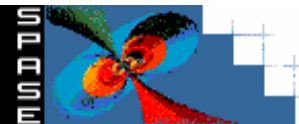
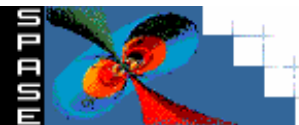


Space Physics Archive Search & Extract



- **SPASE : effort commun pour créer un système de recherche et d'accès aux données**
 - Pour la communauté Space Physics
 - Permettant via une seule interface d'accéder aux divers centres de données internationaux
- **Membres du consortium**
 - » NASA/NSSDC
 - » JHU/APL
 - » IGPP/UCLA
 - » JPL
 - » SWRI
 - » RAL
 - » Augsburg College
 - » CDP



- **Organisation**

- Un steering committee

- 2 sous-groupes

- » Data Model

- » Technical

- Téléconférence environ 2 fois par mois de tout le monde

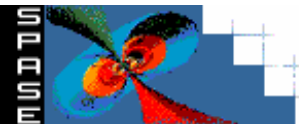
- » Pour l'instant activité très orientée Data Model

- L'objectif étant la création d'un langage commun

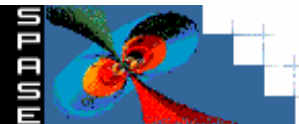
- **Pour l'instant le Data Model est d'un niveau concept**

- Un format XML sera défini

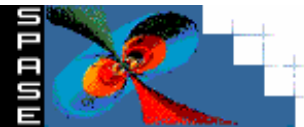
- Le CDPP teste le modèle sur des données Cluster



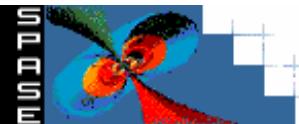
- **Project**
 - The mission or project under which the data was collected.
 - *Example: Cluster*
- **Observatory**
 - The spacecraft, facility or platform that served as host for the instrument that collected the data.
 - *Example: Cluster1, Cluster2, Cluster3, Cluster4*
- **Experiment Type**
 - The type of experiment the instrument performs. This is the technique of observation. For example, an electron density may be measured by a particle detector.
 - *Example: Magnetometer, Particle Detector*
- **Instrument Name**
 - The instrument used to collect the data.
 - *Example: HIC, MAG*
- **Instrument Spatial Region**
 - A characterization of the spatial region the observatory was in while observations were taken.
 - *Example: Magnetosphere, Solar Wind*



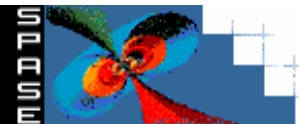
- **Instrument Position**
 - The extent of the region traversed by the observatory while observations were made.
 - *Example: Radial > 5.0, latitude > 10, latitude < 60*
- **Observed Spatial Region**
 - The regime observed by the instrument. This is the physical concept which is observed and measured.
 - *Example: Magnetosphere, Solar Wind, Cloud Tops, Sun, Mars, Mesosphere*
- **Observed Spatial Extent**
 - The extent of the region observed by the instrument. For in-situ instruments this is the same as the Observatory Position Extent, whereas this will differ for remote sensing instruments.
 - *Example: Radial > 5.0, latitude > 10, latitude < 60*
- **Observed Time Span**
 - The span in time over which the observation was taken.
 - *Example: Start > 19990319, Stop < 20030319*
- **Physical Entity**
 - The physical aspect of the observation.
 - *Example: particles, waves, fields, photons*



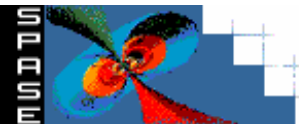
- **Physical Parameter**
 - A property of the observation that is measured by an instrument or derived from observed data.
 - *Example: Velocity, Flux, Energy Range, Wavelength density, temperature, magnitude and direction, fluctuations, spectrum of fluctuations, etc*
- **Measurement Type**
 - A characterization of the type of observation being made by an instrument.
 - *Example: Magnetic Field, Electric Field, Radio and Plasma Waves*
- **Processing Level**
 - The type of processing performed on the observation.
 - *Example: Raw, Calibrated, Detailed, Reduced, Key Parameters, Derived, Catalog*
- **Product Type**
 - A general characterization of the form and source of a product.
 - *Example: Data, Image, Software, Document*
- **Representative Form**
 - The form the observation is represented in.
 - *Example: Image, Tabular, Spectrogram, Magnetogram*



- **Proposition à un Nasa Research Announcement**
 - Récupérer des fonds pour les US members
 - Révèle des divergences "politiques"
 - » Chacun veut le "coeur" du système
 - D'un point de vue techniques le principe du prototype CNES est accepté
 - » Chacun ok pour mettre en place les mécanismes de réponses via des technos AD (Web services, ...)



