



# **CESNET e-Infrastructure**

## **Storage services – vision and plans**

Peter Verčimák  
([peter.vercimak@cesnet.cz](mailto:peter.vercimak@cesnet.cz))

CESNET, Prague  
Czech Republic

## Czech national e-Infrastructure projects

- CESNET
- IT4Innovations
- CERIT-SC

Potential customers of new e-Infrastructure services including storage:

*ESFRI roadmap projects and other national and international R&D groups and projects e.g....*

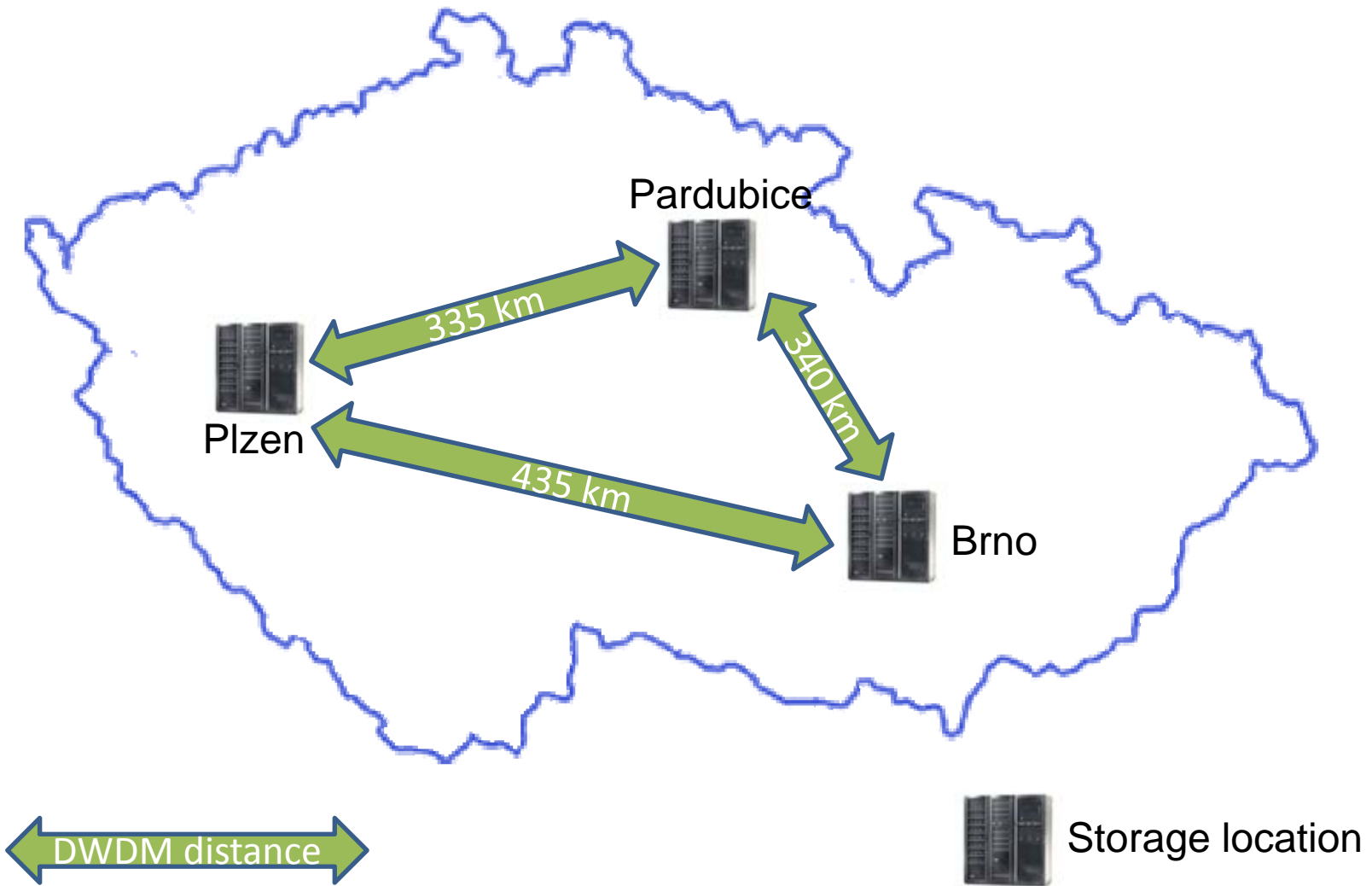
BIOMEDREG, CzechPolar, CZERA, CzechGeo/EPOS, ESS, CANAM, FAIR, ALICE, ELI, PALS, UCNK, CESSDA, ESS-survey, Reaktory Řež, LINDAT/CLARIN, ThALES, SHARE, CzechCOS/ICOS ...

Building of CESNET e-Infrastructure – two projects:

1. The “OP VaVpl” project – in frame of EU Operational Program for R&D (2011-2013)
2. The “Great Infrastructure” project – in frame of national ESFRI Roadmap (2011-2015...)

The goal of both projects in storage area is common – **to build up and to put into service the storage system of three distributed large-scale repositories for saving and sharing of large volume of data including archiving**

# CESNET Distributed Data Repository



The main purpose of the CESNET “national” storage system: easily accessible and redundant data repository for academic and scientific community.

