
dCache у нас.

dCache: обзор

- dCache это система доступа к данным, ориентированная на хранение больших объёмов экспериментальной информации в GRID среде. Может использоваться независимо от GRID.
- Основная идея: представить множество дисковых серверов как дисковое пространство под центральным управлением с автоматической оптимизацией и иерархической системой имён с одним корнем. Отличие от AFS – может включать MSS, акцент на оптимизацию.

- Авторы:

dCache/SRM is a joined effort between the Deutsches Elektronen-Synchrotron (DESY), Fermi National Accelerator Laboratory (FermiLab) and support from the University of California, INFN, Bari, GridPP Rutherford Appleton Laboratory, CERN.

dCache: масштаб (Patrick Fuhrmann et al. 16.10.2007 BE-grid)
<http://www.dcache.org/manuals/d-grid/begrid-dcache-16102007.pdf>

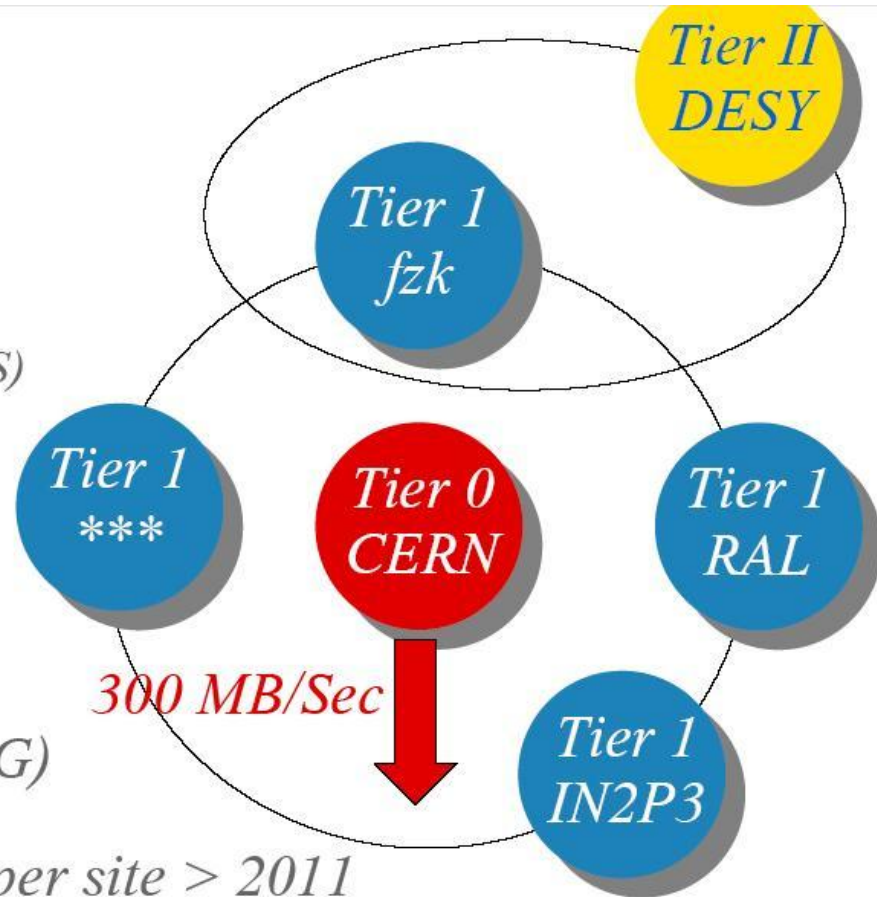
In use at 9 Tier I centers

- fzk (Karlsruhe, GR)
- in2p3 (Lyon, FR)
- RAL (Rutherford, UK)
- BNL (New York, US)
- FERMILab (Chicago, US)
- SARA (Amsterdam, NL)
- PIC (Spain)
- Triumf (Canada)
- NDGF (NordGrid)

About 40 Tier II's

dCache is part of VDT (OSG)