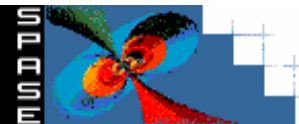
- **Actuel**
 - **Choix :**
 - » **Décrire les requêtes http en XML**
 - Rendre générique l'exécution des requêtes
 - » **Demander des réponses en XML**
 - Présenter les résultats sous une forme homogène
 - **Pour chaque fonctionnalités de chaque centre de données**
 - » **Une description XML**
 - **Exemple : Recherche de jeux de données**
 - URL d'accès
 - Méthodes (GET, POST)
 - Paramètres à passer (nom et type) (Temps début et fin, mots clés, ...)



...

```
<dcRessource type="dataSetSearch">
  <url
    method='GET'
    dtd_res_file='http://idn.cst.cnes.fr:8140/spase/XML/SpaseResult.d
    td'
    xsl_res_file='http://idn.cst.cnes.fr:8140/spase/XML/SpaseResultPr
    esent.xsl'
    >http://sads.cnes.fr:8010/Pub/SadsSeGAD</url>
  <parameters>
    <parameter name="startDate" >
      <urlParameterName>TDeb</urlParameterName>
      <format>YYYY-MM-DDThh:mm:ss</format>
      <valueTest>1999-05-10T01:01:01</valueTest>
      <info>Search on the the real Date and Time of data
        sets</info>
    </parameter>
  </parameters>
</dcRessource>
```

...

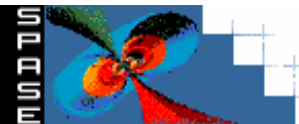


- Un format de réponse commun
- Exemple pour le « dataSetSearch »

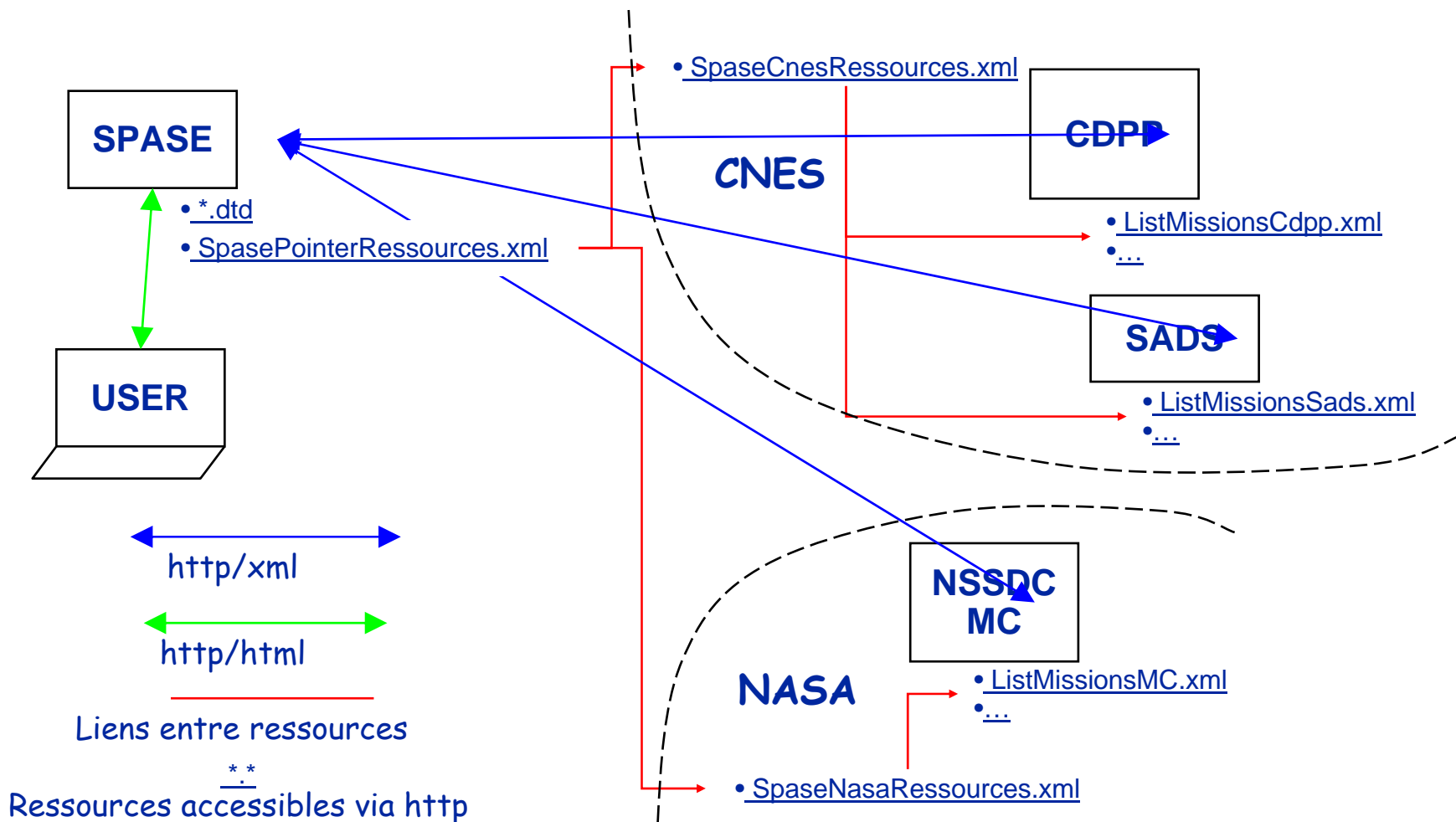
```
<!ELEMENT dataSets (tooManyHits | error | dataSet*)>
```

