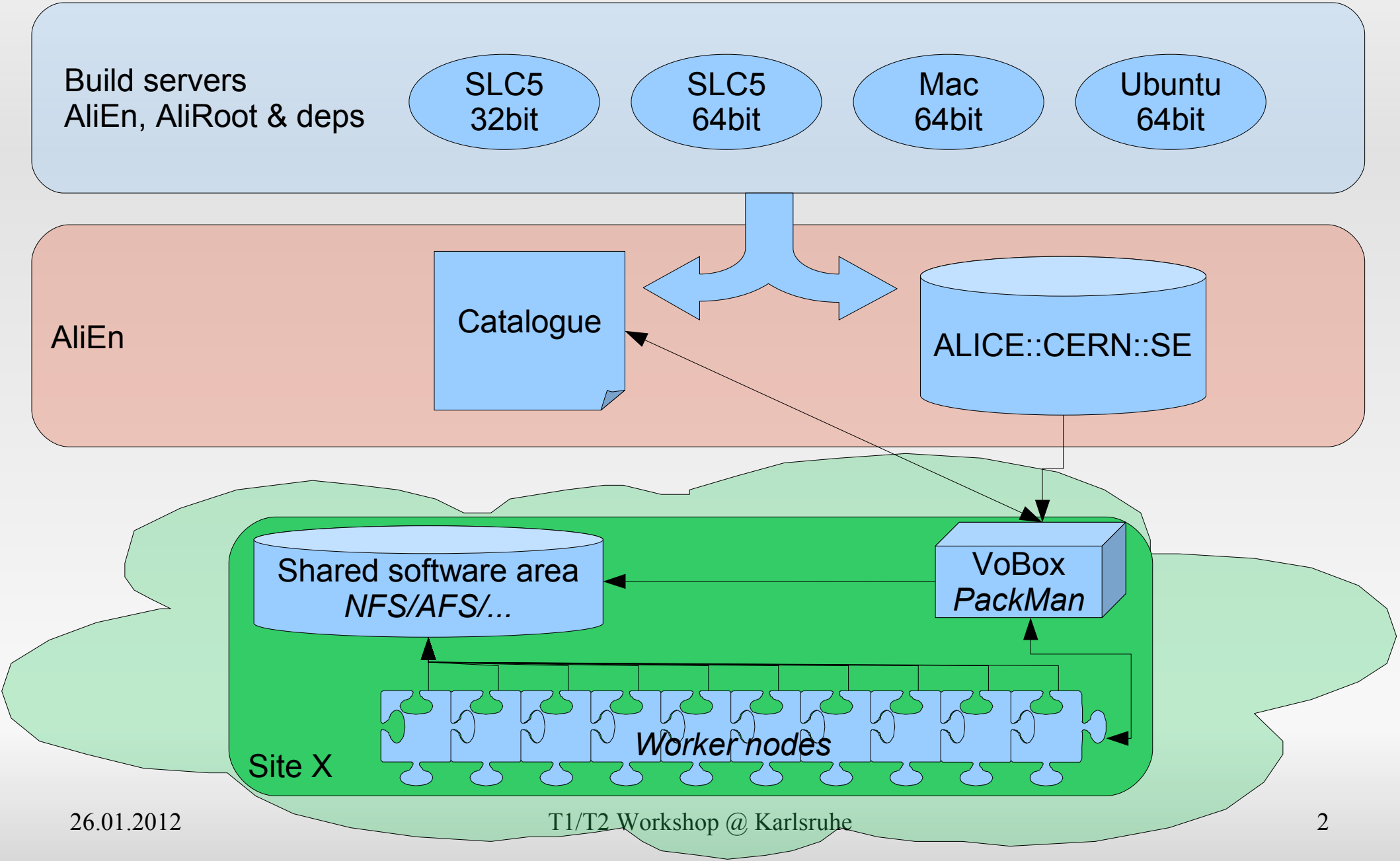
Torrent-based software distribution

Costin.Grigoras@cern.ch

Shared area software deployment



Shared area software deployment

Advantages

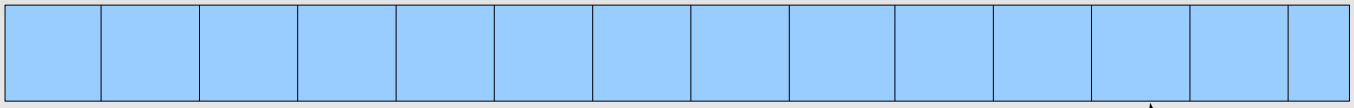
- One service/site managing the installation of required packages

Disadvantages

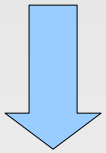
- Shared software area is a single point of failure / bottleneck
- Difficult to redeploy rebuilds of the same version
- Need to keep a short list of active software packages

