GridNEWS: A distributed Grid platform for efficient storage, annotating, indexing and searching of large audiovisual news content

Ioannis Konstantinou

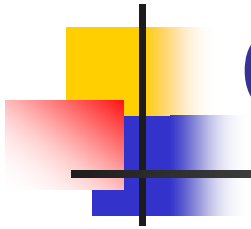
School of ECE

Computing Systems Laboratory

National Technical University of Athens

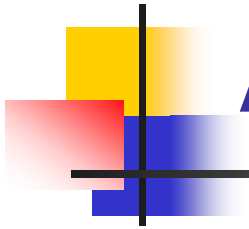
Concept

- Text based search in audiovisual content
- Search results: Portions of video files containing selected keywords
- Example
  - User searches for keyword “Acropolis”
  - Video portions containing the spoken word “Acropolis” are located and presented in the user
- YouTube like functionality



Objectives

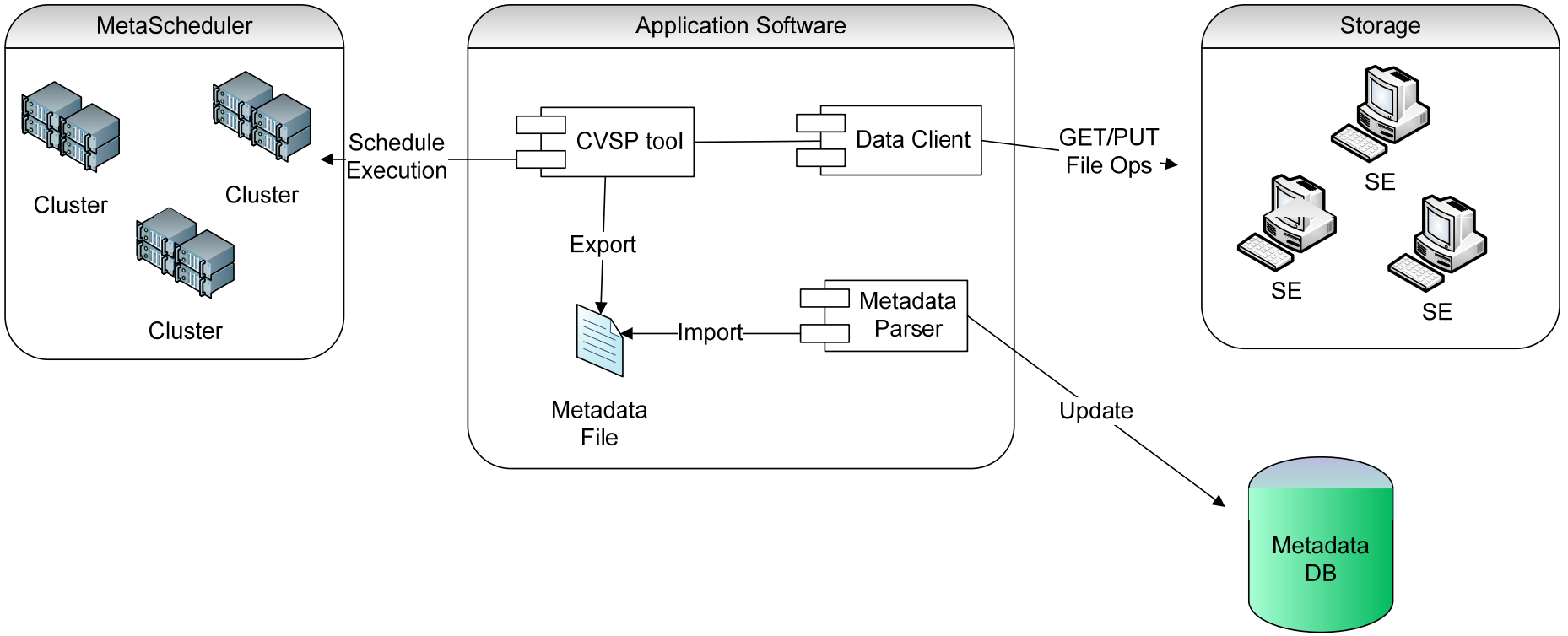
- Keyword extraction from video files using automatic speech recognition algorithms (ASR)
- Efficient and scalable distributed storage of large media content
- Indexing of extracted metadata for efficient keyword search
- YouTube like user interface for video searching/downloading
- Contribution to existing Grid Middleware using GGF standardized components



