Fair Split can also be used as a more generic Congestion Management solution in the case where regulations require all traffic to be treated equal regardless of Application and Service Plan.

Most if not all networks suffer from congestion. Maybe it’s every day at peak time, maybe it’s when the newest iOS release comes out, or when Netflix releases a new original content TV-series. It is believed to be economically unfeasible to keep enough spare capacity to allow for these peaks never to interfere with the quality of traffic. Hence there will be congestion. When unmanaged, congestion is not fair. Some subscribers, typically the subscribers who use a lot of flows (like P2P users, downloads, etc) will consume a lot more than their fair share of the traffic during congestion. The whole business case around Fair Split comes from the fact that the vast majority of subscribers will receive an improved experience during congestion at the cost of very few subscribers who will “only” get their fair share.

When the network is not congested, Fair Split will be completely transparent. But it’s sometimes hard to identify congestion. It might happen spuriously off peak time due to a DOS attack, or just because a few subscribers happened to be consuming lots of bandwidth at the same time. When network utilization is measured as the average load over 5 minutes, these “micro congestion events” are often hard to spot. Fair Split will help with those situations as well.

OBJECTIVES
The objective of Congestion Management using Fair Split Shaping is to transparently handle congestion in the network, providing a fair share of the available network resources, while taking multiple factors into account, such as;

• Location of the subscriber in the Network Topology
• Service/Application, in this context called Traffic Classification
• Subscriber Service Plan or Tier

The result should be an overall higher Quality of Experience and reduction of Round Trip Time (RTT) during congestion.

TRAFFIC CLASSIFICATION
The congestion management solution described in the whitepaper focuses on a number of traffic classes that are differentiated in times of congestion, ensuring real-time traffic such as VoIP is given precedence over bulky services like File Sharing. The traffic classes in this example are using the Procera Networks Categorization;

• Real-Time
  – Categories/Entertainment/Gaming
  – Categories/Messaging and Collaboration/Video
  – Categories/Messaging and Collaboration/VoIP

• Background
  – Categories/File Sharing
  – Categories/File Transfer

• Interactive
  – All other services (Not Equals Real-Time or Background)

This would be considered an Application Aware congestion management solution.
SUBSCRIBER AWARENESS

Of course congestion management can be implemented on a per host basis in networks where subscriber awareness is not available, but there are some immediate advantages of subscriber awareness.

In the case of a fixed network, assigning multiple IP addresses to a single customer, it would be more fair to treat all those addresses as a single subscriber instead of individual hosts. This would also be relevant in the case of a dual stack scenario where a subscriber has multiple IPv4 and/or IPv6 addresses/prefixes. Subscriber awareness is also a fundamental requirement for service plan awareness.

SERVICE PLAN AWARENESS

The congestion management solution can be aware of service plans by utilizing PSM, where each subscribers plan is either statically defined or provisioned by a 3rd party provisioning/billing system. The relevance of a service plan is to ensure that a high value customer gets more bandwidth or a better service during congestion.

TOPOLOGY AWARENESS

Topology awareness is required if there are multiple possible congestion points in the network. This can be based on access network with multiple BRASes (DSL) or CMTSes (Cable). It could also be based on various backhaul capacities that need to be managed. On the other hand if the solution is deployed at an internet exchange, topology awareness may not be required. PSM can provide topology awareness using various data sources, such as Option 82 in DHCP, various RADIUS attributes or in the case of CMTSes, using SNMP or IPDR.

BORROWING

Borrowing is exactly what it sounds like: It allows shaping rules to borrow bandwidth from shaping objects when their “primary” shaping object is full (provided, of course, that there is bandwidth left in the object to borrow from). This is configured by adding multiple shaping objects to one shaping rule.

Packets from connections matching the shaping rule will then be enqueued in all shaping objects. However, the rule will increment the priority number for each step it goes in the shaping object list. That is, if the rule has priority 3 and three shaping objects, the first shaping object has connections assigned with priority 3, the second shaping object priority 4, and the third priority 5.

FAIR SPLIT SHAPING

Fair Split shaping is designed to maximize utilization of the network while ensuring that the congestion management is transparent if there is no congestion.

