

メタバース実現に求められる デジタルインフラの課題と要件

～まだ遠い理想のメタバース～

塚田 学

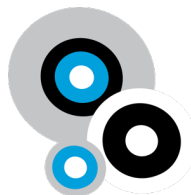
東京大学 情報理工学系研究科



Tsukada Lab.
Computer Network Group

2022年12月14日

第6回「Web3時代に向けたメタバース等の利活用に関する研究会」



- 自己紹介：塚田 学

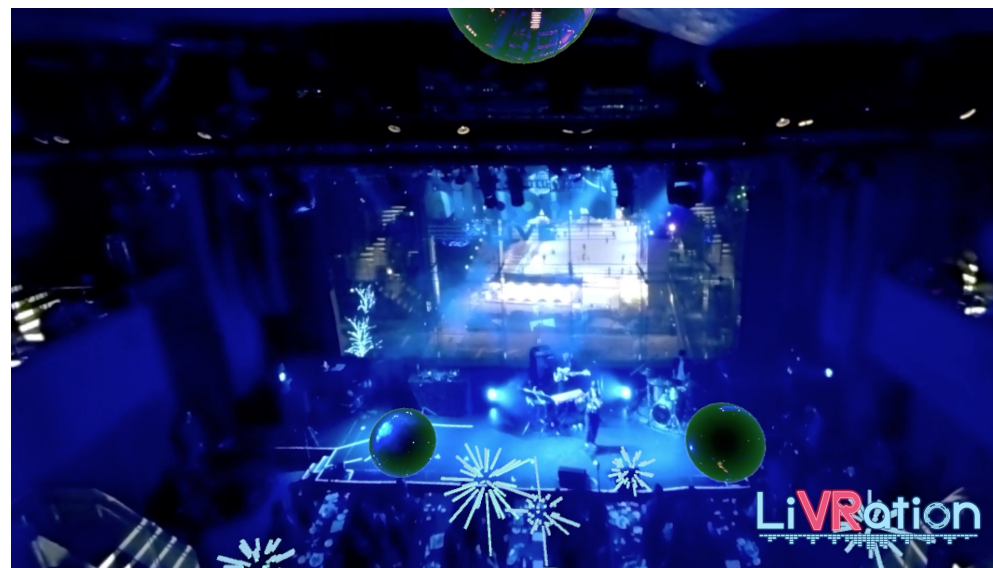
- 慶應義塾大学 SFC学部卒業・修士修了
- フランスMines ParisTech博士号取得・Inriaで研究員
- 2013年～東京大学・2019年研究室の設立

- 専門：コンピュータネットワーク、サイバーフィジカルシステム

ネットワーク型自動運転 (2004～)



インターネット映像音声メディア (2013～)





目次 ～まだ遠い理想のメタバーズ～

- インターネットの発展とメタバーズ
- メタバーズの概要と定義
- メタバーズ実現に求められるデジタルインフラの課題と要件
 - 実現がどれくらい遠いのか？
 - ネットワーク遅延
 - データ量
 - 処理能力



インターネットの発展とメタバーズ

- インターネット利用がデータの、時間的に拡大

1995-2000	2000年代	2010年～現在
インターネットの普及	ブロードバンドの普及	スマートフォンの普及
テキスト中心	画像・音声・動画、SNS	常にインターネットと共にある生活

- それに伴い、オンラインアイデンティも拡大 [1][2]
 - Z世代の33%がオンラインとオフラインの自分に違いはないと回答。または、35%がそれらの自分は非常によく似ていると回答。
 - Z世代の55%は、インターネットはオフラインで経験するよりも創造的な空間であると回答。

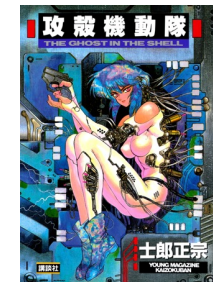
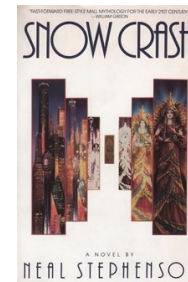
[1] <https://medium.com/@bvatere/generation-z-identity-and-expression-in-the-digital-world-50d2b28df829>

[2] A. Davies, "Decoding Gen Z Identity Construction in Social Networks Through The Paradigm of Branding," May 2020.

- 次世代インターネットとしてのメタバーズ

- 映像音声の共有から、デジタル3次元空間でのインタラクティブな五感体験へ

ニール・スティーブンソンの
SFデストピア小説
『スノウ・クラッシュ』（1992）
に登場する仮想空間サービスの名前
がメタバーズ。アバターなども登場



士郎正宗の漫画
『攻殻機動隊』（1991）
に登場する「電腦空間」

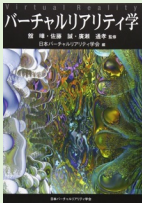
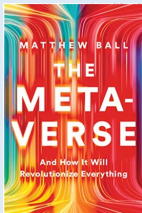
構想はインターネット普及前から存在



