Enabling Cross-Layer Optimizations in Storage Systems with Custom Metadata

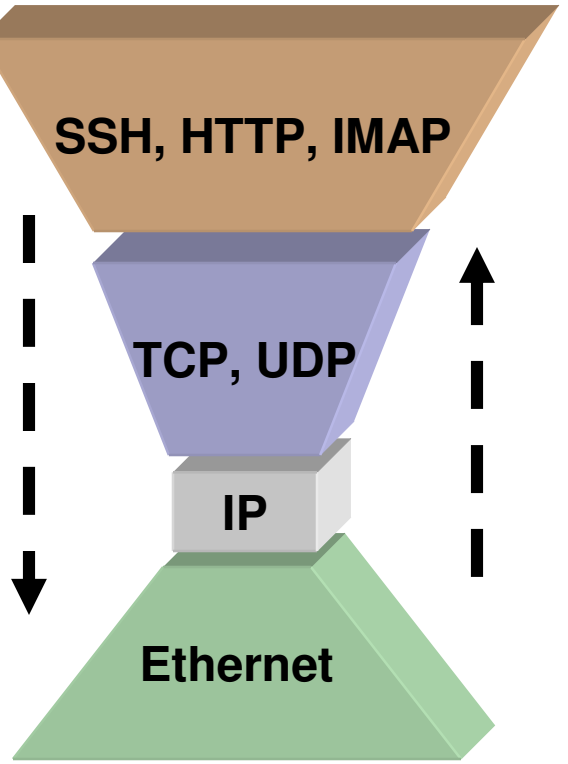
Elizeu Santos-Neto

Samer Al-Kiswany
¥Nazareno Andrade
Sathish Gopalakrishnan
Matei Ripeanu

University of British Columbia
Universidade Federal de Campina Grande¥

Layered Architectures

- TCP/IP
- Good, but…
  - …it limits information flow across layers.



Cross-Layer Communication: TCP/IP

- IP (Optional fields) [Gurtov03]
  - Interplay between the transport and the link layer
  - Communicating Flow QoS requirements (e.g. jitter vs. reliability)

Cross-Layer Communication: Web

- HTTP (Cache directives)
  - Readers can state: *give me a fresh copy*!
  - Publishers can state: *copy expires in T min*

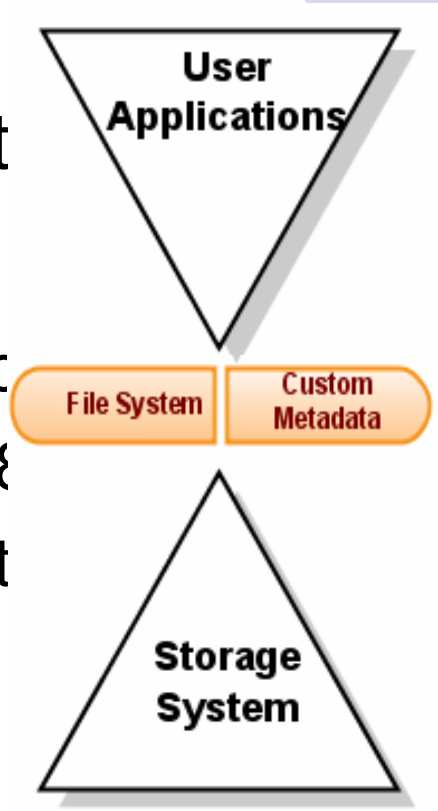
Cross-Layer Communication: Storage

- Applications → Storage System
  - Co-usage patterns
  - QoS requirements
  - Consistency requirements

- Applications ← Storage System
  - Provide storage-level information to applications
  - Notifications about data movements

