ASGC Ceph 系統改善方案

- ASGC CephFS 系統採用 Erasure Coding 8+3 設計，利用分散式架構，容許同時最多損壞三顆硬碟，資料仍可復原，較 RAID5 or RAID6 保障更好，效率也較佳，但可用空間為 8/11。使用者除用於大數據處理分析、儲存研究資料外，亦可用於資料備份。
- 為提升儲存系統可靠度與效能，ASGC 正持續進行以下改善措施，以維護使用者資料安全與資料擷取效率。(大量資料之備份或系統自動修復期間，將暫時影響資料擷取效能)
  - 擴充空間以因應持續增加之需求，同時提升硬碟故障時自動修復等穩定運作之緩衝空間：增加 Ceph 空間 1.3PB (4 月底前已完成，0.8PB 為刪除 dark data), 並採購 1.3PB (五月底前到貨)。
  - 所有使用者資料均自動建立第二份備份：存於 ASGC 的使用者資料(已屬備份資料者除外)，將額外保存一份備份(共二份，此備份使用者無需付費)。此第二備份規劃使用異地磁帶系統(預計 2023.10 建置完成)。2023. 5 已開始實施，先使用 CephFS 空餘資源。
  - 增進 CephFS 系統的穩定：
    - 確保所有子系統均維持 8+3 的保護，同時提高系統自動平衡或修復等之效能
    - 增加整體備轉空間至 25%，提升硬碟故障等復原的效率與空間。
    - 設定各子系統空間使用上限為 75%，使用量達 50% 時將警示。降低因維持系統均衡穩定的資料搬移頻率。
    - 分析硬碟損壞與效能瓶頸等原因，維持硬碟每年更新五分之一的機制
  - Ceph 系統架構調整：為提升 Ceph 分散式系統效能，持續按規劃，改採小容量(磁碟容量與整座 storage server 規模都縮小)、大數量的 storage node，提升效能與容量，減低復原(failure recovery)成本