メタバーズの定義

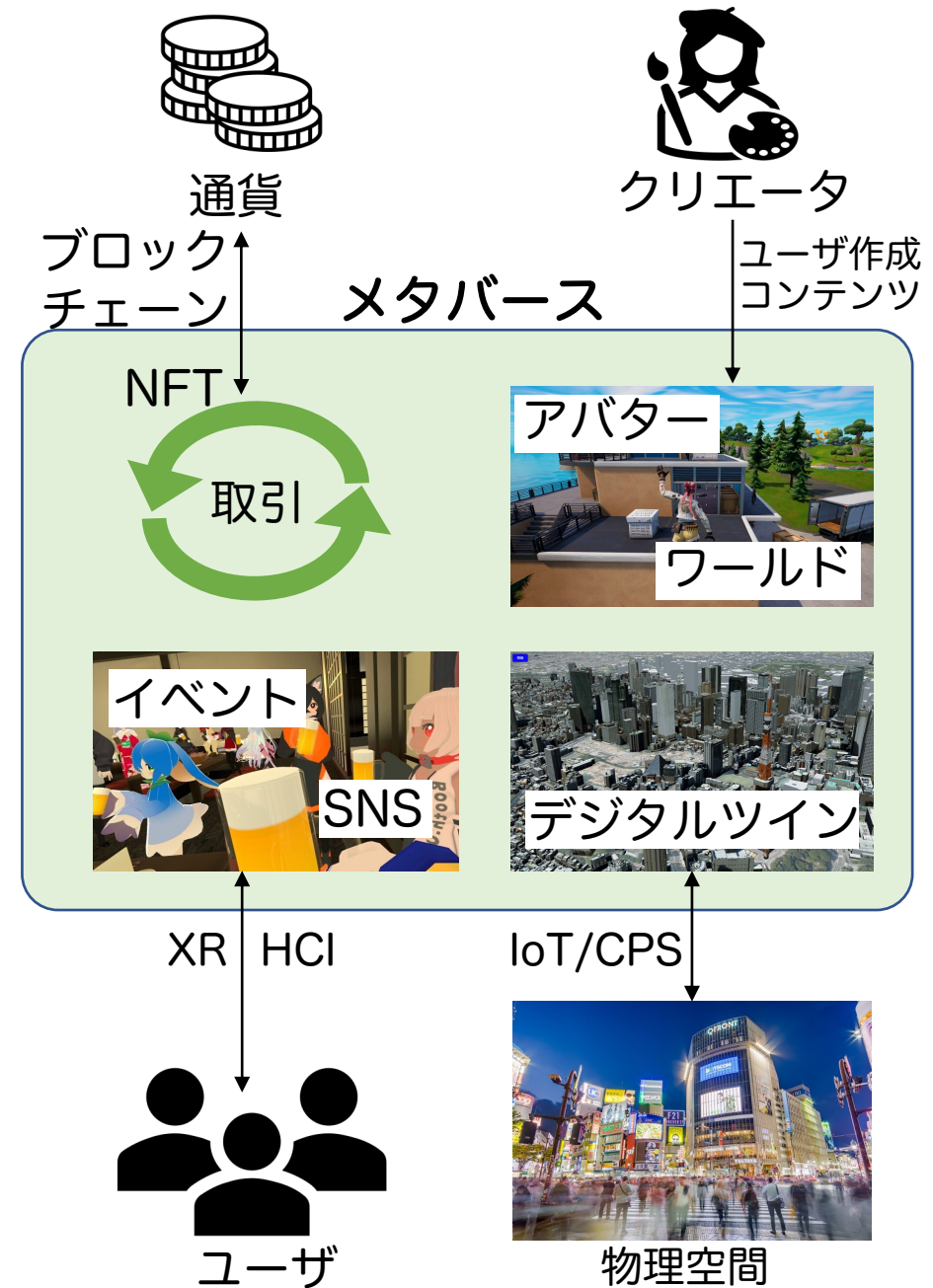
・厳密な定義は定まっていない

1. バーチャルリアリティ学, 舘 暉 (監修), 2010
2. The Metaverse: And How It Will Revolutionize Everything, Matthew Ball (著), 2022
3. メタバーズ進化論—仮想現実の荒野に芽吹く「解放」と「創造」の新世界, バーチャル美少女ねむ (著), 2022

	(1) 空間性： (2) 自己同一性： (3) 大規模同時接続性： (4) 創造性：	三次元の空間の広がりのある世界 自分のアイデンティティを投影した唯一無二の自由なアバターの姿で存在できる世界 大量のユーザーがリアルタイムに同じ場所に集まることのできる世界 ユーザー自身が自由にコンテンツを持ち込んだり創造できる世界
	(3) 同期性： (5) データ連続性： (6) 持続性： (7) 相互運用性：	事実上無制限の数のユーザーがイベントやユーザー間のインタラクションが同期して体験できる 個々の存在感を持ち、アイデンティティ、履歴、資格、オブジェクト、通信、支払いなどのデータの連続性を持つことができる（（2）自己同一性の拡張概念） 空間が永続し、持続的に体験できる世界 異なる運用主体・サービス間で、相互運用可能なネットワーク
	(8) 経済性： (9) アクセス性： (10) 没入性：	ユーザー同士でコンテンツ・サービス・お金を交換でき、現実と同じように経済活動できる スマートフォン・PC・AR/VRなど、目的に応じて最適なアクセス手段を選ぶ事ができ、物理現実と仮想現実が垣根なく繋がる アクセス手段の一つとしてAR/VRなどの没入手段が用意されており、没入感のある体験ができる

メタバース概要とネットワーク要件

- ネットワーク要件に関わる性質
 - (3) 大規模同時接続性：大量のユーザーがリアルタイムに同じ場所に集まることのできる世界
 - (3) 同期性：事実上無制限の数のユーザーがイベントやユーザー間のインタラクションが同期して体験できる
 - (7) 相互運用性：異なる運用主体・サービス間で、相互運用可能なネットワーク
- ネットワーク要件
 - 規模性
 - 遅延
 - データ転送速度





