

# ODCC无损网络测试基准

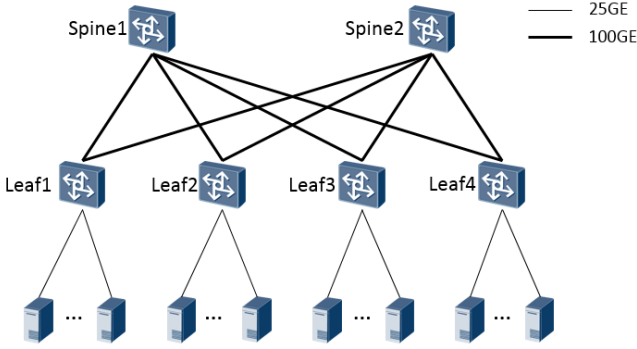
## 目录

1	性能测试.....	3
1.1	RoCEv2打流测试.....	3
1.1.1	RoCE与TCP性能对比.....	3
1.1.2	RoCE流量4打1.....	4
1.1.3	RoCE流量8打1.....	5
1.2	TCP+RoCEv2 混跑测试.....	7
1.2.1	混跑4打1.....	7
1.2.2	混跑8打1.....	11
2	运维功能.....	13
2.1	PFC死锁预防.....	13
2.2	RoCE丢包可视化.....	18
2.3	RoCE性能可视化.....	20
3	分布式存储测试.....	22
3.1	分布式存储性能(DSS).....	22
4	大规模RoCE网络测试.....	23
4.1	跨POD混跑4打1.....	23
4.2	跨POD混跑8打1.....	26
4.3	跨POD的PFC死锁预防.....	28

## 1 性能测试

### 1.1 RoCEv2 打流测试

#### 1.1.1 RoCE 与 TCP 性能对比

测试项目	RoCE与TCP性能对比
测试目的	对带宽、时延、CPU利用率三项指标进行RoCE与TCP对比测试。
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>4) 在同一个Leaf下选择2个Server使用ib_write_bw工具进行1打1带宽测试，消息长度遍历64B、1024B、4096B、65536B，QP数为1。记录带宽利用率、CPU利用率。</li> <li>5) 在同一个Leaf下选择2个Server使用ib_write_lat工具进行1打1时延测试，消息长度遍历64B、1024B、4096B、65536B。记录时延。</li> <li>6) 在同一个Leaf下选择2个Server使用iperf3工具进行TCP 1打1带宽测试，消息长度遍历64B、1024B、4096B、65536B，进程/线程数为1。记录带宽利用率、CPU利用率。</li> <li>7) 在同一个Leaf下选择2个Server使用netperf工具进行TCP 1打1时延测试，消息长度遍历64B、1024B、4096B、65536B。记录时延。有预期结果1。</li> <li>8) 选择跨Leaf的2个Server重复步骤1~4，有预期结果1。</li> </ol>
预期结果	1) 消息长度相同的条件下RoCEv2的吞吐比TCP的更高，时延、CPU利用率比TCP的更低。
测试结果	测试结果

ODCC 无损网络测试基准

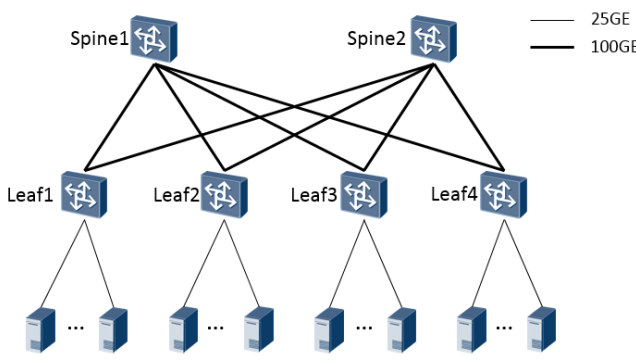
同TOR 1打1												
message size	RDMA(单QP)				TCP(单线程)				提升比例			
	交换机出接口带宽	时延	发端CPU利用率	收端CPU利用率	交换机出接口带宽	时延	发端CPU利用率	收端CPU利用率	带宽	时延	发端CPU利用率	收端CPU利用率
64	38.27%	2.32	2.08%	0%	2.37%	7.79	36.60%	54.70%	1514.77%	70.22%	94.32%	100.00%
1024	100%	3.33	2.08%	0%	29.48%	26.91	37.10%	38.70%	239.21%	87.63%	94.39%	100.00%
4096	100%	5.04	2.08%	0%	63.25%	15.96	28%	56.40%	58.10%	68.42%	92.57%	100.00%
65536	100%	2.31	2.08%	0%	100%	64.88	18.40%	34.30%	0.00%	96.44%	88.70%	100.00%

跨TOR 1打1												
message size	RDMA(单QP)				TCP(单线程)				提升比例			
	交换机出接口带宽	时延	发端CPU利用率	收端CPU利用率	交换机出接口带宽	时延	发端CPU利用率	收端CPU利用率	带宽	时延	发端CPU利用率	收端CPU利用率
64	28.74%	3.79	2.08%	0%	2.22%	11.3	34.90%	31.60%	1194.59%	66.46%	94.04%	100.00%
1024	100%	4.86	2.08%	0%	25.63%	12.34	33.20%	40.20%	290.17%	60.62%	93.73%	100.00%
4096	100%	6.63	2.08%	0%	61.61%	23.46	28%	50.10%	62.31%	71.74%	92.52%	100.00%
65536	100%	27.84	2.08%	0%	100%	70.08	19.70%	16.90%	0.00%	60.27%	89.44%	100.00%

备注

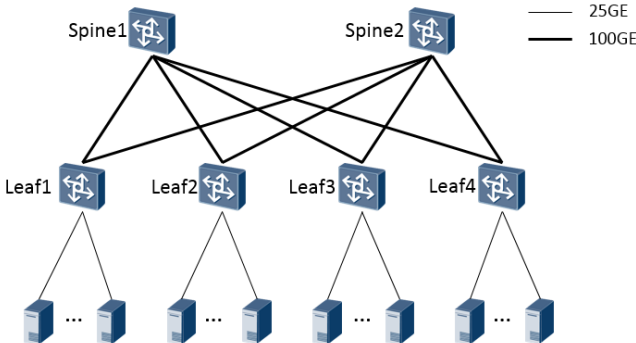
1.1.2 RoCE 流量 4打1

测试项目	RoCE流量4打1
测试目的	测试在4打1流量模型下RoCE的无丢包、高带宽、低时延特性。
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 对同Leaf下的5台服务器使用ib_write_bw工具进行4打1带宽测试，消息长度遍历64B、1024B、4096B、65536B，源端QP数遍历1、32、64、128。记录带宽利用率。</li> <li>2) 使用步骤1)的流量作为背景流的情况下，对同Leaf下的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> <li>3) 对跨Leaf的5台服务器使用ib_write_bw工具进行4打1带宽测试，消息长度遍历64B、1024B、4096B、65536B，源端QP数遍历1、32、64、128。记录带宽利用率。</li> <li>4) 使用步骤3)的流量作为背景流的情况下，对跨Leaf的2台服务器使用</li> </ol>

	ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。																																																																																																																																																																																																																														
预期结果	1) 交换机无丢包，交换机出接口带宽达到95%以上。																																																																																																																																																																																																																														
测试结果	<p>同 Leaf</p> <table border="1"> <thead> <tr> <th rowspan="2">message size</th> <th rowspan="2">收端总QP数</th> <th colspan="2">AI Fabric</th> <th colspan="2">传统PFC/ECN</th> <th rowspan="2">吞吐提升</th> </tr> <tr> <th>吞吐</th> <th>时延</th> <th>吞吐</th> <th>时延</th> </tr> </thead> <tbody> <tr> <td rowspan="4">64</td> <td>4</td> <td>100.00%</td> <td>4.24</td> <td>100.00%</td> <td>4</td> <td>0.00%</td> </tr> <tr> <td>128</td> <td>99.43%</td> <td>71.43</td> <td>80.84%</td> <td>11.78</td> <td>23.00%</td> </tr> <tr> <td>256</td> <td>100.00%</td> <td>237.83</td> <td>81%</td> <td>186.98</td> <td>23.67%</td> </tr> <tr> <td>512</td> <td>99.65%</td> <td>574.51</td> <td>81%</td> <td>547.2</td> <td>23.22%</td> </tr> <tr> <td rowspan="4">1024</td> <td>4</td> <td>99.87%</td> <td>4.97</td> <td>98.92%</td> <td>5.49</td> <td>0.96%</td> </tr> <tr> <td>128</td> <td>99.59%</td> <td>22.83</td> <td>98.45%</td> <td>23.01</td> <td>1.16%</td> </tr> <tr> <td>256</td> <td>96.56%</td> <td>50.75</td> <td>94.82%</td> <td>50.83</td> <td>1.84%</td> </tr> <tr> <td>512</td> <td>99.59%</td> <td>96.25</td> <td>95.09%</td> <td>96.26</td> <td>4.73%</td> </tr> <tr> <td rowspan="4">4096</td> <td>4</td> <td>100%</td> <td>7.24</td> <td>99.78%</td> <td>8.06</td> <td>-0.19%</td> </tr> <tr> <td>128</td> <td>98.66%</td> <td>92.29</td> <td>90.52%</td> <td>102.5</td> <td>8.99%</td> </tr> <tr> <td>256</td> <td>98.45%</td> <td>178.95</td> <td>74.36%</td> <td>238.73</td> <td>32.40%</td> </tr> <tr> <td>512</td> <td>98.43%</td> <td>364.86</td> <td>94.80%</td> <td>369.05</td> <td>3.83%</td> </tr> <tr> <td rowspan="4">65536</td> <td>4</td> <td>98%</td> <td>78.72</td> <td>99.22%</td> <td>63.82</td> <td>-1.04%</td> </tr> <tr> <td>128</td> <td>96.79%</td> <td>1480.18</td> <td>89.39%</td> <td>1596.62</td> <td>8.28%</td> </tr> <tr> <td>256</td> <td>97.85%</td> <td>2857.01</td> <td>75.38%</td> <td>3733.78</td> <td>29.81%</td> </tr> <tr> <td>512</td> <td>95.45%</td> <td>5639.66</td> <td>94.49%</td> <td>5888.62</td> <td>1.02%</td> </tr> </tbody> </table> <p>跨 Leaf</p> <table border="1"> <thead> <tr> <th rowspan="2">message size</th> <th rowspan="2">收端总QP数</th> <th colspan="2">AI Fabric</th> <th colspan="2">传统PFC/ECN</th> <th rowspan="2">吞吐提升</th> </tr> <tr> <th>吞吐</th> <th>时延</th> <th>吞吐</th> <th>时延</th> </tr> </thead> <tbody> <tr> <td rowspan="4">64</td> <td>4</td> <td>99.39%</td> <td>5.2</td> <td>99.78%</td> <td>5.23</td> <td>-0.39%</td> </tr> <tr> <td>128</td> <td>99.86%</td> <td>17.52</td> <td>73.47%</td> <td>15.7</td> <td>35.92%</td> </tr> <tr> <td>256</td> <td>100.00%</td> <td>223.85</td> <td>73%</td> <td>236.99</td> <td>36.89%</td> </tr> <tr> <td>512</td> <td>99.95%</td> <td>548.3</td> <td>87%</td> <td>745.98</td> <td>15.43%</td> </tr> <tr> <td rowspan="4">1024</td> <td>4</td> <td>99.60%</td> <td>6.55</td> <td>97.98%</td> <td>6.89</td> <td>1.65%</td> </tr> <tr> <td>128</td> <td>97.85%</td> <td>102.83</td> <td>93.27%</td> <td>24.9</td> <td>4.91%</td> </tr> <tr> <td>256</td> <td>92.61%</td> <td>48.71</td> <td>91.99%</td> <td>53.6</td> <td>0.67%</td> </tr> <tr> <td>512</td> <td>99.15%</td> <td>95.1</td> <td>85.49%</td> <td>120.34</td> <td>15.98%</td> </tr> <tr> <td rowspan="4">4096</td> <td>4</td> <td>100%</td> <td>8.6</td> <td>99.74%</td> <td>9.48</td> <td>0.03%</td> </tr> <tr> <td>128</td> <td>96.80%</td> <td>171.51</td> <td>90.92%</td> <td>93.65</td> <td>6.47%</td> </tr> <tr> <td>256</td> <td>97.36%</td> <td>199.56</td> <td>81.78%</td> <td>215.25</td> <td>19.05%</td> </tr> <tr> <td>512</td> <td>97.52%</td> <td>363.44</td> <td>78.57%</td> <td>552.94</td> <td>24.12%</td> </tr> <tr> <td rowspan="4">65536</td> <td>4</td> <td>97%</td> <td>78.67</td> <td>98.97%</td> <td>70.6</td> <td>-2.11%</td> </tr> <tr> <td>128</td> <td>95.40%</td> <td>1730.13</td> <td>87.97%</td> <td>1572.12</td> <td>8.45%</td> </tr> <tr> <td>256</td> <td>98.02%</td> <td>3368.3</td> <td>78.66%</td> <td>3644.22</td> <td>24.61%</td> </tr> <tr> <td>512</td> <td>97.84%</td> <td>5827.7</td> <td>79.31%</td> <td>8814.1</td> <td>23.36%</td> </tr> </tbody> </table>	message size	收端总QP数	AI Fabric		传统PFC/ECN		吞吐提升	吞吐	时延	吞吐	时延	64	4	100.00%	4.24	100.00%	4	0.00%	128	99.43%	71.43	80.84%	11.78	23.00%	256	100.00%	237.83	81%	186.98	23.67%	512	99.65%	574.51	81%	547.2	23.22%	1024	4	99.87%	4.97	98.92%	5.49	0.96%	128	99.59%	22.83	98.45%	23.01	1.16%	256	96.56%	50.75	94.82%	50.83	1.84%	512	99.59%	96.25	95.09%	96.26	4.73%	4096	4	100%	7.24	99.78%	8.06	-0.19%	128	98.66%	92.29	90.52%	102.5	8.99%	256	98.45%	178.95	74.36%	238.73	32.40%	512	98.43%	364.86	94.80%	369.05	3.83%	65536	4	98%	78.72	99.22%	63.82	-1.04%	128	96.79%	1480.18	89.39%	1596.62	8.28%	256	97.85%	2857.01	75.38%	3733.78	29.81%	512	95.45%	5639.66	94.49%	5888.62	1.02%	message size	收端总QP数	AI Fabric		传统PFC/ECN		吞吐提升	吞吐	时延	吞吐	时延	64	4	99.39%	5.2	99.78%	5.23	-0.39%	128	99.86%	17.52	73.47%	15.7	35.92%	256	100.00%	223.85	73%	236.99	36.89%	512	99.95%	548.3	87%	745.98	15.43%	1024	4	99.60%	6.55	97.98%	6.89	1.65%	128	97.85%	102.83	93.27%	24.9	4.91%	256	92.61%	48.71	91.99%	53.6	0.67%	512	99.15%	95.1	85.49%	120.34	15.98%	4096	4	100%	8.6	99.74%	9.48	0.03%	128	96.80%	171.51	90.92%	93.65	6.47%	256	97.36%	199.56	81.78%	215.25	19.05%	512	97.52%	363.44	78.57%	552.94	24.12%	65536	4	97%	78.67	98.97%	70.6	-2.11%	128	95.40%	1730.13	87.97%	1572.12	8.45%	256	98.02%	3368.3	78.66%	3644.22	24.61%	512	97.84%	5827.7	79.31%	8814.1	23.36%
	message size			收端总QP数	AI Fabric		传统PFC/ECN		吞吐提升																																																																																																																																																																																																																						
吞吐		时延	吞吐		时延																																																																																																																																																																																																																										
64	4	100.00%	4.24	100.00%	4	0.00%																																																																																																																																																																																																																									
	128	99.43%	71.43	80.84%	11.78	23.00%																																																																																																																																																																																																																									
	256	100.00%	237.83	81%	186.98	23.67%																																																																																																																																																																																																																									
	512	99.65%	574.51	81%	547.2	23.22%																																																																																																																																																																																																																									
1024	4	99.87%	4.97	98.92%	5.49	0.96%																																																																																																																																																																																																																									
	128	99.59%	22.83	98.45%	23.01	1.16%																																																																																																																																																																																																																									
	256	96.56%	50.75	94.82%	50.83	1.84%																																																																																																																																																																																																																									
	512	99.59%	96.25	95.09%	96.26	4.73%																																																																																																																																																																																																																									
4096	4	100%	7.24	99.78%	8.06	-0.19%																																																																																																																																																																																																																									
	128	98.66%	92.29	90.52%	102.5	8.99%																																																																																																																																																																																																																									
	256	98.45%	178.95	74.36%	238.73	32.40%																																																																																																																																																																																																																									
	512	98.43%	364.86	94.80%	369.05	3.83%																																																																																																																																																																																																																									
65536	4	98%	78.72	99.22%	63.82	-1.04%																																																																																																																																																																																																																									
	128	96.79%	1480.18	89.39%	1596.62	8.28%																																																																																																																																																																																																																									
	256	97.85%	2857.01	75.38%	3733.78	29.81%																																																																																																																																																																																																																									
	512	95.45%	5639.66	94.49%	5888.62	1.02%																																																																																																																																																																																																																									
message size	收端总QP数	AI Fabric		传统PFC/ECN		吞吐提升																																																																																																																																																																																																																									
		吞吐	时延	吞吐	时延																																																																																																																																																																																																																										
64	4	99.39%	5.2	99.78%	5.23	-0.39%																																																																																																																																																																																																																									
	128	99.86%	17.52	73.47%	15.7	35.92%																																																																																																																																																																																																																									
	256	100.00%	223.85	73%	236.99	36.89%																																																																																																																																																																																																																									
	512	99.95%	548.3	87%	745.98	15.43%																																																																																																																																																																																																																									
1024	4	99.60%	6.55	97.98%	6.89	1.65%																																																																																																																																																																																																																									
	128	97.85%	102.83	93.27%	24.9	4.91%																																																																																																																																																																																																																									
	256	92.61%	48.71	91.99%	53.6	0.67%																																																																																																																																																																																																																									
	512	99.15%	95.1	85.49%	120.34	15.98%																																																																																																																																																																																																																									
4096	4	100%	8.6	99.74%	9.48	0.03%																																																																																																																																																																																																																									
	128	96.80%	171.51	90.92%	93.65	6.47%																																																																																																																																																																																																																									
	256	97.36%	199.56	81.78%	215.25	19.05%																																																																																																																																																																																																																									
	512	97.52%	363.44	78.57%	552.94	24.12%																																																																																																																																																																																																																									
65536	4	97%	78.67	98.97%	70.6	-2.11%																																																																																																																																																																																																																									
	128	95.40%	1730.13	87.97%	1572.12	8.45%																																																																																																																																																																																																																									
	256	98.02%	3368.3	78.66%	3644.22	24.61%																																																																																																																																																																																																																									
	512	97.84%	5827.7	79.31%	8814.1	23.36%																																																																																																																																																																																																																									
备注																																																																																																																																																																																																																															