Think of a pool of bandwidth, let’s say 100Mbps, being shared between 20 subscribers. If the bandwidth sharing limits were static, each subscriber would get 5 Mbps, no more. But when some subscribers were inactive, this would waste a lot of bandwidth. Then, using PacketLogic™ terminology, the unused bandwidth could be “borrowed” by the active subscribers. If the inactive subscribers come back, they would have priority over their own part of the bandwidth and if all 20 subscribers are active they would each get exactly 5Mbps (if they need and want that much).
But the “borrowed” traffic would not be fairly shared between the active users. So if you set the static limit too low, there would still be a lot of unfairness in this example network. Some subscribers would be using less than 5 Mbps because they were inactive or simply did not need more bandwidth. Some subscribers would be using a lot of bandwidth, maybe 10Mbps or more, because their applications were aggressive. And some subscribers would only get 5Mbps though they actually wanted more.

If you set the 5Mbps limit higher, then you are overcommitting the guaranteed bandwidth, which might mean that a subscriber will not even get his minimum guarantee, and that would also be bad.

The solution to this problem is a more dynamic limit for the subscribers, where the guarantee level is set according to how many subscribers are active. If 10 subscribers are active, the limit will be 10Mbps per sub, and if all 20 are active, this limit will be 5Mbps. Thus a subscriber will always get exactly their fair share, and the bandwidth that they don’t use will be available to subscribers as “borrowing”.

### Fair Factor

In a similar example of a pool of 100Mbps, being shared between 20 subscribers, where 10 subscribers are in a gold service and the other 10 subscribers are in a silver service. The gold service should get three times more than the silver service.

Splitting the bandwidth according to Fair Factor, with all subscribers active would mean that the gold subscribers would get 7.5Mbps each and silver subscribers would get 2.5Mbps each (7.5 * 10 + 2.5 * 10 = 100). Fair Factor can be configured in a range from 1 to 20, where a user with Fair Factor 20 gets twenty times as much as a user with Fair Factor 1.

Note: Fair factor separation is not useful to define a very granular separation. In a setup with, for example, 20 different fair factors, PacketLogic™ will attempt to give a subscriber with fair factor 19 95% of the bandwidth of a subscriber with fair factor 20, but the effects will be so small that other variables in the network can easily cancel them out entirely. Fair factor is recommended as a way to achieve a coarse tiering of subscriber.
EFFECTS OF FAIR SPLIT SHAPING

Fair Split Shaping will not only ensure a fair division of bandwidth during times of congestion, Fair Split Shaping will also improve overall quality (QoE) and reduce round trip time (RTT), especially when differentiating traffic using Fair Factor.

In the following scenario we are using 10 test clients sending stateful TCP traffic in a congested scenario, with two separate service plans (silver and gold) where gold has a Fair Factor of two while silver has a Fair Factor of one.

Clients 00:00:00:00:00:01 to 00:00:00:00:00:05 are in the silver plan while clients 00:00:00:00:06 – 00:00:00:00:10 are in the gold plan.

### REAL-TIME TRAFFIC CONDITIONS WITH A 200 MBPS ShapingObject WITH TAIL DROP

<table>
<thead>
<tr>
<th>Item</th>
<th>Incoming</th>
<th>Outgoing</th>
<th>Total</th>
<th>Connections</th>
<th>In Int Quality</th>
<th>In RTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:00:00:00:01</td>
<td>24.6 Mbps</td>
<td>725.5 kbps</td>
<td>25.3 Mbps</td>
<td>101.0</td>
<td>93.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:02</td>
<td>23.3 Mbps</td>
<td>722.4 kbps</td>
<td>23.9 Mbps</td>
<td>99.0</td>
<td>91.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:03</td>
<td>14.5 Mbps</td>
<td>400.9 kbps</td>
<td>14.9 Mbps</td>
<td>71.0</td>
<td>93.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:04</td>
<td>19.6 Mbps</td>
<td>565.0 kbps</td>
<td>20.2 Mbps</td>
<td>81.0</td>
<td>91.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:05</td>
<td>19.3 Mbps</td>
<td>580.5 kbps</td>
<td>19.8 Mbps</td>
<td>89.0</td>
<td>90.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:06</td>
<td>19.2 Mbps</td>
<td>574.2 kbps</td>
<td>19.8 Mbps</td>
<td>82.0</td>
<td>91.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:07</td>
<td>17.6 Mbps</td>
<td>545.8 kbps</td>
<td>18.1 Mbps</td>
<td>86.0</td>
<td>89.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:08</td>
<td>19.7 Mbps</td>
<td>570.7 kbps</td>
<td>20.2 Mbps</td>
<td>93.0</td>
<td>92.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:09</td>
<td>17.2 Mbps</td>
<td>529.5 kbps</td>
<td>17.7 Mbps</td>
<td>83.0</td>
<td>90.0 %</td>
<td>146.0 ms</td>
</tr>
<tr>
<td>00:00:00:00:00:10</td>
<td>17.9 Mbps</td>
<td>543.0 kbps</td>
<td>18.4 Mbps</td>
<td>92.0</td>
<td>90.0 %</td>
<td>145.0 ms</td>
</tr>
</tbody>
</table>