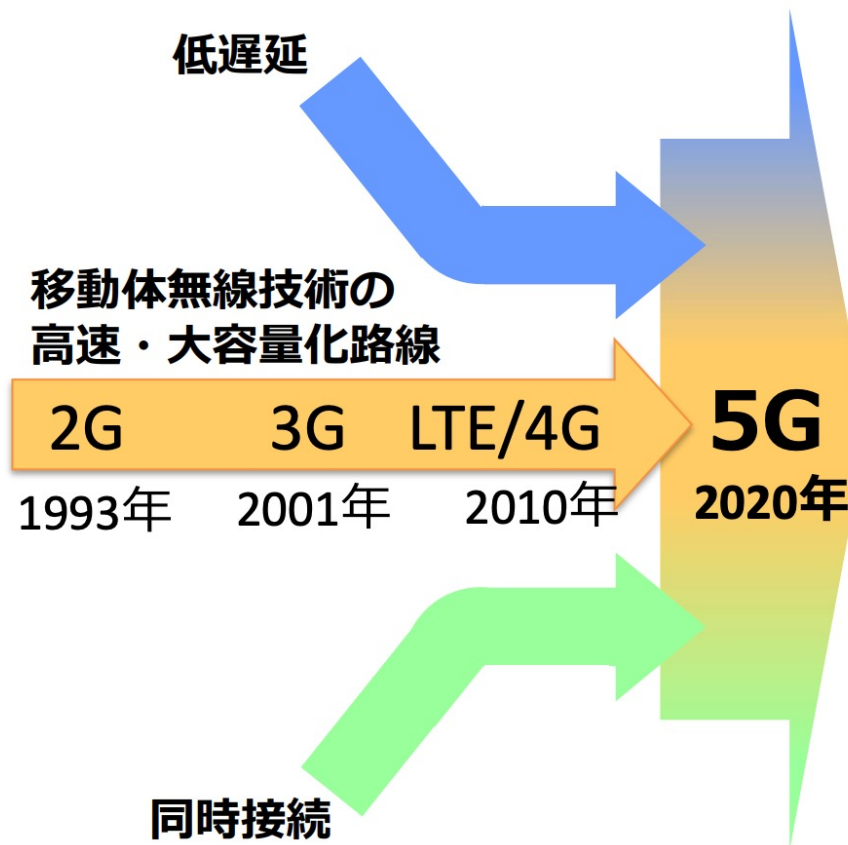
5Gで全て解決？

<5Gの主要性能>

超高速
超低遅延
多数同時接続



最高伝送速度 10Gbps
1ミリ秒程度の遅延
100万台/km²の接続機器数



超低遅延

利用者が遅延(タイムラグ)を意識することなく、リアルタイムに遠隔地のロボット等を操作・制御



ロボットを遠隔制御



東京の病院の専門医が
ヘリ内の医師に指示を
しながら遠隔で処置。

ヘリ内で緊急手術

⇒ ロボット等の精緻な操作 (LTEの10倍の精度) をリアルタイム通信で実現

超高速

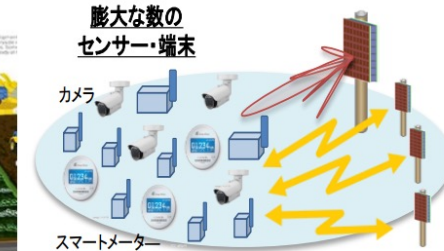
現在の移動通信システムより
100倍速いブロードバンドサービスを提供



⇒ 2時間の映画を3秒でダウンロード (LTEは5分)

多数同時接続

スマホ、PCをはじめ、身の回りのあらゆる機器がネットに接続



膨大な数の
センサー・端末

カメラ

スマートメーター

⇒ 自宅屋内の約100個の端末・センサーがネットに接続
(LTEではスマホ、PCなど数個)

5G 及び Beyond 5G に関する現状

https://www.soumu.go.jp/main_content/000670824.pdf

5G/6G MEC※でカバーできる／できない領域

※ Multi-Access Edge Computing

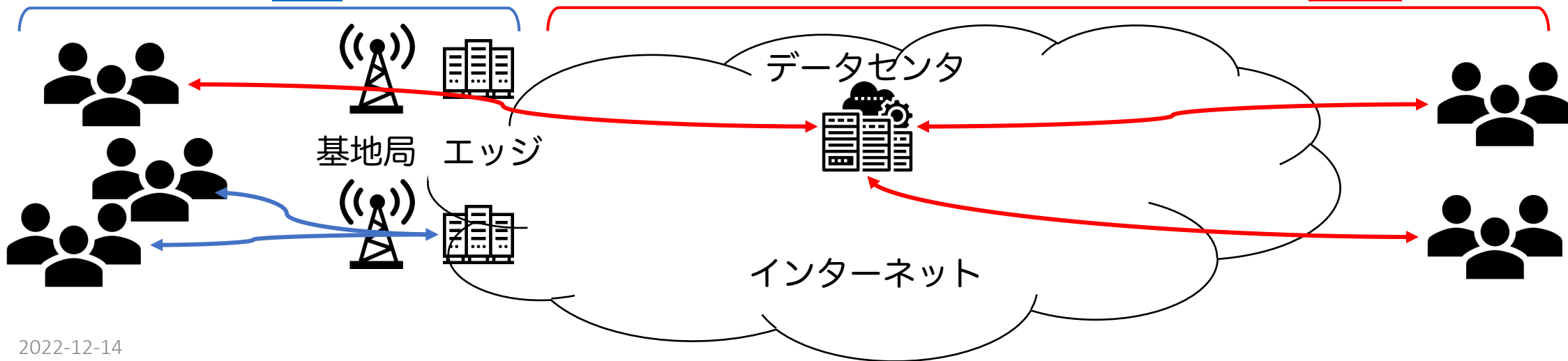
- できる：エッジコンピュータで折り返せるケースはカバー可能
 - 例) スタジアム等の地理的に限定されたケース、特殊なデジタルツイン・ARアプリなど
- できない：インターネット上にユーザが分散するケース
 - 例) メタバース、MMO、オンラインミーティング、遠隔演奏など
 - (緩和できる場合あり→次ページ)

