

高レイノルズ数壁面乱流場の 大規模直接数値シミュレーション

山本 義暢
山梨大学大学院総合研究部

yamamotoy@yamanashi.ac.jp

本日の発表内容

▶ 名古屋大学の新システム：
FX1000(富岳と同一) 2048ノード
を用いて、世界最大Re条件下で壁
面乱流直接数値計算コードのベン
チマークを実施

▶ 発表項目：

- 1) 壁面乱流場（チャンネル乱流場）
とは
- 2) 直接数値計算の概要
- 3) 直接数値計算コードの概要
- 4) ベンチマーク条件（世界最大Re条
件）
- 5) ベンチマーク結果と考察
- 6) 今後の課題
- 7) その他

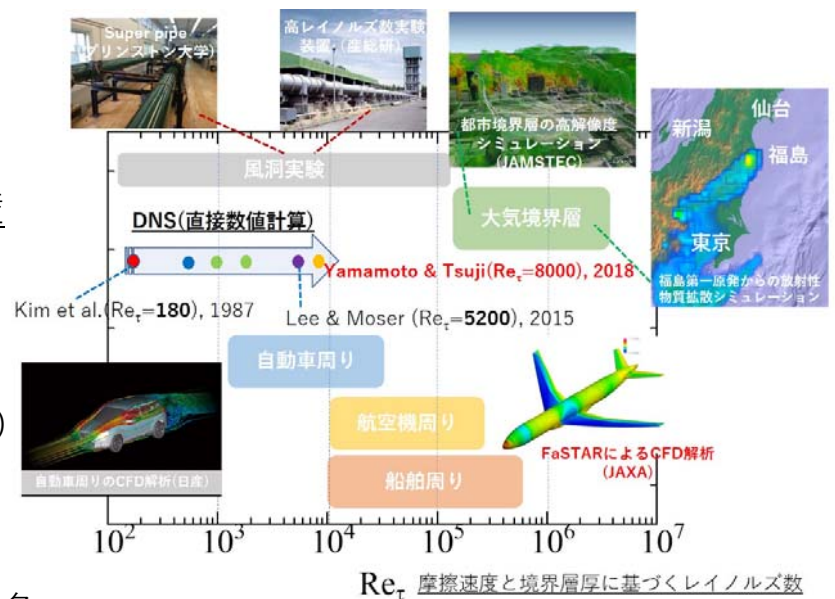


FIG. 工学現象におけるRe規模

DNSにおいては、
対象とするReを上げることが課題

1)チャンネル流(平行平板間流)

2枚の平板間の流れ：
壁面に沿った流れ（壁乱流）のカノニカル流
乱流モデル評価等で必ず対象とされる流動場

名古屋大学乱流風洞
(辻研究室)

