JLUG 2021

Intel® Optane™ Technology, DAOS update and Lustre Persistent Client Cache with Intel® Optane™ Persistent Memory

Intel Optane Group Fumiyasu Ishibashi

第3世代インテル® Xeon® スケーラブルプロセッサ対応 インテル® Optane™ テクノロジー製品

第３世代インテル® Xeon® スケーラブル・プロセッサ・ファミリー

インテル® Optane™ パーシステント・メモリー **200**シリーズ (2020年12月発表)

8チャンネル メモリー	128~512GB モジュール対応
ソケットあたり最大 4TB	メモリー帯域 3,200MT/s
eADR ADモード対応アプリケーションの性能向上	-17% 前世代と比べて消費電力を削減

メモリー帯域が最大 **32%** 向上

2ソケットサーバーで最大**8TB**を実現

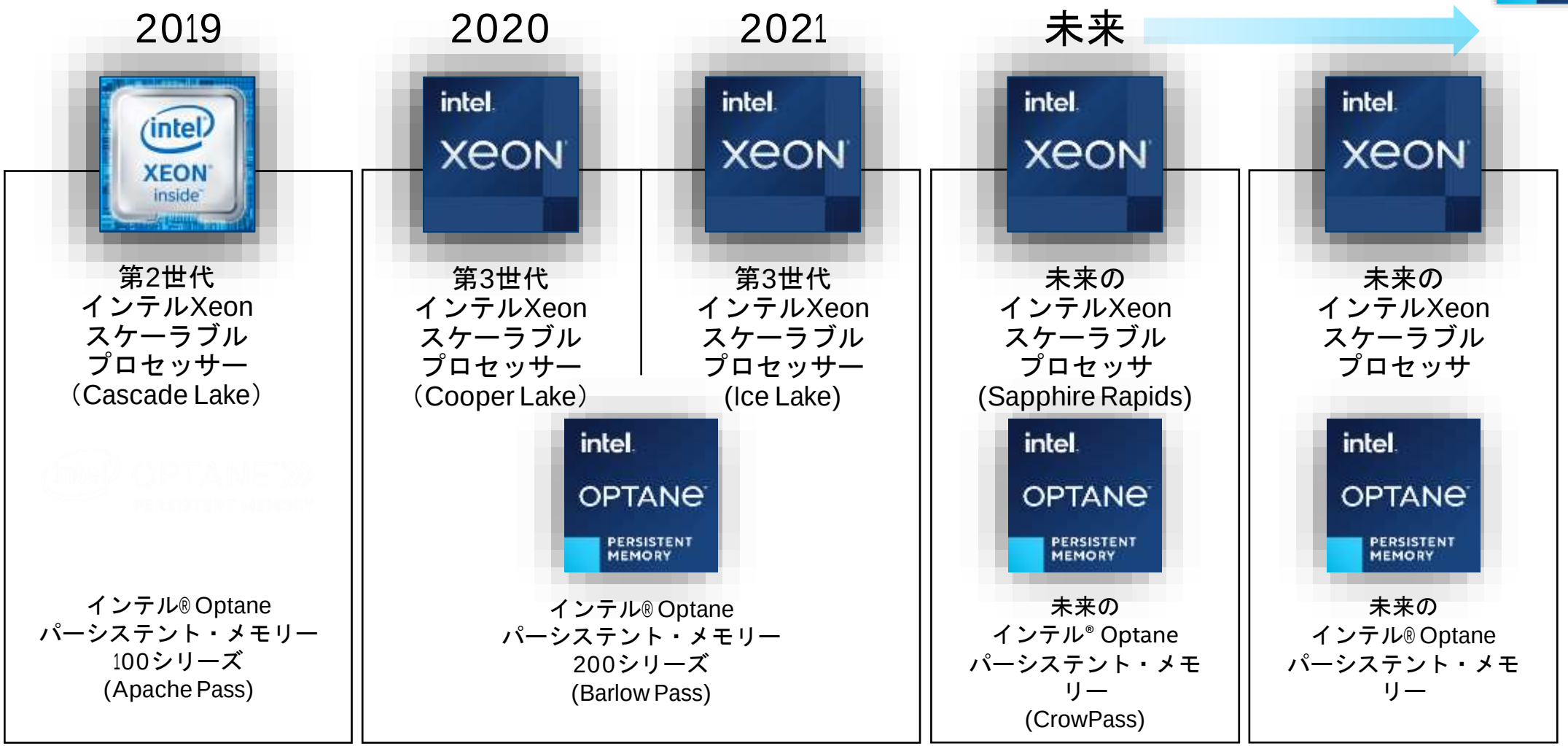
インテル® Optane™ SSD **P5800X** シリーズ (2020年12月発表)