...

```
<!ELEMENT dataSet (startDateTime, stopDateTIme , mission, experiment,  
dataSetName, description, accessType)>
```



Le prototype



Welcome to the "SPASE Prototype"

What is SPASE ?

SPASE aims to provide an interoperable search tool for the Space Physics community.
Get More information at [SPASE Home Page](#).

What does SPASE prototype provide to the user ?

SPASE queries all the participating data centers for you with your search criteria, and produces a page with all the replies presented in a homogeneous way.

What is a participating Data Center ?

It is a data center which provides the ressources needed to be queried by SPASE.

May I participate in SPASE prototype ?

Any space physics data holding organization can participate. See the [documentation](#) area to learn how to be a part of SPASE prototype.

What are the future developments ?

- More search criteria : (scientific keyword, parameter, location, investigator name, type of instrument, ...)

Choose which data center you want to query :

Participating Data centers :

- SADS(CNES - Space Data Access System)
- CDPP(CNES - Space Physics Data Center)
- NMC(NASA - NSSDC Master Catalog)

Sources Selected

None



Search Query

Please Enter your query

Sources Selected

- SADS
- CDPP
- NMC

| | | SADS | CDPP | NMC |
|--|---------------------|-------------------------------------|-------------------------------------|--------------------------|
| Start date & time time (YYYY-MM-DDThh:mm:ss) : | 1973-09-26T00:00:00 | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Stop date & time time (YYYY-MM-DDThh:mm:ss) : | 2003-09-26T00:00:00 | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Mission or Spacecraft : | ALL ARCAD-3 | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Experiment : | ALL AUREOL-3 | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| Free Text Search : | | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

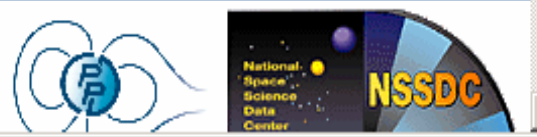
(click on the name of the data center to know more about it)

(position your cursor over the check box for information on how the criteria is applied)

Search Reset

Thierry.Levor@cn.es.fr

NASA Official: J. Thieman (james.thieman@gssc.nasa.gov)