5G/6G

超高速 10Gbps
超低遅延 1ms

インターネット

例えば、Tokyo-NYは距離12,500km。光の速度で45msかかる。
光ファイバーは光速の66%ほどの伝搬速度なので68msほどかかる。

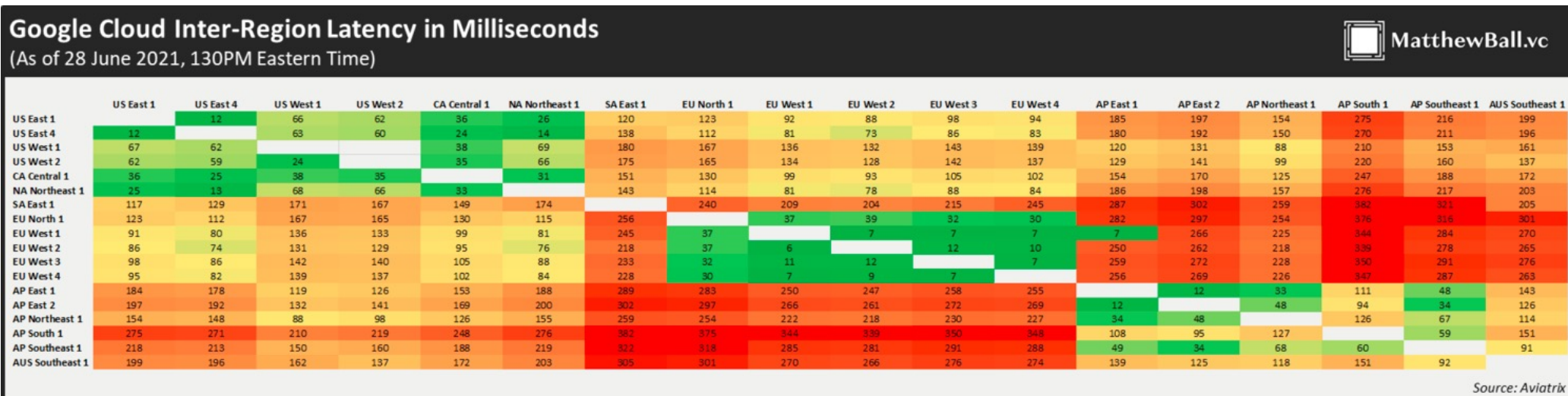




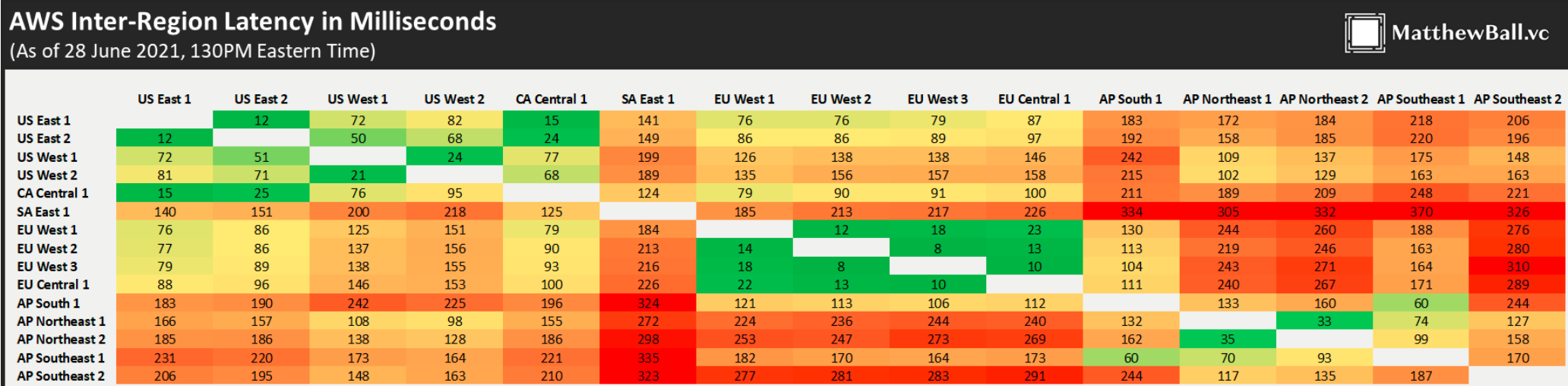
インターネットの遅延 (2021年)

