

Praktikum: Datenmanagement und Datenanalyse

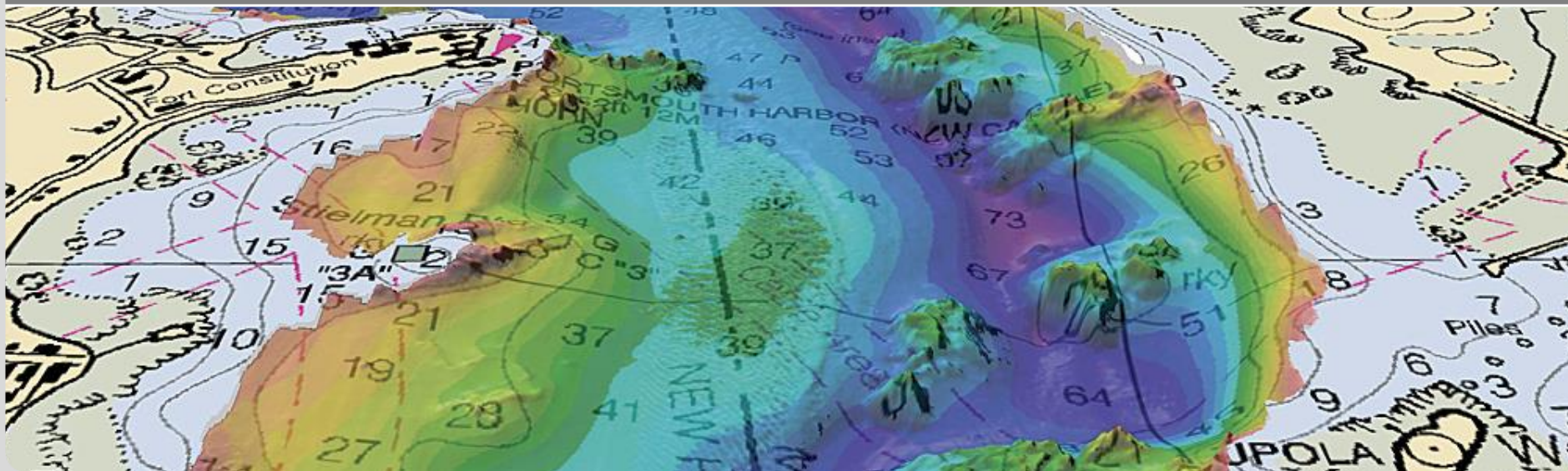
Projekt: Skalier- und erweiterbare Online-Plattform für Geodaten

EasyThredds: einfacher Datenzugriff und Protokolloptimierung

Maximilian Czerny

STEINBUCH CENTER FOR COMPUTING (SCC), KIT

Betreuer: Uğur Çayoğlu



- Aktuell: technologische Vielfalt und Komplexität
 - Diverse Daten-, Anfrage- und Übermittlungsformate
- Ziel: Geowissenschaftlern Zugriff, Analyse und Visualisierung erleichtern
 - Einheitlicher Zugriff
 - Plattform wählt selbstständig beste Formate / Protokolle

Verwirrungspunktestand —→ 0

THREDDS

Thematic Real-time Environmental Distributed Data Services

- *“middleware to bridge the gap between data providers and data users” – [TDS-facts]*

THREDDS Data Server (TDS)

- Katalog mit Metadaten
- Datenaggregation
- Datenzugriff: OPeNDAP, CdmRemote, DAP4, NCSS, HTTP, FTP, ...

Common Data Model (CDM)

- Formate: netCDF, HDF, GRIB
- Erweiterbar durch eigene Leser mit netCDF-Java API
- Datentypen: Gitter, Punkte, Radiär

THREDDS

Thematic Real-time Environmental Distributed Data Services

- *“middleware to bridge the gap between data providers and data users” – [TDS-facts]*