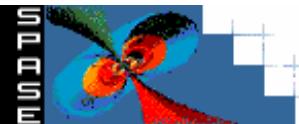
Results from : SADS

| 4 | Start Time | Stop Time | Description |
|--|---------------------|--|-------------------------------------|
| Mission > Experiment > Data Set Name (Access Type) | | | |
| 1 | 1981/09/30-14:22:01 | 1986/03/26-13:39:57 | DENSITY data (N1) (N1) |
| | | ARCAD-3-AUREOL-3 > ISOPROBE > DENSITE (Public) | |
| 2 | 1981/09/29-00:00:00 | 1986/05/15-00:00:00 | ISO Electric Waves Data (N2) (data) |
| | | ARCAD-3-AUREOL-3 > ISOPROBE > TRANSFERT_ISO (Public) | |
| 3 | 1981/01/01-00:00:00 | 1986/12/31-00:00:00 | SPECTRO data (N2) (data) |
| | | ARCAD-3-AUREOL-3 > SPECTRO > TRANSFERT_SPECTRO (Public) | |
| 4 | 1981/01/01-00:00:00 | 1986/12/31-00:00:00 | Spectro Data (N2) (data) |
| | | ARCAD-3-AUREOL-3 > SPECTRO > TRANSFERT_SPECTRO_V2 (Public) | |



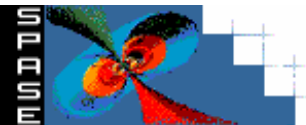
Results from : CDPP

| 10 | Start Time | Stop Time | Description |
|--|-------------------------|--|-------------|
| Mission > Experiment > Data Set Name (Access Type) | | | |
| 1 | 1981/09/29-16:27:05:000 | 1986/05/17-03:06:58:000 | description |
| | | ARCAD-3 > ISOPROBE > ARCAD3 ISOPROBE probe sweeps (Public) | |
| 2 | 1981/09/29-16:27:14:000 | 1986/05/17-03:06:58:000 | description |
| | | ARCAD-3 > ISOPROBE > ARCAD3 ISOPROBE high resolution density waveform (Public) | |
| 3 | 1981/09/29-14:22:01:000 | 1986/03/26-13:39:57:000 | |
| | | ARCAD-3-AUREOL-3 > ISOPROBE > DENSITE (Public) | |



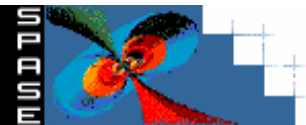
- Perspectives :

- Description des ressources en XML Schema en lieu et place des DTDs, notamment pour le typage et le contrôle des paramètres.
- Ajout de la fonction DataFileSearch
- Ajout de Web Services (SOAP) pour les 2 méthodes DataSetSearch et DataFileSearch avec publication de la description WSDL
- Modification du serveur CDPP pour prendre en charge les requêtes SOAP

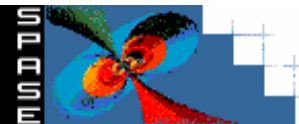


•

FIN



- **Transparents Complémentaires**



- **Les concepts importants**

- Spase peut être

- » un système (serveur) indépendant

- Ou

- » Une option de requête sur un serveur existant (plug-in Spase)

- Spase accède aux « Participating Data Center »

- » Sites ayant décrit leurs ressources

- **Le scénario type**

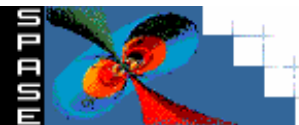
- L'utilisateur saisit une requête

- Spase

- » Interroge tous les sites enregistrés

- » Récupère les résultats

- » Les présente sous une forme homogène



- **Tous les sites n'ont pas les mêmes capacités d'interrogation**
 - Intervalles de temps
 - Nom de mission
 - Paramètres
 - ...
- **Spase gère cela uniquement en informant l'utilisateur :**
 - Des paramètres utilisés pour les requêtes sur chaque serveur
 - Du périmètre de l'interrogation :
 - » Exemple : pour un intervalle de temps demandé, un site peut retourner tous les jeux de données qui sont :
 - Complètement inclus dedans
 - L'intersection est non nulle
 -
 - » Cette information est fournie par le site dans la définition de ses ressources



- **Plusieurs niveaux**

- Niveau Configuration de SPASE

- » Pointeurs sur toutes les ressources

- Centralisé pour l'instant, mais peut-être à répartir

- Niveau Domaine

- » Une description de ressources par entité (ex CNES ou MPI/PS)

- Niveau Centre de données

- » Description d'un centre (ex : CDPP)