PCI 4.0 対応	1.5M IOPS 最大
シーケンシャル･リード 7.2GB/s	超低レイテンシー 6us
100 DWPD 一般的なNAND SSDの30~100倍の耐久性	+40% 前世代と比べてQoSが向上

前世代と比較して **約3倍** 高性能

世界最速のデータセーンター向けSSD

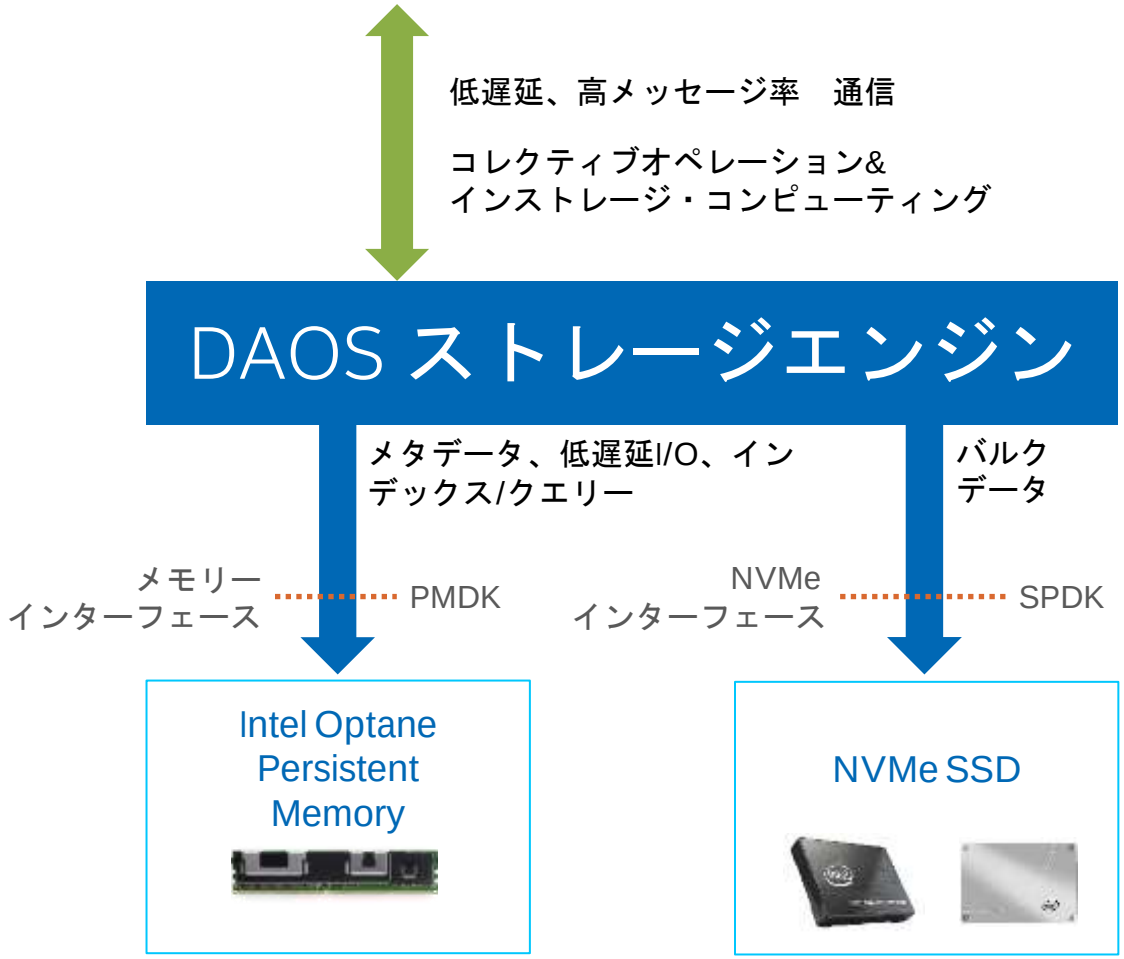
強力なメモリとストレージの未来



DAOSとは？