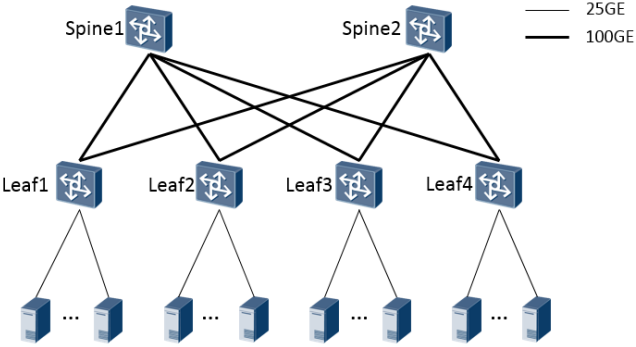
## 1.1.3 RoCE 流量 8 打 1

测试项目	RoCE流量8打1
测试目的	测试在8打1流量模型下RoCE的无丢包、高带宽、低时延特性。
测试环境	测试组网：

	 <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。</li> </ol>																																																																																																																
<p>测试步骤</p>	<ol style="list-style-type: none"> <li>1) 选择9台服务器使用ib_write_bw工具进行8打1带宽测试,其中4个源端和目的端处于一个Leaf下,另外4个源端和目的端处于不同Leaf下。消息长度遍历64B、1024B、4096B、65536B,发流端的QP数遍历1、32、64、128。记录带宽利用率。</li> <li>2) 使用步骤2)的流量作为背景流的情况下,对跨Leaf的2台服务器使用ib_write_lat工具进行时延测试,消息长度与背景流的消息长度保持一致,目的端为同一服务器。记录时延。</li> </ol>																																																																																																																
<p>预期结果</p>	<ol style="list-style-type: none"> <li>1) 交换机无丢包,交换机出接口带宽达到95%以上。</li> </ol>																																																																																																																
<p>测试结果</p>	<table border="1"> <thead> <tr> <th rowspan="2">message size</th> <th rowspan="2">收端总QP数</th> <th colspan="2">AI Fabric</th> <th colspan="2">传统PFC/ECN</th> <th rowspan="2">吞吐提升</th> </tr> <tr> <th>吞吐</th> <th>时延</th> <th>吞吐</th> <th>时延</th> </tr> </thead> <tbody> <tr> <td rowspan="4">64</td> <td>8</td> <td>99.94%</td> <td>11.57</td> <td>99.03%</td> <td>4.92</td> <td>0.92%</td> </tr> <tr> <td>256</td> <td>100.00%</td> <td>37.82</td> <td>81.52%</td> <td>20.74</td> <td>22.67%</td> </tr> <tr> <td>512</td> <td>100.00%</td> <td>425.37</td> <td>90%</td> <td>446.15</td> <td>10.98%</td> </tr> <tr> <td>1024</td> <td>99.97%</td> <td>1578.28</td> <td>100%</td> <td>1578.58</td> <td>0.11%</td> </tr> <tr> <td rowspan="3">1024</td> <td>8</td> <td>97.22%</td> <td>6.9</td> <td>99.47%</td> <td>7.31</td> <td>-2.26%</td> </tr> <tr> <td>256</td> <td>98.12%</td> <td>76.99</td> <td>99.26%</td> <td>50.12</td> <td>-1.15%</td> </tr> <tr> <td>512</td> <td>99.78%</td> <td>88.32</td> <td>94.23%</td> <td>104.63</td> <td>5.89%</td> </tr> <tr> <td rowspan="3">4096</td> <td>1024</td> <td>100.00%</td> <td>6043.36</td> <td>99.60%</td> <td>7490.3</td> <td>0.40%</td> </tr> <tr> <td>8</td> <td>99%</td> <td>9.02</td> <td>100.00%</td> <td>11.48</td> <td>-0.57%</td> </tr> <tr> <td>256</td> <td>99.43%</td> <td>182.75</td> <td>89.11%</td> <td>206.66</td> <td>11.58%</td> </tr> <tr> <td rowspan="4">65536</td> <td>512</td> <td>97.61%</td> <td>318.78</td> <td>83.12%</td> <td>733.14</td> <td>17.43%</td> </tr> <tr> <td>1024</td> <td>99.79%</td> <td>23469.64</td> <td>99.13%</td> <td>7055.05</td> <td>0.67%</td> </tr> <tr> <td>8</td> <td>99%</td> <td>146.64</td> <td>99.63%</td> <td>139.75</td> <td>-0.62%</td> </tr> <tr> <td>256</td> <td>96.82%</td> <td>2740.48</td> <td>87.73%</td> <td>3390.41</td> <td>10.36%</td> </tr> <tr> <td rowspan="2"></td> <td>512</td> <td>97.38%</td> <td>5628.56</td> <td>82.88%</td> <td>11651.84</td> <td>17.50%</td> </tr> <tr> <td>1024</td> <td>96.94%</td> <td>8803.7</td> <td>97.48%</td> <td>62019.8</td> <td>-0.55%</td> </tr> </tbody> </table>	message size	收端总QP数	AI Fabric		传统PFC/ECN		吞吐提升	吞吐	时延	吞吐	时延	64	8	99.94%	11.57	99.03%	4.92	0.92%	256	100.00%	37.82	81.52%	20.74	22.67%	512	100.00%	425.37	90%	446.15	10.98%	1024	99.97%	1578.28	100%	1578.58	0.11%	1024	8	97.22%	6.9	99.47%	7.31	-2.26%	256	98.12%	76.99	99.26%	50.12	-1.15%	512	99.78%	88.32	94.23%	104.63	5.89%	4096	1024	100.00%	6043.36	99.60%	7490.3	0.40%	8	99%	9.02	100.00%	11.48	-0.57%	256	99.43%	182.75	89.11%	206.66	11.58%	65536	512	97.61%	318.78	83.12%	733.14	17.43%	1024	99.79%	23469.64	99.13%	7055.05	0.67%	8	99%	146.64	99.63%	139.75	-0.62%	256	96.82%	2740.48	87.73%	3390.41	10.36%		512	97.38%	5628.56	82.88%	11651.84	17.50%	1024	96.94%	8803.7	97.48%	62019.8	-0.55%
message size	收端总QP数			AI Fabric		传统PFC/ECN			吞吐提升																																																																																																								
		吞吐	时延	吞吐	时延																																																																																																												
64	8	99.94%	11.57	99.03%	4.92	0.92%																																																																																																											
	256	100.00%	37.82	81.52%	20.74	22.67%																																																																																																											
	512	100.00%	425.37	90%	446.15	10.98%																																																																																																											
	1024	99.97%	1578.28	100%	1578.58	0.11%																																																																																																											
1024	8	97.22%	6.9	99.47%	7.31	-2.26%																																																																																																											
	256	98.12%	76.99	99.26%	50.12	-1.15%																																																																																																											
	512	99.78%	88.32	94.23%	104.63	5.89%																																																																																																											
4096	1024	100.00%	6043.36	99.60%	7490.3	0.40%																																																																																																											
	8	99%	9.02	100.00%	11.48	-0.57%																																																																																																											
	256	99.43%	182.75	89.11%	206.66	11.58%																																																																																																											
65536	512	97.61%	318.78	83.12%	733.14	17.43%																																																																																																											
	1024	99.79%	23469.64	99.13%	7055.05	0.67%																																																																																																											
	8	99%	146.64	99.63%	139.75	-0.62%																																																																																																											
	256	96.82%	2740.48	87.73%	3390.41	10.36%																																																																																																											
	512	97.38%	5628.56	82.88%	11651.84	17.50%																																																																																																											
	1024	96.94%	8803.7	97.48%	62019.8	-0.55%																																																																																																											
<p>备注</p>																																																																																																																	

## 1.2 TCP+RoCEv2 混跑测试

## 1.2.1 混跑4打1

测试项目	混跑4打1
测试目的	测试TCP、RoCE混跑场景下调度的精准性、出接口带宽，同时测试RoCE时延不受TCP影响保持低时延。
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。TCP流量映射到队列0，RoCE流量映射到队列3，配置0、3队列调度方式为DRR。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 对同Leaf下的5台服务器使用ib_write_bw工具进行4打1带宽测试，同时使用netperf工具进行TCP4打1，消息长度遍历64B、512B、1024B、4096B，发端QP数遍历1、2、4、8、16，DRR调度比重覆盖1:9，3:7，5:5，7:3，9:1，记录带宽利用率和带宽比例。</li> <li>2) 使用步骤1)的流量作为背景流的情况下，对同Leaf下的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> <li>3) 对跨Leaf的5台服务器使用ib_write_bw工具进行4打1带宽测试，同时使用netperf工具进行TCP4打1，消息长度遍历64B、512B、1024B、4096B，发端QP数遍历1、2、4、8、16，DRR调度比重覆盖1:9，3:7，5:5，7:3，9:1，记录带宽利用率和带宽比例。</li> <li>4) 使用步骤3)的流量作为背景流的情况下，对跨Leaf的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> </ol>
预期结果	<ol style="list-style-type: none"> <li>1) 交换机上RoCEv2流量无丢包，除64B外交换机出接口带宽达到95%以上。吞吐比例控制在5%误差范围内。</li> <li>2) RoCE时延在TCP:RoCE=9:1比例下控制在1ms内。</li> </ol>