- Niveau Fonctionnalités

- » Recherche de jeux de données

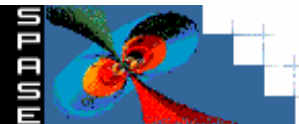
- URL d'accès
- Méthodes
- Paramètres à passer (nom et type)
- DTD de la réponse (pour l'instant même DTD pour tout le monde)

- » Commande de données

- Idem

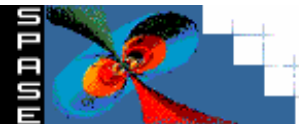
- Niveau « Thesaurus » (plutôt dictionnaire)

- » Nom de missions, expériences, paramètres gérés dans un serveur



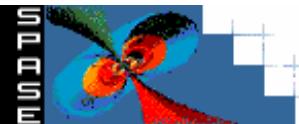
La Configuration SPASE

```
<?xml version="1.0"?>
<!-- Spase ressources -->
<!-- 17/01/2001 -->
<!-- Thierry.Levoir@cnes.fr -->
<!DOCTYPE sourcesDonnee SYSTEM "http://lumiere.cnes.fr:8140/XML/SpaseRessources.dtd">
<participatingInstitutes>
  <participatingInstitute>
    <institute>CNES</institute>
    <UrlSpaseDataSources>
      http://sads.cnes.fr:8010/XML/SpaseDataSources_cnes.xml
    </UrlSpaseDataSources>
  </participatingInstitute>
  <participatingInstitute>
    <institute>NASA</institute>
    <UrlSpaseDataSources>
      http://nssdc.gsfc.nasa.gov:8080/XML/SpaseDataSources_nasa.xml
    </UrlSpaseDataSources >
  </participatingInstitute>
</participatingInstitutes>
```

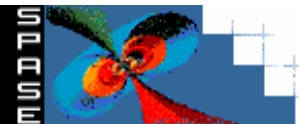


Configuration d'un domaine

```
<?xml version="1.0"?>
<!-- Data Center ressources CNES -->
<!-- 17/01/2001 -->
<!-- Thierry.Levoir@cnes.fr -->
<!DOCTYPE SpaseDataSources SYSTEM
"http://idn.cnes.fr:8140/XML/SpaseDataSources.dtd">
<dataCenters>
  <origin
    institute="CNES"
    author="Thierry Levoir"
    authorEMail="Thierry.Levoir@cnes.fr"
    lastUpdated="2001-01-17"/>
  <dataCenter name="SADS" lastUpdated="2000-08-26">
    . . .
  </dataCenter>
  <dataCenter name="CDPP" lastUpdated="2000-08-26">
    . . .
  </dataCenter>
</dataCenters>
```



```
<dataCenter name="SADS" lastUpdated="2000-08-26">
  <dcTitle>CNES - Space Data Access System </dcTitle>
  <dcLogoURL>http://sads.cnes.fr:8010/icons/sigle_s0.gif</dcLogoURL>
  <dcURL>http://sads.cnes.fr:8010/</dcURL>
  <dcDescription>
    The SADS system provides a user-friendly access to many space physics data stored at the CNES space center
    in Toulouse, France. .
  </dcDescription>
  <dcRessource type="dataSetSearch">
    <url
      method='GET'
      dtd_res_file='http://idn.cst.cnes.fr:8140/spase/XML/SpaseResult.dtd'
      xsl_res_file='http://idn.cst.cnes.fr:8140/spase/XML/SpaseResultPresent.xsl'
      >http://sads.cnes.fr:8010/Pub/SadsSeGAD</url>
    <parameters>
      <parameter name="startDate" >
        <urlParameterName>TDeb</urlParameterName>
        <format>YYYY-MM-DDThh:mm:ss</format>
        <valueTest>1999-05-10T01:01:01</valueTest>
        <info>Search on the the real Date and Time of data sets</info>
      </parameter>
      . . .
      <parameter name="mission" >
        <urlParameterName>mission</urlParameterName>
        <valueTest>INTERBALL</valueTest>
        <list>http://sads.cnes.fr:8010/Pub/SadsSpaseListMission</list>
        <info>Search on mission or spacecraft name as defined in SADS (case sensitive)</info>
      </parameter>
    </dcRessource>
    . . .
  </dataCenter>
```



La DTD des ressources (1)