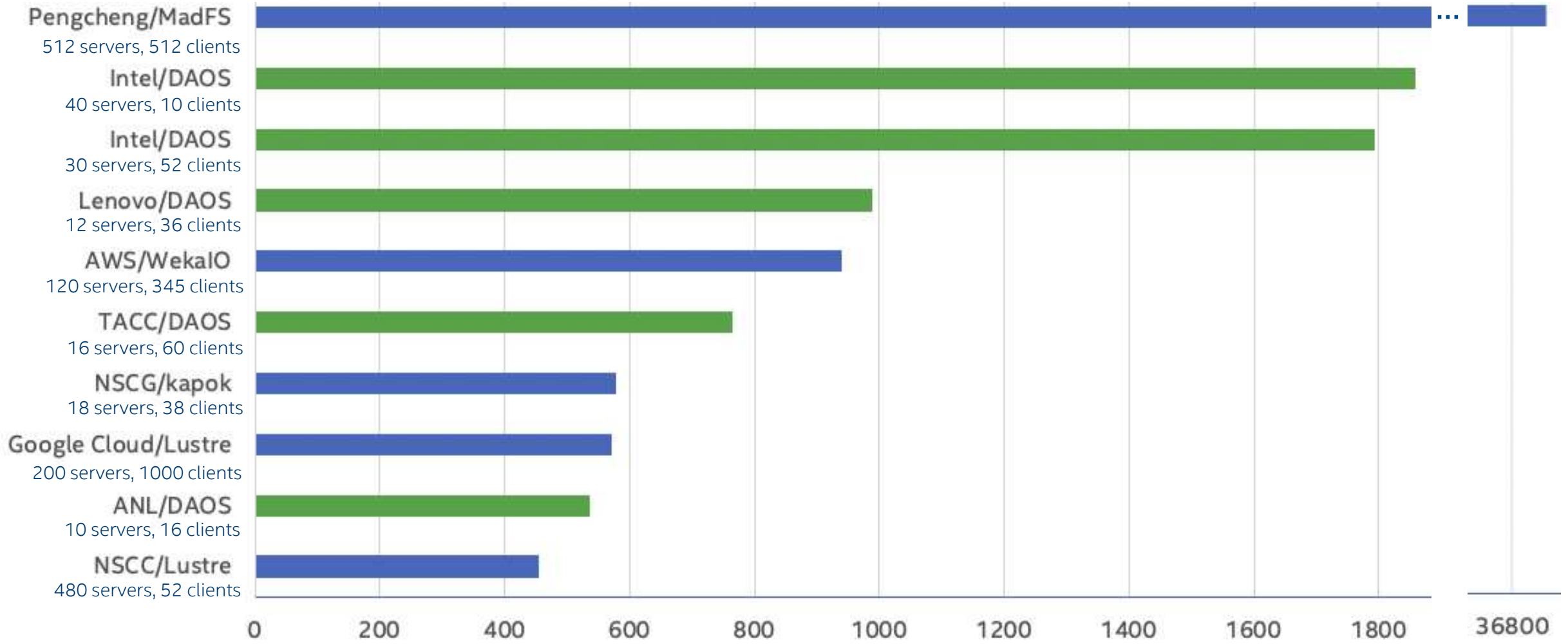
- インテル® Optane™ パーシステント・メモリーとNVMe SSD をベースにした、新しい革新的な分散型並列オブジェクトストア
- ブロックベースのIOとPOSIXにおける業界のボトルネックを克服
- コモディティサーバー上で非常に高い帯域幅とIOPSを実現し、HPC、AI、高性能データ分析の要求に応える
- スタンドアロンのファイルシステムとしても、既存のストレージシステムと統合したパフォーマンス層としても利用可能

