Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

パブリッククラウドを利用した ALMA観測データの品質保証実験

小杉 城治、森田英輔、中里剛、林洋平、ミエル ルノー（国立天文台: NAOJ）
合田憲人、吉田浩（国立情報学研究所: NII）

2020/2/14 宇宙科学情報解析シンポジウム

Contents

- ALMA概要とその運用
- ALMAアーカイブ
- クラウドアーカイブ
- ALMAパイプライン
- クラウドパイプライン
- まとめ

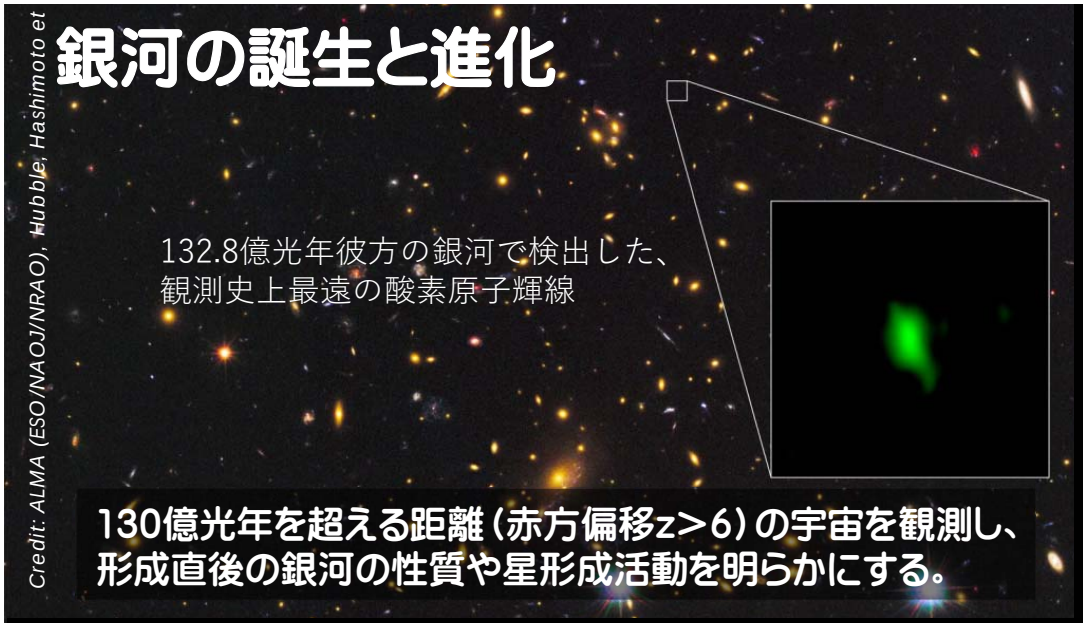
ALMAとは

ALMA望遠鏡：人類の新しい眼

- 南米チリ、標高5000mのアタカマ高地に建設、2011年に科学観測を開始。
- 東アジア（日本・台湾・韓国）、北米（アメリカ・カナダ）、欧州南天天文台（16カ国）とチリの国際協力。
- 66台のアンテナ（12m×54台＋7m×12台）を結合してひとつの巨大電波望遠鏡を構成。アンテナの展開範囲は最大16km、山手線大の望遠鏡に匹敵する分解能（0.01秒角=視力6000）を実現。
- 観測波長はミリ波・サブミリ波（波長 0.35～3.6 mm、周波数 86～950 GHz）。

Credit: Clem & Adri Bacri-Normier (wingsforscience.com)/ESO

銀河の誕生と進化

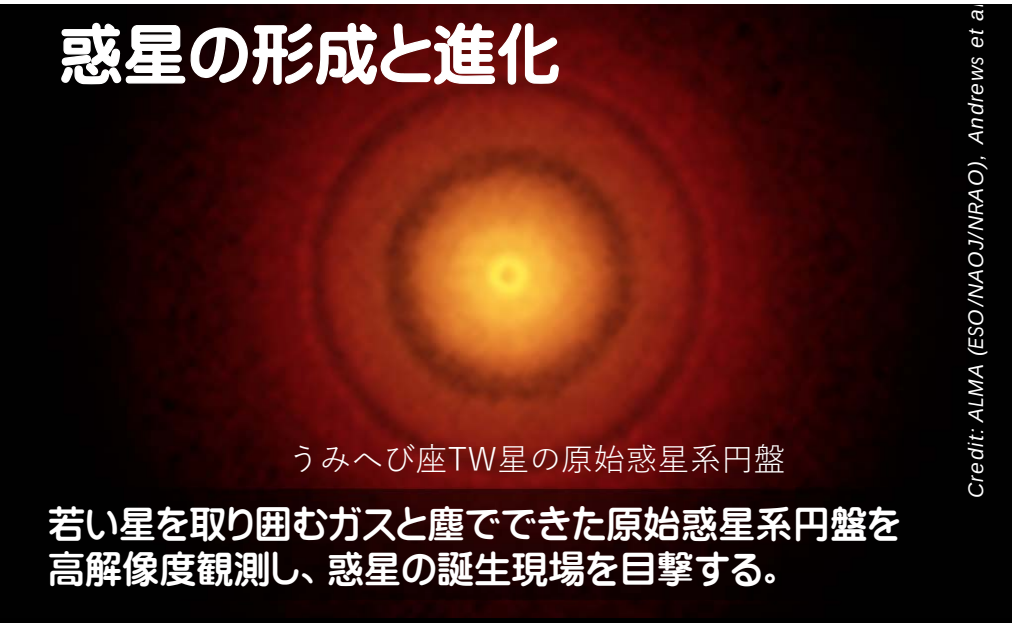


132.8億光年彼方の銀河で検出した、
観測史上最遠の酸素原子輝線

130億光年を超える距離（赤方偏移z>6）の宇宙を観測し、
形成直後の銀河の性質や星形成活動を明らかにする。

Credit: ALMA (ESO/NAOJ/NRAO), Hubble, Hashimoto et

惑星の形成と進化



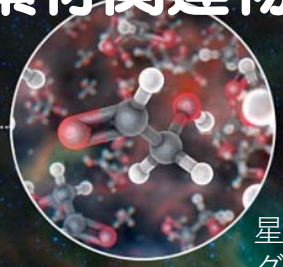
うみへび座TW星の原始惑星系円盤

若い星を取り囲むガスと塵でできた原始惑星系円盤を
高解像度観測し、惑星の誕生現場を目撃する。

Credit: ALMA (ESO/NAOJ/NRAO), Andrews et al.

ALMA望遠鏡が挑む宇宙の謎

生命素材関連物質探査



星形成領域で発見した有機分子
グリコールアルデヒドの想像図

生命の材料となりうるアミノ酸など複雑な有機分
子を、宇宙の様々な場所に探す。

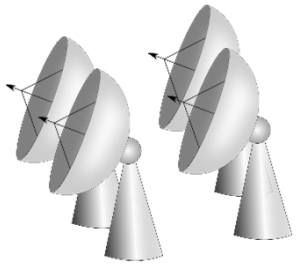
Credit: ESO/L. Calçada & NASA/JPL-Caltech/WISE Team

