

スパコン大規模通信のための QoSモニター試験

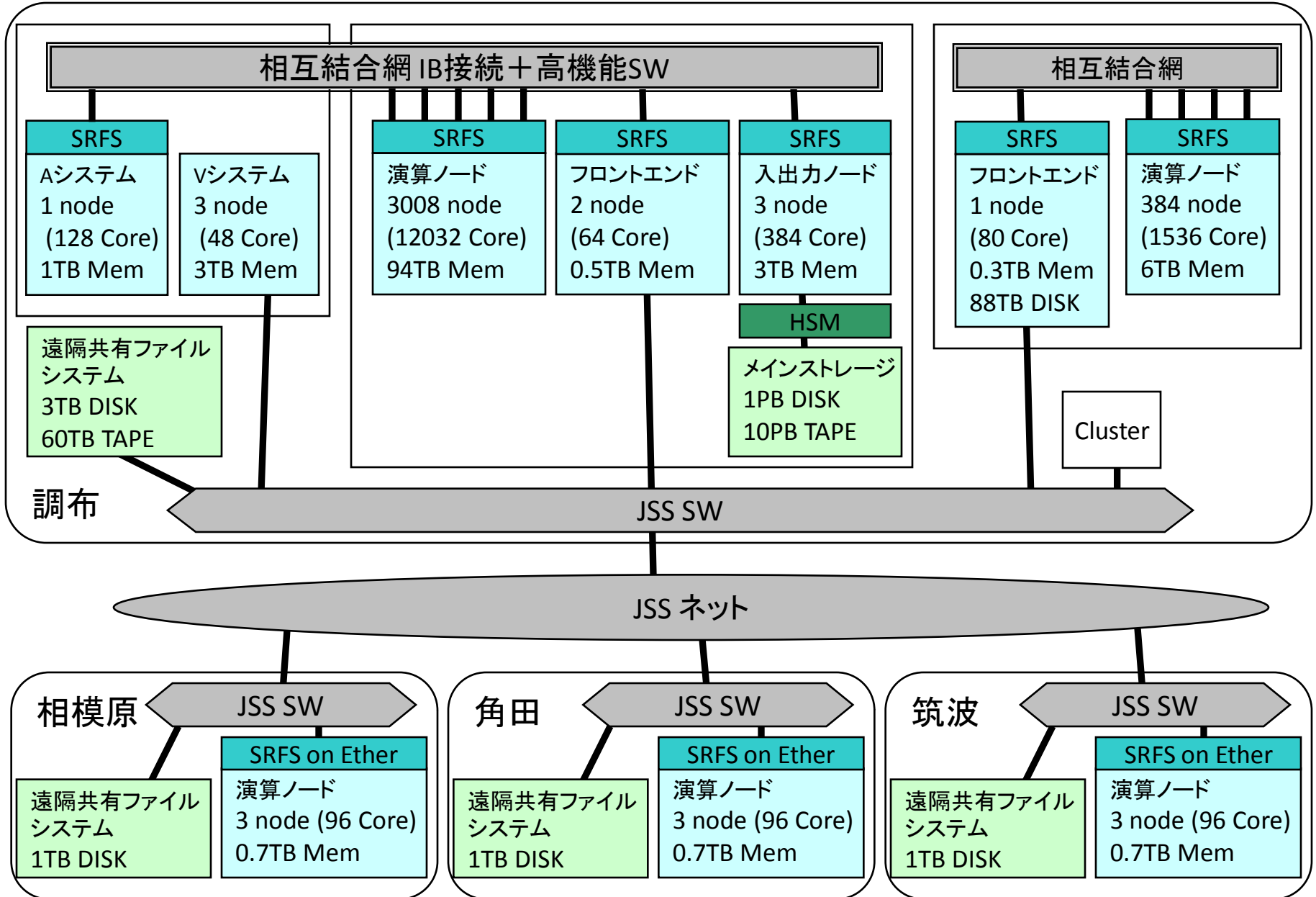
宇宙航空研究開発機構
研究開発本部 数値解析G

大川博文(ookawa@chofu.jaxa.jp)

