

JLDG/ILDG講習会・利用検討会

ILDGの概要と利用法

天笠俊之(筑波大学計算科学研究センター)

内容

1. ILDG構築の経緯と現在の状況
2. ILDGのデータ, メタデータ (QCDml)
3. ILDGのハードウェア, ミドルウェア
 - ▶ MDC, FC, SE, VOMS
4. ファセット検索インタフェースを利用したensemble検索
5. ユーザ登録の手順
6. クライアントパッケージ (ildg-get, ildg_cksum, LIME)
 - ▶ インストール方法
 - ▶ 利用法
7. ファイルフォーマット

ILDG (Int'l Lattice Data Grid)について

▶ ILDGとは

- ▶ 計算素粒子物理学分野のLattice QCD計算の計算結果である配位データを国際的に共有するためのデータグリッド

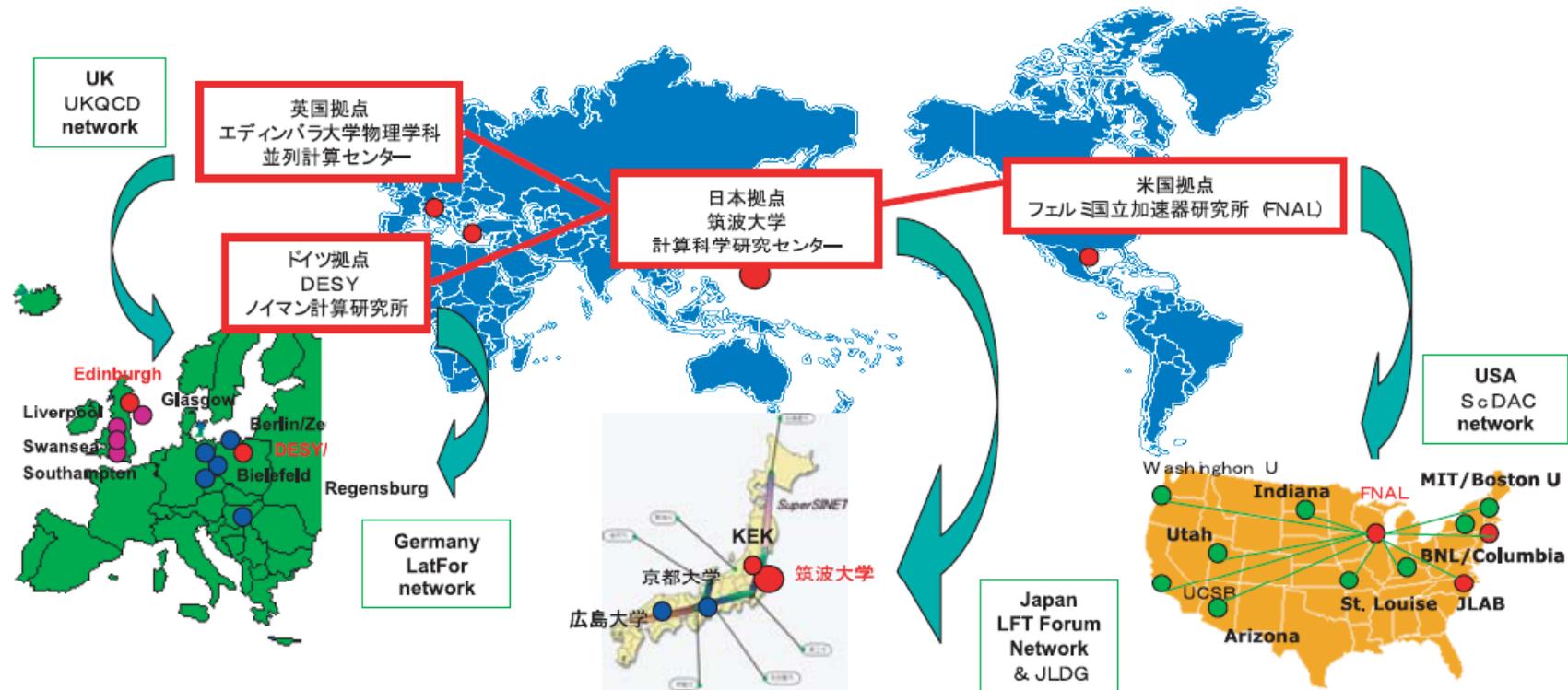
▶ 目的

- ▶ 2002年 R. Kenway (EPCC, Edinburg) により提案
- ▶ 過去に計算された配位データ(ensemble, configuration)の国際的な共有基盤
 - ▶ 格納, 検索, 取得