...

```
<!ELEMENT dataCenters (origin, dataCenter+)>
```

```
<!ELEMENT origin EMPTY>
```

```
<!ATTLIST origin institute CDATA #REQUIRED>
```

```
<!ATTLIST origin author CDATA #REQUIRED>
```

```
<!ATTLIST origin authorEMail CDATA #REQUIRED>
```

```
<!ATTLIST origin lastUpdated CDATA #REQUIRED>
```

```
<!ELEMENT dataCenter (dcTitle, dcLogoURL, dcURL, dcDescription, dcRessource*)>
```

```
<!ATTLIST dataCenter name CDATA #REQUIRED>
```

```
<!-- short name of the data center -->
```

```
<!ATTLIST dataCenter lastUpdated CDATA #REQUIRED>
```

```
<!-- date of the last modification of the data center record -->
```

```
<!ELEMENT dcTitle (#PCDATA)>
```

```
<!-- long Name of the data center -->
```

```
<!ELEMENT dcLogoURL (#PCDATA)>
```

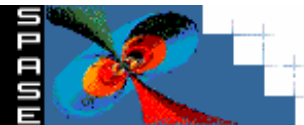
```
<!-- logo (gif or jpg) URL of the data Center -->
```

```
<!ELEMENT dcURL (#PCDATA)>
```

```
<!-- data center URL -->
```

```
<!ELEMENT dcDescription (#PCDATA)>
```

```
<!-- data center quick description (max 6 lines) -->
```

La DTD des ressources (2)

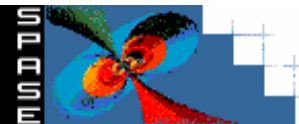
```
<!ELEMENT dcRessource (url, parameters, infoRes? )>
<!-- one of the data center ressource -->
<!ATTLIST dcRessource type (dataSetSearch | dataGranuleSearch) #REQUIRED >
<!-- type of the ressource (at the moment only dataSetSearch is implemeted) -->

<!ELEMENT url (#PCDATA)>
<!-- url of the ressource -->
<!ATTLIST url method (GET | POST) #REQUIRED >
<!-- http method to use (at the moment only GET is implemeted) -->
<!ATTLIST url dtd_res_file CDATA #FIXED 'http://idn.cst.cnes.fr:8140/spase/XML/SpaseResult.dtd' >
<!-- dtd of of the result : FIXED -->
<!ATTLIST url xsl_res_file CDATA #FIXED 'http://idn.cst.cnes.fr:8140/spase/XML/SpaseResultPresent.xsl' >
<!-- xsl to present the result : FIXED -->

<!ELEMENT parameters (paramater* )>
<!-- parameter to be passed to the URL i.e. : http://.../cgi-bin/dataSetQuery?par1=....&par2=... -->

<!ELEMENT parameter (urlParamaterName, format?, valueTest, list?, info)>
<!ATTLIST parameter name (startDate | stopDate | mission | experiment | freeTextSearch ) #REQUIRED >

<!ELEMENT urlParamaterName (#PCDATA)>
...
```



- Pour une même information les serveurs utilisent des noms différents

- Exemple pour les parameters:

- » Pour une propriété « vitesse »

| CDPP | NSSDC-MC | SWRI |
|--|--|---|
| Charged particles > Velocity* Neutral particles > Velocity* | Neutral particles > Speed* Charged particles > Speed* | Neutrals > Speed* Neutrals > Winds* Neutrals > Velocity Fields* Charged particles > Speed* |

- Exemple pour les missions

- » Interball Auroral Probe (NSSDC)

- » Interball-Auroral (CDPP)

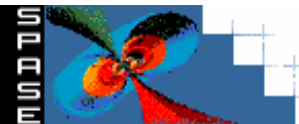
- Utilisation d'un Identifieur unique pour chaque info

- Les centres fournissent une liste de

- » ID

- » Common name (nom utilisé par le serveur)

- » Synonymes (s'il y en a)



- **Spase**

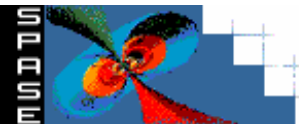
- Permet à l'utilisateur de choisir parmi Les « common name » ou synonymes
- Fait le lien entre les infos des différents centres grâce à l'ID.
- Interroge les sites avec leurs « common name »

- **Avantage de ce mécanisme**

- Pas de modification pour les centres de données, SPASE utilise leurs « common name » pour les interroger.