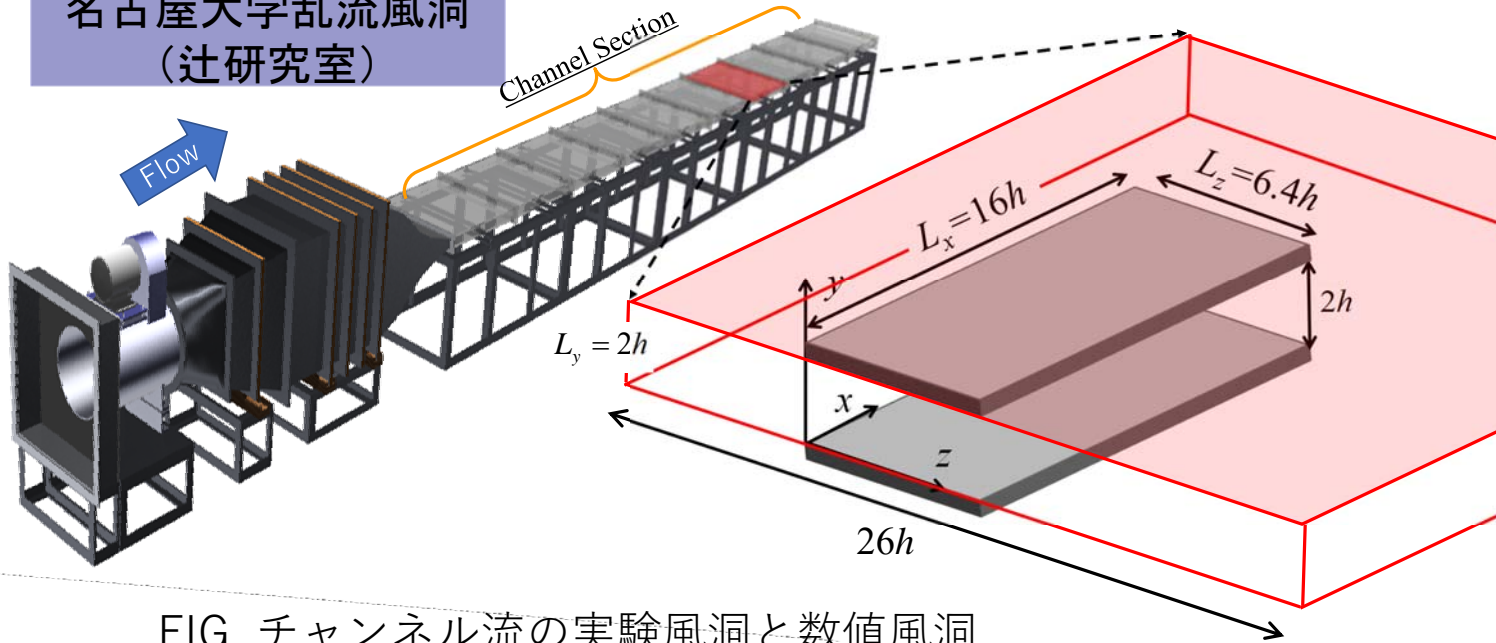


FIG. チャンネル流の実験風洞と数値風洞

2)直接数値計算の概要

最大渦を包括する計算領域
最小渦を解像する格子分解能
→数百万ステップに及ぶ長時間積分

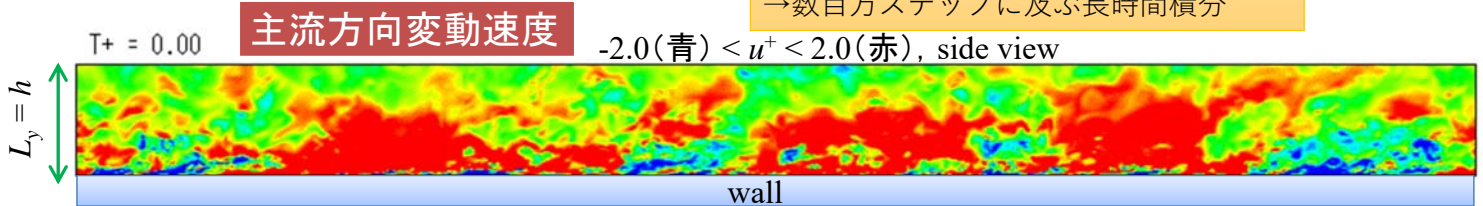


FIG. 時間積分： $T^+ = 680$ 間の流れ場の可視化 ($Re_\tau = 1000$): $T^+/Re_\tau = 0.68$

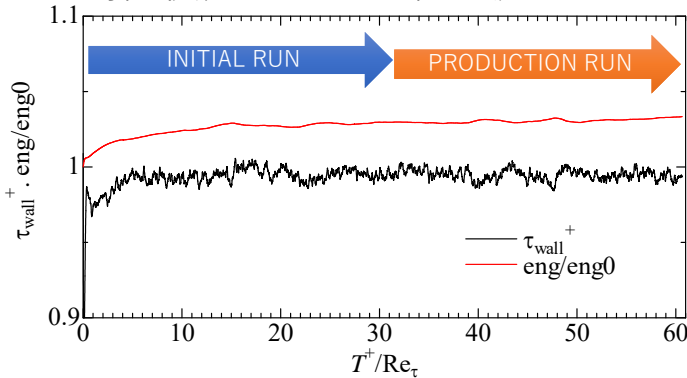


FIG. せん断応力・運動エネルギーの時間変化

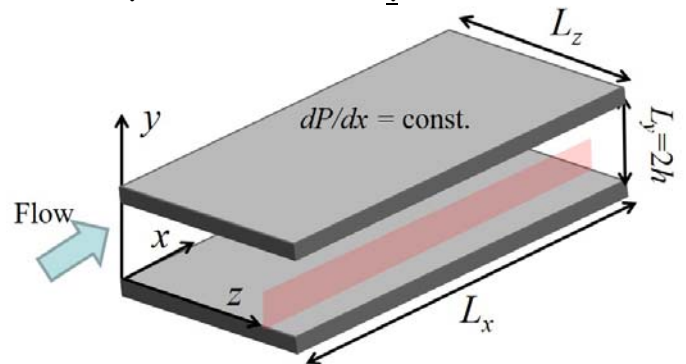


FIG. チャンネル流

INITIAL RUN: 助走計算 PRODUCTION RUN: 本計算(長時間平均)

3)DNSコードの概要 I(all-to-all版) *

- 基礎方程式
非圧縮性のNavier-Stokes式及び連続式