ODCC 无损网络测试基准

TCP:ROCE		message size	发端QP数	AI Fabric				传统PFC/ECN			
				吞吐	时延	比例	比例误差	吞吐	时延	比例	比例误差
同 Leaf	1:9	64	1	100%	4.55	86.29%	3.71%	100%	4.14	81.78%	8.22%
			2	100.00%	8.29	87.60%	2.40%	99.98%	4.04	82.35%	7.65%
			4	100%	7.65	87.53%	2.47%	100%	4.35	81.71%	8.29%
			8	99.62%	8.29	87.65%	2.35%	100.00%	4.24	80.28%	9.72%
			16	100%	8.03	88.44%	1.56%	100%	4.24	77.71%	12.29%
		512	1	100.00%	19.07	89.53%	0.47%	99.89%	5.02	87.10%	2.90%
			2	100%	24.92	89.59%	0.41%	100%	5.03	86.84%	3.16%
			4	100%	25.83	89.58%	0.42%	100%	5.21	86.37%	3.63%
			8	100.00%	30.40	89.62%	0.38%	99.82%	5.10	85.72%	4.28%
			16	100%	34.34	89.61%	0.39%	100%	5.59	83.63%	6.37%
		1024	1	100%	6.04	89.27%	0.73%	99%	5.73	88.81%	1.19%
			2	99.95%	6.53	89.34%	0.66%	99.67%	5.72	88.62%	1.38%
			4	100%	6.27	89.11%	0.89%	100%	5.92	88.13%	1.87%
			8	100%	7.27	88.33%	1.67%	99%	5.68	87.42%	2.58%
			16	99.91%	13.50	87.84%	2.16%	99.95%	16.11	77.71%	12.29%
		4096	1	100%	8.86	89.02%	0.98%	100%	8.37	85.37%	4.63%
			2	99.97%	15.95	87.32%	2.68%	99.91%	8.42	82.61%	7.39%
			4	100.00%	10.86	88.26%	1.74%	100.00%	8.04	77.22%	12.78%
			8	100.00%	23.67	87.89%	2.11%	100.00%	29.97	75.82%	14.18%
			16	100.00%	49.42	87.51%	2.49%	99.04%	57.21	76.90%	13.10%
测试结果	3:7	64	1	99.84%	6.87	66.96%	3.04%	99.10%	4.32	65.26%	4.74%
			2	100.00%	12.61	67.15%	2.85%	99.24%	4.37	64.55%	5.45%
			4	99.84%	11.84	67.06%	2.94%	98.22%	4.23	64.94%	5.06%
			8	100.00%	10.85	67.03%	2.97%	97.58%	4.31	64.20%	5.80%
			16	99.24%	14.11	67.05%	2.95%	97.91%	4.25	64.76%	5.24%
		512	1	100%	33.11	69.30%	0.70%	100%	5.60	68.66%	1.34%
			2	99.91%	39.18	69.28%	0.72%	99.92%	5.66	68.47%	1.53%
			4	99.87%	49.00	69.29%	0.71%	100.00%	5.90	68.06%	1.94%
			8	100%	45.50	69.26%	0.74%	100%	5.69	66.92%	3.08%
			16	100%	22.03	69.22%	0.78%	99%	5.72	66.75%	3.25%
		1024	1	99.97%	5.62	69.02%	0.98%	100.00%	6.62	69.34%	0.66%
			2	100.00%	5.69	69.01%	0.99%	99.63%	6.05	69.24%	0.76%
			4	99.01%	5.76	68.41%	1.59%	99.74%	6.44	68.98%	1.02%
			8	100%	6.45	68.89%	1.11%	100%	6.25	68.36%	1.64%
			16	100.00%	18.14	68.36%	1.64%	98.98%	18.00	63.27%	6.73%
		4096	1	100%	8.99	68.48%	1.52%	100%	11.20	69.20%	0.80%
			2	100%	9.52	68.48%	1.52%	100%	10.40	68.65%	1.35%
			4	100%	11.68	68.13%	1.87%	100%	18.25	67.93%	2.07%
			8	100.00%	74.91	68.46%	1.54%	99.59%	33.62	68.19%	1.81%
			16	100.00%	87.65	65.68%	4.32%	100.00%	72.43	65.62%	4.38%
测试结果	5:5	64	1	85.87%	3.94	59.97%	9.97%	85.73%	3.77	59.51%	9.51%
			2	87.98%	4.25	59.44%	9.44%	84.87%	3.76	58.66%	8.66%
			4	87.14%	3.79	60.19%	10.19%	83.52%	3.64	59.39%	9.39%
			8	87.23%	4.08	59.99%	9.99%	83.65%	3.87	58.59%	8.59%
			16	85.11%	3.94	59.99%	9.99%	84.07%	3.99	59.18%	9.18%
		512	1	99%	5.55	49.04%	0.96%	100%	6.35	48.94%	1.06%
			2	99%	5.74	49.01%	0.99%	100%	6.61	48.70%	1.30%
			4	100.00%	5.43	48.50%	1.50%	99.87%	6.93	48.03%	1.97%
			8	100%	6.38	48.29%	1.71%	100%	7.14	47.00%	3.00%
			16	100%	12.67	48.54%	1.46%	99.10%	13.66	46.64%	3.36%
		1024	1	100.00%	6.68	49.52%	0.48%	99.95%	7.32	49.49%	0.51%
			2	99.96%	6.42	49.50%	0.50%	100.00%	8.32	49.41%	0.59%
			4	99.96%	6.72	49.18%	0.82%	100.00%	7.49	49.13%	0.87%
			8	100%	19.17	48.32%	1.68%	100%	12.48	48.26%	1.74%
			16	100%	23.87	48.51%	1.49%	100%	23.57	48.03%	1.97%
		4096	1	100%	10.02	49.38%	0.62%	100%	14.49	49.71%	0.29%
			2	100%	23.01	48.82%	1.18%	100%	14.46	49.66%	0.34%
			4	99.81%	26.67	47.28%	2.72%	99.92%	25.43	49.49%	0.51%
			8	100%	52.09	48.09%	1.91%	100%	49.45	49.13%	0.87%
			16	100%	95.52	48.49%	1.51%	100%	92.47	49.26%	0.74%



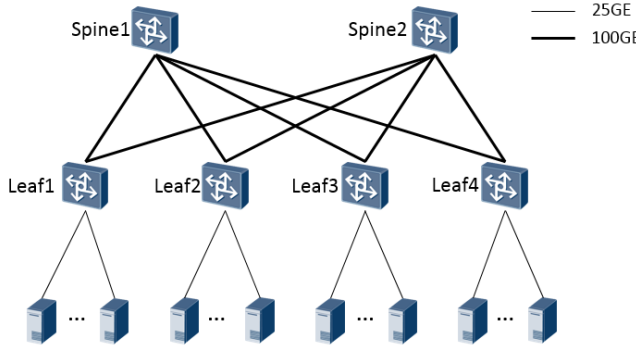
ODCC 无损网络测试基准

73	64	1	66.93%	3.80	48.28%	18.28%	75.34%	3.49	55.44%	25.44%
		2	65.36%	3.20	47.25%	17.25%	73.31%	3.55	53.16%	23.16%
		4	64.58%	3.27	47.72%	17.72%	72.82%	3.44	53.56%	23.56%
		8	65.81%	3.56	48.65%	18.65%	74.14%	3.54	55.04%	25.04%
		16	65.83%	3.35	48.15%	18.15%	74.31%	3.65	55.09%	25.09%
	512	1	100%	7.21	29.33%	0.67%	100%	8.50	29.24%	0.76%
		2	100%	7.27	29.29%	0.71%	100%	8.00	29.12%	0.88%
		4	100%	7.02	29.19%	0.81%	100%	9.13	28.76%	1.24%
		8	100.00%	11.51	28.91%	1.09%	99.72%	11.20	29.23%	0.77%
		16	99.89%	20.73	28.72%	1.28%	100.00%	20.52	29.15%	0.85%
	1024	1	99.99%	9.47	29.67%	0.33%	100.00%	14.71	29.64%	0.36%
		2	100.00%	8.98	29.65%	0.35%	100.00%	14.93	29.58%	0.42%
		4	99.98%	9.11	29.53%	0.47%	100.00%	14.89	29.39%	0.61%
		8	99.24%	20.05	29.53%	0.47%	100.00%	19.80	29.63%	0.37%
		16	100.00%	37.79	29.42%	0.58%	100.00%	39.53	29.59%	0.41%
	4096	1	100.00%	14.93	29.83%	0.17%	100.00%	22.77	29.84%	0.16%
		2	97%	20.54	29.79%	0.21%	100%	25.77	29.80%	0.20%
		4	99.74%	41.62	29.72%	0.28%	99.86%	46.63	29.71%	0.29%
		8	100%	72.50	29.69%	0.31%	100%	5477.23	29.88%	0.12%
		16	100.00%	154.62	27.68%	2.32%	99.53%	158.40	29.78%	0.22%
9:1	64	1	50.00%	2.99	32.56%	22.56%	65.87%	3.36	49.07%	39.07%
		2	49.99%	3.15	34.62%	24.62%	64.37%	3.39	47.79%	37.79%
		4	51.06%	3.05	33.35%	23.35%	63.87%	3.35	46.78%	36.78%
		8	51.24%	3.04	33.54%	23.54%	65.35%	3.41	47.77%	37.77%
		16	51.74%	3.00	35.67%	25.67%	69.20%	3.49	49.79%	39.79%
	512	1	100%	19.31	9.73%	0.27%	100%	38.46	9.73%	0.27%
		2	100.00%	20.70	9.72%	0.28%	99.94%	39.74	9.70%	0.30%
		4	100.00%	22.45	9.73%	0.27%	100.00%	42.52	9.73%	0.27%
		8	100.00%	34.65	9.72%	0.28%	99.99%	2036.48	9.74%	0.26%
		16	99.45%	202.77	9.73%	0.27%	99.86%	5084.19	9.74%	0.26%
	1024	1	100%	28.77	9.86%	0.14%	100%	61.75	9.86%	0.14%
		2	100%	29.18	9.86%	0.14%	100%	64.72	9.86%	0.14%
		4	100%	35.89	9.86%	0.14%	100%	78.10	9.86%	0.14%
		8	100.00%	61.25	9.87%	0.13%	100.00%	4597.05	9.87%	0.13%
		16	100%	222.58	9.87%	0.13%	100%	9574.09	9.88%	0.12%
	4096	1	100.00%	51.78	9.94%	0.06%	99.93%	2331.53	9.95%	0.05%
		2	100%	76.24	9.94%	0.06%	100%	5479.76	9.95%	0.05%
		4	100.00%	128.82	9.94%	0.06%	100.00%	10952.87	9.95%	0.05%
		8	99.96%	247.05	9.78%	0.22%	99.97%	21868.21	9.96%	0.04%
		16	100%	510.49	9.95%	0.05%	100%	17683.30	9.95%	0.05%
跨 Leaf										
TCP:ROCE	Message size	发端QP数	AI Fabric				传统PFC/ECN			
			吞吐	时延	比例	比例误差	吞吐	时延	比例	比例误差
1:9	64	1	100%	6.01	0.8340	6.60%	100%	5.72	0.7558	14.42%
		2	99.27%	9.07	0.8874	1.26%	99.96%	5.78	0.8204	7.96%
		4	100%	10.56	0.8863	1.37%	100%	5.82	0.8154	8.46%
		8	99.96%	10.36	0.8865	1.35%	99.56%	5.85	0.8054	9.46%
		16	100%	10.66	0.8784	2.16%	100%	5.77	0.7911	10.89%
	512	1	99.55%	21.93	0.8957	0.43%	99.98%	6.50	0.8684	3.16%
		2	100%	23.86	0.8964	0.36%	100%	6.66	0.8687	3.13%
		4	100%	24.78	0.8964	0.36%	100%	6.77	0.8648	3.52%
		8	100.00%	33.41	0.8965	0.35%	99.79%	6.74	0.8598	4.02%
		16	100%	32.68	0.8964	0.36%	100%	6.78	0.8430	5.70%
	1024	1	100%	7.73	0.8939	0.61%	100%	7.11	0.8876	1.24%
		2	100.00%	7.73	0.8934	0.66%	99.92%	7.20	0.8850	1.50%
		4	100%	7.98	0.8782	2.18%	100%	7.24	0.8814	1.86%
		8	100%	7.56	0.8719	2.81%	100%	7.27	0.8759	2.41%
		16	99.81%	18.35	0.8812	1.88%	99.81%	7.31	0.8243	7.57%
	4096	1	100%	13.66	0.8810	1.90%	100%	9.80	0.8625	3.75%
		2	100.00%	14.77	0.8804	1.96%	99.80%	9.67	0.8373	6.27%
		4	100.00%	9.99	0.8813	1.87%	99.90%	9.92	0.7863	11.37%
		8	100.00%	11.55	0.8779	2.21%	99.98%	10.66	0.7796	12.04%
		16	100.00%	49.99	0.8746	2.54%	99.97%	53.23	0.8145	8.55%

ODCC 无损网络测试基准

3:7	64	1	100.00%	7.04	0.7262	2.62%	96.60%	5.50	0.7184	1.84%					
		2	99.96%	11.97	0.7275	2.75%	94.45%	5.45	0.7071	0.71%					
		4	99.47%	12.88	0.7297	2.97%	93.78%	5.50	0.7093	0.93%					
		8	99.74%	12.45	0.7261	2.61%	93.20%	5.42	0.7020	0.20%					
		16	99.83%	13.51	0.7249	2.49%	93.58%	5.37	0.7051	0.51%					
		512	1	100%	32.17	0.6930	0.70%	100%	7.08	0.6853	1.47%				
			2	100.00%	41.17	0.6932	0.68%	99.98%	7.26	0.6850	1.50%				
			4	99.81%	45.08	0.6932	0.68%	99.71%	7.48	0.6814	1.86%				
			8	100%	51.25	0.6919	0.81%	100%	7.48	0.6704	2.96%				
		1024	16	100%	55.14	0.6852	1.48%	100%	7.48	0.6678	3.22%				
			1	100.00%	7.24	0.6908	0.92%	99.87%	7.99	0.6931	0.69%				
			2	100.00%	7.23	0.6903	0.97%	99.98%	8.33	0.6930	0.70%				
			4	100.00%	7.29	0.6880	1.20%	99.62%	8.31	0.6904	0.96%				
		4096	8	100%	8.21	0.6923	0.77%	100%	8.21	0.6847	1.53%				
			16	100.00%	9.08	0.6847	1.53%	99.97%	7.85	0.6593	4.07%				
			1	100%	16.74	0.6804	1.96%	100%	11.89	0.6922	0.78%				
	2		100%	10.79	0.6871	1.29%	100%	12.17	0.6884	1.16%					
		4	100%	12.68	0.6865	1.35%	100%	13.35	0.6855	1.45%					
		8	99.81%	53.61	0.6672	3.28%	99.57%	14.70	0.6760	2.40%					
		16	99.94%	63.61	0.6758	2.42%	99.95%	60.78	0.6709	2.91%					
	5:5	64	1	80.85%	4.76	0.6621	16.21%	91.58%	5.23	0.7081	20.81%				
			2	82.11%	5.57	0.6663	16.63%	84.89%	5.12	0.6789	17.89%				
			4	81.55%	5.39	0.6728	17.28%	85.33%	5.08	0.6918	19.18%				
			8	81.51%	5.67	0.6663	16.63%	83.66%	5.17	0.6765	17.65%				
			16	81.13%	5.35	0.6632	16.32%	84.72%	5.29	0.6714	17.14%				
			512	1	99%	7.01	0.4904	0.96%	100%	8.73	0.4885	1.15%			
				2	100%	7.19	0.4904	0.96%	100%	8.67	0.4871	1.29%			
				4	99.89%	7.17	0.4861	1.39%	99.86%	8.72	0.4814	1.86%			
				8	100%	19.80	0.4722	2.78%	100%	8.55	0.4733	2.67%			
			1024	16	100.00%	29.48	0.4792	2.08%	99.78%	7.97	0.4753	2.47%			
				1	100.00%	8.20	0.4955	0.45%	99.96%	9.28	0.4947	0.53%			
				2	100.00%	7.95	0.4953	0.47%	99.97%	9.29	0.4943	0.57%			
				4	99.79%	8.23	0.4944	0.56%	99.96%	9.78	0.4918	0.82%			
			4096	8	100%	42.05	0.4833	1.67%	100%	9.81	0.4892	1.08%			
				16	100%	23.18	0.4868	1.32%	100%	23.79	0.4843	1.57%			
				1	100%	11.54	0.4947	0.53%	100%	15.74	0.4971	0.29%			
		2		100%	11.99	0.4891	1.09%	100%	16.18	0.4968	0.32%				
			4	99.78%	22.57	0.4783	2.17%	99.97%	17.37	0.4955	0.45%				
			8	100%	48.61	0.4804	1.96%	100%	43.11	0.4900	1.00%				
			16	100%	131.42	0.4802	1.98%	100%	93.69	0.4894	1.06%				
		7:3	64	1	59.48%	4.51	0.5482	24.82%	87.97%	5.03	0.6974	39.74%			
				2	60.41%	4.39	0.5756	27.56%	78.76%	4.94	0.6641	36.41%			
				4	61.89%	4.53	0.5770	27.70%	76.25%	4.90	0.6538	35.38%			
				8	60.95%	4.53	0.5639	26.39%	77.79%	5.10	0.6555	35.55%			
				16	63.00%	4.58	0.5849	28.49%	78.57%	5.04	0.6650	36.50%			
				512	1	100%	8.52	0.2935	0.65%	100%	11.03	0.2927	0.73%		
					2	100%	8.87	0.2932	0.68%	100%	10.66	0.2912	0.88%		
					4	100%	8.68	0.2922	0.78%	100%	10.65	0.2880	1.20%		
					8	99.84%	8.99	0.2912	0.88%	99.83%	11.01	0.2899	1.01%		
				1024	16	99.77%	22.43	0.2842	1.58%	99.31%	19.30	0.2878	1.22%		
					1	99.94%	10.69	0.2969	0.31%	99.84%	16.86	0.2966	0.34%		
					2	100.00%	10.69	0.2966	0.34%	99.82%	16.76	0.2960	0.40%		
					4	99.95%	10.53	0.2956	0.44%	99.75%	15.54	0.2946	0.54%		
				4096	8	100.00%	11.39	0.2946	0.54%	99.95%	20.71	0.2948	0.52%		
					16	100.00%	42.84	0.2900	1.00%	99.98%	41.26	0.2929	0.71%		
					1	100.00%	16.23	0.2985	0.15%	99.95%	24.90	0.2985	0.15%		
			2		100%	17.34	0.2980	0.20%	100%	24.12	0.2980	0.20%			
				4	99.83%	40.71	0.2953	0.47%	99.97%	38.38	0.2973	0.27%			
				8	100%	70.09	0.2962	0.38%	100%	74.63	0.2971	0.29%			
				16	100.00%	178.58	0.2809	1.91%	99.83%	157.58	0.2947	0.53%			
			9:1	64	1	46.95%	4.46	0.4432	34.32%	78.95%	5.03	0.6600	56.00%		
					2	49.39%	4.45	0.4769	37.69%	74.20%	4.99	0.6374	53.74%		
					4	50.35%	4.47	0.4815	38.15%	73.22%	4.85	0.6353	53.53%		
					8	53.39%	4.45	0.5116	41.16%	74.31%	4.91	0.6384	53.84%		
					16	55.35%	4.52	0.5288	42.88%	71.76%	5.19	0.6428	54.28%		
					512	1	100%	21.23	0.0973	0.27%	100%	42.47	0.0972	0.28%	
						2	99.83%	21.42	0.0973	0.27%	99.80%	42.68	0.0972	0.28%	
						4	100.00%	23.06	0.0967	0.33%	99.78%	47.33	0.0970	0.30%	
						8	98.60%	35.15	0.0973	0.27%	99.61%	1456.97	0.0973	0.27%	
					1024	16	100.00%	85.42	0.0973	0.27%	99.63%	4802.43	0.0974	0.26%	
						1	100%	31.18	0.0987	0.13%	100%	68.48	0.0986	0.14%	
						2	100%	32.31	0.0987	0.13%	100%	71.87	0.0987	0.13%	
						4	100%	37.58	0.0986	0.14%	100%	75.84	0.0987	0.13%	
					4096	8	99.65%	61.85	0.0987	0.13%	99.22%	3009.74	0.0987	0.13%	
						16	100%	155.79	0.0987	0.13%	100%	7956.02	0.0988	0.12%	
						1	99.44%	56.85	0.0994	0.06%	99.90%	2210.73	0.0995	0.05%	
				2		100%	74.35	0.0994	0.06%	100%	3203.22	0.0995	0.05%		
					4	99.77%	150.06	0.0994	0.06%	99.96%	6486.22	0.0995	0.05%		
					8	99.95%	255.86	0.0974	0.26%	99.89%	8029.71	0.0995	0.05%		
					16	100%	562.61	0.0994	0.06%	100%	374.67	0.0994	0.06%		
				备注											