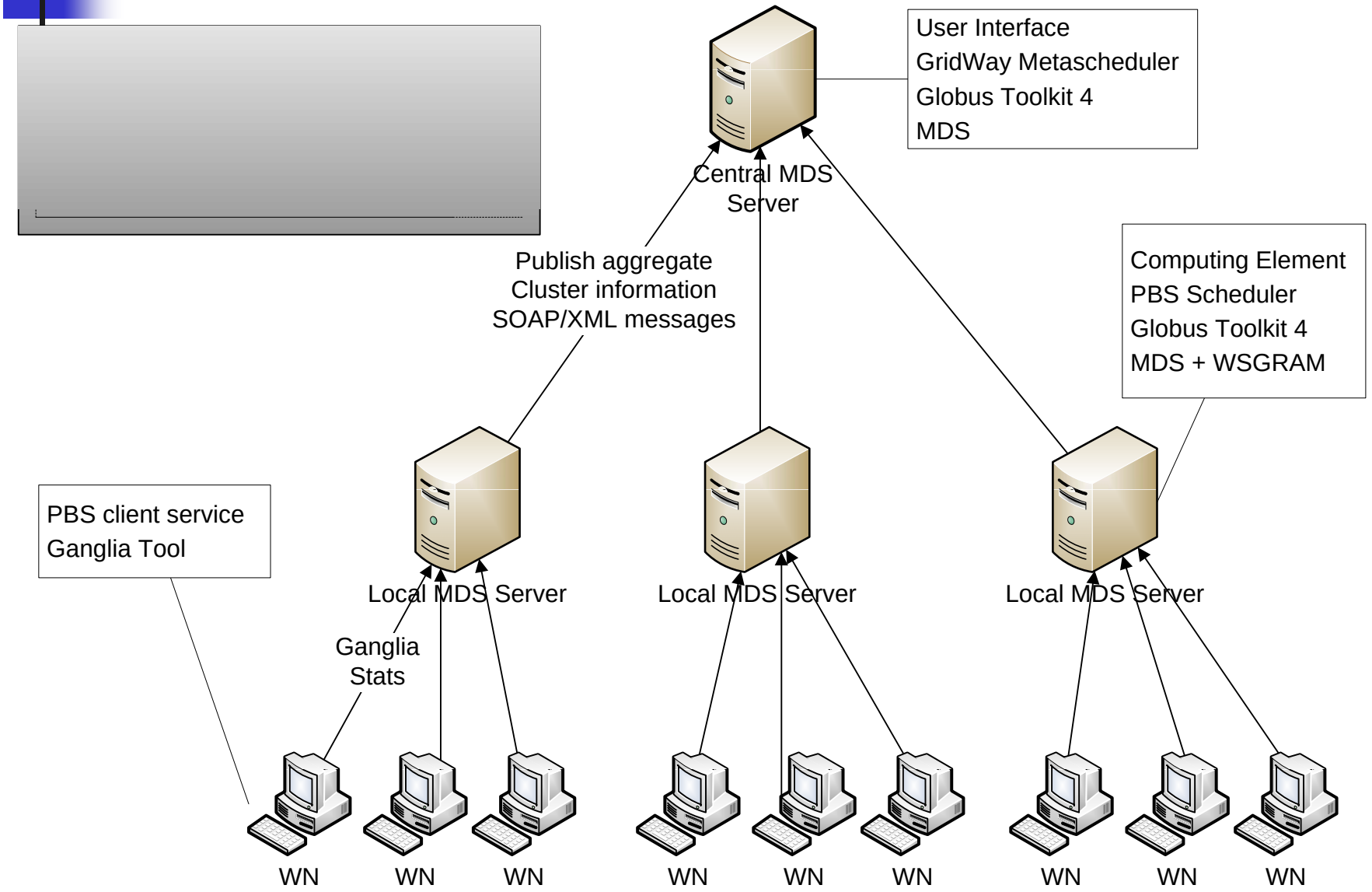
Addressed Issues

- Execution
  - Distributed execution of CPU/Data intensive Speech Recognition Algorithms
- Storage
  - Server load balancing using performance metrics
  - Client transfer time optimization using bittorrent like algorithms
  - Increase data availability
  - Multi-organizational data storage support using Virtual Organizations (VOs)

