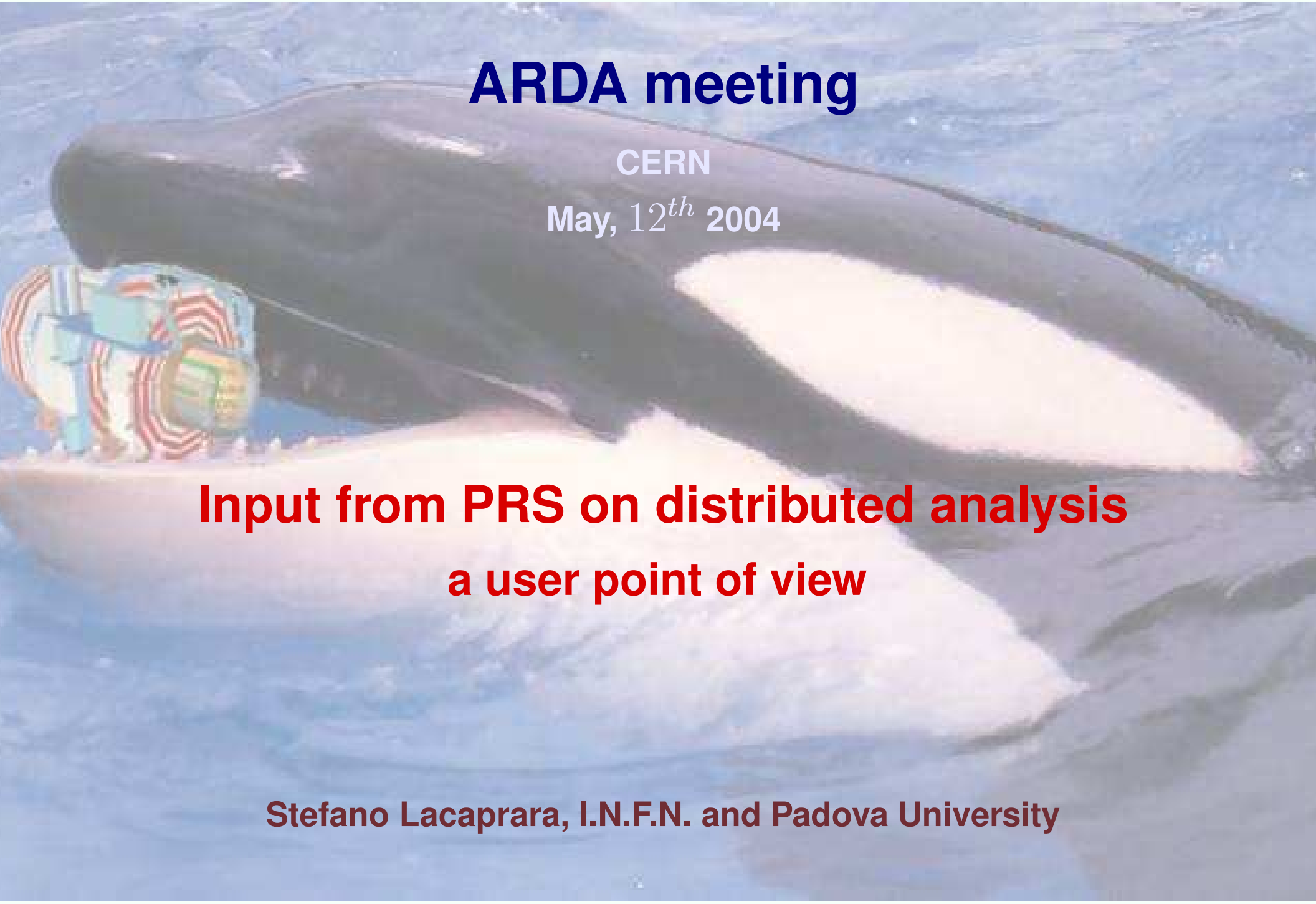

ARDA meeting

CERN

May, 12^{th} 2004

Input from PRS on distributed analysis

a user point of view

Stefano Lacaprara, I.N.F.N. and Padova University

Building Blocks

- Want to access data easily and efficiently,
- Data access should not be different from interactive access (eg as shown on ORCA tutorial),

What do we need

- Information: know what is avaialble and where,
- Software: up–to–date software installed everywhere,
- Data availability: access data early,
- Data access: full access to data produced,
- Use cases: private production, access to chunk of events,...