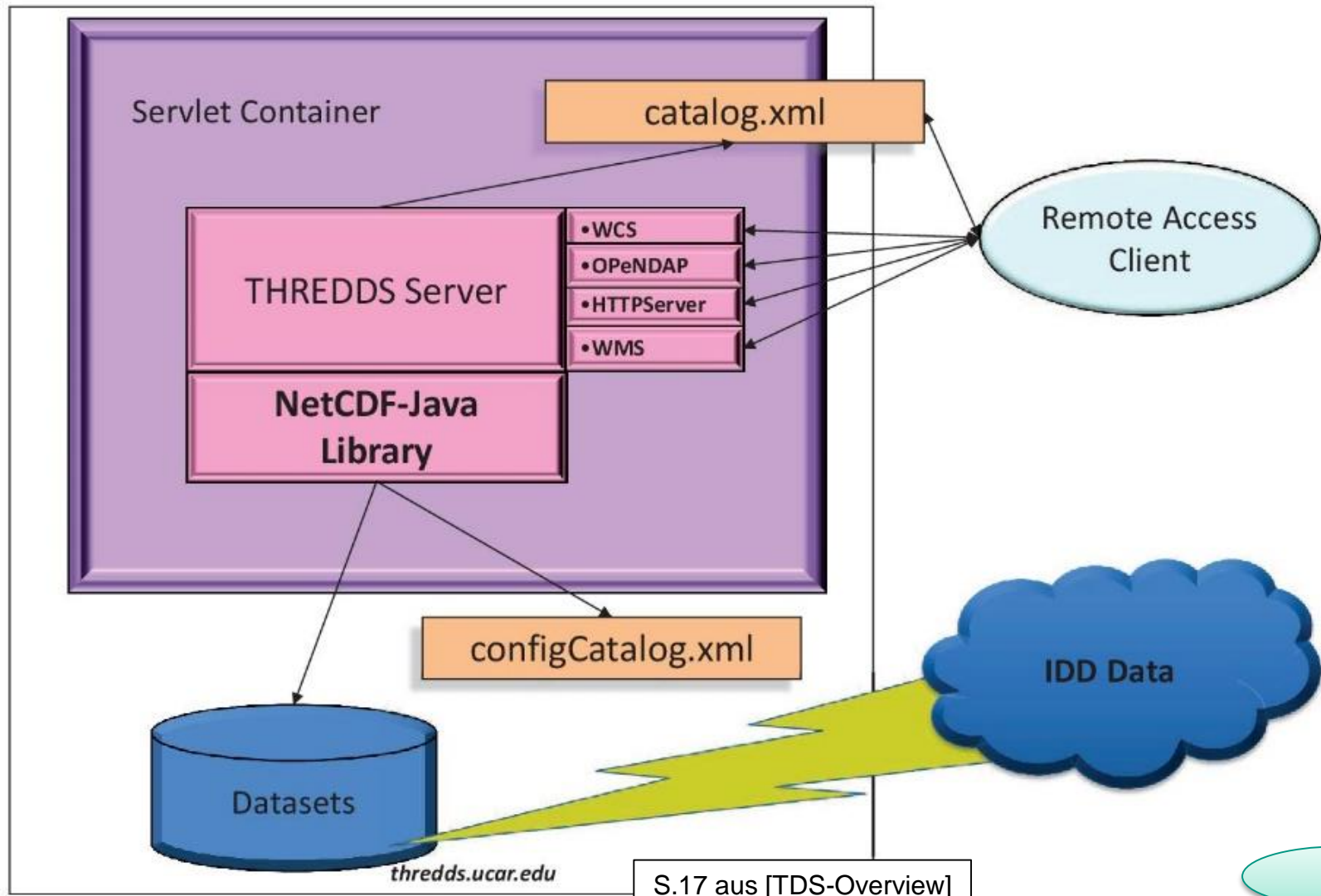
THREDDS Data Server (TDS)

- Katalog mit Metadaten
- Datenaggregation
- Datenzugriff: OPeNDAP, CdmRemote, DAP4, NCSS, HTTP, FTP, ...

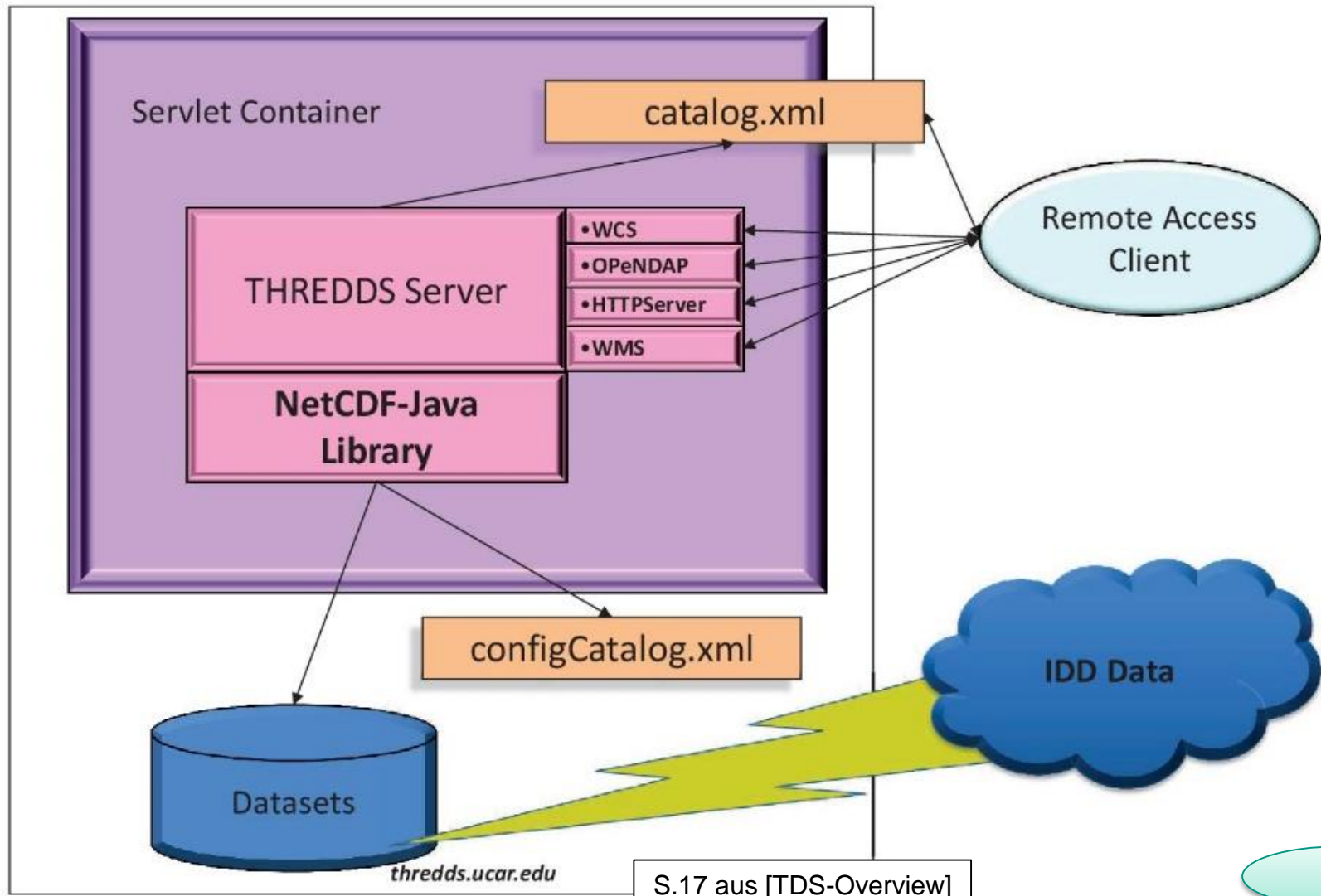
Common Data Model (CDM)