Credit: ESO/C. Malin

ALMAの運用

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

観測提案受付・審査



観測・データ取得

データアーカイブ

データ解析
パイプライン

データ品質保証

観測提案者への
データ配布・
1年後に公開

ALMAのデータレート

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

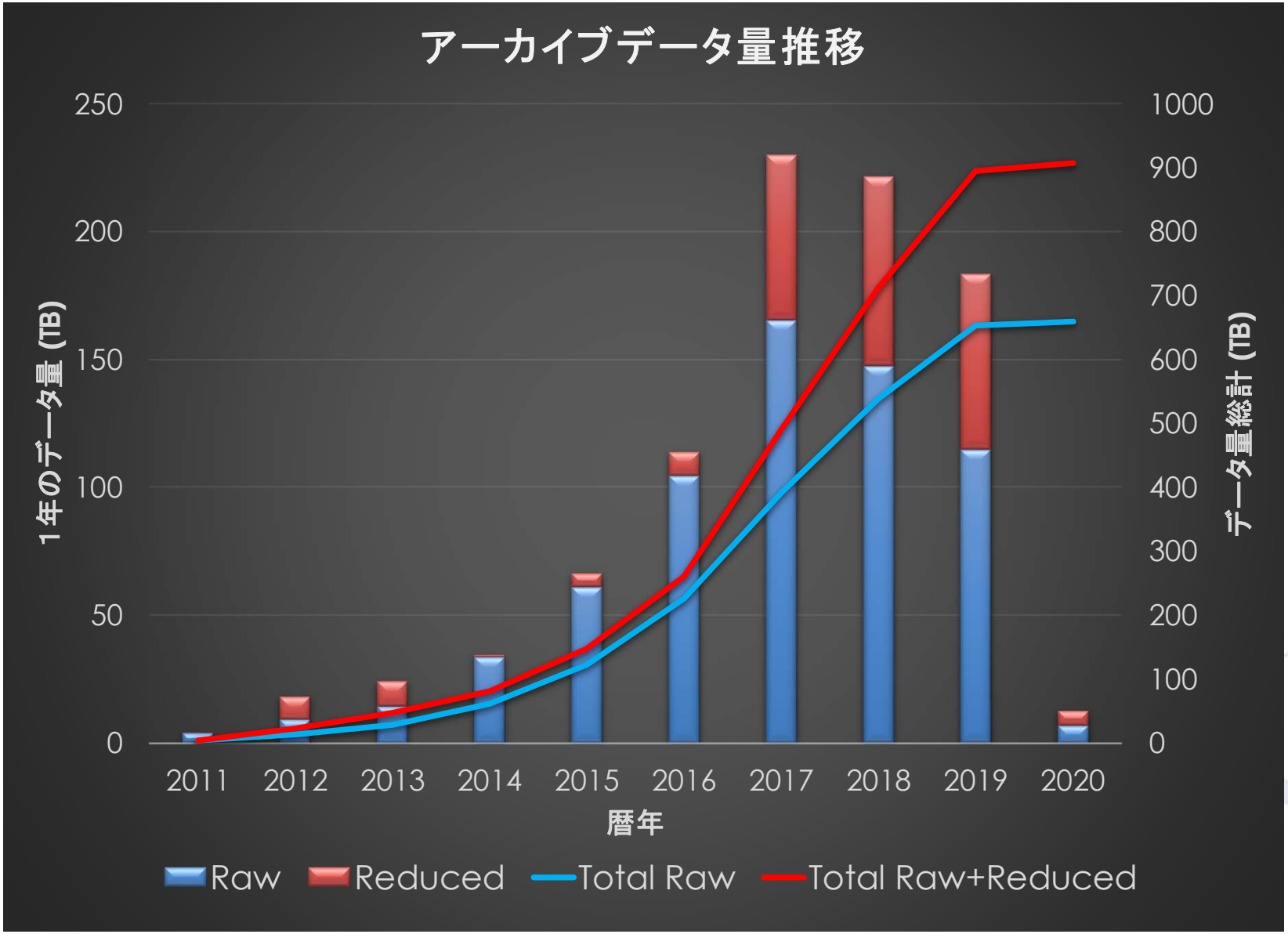
- 最大データレート：60 MBytes/s
- 平均データレート：6 MBytes/s
- 年間データ量の目安: 200TB/year程度

- 最大データレートや年間データ量は数年後にアップグレードされる予定。

→ ～700TB/year?

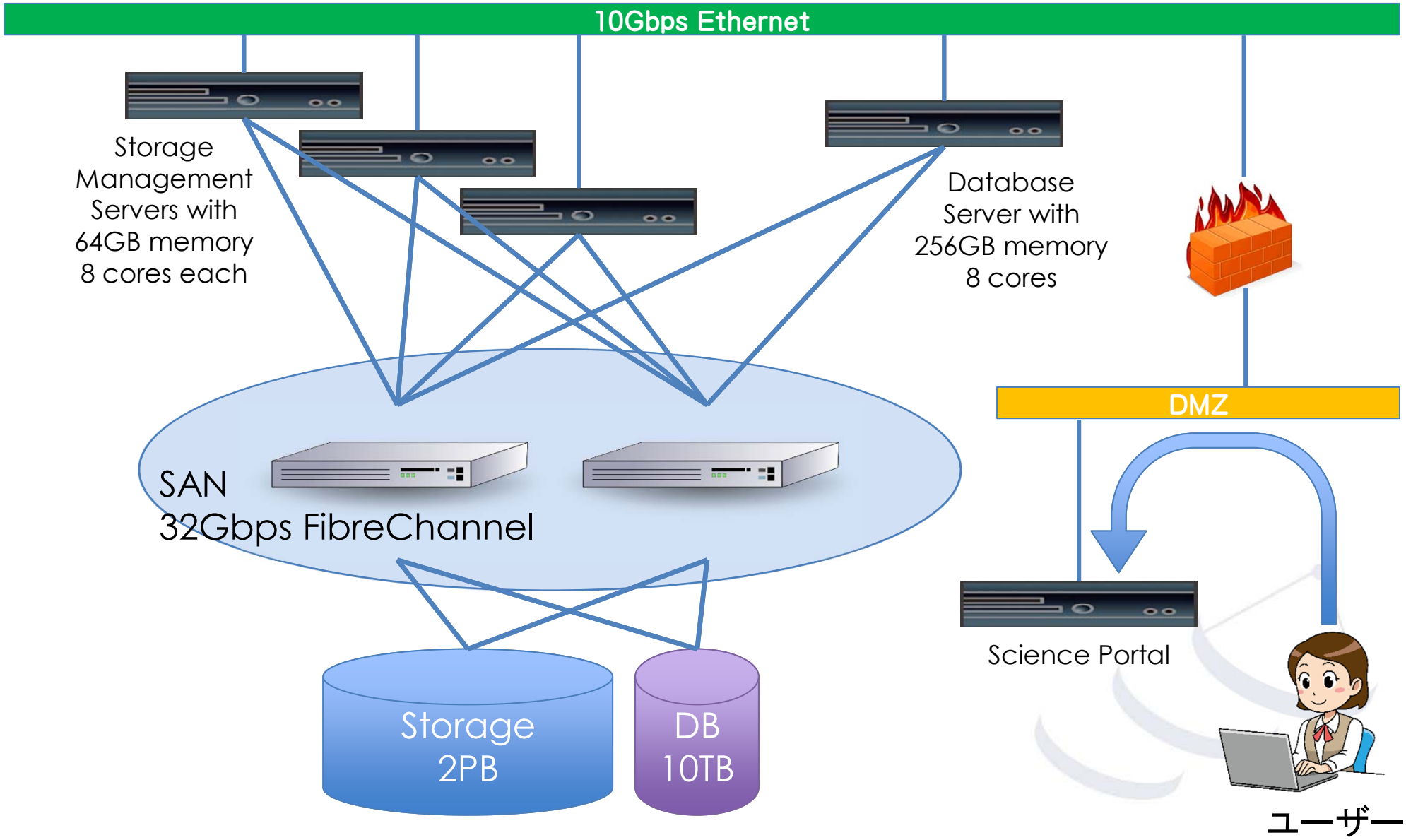
ALMAアーカイブデータ量

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



オンプレミスALMAアーカイブ

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



クラウドコールドストレージ

- コールドストレージ
  - 通常のストレージより低いコストで、頻繁にはアクセスされないデータ(コールドデータ)を格納する。性能(アクセス速度など)を多少犠牲にするが、比較的低コストで長期データ保管に有利
- パブリッククラウドのコールドストレージサービス
  - データの保管料金が比較的安い
  - 但し
    - データの復元処理(長時間を要する)が必要な場合がある
    - データアクセスの課金が割高などの制約がある
- アクセス頻度の少ないアーカイブデータに適している