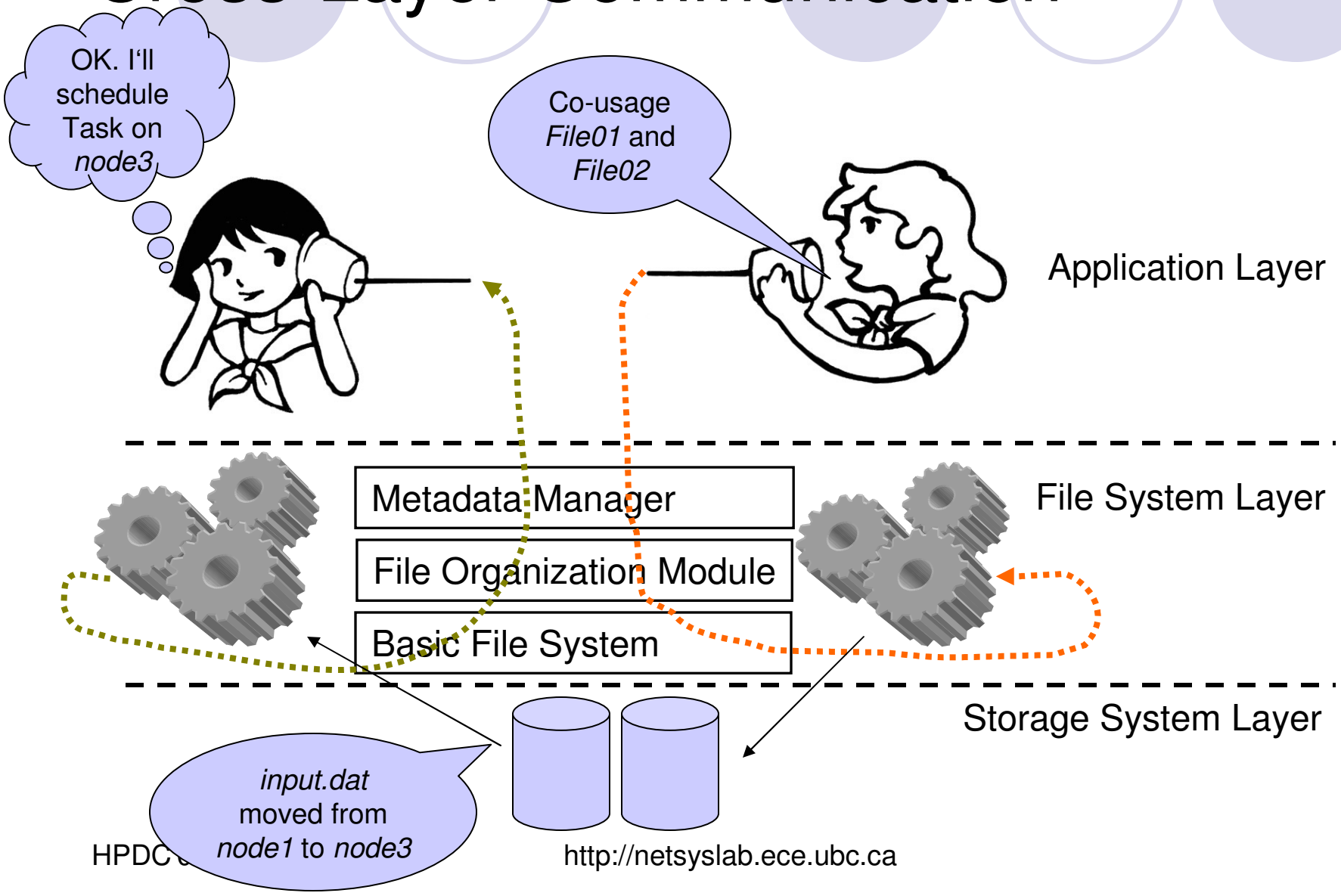
Exploiting Custom Metadata

- Custom metadat … files
- Past Uses: Applic … ation
  - MetaFS [Berry0… dar99]
  - Linking File Syst … es05]



- **Idea:** *custom metadata for cross-layer optimizations!*

Cross-Layer Communication



Opportunities

- POSIX interface already supports extended attributes
  - Backward compatibility is important
- Application-agnostic cross-layer communication
  - *<key, value>* pair provide good flexibility
- Incremental transition
  - Initially custom metadata should be a bonus
- New opportunities for usage-based optimization
  - Direct communication from the application