- Formate: netCDF, HDF, GRIB
- Erweiterbar durch eigene Leser mit netCDF-Java API
- Datentypen: Gitter, Punkte, Radiär

THREDDS Data Server (TDS)



THREDDS Data Server (TDS)



TDS Beispiel



TDS Installation of IMK-ASF @ KIT

THREDDS Data Server

Catalog http://nc-catalogue.scc.kit.edu/thredds/imk_rc1sdb8.html

Dataset: THREDDS IMK Server!/cloud.nc

- ID: RC1SD-base-08_cloud

Access:

1. **OPENDAP:** </thredds/dodsC/RC1SD-base-08/cloud>
2. **DAP4:** </thredds/dap4/RC1SD-base-08/cloud>
3. **NetcdfSubset:** </thredds/ncss/RC1SD-base-08/cloud>
4. **CdmRemote:** </thredds/cdmremote/RC1SD-base-08/cloud>

Viewers:

- [NetCDF-Java ToolsUI \(webstart\)](#)

Screenshot [CLOUD]

Datenzugriffsprotokolle

OPeNDAP

DAP4

CdmRemote

NCSS

Datenzugriffsprotokolle

	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc

Datenzugriffsprotokolle

	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc
Metadaten	.dds	.dds	header ncstream	.nc
Dimensions- daten	Enthalten	Enthalten	data ncstream	XML

Datenzugriffsprotokolle

	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc
Metadaten	.dds	.dds	header ncstream	.nc
Dimensions- daten	Enthalten	Enthalten	data ncstream	XML
netCDF- konform	✓	✓	✗	✓

Datenzugriffsprotokolle

	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc
Metadaten	.dds	.dds	header ncstream	.nc
Dimensions- daten	Enthalten	Enthalten	data ncstream	XML
netCdf- konform	✓	✓	✗	✓
Subsetting	✓	✓	✓	✗
absolute Werte	✗	✗	✗	✓

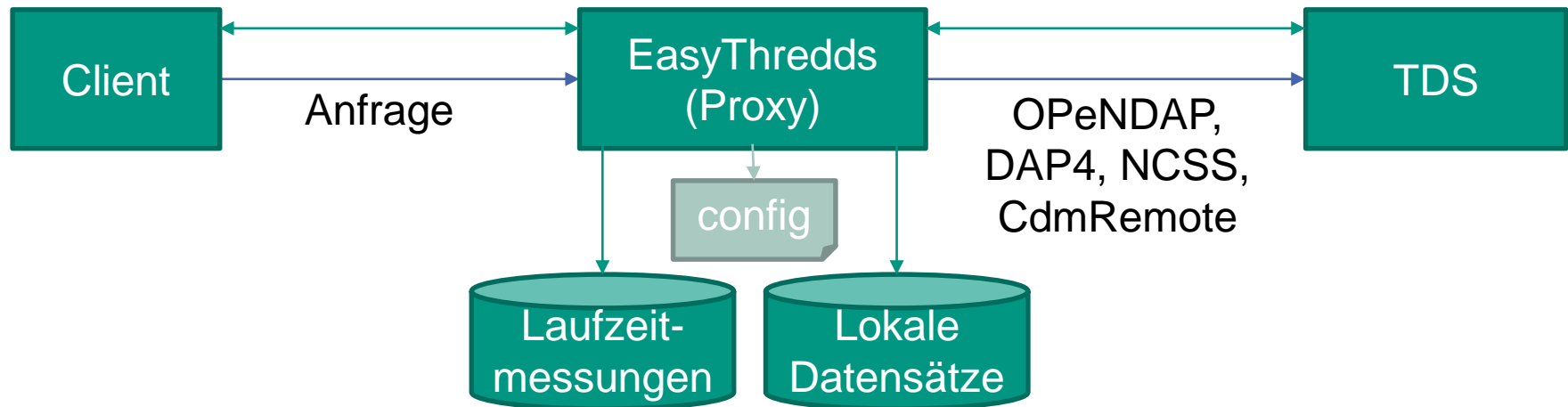
Datenzugriffsprotokolle

	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc
Metadaten	.dds	.dds	header ncstream	.nc
Dimensions- daten	Enthalten	Enthalten	data ncstream	XML
netCdf- konform	✓	✓	✗	✓
Subsetting	✓	✓	✓	✗
absolute Werte	✗	✗	✗	✓
Auswahlhilfe	✓	✗	✗	✓