http://daos.io/



価格あたりのIOPsと帯域幅を増やし、記録的なパフォーマンスを実現

IO-500 Full List Overall Score (ISC'21)



ベンダー採用の増加 – フルリストのトップ10のうち5つがDAOS

Source: io500.org

DAOS エコシステム

ハードウェア パートナー	Hewlett Packard Enterprise, intel, SUPERMICRO, Atos, Orchestrating a brighter world NEC, Lenovo, inspur, QCT
リセラー パートナー	COLFAX Customized Solutions, RSC, 2crsi, MEGWARE, E4 COMPUTER ENGINEERING, ScaleWorX
ソフトウェア開発 サードパーティー サポート	croit, Brightskies, Nettrix宁畅, SEAGATE, Google Cloud, H3C 数字化解决方案领导者
エンドユーザー	JINR, Argonne NATIONAL LABORATORY, Sechenov First Moscow State Medical University, UNIVERSITY OF CAMBRIDGE, lrz

DAOS ロードマップ

DAOS 1.0
Released June 18, 2020

DAOS 1.0
2020年6月18日リリース
NVMe & DCPMMのサポート
プール毎のACL
DAOSのdfuse経由でのUNS
MPI-IOドライバー
HDF5 DAOSコネクター
基本的なPOSIX I/Oのサポート
CentOS 7対応
レプリケーション＆セルフヒーリング（プレビュー）

DAOS 2.0

イレイジャーコード
静止状態のサーバードレイン
テレメトリーとジョブごとの統計
Libfabricの安定化
操作性の向上
CentOS 8のサポート

DAOS 2.4

チェックサム・スクラビング
分散型トランザクション
コンテナパーキング/シリアル化
マルチホーム・ネットワークのサポート
キャタストロフィックリカバリー
DSAの最適化

DAOS 3.0

プログレッシブレイアウト
インライン重複排除
インラインでの暗号化(QAT)
述語のプッシュダウンによるSQLサポート
メタデータのSSDへのシリアライズ

DAOS 1.2
Released April 30, 2021

エンド・ツー・エンドのデータ整合性
コンテナーごとのACL
レプリケーション＆セルフヒーリング
休止中のサーバーの再統合
HPD5 DAOS VOLコネクタ
POSIX I/Oの改善
POSIXデータムーバー
Apache Sparkのサポート
Leap 15のサポート
イレージャーコード（プレビュー）

DAOS 2.2

Lustreの統合
オンラインドレイン、再統合＆追加
インテルVMDのサポート
Libfabric TCP Providerのサポート
HPE Slingshotのサポート
MLX UCXのサポート