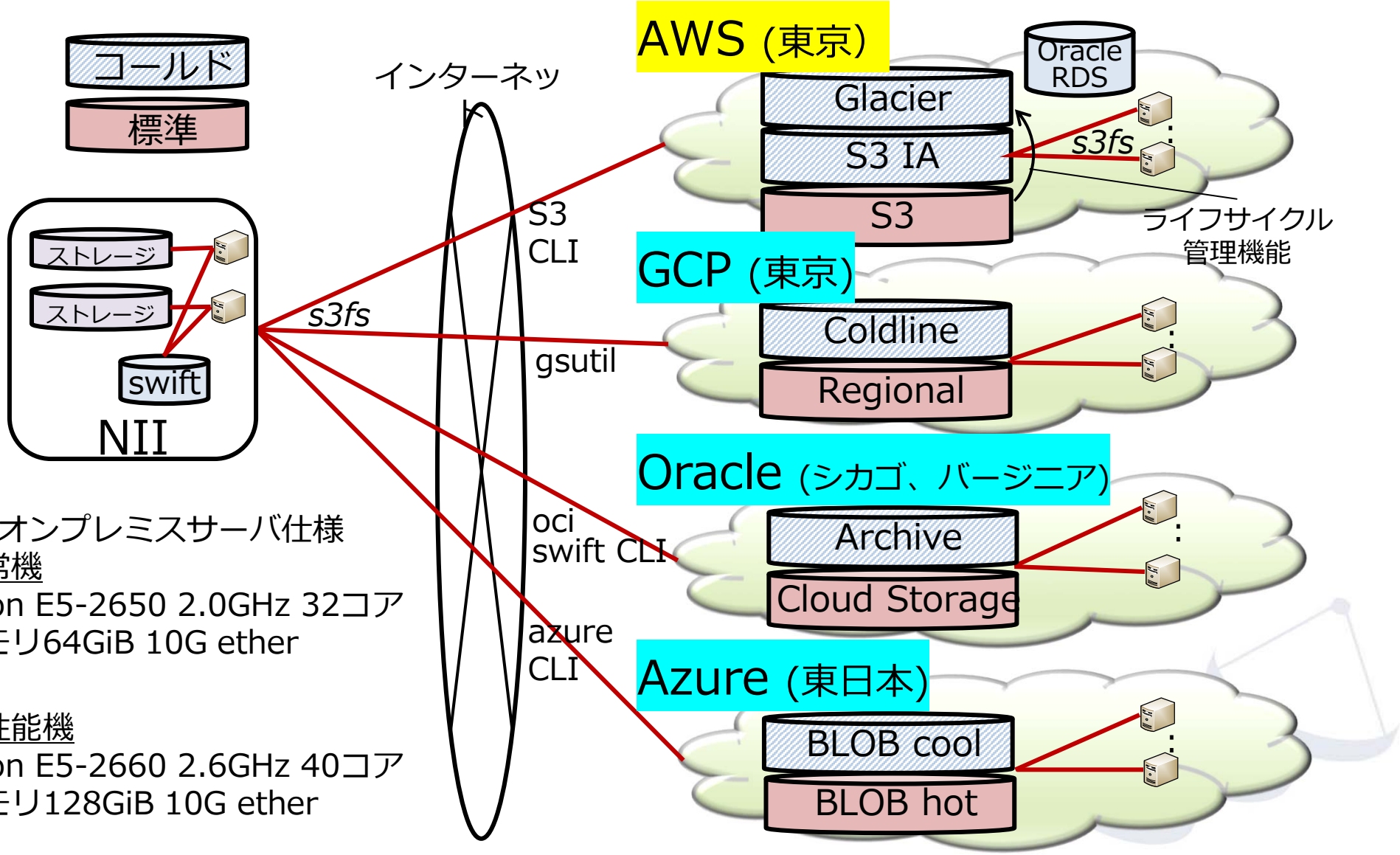
クラウドストレージの料金体系

- Amazon Web Service (AWS)の参考料金 (2018/6)

	課金内容	金額(単位USD)
AWS S3	データ保管(GB当り月額)	0.025
AWS S3	データアクセスリクエスト料金(GB換算)	0.0037
AWS S3 IA	データ保管(GB当り月額)	0.019
AWS S3 IA	データアクセスリクエスト料金(GB換算)	0.010
AWS S3 IA	データ取り出し料金(GB換算)	0.010
AWS Glacier	データ保管(GB当り月額)	0.005
AWS Glacier	データリストア(標準)リクエスト料金(回)	0.000057
AWS Glacier	データリストア用スペース料金(1日)	0.000852
AWS Glacier	データリストア後アクセスリクエスト料金(GB換算)	0.0037
AWS Glacier	データリストア後取り出し料金(GB換算)	0.011
共通経費	対インターネットデータ転送料金	0.14
共通経費	対SINETデータ転送料金	0.042

クラウドコールドストレージの実証実験

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



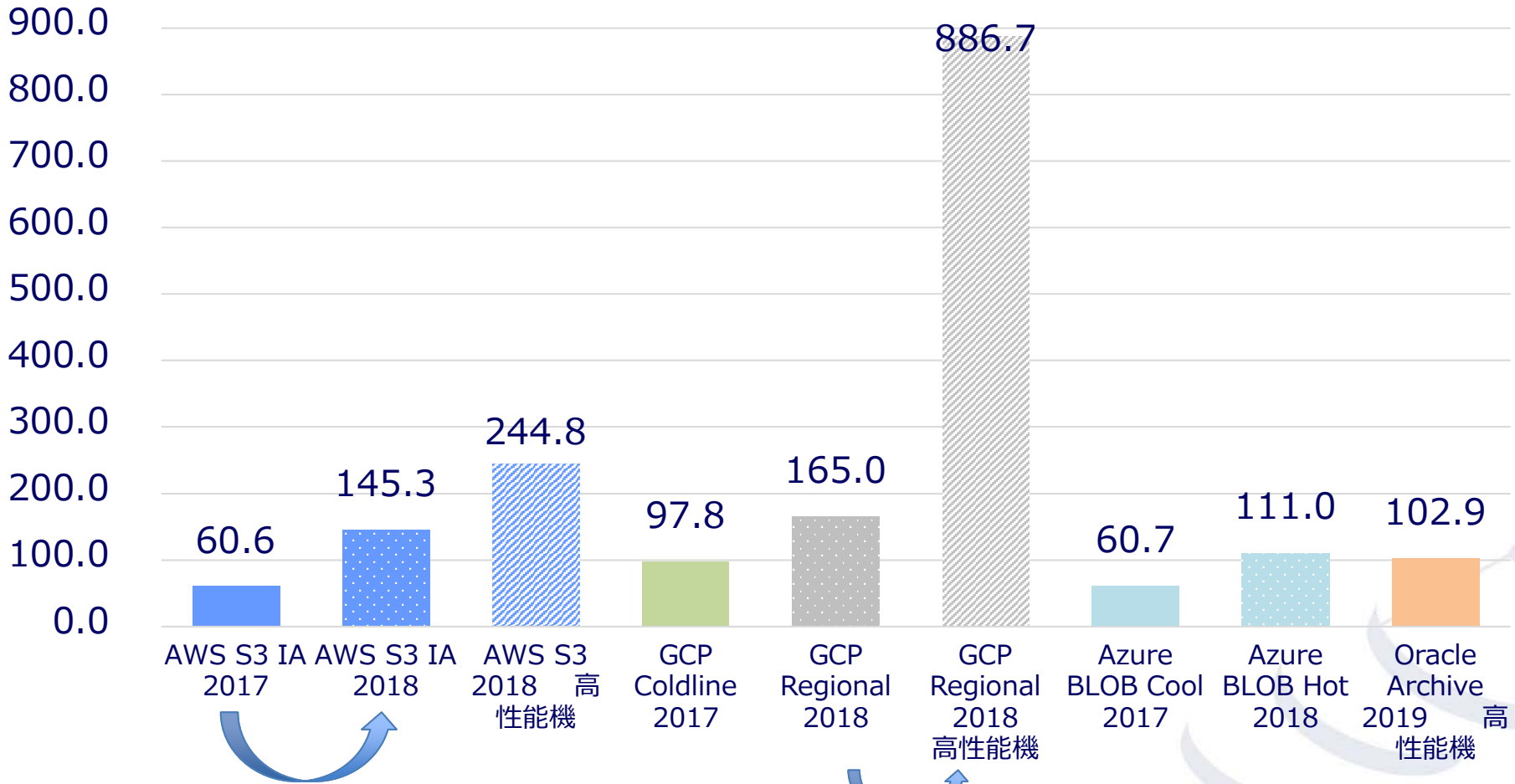
NIIオンプレミスサーバ仕様
通常機
Xeon E5-2650 2.0GHz 32コア
メモリ64GiB 10G ether

高性能機
Xeon E5-2660 2.6GHz 40コア
メモリ128GiB 10G ether

ALMAデータのアップロード性能

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

2017-2019年に測定
（単位MiB/s）



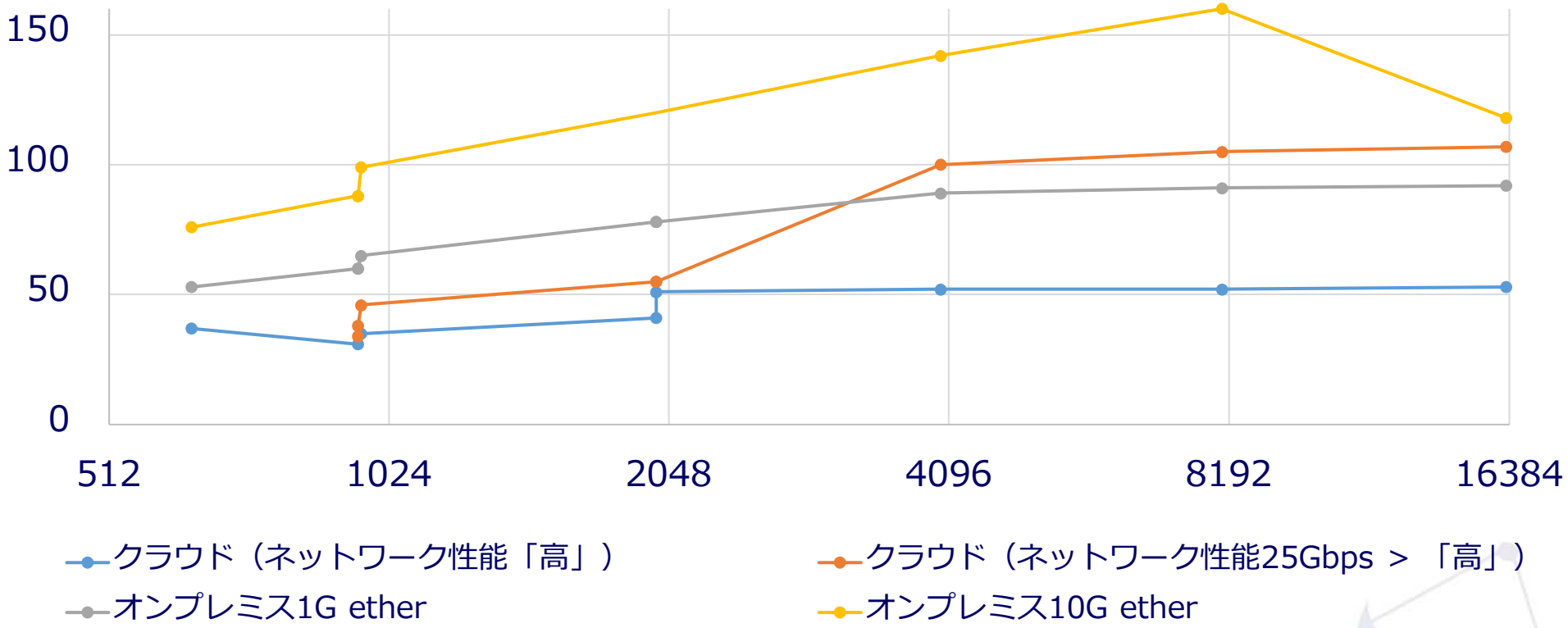
同じ条件でも年々高速になる

ローカル計算機が速いとアップロード性能上がる

ALMAアーカイブからのダウンロード性能

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

スループット(MiB/s)

