



ResourceSync:
A Web-Based Resource
Synchronization Framework

ResourceSync Problem Statement

Consideration:

- Source (server) A has resources that change over time: they get created, modified, deleted
- Destination (servers) X, Y, and Z leverage (some) resources of Source A.

Problem:

- Destinations want to keep in step with the resource changes at Source A: resource synchronization.

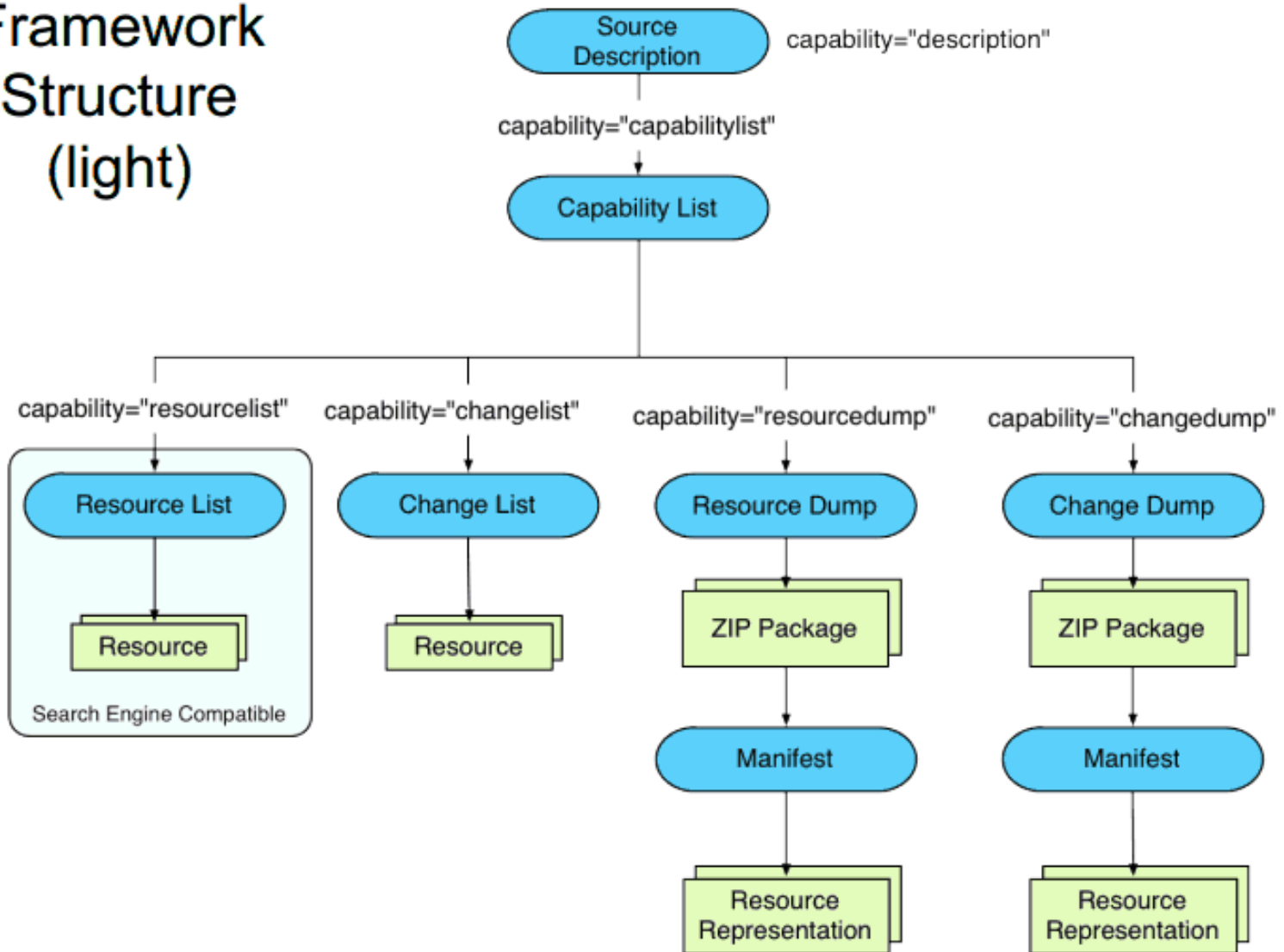
Goal:

- Design an approach for resource synchronization aligned with the Web Architecture that has a fair chance of adoption by different communities.
- The approach must scale better than recurrent HTTP HEAD/GET on resources.

Synchronization

1. Baseline synchronization – A destination must be able to perform an initial load or catch-up with a source
2. Incremental synchronization – A destination must have some way to keep up-to-date with changes at a source
3. Audit – A destination should be able to determine whether it is synchronized with a source

Framework Structure (light)



Framework Technical Details

1. Sitemaps
2. Pull method
3. Linking between resources
4. Discovery
5. Push method
6. Archives

Use Case 1: arXiv.org



Cornell University
Library

We gratefully acknowledge support from
the Simons Foundation
and member institutions

arXiv.org

[Login](#)

Search or Article-id

[\(Help | Advanced search\)](#)

All papers ▾

Go!