Datenzugriffsprotokolle

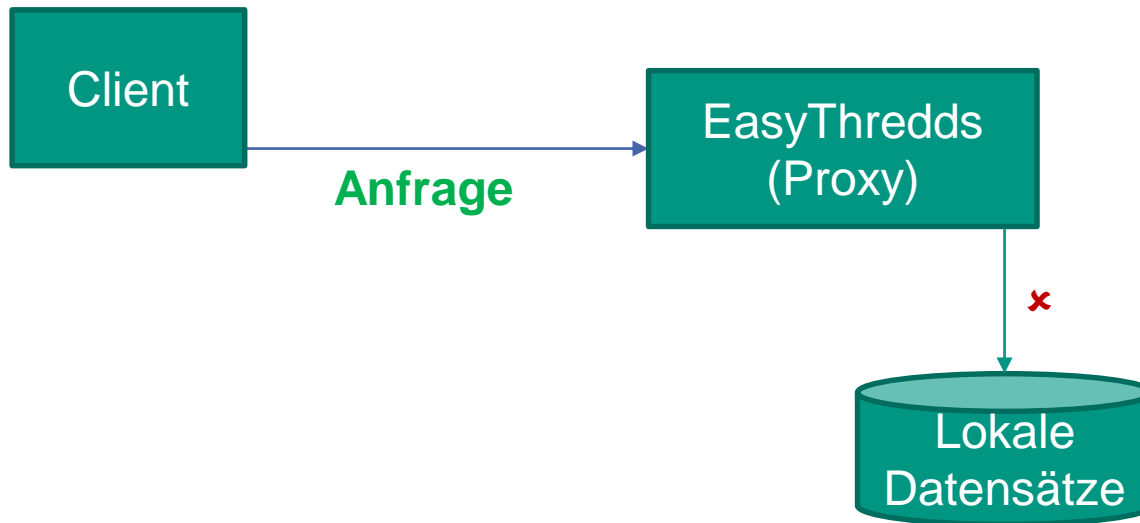
	OPeNDAP	DAP4	CdmRemote	NCSS
Daten	.dods	.dap	data ncstream	.nc
Metadaten	.dds	.dds	header ncstream	.nc
Dimensions- daten	Enthalten	Enthalten	data ncstream	XML
netCdf- konform	✓	✓	✗	✓
Subsetting	✓	✓	✓	✗
absolute Werte	✗	✗	✗	✓
Auswahlhilfe	✓	✗	✗	✓

EasyThredds als TDS-Proxy



- Tomcat Servlet als TDS-Proxy (lokal oder remote)
- Einheitliche Anfragesprache für Längen- und Breitengrad sowie Höhen- und Zeitangaben inklusive Subsetting

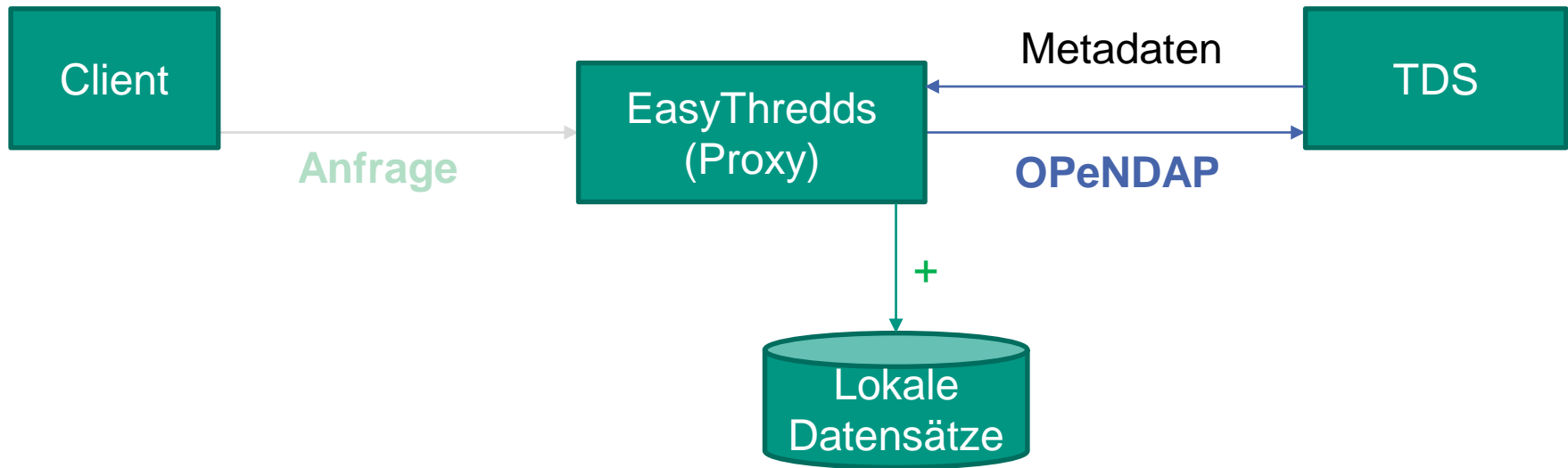
EasyThredds: neue Anfrage



[http://localhost:8080/EasyThredds/.opendap/translate/RC1SD-base-08/cloud?
var=ac1c&lev=\[0:1:5\]&lat=\[12.4;1;17.8\]&lon=\[4.0;;18.9\]&
time=\[20/08/1992-20:00:00;1;22/02/1999-12:00:00\]](http://localhost:8080/EasyThredds/.opendap/translate/RC1SD-base-08/cloud?var=ac1c&lev=[0:1:5]&lat=[12.4;1;17.8]&lon=[4.0;;18.9]&time=[20/08/1992-20:00:00;1;22/02/1999-12:00:00])

