximum datagram size.

The human-readable option lets the user postfix the value with K (ilobyte), M(egabyte) or G(igabtye). For example, 10M results in 10485760 bytes.

3.5.5 MaxBurstSize

Amount in bytes to be sent at maximum every 'Resolution' milliseconds. The default value is set to 1GB per resolution tick. This can be regarded as effectively unlimited, as it far exceeds the capacity of current physical networks.

The human-readable option lets the user postfix the value with K (ilobyte), M(egabyte) or G(igabtye). For example, 10M results in 10485760 bytes.

With this parameter and resolution we are able to limit the channel throughput. For WAN this limit should be quite low.

3.5.6 ThrottleLimit & ThrottleThreshold

Throttling will enable you to further limit (below MaxBurstSize) the amount of data that is sent every Resolution interval. This happens if one of the receiving nodes in the network indicates that it has trouble processing all incoming data. Throttle limit is the lower boundary of the range over which the throttling can adapt the limit. If this value is set to the same value (or higher) as MaxBurstSize throttling is disabled.

ThrottleThreshold is the number of unprocessed network fragments that a node will store before it will inform the other nodes in the network that it has trouble processing the incoming data. Those other nodes can use this information to adjust their throttle values, effectively reducing the amount of incoming data in case of a temporary overflow, and increasing again when the node is able to catch up.

It is considered good practice to specify the ThrottleTreshold consistently throughout the system.

Throttle is a mechanism that protects nodes with less CPU power from being overwhelmed with data, thus is not used in WAN communications since the network bandwidth is so low that this is very unlikely to happen.

3.5.7 MaxRetries

The number of retransmissions the service has to execute before considering the addressed node as not responding. Once a node is classified as not responding no more data is sent to it, unless it is rediscovered.

3.5.8 RecoveryFactor

A lost message is resent after Resolution * RecoveryFactor milliseconds.

3.6 Reconnection

This element specifies the desired networking-behavior with respect to the validity of restoring lost connectivity with remote nodes. Here 'lost connectivity' means a prolonged inability to communicate with a known and still active remote node (typically because of network issues) that has resulted in such a node being declared 'dead' either by the topology-discovery or lost-reliability being detected by a reliable channel's reactivity-checking mechanism.

3.7 NetworkInterfaceAddress

Every Networking service is bound to only one network interface card. The card can be uniquely identified by its corresponding IP address or by its symbolic name (e.g. eth0). If the value "first available" is entered here, OpenSplice will try to look up an interface that has the required capabilities.

4 OSPL_WAN.XML

```
<OpenSplice>
  <Domain>
    <Name>99</Name>
    <Database>
      <Size>30485760</Size>
    </Database>
    <Lease>
      <ExpiryTime update_factor="0.9">600000.0</ExpiryTime>
    </Lease>
    <Service name="networking">
      <Command>networking</Command>
      <FailureAction>restart</FailureAction>
    </Service>
    <Service enabled="false" name="durability">
      <Command>durability</Command>
    </Service>
    <Role>Simware</Role>
  </Domain>
  <NetworkService name="networking">
    <Partitioning>
      <GlobalPartition Address="239.255.1.1" MulticastTimeToLive="32"/>
      <NetworkPartitions>
        <NetworkPartition   Address="172.30.2.102"   Compression="false"   Connected="true"
MulticastTimeToLive="32" Name="STAGE"/>
      </NetworkPartitions>
      <IgnoredPartitions>
        <IgnoredPartition DCPSPartitionTopic="BPStest.*"/>
        <IgnoredPartition DCPSPartitionTopic="BPSarmas.*"/>
      </IgnoredPartitions>
    <Channel enabled="true" name="Reliable" priority="10" reliable="true">
      <PortNr>53380</PortNr>
      <Sending>
        <MaxRetries>20</MaxRetries>
        <TimeToLive>128</TimeToLive>
        <RecoveryFactor>20</RecoveryFactor>
        <MaxBurstSize>1024</MaxBurstSize>
        <DontRoute>false</DontRoute>
      </Sending>
      <Receiving>
        <ReceiveBufferSize>10000000</ReceiveBufferSize>
        <SMPOptimization enabled="true"/>
        <DefragBufferSize>50000</DefragBufferSize>
      </Receiving>
      <Resolution>30</Resolution>
      <FragmentSize>1024</FragmentSize>
    </Channel>
  </Channels>
  <Discovery Scope="Simware" enabled="true">
    <PortNr>53390</PortNr>
    <Sending>
      <Interval>60000</Interval>
      <TimeToLive>128</TimeToLive>
      <DontRoute>false</DontRoute>
      <SafetyFactor>0.9</SafetyFactor>
    </Sending>
    <Receiving>
      <DeathDetectionCount>10</DeathDetectionCount>
    </Receiving>
    <ProbeList>172.30.2.102</ProbeList>
  </Discovery>
```