We are expecting > 20 PB per site > 2011



dCache will hold the largest share of the LHC data.

dCache: внутренняя структура

- Несколько десятков или сотен дисковых партиций объединяются в одно пространство и разделяется на именованные пулы, которые скрыты от пользователя.
- Имена файлов объединяются в одно пространство с единственным корнем (это скорее недостаток), для хранения метайнформации используется база данных.
- Файл может иметь несколько копий, это выравнивает нагрузку, повышает надёжность. Использование пространства оптимизируется.
- Пулы привязаны к IP с которого производится доступ к файлу и к директории в которую попадает файл.
- Доступ к системе производится через выделенные серверы (двери)

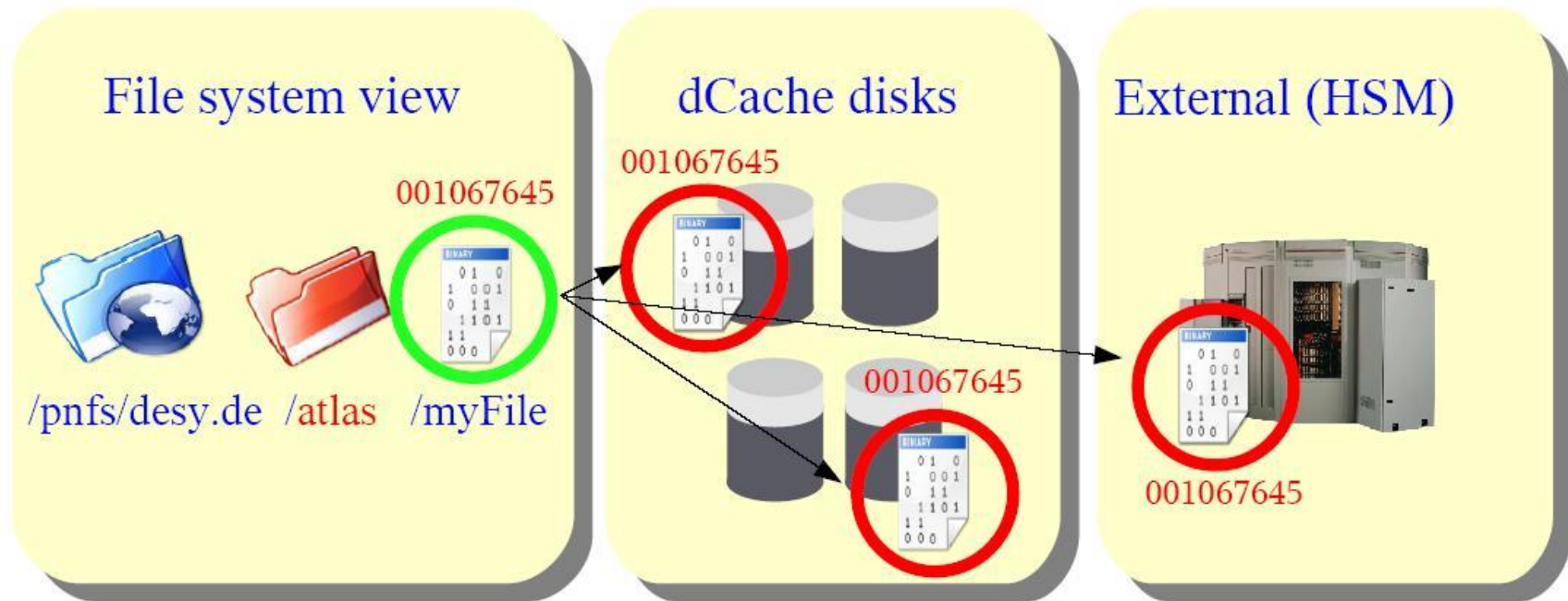
dCache: с точки зрения пользователя

- Основной принцип: информация один раз записывается, много раз читается, никогда не модифицируется, так в GRID.
- Прозрачно поддерживаются команды LINUX, работающие с метайнформацией (ls, chmod, du...).
- Для чтения файла средствами ОС, он должен быть скопирован в локальное дисковое пространство.
- Приложение собранное с библиотекой dcar, которая обеспечивает POSIX - доступ (open, read, write...) работает с файлами без промежуточного копирования.
- dCache автоматически оптимизирует использование дискового пространства, может хранить несколько копий файла.