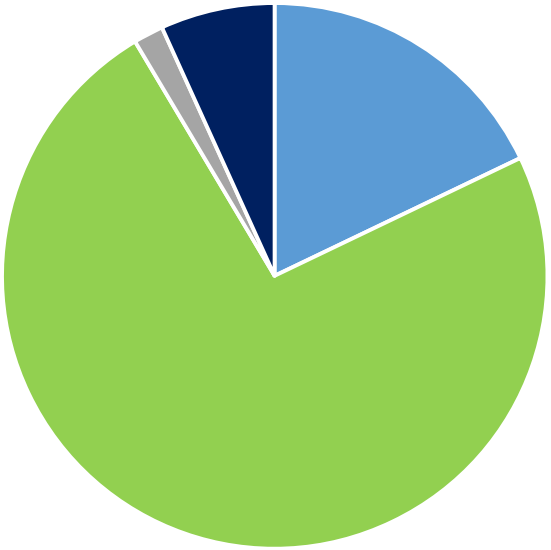


クラウド上のALMAアーカイブは、オンプレミスに比べて性能が低下している
→ストレージの性能差、オーバーヘッド、データベース性能差によるものと考えられる
ただし、インターネット経由の場合、その帯域でリミットされる

クラウドALMAアーカイブの運用費試算

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

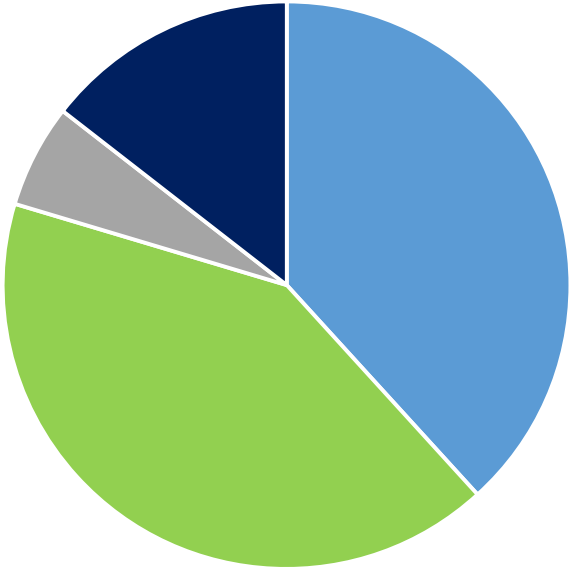
- S3 IAにアーカイブデータを格納
(500TiB、年間250TiBを読出し)



■ インスタンス/DB ■ s3 IAデータ保管
■ s3 IA読出し ■ データ転送

- 年間総運用コスト約$160,000-
- 保管コストが全体コストの2/3

- Glacierにアーカイブデータを格納
(500TiB、年間250TiBを読出し)



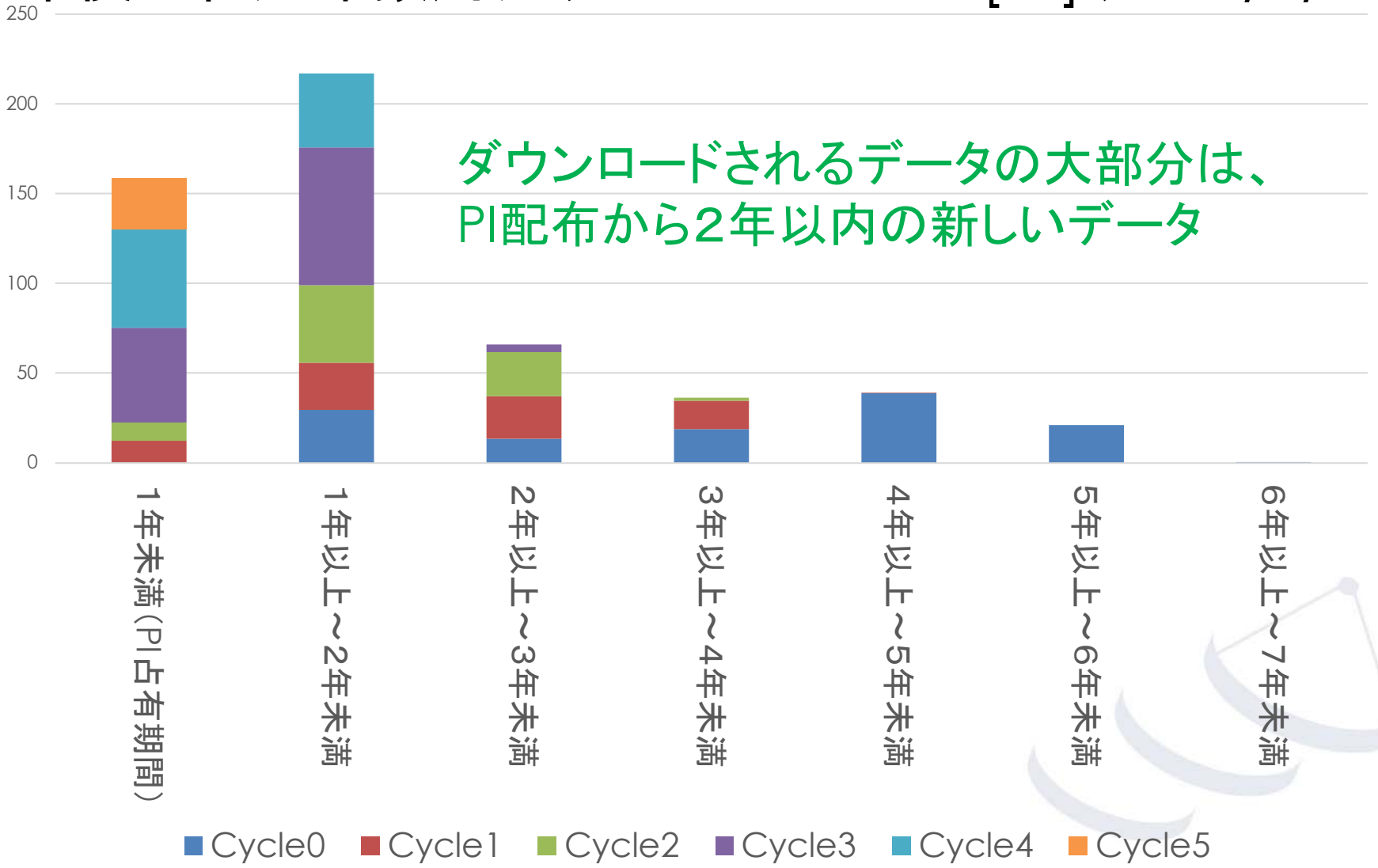
■ インスタンス/DB ■ Glacier データ保管
■ Glacier標準復元・読出し ■ データ転送