- 都市間の通信遅延の中央値は100~200ms

- Google Cloud

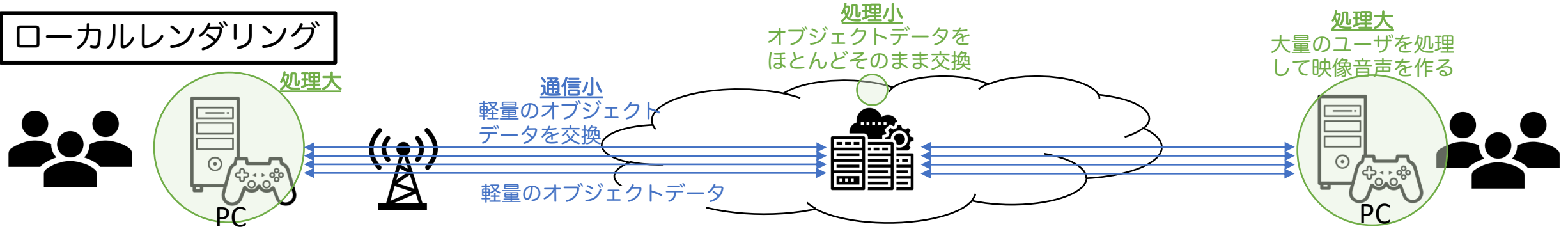


- AWS

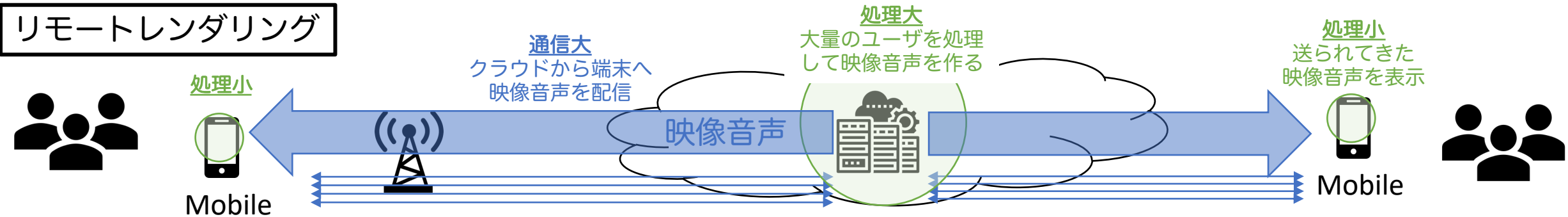


処理とネットワークのトレードオフ

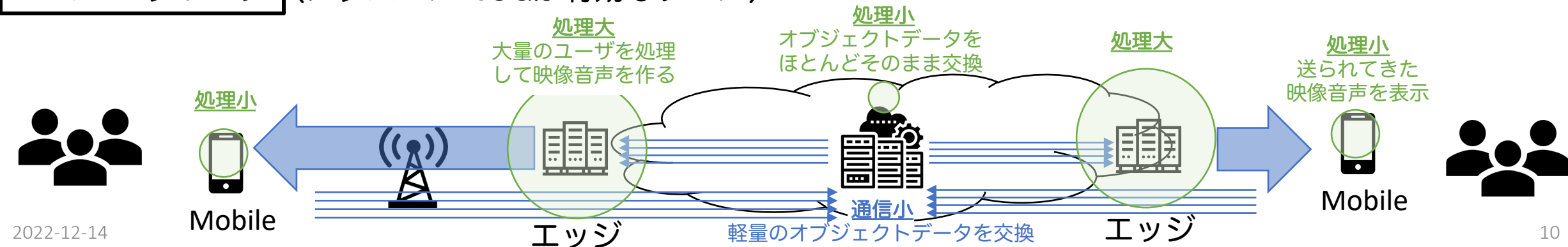
ローカルレンダリング



リモートレンダリング



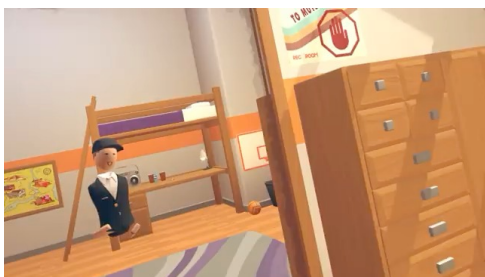
MECレンダリング (メタバースで5Gが有効なケース)





Preメタバーズのネットワーク計測[3]

- ・5つのソーシャルVRを計測
 - ・VRヘッドセットを利用したサービスを計測



Rec Room
<https://youtu.be/62t1ysH-umc>



AltspaceVR
<https://youtu.be/W6tAt34yC2w>



VRChat
<https://youtu.be/PWLpw4BE9Iq>

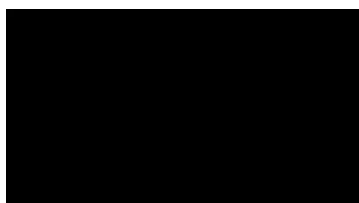


Horizon Worlds
<https://youtu.be/02kCEurWkgU>



Mozilla Hubs
<https://youtu.be/5QnOsyvbeEQ>

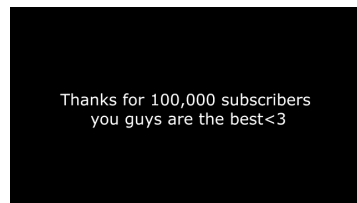
- ・2次元ディスプレイを使ったサービスを除く



Roblox



Fortnite
<https://youtu.be/ZyEafiMivUA>



Minecraft
<https://youtu.be/rExtPFCIbvQ>

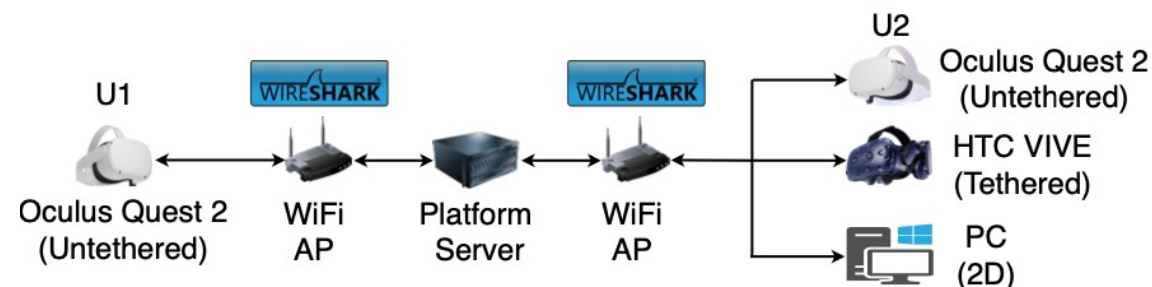


Figure 1: Measurement setup with two users U1 and U2.
送受信のパケットログを用いて計測

現在のソーシャルVRプラットフォーム^[3]