## 1.2.2 混跑8打1

测试项目	混跑8打1
测试目的	测试TCP、RoCE混跑场景下调度的精准性、出接口带宽，同时测试RoCE时延不受TCP影响保持低时延。
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。TCP流量映射到队列0，RoCE流量映射到队列3，配置0、3队列调度方式为DRR。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 选择9台服务器使用ib_write_bw工具进行8打1带宽测试，同时使用netperf工具进行TCP8打1，其中4个源端和目的端处于一个Leaf下，另外4个源端和目的端处于不同Leaf下。发端QP数遍历1、2、4、8、16，DRR调度比重覆盖1:9，3:7，5:5，7:3，9:1，记录带宽利用率和带宽比例。</li> <li>2) 使用步骤1)的流量作为背景流的情况下，对跨Leaf的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> </ol>
预期结果	<ol style="list-style-type: none"> <li>1) 交换机上RoCEv2流量无丢包，除64B外交换机出接口带宽达到95%以上。吞吐比例控制在5%误差范围内。</li> <li>2) RoCE时延在TCP:RoCE=9:1比例下控制在1ms内。</li> </ol>

ODCC 无损网络测试基准

测试结果

TCP:ROCE	Message size	发端QP数	AI Fabric				传统PFC/ECN			
			吞吐	时延	比例	比例误差	吞吐	时延	比例	比例误差
1:9	64	1	100%	13.89	0.8834	1.66%	100%	5.60	0.8548	4.52%
		2	100.00%	40.21	0.8853	1.47%	99.96%	5.66	0.8494	5.06%
		4	99%	16.22	0.8868	1.32%	100%	5.68	0.8388	6.12%
		8	100.00%	17.15	0.8855	1.45%	99.95%	5.76	0.8223	7.77%
		16	100%	16.88	0.8855	1.45%	99%	5.85	0.8339	6.61%
	512	1	99.71%	54.44	0.8964	0.36%	100.00%	6.57	0.8887	1.13%
		2	100%	51.37	0.8959	0.41%	100%	6.95	0.8868	1.32%
		4	100%	185.16	0.8964	0.36%	100%	6.71	0.8808	1.92%
		8	99.79%	538.38	0.8965	0.35%	100.00%	6.74	0.8630	3.70%
		16	100%	104.91	0.8968	0.32%	100%	6.64	0.8203	7.97%
	1024	1	99%	7.52	0.8932	0.68%	100%	7.43	0.8941	0.59%
		2	100.00%	7.61	0.8917	0.83%	100.00%	7.33	0.8916	0.84%
		4	100%	8.75	0.8908	0.92%	100%	7.49	0.8885	1.15%
		8	100%	11.67	0.8880	1.20%	100%	8.37	0.8583	4.17%
		16	100.00%	30.18	0.8909	0.91%	100.00%	24.76	0.8857	1.43%
	4096	1	99%	10.64	0.8883	1.17%	100%	12.44	0.8963	0.37%
2		99.96%	10.65	0.8846	1.54%	100.00%	12.39	0.8952	0.48%	
4		100.00%	21.09	0.8818	1.82%	100.00%	12.31	0.8903	0.97%	
8		100.00%	65.68	0.8717	2.83%	99.87%	51.91	0.8472	5.28%	
16		100.00%	106.82	0.8722	2.78%	100.00%	98.58	0.8631	3.69%	
3:7	64	1	100.00%	16.26	0.6718	2.82%	100.00%	5.89	0.6596	4.04%
		2	99.93%	23.44	0.6719	2.81%	100.00%	5.95	0.6527	4.73%
		4	99.98%	27.94	0.6718	2.82%	100.00%	6.10	0.6352	6.48%
		8	100.00%	27.98	0.6719	2.81%	99.99%	6.19	0.6486	5.14%
		16	99.77%	23.59	0.6718	2.82%	99.86%	6.50	0.6341	6.59%
	512	1	100%	69.03	0.6932	0.68%	100%	7.50	0.6902	0.98%
		2	99.80%	84.86	0.6929	0.71%	99.82%	7.60	0.6873	1.27%
		4	99.70%	17.17	0.6431	5.69%	99.53%	7.50	0.6799	2.01%
		8	100%	23.42	0.6777	2.23%	100%	7.41	0.6825	1.75%
		16	100%	34.34	0.6749	2.51%	100%	14.80	0.6592	4.08%
	1024	1	99.65%	7.63	0.6954	0.46%	99.93%	9.68	0.6949	0.51%
		2	100.00%	7.52	0.6946	0.54%	100.00%	9.32	0.6936	0.64%
		4	99.75%	24.94	0.6797	2.03%	99.68%	9.30	0.6915	0.85%
		8	100%	20.65	0.6850	1.50%	100%	10.00	0.6887	1.13%
		16	99.63%	31.39	0.6866	1.34%	100.00%	35.09	0.6893	1.07%
	4096	1	98%	10.20	0.6882	1.18%	100%	14.27	0.6971	0.29%
2		100%	17.36	0.6781	2.19%	100%	13.98	0.6959	0.41%	
4		100%	12.10	0.6831	1.69%	100%	14.39	0.6936	0.64%	
8		99.61%	61.55	0.6803	1.97%	99.80%	59.45	0.6882	1.18%	
16		99.98%	132.79	0.6824	1.76%	100.00%	127.42	0.6865	1.35%	
5:5	64	1	98.67%	30.84	0.4679	3.21%	100.00%	6.95	0.4538	4.62%
		2	99.79%	43.55	0.4679	3.21%	100.00%	6.91	0.4407	5.93%
		4	100.00%	44.16	0.4682	3.18%	99.96%	6.82	0.4254	7.46%
		8	100.00%	40.60	0.4680	3.20%	100.00%	6.71	0.4362	6.38%
		16	99.66%	46.28	0.4680	3.20%	99.98%	7.27	0.4399	6.01%
	512	1	100%	7.12	0.4917	0.83%	100%	8.48	0.4900	1.00%
		2	100%	7.43	0.4880	1.20%	100%	9.26	0.4872	1.28%
		4	100.00%	7.73	0.4805	1.95%	99.58%	8.97	0.4822	1.78%
		8	100%	20.50	0.4873	1.27%	100%	9.62	0.4886	1.14%
		16	99.95%	24.48	0.4862	1.38%	99.64%	27.61	0.4898	1.02%
	1024	1	99.38%	8.89	0.4955	0.45%	100.00%	12.44	0.4953	0.47%
		2	99.66%	8.60	0.4949	0.51%	100.00%	11.49	0.4941	0.59%
		4	99.90%	8.88	0.4925	0.75%	99.97%	12.17	0.4919	0.81%
		8	100%	19.30	0.4865	1.35%	100%	21.05	0.4931	0.69%
		16	100%	48.80	0.4873	1.27%	100%	50.84	0.4937	0.63%
	4096	1	100%	13.02	0.4962	0.38%	100%	17.50	0.4977	0.23%
2		100%	13.27	0.4956	0.44%	100%	20.46	0.4971	0.29%	
4		99.21%	49.20	0.4842	1.58%	99.99%	43.42	0.4939	0.61%	
8		100%	92.50	0.4798	2.02%	100%	89.39	0.4963	0.37%	
16		100%	183.31	0.4856	1.44%	99%	198.84	0.4881	1.19%	
7:3	64	1	90.00%	6.66	0.3337	3.37%	88.67%	6.43	0.3338	3.38%
		2	88.73%	6.30	0.3303	3.03%	89.87%	6.37	0.3336	3.36%
		4	89.42%	6.13	0.3298	2.98%	89.45%	6.44	0.3333	3.33%
		8	89.80%	6.91	0.3331	3.31%	88.71%	6.72	0.3299	2.99%
		16	88.49%	7.98	0.3307	3.07%	89.06%	6.58	0.3404	4.04%
	512	1	100%	9.83	0.2934	0.66%	100%	14.05	0.2928	0.72%
		2	100%	9.86	0.2927	0.73%	100%	14.02	0.2911	0.89%
		4	100%	10.13	0.2927	0.73%	100%	15.30	0.2926	0.74%
		8	100.00%	20.98	0.2914	0.86%	100.00%	22.61	0.2916	0.84%
		16	99.92%	40.79	0.2896	1.04%	99.98%	44.14	0.2916	0.84%
	1024	1	100.00%	12.20	0.2968	0.32%	99.84%	20.43	0.2966	0.34%
		2	99.11%	12.16	0.2957	0.43%	99.83%	21.29	0.2956	0.44%
		4	99.73%	21.46	0.2963	0.37%	100.00%	22.39	0.2965	0.35%
		8	99.59%	44.79	0.2950	0.50%	100.00%	44.66	0.2966	0.44%
		16	100.00%	81.98	0.2892	1.08%	99.87%	5546.48	0.2971	0.29%
	4096	1	99.79%	19.79	0.2985	0.15%	99.99%	33.04	0.2985	0.15%
2		100%	41.30	0.2981	0.19%	100%	41.11	0.2984	0.16%	
4		100.00%	76.19	0.2982	0.18%	99.81%	77.53	0.2986	0.14%	
8		100%	183.00	0.2805	1.95%	100%	5915.33	0.2988	0.12%	
16		99.93%	286.59	0.2810	1.90%	99.67%	350.61	0.2960	0.40%	

	64	1	70.87%	5.76	0.1927	9.27%	75.99%	6.47	0.2327	13.27%
		2	73.28%	6.08	0.1817	8.17%	75.47%	6.42	0.2323	13.23%
		4	72.00%	6.04	0.1917	9.17%	76.37%	6.67	0.2396	13.96%
		8	72.04%	5.76	0.1960	9.60%	78.47%	6.80	0.2447	14.47%
	512	16	80.00%	25.18	0.2650	16.50%	81.05%	7.24	0.2800	18.00%
		1	100%	29.73	0.0973	0.27%	99%	66.50	0.0973	0.27%
		2	99.97%	31.79	0.0973	0.27%	100.00%	75.54	0.0973	0.27%
		4	99.87%	38.33	0.0973	0.27%	100.00%	84.33	0.0973	0.27%
	1024	8	100.00%	294.97	0.0973	0.27%	99.98%	5111.64	0.0974	0.26%
		16	99.73%	526.49	0.0973	0.27%	100.00%	11192.42	0.0974	0.26%
		1	100%	47.37	0.0987	0.13%	100%	118.56	0.0987	0.13%
		2	100%	50.41	0.0987	0.13%	100%	132.95	0.0986	0.14%
	4096	4	100%	66.85	0.0987	0.13%	100%	3822.97	0.0987	0.13%
		8	99.96%	198.53	0.0987	0.13%	100.00%	9867.41	0.0987	0.13%
		16	99%	579.97	0.0987	0.13%	100%	20674.71	0.0988	0.12%
		1	100.00%	84.45	0.0995	0.05%	100.00%	5135.91	0.0995	0.05%
		2	100%	170.69	0.0995	0.05%	100%	10810.08	0.0995	0.05%
		4	99.67%	262.28	0.0995	0.05%	99.82%	21786.30	0.0995	0.05%
		8	99.63%	734.16	0.0995	0.05%	98.98%	33629.22	0.0995	0.05%
		16	100%	954.12	0.0997	0.03%	100%	227093.88	0.0996	0.04%
备注										