When treating all traffic equal without any Advanced Queue Management (AQM) or Fair Split Shaping, clients will fight equally over the bandwidth but due to the nature of TCP Congestion backoff, bandwidth will not be fairly divided between the subscribers and Quality will be poor.

### REAL-TIME TRAFFIC CONDITIONS WITH A 200 MBPS ShapingObject WITH TAIL DROP

<table>
<thead>
<tr>
<th>Item</th>
<th>Incoming</th>
<th>Outgoing</th>
<th>Total</th>
<th>Connections</th>
<th>In Int Quality</th>
<th>In RTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>gold</td>
<td>90.3 Mbps</td>
<td>2,684.6 kbps</td>
<td>92.9 Mbps</td>
<td>418.0</td>
<td>91.0 %</td>
<td>144.0 ms</td>
</tr>
<tr>
<td>silver</td>
<td>99.5 Mbps</td>
<td>2,939.4 kbps</td>
<td>102.4 Mbps</td>
<td>445.0</td>
<td>91.0 %</td>
<td>144.0 ms</td>
</tr>
</tbody>
</table>

Broken down by plan, there is no differentiation between silver and gold subscribers, as a matter of fact, silver subscribers manage to get more throughput than gold subscribers during congestion when the traffic is unmanaged.
When enabling Fair Split Shaping, the bandwidth is distributed among the subscribers according to the Fair Factor, gold subscribers get twice as much as silver subscribers. Quality is better and RTT is lower for gold subscribers.

When looking at the same comparison using historical statistics the impact is even clearer, in this case traffic is unmanaged between 10:00 and noon, and Fair Split Shaping with Fair Factor is being used between noon and 14:00.

### Figure 3

**REAL-TIME TRAFFIC CONDITIONS WITH FAIR SPLIT SHAPING**

<table>
<thead>
<tr>
<th>Item</th>
<th>Time</th>
<th>Incoming</th>
<th>Outgoing</th>
<th>Total</th>
<th>Connections</th>
<th>In Int Quality</th>
<th>In RTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00:00:00:00:01</td>
<td>13.5 Mbps</td>
<td>616.3 kbps</td>
<td>14.1 Mbps</td>
<td>74.0</td>
<td>87.0 %</td>
<td>113.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:02</td>
<td>13.4 Mbps</td>
<td>611.7 kbps</td>
<td>14.0 Mbps</td>
<td>74.0</td>
<td>86.0 %</td>
<td>188.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:03</td>
<td>12.4 Mbps</td>
<td>525.5 kbps</td>
<td>12.9 Mbps</td>
<td>68.0</td>
<td>89.0 %</td>
<td>118.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:04</td>
<td>13.9 Mbps</td>
<td>629.7 kbps</td>
<td>14.5 Mbps</td>
<td>78.0</td>
<td>87.0 %</td>
<td>140.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:05</td>
<td>13.1 Mbps</td>
<td>594.5 kbps</td>
<td>13.7 Mbps</td>
<td>76.0</td>
<td>87.0 %</td>
<td>167.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:06</td>
<td>26.5 Mbps</td>
<td>984.5 kbps</td>
<td>27.5 Mbps</td>
<td>103.0</td>
<td>92.0 %</td>
<td>115.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:07</td>
<td>26.7 Mbps</td>
<td>1,046.9 kbps</td>
<td>27.7 Mbps</td>
<td>103.0</td>
<td>92.0 %</td>
<td>127.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:08</td>
<td>26.1 Mbps</td>
<td>988.8 kbps</td>
<td>27.1 Mbps</td>
<td>108.0</td>
<td>92.0 %</td>
<td>114.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:09</td>
<td>25.4 Mbps</td>
<td>987.4 kbps</td>
<td>26.4 Mbps</td>
<td>104.0</td>
<td>94.0 %</td>
<td>168.0 ms</td>
<td></td>
</tr>
<tr>
<td>00:00:00:00:00:10</td>
<td>25.4 Mbps</td>
<td>975.7 kbps</td>
<td>26.4 Mbps</td>
<td>103.0</td>
<td>95.0 %</td>
<td>143.0 ms</td>
<td></td>
</tr>
</tbody>
</table>