- 年間総運用コスト約$74,000-
- データアクセス前の復元時間
(標準で3.3時間)が常に必要
- データベースのインスタンスが高価

ALMAアーカイブデータのアクセスパターン

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

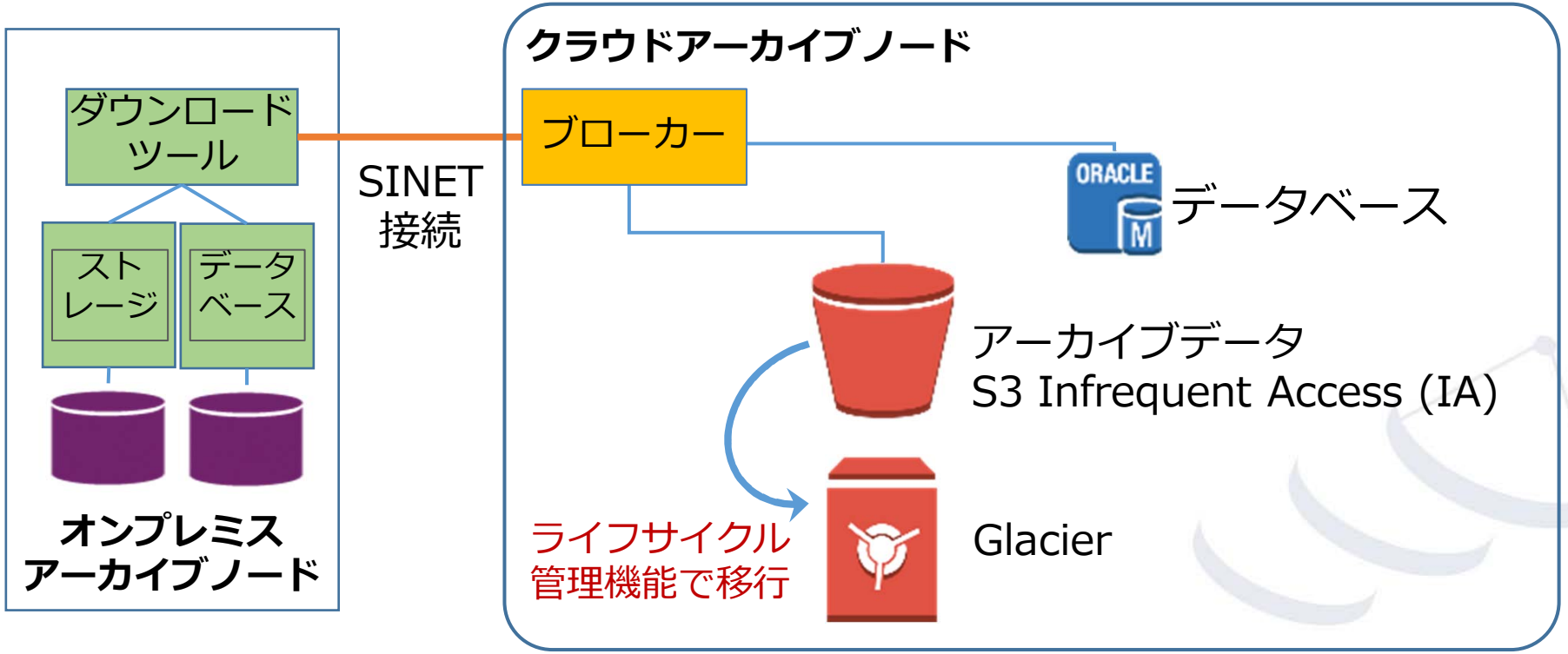
PI配布後の経過年数別ダウンロードサイズ [TB] (2018/5/23)



クラウドALMAアーカイブの最適化：ハイブリッド

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

- 新しくダウンロード頻度が高いデータはオンプレミスに置く（2年以内）
- 少し古いデータはクラウド上の即座にアクセスできる領域に置く
- より古いデータはクラウドコールドストレージに置く。ただし、データリストアに時間がかかる。



クラウドアーカイブはペイするか？試算

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

アーカイブデータ総量	1,000TiB
内　オンプレミスデータ量	500TiB
S3 IA上のデータ量	400TiB
Glacier上のデータ量	100TiB
年間ダウンロード量	250TiB
内　S3 IAからダウンロード	50TiB（20%）
Glacierからダウンロード	25TiB（10%）

クラウドアーカイブ年間運用費　約$130,000-

データベースをオンプレミスだけで賄えると約$110,000-

クラウドアーカイブのリスク

- クラウド業者がいつまでサービスを維持するか？
- 撤退時に別業者のサービスにデータ移行が（安価に）できるのか？
- 競争から寡占状態に移行したときに、価格が安く維持されるか？

ALMAデータの品質保証QA

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

- ALMA望遠鏡の理念
  - 電波天文学者以外でも、更に、天文研究者以外でも、ALMAのデータを利用して科学研究ができる
- そのために
  - 装置や望遠鏡固有の特性を除去し、データをすぐに使える状態まで整約
  - 観測提案者PIが求めるデータ品質を確認してからデータ配布：品質保証(Quality Assurance: QA)
- そんな解析済みデータをアーカイブ公開

QAオンプレミス解析環境

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

- 12パイプラインQA解析クラスターノード
  - 64～256GBメモリ
  - 1CPU with 4～6 物理cores (高クロック周波数)
  - 高速共有ファイルシステム（200TB程度）
- 12マニュアルQA解析計算機
  - 64GBメモリ
  - 1CPU with 4～6 物理cores (高クロック周波数)
  - 12～48TBローカル高速ファイルシステム
- ALMAアーカイブから直接生データを転送

QAクラウド解析環境(AWS)

- 複数の異なるインスタンスで処理性能と課金を比較
  - 8GB, 32GB, 61～64GB, 244GBメモリ
    - 但しメモリ割当量によって使えるCPUが異なる
  - 物理Coreの割当は4個
  - 中間データはクラウド上の高性能アクセス領域(EBS: Elastic Block Store)に置く（SSD或いはHDD）
  - 生データと解析済みデータはS3 Standardに置く

QAクラウド解析環境の価格（単位はUSD/月）

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

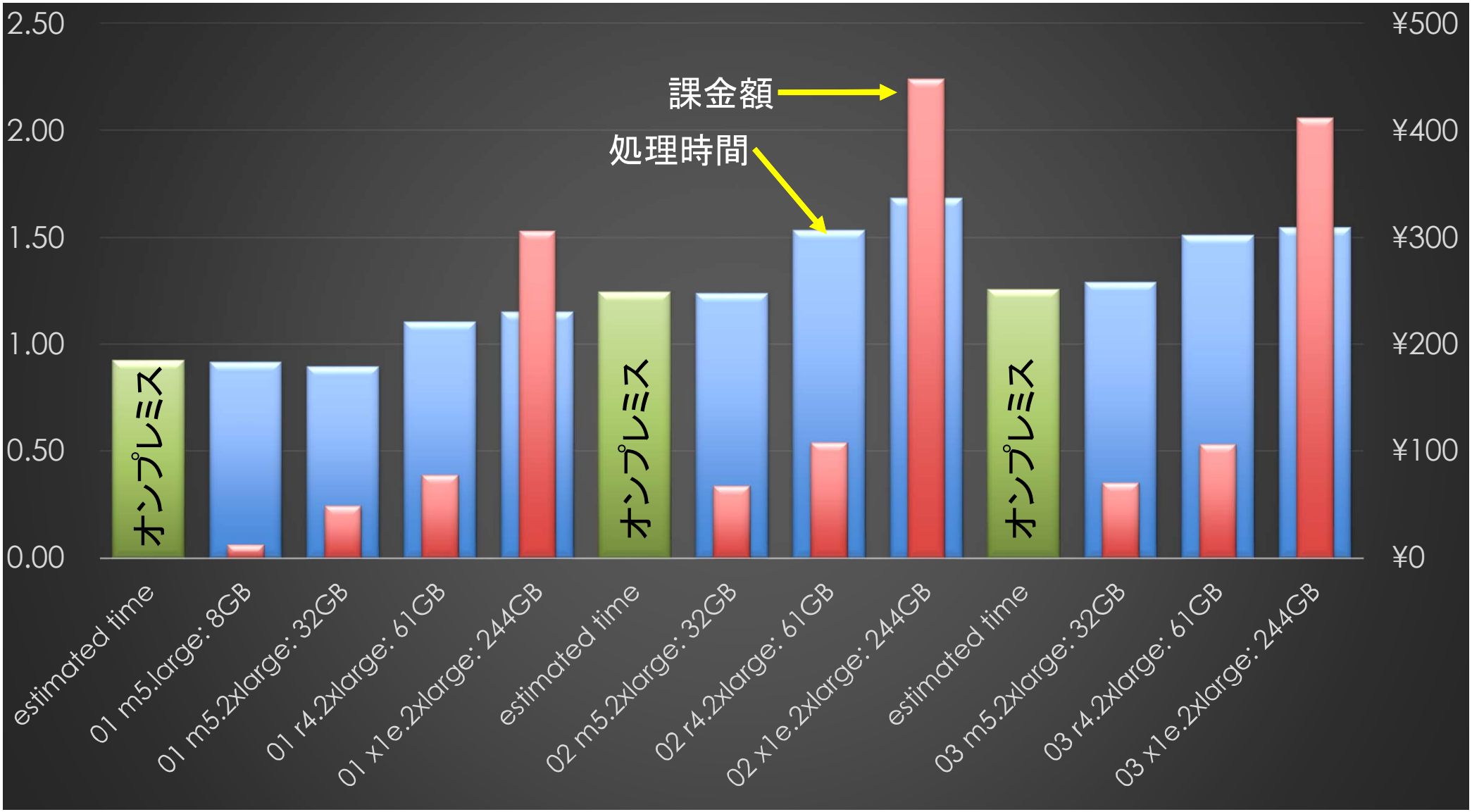
- 32GBメモリのインスタンス
  - m4.2xlarge（第４世代）物理4コア　約 $385-
  - m5.2xlarge（第５世代）物理4コア　約 $370-
- 61～64GBメモリのインスタンス
  - R4.2xlarge（第４世代）物理4コア　約 $480-
  - R5.2xlarge（第５世代）物理4コア　約 $455-
- 244GBメモリのインスタンス
  - x1e.2xlarge物理4コア　約 $1,800-
- EBS高性能アクセス領域
  - SSD 6TB　約 $720-
  - HDD 6TB　約 $320-
- 通常データ領域 S3 Standard 6TB　約 $150-