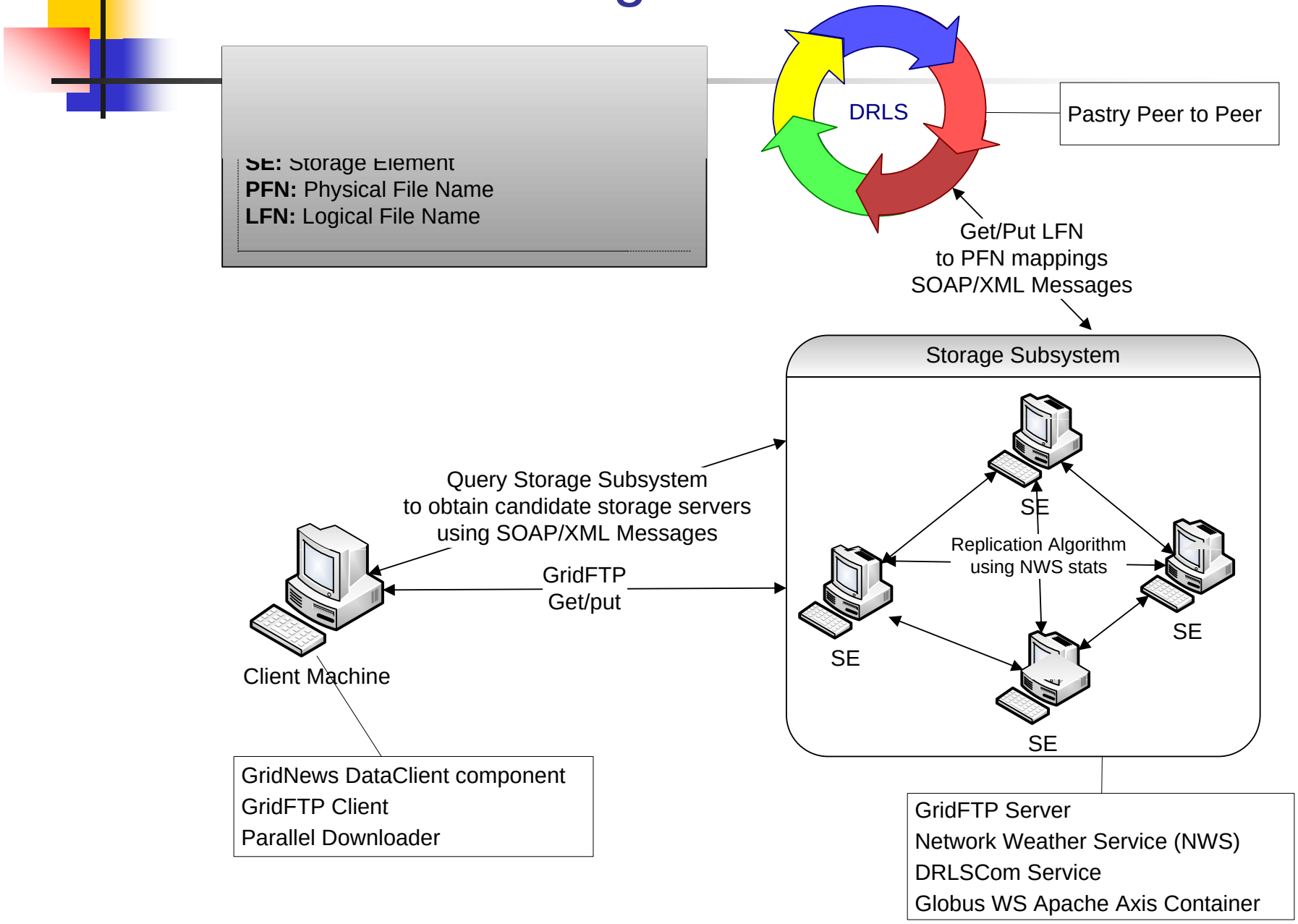
Video file keyword extraction procedure

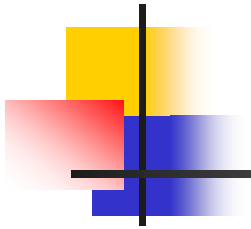


Distributed Execution Platform Architecture



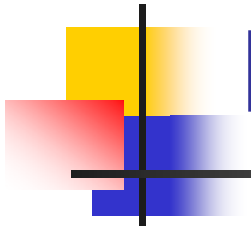
Distributed Storage Platform Architecture





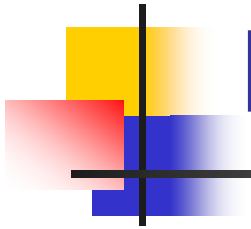
Distributed Replica Location Service

- Contains mappings of LFNs to PFNs
- DHT used: Pastry P2P
- Logarithmic routing
  - In a network with n nodes, a query needs only log(n) messages (hops)
  - Plaxton's algorithm minimizes query latencies
- Redundancy through replication
  - Eliminates single point of failure situations
- Inherent load balancing capabilities
  - Consistent hashing algorithms



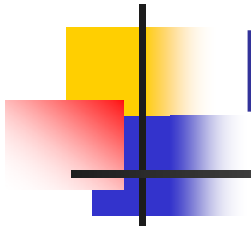
Load Balancing

- Servers exchange load metrics
  - CPU
  - Bandwidth
  - Free Disk Space
- Prediction algorithms (e.g. Linear regression) forecast future metrics from history data
- Weighted Normalized Metric WNM : $W_m X(M_t/M_{max})$
- Total Server Load (TSL): $Sum(WNM_i)_{i=1..n}$
- Servers maintain numerically sorted TSL list: $[TSL_1..TSL_n]$
- TSL list periodically refreshed



Replication

- Upon a STORE client request:
- Top k servers are selected from WNM list
  - k: configurable static replication factor
- Most suitable server is returned to the client
- Client initiates a **single** GridFTP file upload
- Server replicates the new file according to WNM list and factor k
- DRLS is informed about the new LFN->PFN mappings
- Client is informed Upon completion