- Angabe von Protokoll (optional), Query und Datensatz

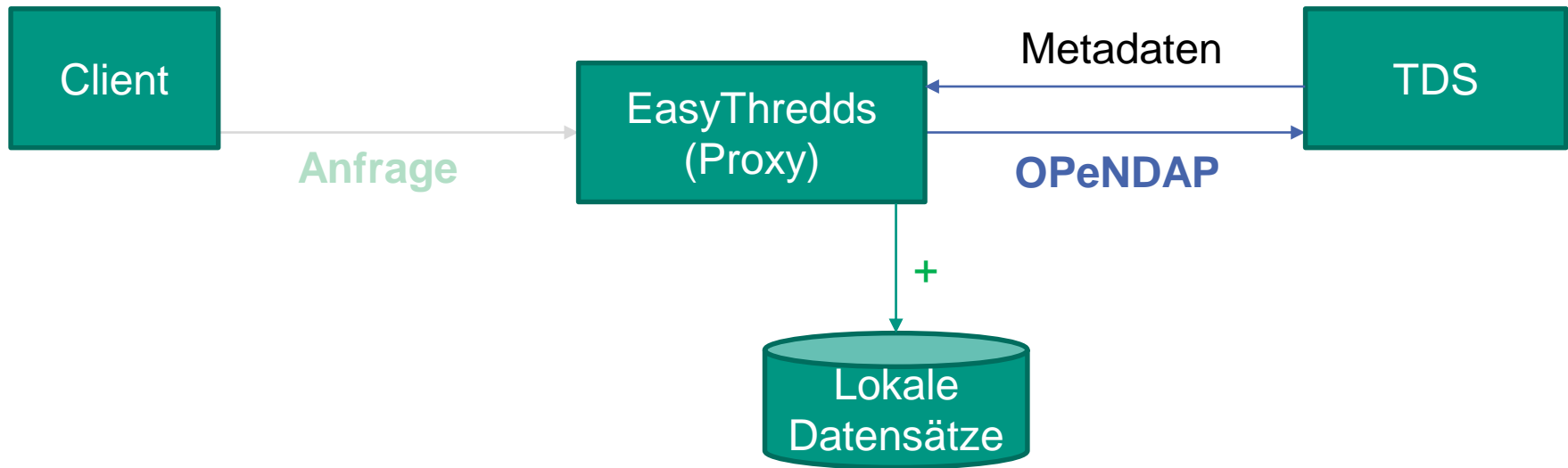
EasyThredds: neue Anfrage



[http://localhost:8080/EasyThredds/opeNDAP/translate/RC1SD-base-08/cloud?var=ac1c&lev=\[0:1:5\]&lat=\[12.4;1;17.8\]&lon=\[4.0;;18.9\]&time=\[20/08/1992-20:00:00;1;22/02/1999-12:00:00\]](http://localhost:8080/EasyThredds/opeNDAP/translate/RC1SD-base-08/cloud?var=ac1c&lev=[0:1:5]&lat=[12.4;1;17.8]&lon=[4.0;;18.9]&time=[20/08/1992-20:00:00;1;22/02/1999-12:00:00])