### Figure 4

**REAL-TIME TRAFFIC CONDITIONS WITH FAIR SPLIT SHAPING**

<table>
<thead>
<tr>
<th>Item</th>
<th>Incoming</th>
<th>Outgoing</th>
<th>Total</th>
<th>Connections</th>
<th>In Int Quality</th>
<th>In RTT</th>
</tr>
</thead>
<tbody>
<tr>
<td>gold</td>
<td>129.4 Mbps</td>
<td>5,002.3 kbps</td>
<td>134.4 Mbps</td>
<td>493.0</td>
<td>93.0 %</td>
<td>107.0 ms</td>
</tr>
<tr>
<td>silver</td>
<td>68.6 Mbps</td>
<td>3,116.3 kbps</td>
<td>71.7 Mbps</td>
<td>363.0</td>
<td>87.0 %</td>
<td>169.0 ms</td>
</tr>
</tbody>
</table>

Broken down by plan, we can see a clear differentiation between silver and gold subscribers. Gold subscribers get significantly better quality and lower RTT.
When the traffic is unmanaged, there is little differentiation between gold and silver subscribers, once Fair Split Shaping is enabled Fair Factor will ensure that the gold subscribers gets twice the bandwidth of silver subscribers.

When the traffic is unmanaged, there is no differentiation between gold and silver subscribers for both Quality and RTT, once Fair Split Shaping is enabled Fair Factor will that gold subscribers gets better Quality and lower RTT.
PARALLEL QUEUEING
Fair Split Shaping can also be combined with other policies by leveraging parallel queueing, allowing bandwidth limitations on a per plan basis or such. For more information on Parallel Queueing please refer to the PacketLogic™ Product Guide Section 5 (PacketLogic™ Traffic Shaping).

PSM NETOBJECT TREE
When combining topology and service plan in the context of Fair Split Shaping, we have to be able to match on a unique combination of topology and plan, therefore plans must be nested in the topology tree to provide a single NetObject as a condition for a shaping rule.

Figure 7

SIDEBAR

```
/By Topology
  /<subscriber.site>
    /<subscriber.service> !
    /All !
```

This type of NetObject tree allows matching Gold users under Site1 separately from Silver users under Site1, at the same time Gold users under Site1 can be matched separately from Gold users under Site2.

This can be greatly simplified with the introduction of ConditionObjects in Version 16.0 (Ruleset Enhancements).
CONFIGURING FAIR SPLIT SHAPING

Fair Split is implemented using a method similar to the “Split by Local Host”-feature in PacketLogic™. I.e one Shaping Object is allocated per active subscriber, and the limit of that Shaping Object is dynamically set according to the number of active subscribers. So when checking for availability for a packet, the PacketLogic™ Shaping Engine will first check if there is “room” in the per subscriber shaping object. If there is, then the packet will be queued there and eventually transmitted. If there is not room, then the total rule will be consulted to see if there is room there. Any packets borrowing from this “total rule” will have lower priority, and will compete with any other subscribers wanting to use more than their “fair share”.

To enable Fair Split Shaping, configure a Shaping Object and select a Split by, typically Subscriber, or Localhost / Local Network Prefix if there is no subscriber awareness from PSM. Further Host fairness must be set to “Fair Split”.

When configuring Fair Split, the Shaping Object bandwidth defines the total bandwidth to be shared by all subscribers in the shaping object, this behavior differs from how a normal Split by Shaping Object would function (each object copy would get the configured bandwidth).

In this example we configure a Shaping Object named “Site1 Background”, Split by Subscriber with Fair Split where the total available bandwidth is 1Mbps to be shared by all subscribers in this Shaping Object.

PRE-REQUISITES

Fair Split shaping is available from release 14.1 onwards, additional features for service plan awareness was added in release 15.0, hence if service plan awareness is required, use 15.0, it not, 14.1 will suffice. This whitepaper assumes version 15.0 and PSM for subscriber, service plan and topology awareness.
**POLICY CONFIGURATION**

By utilizing a combination of Fair Split Shaping and Borrowing, we can accomplish a configuration where the bandwidth within each Application or Traffic Class is fairly divided between the subscribers, and also the amount of bandwidth per Traffic Class can be specified to ensure available bandwidth for Real-Time and Interactive Traffic Classes during congestion. This is done by allowing each Traffic Class to borrow from the total available bandwidth in a given part of the network.