- 離散化手法
(x, z):フーリエスペクトル法
→ 2D-FFT
 y :2次精度中心差分法
→差分演算+3重対角行列解法(TDMA)
- 領域分割
 - ・1次元領域分割($x, z, y/nproc$)
 - ・2次元領域分割($x, z/nproc_z, y/nproc_y$)
- 通信タイプ
 - ・all-to-all (TDMA)
 - ・隣接シフト通信(差分テンソル)

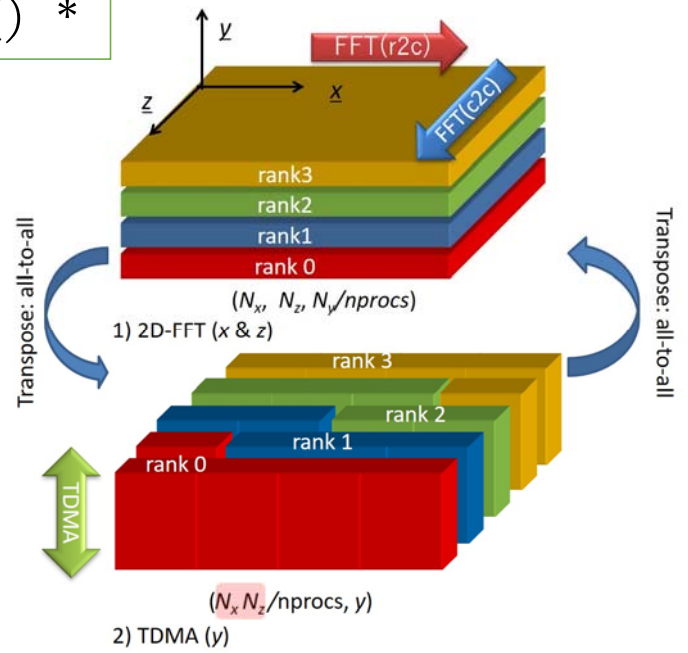


FIG. 1次元領域分割 (4並列)
チャンネル流では、 $N_x \neq N_y \neq N_z$
→TDMA時は(x, z)方向を1次元化

*京を対象として開発

3)DNSコードの概要 II(shift版)**

空間離散化方法及び領域分割(y)は同じ

格子数: $N_{x_small} * N_z * N_y$
TDMA(y 方向格子当たりの)演算量: M
全演算量: $M * N_{x_small} * N_z * N_y$
MPI並列数: $nprocs$