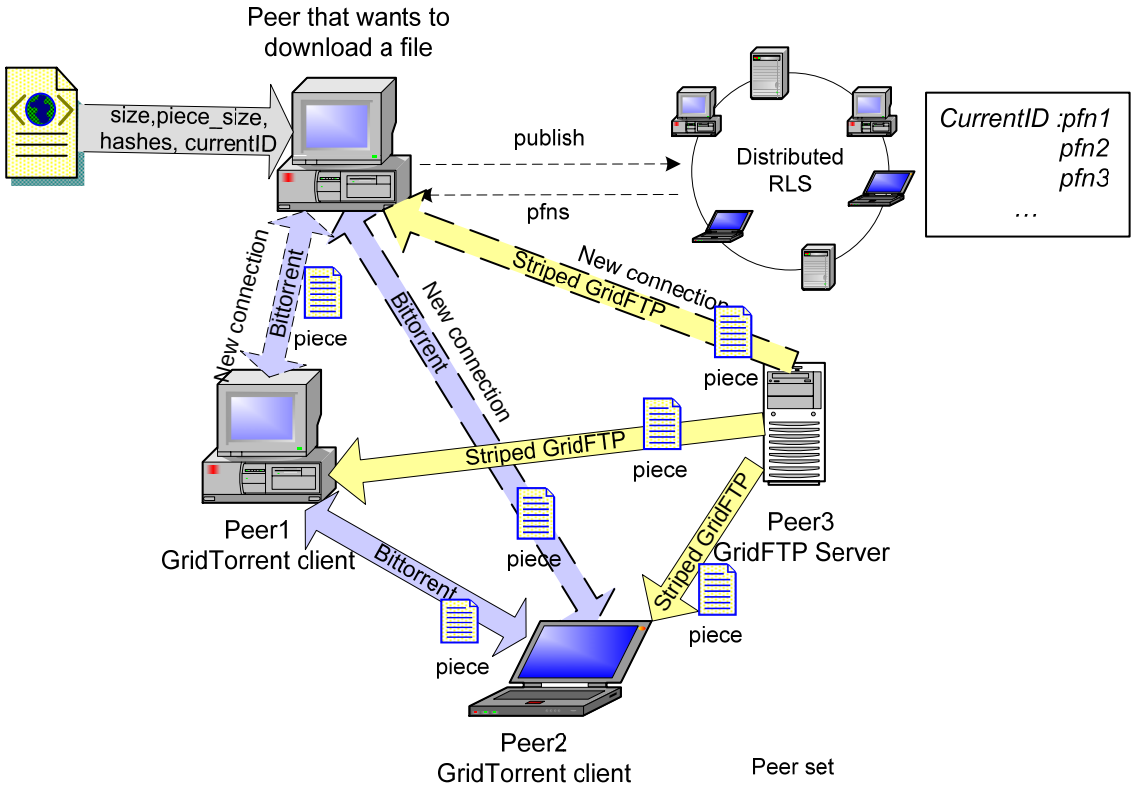


Parallel Downloader

- Upon a GET client request:
- Server contacts DRLS and retrieves replica locations
- Client establishes N GridFTP connections
- Client initiates N parallel (threaded) small data chunk requests
- After each successful retrieval, client re-initiates another request
- Optimum file transfer time:
  - The greater file portion is retrieved from the faster storage nodes
- To be replaced by GridTorrent

GridTorrent

- Metadata fields
  - Current Id
  - File size
  - Piece size
  - Hashes
- Distributed RLS instead of Tracker
- Partial GridFTP for actual transfer
- BitTorrent replica selection and tit-for-tat algorithm.
- Compatible with plain GridFTP servers
- PFN's prefix determines protocol (*gtp://site.fully.qualified.domain.name/path/to/file)*



