Potential usage of Bittorrent for sharing weather data within OpenWIS

Openswis Association meeting

30 May 2018 – FMI, Helsinki, Finland

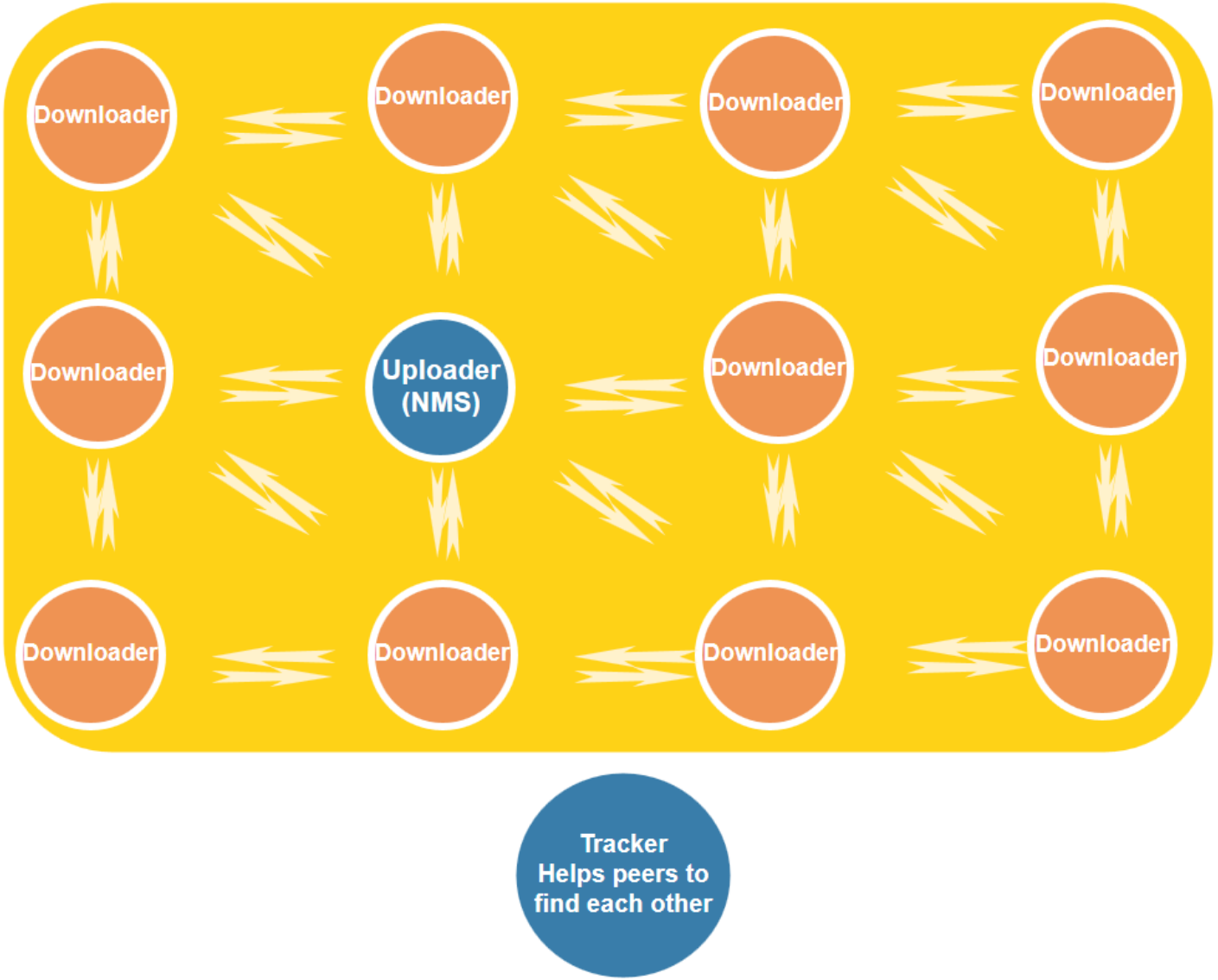
david.podeur@meteo.fr

Météo-France

Alternative Data Dissemination Methods:
“Custom” vs “Bulk” data delivery

- Make large amounts of data reach end users
- Two strategies:
  - Custom delivery:
    - Allow users to specify custom datasets
      - Specific parameters: humidiy at 850Hpa
      - Specific domains: e.g. [10N, 20S, 50W, 80W]
      - Specific timeranges
    - Techniques on server side: OGC Webservices + underlying layers
      => Heavy development costs & processing + Maintenance
  - Bulk delivery:
    - Allow users to get whole raw datasets
    - Associated techniques / protocols:
      - HTTP/FTP file delivery
      - Cloud delivery: e.g. Amazon S3
      - Peer to peer delivery: e.g. Bittorrent
    - Possible to allow some custom delivery mechanisms

Alternative Data Dissemination Methods:
Focus on Bittorrent data dissemination



- Downloader are also Uploaders
- Bandwidth and CPU usage is limited at NMS
- One tracker needs to be implemented
- Files remain available until all Downloaders stop sharing (possible long lifetime)
- Somes downloads do not affect initial uploader at all

Alternative Data Dissemination Methods:
Focus on Bittorrent data dissemination

- In order to start a download, a bandwidth client needs:
  - A torrent file: e.g MeteoFrance_AROME01_20180422_0600.torrent
  - Size: ~ 20KB
- With this file, a bittorrent client is able to:
  - Connect to the specified tracker server and know which peers are already sharing the file
    - Tracker example: bittorrent.meteo.fr:6969
  - Initiate connection to several other peers to download several file pieces simultaneously.
  - Inform the tracker about the file parts already downloaded (and available for others).

    Note: It is common to download from 20+ peers at a given time.
- The Tracker:
  - Serves list of peers having the requesting file pieces
  - Is able to favour peers having a high upload/download ratio (in order to encourage sharing): **“the more you share, the faster you download.”**