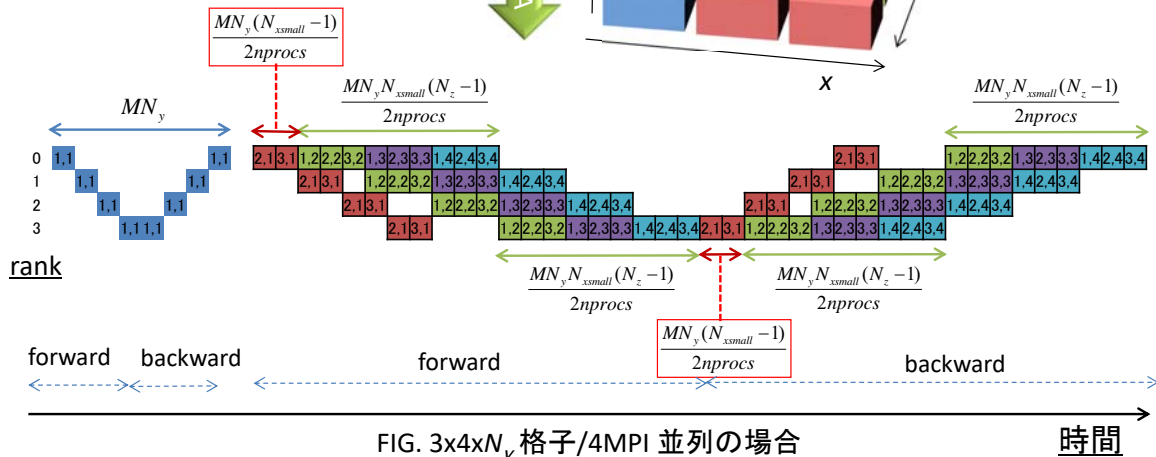


FIG. $3 \times 4 \times N_y$ 格子/4MPI 並列の場合

$$T_c[s] = \left\{ MN_y + \frac{2MN_y N_{x_small} (N_z - 1)}{nprocs} + \frac{MN_y (N_{x_small} - 1)}{nprocs} \right\} / V_c$$

TDMA時のall-to-all通信が発生せず、
隣接シフト通信のみで計算可能

**SX-ACEを対象として開発

4) ベンチマーク条件

レイノルズ数	計算領域		格子点数 (解像度)			メモリ量	演算量
Re_τ	L_x /h	L_z /h	N_x (Δx^+)	N_y (Δy^+)	N_z (Δz^+)	Mem [TB]	Tflop /step
8000	16.0	6.4	6912 (18.5)	4096 (0.6-8.0)	5760 (8.9)	54	218

・計測システム

理論性能 (倍精度)

	system	#node	Mem	Tflop/s
Nagoya U.	FX1000	2048	56	6921
NIFS	FX100	2048	56	2068
Tohoku U.	SX-ACE	1024	60	262

■ all-to-all通信

FXシステム: mpi_alltoallv
SX-ACE: mpi_put + スケ
ジューリング

■ shift通信(TDMA):

mpi_sendrecv
により実装

・FFTライブラリ

FX1000/FX100 FFTW(OpenMP版)

SX-ACE ASL (自動並列版)

5) ベンチマーク結果

・計測結果 表スペクトル法:6912x4096x5760 格子

	system	ver.	#node	s/step	Tflop/s*	efficiency
Nagoya U.	FX1000	shift	2048	2.21	101.0	1.5%
NIFS	FX100**	alltoall	2048	3.18	68.3	3.3%
Tohoku U.	SX-ACE	shift	1024	9.51	22.8	8.7%
	SX-ACE	alltoall	1024	13.57	16.0	6.1%

*IO 部分を含まない **tofu2 :4レーン使用時

・当初の期待

➤ SX-ACE(1B/F)におけるFFT部分の実行効率 は理論性能の16%程度
→FX1000(0.3B/F) 理論性能の5%程度、通信比率を40%とすると、3%ぐ
らいの実行効率 (実測:1.46%)

➤ FX1000とFX100の理論性の比は3.35倍

→本コードの実行速度比は1.44倍

6)今後の課題(1)

- ・原因と対策：単体性能
 - そもそも単体性能があまり良くない

表シングルノード性能：384x384x384 格子

	system	ver.	#node	#MPI	#AP	Gflop/s	efficiency
Nagoya U.	FX1000	shift	1	48	1	140.5	4.2%
	CX2570	shift	1	32	1	78.9	3.0%
Tohoku U.	SX-ACE	shift	1	4	1	38.7	15.1%
Kyoto U.	XC40	shift	1	64	1	80.2	2.6%
	CS400	shift	1	32	1	32.9	2.7%