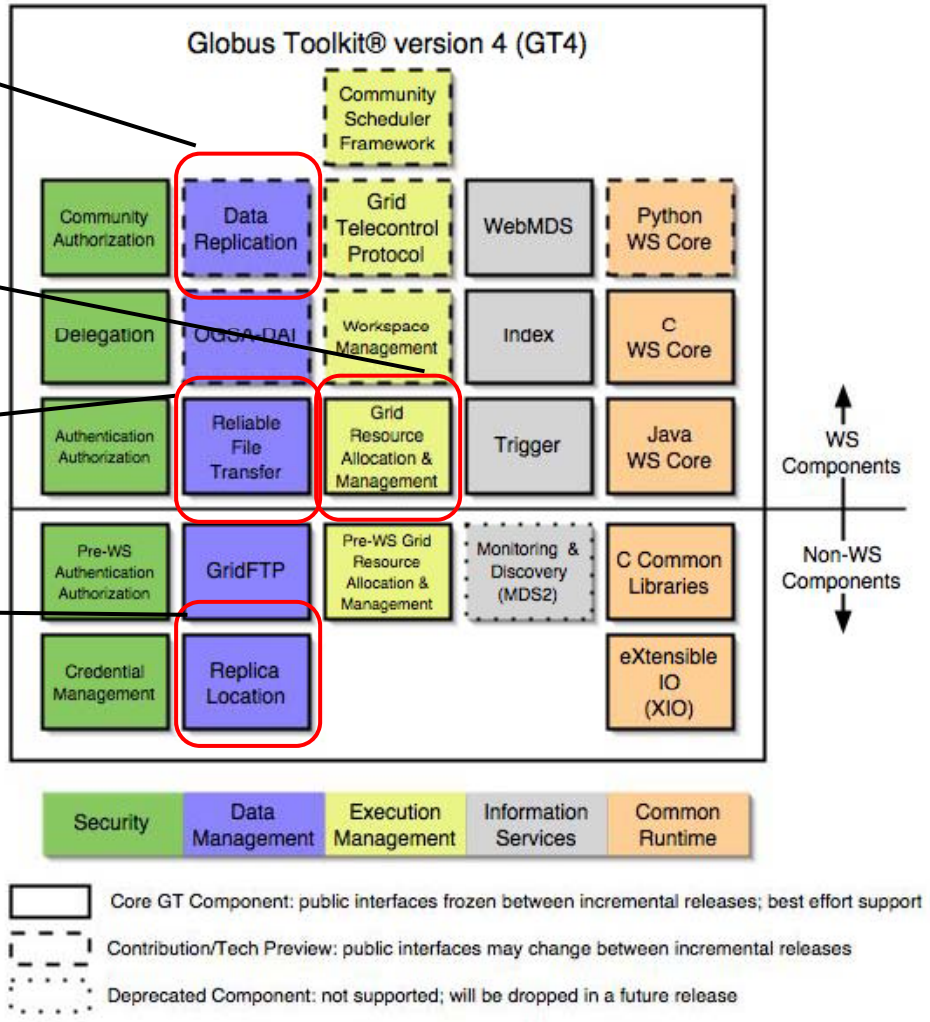
EXISTING MIDDLEWARE CONTRIBUTION

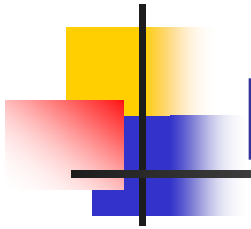
Design and development of a dynamic replica selection/placement algorithm

Added support for multiple clusters using Gridway Metascheduler

Enhanced bittorrent like file downloading from multiple sources

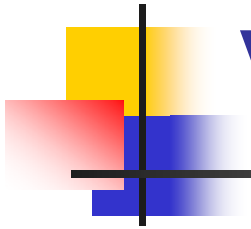
Replace centralized replica location service with a scalable distributed peer to peer solution





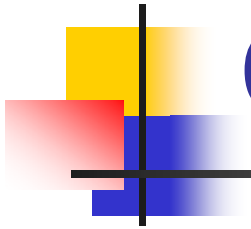
Development Testbed

- Hardware
  - 4X Dual Core AMD Opteron(tm) Processor 875 2.2GHz – 8 virtual CPU 16Gb Ram
  - Deployment of 5 Xen virtual machines, 2GB ram
- Software
  - Globus Toolkit v4.0
  - Globus WS Core (Apache AXIS WS Container)
  - Rice Pastry P2P (Java)
  - Network Weather Service
  - Torque (OpenPBS) scheduler
  - GridWay Metascheduler



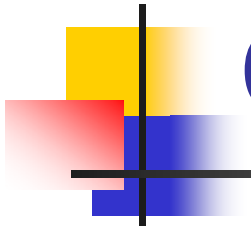
Virtualization

- Xen Hypervisor
  - Paravirtualization tequnique
  - Guest OS use special “xen aware” kernel
  - Direct utilization of special CPU instructions
  - Faster than full virtualization (VmWare)
- Use of Xen Hypervisor
  - Easy prototype management/administration
  - Simple control of the node lifecycle
  - Facilitate prototype deployment in many actual nodes