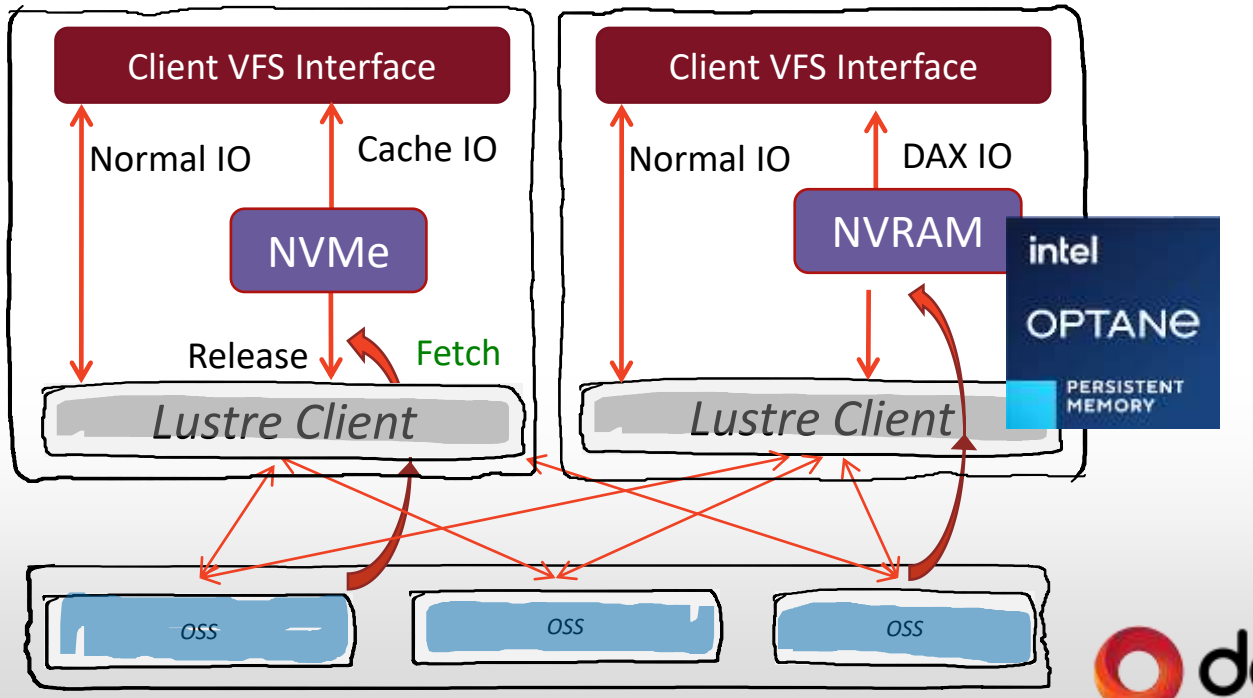
DAOS 2.6

ネットワークトポロジーの認識
DAOSプールのサイズ変更
ティアリングサポート
インライン圧縮(QAT)

注：このロードマップに記載されている情報は、予告なしに変更されることがあります。

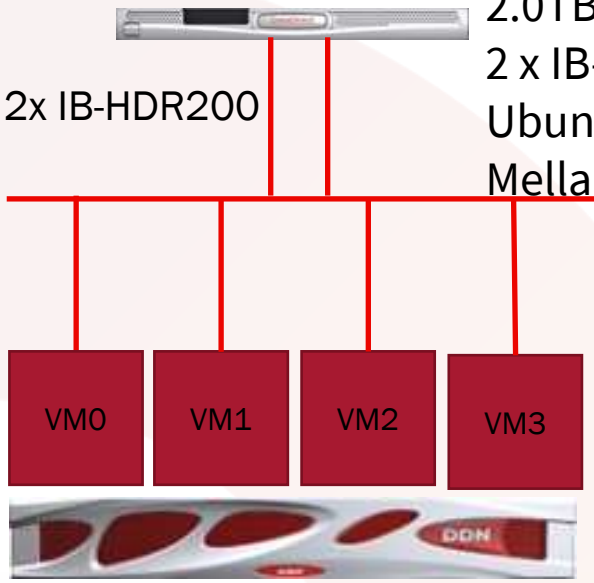
DDN EXA6 永続的クライアントキャッシュ (PCC: Persistent Client Cache)

- クライアントのNVMe/NVRAM領域を透過的にLustreのキャッシュ領域に活用
  - ファイルアクセス時に非同期でファイルをPCC(ローカルキャッシュ)に保存
  - 同一ファイルへの再アクセスの高速化とネットワークリソースの軽減
  - Optane Persistent MemoryもDAXモードとして利用可能