## 2 运维功能

### 2.1 PFC 死锁预防

测试项目	PFC死锁预防
测试目的	验证设备在同POD内支持预防PFC死锁。
测试环境	<p>测试组网：</p> <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。其中RoCE映射到3队列，开启3、4队列PFC功能，并配置3、4队列DRR调度，调度比例为5:5。</li> <li>4) 交换机上不配置PFC死锁检测功能。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 为了触发更多PFC复现Deadlock，交换机关闭AI Fabric，关闭DCQCN；</li> <li>2) 按照以上拓扑shutdown两条链路。</li> <li>3) 如图所示打四条DSCP为24的流（ROCEv2），造成队列3的CBD，触发Deadlock。（调整流量路径的方法：配置静态主机路由，或调整传输层port号等）</li> </ol>

	<p>4) 如果流量不够大没有触发足够的PFC，可以对leaf2或leaf3的上行流量限速，触发更多PFC，最终触发Deadlock。</p> <p>5) 停止所有流量，观察PFC Deadlock仍然持续。</p> <p>6) 关闭CBD中任一port的PFC，从而切断CBD，解除Deadlock。</p> <p>7) 设置死锁预防配置，重新打上述流量，Deadlock不再发生。</p> <p>8) 观察PFC的变化，观察各端口流量cos队列的变化。</p>
预期结果	<p>1) 复现 PFC Deadlock。</p> <p>2) 设置死锁预防配置后，不能复现 PFC Deadlock。</p>
测试结果	<p>H10&lt;-&gt;h7 H4&lt;-&gt;H3</p> <pre>[root@h7 ~]# ./send_traffic.sh -c h10,h7 -n 1 -s 4096 -l 120000 &amp; [1] 444287 [root@h7 ~]# ssh h10 Last login: Fri Jul 19 15:08:22 2019 from 192.89.52.195 [root@h10 ~]# ./send_traffic.sh -c h10,h7 -n 1 -s 4096 -l 120000 &amp; [1] 163652 [root@h10 ~]# exit logout Connection to h10 closed. [root@h7 ~]# ssh h4 Last login: Fri Jul 19 14:54:14 2019 from 192.89.52.141 [root@h4 ~]# ./send_traffic.sh -c h3,h4 -n 1 -s 4096 -l 120000 &amp; [1] 23434 [root@h4 ~]# exit logout Connection to h4 closed. [root@h7 ~]# ssh h3 Last login: Fri Jul 19 14:55:41 2019 from 192.89.52.195 [root@h3 ~]# ./send_traffic.sh -c h3,h4 -n 1 -s 4096 -l 120000 &amp; [1] 19708 [root@h3 ~]#</pre>

```

<Leaf1>dis dcb pfc
-----
Interface      Queue      Received(Frames)  ReceivedRate(pps)  DeadLockNum
              Transmitted(Frames)  TransmittedRate(pps)  RecoveryNum
-----
100GE1/0/1     3           3020027            0                   0
              4             0                  0                   0
100GE1/0/2     3             0                  0                   0
              4             0                  0                   0
25GE1/0/1     3           754944             0                   0
              4             0                  0                   0
25GE1/0/2     3             0                  0                   0
              4             0                  0                   0
25GE1/0/3     3             0                  0                   0
              4             0                  0                   0
-----
<Leaf1>dis dcb pfc
-----
Interface      Queue      Received(Frames)  ReceivedRate(pps)  DeadLockNum
              Transmitted(Frames)  TransmittedRate(pps)  RecoveryNum
-----
100GE1/0/1     3           3117125            0                   0
              4             0                  0                   0
100GE1/0/2     3             0                  0                   0
              4             0                  0                   0
25GE1/0/1     3           779212             0                   0
              4             0                  0                   0
25GE1/0/2     3             0                  0                   0
              4             0                  0                   0
25GE1/0/3     3             0                  0                   0
              4             0                  0                   0
-----
<Leaf1>

```

```

<Leaf1> dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface      PHY      Protocol  InUti  OutUti  inErrors  outErrors
100GE1/0/1     up      up        0.05%  0.01%  0         0
25GE1/0/1     up      up        0%     0.05%  0         0
25GE1/0/2     up      up        0%     0.01%  0         0
Meth0/0/0     up      up        0.01%  0.01%  0         0
NULL0         up      up(s)    0%     0%     0         0
Vlanif10     up      up        --     --     0         0
<Leaf1>

```

配置PFC dead lock free流量不受影响

```
<Leaf2> dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDAP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface          PHY          Protocol    InUti  OutUti    inErrors  outErrors
100GE1/0/1         up           up          0.01% 46.76%    0         0
100GE1/0/2         up           up          46.76% 0.01%    0         0
25GE1/0/1          up           up          99.98% 99.97%    0         0
LoopBack0          up           up(s)       0%     0%        0         0
MEth0/0/0          up           up          0.01% 0.01%    0         0
NULL0              up           up(s)       0%     0%        0         0
Vlanif20           up           up          --     --        0         0
```

出现4队列报文

```
<Leaf3> dis q q s int 100g 1/0/2
Queue  CIR/PIR          Passed          Pass Rate          Dropped          Drop Rate          Drop Time
(% or kbps) (Packets/Bytes) (pps/bps) (Packets/Bytes) (pps/bps)
-----
0      0                0               0                 0                 0                 -
100000000
-----
1      0                0               0                 0                 0                 -
100000000
-----
2      0                0               0                 0                 0                 -
100000000
-----
3      0                73844526       1447931           0                 0                 -
100000000      102404595868  16063466016
-----
4      0                55720164       1092552           0                 0                 -
100000000      86852595976  13623936616
-----
5      0                0               0                 0                 0                 -
100000000
-----
6      0                60              1                 0                 0                 -
100000000      4840          736
-----
7      0                0               0                 0                 0                 -
100000000
```

停流后  
Pfc计数不增加



```

<Leaf2>dis dcb pfc
-----
Interface      Queue      Received(Frames)
              Transmitted(Frames)  ReceivedRate(pps)
              TransmittedRate(pps)  DeadlockNum
              RecoveryNum
-----
100GE1/0/1     3           0
              0
              0
              0
              0
              0
              0
100GE1/0/2     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/1      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
-----
<Leaf2>dis dcb pfc
-----
Interface      Queue      Received(Frames)
              Transmitted(Frames)  ReceivedRate(pps)
              TransmittedRate(pps)  DeadlockNum
              RecoveryNum
-----
100GE1/0/1     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
100GE1/0/2     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/1      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
-----
<Leaf2>

```

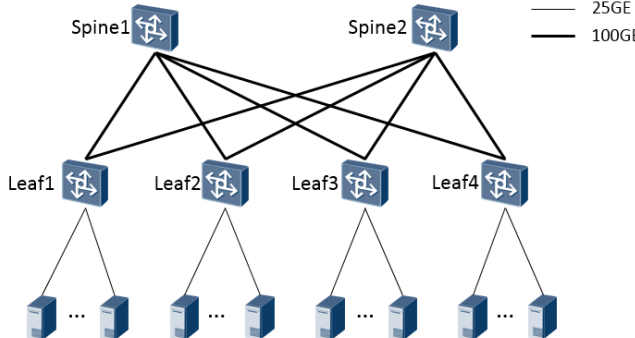
```

<Leaf3> dis dcb pfc
-----
Interface      Queue      Received(Frames)
              Transmitted(Frames)  ReceivedRate(pps)
              TransmittedRate(pps)  DeadlockNum
              RecoveryNum
-----
100GE1/0/1     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
100GE1/0/2     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/1      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/2      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/3      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/4      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/5      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/6      3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
25GE1/0/24     3           0
              0
              0
              0
              0
              0
              0
              4           0
              0
              0
              0
              0
              0
-----
<Leaf3>

```

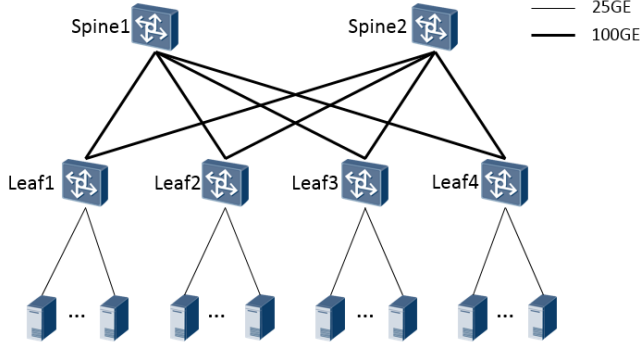
备注	
----	--

## 2.2 RoCE 丢包可视化

测试项目	丢包可视化
测试目的	测试设备支持RoCEv2报文的丢弃报文信息查看
测试环境	<p>测试组网：</p> 
测试步骤	<ol style="list-style-type: none"> <li>1) 使能丢弃报文捕获功能。</li> <li>2) 所有交换机的使用的端口下去使能PFC, ECN等无损网络功能, 对9个节点使用ib_write_bw打流工具产生8打1的RoCE流量, 有预期结果1。</li> </ol>
预期结果	1) 通过命令行可以查看到交换机丢弃报文的信息, 包括丢包时间、丢包数、源IP、目的IP、源MAC、目的MAC等信息。
测试结果	<p>端口下去使能 PFC、ECN</p> <pre>&lt;Leaf1&gt;dis curr int 25g 1/0/1 # interface 25GE1/0/1 port default vlan 10 trust dscp device transceiver 25GBASE-COPPER # return</pre> <p>2 打 1 RoCE 流量</p>

	<pre> &lt;Leaf1&gt;dis int b   i up PHY: Physical *down: administratively down ^down: standby (l): loopback (s): spoofing (b): BFD down (e): ETHOAM down (d): Dampening Suppressed (p): port alarm down (dl): DLDAP down (c): CFM down (sd): STP instance discarding InUti/OutUti: input utility rate/output utility rate Interface      PHY      Protocol  InUti  OutUti  inErrors  outErrors 100GE1/0/1     up      up        13.31% 0.17%   0         0 100GE1/0/2     up      up        13.98% 0.18%   0         0 25GE1/0/1      up      up        1.38% 99.83%  0         0 Meth0/0/0      up      up        0.01% 0.01%   0         0 NULL0          up      up(s)     0%     0%     0         0 Vlanif10      up      up        --     --     0         0 &lt;Leaf1&gt;  &lt;Leaf1&gt;dis q q s int 25g 1/0/1 Queue  CIR/PIR      Passed      Pass Rate      Dropped      Drop Rate  Drop (% or kbps) (Packets/Bytes) (pps/bps) (Packets/Bytes) (pps/bps) ----- 0      0            0            0              0              0 25000000 ----- 1      0            0            0              0              0 25000000 ----- 2      0            0            0              0              0 25000000 ----- 3      0            119520010    1085259        7243186        52679 2019.0 25000000      253132935828 18387714640    15340865292    892590800 15:3 ----- 4      0            0            0              0              0 25000000 ----- 5      0            0            0              0              0 25000000 ----- 6      0            22           0              0              0 25000000      3310          168 -----  &lt;Leaf1&gt;dis qos capture statistics drop-packet ip tcb slot 1 Time: 2019-07-17 15:33:04 Total Drop Packets: 1024 ----- Num   Time           Source-MAC      Destination-MAC  Vlan Queue Source-IP      Destination-IP   Source-Port      Destination-Port IfName PacketHeader ----- 0     2019-07-17 15:32:02 0012-2545-4102 60fa-9df9-4a82 0 10.0.4.3      10.0.1.4        49649           4791 3     25GE1/0/1     Egress admission drop can cause a trigger 60fa9df94a820012254541020800456a083cf7ef40003e1123510a0004030a000104c1f112b70828 00000640ffff000005dd00af048a00007fc4d7aea000011d5ad00001000b748fa3d3679c8a4e335 1     2019-07-17 15:32:02 0012-2545-4102 60fa-9df9-4a82 0 10.0.4.3      10.0.1.4        49649           4791 3     25GE1/0/1     Egress admission drop can cause a trigger 60fa9df94a820012254541020800456a082cf7f240003e11235e0a0004030a000104c1f112b70818 00000840ffff000005dd80af048d3cb879877fe5bc0ed269d3080b9ee9bc7b5b203d70e71dff0a06 2     2019-07-17 15:32:02 0012-2545-4102 60fa-9df9-4a82 0 10.0.4.3      10.0.1.4        49649           4791 ----- More ----- </pre>
备注	

## 2.3 RoCE 性能可视化

测试项目	RoCE性能可视化
测试目的	测试设备支持RoCEv2报文的基于流的路径、吞吐、RTT可视化
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 在Leaf上配置丢包可视和性能可视的RoCEv2智能流量分析功能。</li> <li>2) 同一对Leaf的2台Server发送双向RoCE流量，有预期结果1。</li> <li>3) 跨Leaf的2台Server发送双向RoCE流量，有预期结果1。</li> </ol>
预期结果	1) 通过命令行可以查看到基于流的吞吐、RTT、流量在交换机的出入端口。
测试结果	<p>配置：</p> <pre>[Leaf4] dis curr   sec in assign # assign forward enp netstream enable slot 1 #</pre> <pre>[Leaf4]dis curr   sec in analy # traffic-analysis rocev2 global inbound traffic-analysis rocev2 drop global [Leaf4]</pre> <p>同 TOR</p>

```
[Leaf4]dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDAP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface      PHY      Protocol  InUti  OutUti  inErrors  outErrors
100GE1/0/1     up       up        0.01%  0.01%   0         0
100GE1/0/2     up       up        0.01%  0.01%   0         0
25GE1/0/1      up       up        2.00%  99.76%  0         0
25GE1/0/2      up       up        99.76%  2.00%   0         0
Meth0/0/0      up       up        0.01%  0.01%   0         0
NULL0          up       up(s)     0%     0%      0         0
Vlanif40       up       up        --     --      0         0
```

```
[Leaf4] display traffic-analysis rocev2 cache sl 1
NOTE: S2C: server to client C2S: client to server
T: throughput(kB/s) RTT: round trip time(ns)
Traffic analysis cache information:
-----
Flow Created Time
ClientIP      ClientQP      C2S Read Pkt T      C2S Send/Write Pkt T
C2S RTT      C2S NAK Pkts  C2S Interface
ServerIP      ServerQP      S2C Read Pkt T      S2C Send/Write Pkt T
S2C RTT      S2C NAK Pkts  S2C Interface
-----
2019-07-17 23:58:51
10.0.4.6      11924         0                    0
0             0             25GE1/0/2
10.0.4.3      11164         0                    0
0             0             25GE1/0/1
2019-07-17 23:58:52
10.0.4.6      11925         0                    2998000
2834         0             25GE1/0/2
10.0.4.3      11165         0                    0
0             0             25GE1/0/1
-----
```