パイプライン処理用ALMA観測データ

- 小データ
  - サイズ：500～700MB
  - 典型的な処理時間：1時間程度
- 中データ
  - サイズ：3～5GB
  - 典型的な処理時間：5時間程度
- 大データ
  - サイズ：10～30GB
  - 典型的な処理時間：1日程度
- 特大データ
  - サイズ：～100GB
  - 典型的な処理時間：一週間程度

処理時間と課金額：小データ（処理時間１時間程度）

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

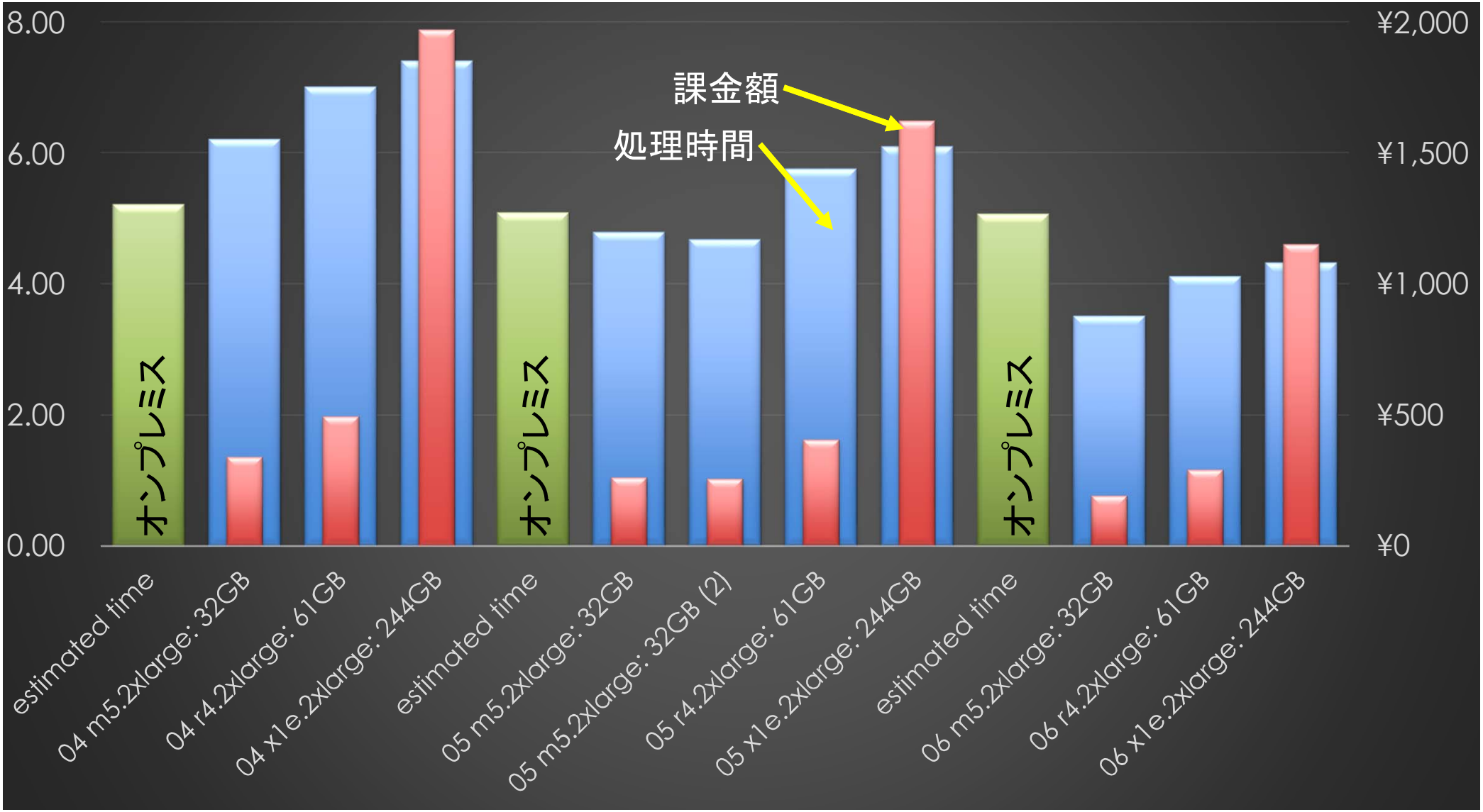


処理時間と課金額：中データ（処理時間5時間程度）

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

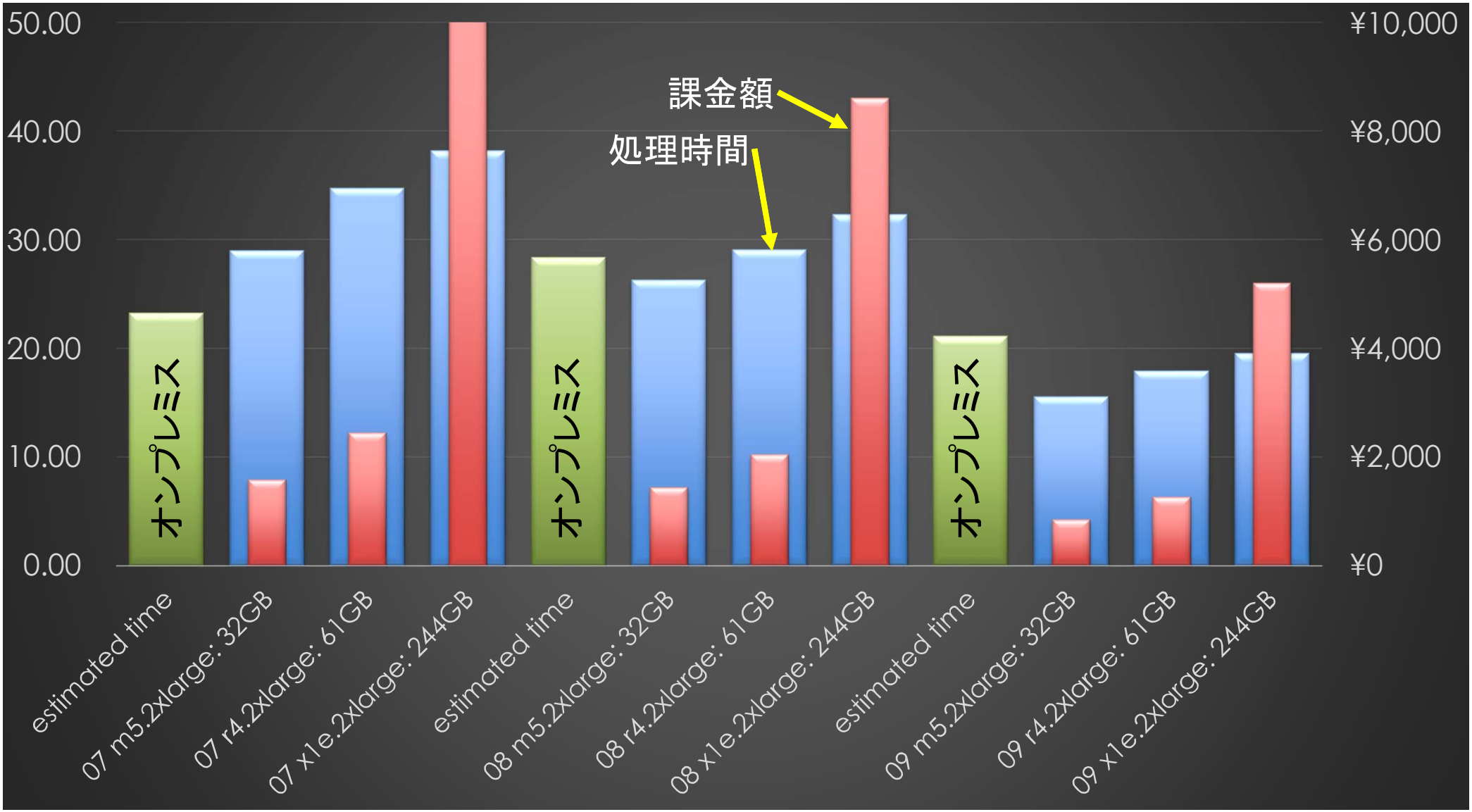
処理時間(h)

課金額(¥ ($1=¥110))