- ・制御チャンネル（ゲームのメニュー操作やクロック同期など）とデータチャンネル（アバターの体現や音声データ）のプロトコルが別
 - ・制御チャンネルはHTTPS、データチャンネルはUDP(またはRTP)
 - ・両チャンネルに別のサーバを割り当てるケースもある／適切に割り当てられない場合、遅延が増大

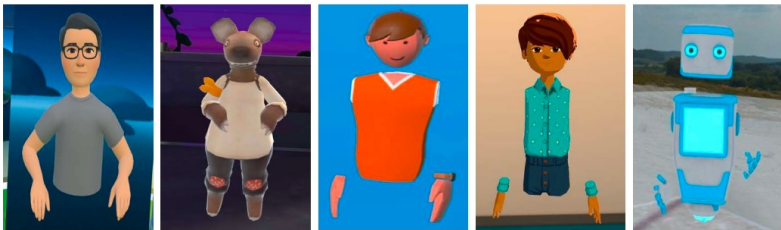
Platform	Control Channel				Data Channel			
	Protocol	Server Loc. / Owner	Anycast?	RTT (ms)	Protocol	Server Loc. / Owner	Anycast?	RTT (ms)
AltspaceVR	HTTPS	- / Microsoft	✓	3.08/0.1	UDP	Western U.S. / Microsoft	✗	72.1/0.2
Hubs	HTTPS	Western U.S. / AWS	✗	74.1/0.3	RTP/RTCP HTTPS	Western U.S. / AWS	✗	73.5/0.2 74.1/0.3
Rec Room	HTTPS	- / ANS	✓	2.21/0.1	UDP	- / Cloudflare	✓	2.97/0.1
VRChat	HTTPS	Eastern U.S. / AWS	✗	2.32/0.3	UDP	- / Cloudflare	✓	3.24/0.3
Worlds	HTTPS	Eastern U.S. / Meta	✗	2.23/0.2	UDP	Eastern U.S. / Meta	✗	2.71/0.1

表2：ソーシャルVRプラットフォーム5社のネットワークプロトコルとインフラ（実験は米国東海岸で実施）。サーバーの位置は、エニーキャストを使用した場合、-と表記している。RTTの2つの数値は、平均と標準偏差。

- ・ローカルレンダリングが利用されている
- ・スループットが低い
 - ・静的バーチャル背景はプレイ前にダウンロードされ、プレイ中はアバター情報がやり取りされる
 - ・一番データ量が多いWorldsでもUp 752Kbps/ Down 413Kbps
 - ・Worldsはアバターが他より詳細で、スループットほぼ10倍

Platform	Tput (Kbps)		Resolution	Avatar (Kbps)
	Up	Down		
VRChat	31.4/2.6	31.3/3.3	1440×1584	24.7/1.5
AltspaceVR	41.3/2.1	40.4/3.2	2016×2224	11.1/1.2
Rec Room	41.7/3.8	41.5/3.0	1224×1346	35.2/4.1
Hubs	83.3/5.6	83.1/6.4	1216×1344	77.4/7.7
Worlds	752/12	413/8.3	1440×1584	332/7.5

表3：5つのプラットフォームのアバター具現化に関するデータの総合スループット、コンテンツ解像度、スループット（クエスト2のユーザーは2名）。スループットの2つの数値は、平均結果と標準偏差。



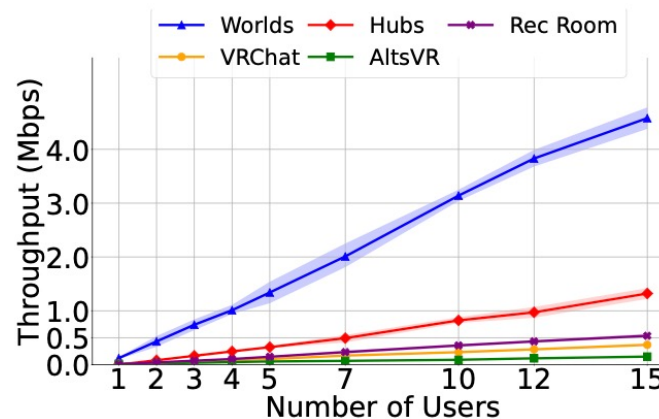
Worlds / VRChat / Rec Room/ AltspaceVR / Hubs



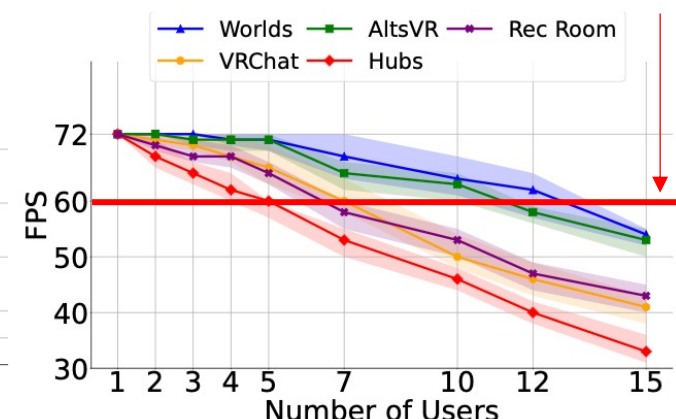
規模性[3]

- イベント参加ユーザ数が増加すると
1) スループットが増加し、
2) FPSが低下する
- イベント参加ユーザ数が増加すると
3) CPU、4) GPU、5) メモリ
の使用量が増加

VRの最小要件は60FPS[4]