目的

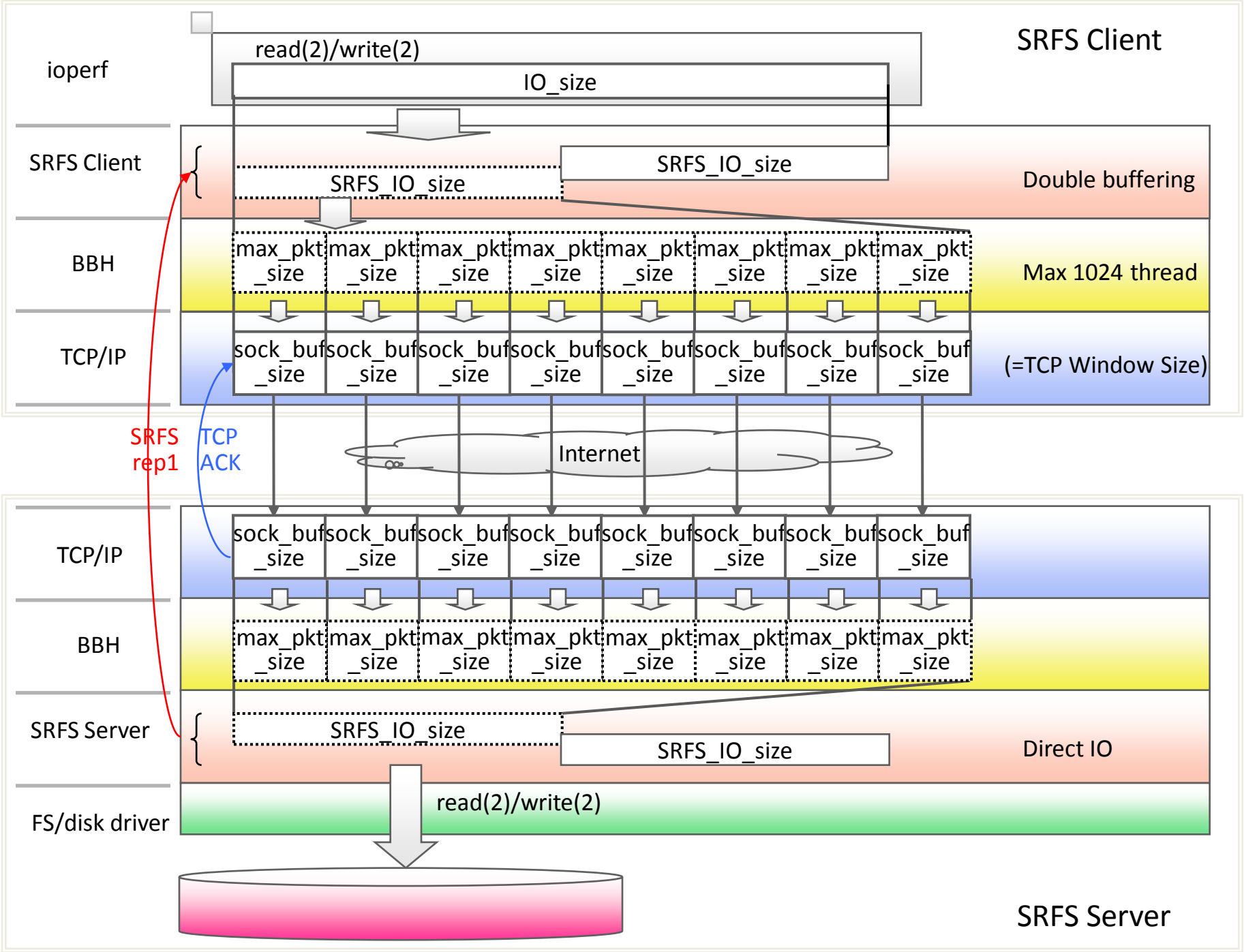
- 平成21年4月より、JAXAスーパーコンピュータシステム(JSS)が本格稼働を開始する。
- JSSでは、JAXA主要4事業所をSINET3経由で市販のVPNルータを用いてVPN接続した、JSSネットでシステム間の通信を行う。
- そのため今後、大規模なデータ通信が発生することが予測される。
- 今回のQoSサービスモニター試験では、調布-相模原間に新規に、SINET3が提供するL2VPN網を構築、SRFS on Etherを用いたIO性能の測定を実施し、QoSサービスの有効性を確認する。