Information

- Today the only way to get info about what has been produced is the RefDB
- RefDB is a production tool, not a end–user information source
- Has too much info for the final user, and sometime too few
- Need for a PRS user
  - ⋆ name of dataset/owner(s), with flow for hits/digi/dst owner
  - ⋆ how many events
  - ⋆ type of events (cards+software version)
  - ⋆ integrated luminosity
  - ⋆ ...
- All available in RefDB: need a PRS user interface

CMS MC Production Page

DatasetName	GeneratedNbOfEvts	GenSelectedNbOfEvts	Sim.NbOfEvts	CollectionName
mu03_tt2mu	24212579	600000	0	

Collection ID/Name	User Federation(s)	Nb. Valid Evts
3743: Generation genPYT102, valid runs	Done but Not yet available	0
3962: /System/mu_Hit241_g133/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	552531
4279: /System/sw_Hit2451_g133/mu03_tt2mu/mu03_tt2mu, valid runs	Production not yet started	0
4429: /System/mu_Hit245_2_g133/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	99736
4931: /System/mu_2x1033PU761_TkMu_2_g133_OSC/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	90994
5171: /System/mu_DST771_2_3_g133_CMS/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	0
5181: noname, valid runs	Production not yet started	0
5182: /System/mu_DSTs800_2_3_g133_CMS/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	545960
5234: /System/mu_DSTs801_5_g133_OSC/mu03_tt2mu/mu03_tt2mu, valid runs	Production in progress	545120

Production

- Up to now production done in huge bunches, not continuously
- Software availability has always been a problem (maybe the major)
- We should aim to a continuos MC production: user (PRS) ask for a dataset with given sw and cards and have results after a *short* delay (days? weeks? not months!)
- Use huge computing power: eventual priority for analysis

- How to put a MC request: now is rather complex!
- Agree on simplified procedure, accessible to generic user
- or centralized the request (per PRS, as today) so that user should ask the PRS to submit request what he needs

- Foresee *private* production.
- A user (typically phd students...) needs to produce quickly small amounts of events: need official PU, not to re–invent production tools, official procedure in order to obtain good simulation, etc
- Events produced can be used by the community: publication

Data Availability

- In order to run on data, we need dataset with runs attached to COBRA MetaData
- A dataset is really *"produced"* only when the MetaData are attached
- Winter mode access is a very specific and complex way to access data, not suitable for end user
- If producing (digis, DST, ...) a dataset takes time (as is now), not want to wait for accessing partial data
- Fraction of dataset (with fully attached MetaData) should be available soon: not real-time, but not only at the end of the production either. Say every few days or a week.
- This would allow also for early data validation!
- User need also a Pool catalog

- Proposal: create local Pool catalog(s) on Tn where data are shipped
- When a given Tn get a dataset (or a fraction), create a local catalog (xml or mysql)
  - ⋆ contains the **lfn** and **metadata** of all transferred files
  - ⋆ **pfn** related to local filesystem (including eventual access protocol rfio: dcache: etc...)
- The local catalogs is published and can be used on the local farm to access local data
- Publication can be web page or (better) RLS

- If RLS, can use local catalog for automatic data discovery
- Can publish also (as RPS metadata) more information about the size of collection and which data type ara available

lfn=PoolCatalog-tt2mu-DST_812-LNL.xml

pfn=/data/catalog/PoolCatalog-tt2mu-DST_812-LNL.xml

metadata dataset=tt2mu

metadata owner=DST_812-LNL

metadata eventRange=1-100000

metadata content=DST

metadata Tn=LNL

...

- Example:
  - User want study $t\bar{t} \rightarrow 2\mu$
  - Look for suitable dataset in datasets list: find dataset name and owner name
  - Search for available data for that dataset/owner: query for a local catalog on RLS (query for few files)
  - Result: LNL events 1-10000, CNAF 1-20000, etc...
  - Want to run on 10000 events, so correct site is LNL
  - Use **edg** to submit job to LNL (or where a suitable local catalog is available)
  - User `.orcarc` contains `InputFileCatalogURL` from result of RLS query
  - Tested personally on $\mathcal{O}(1000)$ from PD UI to LNL: it works!
  - No major problems found, can give real feedback to edg people to improve things, but is definitively usable right now!

- Needs:
  - Tn must provide local catalog when data arrives
  - Catalog must be kept up–to–date (in case of data movement)
- Pros:
  - ◇ works **now** (data discovery not yet tested)
  - ◇ Does not need to create a catalog for every jobs which is submitted: use a *cached* one
  - ◇ Can use data discovery using local catalog as "tag" for event collection, instead of looking for $\mathcal{O}(1000)$ files, look for just fews
- Cons:
  - ⋆ What if a file is missing? Job crashes!
  - ⋆ Future: if a file is required by COBRA but is not available locally, trigger (only for that file) a query on a file catalog and copy it from somewhere else.
  - ⋆ Not needed for all accessed file, not needed at all if dataset integrity is guaranteed

Use Case

- Want to look (via visualization) at a specific event which is somewhere
- Need to access full info (SimHits, Digis)
- Not need to copy locally all EVD files, just to read them once
- WAN access to (small) fraction of data, without full local copy (as on good–old AMS/Objectivity days...)