- Abfrage von Koordinaten zur Indexauflösung
- Metadaten speichern

EasyThredds: neue Anfrage

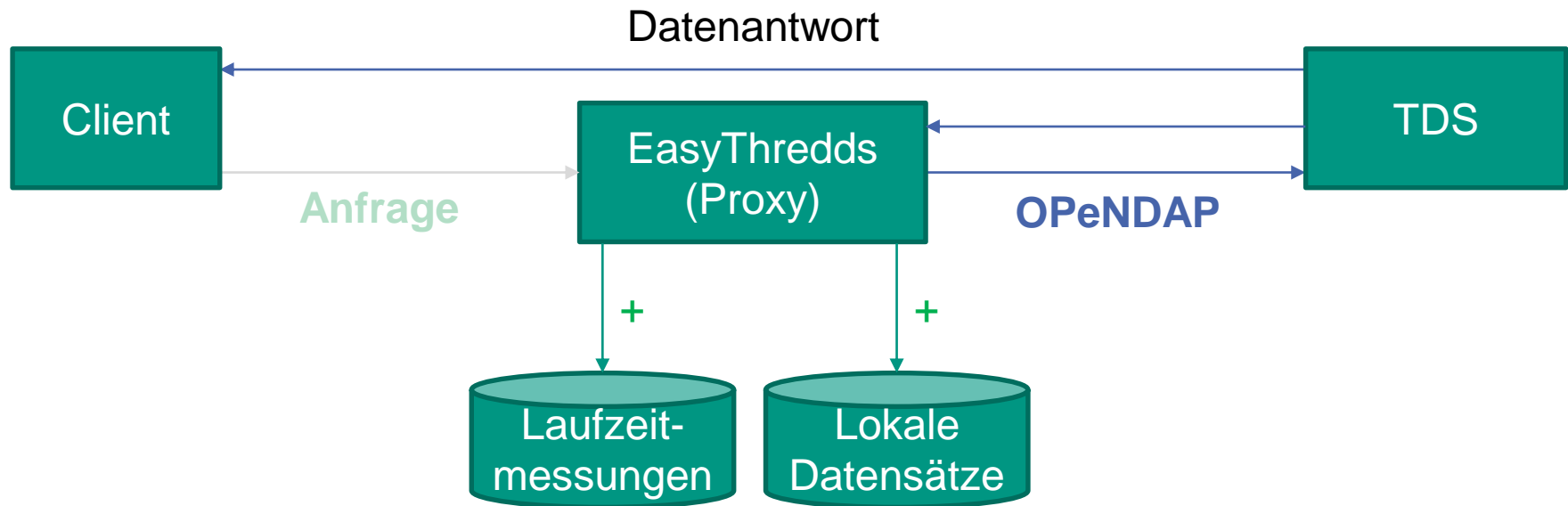


[http://localhost:8080/EasyThredds/ope ndap/translate/RC1SD-base-08/cloud?var=ac lc&lev=\[0:1:5\]&lat=\[12.4;1;17.8\]&lon=\[4.0;;18.9\]&time=\[20/08/1992-20:00:00;1;22/02/1999-12:00:00\]](http://localhost:8080/EasyThredds/ope ndap/translate/RC1SD-base-08/cloud?var=ac lc&lev=[0:1:5]&lat=[12.4;1;17.8]&lon=[4.0;;18.9]&time=[20/08/1992-20:00:00;1;22/02/1999-12:00:00])

↓ Indexübersetzung

[http://nc-catalogue.scc.kit.edu/thredds/dodsC/RC1SD-base-08/cloud.dods?ac lc\[11954:1:17658\]\[0:1:4\]\[26:1:28\]\[2:1:7\]](http://nc-catalogue.scc.kit.edu/thredds/dodsC/RC1SD-base-08/cloud.dods?ac lc[11954:1:17658][0:1:4][26:1:28][2:1:7])

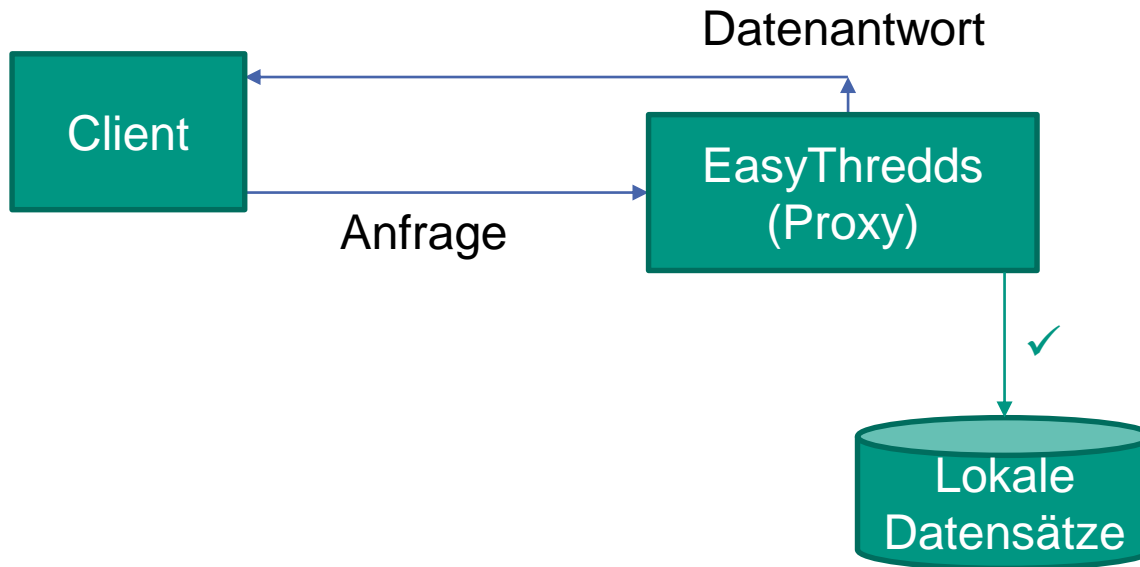
EasyThredds: neue Anfrage



[http://nc-catalogue.scc.kit.edu/thredds/dodsC/RC1SD-base-08/cloud.dods?
aclc\[11954:1:17658\]\[0:1:4\]\[26:1:28\]\[2:1:7\]](http://nc-catalogue.scc.kit.edu/thredds/dodsC/RC1SD-base-08/cloud.dods?aclc[11954:1:17658][0:1:4][26:1:28][2:1:7])