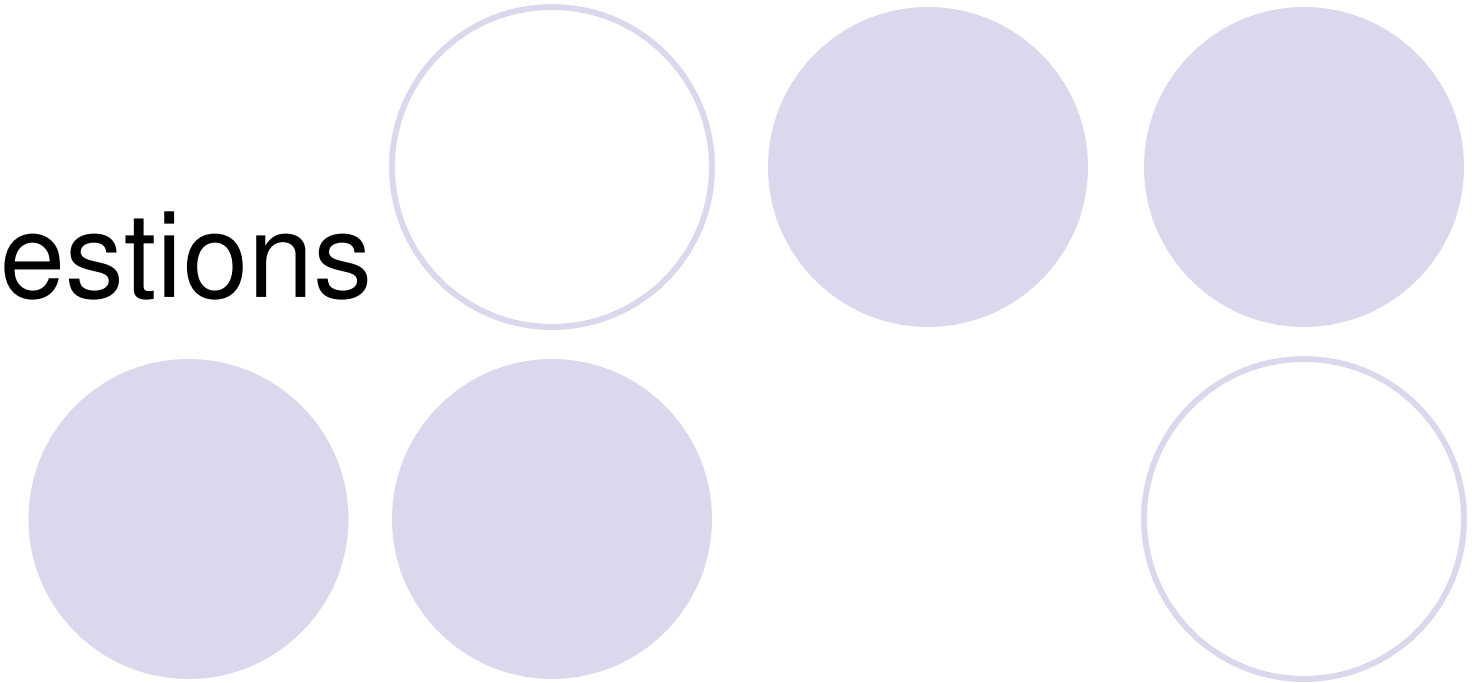
Challenges

- Design
  - Layered design is good, we should not break it
  - Efficient metadata management
- Will standards emerge?
- Policy vs. Mechanism
  - The cross-layer communication is not everything
- Incentive-compatible adoption
  - The lesson from IP optional fields [Fonseca05]

Summary and Future Work

- Custom metadata enables cross-layer optimizations in storage systems
- Build the cross-layer communication mechanism
  - Target cluster-based storage systems
  - e.g. FreeLoader [Vazhkudai05]
- Experimental evaluation

Questions



http://netsyslab.ece.ubc.ca

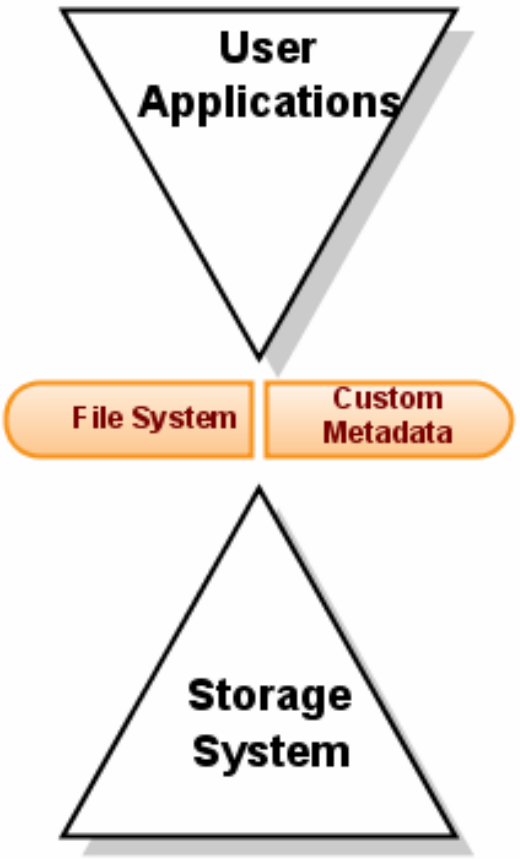
Use cases

- Application → Storage System
  - Custom metadata describing co-usage of files
  - Caching mechanism can consider file bundling