メモリバンド幅から考えれば、単体性能で理論性能の5-6%程度は出るのでは？
→コンパイルオプション等の検討; zfill

- ノード内のOpenMP部分の効率が悪い
→原因は調査中 fjomplib に変更？
- ノード内コア数(48)に対し、格子数のバランスが悪い、FX100 (32コア)

6)今後の課題(2)

- ・原因と対策（ノード間）
 - ノード配置の最適化
 - ・tofu構成単位（12ノード）
 - ・隣接シフト通信用にノード配置を指定（MAP使用）
→今回未対応
 - 通信ソルバの性能評価
 - ・通信部分を切り出し、実効通信速度を評価（場合によっては置き換え）
 - 2次元領域分割版の性能評価と最適化（富岳への接続）

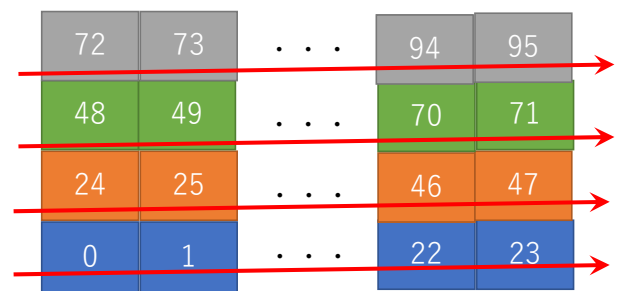


FIG. 通常のノード配置

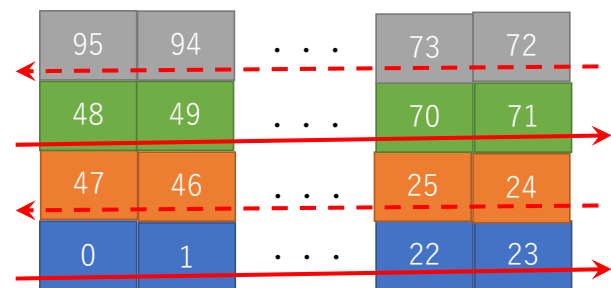
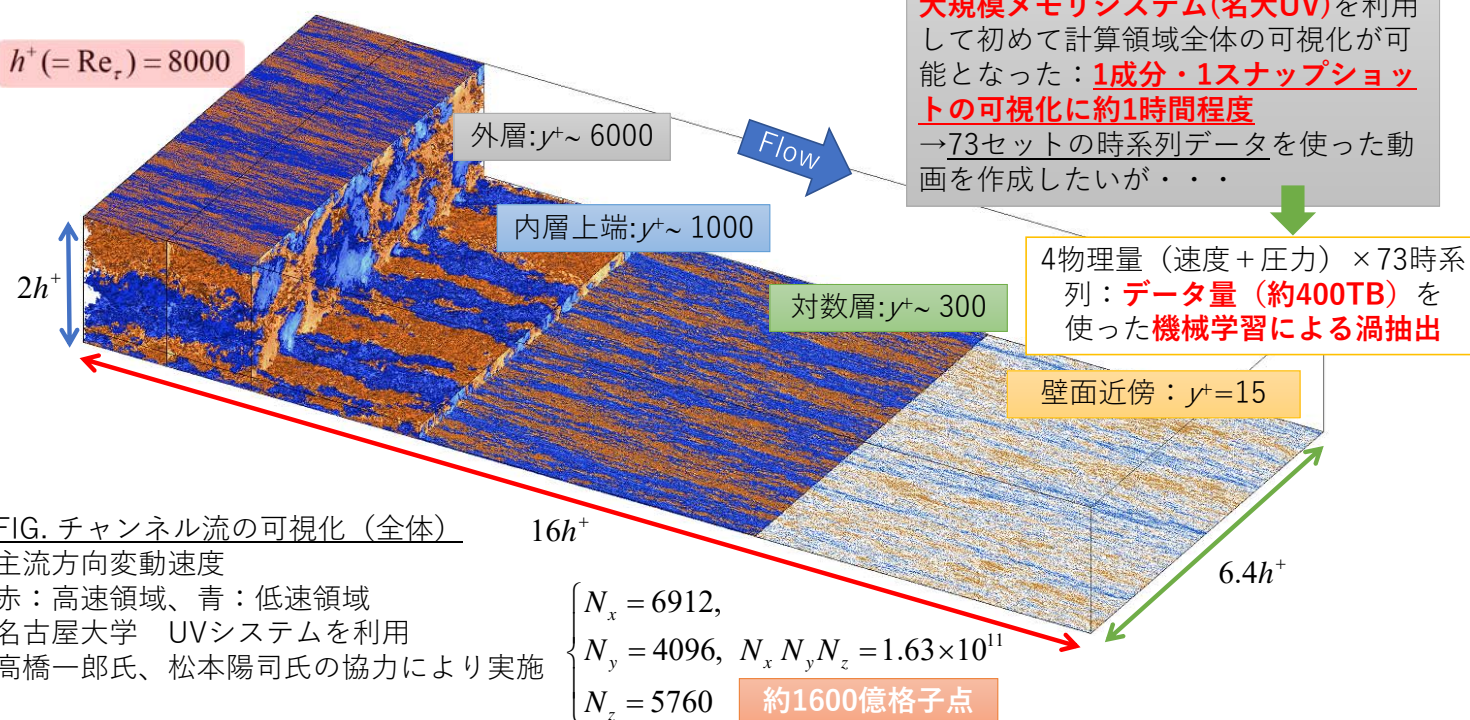