Open access to 872,445 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics

Subject search and browse:

See cumulative ["What's New"](#) pages. Read [robots beware](#) before attempting any automated download

Physics

- **Astrophysics** ([astro-ph new](#), [recent](#), [find](#))
includes: [Cosmology and Extragalactic Astrophysics](#); [Earth and Planetary Astrophysics](#); [Galaxy Astrophysics](#); [High Energy Astrophysical Phenomena](#); [Instrumentation and Methods for Astrophysics](#); [Solar and Stellar Astrophysics](#)
- **Condensed Matter** ([cond-mat new](#), [recent](#), [find](#))
includes: [Disordered Systems and Neural Networks](#); [Materials Science](#); [Mesoscale and Nanoscale Physics](#); [Other Condensed Matter](#); [Quantum Gases](#); [Soft Condensed Matter](#); [Statistical Mechanics](#); [Strongly Correlated Electrons](#); [Superconductivity](#)
- **General Relativity and Quantum Cosmology** ([gr-qc new](#), [recent](#), [find](#))
- **High Energy Physics - Experiment** ([hep-ex new](#), [recent](#), [find](#))
- **High Energy Physics - Lattice** ([hep-lat new](#), [recent](#), [find](#))
- **High Energy Physics - Phenomenology** ([hep-ph new](#), [recent](#), [find](#))
- **High Energy Physics - Theory** ([hep-th new](#), [recent](#), [find](#))
- **Mathematical Physics** ([math-ph new](#), [recent](#), [find](#))
- **Nonlinear Sciences** ([nlin new](#), [recent](#), [find](#))
includes: [Adaptation and Self-Organizing Systems](#); [Cellular Automata and Lattice Gases](#); [Chaotic Dynamics](#); [Exactly Solvable and Integrable Systems](#); [Pattern Formation and Solitons](#)
- **Nuclear Experiment** ([nucl-ex new](#), [recent](#), [find](#))
- **Nuclear Theory** ([nucl-th new](#), [recent](#), [find](#))
- **Physics** ([physics new](#), [recent](#), [find](#))
includes: [Accelerator Physics](#); [Atmospheric and Oceanic Physics](#); [Atomic Physics](#); [Atomic and Molecular Clusters](#); [Biological Physics](#); [Chemical Physics](#); [Classical Physics](#); [Computational Physics](#); [Data Analysis, Statistics and Probability](#); [Fluid Dynamics](#); [General Physics](#); [Geophysics](#); [History and Philosophy of Physics](#); [Instrumentation and Detectors](#); [Medical Physics](#); [Optics](#); [Physics Education](#); [Physics and Society](#); [Plasma Physics](#); [Popular](#)

Use Case 1: arXiv Mirroring and Data Sharing

- Repository of scholarly articles in physics, mathematics, computer science, etc.
- > 850k articles
- approx. 1.5 revisions per article on average
- approx. 75k new articles per year
- Each article has full-text and separate metadata record

Use Case 1: arXiv.org

arXiv mirror sites



- cn.arXiv.org (China)
- fr.arXiv.org (France)
- de.arXiv.org (Germany)
- in.arXiv.org (India)
- jp.arXiv.org (Japan)
- es.arXiv.org (Spain)
- uk.arXiv.org (U.K.)
- lanl.arXiv.org (née xxx.lanl.gov, U.S. mirror at Los Alamos)
- arXiv.org (U.S. primary site at Cornell University)

2012-06: We are in the process of reducing the size of the arXiv mirror network. Our goal is to support just a few of the more heavily used mirrors to provide geographic redundancy, but otherwise focus our attention on the main site. Geographic locality is much less useful than it was when the mirror network was established. Most users geographically close to mirrors, perhaps in neighboring countries, have better access to the main site than to a "nearby" mirror.

2012-10: The Israeli mirror (formerly il.arxiv.org) has been discontinued.

2012-08: The Taiwan mirror (formerly tw.arxiv.org) has been discontinued.

2012-04: The Brazilian and Russian mirrors (formerly br.arxiv.org and ru.arxiv.org) have been discontinued.

2012-02: The APS and Australian mirrors (formerly aps.arxiv.org and au.arxiv.org) have been discontinued.

2009-01: The Italian mirror (formerly it.arxiv.org) has been discontinued.

2008-12: The South Korean mirror (formerly kr.arXiv.org) has been discontinued.

2007-06: The South Africa mirror (formerly za.arXiv.org) has been discontinued.

Other interfaces to arXiv articles

- The Front for the arXiv
- The [NASA Astrophysics Data System \(ADS\)](http://NASA Astrophysics Data System (ADS))
- The IOP's eprintweb.org

Use Case 1: arXiv.org Goals

- Goal #1: Keep mirror sites synchronized with daily changes
- Goal #2: Make resources and update information publicly available so that any other service may synchronize at the frequency it needs, e.g. Math Front at UC Davis, printWeb from IOP in UK, Data for bibliometric and scientometric analysis

Timeline

August 2013

- Release of version 0.9.2 of ResourceSync framework
- Specification Version 0.x of Push-based methods for ResourceSync
- Soliciting broad feedback

Fall 2013

- Specification becomes NISO standard

ResourceSync & ADR

- Is ResourceSync appropriate for the ADR?
- How could ResourceSync help the ADR build a better digital repository?