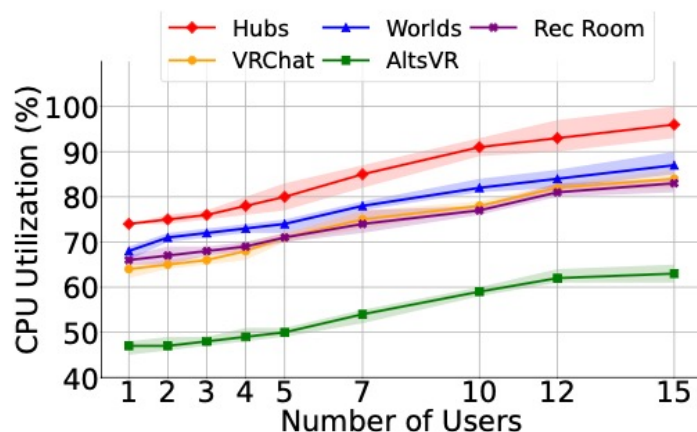


1) スループット

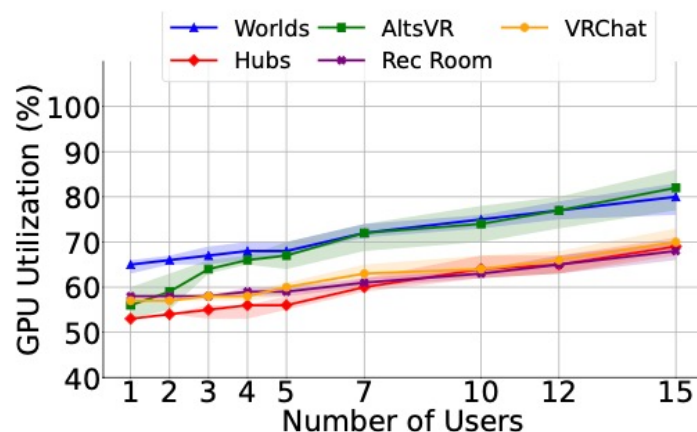


2) FPS

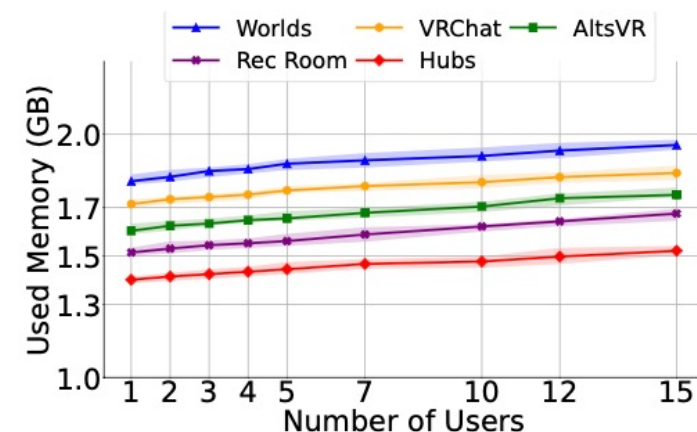
[4] Z. Lai, Y. C. Hu, Y. Cui, L. Sun, and N. Dai, "Furion: Engineering High-Quality Immersive Virtual Reality on Today's Mobile Devices," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, Snowbird, Utah, USA, Oct. 2017, pp. 409–421.



3) CPU利用率



4) GPU利用率



5) メモリ使用量

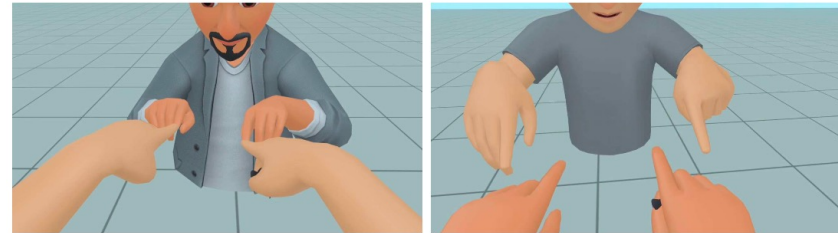


End-to-End遅延^[3]

Motion-to-Photon遅延

- シューティングゲームの熟達したゲーマー^[5]
 - 20 ms：メーカー推奨
 - 50 ms：フラストレーションがたまる
- カジュアルゲーマー^[5]
 - 110 msで失敗を遅延のせいにする
 - **150 ms**で素早い反応が必要なゲーム不能
- マルチモーダル・コミュニケーションシステム^[6]
 - **150 ms**は遅延の安全な目標値
- 映像・音声の同期^[7]
 - 175ms以内のずれに気づかない
 - 275ms以内のずれを許容する
- 楽器演奏^[8]
 - 30ms 以上の遅延だと認知され、
 - 50ms 以上の遅延だと演奏が困難

[5] E. Banatt, S. Uddenberg, and B. Scholl, "Input Latency Detection in Expert-Level Gamers," Yale University, 2017.
 [6] D. Roberts, T. Duckworth, C. Moore, R. Wolff, and J. O'Hare, "Comparing the End to End Latency of an Immersive Collaborative Environment and a Video Conference," in 2009 13th IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications, Oct. 2009, pp. 89–94.
 [7] The Metaverse: And How It Will Revolutionize Everything, Matthew Ball (著), 2022
 [8] 西堀佑, 多田幸生, 曽根卓朗, and Others, "遅延のある演奏系での遅延の認知に関する実験とその考察," 情報処理学会研究報告音楽情報科学 (MUS), vol. 2003, no. 127 (2003-MUS-053), pp. 37–42, 2003.