▶ 活動内容

- ▶ ILDG Board
- ▶ 二つのWG
 - ▶ ミドルウェアWG
 - ▶ メタデータWG
- ▶ 年2回のワークショップ (ILDG15, Dec. 4, 2009)

ILDGの主要拠点



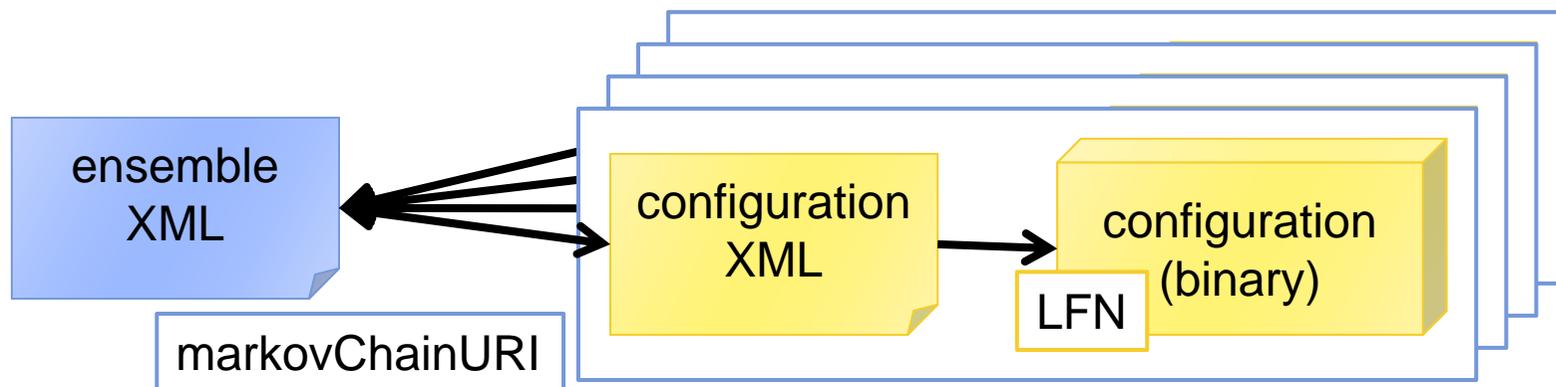
CSSM, Adelaide, Australia
 JLDG, Tsukuba, Japan
 LDG, DESY, Germany / France / Italy
 UKQCD, Edinburgh, UK
 USQCD, Fermilab/JLab, USA

ILDGの基本的な概念

- ▶ 共有する配位データの設計と合意
 - ▶ ファイルフォーマット
 - ▶ アクセスの仕組み
- ▶ 各地域グリッドを連携させる仕組みの設計と合意
 - ▶ 相互接続のためのインタフェース (Grid-of-Grids)
 - ▶ 2階層からなるデータグリッド
 - ▶ 地域グリッド
 - 研究グループにおけるデータ共有により, 研究リソースを有効活用
 - ▶ 国際グリッド
 - 世界規模での配位データの共有により, 分野全体の進歩を加速
 - ▶ Webサービス技術の利用