Currently working

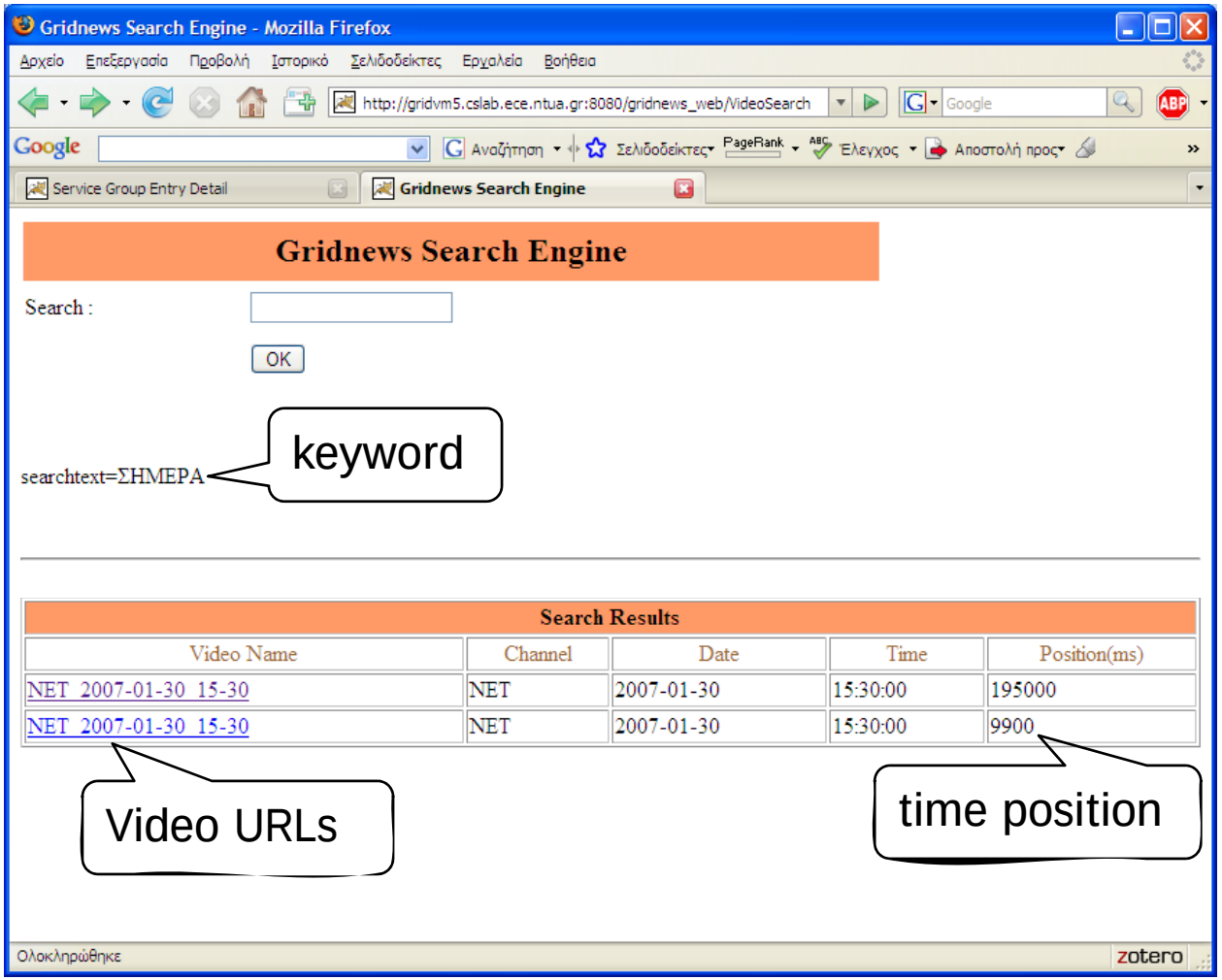
- Replace ParallelDownloader with GridTorrent
- Deploy prototype in the PlanetLab testbed
- Run experiments
- Fine-tune designed algorithms



Gridnews Portal

- Users can:
  - Perform keyword search in the auto-annotated multimedia content
  - View the video from their browser in a youTube style
  - Download only a fragment of the video where this keyword exists

Screenshots



Questions