

## 添付資料 4 調査委員会報告書の概要

### 1. 調査委員会のメンバー構成

調査委員会は、山中卓（大阪大学大学院理学研究科教授、委員長）、坂本宏（東京大学素粒子物理国際研究センター教授）、堺井義秀（KEK 素粒子原子核研究所教授、Belle実験グループ共同代表者）、佐々木節（KEK 共通基盤研究施設計算科学センター教授）の4名で構成された。

### 2. 移行したデータの概要

1999年6月から2010年6月にかけて行われた Belle 実験の結果得られた実験データは、シミュレーションによって作成した疑似データ、ユーザー作成のデータとともに、KEKに設置された計算機システム（B計算機）に保存されていた。B計算機のデータ記録装置は、容量3000TBの磁気テープを主としたHSMシステムから成る大容量記憶装置であった。

### 3. データ移行作業の概要

B計算機のリース契約は2012年2月までであったため、ここに保存されていたデータを次期計算機システム（新中央計算機）に移行する必要が生じた。このため、2011年6月から翌年1月までに、B計算機にあるデータを一旦、旧共通計算機に移動し、さらに旧共通計算機のリース契約を数ヶ月延長して新中央計算機と並行して運用、その間にデータを旧共通計算機から新中央計算機に移動する一連の作業が行われた。

### 4. データ損失の経緯と原因

B計算機から旧共通計算機へのデータ移行の過程で、データの一部がコピーされず、さらにその一部は回復不可能であるという事態が発生した。HSMシステムにおかれていた全データについては29%、実験で収集した生データについては18%がコピーされなかったが、復旧が行われた結果、生データの5%が最終的に失われた。物理解析で使用するために抽出した物理データは国内外3箇所（KEK、名古屋大学、米国PNNL研究所）にもコピーが存在し本質的な部分での損失はなかったが、物理データのうち、ハドロン共鳴の研究に用いられる特殊な物理データの一部はKEK外にコピーされていなかったため、失われた。疑似データなどのその他のデータも一部失われたが、作り直すことができたものであった。

データ損失の直接の原因は移行すべきファイルのリストに間違いがあったためである。この過程に十分なチェックが行われるような体制をとっていれば防ぐことができたはずであるが、Belleグループ、素核研および計算科学センターのデータ移行プロジェクトに対する体制の不備などにより、十分な措置がされず事案の発生に至った。具体的な技術的問題として、作業に用いられたスクリプトの動作結果の確認が不十分だったことがあったが、加

えて、プロジェクトの規模に比して作業人数が少なかったこと、移行すべきファイルのリストを作る段階についてクロスチェックされていなかったなどの問題が指摘された。さらにそれらの背景として、データ移行の期間が約7ヶ月と短く、計算機関係の人員不足があったことが推察された。今後の対策としては、重要な作業に関しては、部内における作業担当者以外のチェックが必要であること、プロジェクトの遂行に十分な人員と時間を充てること、より一層データの保護の重要性を認識すべきであることなどが指摘された。

---

<sup>i</sup> HSM (Hierarchical Storage Management) システムとは階層型ストレージ管理システムの略こと。参照頻度が高いデータは一次媒体、参照頻度が低いデータは二次媒体に置かれ、仮想ファイルシステムが組まれる。この仕組みを自動化したシステム。