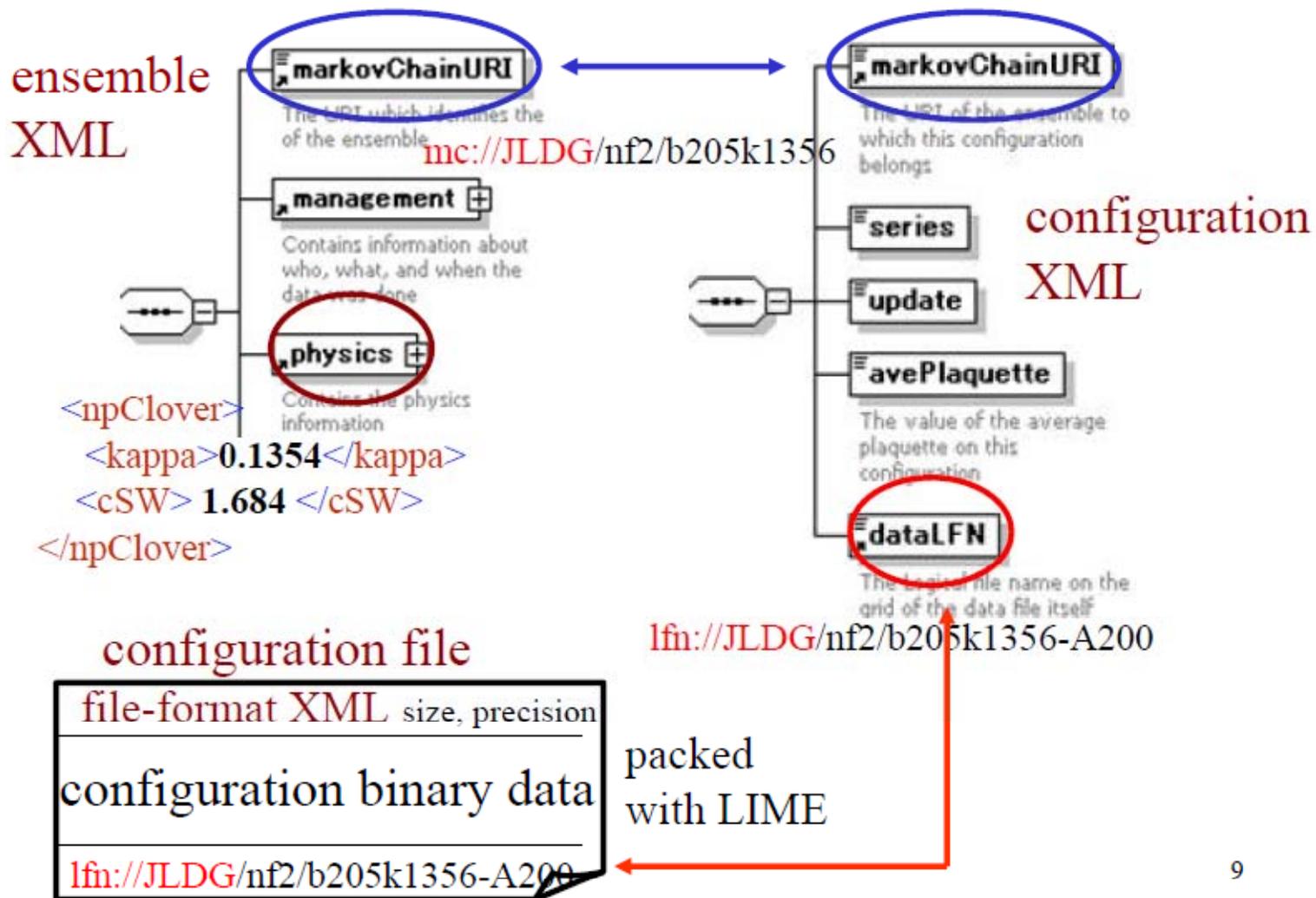
ファイルフォーマット

- ▶ 配位データのバイナリフォーマット
 - ▶ LIME (Lattice QCD Interchange Message Encapsulation)
- ▶ メタデータ (QCDml)
 - ▶ データがどのような内容であることを記述したXMLデータ
 - ▶ ensemble XML: どのような物理シミュレーションか
 - ▶ configuration XML: どのようなパラメータ(系列)のデータか
 - ▶ LFN (Logical File Name)
 - ▶ 配位データのバイナリファイルの識別子