処理時間と課金額：大データ（処理時間１日程度）

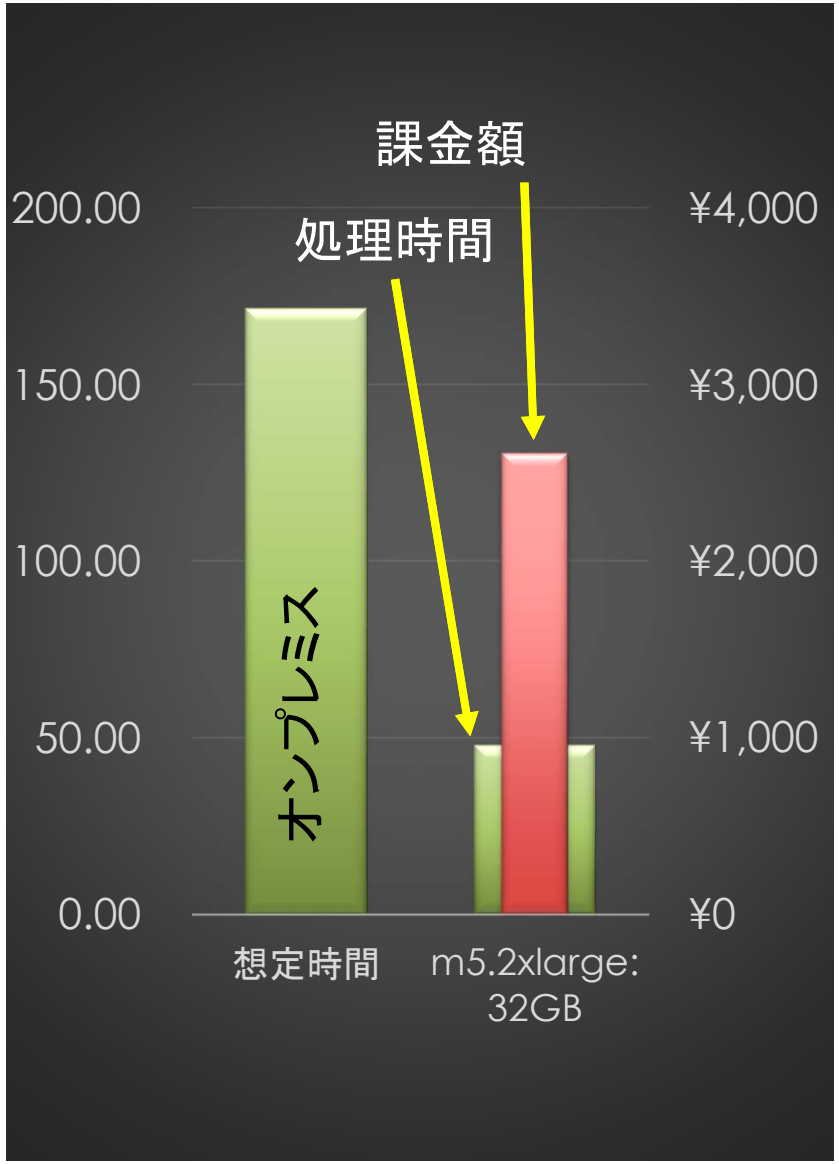
Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



処理時間と課金額：特大データ（処理時間１週間）

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

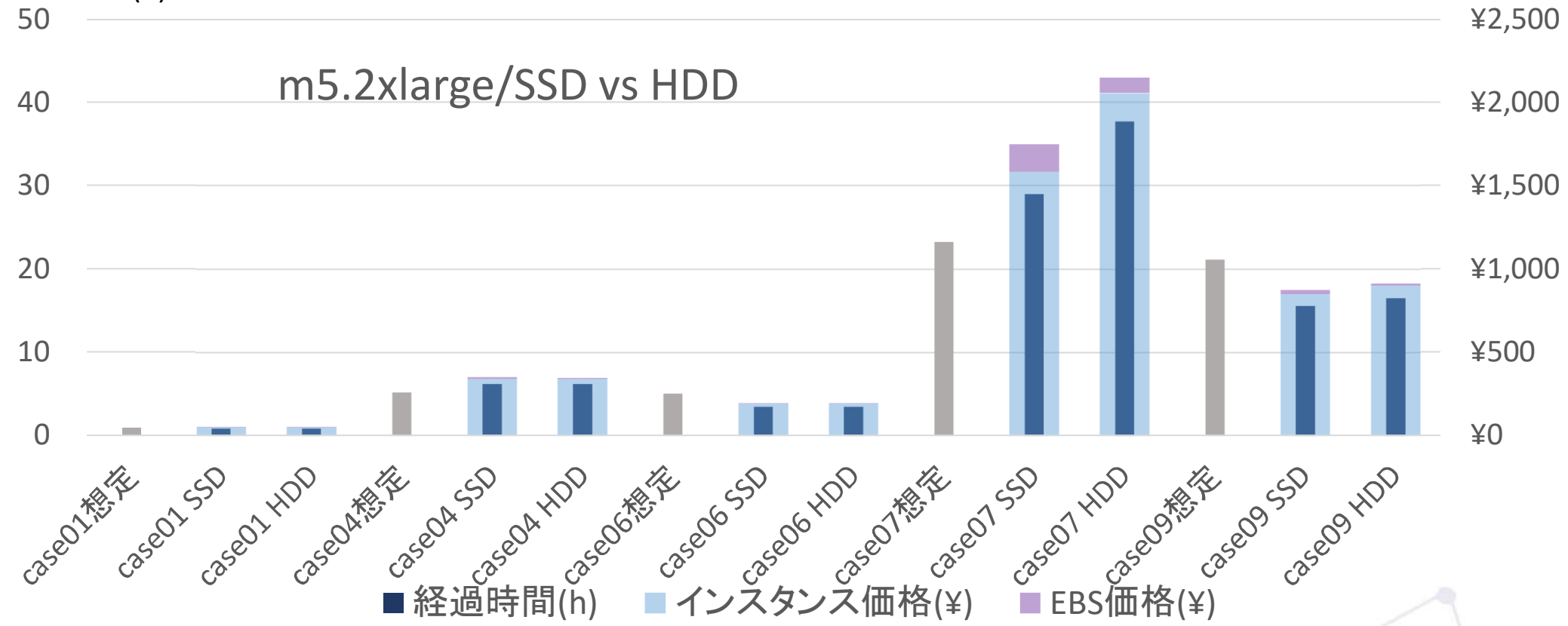
処理時間(h)　　課金額(¥ ($1=¥110))



大メモリのインスタンスは
高価なので32GBのみ使用

EBSの選択肢 SSD or HDD?

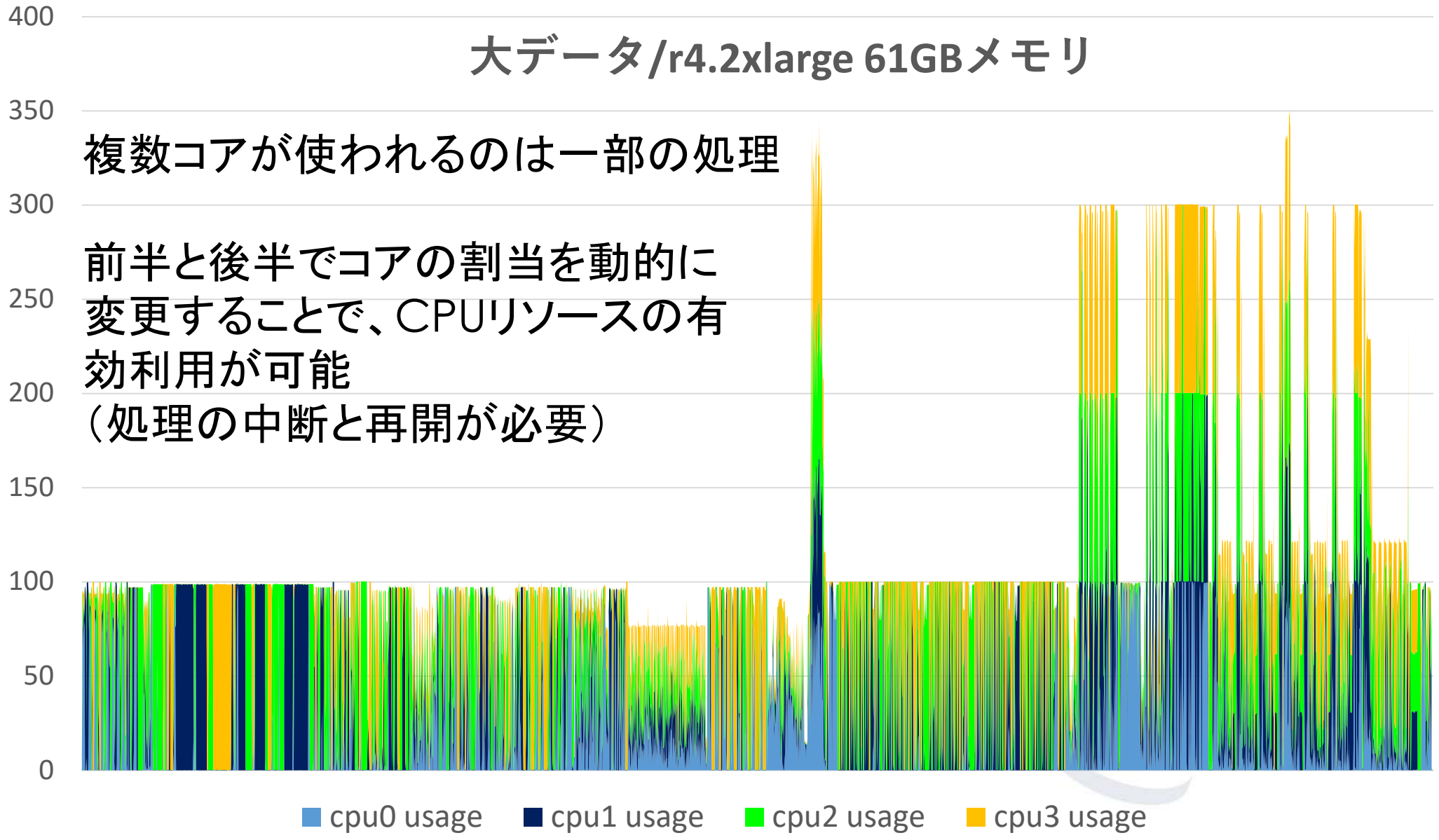
Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