In this example we reserve 1 Mbps for Real-Time traffic, 1 Mbps for Background traffic and 8 Mbps for Interactive traffic, borrowing from Site1 at 10 Mbps. If there is bandwidth available in Site1 any Traffic Class can borrow additional bandwidth.

![Figure 9](image1)

**ILLUSTRATION OF TRAFFIC CLASSES AND BORROWING**

![Figure 10](image2)

**ILLUSTRATION OF TRAFFIC CLASSES WITH FAIR FACTOR**

Within each Traffic Class the bandwidth will be Fair Split between the subscribers based on the Fair Factor. In this case, Gold gets twice as much as Silver.
The configuration of the traffic classes is done using Fair Split Shaping.

**Figure 11**

**Configuration of the Real-Time Traffic Class**

![Real-Time Traffic Class Configuration](image)

**Figure 12**

**Configuration of the Interactive Traffic Class**

![Interactive Traffic Class Configuration](image)
The configuration of the Site1 Pipe is done using Split by None, as this bandwidth is to be shared by all subscribers in Site 1.
With the Shaping Objects configured we can move on to the Shaping Rules. The policy requires one Shaping Rules per Site per Traffic Class per Service Plan, where the ServiceObject conditions will differ for the classes, NetObject conditions will differ for Topology and Service plan and the Fair Factor for the Service Plans. Each Shaping Rule will use the Traffic Class as the primary Shaping Object and the Site1 Pipe as a secondary object with Borrowing.

**Figure 15**

**CONFIGURATION OF SHAPING RULE FOR SITE1, BACKGROUND TRAFFIC CLASS, SILVER SERVICE PLAN**

Site1 Silver Background traffic will use a default priority of 5 and Fair Factor 1. Background traffic is configured as Categories/File Sharing and Categories/File Transfer.

**Figure 16**

**CONFIGURATION OF BORROWING CHAIN FOR SITE1 SILVER BACKGROUND**

Site1 Silver Background will use the Site1 Background Shaping Object as the primary object and Site1 Pipe as secondary object for borrowing.
For Site1 Gold Background, the borrowing chain will be identical, but the Shaping Rules will be configured with a Fair Factor of 2, assuring Gold subscribers get twice as much as Silver subscribers during congestion.

**Figure 17**

Site1 Gold Background traffic will use a default priority of 5 and Fair Factor 2. Background traffic is configured as Categories/File Sharing and Categories/File Transfer.

Configuration for the Interactive and Real-Time traffic classes will be identical, except the matching ServiceObjects.

**Figure 18**

**CONFIGURATION OF SHAPING RULE FOR SITE1, INTERACTIVE TRAFFIC CLASS, SILVER SERVICE PLAN**

Site1 Silver Interactive traffic will use a default priority of 5 and Fair Factor 1. Interactive traffic is configured as Not Equals Categories/File Sharing, Categories/File Transfer, Categories/Entertainment/Gaming, Categories/Messaging and Collaboration Video, Categories/Messaging and Collaboration/VoIP.
The same configuration needs to be repeated for remaining traffic classes in the Gold Service plan. Similarly configuration has to be repeated for additional network locations / sites.

**SYSTEM CONFIGURATION**

Fair Split Shaping require certain System Configuration settings which are non-default:

- **Shaping/SHAPING_OR_BORROWING** should be set to True
- Also do ensure that the following settings are correct (although they are default in 14.1/15.0):
  - **Shaping/PRIORITIZE** should be set to True
  - **Shaping/PRIORITIZE_RETRANS** should be set to False

Site1 Silver Real-Time traffic will use a default priority of 5 and Fair Factor 1. Real-Time traffic is configured as Categories/Entertainment/Gaming, Categories/Messaging and Collaboration Video and Categories/Messaging and Collaboration/VoIP.
NET NEUTRALITY CONSIDERATIONS

With recent developments both on the American and European markets in regards to Net Neutrality, it may be impossible to implement an Application Aware or Plan Aware Congestion Management solution.

In these cases, the benefits of Fair Split Shaping is still valid, although all subscribers may have to be considered equal.

This easily be achieved by simply removing the Traffic Classes and just using a single class for all traffic, and also removing the Fair Factor differentiating the service plans, resulting in a Service and Application Agnostic Congestion Management solution conforming to Net Neutrality regulations.