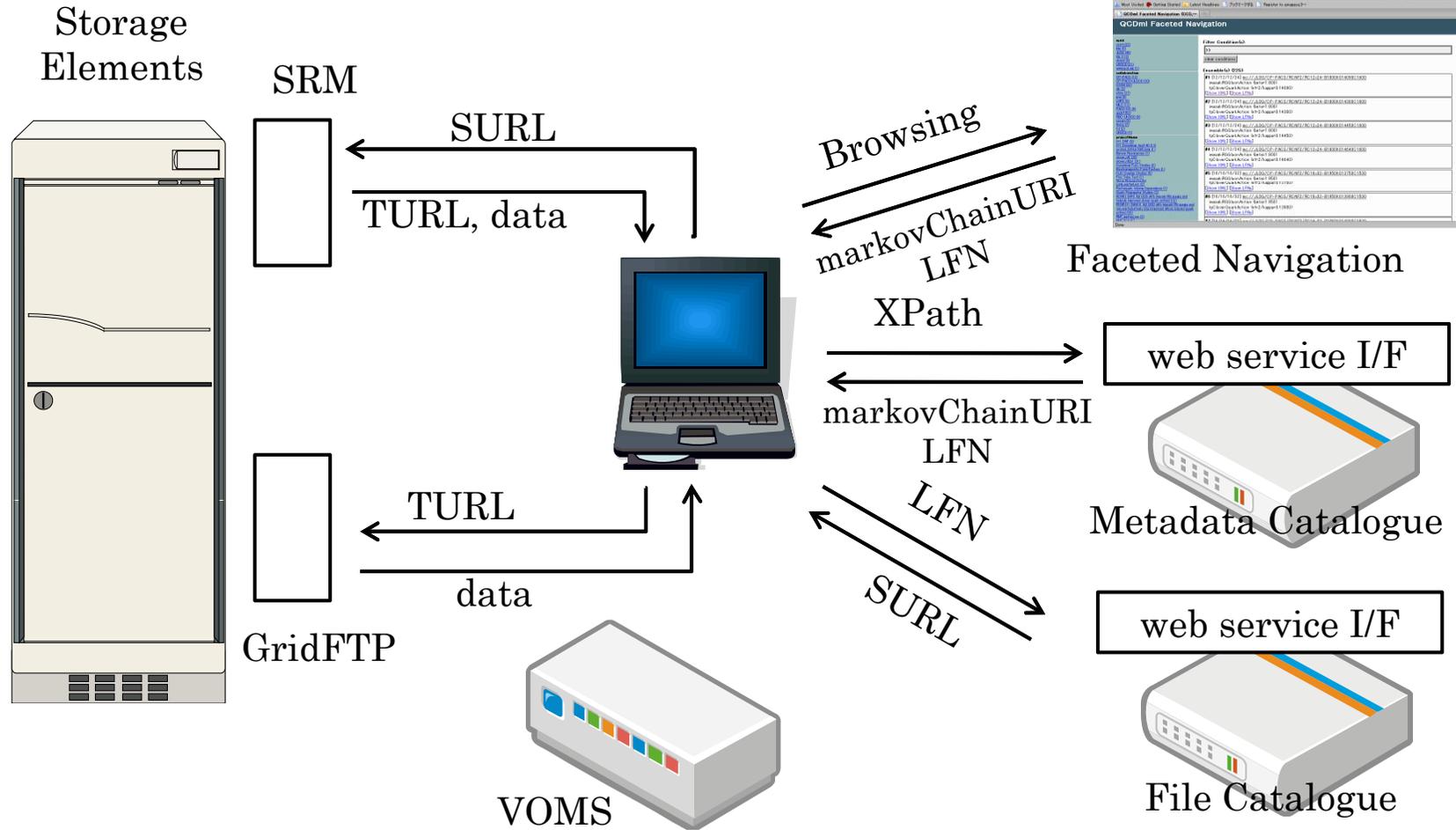
QCDmlの構造

ensemble XML / configuration XML

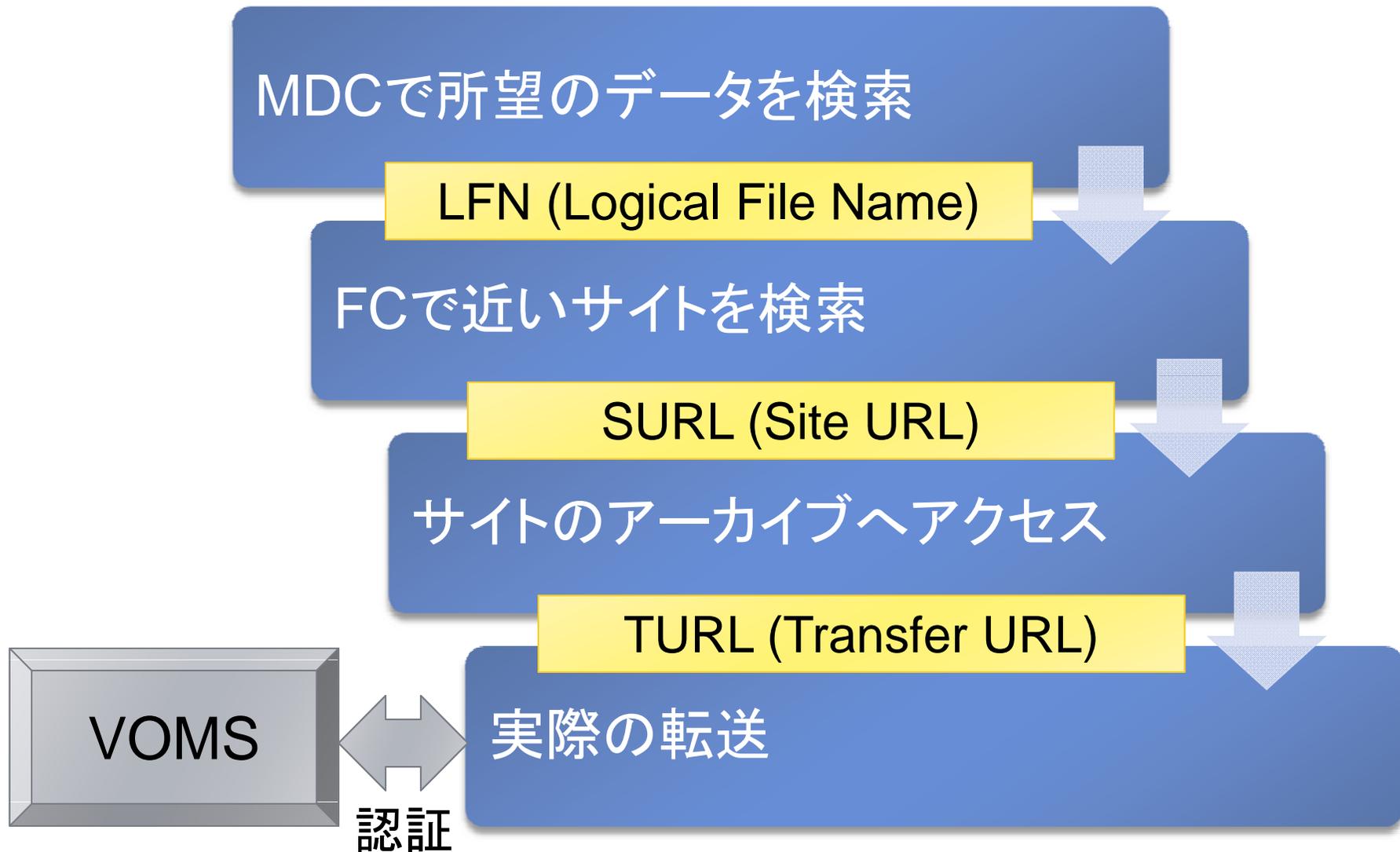


ILDGの構成要素

Middleware Components



ILDGの処理フロー

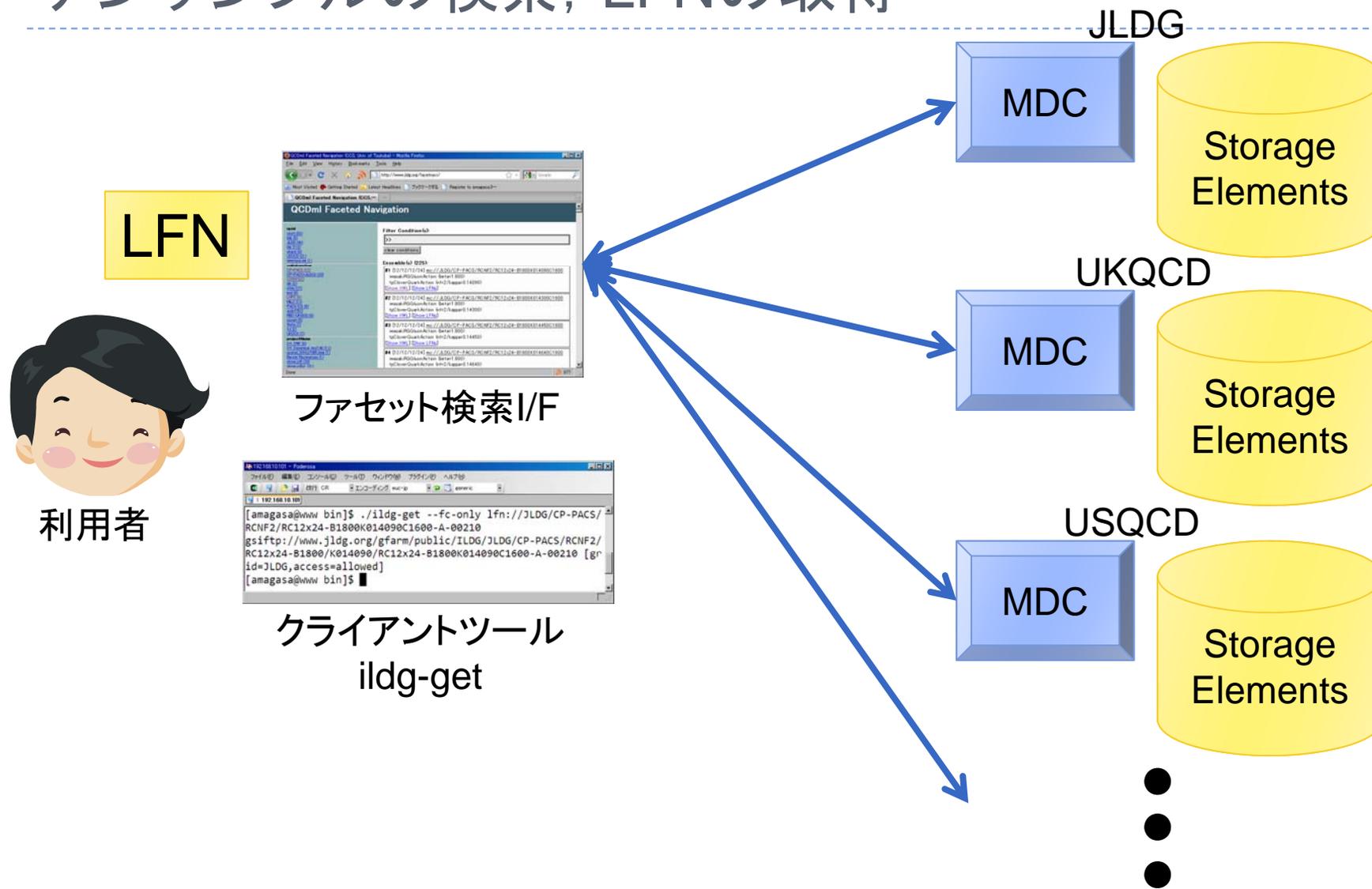


ILDGの構成要素 まとめ

- ▶ MDC (Metadata Catalogue)
 - ▶ 配位データのメタデータ(QCDml)を格納したデータベース
 - ▶ LFN (Logical File Name): 配位データの識別子
- ▶ FC (File Catalogue)
 - ▶ 配位データの複製を管理
 - ▶ SURL (Site URL): 複製を持つサイトの識別子
- ▶ SRM (Storage Resource Manager)
 - ▶ 各サイトのデバイスへのアクセス方法を管理
- ▶ データ転送レイヤ
 - ▶ 実際の転送を行うレイヤ (gftp, ftp, http, ...)
 - ▶ TURL (Transfer URL): 配位データへのアクセスURL
- ▶ VOMS (Virtual Organization Membership Service)
 - ▶ アクセスされるデータへの認証を管理

ILDGの利用 (1)