Preparing for torrent

package.tar.bz2



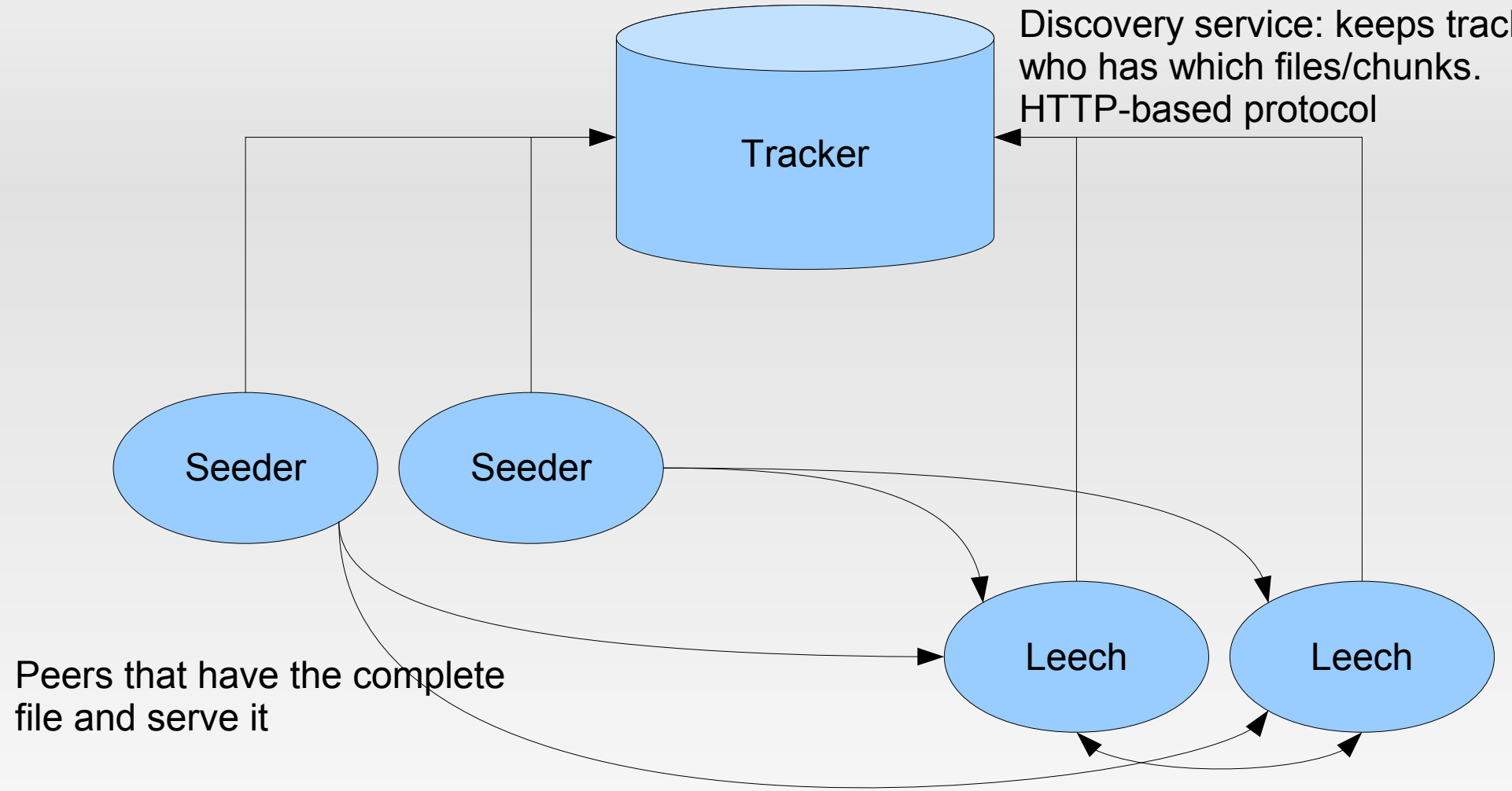
Chunks of equal size



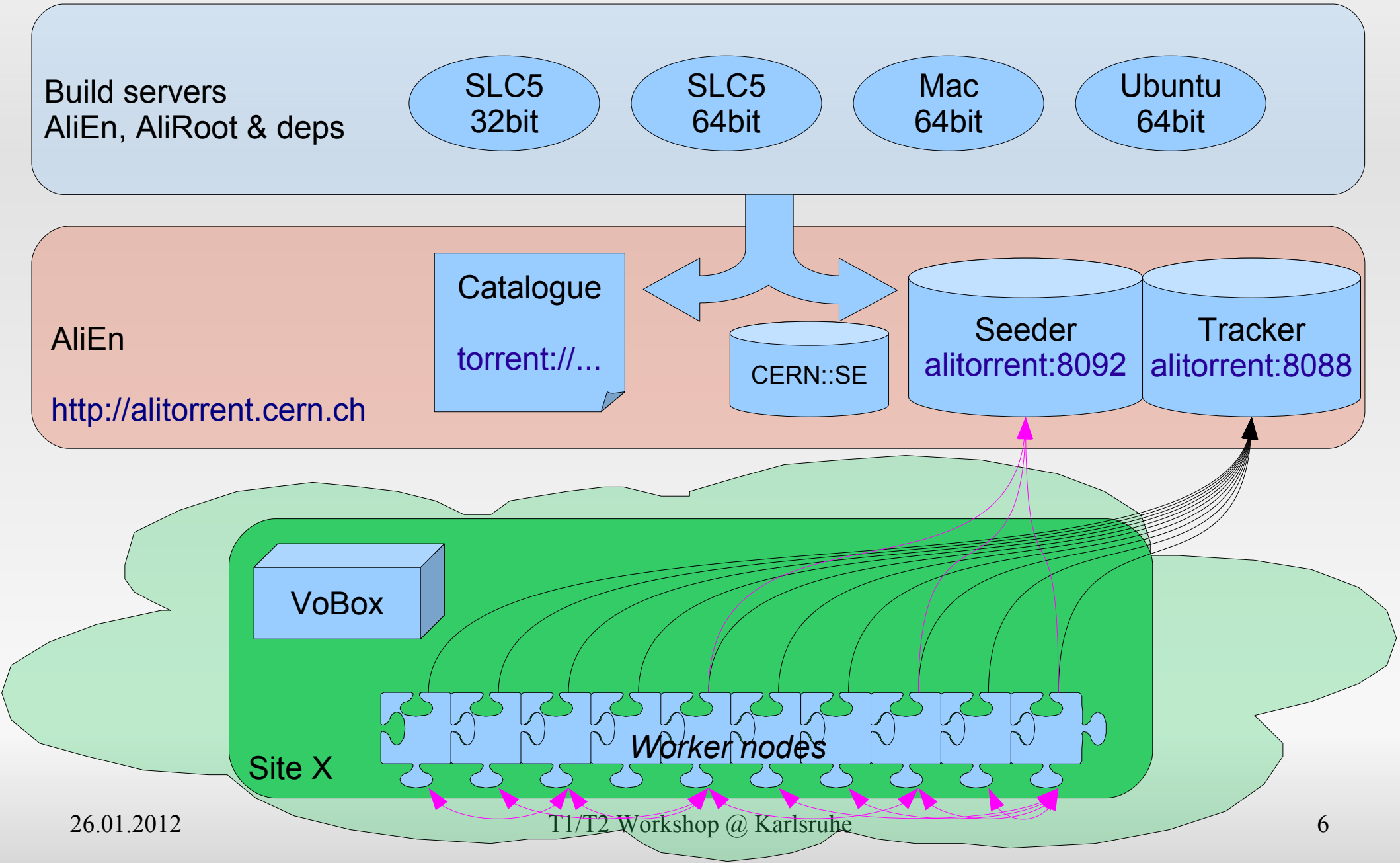
package.tar.bz2.torrent (tens of KB)

Metadata info of the original file:
- SHA1 hashes of chunks
- SHA1 hash of the entire file
  * uniquely identifies the file
- Tracker location (entry point)

Data flow in torrent networks



Implementation in AliEn



Implementation in AliEn

- Worker nodes keep seeding the packages that they have downloaded
  - Other worker nodes will fetch the content mostly from local nodes, if available
  - Worker nodes from site A are usually firewalled from site B, so no inter-site traffic
  - If initial download is not possible via torrent, fall back to wget and then seed the fetched files

Current status

- AliEn itself is packaged in a small (35MB) archive
- AliRoot, Root & deps. : max. 300MB/job
- Activated by a flag in LDAP
  - LDAP flag to switch modes:
    name=<CE_NAME>,ou=CE,ou=Services,ou=<SITE>,ou=Sites,o=alice,dc=cern,dc=ch
    installMethod=Torrent
  - Stable running at CERN, 2 T1s and a few T2s
- Integrated in the *alien-installer* and http://alimonitor.cern.ch/packages/

Special features of aria2c we use

- DHT (Distributed Hash Table)
  - Decentralized distributed lookup system
- Peer exchange
  - Information about local peers will be quickly propagated between peers
- LPD (Local Peer Discovery)
  - Multicast mechanism to find out other peers in the local network
- So the system can work even without access to the central seeder and tracker

Firewall requirements

- Outgoing access from the WNs to alitorrent.cern.ch:{8088, 8092}
- Please don't allow incoming connection requests from the world to the WNs
  - But don't be surprised if they do talk to other outside nodes (users that have the package...)
- Allow WN-to-WN connections on at least
  - TCP,UDP/6881:6999 – aria2c listening ports
  - UDP,IGMP → 224.0.0.0/4 – local peer discovery

Other considerations

- You can install your own aria2c client on the system to avoid downloading it for bootstrap
- In the default configuration file you can tune it more to your liking
  - Enabling JSON/XML-RPC for local activity monitoring …
  - Custom ports
  - HTTP proxies
  - Bandwidth limits

FAQ: Bandwidth worries

- CERN seeder limited to 50MB/s
- In practice the machine has an average of 8MB/s outgoing
- So the fraction that goes to any particular site is negligible
- Even the internal network traffic should reduce and not saturate any particular segment any more

Plans

- Very soon enable this option for all sites
- One site at a time
  - We will contact in advance
- Please ensure the firewall allows smooth running of this
- Please be ready to support one service less for ALICE :)