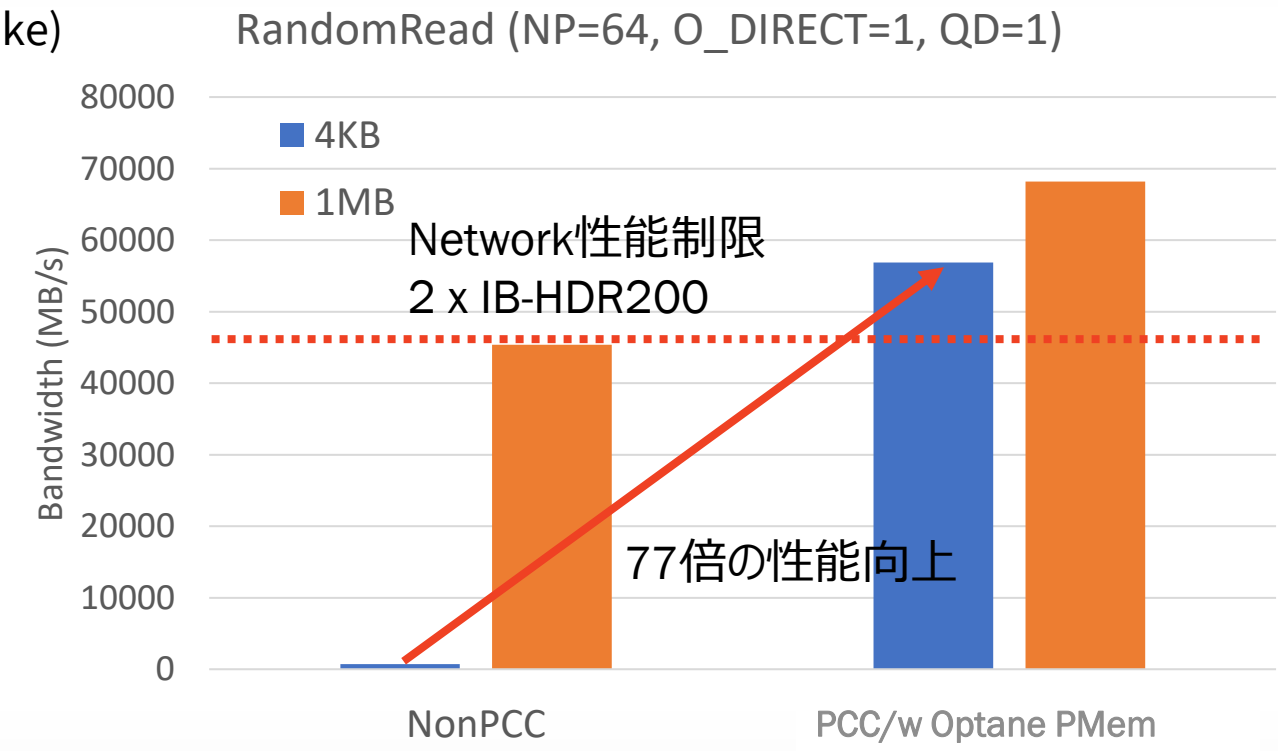
© 2021 DDN

Optane Persistent Memoryを使ったPCC(PCC-PMem)性能



EXAScaler(Lustre)6 Client
2 x Gold 6338 CPU @ 2.00GHz(IceLake)
512GB DDR4(3200MHz)
2.0TB Optane PMEM 200
2 x IB-HDR200
Ubuntu20.04 (5.4.0-91-generic)
Mellanox OFED-5.4-3.1.0.0

ES400NVX2
DDN EXAScaler Appliance
4 x VM(OSS/MDS)
20 x 3.84TB NVMe
4 x IB-HDR200



- DAXを使ってLinux kernelのPageCacheをバイパス
- QD=1の4KB IOサイズにおいてPCC-PMemは圧倒的な性能実現
- PCC-PMemでは、ネットワーク性能(2xIB-HDR200)の性能

© 2021 DDN

intel