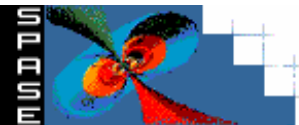
- **Contrainte**

- Les centres de données doivent fournir les « listes de conversions » (Peut être statique)
- Pour les paramètres scientifiques, cela nécessite un travail de la part des chercheurs, pour éviter des interprétations erronées (travail en cours)

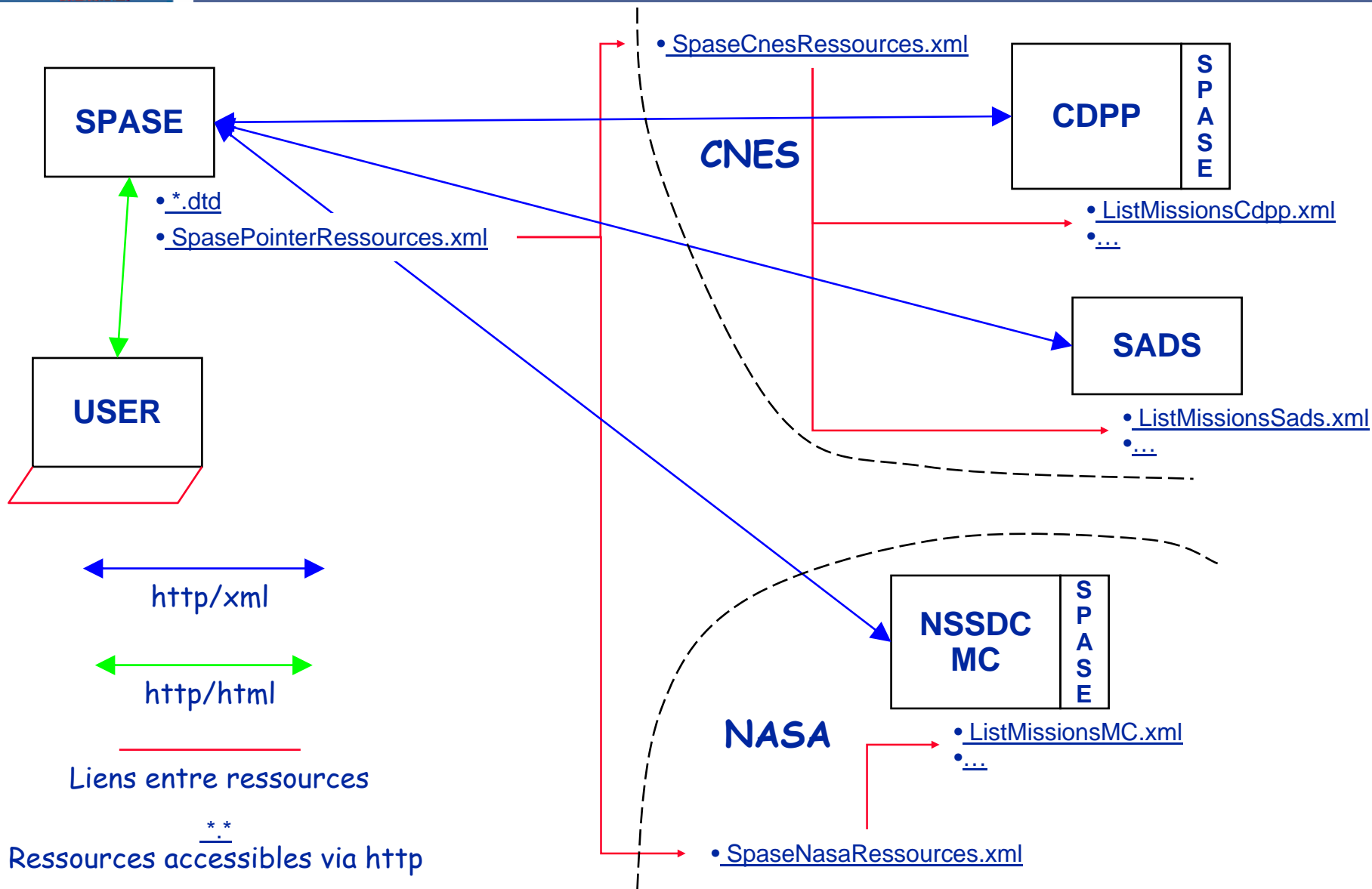


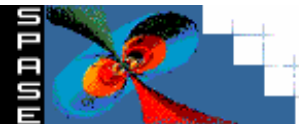
Exemple de « dictionnaire »

```
<?xml version="1.0"?>
<!-- Mission Value List for SADS -->
<!-- 18/12/2000 -->
<!-- Thierry.Levoir@cnes.fr -->
<!DOCTYPE sourcesDonnee SYSTEM
  "http://128.183.190.63:8080/XML/ListParameter.dtd">
<list>
  <parameter>
    <ID>1986-019B</ID>
    <commonName>VIKING</commonName>
    <synonyme>V4</synonyme>
  </parameter>
  <parameter>
    <ID>1996-050C</ID>
    <commonName>AURORAL</commonName>
    <synonyme>INTERBALL</synonyme>
    <synonyme>Auroral Probe</synonyme>
    <synonyme>Prognoz 2M</synonyme>
    <synonyme>Interball 2</synonyme>
    <synonyme>Prognoz 12</synonyme>
    <synonyme>24293</synonyme>
  </parameter>
</list>
```



SPASE : architecture



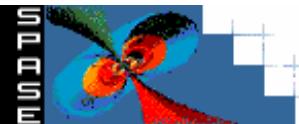


- Utilisation de l'API DOM et SAX Xerces d'Apache
 - DOM pour Document Object Model
 - SAX pour Simple API for XML
 - Fournit des méthodes d'accès à l'arbre XML

```
DOMParser parser = new DOMParser();
try {
    parser.parse(uri);
    // Récupération de la racine
    Document document = parser.getDocument();
    Element root = document.getDocumentElement();

    // Récupération de la liste des élément fils
    NodeList l = root.getElementsByTagName("dataCenter");

    . . .
```



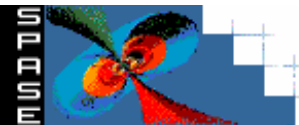
- Les centres de données respectent une DTD pour fournir leur réponse au requête.
 - Pour le « dataSetSearch »

```
<!ELEMENT dataSets (tooManyHits | error | dataSet*)>
```

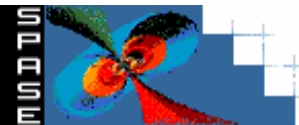
...

```