指を素早く動かして、反映までの時間を動画分析する

Platform	E2E	Sender	Receiver	Server
Rec Room	101.7/8.7	25.9/8.6	39.9/7.8	29.9/6.4
VRChat	104.3/9.3	27.3/6.2	37.4/6.4	33.5/9.5
Worlds	128.5/11	26.2/4.5	49.1/9.1	40.2/11
AltspaceVR	209.2/13	24.5/5.2	36.1/9.9	68.6/12
Hubs	239.1/7.3	42.4/6.3	60.1/6.5	52.2/7.7
Hubs*	130.7/6.3	40.3/5.2	61.5/5.7	16.2/2.4

表4：エンドツーエンド (E2E) の遅延時間 (単位: ms)、送信、受信、サーバ処理の遅延時間。各セル内の2つの数値は、平均結果と標準偏差。Hubs*は、我々のプライベートHubsサーバーを指します。

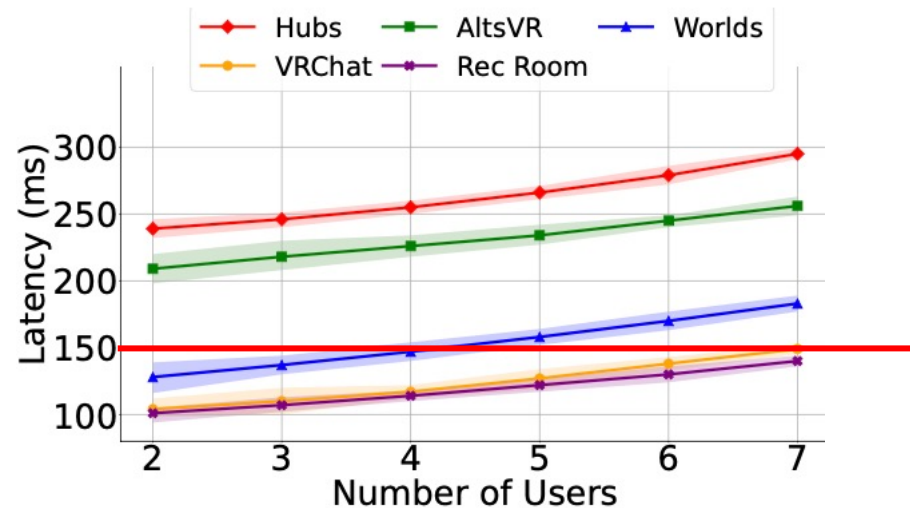


図11：5つのプラットフォームにおける、ユーザー数の異なる平均エンド・ツー・エンド遅延。帯は95%信頼区間を表す。



クラウドゲーミング（リモートレンダリング）

- 3つのクラウドゲーミングを比較^[9]
 - Google Studio （サービス終了）
 - Nvidia GeForce Now
 - Sony PlayStation Now
 - 以下は未検証
 - Amazon Luna
 - Microsoft Xbox Cloud Gaming (Beta)
- StudioとGeForce NowはWebRTCを利用し、最大45Mbpsの帯域幅を消費（4K動画）
 - Xbox Cloud GamingもWebRTCを利用^[10]
- PS Nowは独自のUDPを使い、13Mbpsを超えることはない

[9] A. Di Domenico, G. Perna, M. Trevisan, L. Vassio, and D. Giordano, “A Network Analysis on Cloud Gaming: Stadia, GeForce Now and PSNow,” *Network*, vol. 1, no. 3, pp. 247–260, Oct. 2021.

[10] P. Graff, X. Marchal, T. Cholez, S. Tuffin, B. Mathieu, and O. Festor, “An Analysis of Cloud Gaming Platforms Behavior under Different Network Constraints,” in 2021 17th International Conference on Network and Service Management (CNSM), Oct. 2021, pp. 551–557.



まとめ

理想

- 事実上無制限の数のユーザーがイベントやユーザ間のインタラクションが同期して体験できる

現実

- 15人ぐらいで、VR最低要件（60FPS）を割り込む
- 15人ぐらいで、ゲーム不能レベルの遅延が発生

将来

- バーチャル背景とのインタラクションを遅延が増大
- アバターを詳細化すると遅延が増大
- Holoportation^[11]では、1人の転送で1Gbps



<https://youtu.be/84ngPFBWLVI>

[11] S. Orts-Escalano et al., “Holoportation: Virtual 3D Teleportation in Real-time,” in Proceedings of the 29th Annual Symposium on User Interface Software and Technology, Tokyo, Japan, Oct. 2016, pp. 741–754.

レンダリング	処理			ネットワーク		利点	欠点
	端末	MEC	DC	インターネット	5G		
ローカル	大	--	小	負荷小（解像度に非依存） 1Mbps程度		ネットワーク負荷が小さい	端末の省力化が難しい
リモート	小	--	大	負荷大（解像度に依存） 最大45Mbps程度		- 端末が省力化できる - 処理の整合性が取りやすい	- ネットワーク負荷が大きい - motion-to-photon遅延が大きくなる
MEC	小	大	小	負荷小	負荷大	- 端末が省力化できる - 処理の負荷分散ができる - インターネットの負荷が小さい	エッジの割り当てが鍵（ユーザに近く、負荷分散できる場所）



課題

- 通信事業者、データセンタ、OEM、サービス事業者、政策立案者、より広い業界の連携
-
- 調和の取れたネットワークアーキテクチャ
 - 計算資源に応じた MEC の効率的配備
 - ローカル、リモート、MECのハイブリッドリ・レンダリング
 - 大規模End-to-End性能測定



Internet of Realitiesでは教員/研究者を募集しています

CREST研究課題「多様な形態の現実を安心・安全に創り・繋ぐ
Trusted Inter-Reality 基盤」(研究代表者: 米澤拓郎准教授)

専門性: メタバース, デジタルツイン, VR/AR/MR/XR, ネットワーク, HMI, AI, セキュリティ、プライバシー、ブロックチェーン

Unity, Unreal Engineなどゲームエンジンの開発経験があることが望ましい。



<https://www.internet-of-realities.org/>

<https://tlab.hongo.wide.ad.jp/>