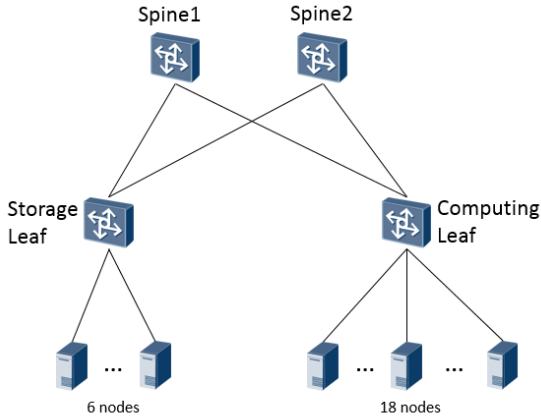
跨 TOR:

```
[Leaf4]dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDAP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface      PHY      Protocol  InUti  OutUti  inErrors  outErrors
100GE1/0/1     up       up        25.11%  0.50%   0         0
100GE1/0/2     up       up        0.01%  0.01%   0         0
25GE1/0/1      up       up        2.02%  100%    0         0
25GE1/0/2      up       up        0%     0.01%   0         0
Meth0/0/0      up       up        0.01%  0.01%   0         0
NULL0          up       up(s)     0%     0%      0         0
Vlanif40       up       up        --     --      0         0
[Leaf4]
```

	<pre>[Leaf4] display traffic-analysis rocev2 cache sl 1 NOTE: S2C: server to client C2S: client to server T: throughput(kB/s) RTT: round trip time(ns) Traffic analysis cache information: ----- Flow Created Time ClientIP          ClientQP          C2S Read Pkt T   C2S Send/Write Pkt T C2S RTT           C2S NAK Pkts     C2S Interface ServerIP          ServerQP          S2C Read Pkt T   S2C Send/Write Pkt T S2C RTT           S2C NAK Pkts     S2C Interface ----- 2019-07-18 00:02:02 10.0.1.4          10018            0                 2972366 2612              0                 100GE1/0/1 10.0.4.3          11167            0                 0 0                 0                 25GE1/0/1 ----- [Leaf4]</pre>
备注	

### 3 分布式存储测试

#### 3.1 分布式存储性能(DSS)

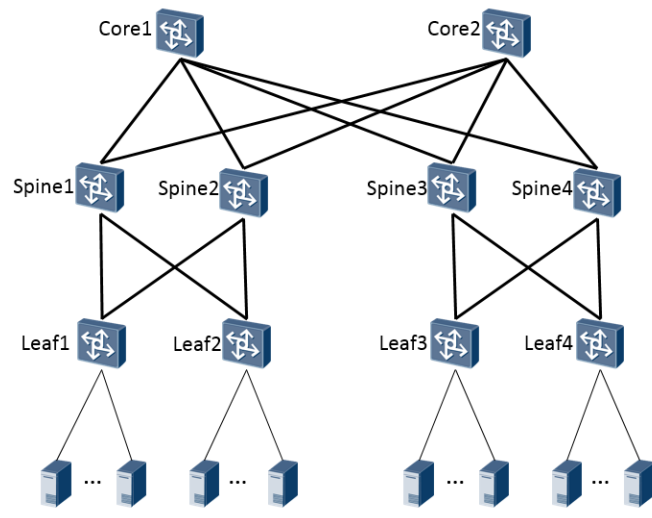
测试项目	分布式存储性能测试
测试目的	分布式存储性能测试
测试环境	<p>测试组网：</p>  <pre> graph TD     Spine1 --- StorageLeaf     Spine1 --- ComputingLeaf     Spine2 --- StorageLeaf     Spine2 --- ComputingLeaf     StorageLeaf --- N1[ ]     StorageLeaf --- N2[ ]     StorageLeaf --- N3[ ]     StorageLeaf --- N4[ ]     StorageLeaf --- N5[ ]     StorageLeaf --- N6[ ]     ComputingLeaf --- C1[ ]     ComputingLeaf --- C2[ ]     ComputingLeaf --- C3[ ]     ComputingLeaf --- C4[ ]     ComputingLeaf --- C5[ ]     ComputingLeaf --- C6[ ]     ComputingLeaf --- C7[ ]     ComputingLeaf --- C8[ ]     ComputingLeaf --- C9[ ]     ComputingLeaf --- C10[ ]     ComputingLeaf --- C11[ ]     ComputingLeaf --- C12[ ]     ComputingLeaf --- C13[ ]     ComputingLeaf --- C14[ ]     ComputingLeaf --- C15[ ]     ComputingLeaf --- C16[ ]     ComputingLeaf --- C17[ ]     ComputingLeaf --- C18[ ]   </pre> <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) Spine和Leaf之间配置动态路由，多条链路形成ECMP，其中Leaf1下6台存储节点安装NVME SSD硬盘。</li> <li>2) 部署分布式存储系统，将所有存储节点池化，计算节点安装FIO工具</li> <li>3) 在Leaf2下的18个计算节点安装FIO工具以测试IO读写性能。</li> <li>4) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。</li> </ol>

测试步骤	1) 每个计算节点挂载 4 个 LUN，每个 LUN100G，一共 72 个 LUN。使用 fio 工具同时对这个 72 个 LUN 测试以下典型存储流量模型，记录带宽、IOPS、平均时延、拖尾指标。					
	业务类型	IO 类型		典型读写比例		IO 大小
		随机	顺序	读	写	
	OLTP	Y		50%	50%	8K-64K
	OLAP		Y	50%	50%	256K-4M
	VDI	Y		20%	80%	1K-64K
	Exchange Server	Y		60%	40%	32K-
		Y		30%	70%	512K
	视频		Y	20%	80%	64K
			Y	20%	80%	16K-64K
		Y	40%	60%	16K-64K	
预期结果	记录 IOPS、BW 和平均时延、拖尾时延。					
测试结果						
备注						

## 4 大规模 RoCE 网络测试

### 4.1 跨 POD 混跑 4 打 1

测试项目	跨POD混跑4打1
测试目的	测试TCP、ROCE混跑场景下调度的精准性、出接口带宽，同时测试RoCE时延不受TCP影响保持低时延。
测试环境	测试组网：

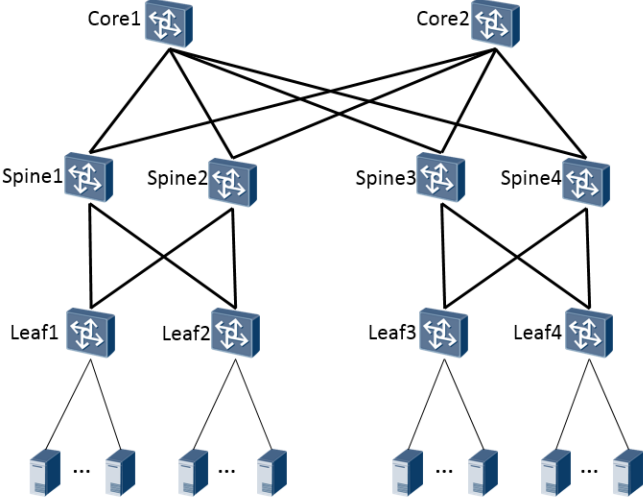
	 <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。TCP流量映射到队列0，RoCE流量映射到队列3，配置0、3队列调度方式为DRR。</li> </ol>																																																																																																																																																																																																												
<p>测试步骤</p>	<ol style="list-style-type: none"> <li>1) 对跨POD的5台服务器使用ib_write_bw工具进行4打1带宽测试，同时使用netperf工具进行TCP4打1，消息长度遍历64B、512B、1024B、4096B，发端QP数遍历1、2、4、8、16，DRR调度比重覆盖1:9，3:7，5:5，7:3，9:1，记录带宽利用率和带宽比例。</li> <li>2) 使用步骤3)的流量作为背景流的情况下，对跨POD的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> </ol>																																																																																																																																																																																																												
<p>预期结果</p>	<ol style="list-style-type: none"> <li>1) 交换机上RoCEv2流量无丢包，除64B外交换机出接口带宽达到95%以上。吞吐比例控制在5%误差范围内。</li> <li>2) RoCE时延在TCP:RoCE=9:1比例下控制在10ms内。</li> </ol>																																																																																																																																																																																																												
<p>测试结果</p>	<table border="1"> <thead> <tr> <th rowspan="2">TCP:RoCE</th> <th rowspan="2">Message siz</th> <th rowspan="2">发端QP数</th> <th colspan="4">AI Fabric</th> <th colspan="4">传统PFC/ECN</th> </tr> <tr> <th>吞吐</th> <th>时延</th> <th>比例</th> <th>比例误差</th> <th>吞吐</th> <th>时延</th> <th>比例</th> <th>比例误差</th> </tr> </thead> <tbody> <tr> <td rowspan="16">1:9</td> <td rowspan="4">64</td> <td>1</td> <td>100%</td> <td>7.01</td> <td>0.7670</td> <td>13.30%</td> <td>100%</td> <td>6.75</td> <td>0.7371</td> <td>16.29%</td> </tr> <tr> <td>2</td> <td>99.94%</td> <td>10.60</td> <td>0.8873</td> <td>1.27%</td> <td>100.00%</td> <td>6.91</td> <td>0.8185</td> <td>8.15%</td> </tr> <tr> <td>4</td> <td>99%</td> <td>33.26</td> <td>0.8875</td> <td>1.25%</td> <td>100%</td> <td>6.79</td> <td>0.8113</td> <td>8.87%</td> </tr> <tr> <td>8</td> <td>99.28%</td> <td>11.15</td> <td>0.8795</td> <td>2.05%</td> <td>99.37%</td> <td>7.04</td> <td>0.8000</td> <td>10.00%</td> </tr> <tr> <td>16</td> <td>100%</td> <td>11.27</td> <td>0.8788</td> <td>2.12%</td> <td>100%</td> <td>6.85</td> <td>0.7865</td> <td>11.35%</td> </tr> <tr> <td rowspan="4">512</td> <td>1</td> <td>100.00%</td> <td>22.67</td> <td>0.8964</td> <td>0.36%</td> <td>99.99%</td> <td>7.92</td> <td>0.8683</td> <td>3.17%</td> </tr> <tr> <td>2</td> <td>100%</td> <td>25.50</td> <td>0.8964</td> <td>0.36%</td> <td>100%</td> <td>7.99</td> <td>0.8636</td> <td>3.64%</td> </tr> <tr> <td>4</td> <td>100%</td> <td>29.53</td> <td>0.8961</td> <td>0.39%</td> <td>100%</td> <td>7.87</td> <td>0.8611</td> <td>3.89%</td> </tr> <tr> <td>8</td> <td>99.97%</td> <td>34.80</td> <td>0.8963</td> <td>0.37%</td> <td>100.00%</td> <td>7.86</td> <td>0.8569</td> <td>4.31%</td> </tr> <tr> <td>16</td> <td>100%</td> <td>36.94</td> <td>0.8964</td> <td>0.36%</td> <td>100%</td> <td>7.92</td> <td>0.8409</td> <td>5.91%</td> </tr> <tr> <td rowspan="4">1024</td> <td>1</td> <td>100%</td> <td>8.58</td> <td>0.8935</td> <td>0.65%</td> <td>100%</td> <td>8.32</td> <td>0.8871</td> <td>1.29%</td> </tr> <tr> <td>2</td> <td>99.91%</td> <td>8.82</td> <td>0.8936</td> <td>0.64%</td> <td>99.94%</td> <td>8.46</td> <td>0.8843</td> <td>1.57%</td> </tr> <tr> <td>4</td> <td>100%</td> <td>9.65</td> <td>0.8838</td> <td>1.62%</td> <td>100%</td> <td>8.41</td> <td>0.8809</td> <td>1.91%</td> </tr> <tr> <td>8</td> <td>100%</td> <td>9.99</td> <td>0.8885</td> <td>1.15%</td> <td>100%</td> <td>8.64</td> <td>0.8748</td> <td>2.52%</td> </tr> <tr> <td>16</td> <td>100.00%</td> <td>9.43</td> <td>0.8751</td> <td>2.49%</td> <td>100.00%</td> <td>8.72</td> <td>0.8257</td> <td>7.43%</td> </tr> <tr> <td rowspan="4">4096</td> <td>1</td> <td>99%</td> <td>11.28</td> <td>0.8805</td> <td>1.95%</td> <td>100%</td> <td>10.82</td> <td>0.8629</td> <td>3.71%</td> </tr> <tr> <td>2</td> <td>99.98%</td> <td>11.40</td> <td>0.8805</td> <td>1.95%</td> <td>99.78%</td> <td>10.75</td> <td>0.8380</td> <td>6.20%</td> </tr> <tr> <td>4</td> <td>100.00%</td> <td>12.65</td> <td>0.8832</td> <td>1.68%</td> <td>100.00%</td> <td>11.40</td> <td>0.7862</td> <td>11.38%</td> </tr> <tr> <td>8</td> <td>99.84%</td> <td>14.21</td> <td>0.8761</td> <td>2.39%</td> <td>99.90%</td> <td>12.96</td> <td>0.7788</td> <td>12.12%</td> </tr> <tr> <td>16</td> <td>99.82%</td> <td>46.82</td> <td>0.8791</td> <td>2.09%</td> <td>99.77%</td> <td>53.43</td> <td>0.8098</td> <td>9.02%</td> </tr> </tbody> </table>	TCP:RoCE	Message siz	发端QP数	AI Fabric				传统PFC/ECN				吞吐	时延	比例	比例误差	吞吐	时延	比例	比例误差	1:9	64	1	100%	7.01	0.7670	13.30%	100%	6.75	0.7371	16.29%	2	99.94%	10.60	0.8873	1.27%	100.00%	6.91	0.8185	8.15%	4	99%	33.26	0.8875	1.25%	100%	6.79	0.8113	8.87%	8	99.28%	11.15	0.8795	2.05%	99.37%	7.04	0.8000	10.00%	16	100%	11.27	0.8788	2.12%	100%	6.85	0.7865	11.35%	512	1	100.00%	22.67	0.8964	0.36%	99.99%	7.92	0.8683	3.17%	2	100%	25.50	0.8964	0.36%	100%	7.99	0.8636	3.64%	4	100%	29.53	0.8961	0.39%	100%	7.87	0.8611	3.89%	8	99.97%	34.80	0.8963	0.37%	100.00%	7.86	0.8569	4.31%	16	100%	36.94	0.8964	0.36%	100%	7.92	0.8409	5.91%	1024	1	100%	8.58	0.8935	0.65%	100%	8.32	0.8871	1.29%	2	99.91%	8.82	0.8936	0.64%	99.94%	8.46	0.8843	1.57%	4	100%	9.65	0.8838	1.62%	100%	8.41	0.8809	1.91%	8	100%	9.99	0.8885	1.15%	100%	8.64	0.8748	2.52%	16	100.00%	9.43	0.8751	2.49%	100.00%	8.72	0.8257	7.43%	4096	1	99%	11.28	0.8805	1.95%	100%	10.82	0.8629	3.71%	2	99.98%	11.40	0.8805	1.95%	99.78%	10.75	0.8380	6.20%	4	100.00%	12.65	0.8832	1.68%	100.00%	11.40	0.7862	11.38%	8	99.84%	14.21	0.8761	2.39%	99.90%	12.96	0.7788	12.12%	16	99.82%	46.82	0.8791	2.09%	99.77%	53.43	0.8098	9.02%
TCP:RoCE	Message siz				发端QP数	AI Fabric				传统PFC/ECN																																																																																																																																																																																																			
		吞吐	时延	比例		比例误差	吞吐	时延	比例	比例误差																																																																																																																																																																																																			
1:9	64	1	100%	7.01	0.7670	13.30%	100%	6.75	0.7371	16.29%																																																																																																																																																																																																			
		2	99.94%	10.60	0.8873	1.27%	100.00%	6.91	0.8185	8.15%																																																																																																																																																																																																			
		4	99%	33.26	0.8875	1.25%	100%	6.79	0.8113	8.87%																																																																																																																																																																																																			
		8	99.28%	11.15	0.8795	2.05%	99.37%	7.04	0.8000	10.00%																																																																																																																																																																																																			
	16	100%	11.27	0.8788	2.12%	100%	6.85	0.7865	11.35%																																																																																																																																																																																																				
	512	1	100.00%	22.67	0.8964	0.36%	99.99%	7.92	0.8683	3.17%																																																																																																																																																																																																			
		2	100%	25.50	0.8964	0.36%	100%	7.99	0.8636	3.64%																																																																																																																																																																																																			
		4	100%	29.53	0.8961	0.39%	100%	7.87	0.8611	3.89%																																																																																																																																																																																																			
		8	99.97%	34.80	0.8963	0.37%	100.00%	7.86	0.8569	4.31%																																																																																																																																																																																																			
	16	100%	36.94	0.8964	0.36%	100%	7.92	0.8409	5.91%																																																																																																																																																																																																				
	1024	1	100%	8.58	0.8935	0.65%	100%	8.32	0.8871	1.29%																																																																																																																																																																																																			
		2	99.91%	8.82	0.8936	0.64%	99.94%	8.46	0.8843	1.57%																																																																																																																																																																																																			
		4	100%	9.65	0.8838	1.62%	100%	8.41	0.8809	1.91%																																																																																																																																																																																																			
		8	100%	9.99	0.8885	1.15%	100%	8.64	0.8748	2.52%																																																																																																																																																																																																			
	16	100.00%	9.43	0.8751	2.49%	100.00%	8.72	0.8257	7.43%																																																																																																																																																																																																				
	4096	1	99%	11.28	0.8805	1.95%	100%	10.82	0.8629	3.71%																																																																																																																																																																																																			
2		99.98%	11.40	0.8805	1.95%	99.78%	10.75	0.8380	6.20%																																																																																																																																																																																																				
4		100.00%	12.65	0.8832	1.68%	100.00%	11.40	0.7862	11.38%																																																																																																																																																																																																				
8		99.84%	14.21	0.8761	2.39%	99.90%	12.96	0.7788	12.12%																																																																																																																																																																																																				
16	99.82%	46.82	0.8791	2.09%	99.77%	53.43	0.8098	9.02%																																																																																																																																																																																																					