SSDとHDDによる処理時間の差はデータ依存だが、処理時間に顕著な違いが見られる場合もあり、インスタンスの課金額が大きく異なる。
EBS(ストレージ)の課金額はインスタンスに比較して小さいため、
今のところ、ALMAデータ処理には常にSSDを用いるのが良い。

ALMAパイプライン処理のプロファイル（CPU）

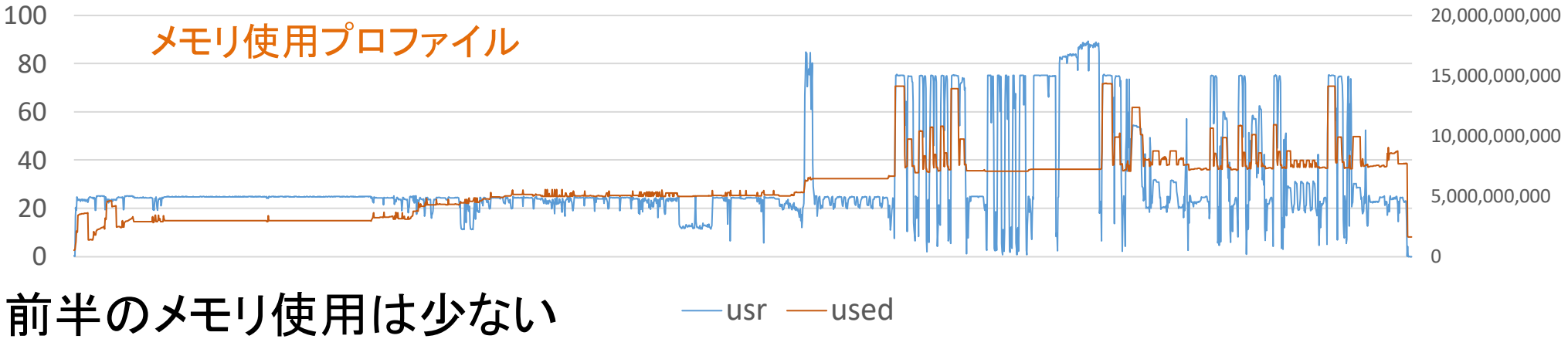
Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



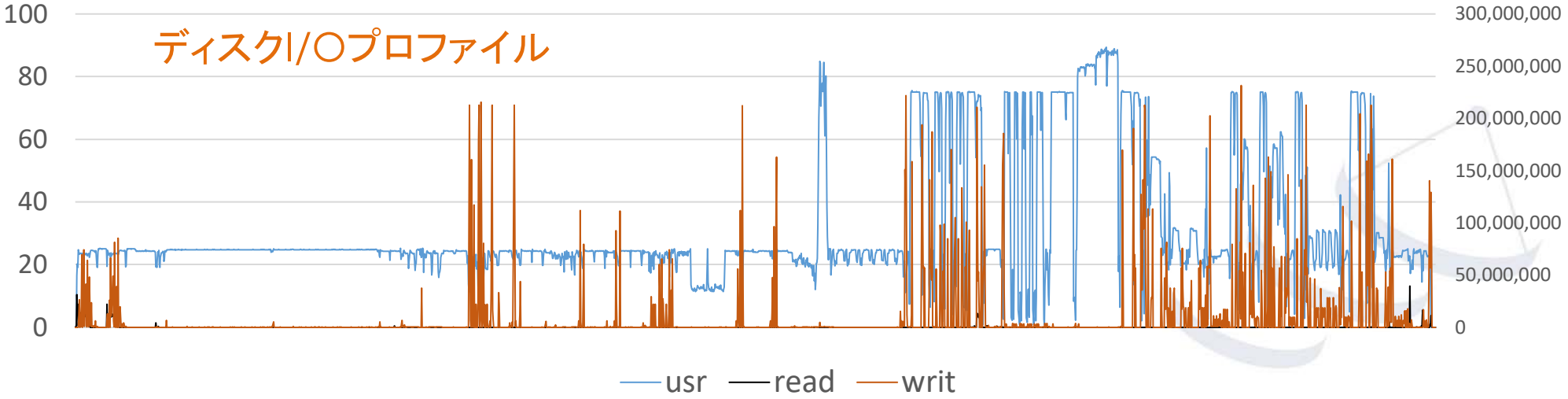
ALMAパイプライン処理のプロファイル（メモリ、ディスク）

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins

中データ / 61GBメモリ

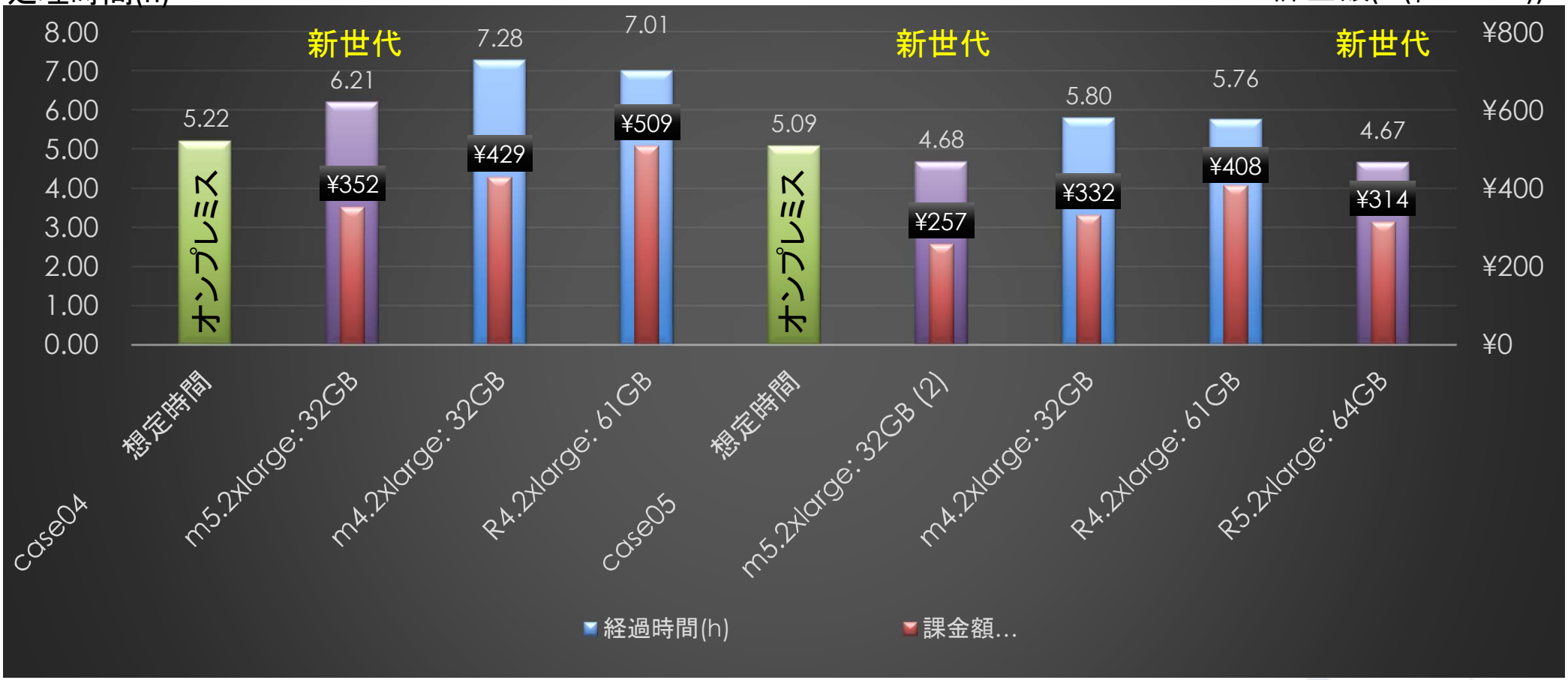


前半のメモリ使用は少ない
→ 前半後半でメモリ割当量を変更する
ことで、リソース使用を最適化可能



インスタンスの世代による違い

Atacama Large Millimeter/submillimeter Array
In search of our Cosmic Origins



新しい世代のインスタンスの方が単価が安く、また性能が高いために処理時間も短縮し、結果として課金額が抑えられる。

QAクラウド解析環境、今後の検討課題

- パイプライン処理に必要なリソースの予測
  - パイプラインは毎年バージョンアップする。毎年プロファイルを取る必要があるかもしれない。
- 最適リソース（CPU,コア,メモリ,ストレージ）の動的な割当の仕組み
- オンプレミスとの価格比較（電気代、ハードウェアメンテナンス人員コスト等を含める）
- クラウドALMAアーカイブとの接続も考慮した全体最適化

Questions?

Thank you…