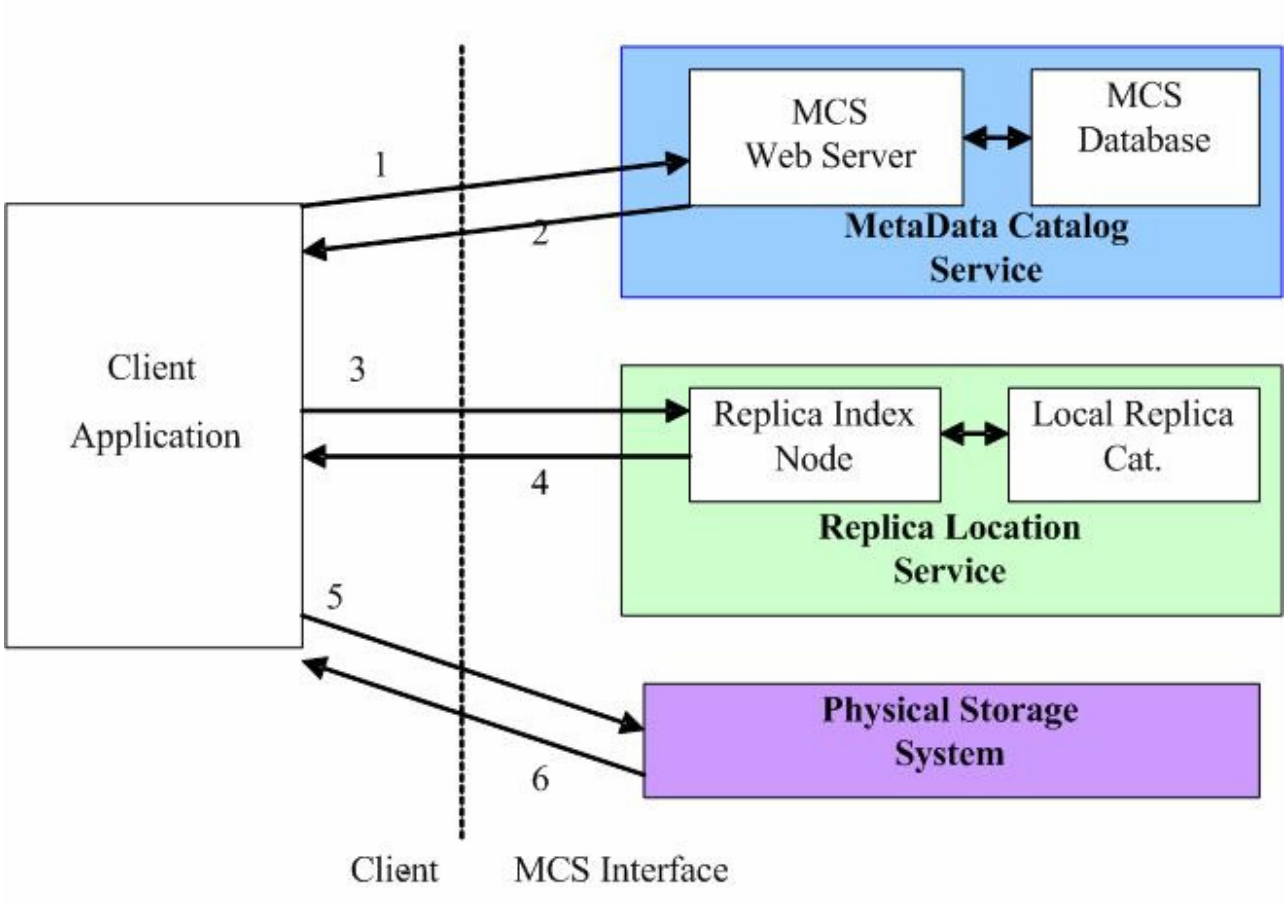
- Storage System → Application
  - Custom metadata describing file location characteristics
  - Application scheduler can exploit this information

Design Considerations

- Modular and Extendable: new mechanisms and attributes.
- Interfaces should be stable
- Components (layers) should be confined



Globus Metadata Catalog Service



Source: http://www.isi.edu/~deelman/MCS/

Extended Attributes

- Application → storage system communication
- *e.g.* `Ext2fs` tools (`lsattr` and `chattr`)
  - Synchronous writes (`S`)
  - Immutable files (`i`)
  - Undelete support (`u`)
- Mechanisms may not honor the attributes semantics

Traditional Use of Custom Metadata