FIG. MAPを使い一筆書きに設定したノード配置：隣接シフト通信が効率的

7) その他：ポスト処理（可視化・統計解析）



まとめ

- 名古屋大学の新システム：FX1000(富岳と同一) **2048ノード**を用いて、世界最大Re条件下で壁面乱流直接数値計算コードのベンチマークを実施
- 単純なコード移植により約**100TFlop/s**の実行演算速度を確認
- 今後は、理論性能の3%程度の実行効率を目標としてコードの最適化を行う予定
- 2048ノードクラスの計算は富岳との接続において有用、またポスト処理においても多様なシステムを活用したい

本研究は、HPCIシステム利用研究課題(課題番号：hp200107)を通じて実施しました。

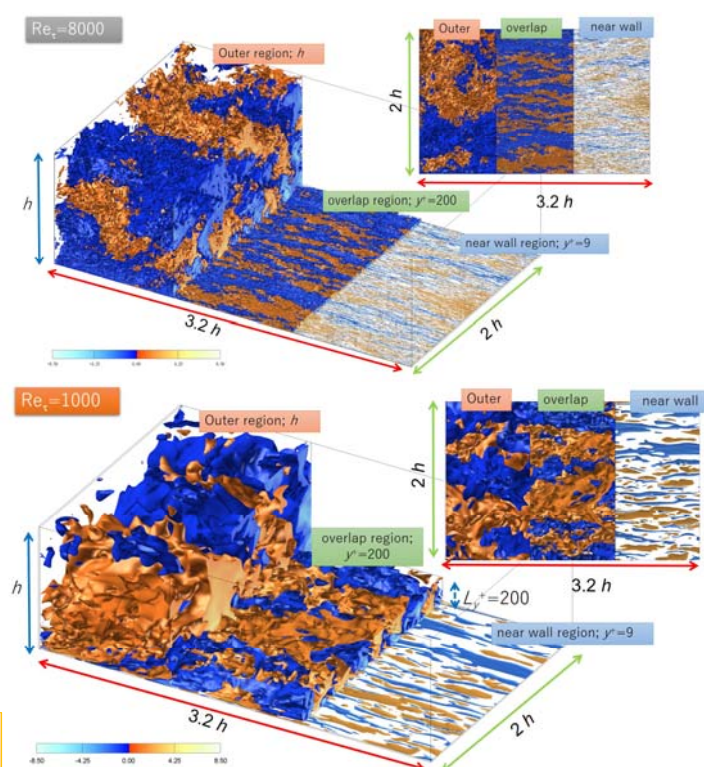


FIG. チャンネル流の可視化(Re効果)