アンサンブルの検索, LFNの取得

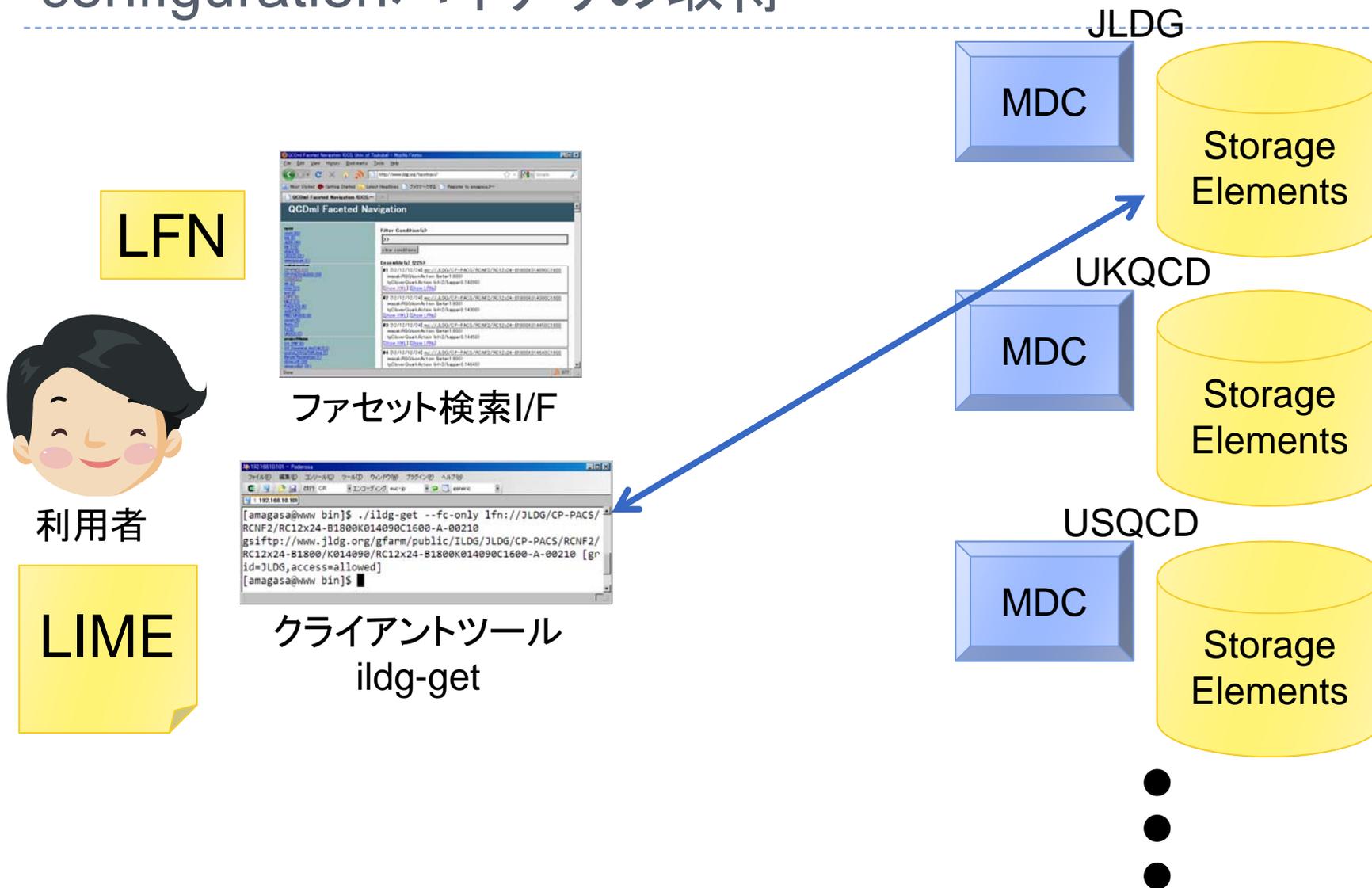


QCDmlファセット探索インタフェース

<http://www.jldg.org/facetnavi/>

- ▶ **ファセット検索とは？**
 - ▶ オブジェクト検索の一手法
 - ▶ 近年多く利用されるようになった(例: Apple iTunes等)
- ▶ **手順**
 - ▶ ファセットを選び, 一つ(複数の)ノードを選択
 - ▶ オブジェクトの絞り込み
 - ▶ 絞り込み条件に応じて, オブジェクトのリストを更新
 - ▶ (繰り返し)
- ▶ **利点**
 - ▶ どの属性を使って絞り込むかは, 利用者が自由にできる
 - ▶ データ管理コストの低減化
 - ▶ オブジェクトの追加が大域的な変化を与えない
 - ▶ 大量のオブジェクトを効率的に分類
 - ▶ Busch's Law
 - 10,000オブジェクトの分類には, 10ノードからなる四つのファセットで十分

ILDGの利用 (2) configurationバイナリの取得



ここまでのまとめ

▶ ILDGの概要

- ▶ 経緯
- ▶ 構成要素
 - ▶ MDC, FC, SRM, VOMS, ...
- ▶ ファイルフォーマット
 - ▶ LIME, ensemble XML, configuration XML
- ▶ ファセット検索インタフェース

▶ この後

- ▶ クライアントツールによる配位データの取得
- ▶ コマンドラインツールを用いた配位データの操作
- ▶ ファイルフォーマットの解説

ご静聴ありがとうございました

質問は amagasa@cs.tsukuba.ac.jp まで