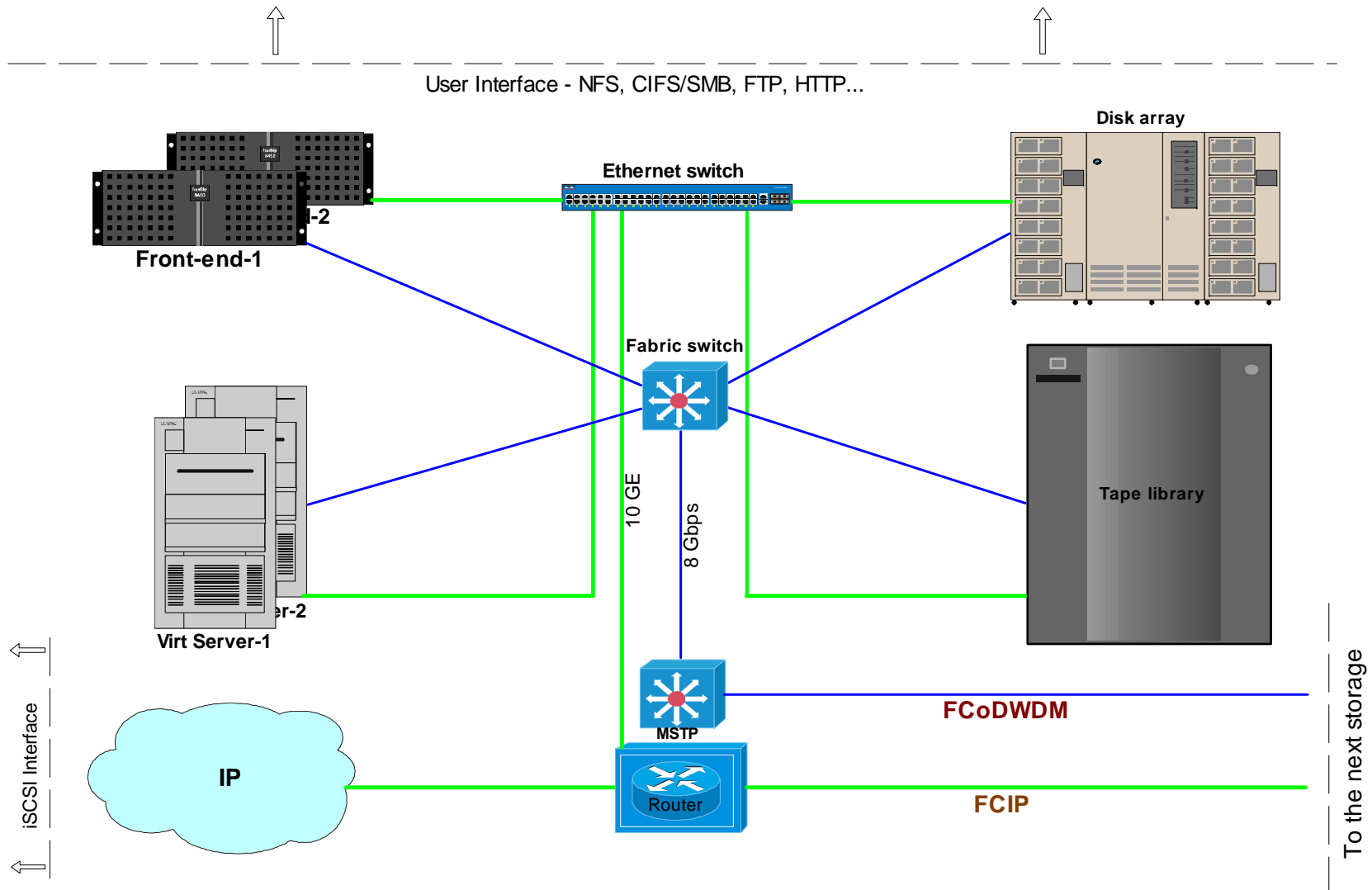
- From user point of view: “unlimited storage capacities”, otherwise overall capacity - about 10 PB at the three locations
- Technical concept: HSM system – composed of disk arrays and tape libraries (or equivalent systems – MAID, VTL and the like)
- From data access method and communication protocol point of view – combined NAS/SAN system

### **Data repository services:**

- Storage element
- File system services (including backup)
- Data-block access services

- **Storage element** – data repository oriented to capacity
  - Performance is important but not critical parameter
  - Accessible via protocols for large volume data transfer – gridFTP (or just via common scp, ftp, rsync,...)
  - Locally connected file system
- **File system services** – from HSM point of view again locally connected FS but through NFS or SMB protocols
  - Could be used for backup services, too – both just as raw capacity for “customers administrated” backup or as back-end repository providing the server part of backup SW on the storage system
- **Data-block access services** – limited size of customers would have possibility to access their data via iSCSI or even FC

# Simplified scheme of data repository in one storage location



## Datacenter composition

- Disk array (Tier1, Tier2) – about 300 TB
  - Faster (15k FC, SAS) and slower (7k SATA) disk tiers
- Tape library (Tier3) – about 3 PB
- Fabric and Ethernet network infrastructure
  - Redundant SAN (8Gb) and LAN (10GE) switches and lines
- User interface (Front-end) servers
- Virtualization farm (hypervisor platform)
  - The servers for application support of specific users' OS and other (application SW) requirements

- Management system (SW and its platform according to proprietary vendors' solution)
  - HSM system
  - NAS Heads
  - Metadata server
  
- Superstructural application servers
  - Platforms for data administration systems (iRODS) or advanced storage services (webDAV, ePosteRestante)
  
- Replication interface components
  - Dedicated lines between storage locations will use FCoDWDM or FCIP to ensure backup for DR purposes





**Thanks for your attention.**

**Any questions? Not required, but tolerable...**