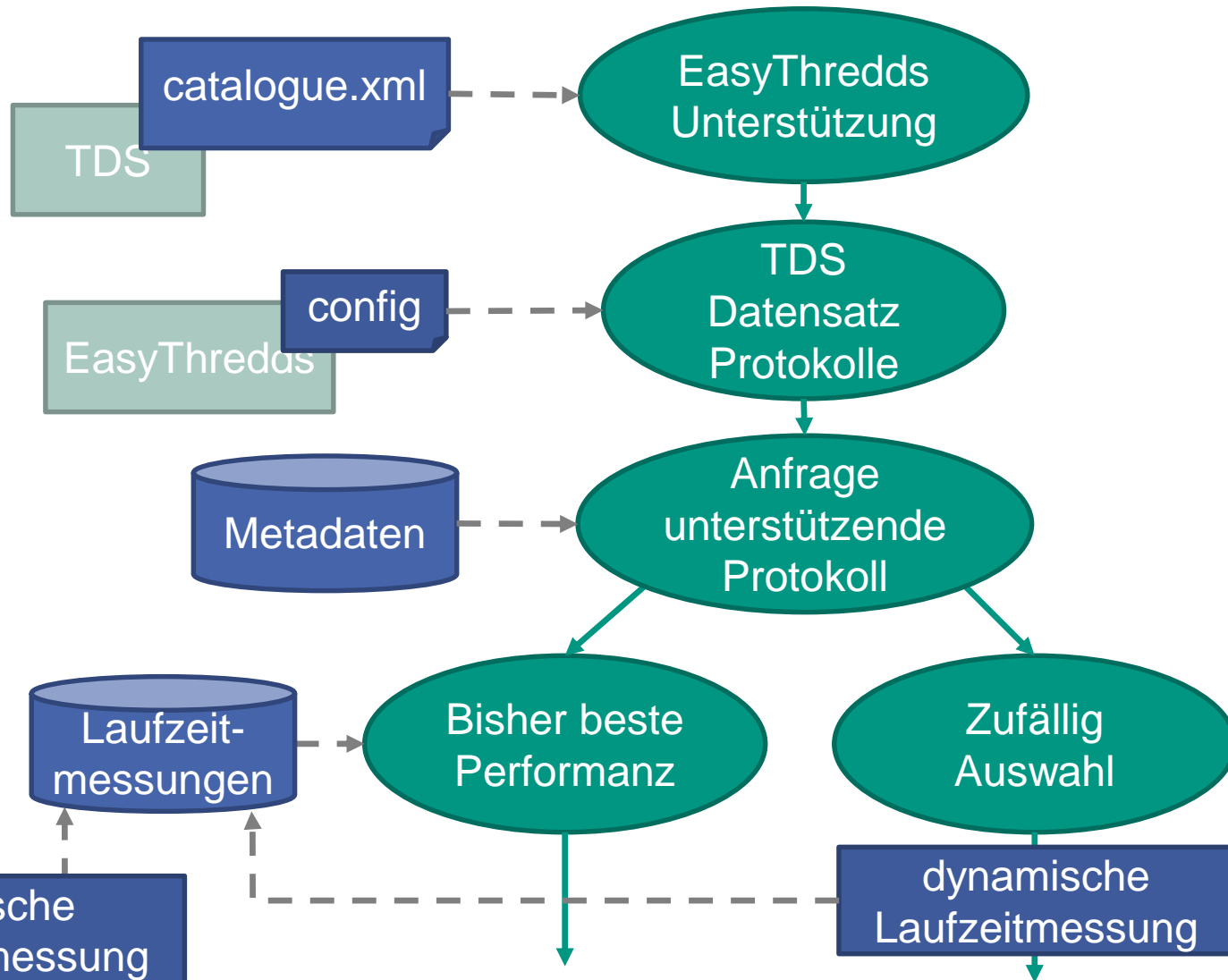
- OPeNDAP-Datenanfrage
- Antwort an Client weiterleiten
- Antwort im Proxy speichern (optional)