<!ELEMENT dataSet (startDateTime, stopDateTIme , mission, experiment, dataSetName, description, accessType)>
```

...
- Spase récupère toutes les réponses et les affiche en HTML en utilisant une transformation XSLT
 - Intérêt :
 - » À l'insertion d'une nouvelle ressource on peut gérer la modification de la DTD (ajout d'un élément dans la réponse) sans changement de code : seul le XSL est à mettre à jour.
 - » Cela permet aux utilisateurs du code SPASE (dans le cas de l'utilisation de SPASE comme un plug-in) de présenter les réponses comme ils ont l'habitude de le faire.



```
<?xml version="1.0"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
version="1.0" > <xsl:output method="html"/>
<td class="normal">
  <xsl:attribute name="nowrap"/>
  <xsl:apply-templates select="startDateTime"/>
</td>
. . .
<xsl:template match="startDateTime">
  <xsl:value-of select="."/>
</xsl:template>
. . .
<xsl:template match="mission">
  <xsl:choose>
    <xsl:when test="./@url!=''">
      <A style="text-decoration: underline" HREF="{@url}">
        <xsl:value-of select="."/></A>
    </xsl:when>
    <xsl:otherwise>
      <xsl:value-of select="."/>
    </xsl:otherwise>
  </xsl:choose>
</xsl:template>
```

- Utilisation du processeur Xalan du projet Apache

```
public static void init() {
    try {
        _processor = XSLTProcessorFactory.getProcessor();
        . . .

    /** Convert XML String to HTML String */
    public static String convert(String chaineXML, String fichierStyle) {
        . . .
        XSLTInputSource xmldoc = new XSLTInputSource(new
StringReader(chaineXML));
        ByteArrayOutputStream baos = new ByteArrayOutputStream();
        XSLTResultTarget outBuffer = new XSLTResultTarget(baos);

        try {
            _processor.process(xmldoc, new XSLTInputSource(fichierStyle), outBuffer);
        }
        . . .
        return(baos.toString());
        . . .
    }
}
```