JSS 概要



SRFS on Ether 概要

- JAXAの前スパコンシステム NS III において、ノード間分散ファイルシステムとして提供されていたSRFSを、Eternet経由で利用できるようにJAXAが追加したインタフェース。
- JSSでは、富士通(株)の製品として、同様の機能を有するSRFS on Etherが提供されている。(今回の試験ではNS IIIのものを利用した)
- NFSと同様の使い方ができるリモートファイルシステムであり、並列転送による高速性、データ整合性の確保、独自の権限管理など、NFSの問題点の解決をしている。
- ファイルIOが大きく、シーケンシャルな入出力で特に性能を発揮する。

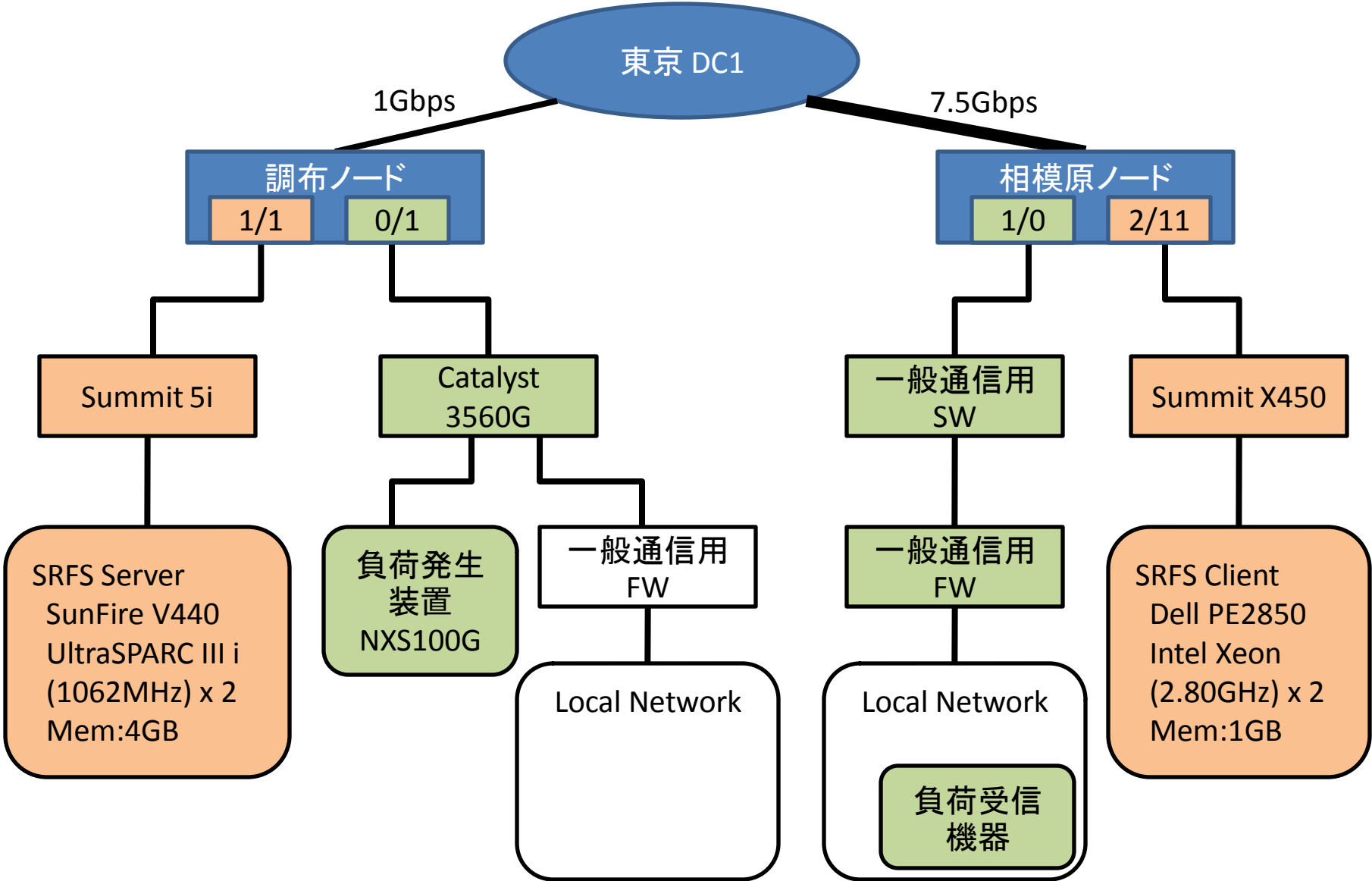


QoSサービス試験 概要

UDPによる負荷を既存回線側にかけた状態で、SRFS on Etherを用いたIO性能測定を実施する。

1. テスト環境の構築
2. 事前測定 (RTTの測定、パラメータ決定)
3. 対象試験 (QoS設定なしでの測定)
4. 本試験 (L2VPN網にAF-30%のQoS設定)
5. 考察

テスト環境



事前測定

- 調布-相模原間(L2VPN)遅延測定

pingを双方向から10回ずつ、間隔をあけて適時実行し、5セットの平均時間を取得

平均RTT = 8.0 [msec]

- SRFS on Etherパラメータ探索

SRFS_IO=(4,8,16,32) [MB]

並列数(SRFS_IO/max_pkt_size)=(4,8,16,32)

sock_buf_size(TCP_window_size)=128 [KB] (固定)

Read	[MB/sec]			
	4	8	16	32
4MB	44.860	47.077	50.498	18.409
8MB	61.595	65.895	65.738	67.259
16MB	74.511	79.952	80.563	79.933
32MB	79.025	86.320	87.545	67.881

Write	[MB/sec]			
	4	8	16	32
4MB	23.257	15.392	5.810	2.373
8MB	29.669	26.021	5.673	2.520
16MB	34.186	31.567	10.324	3.786
32MB	37.589	41.671	18.833	6.280

対象試験および本試験

SRFS on Ether パラメータ

SRFS_IO = 32 [MB]

max_pkt_size = 4 [MB] (並列数 = 8)

sock_buf_size = 128 [KB]

実行アプリケーション

ioperf (シーケンシャルリード)

IO長 = 64 [MB]