EasyThredds: erneute Anfrage



- Koordinaten protokollübergreifend verwendbar
- Datenantwort cachebar

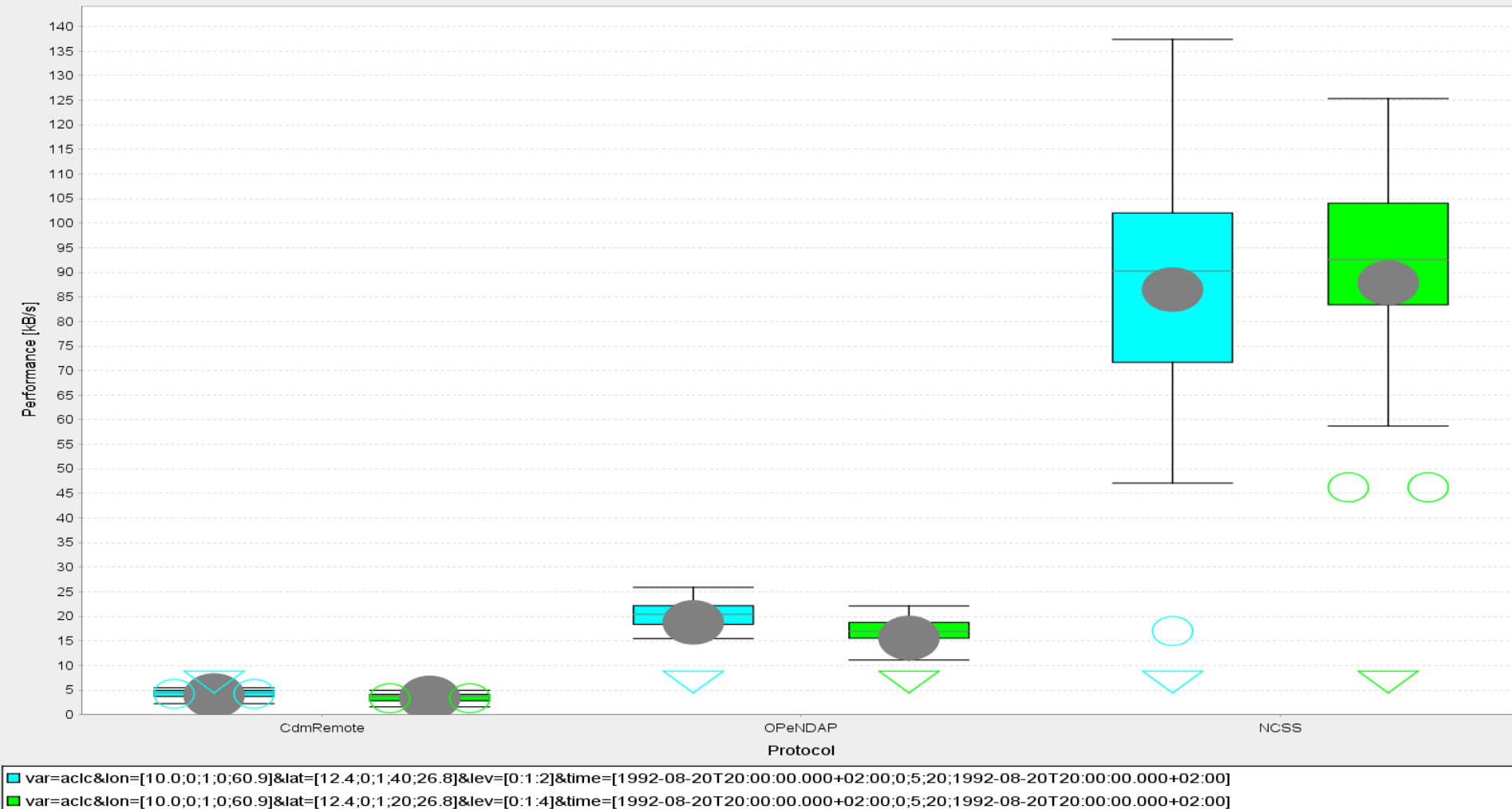
Dynamische Protokollauswahl



Statische Laufzeitmessung

Protocol Performance

Performance for different queries using different protocols



Zusammenfassung

- Einheitliche Anfragesprache für Variablen und Subsetting
- Durch Proxy kein Unterschied für Geowissenschaftler
 - RESTful Design leicht durch Frontend erweiterbar
- Integrierte Protokollwahl
 - Einfach zu Daten gelangen
 - Leicht erweiterbar für weitere Protokoll
 - Dynamische Laufzeitoptimierung

Quellenangaben

[TDS-Overview] - <https://www.unidata.ucar.edu/software/thredds/v5.0/tds/tutorial/TDSOverview.pdf>

[TDS-Facts] - https://www.unidata.ucar.edu/publications/factsheets/current/factsheet_tds.pdf

[CLOUD] - http://nc-catalogue.scc.kit.edu/thredds/imk_rc1sdb8.html?dataset=RC1SD-base-08_cloud

[GIT] - <https://github.com/mcxme/EasyThredds>

BACKUP

Übersetzungen einer Anfrage

`http://localhost:8080/EasyThredds/translate/RC1SD-base-08/cloud?
var=aclc&lev=[0:1:5]&lat=[12.4;1;17.8]&lon=[4.0;;18.9]&
time=[20/08/1992-20:00:00;1;22/02/1999-12:00:00]`



`http://nc-catalogue.scc.kit.edu/thredds/dodsC/RC1SD-base-08/cloud.dods?
aclc[11954:1:17658][0:1:4][26:1:28][2:1:7]`

`http://nc-catalogue.scc.kit.edu/thredds/cdmremote/RC1SD-base-08/cloud?
req=data&var=aclc(11954:17658:1,0:4:1,26:28:1,2:7:1)&accept=ncstream`

`http://nc-catalogue.scc.kit.edu/thredds/ncss/RC1SD-base-08/cloud?
var=aclc&south=12.4&north=17.8&west=4.0&east=18.9&horizStride=1&vertStride=1&
time_start=1992-08-20T20:00:00.000+02:00&time_end=1999-02-
22T12:00:00.000+01:00&time_stride=1&accept=netcdf&addLatLon=true`

netCDF-Java API

- `NetcdfDataset.openFile(...);`
- netCdf-Streaming-Bibliothek fehlerhaft
 - Zugriff nur über Generics
- Subsetting: lokal
- Datenabfrage: remote / lokal

Anfragesprache

Name	Condition	Example
vars	at least one variable is required	vars=aclc
lat	The range has to be within the boundaries 0.0 to 360.0	lat=[12.4;1;17.8]
lon	The range has to be within the boundaries - 90.0 to 90.0	lat=[4.0;;18.9]
lev	-	lev=[0:1:5]
time	The date time format is 'dd/MM/yyyy-HH:mm:ss'	lat=[20/08/1992-20:00:00;1; 22/02/1999-12:00:00]