



Enabling Grids for E-scienceE

Data Management for Grid Environment

Marek Ciglan

Institute of Informatics, Slovak Academy of Sciences



www.eu-egEE.org



INFSO-RI-508833



Data management tasks

- **Safe and reliable file transfer between grid nodes**
 - GridFTP
- **Mechanism for locating data sources in the grid**
 - Replica Location Service
- **Mechanism for data content identification**
 - Metadata services



GridFTP – safe and reliable file transfer protocol

- **Based on FTP protocol**

FTP

- **Widely used**
- **Well-understood standard protocol**
- **Well-defined architecture for protocol extensions**
- **Transfers between client and server support**
- **Third party transfers between two servers support**

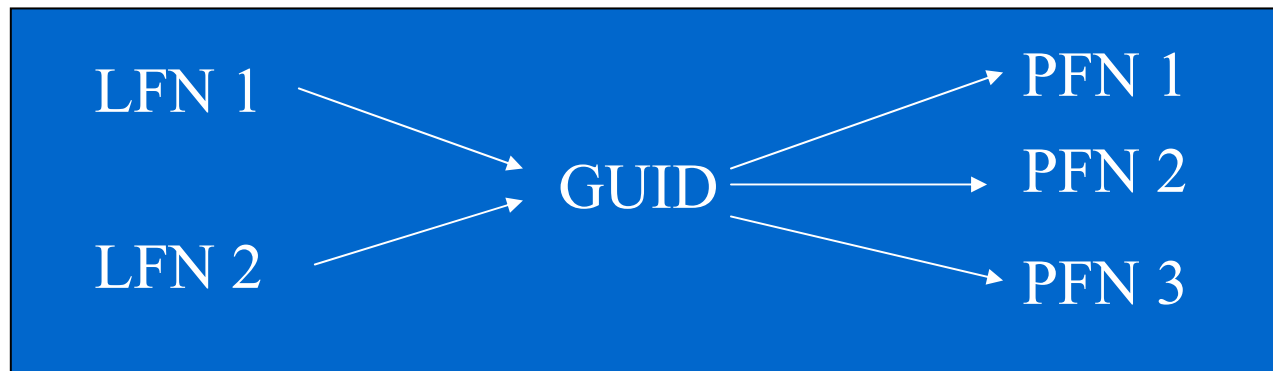


GridFTP - extension of FTP

- **Grid Security Infrastructure (GSI) support**
- **Third-party control of data transfer**
- **Parallel data transfer**
- **Partial file transfer**
- **Striped data transport**
- **Support for reliable data transfer**



- **Provide information about replica location**
- **Replica Naming**
 - LFN – Logical File Name
 - GUID – Grid Unique Identifier
 - PFN – Physical file name



- **Stores information about files stored at the grid node**
- **Provides mapping between LFN and PFN**



- **descriptive information about data**
- **lot of metadata types**
 - general metadata (file type, file size, creation date)
 - security metadata
 - application specific metadata



```
AF01-BE1227-3318A1  
CC58-48FA03-0054A0  
0322-BB1829-101504
```





Thank you for your attention.