ODCC 无损网络测试基准

	3.7	64	1	99.75%	8.06	0.7215	2.15%	95.79%	6.70	0.7122	1.22%		
			2	99.70%	13.31	0.7239	2.39%	94.46%	6.65	0.7105	1.05%		
			4	98.73%	12.84	0.7257	2.57%	93.98%	6.73	0.7045	0.45%		
			8	100.00%	13.32	0.7264	2.64%	93.02%	6.66	0.7060	0.60%		
		16	99.86%	14.44	0.7236	2.36%	92.55%	6.70	0.7040	0.40%			
		512	1	100%	32.81	0.6931	0.69%	100%	8.48	0.6852	1.48%		
			2	99.89%	48.18	0.6930	0.70%	100.00%	8.44	0.6848	1.52%		
			4	100.00%	44.47	0.6931	0.69%	100.00%	8.46	0.6812	1.88%		
			8	100%	51.83	0.6922	0.78%	100%	8.92	0.6711	2.89%		
		1024	16	100%	63.84	0.6825	1.75%	100%	8.59	0.6673	3.27%		
			1	100.00%	15.49	0.6858	1.42%	99.93%	9.26	0.6931	0.69%		
			2	99.94%	8.45	0.6854	1.46%	99.48%	9.13	0.6924	0.76%		
			4	99.50%	27.14	0.6769	2.31%	99.19%	9.16	0.6904	0.96%		
		4096	8	100%	8.99	0.6748	2.52%	100%	9.40	0.6845	1.55%		
			16	100.00%	10.15	0.6776	2.24%	99.96%	9.85	0.6573	4.27%		
			1	100%	11.04	0.6855	1.45%	100%	12.51	0.6927	0.73%		
			2	100%	12.13	0.6839	1.61%	100%	12.44	0.6878	1.22%		
					4	100%	24.90	0.6859	1.41%	100%	14.01	0.6856	1.44%
					8	99.66%	15.40	0.6787	2.13%	99.56%	15.28	0.6763	2.37%
					16	100.00%	63.73	0.6792	2.08%	99.76%	65.08	0.6702	2.98%
	5.5	64	1	76.99%	6.63	0.6501	15.01%	89.31%	6.50	0.6932	19.32%		
			2	85.15%	5.96	0.6837	18.37%	85.01%	6.36	0.6789	17.89%		
			4	81.67%	6.02	0.6669	16.69%	84.90%	6.33	0.6852	18.52%		
			8	83.79%	6.48	0.6743	17.43%	84.01%	6.43	0.6746	17.46%		
		16	79.12%	6.30	0.6576	15.76%	84.90%	6.38	0.6835	18.35%			
		512	1	100%	8.27	0.4905	0.95%	100%	9.29	0.4887	1.13%		
			2	99%	8.32	0.4901	0.99%	100%	9.86	0.4870	1.30%		
			4	100.00%	8.12	0.4744	2.56%	99.93%	9.85	0.4816	1.84%		
			8	100%	55.78	0.4861	1.39%	100%	9.65	0.4734	2.66%		
		1024	16	99.94%	30.30	0.4793	2.07%	99.76%	9.10	0.4751	2.49%		
			1	99.95%	9.14	0.4955	0.45%	99.81%	10.11	0.4947	0.53%		
			2	98.86%	9.26	0.4953	0.47%	99.77%	9.97	0.4943	0.57%		
			4	99.24%	9.53	0.4938	0.62%	100.00%	10.91	0.4918	0.82%		
		4096	8	100%	34.78	0.4808	1.92%	100%	10.05	0.4891	1.09%		
			16	100%	24.39	0.4856	1.44%	100%	22.87	0.4815	1.85%		
			1	100%	12.72	0.4948	0.52%	100%	16.86	0.4971	0.29%		
			2	100%	13.28	0.4891	1.09%	100%	17.10	0.4968	0.32%		
					4	99.98%	23.31	0.4758	2.42%	100.00%	18.29	0.4956	0.44%
					8	100%	49.85	0.4821	1.79%	100%	39.30	0.4900	1.00%
					16	100%	100.94	0.4842	1.58%	100%	80.74	0.4894	1.06%
	7.3	64	1	60.59%	5.57	0.5695	26.95%	86.10%	6.37	0.6834	38.34%		
			2	59.97%	5.71	0.5551	25.51%	77.79%	6.20	0.6560	35.60%		
			4	61.49%	5.69	0.5705	27.05%	78.43%	6.31	0.6522	35.22%		
			8	60.59%	5.83	0.5614	26.14%	77.20%	6.21	0.6539	35.39%		
		16	59.54%	5.76	0.5626	26.26%	76.40%	6.37	0.6457	34.57%			
		512	1	99%	9.72	0.2935	0.65%	100%	11.12	0.2928	0.72%		
			2	100%	10.03	0.2930	0.70%	100%	12.04	0.2913	0.87%		
			4	100%	10.08	0.2926	0.74%	100%	12.16	0.2880	1.20%		
			8	99.68%	10.19	0.2913	0.87%	99.87%	12.76	0.2898	1.02%		
		16	100.00%	20.68	0.2851	1.49%	100.00%	21.98	0.2879	1.21%			
		1024	1	100.00%	11.86	0.2968	0.32%	99.84%	17.24	0.2966	0.34%		
			2	99.91%	11.91	0.2966	0.34%	99.99%	17.13	0.2961	0.39%		
			4	99.84%	12.30	0.2959	0.41%	99.43%	17.35	0.2946	0.54%		
			8	99.83%	15.88	0.2949	0.51%	100.00%	21.34	0.2948	0.52%		
		16	99.81%	40.52	0.2908	0.92%	99.83%	40.58	0.2930	0.70%			
		4096	1	99.70%	17.54	0.2985	0.15%	99.89%	24.00	0.2986	0.14%		
			2	99%	17.86	0.2980	0.20%	100%	25.93	0.2981	0.19%		
			4	100.00%	35.24	0.2954	0.46%	100.00%	32.07	0.2973	0.27%		
			8	100%	76.73	0.2964	0.36%	100%	73.71	0.2971	0.29%		
		16	99.92%	158.77	0.2848	1.52%	99.62%	159.68	0.2946	0.54%			
	9.1	64	1	49.50%	5.67	0.4658	36.58%	79.25%	6.20	0.6607	56.07%		
			2	49.76%	5.66	0.4782	37.82%	72.19%	6.12	0.6298	52.98%		
			4	50.36%	5.75	0.4753	37.53%	75.17%	6.24	0.6558	55.58%		
			8	49.24%	5.70	0.4685	36.85%	73.79%	6.28	0.6348	53.48%		
		16	57.42%	5.73	0.5568	45.68%	75.54%	6.29	0.6478	54.78%			
		512	1	99%	22.73	0.0972	0.28%	100%	42.64	0.0973	0.27%		
			2	100.00%	23.57	0.0973	0.27%	99.98%	44.36	0.0972	0.28%		
			4	100.00%	24.72	0.0971	0.29%	99.93%	48.52	0.0970	0.30%		
			8	99.79%	34.70	0.0973	0.27%	100.00%	52.81	0.0973	0.27%		
		16	99.10%	82.31	0.0972	0.28%	100.00%	4369.16	0.0974	0.26%			
		1024	1	98%	31.62	0.0987	0.13%	100%	67.69	0.0987	0.13%		
			2	100%	33.66	0.0987	0.13%	100%	70.42	0.0986	0.14%		
			4	100%	38.16	0.0987	0.13%	100%	76.90	0.0986	0.14%		
			8	99.92%	62.92	0.0987	0.13%	99.89%	87.43	0.0987	0.13%		
		16	100%	463.44	0.0987	0.13%	100%	8781.78	0.0988	0.12%			
		4096	1	100.00%	455.03	0.0995	0.05%	100.00%	2215.76	0.0995	0.05%		
			2	100%	75.17	0.0995	0.05%	100%	4486.23	0.0995	0.05%		
			4	98.57%	130.04	0.0994	0.06%	100.00%	9308.52	0.0995	0.05%		
			8	99.97%	758.02	0.0996	0.04%	99.88%	6761.53	0.0995	0.05%		
		16	100%	940.56	0.0997	0.03%	100%	578.39	0.0997	0.03%			
	备注												

## 4.2 跨POD混跑8打1

测试项目	跨POD混跑8打1
测试目的	测试TCP、RoCE混跑场景下调度的精准性、出接口带宽，同时测试RoCE时延不受TCP影响保持低时延。
测试环境	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置PFC/ECN等特性，所有用例中PFC/ECN配置保持固定。TCP流量映射到队列0，RoCE流量映射到队列3，配置0、3队列调度方式为DRR。</li> </ol>
测试步骤	<ol style="list-style-type: none"> <li>1) 选择9台服务器使用ib_write_bw工具进行8打1带宽测试，同时使用netperf工具进行TCP8打1，其中4个源端和目的端处于一个POD下，另外4个源端和目的端处于不同POD下。发端QP数遍历1、2、4、8、16，DRR调度比重覆盖1:9，3:7，5:5，7:3，9:1，记录带宽利用率和带宽比例。</li> <li>2) 使用步骤1)的流量作为背景流的情况下，对跨POD的2台服务器使用ib_write_lat工具进行时延测试，消息长度与背景流的消息长度保持一致，目的端为同一服务器。记录时延。</li> </ol>
预期结果	<ol style="list-style-type: none"> <li>1) 交换机上RoCEv2流量无丢包，除64B外交换机出接口带宽达到95%以上。吞吐比例控制在5%误差范围内。</li> <li>2) RoCE时延在TCP:RoCE=9:1比例下控制在10ms内。</li> </ol>