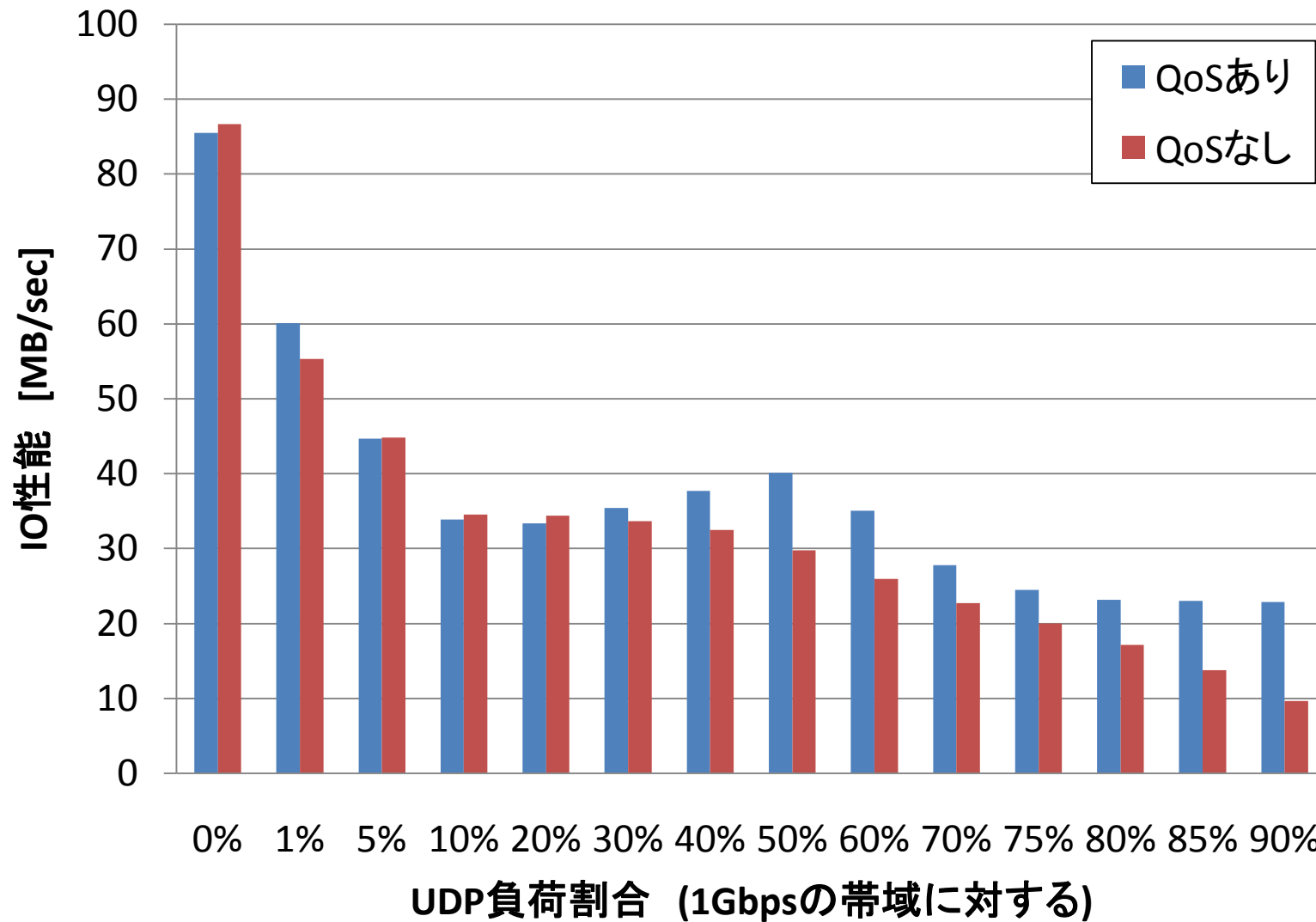
ファイルサイズ = 1 [GB]

帯域1Gbpsに対するUDP負荷

送信フレームサイズ = 1280 [B]

負荷割合 = (0~90) [%]

結果



結果の考察

現在、結果の考察にまで至っていない。

- QoS設定有りでは、75%以上の負荷でもIO性能が確保されている。
- ? QoS設定の有無によらず、少量のUDP負荷でIO性能が急激に低下している。
 - SRFS on Etherの通信性能は、TCP通信の再送に敏感であるが、低下が極端すぎる
 - ピーク性能を求めるパラメータ以外で結果が向上するか
- ? QoS設定無しの場合、10~40%の負荷でIO性能にほぼ変化がないが、それ以降では直線的にIO性能が低下していく。
- ? QoS設定有りの場合、30~50%の間で負荷の増加とは逆にIO性能が向上している。
 - 負荷発生装置のUDP負荷の発生方法に問題はないか

今後の課題

- 今回の試験では、1Gbpsの帯域を瞬間的には完全に利用できる、ピーク性能を求めるパラメータを設定したが、実運用ではJSSの他の通信とのバランスを考慮した設定で再検討する必要がある。
- 今回は負荷としてUDP通信を利用したが、実際にはUDP通信が大量に発生することはほとんど無い。TCP通信での定量的な負荷の発生方法について考慮が必要である。
- L2VPN網をJSSネットの幹線として利用する場合、一般線側に今回のような大規模通信が発生することはあまり考えられない。QoS設定はJSSネット内の特定ホスト間に付与し、JSSネット内で負荷を与えるテスト環境を検討する必要がある。