

VO-Theory



2014.07.07 - Observatoire de Paris

The original goal

- Access state of the art numerical models to interpret observations

The original goal

- Access state of the art numerical models to interpret observations
- Re-use big simulations results

First step

- Each team publishes its simulation output files with a documentation on a webpage

First step

- Each team publishes its simulation output files with a documentation on a webpage
- Sometime, the big teams develop specific small tools to handle their output files

Good but

- No advanced use / analysis of results

Good but

- No advanced use / analysis of results
- No common result format for easy exchange

Good but

- No advanced use / analysis of results
- No common result format for easy exchange
- No re-use of past efforts

It's getting worse

- New instruments produce more & more data

It's getting worse

- New instruments produce more & more data
- New simulations produce more & more data

It's getting worse

- New instruments produce more & more data
- New simulations produce more & more data
- More teams involved

It's getting worse

- New instruments produce more & more data
- New simulations produce more & more data
- More teams involved
- User wants interactive services & tools

It's getting worse

- New instruments produce more & more data
- New simulations produce more & more data
- More teams involved
- User wants interactive services & tools
- User wants accurate data description

We used to have that

parameter_0	<input type="text"/>
parameter_1	<input type="text"/>
parameter_2	<input type="text"/>
parameter_3	<input type="text"/>
parameter_4	<input type="text"/>
parameter_5	<input type="text" value="5.306602734233685e+1;"/>
parameter_6	<input type="text"/>
parameter_7	<input type="text"/>
parameter_8	<input type="text"/>
parameter_9	<input type="text"/>

But have already reached that

parameter_715	<input type="text"/>
parameter_716	<input type="text"/>
parameter_717	<input type="text" value="9.155692234465601e+17"/>
parameter_718	<input type="text"/>
parameter_719	<input type="text"/>
parameter_720	<input checked="" type="text" value="9.379443694351487e+17"/> <input type="text" value="6.068533287565551e+17"/> <input type="text" value="4.593911530233026e+17"/>
parameter_721	<input type="text"/>
parameter_722	<input type="text"/>
parameter_723	<input type="text"/>
parameter_724	<input type="text"/>
parameter_725	<input type="text" value="4.540480604600287e+17"/>
parameter_726	<input type="text"/>
parameter_727	<input type="text"/>
parameter_728	<input type="text"/>



Current goal

- **Easy exchange:** common description format

Current goal

- **Easy exchange**: common description format
- **Fast discovery** of data thanks to rich metadata

Current goal

- **Easy exchange**: common description format
- **Fast discovery** of data thanks to rich metadata
- **Interactive analyse** of huge data

By defining standards

- Reach **interoperability** models \leftrightarrow observations

By defining standards

- Reach **interoperability** models <-> observations
- Have a **common and thorough description of the data**, for later analysis

By defining standards

- Reach **interoperability** models <-> observations
- Have a **common and thorough description of the data**, for later analysis
- Make **standard interfaces to services** to efficiently **use results**

Challenges

- Smart handling of big data
 - storage
 - access

Challenges

- Smart handling of big data
 - storage
 - access
- Smart handling of big metadata
 - mining
 - interactive use

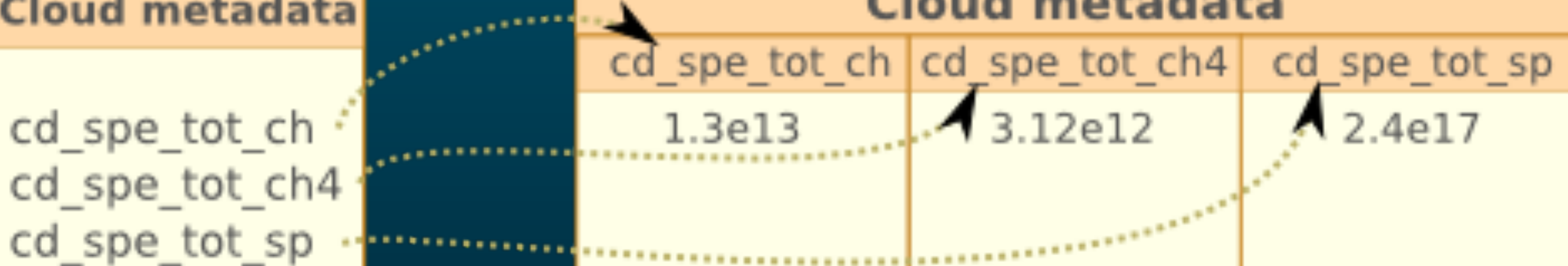
Simulated cloud some years ago

object

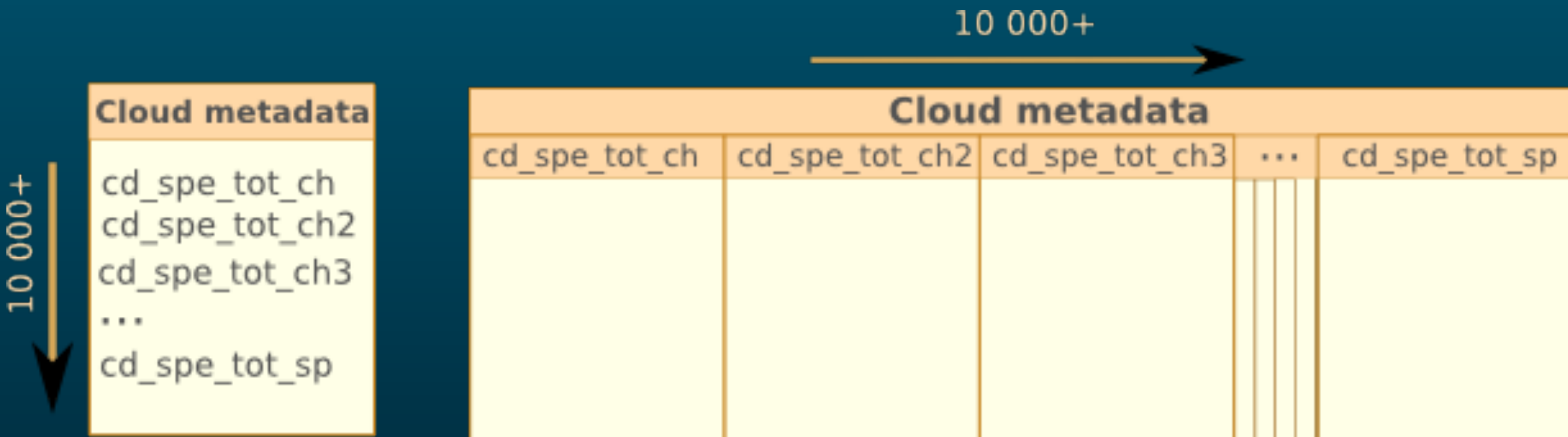
Cloud metadata
cd_spe_tot_ch
cd_spe_tot_ch4
cd_spe_tot_sp

table

Cloud metadata		
cd_spe_tot_ch	cd_spe_tot_ch4	cd_spe_tot_sp
1.3e13	3.12e12	2.4e17



Simulated cloud today



how to deal with so many columns ?

Solutions

- R&D, **prototyping** (a lot !)
- National effort: OV-Paris, ...
- International effort: IVOA
- Right use of technologies
- **Expose work as services**