ODCC 无损网络测试基准

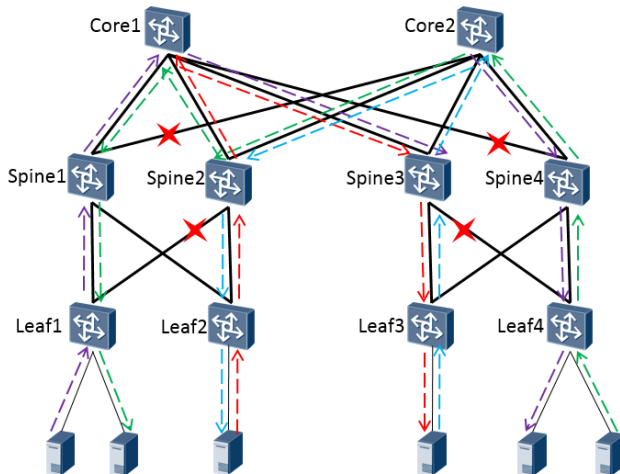
TCP:ROCE	Message size	发端QP数	AI Fabric				传统PFC/ECN					
			吞吐	时延	比例	比例误差	吞吐	时延	比例	比例误差		
测试结果	1:9	64	1	100%	17.36	0.8833	1.67%	100%	6.74	0.8548	4.52%	
			2	100.00%	15.34	0.8868	1.32%	99.90%	6.85	0.8485	5.15%	
			4	100%	25.25	0.8848	1.52%	100%	6.80	0.8389	6.11%	
			8	100.00%	16.32	0.8867	1.33%	99.71%	6.85	0.8212	7.88%	
		16	100%	52.96	0.8870	1.30%	100%	6.90	0.8320	6.80%		
		512	1	99.86%	81.44	0.8963	0.37%	100.00%	7.73	0.8891	1.09%	
			2	100%	18.51	0.8945	0.55%	100%	8.05	0.8870	1.30%	
			4	98%	175.05	0.8963	0.37%	100%	7.99	0.8806	1.94%	
			8	99.74%	144.40	0.8960	0.40%	100.00%	7.94	0.8620	3.80%	
		16	100%	95.91	0.8966	0.34%	100%	7.95	0.8180	8.20%		
		1024	1	100%	8.01	0.8933	0.67%	100%	8.48	0.8943	0.57%	
			2	100.00%	8.90	0.8931	0.69%	100.00%	8.54	0.8915	0.85%	
	4		100%	24.18	0.8900	1.00%	100%	8.60	0.8881	1.19%		
	8		100%	84.88	0.8912	0.88%	100%	8.81	0.8573	4.27%		
	16	100.00%	23.69	0.8870	1.30%	100.00%	25.84	0.8862	1.38%			
	4096	1	100%	11.68	0.8877	1.23%	100%	13.09	0.8965	0.35%		
		2	99.81%	10.60	0.8727	2.73%	100.00%	13.39	0.8953	0.47%		
		4	100.00%	12.09	0.8795	2.05%	100.00%	13.53	0.8909	0.91%		
		8	100.00%	48.72	0.8875	1.25%	99.97%	48.31	0.8470	5.30%		
	16	99.53%	99.57	0.8811	1.89%	100.00%	101.35	0.8615	3.85%			
	3:7	64	64	1	99.64%	16.54	0.6720	2.80%	100.00%	7.27	0.6596	4.04%
				2	99.99%	24.59	0.6719	2.81%	99.52%	7.32	0.6529	4.71%
				4	100.00%	28.05	0.6719	2.81%	100.00%	7.33	0.6351	6.49%
				8	99.96%	24.62	0.6721	2.79%	99.90%	7.45	0.6496	5.04%
			16	99.94%	28.53	0.6721	2.79%	100.00%	7.52	0.6336	6.64%	
			512	1	100%	66.43	0.6930	0.70%	100%	8.93	0.6903	0.97%
				2	99.99%	85.38	0.6929	0.71%	100.00%	9.12	0.6873	1.27%
				4	99.48%	9.54	0.6845	1.55%	99.99%	9.04	0.6801	1.99%
				8	100%	22.42	0.6739	2.61%	100%	9.01	0.6824	1.76%
			16	100%	41.97	0.6743	2.57%	100%	17.70	0.6611	3.89%	
			1024	1	99.92%	8.80	0.6956	0.44%	100.00%	9.77	0.6951	0.49%
				2	100.00%	8.87	0.6926	0.74%	100.00%	10.14	0.6938	0.62%
		4		99.97%	8.84	0.6905	0.95%	99.87%	10.60	0.6916	0.84%	
		8		100%	10.09	0.6853	1.47%	100%	11.08	0.6888	1.12%	
		16	99.97%	30.46	0.6875	1.25%	99.77%	34.60	0.6893	1.07%		
		4096	1	100%	21.74	0.6890	1.10%	100%	15.09	0.6972	0.28%	
			2	100%	13.58	0.6933	0.67%	100%	15.58	0.6962	0.38%	
			4	100%	14.99	0.6810	1.90%	100%	16.49	0.6935	0.65%	
			8	99.99%	58.41	0.6790	2.10%	100.00%	64.24	0.6882	1.18%	
		16	99.78%	129.57	0.6795	2.05%	99.99%	133.58	0.6857	1.43%		
		64	64	1	99.78%	31.88	0.4678	3.22%	99.42%	8.11	0.4556	4.44%
				2	100.00%	44.25	0.4678	3.22%	99.73%	8.16	0.4406	5.94%
				4	98.98%	53.06	0.4678	3.22%	100.00%	7.93	0.4259	7.41%
				8	99.67%	47.81	0.4678	3.22%	100.00%	8.29	0.4360	6.40%
			16	100.00%	51.01	0.4680	3.20%	100.00%	8.60	0.4383	6.17%	
			512	1	100%	8.54	0.4916	0.84%	100%	10.07	0.4901	0.99%
				2	99%	8.88	0.4912	0.88%	100%	10.01	0.4873	1.27%
				4	99.97%	8.85	0.4873	1.27%	100.00%	10.30	0.4823	1.77%
8				100%	9.60	0.4868	1.32%	100%	10.36	0.4887	1.13%	
16			99.56%	28.07	0.4876	1.24%	99.86%	25.63	0.4896	1.04%		
1024			1	99.13%	9.93	0.4958	0.42%	99.91%	13.74	0.4953	0.47%	
			2	99.70%	10.33	0.4952	0.48%	100.00%	13.73	0.4942	0.58%	
		4	99.54%	10.29	0.4932	0.68%	99.95%	12.64	0.4920	0.80%		
		8	100%	23.14	0.4890	1.10%	100%	27.32	0.4933	0.67%		
16		100%	47.27	0.4862	1.38%	100%	47.39	0.4937	0.63%			
4096		1	100%	13.65	0.4963	0.37%	100%	20.41	0.4978	0.22%		
		2	100%	14.00	0.4958	0.42%	99%	20.39	0.4970	0.30%		
		4	100.00%	39.86	0.4864	1.36%	99.93%	47.25	0.4937	0.63%		
		8	99%	65.57	0.4833	1.67%	100%	86.63	0.4963	0.37%		
16		100%	192.72	0.4825	1.75%	100%	194.18	0.4905	0.95%			

ODCC 无损网络测试基准

	7:3	64	1	89.88%	6.95	0.3288	2.88%	89.97%	7.85	0.3266	2.66%		
			2	88.86%	7.46	0.3341	3.41%	89.84%	7.80	0.3309	3.09%		
			4	88.93%	9.16	0.3214	2.14%	88.68%	7.90	0.3300	3.00%		
			8	88.59%	7.89	0.3317	3.17%	89.31%	7.81	0.3330	3.30%		
			16	89.59%	9.53	0.3329	3.29%	89.55%	8.13	0.3409	4.09%		
		512	1	100%	11.10	0.2934	0.66%	100%	15.54	0.2929	0.71%		
			2	100%	11.46	0.2929	0.71%	100%	15.21	0.2911	0.89%		
			4	100%	11.21	0.2932	0.68%	100%	16.27	0.2927	0.73%		
			8	99.97%	23.63	0.2923	0.77%	100.00%	24.72	0.2920	0.80%		
			16	99.98%	43.43	0.2919	0.81%	100.00%	48.36	0.2916	0.84%		
		1024	1	99.97%	13.32	0.2968	0.32%	99.96%	20.42	0.2967	0.33%		
			2	100.00%	13.75	0.2963	0.37%	99.95%	22.75	0.2955	0.45%		
			4	99.60%	13.97	0.2967	0.33%	99.90%	23.40	0.2965	0.35%		
			8	100.00%	41.75	0.2960	0.40%	100.00%	41.65	0.2956	0.44%		
			16	99.41%	76.86	0.2944	0.56%	100.00%	4362.05	0.2971	0.29%		
		4096	1	99.96%	20.81	0.2985	0.15%	99.64%	35.36	0.2985	0.15%		
			2	99%	44.27	0.2983	0.17%	100%	41.25	0.2984	0.16%		
			4	99.86%	79.12	0.2987	0.13%	99.93%	78.59	0.2985	0.15%		
			8	100%	182.65	0.2948	0.52%	100%	5914.01	0.2988	0.12%		
			16	99.31%	294.72	0.2815	1.85%	100.00%	355.44	0.2956	0.44%		
		64	1	71.66%	7.22	0.1876	8.76%	75.26%	7.99	0.2400	14.00%		
			2	71.74%	7.42	0.1963	9.63%	76.73%	8.09	0.2339	13.39%		
			4	73.09%	7.44	0.1924	9.24%	76.38%	8.10	0.2295	12.95%		
			8	71.08%	7.24	0.2000	10.00%	77.96%	8.02	0.2473	14.73%		
			16	80.50%	25.11	0.2673	16.73%	81.37%	8.73	0.2775	17.75%		
		512	1	100%	31.14	0.0972	0.28%	100%	67.88	0.0973	0.27%		
			2	100.00%	32.97	0.0973	0.27%	100.00%	76.18	0.0973	0.27%		
			4	100.00%	55.95	0.0973	0.27%	100.00%	87.67	0.0973	0.27%		
			8	99.89%	1386.98	0.0974	0.26%	100.00%	4986.55	0.0973	0.27%		
			16	100.00%	1350.96	0.0974	0.26%	99.96%	11328.60	0.0974	0.26%		
		1024	1	100%	47.66	0.0987	0.13%	100%	121.36	0.0987	0.13%		
			2	100%	54.21	0.0986	0.14%	100%	135.53	0.0987	0.13%		
			4	100%	67.29	0.0987	0.13%	100%	1940.71	0.0987	0.13%		
			8	99.84%	1361.54	0.0987	0.13%	99.99%	9511.42	0.0987	0.13%		
			16	100%	1555.35	0.0988	0.12%	100%	22251.30	0.0988	0.12%		
		4096	1	99.93%	1354.23	0.0995	0.05%	100.00%	5183.19	0.0995	0.05%		
			2	100%	1647.44	0.0995	0.05%	100%	10859.53	0.0995	0.05%		
			4	100.00%	1827.03	0.0995	0.05%	99.41%	23474.84	0.0996	0.04%		
			8	100.00%	1804.94	0.0995	0.05%	100.00%	36574.85	0.0996	0.04%		
			16	100%	2588.16	0.0995	0.05%	100%	129066.56	0.0996	0.04%		
		备注											

4.3 跨 POD 的 PFC 死锁预防

测试项目	跨POD的PFC死锁预防
测试目的	验证设备在跨POD流量下支持预防PFC死锁功能。

<p>测试环境</p>	<p>测试组网：</p>  <p>前置条件：</p> <ol style="list-style-type: none"> <li>1) 根据以上拓扑设备正常工作。</li> <li>2) 网卡配置PFC、DCQCN。</li> <li>3) 交换机上配置AI Fabric特性。其中RoCE映射到3队列，开启3、4队列PFC功能，并配置3、4队列DRR调度，调度比例为5:5。</li> <li>4) 交换机上不配置PFC死锁检测功能。</li> </ol>
<p>测试步骤</p>	<ol style="list-style-type: none"> <li>1) 为了触发更多PFC复现Deadlock，关闭AI Fabric，关闭DCQCN；</li> <li>2) 按照以上拓扑shutdown4条链路。</li> <li>3) 如图所示打四条dscp为24的流(TCP或ROCEv2)，造成队列4的CBD，触发Deadlock。 (调整流量路径的方法：配置静态主机路由，或调整传输层port号)</li> <li>4) 停止所有流量，观察PFC Deadlock仍然持续。</li> <li>5) 关闭CBD中任一port的PFC，从而切断CBD，解除Deadlock。</li> <li>6) 设置死锁预防配置，重新打上上述流量，Deadlock不再发生。</li> <li>7) 观察PFC的变化，观察各端口流量cos队列的变化。</li> </ol>
<p>预期结果</p>	<ol style="list-style-type: none"> <li>1) 复现PFC Deadlock。</li> <li>2) 设置死锁预防配置后，不能复现PFC Deadlock。</li> </ol>
<p>测试结果</p>	<pre>[root@h7 ~]# ./send_traffic.sh -c h10,h7 -n 1 -s 4096 -l 120000 &amp; [1] 336614 [root@h7 ~]# ssh h10 Last login: Mon Jul 22 19:57:08 2019 from 192.89.52.195 [root@h10 ~]# ./send_traffic.sh -c h10,h7 -n 1 -s 4096 -l 120000 &amp; [1] 4445 [root@h10 ~]# exit logout Connection to h10 closed. [root@h7 ~]# ssh h4 Last login: Mon Jul 22 19:37:08 2019 from 192.89.52.195 [root@h4 ~]# ./send_traffic.sh -c h3,h4 -n 1 -s 4096 -l 120000 &amp; [1] 10114 [root@h4 ~]# ssh h3 Last login: Mon Jul 22 19:38:50 2019 from 192.89.52.195 [root@h3 ~]# ./send_traffic.sh -c h3,h4 -n 1 -s 4096 -l 120000 &amp; [1] 22155</pre>

```

<Spine1>dis dcb pfc
-----
Interface      Queue      Received(Frames)      ReceivedRate(pps)      DeadlockNum
              Transmitted(Frames)      TransmittedRate(pps)      RecoveryNum
-----
100GE1/0/1      3           0                      0                      0
              4           0                      0                      0
100GE1/0/2      3           2391851                0                      0
              4           0                      0                      0
100GE1/0/3      3           0                      0                      0
              4           0                      0                      0
100GE1/0/4      3           0                      0                      0
              4           0                      0                      0
100GE1/0/5      3           2392035                0                      0
              4           0                      0                      0
100GE1/0/6      3           2392022                0                      0
              4           0                      0                      0
-----
<Spine1>dis dcb pfc
-----
Interface      Queue      Received(Frames)      ReceivedRate(pps)      DeadlockNum
              Transmitted(Frames)      TransmittedRate(pps)      RecoveryNum
-----
100GE1/0/1      3           0                      0                      0
              4           0                      0                      0
100GE1/0/2      3           2490660                41511                 0
              4           0                      0                      0
100GE1/0/3      3           0                      0                      0
              4           0                      0                      0
100GE1/0/4      3           0                      0                      0
              4           0                      0                      0
100GE1/0/5      3           2490841                41514                 0
              4           0                      0                      0
100GE1/0/6      3           2490827                41513                 0
              4           0                      0                      0
-----

```

```
<Spine1> dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDAP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface      PHY      Protocol  InUti  OutUti   inErrors  outErrors
100GE1/0/2     up       up        0%    0.05%    0         0
100GE1/0/5     up       up        0.05% 0%      0         0
100GE1/0/6     up       up        0.01% 0.05%    0         0
MEth0/0/0      up       up        0.01% 0.01%    0         0
NULL0          up       up(s)     0%    0%      0         0
```

Dead lock free使能后

```
<Spine1>dis int b | i up
PHY: Physical
*down: administratively down
^down: standby
(l): loopback
(s): spoofing
(b): BFD down
(e): ETHOAM down
(d): Dampening Suppressed
(p): port alarm down
(dl): DLDAP down
(c): CFM down
(sd): STP instance discarding
InUti/OutUti: input utility rate/output utility rate
Interface      PHY      Protocol  InUti  OutUti   inErrors  outErrors
100GE1/0/2     up       up        24.99% 24.99%   0         0
100GE1/0/5     up       up        0.01% 46.69%   0         0
100GE1/0/6     up       up        46.69% 0%       0         0
MEth0/0/0      up       up        0.01% 0.01%    0         0
NULL0          up       up(s)     0%    0%      0         0
```

```
<Spine1> dis q q s int 100g 1/0/5
Queue  CIR/PIR      Passed      Pass Rate      Dropped      Drop Rate      Drop Time
(% or kbps) (Packets/Bytes) (pps/bps) (Packets/Bytes) (pps/bps)
-----
0      0           0           0             0             0             -
100000000 0           0           0             0             0
-----
1      0           0           0             0             0             -
100000000 0           0           0             0             0
-----
2      0           0           0             0             0             -
100000000 0           0           0             0             0
-----
3      0           187433408   389227        0             0             -
100000000 272152796594 4690020528 0             0
-----
4      0           145012255   190922        0             0             -
100000000 203880070252 1908760840 0             0
-----
5      0           0           0             0             0             -
100000000 0           0           0             0             0
-----
6      0           25          0             0             0             -
100000000 3728        160          0             0             0
-----
7      0           0           0             0             0             -
100000000 0           0           0             0             0
-----
```

	<pre>&lt;Spine1&gt;  dis dcb pfc ----- Interface      Queue      Received(Frames)  ReceivedRate(pps)  DeadlockNum               Transmitted(Frames)  TransmittedRate(pps)  RecoveryNum -----</pre>				
	<pre>100GE1/0/1      3           0                0                0                 4           0                0                0 100GE1/0/2      3           0                0                0                 4           0                0                0 100GE1/0/3      3           0                0                0                 4           0                0                0 100GE1/0/4      3           0                0                0                 4           0                0                0 100GE1/0/5      3           0                0                0                 4           0                0                0 100GE1/0/6      3           0                0                0                 4           0                0                0 -----</pre>				
备注					