  - **Should OpenWIS use bittorrent as a dissemination method, it will just need to provide the user with the corresponding bittorrent file.**

Alternative Data Dissemination Methods:
Focus on Bittorrent data dissemination

- Keep in mind:
  - Bittorrent in itself is legal
  - However, many illegal usage of bittorrent to distribute copyrighted content:
    - improved the protocol
    - provided very stable and robust implementations
  - It is widely used:
    - By NASA (Visible Earth program)
    - By Facebook & Twitter (Distribute large files between servers)
    - Linux ISOs distribution

Alternative Data Dissemination Methods:
“Custom” vs “Bulk” data delivery

	HTTP / FTP Delivery	Cloud delivey	Peer to Peer	OGC Webservices & REST
Description	Files are hosted in a single server	Files are sent once to one “cloud” service, which takes care of final delivery	Files are sent using bit torrent protocol. Users are both downloading and uploading	Data is processed and only relevand parameters are delivered to user
Infrastructure costs	**€€**	**€?**	**€**	**€€€**
Telecom costs	**€€€**	**€?**	**€**	**€€**
Set up costs	**€**	**€€**	**€**	**€€€**
Maintenance costs	**€€**	**€€**	**€**	**€€€€**
Custom queries	**limited**	**depends**	**limited**	**large possibilities**
User access control	**limited**	**limited**	**limited**	**fine control but costly**
Some examples	· NOAA NOMADS (GFS) · DWD ICON Model	· FMI HIRLAM on Amazon S3	· We could be pionneers	· Meteo France Models DCPC · AEMET, Spain
Notes	· “Provider pay for users bandwidth.”	· Cloud provider may change pricing policy	· Suitable mostly for large data sets (models) · “Users pay for their bandwidth”	· Heavy maintenance costs when data evolves

Possibility offered with Bittorrent

-

- Great speed up of data delivery & low telecom footprint
- Reduced telecom capacity planning
- Reduced DdoS risks
- Allow Partial Downloads
- Suitable for all data set sizes > 100MB
- Possible to interface with data anouncement protocols:
  - AMQP
  - RSS

OpenWIS 2 Bittorrent candidate architecture