dCache: методы доступа

- Локальное управление пространством имён через nfs 2/4.
- Из локальной сети - dCap и его клоны (обычный, защищённый Kerberos, ssl, gsi), очень шустрый протокол.
- Из WAN - ftp (обычный, защищённый gsi gFTP).
- Storage Resource Manager 2.2 (Storage Element на основе dCache), это не протокол, а организатор передачи между сайтами.
- XROOT – пока толком не работает.

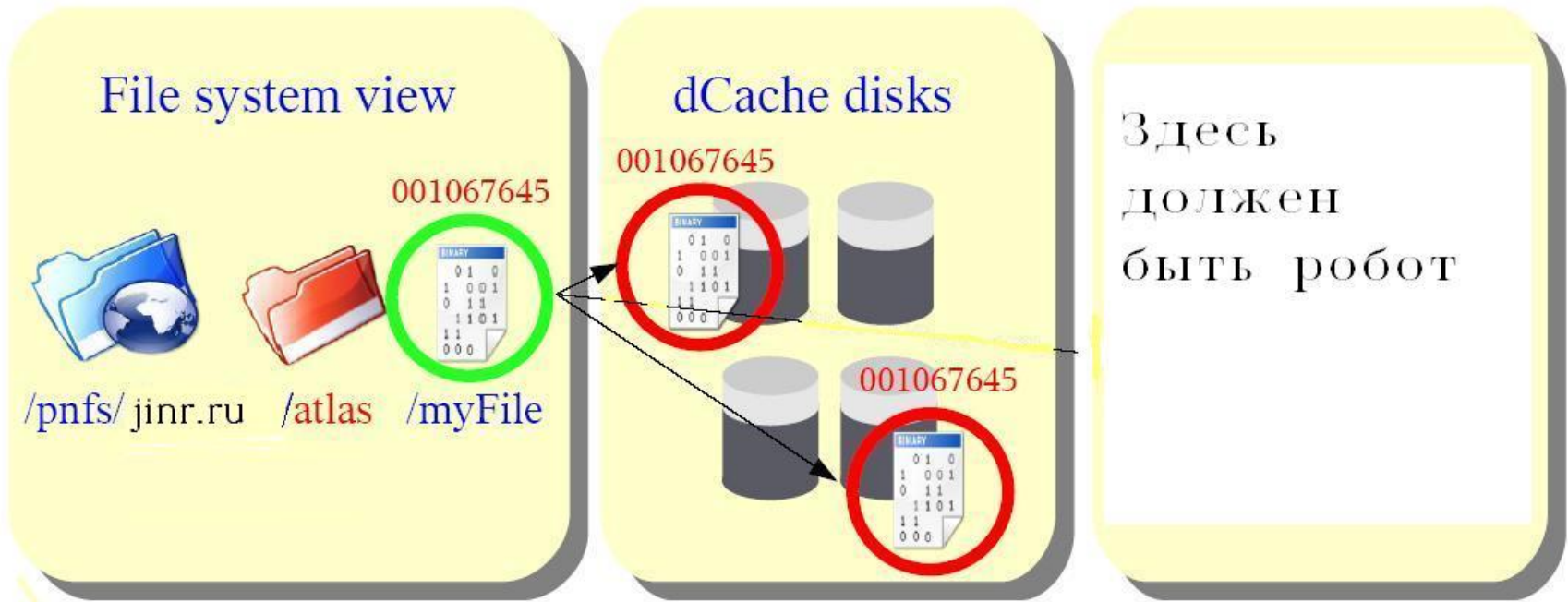
dCache: как это у белых людей



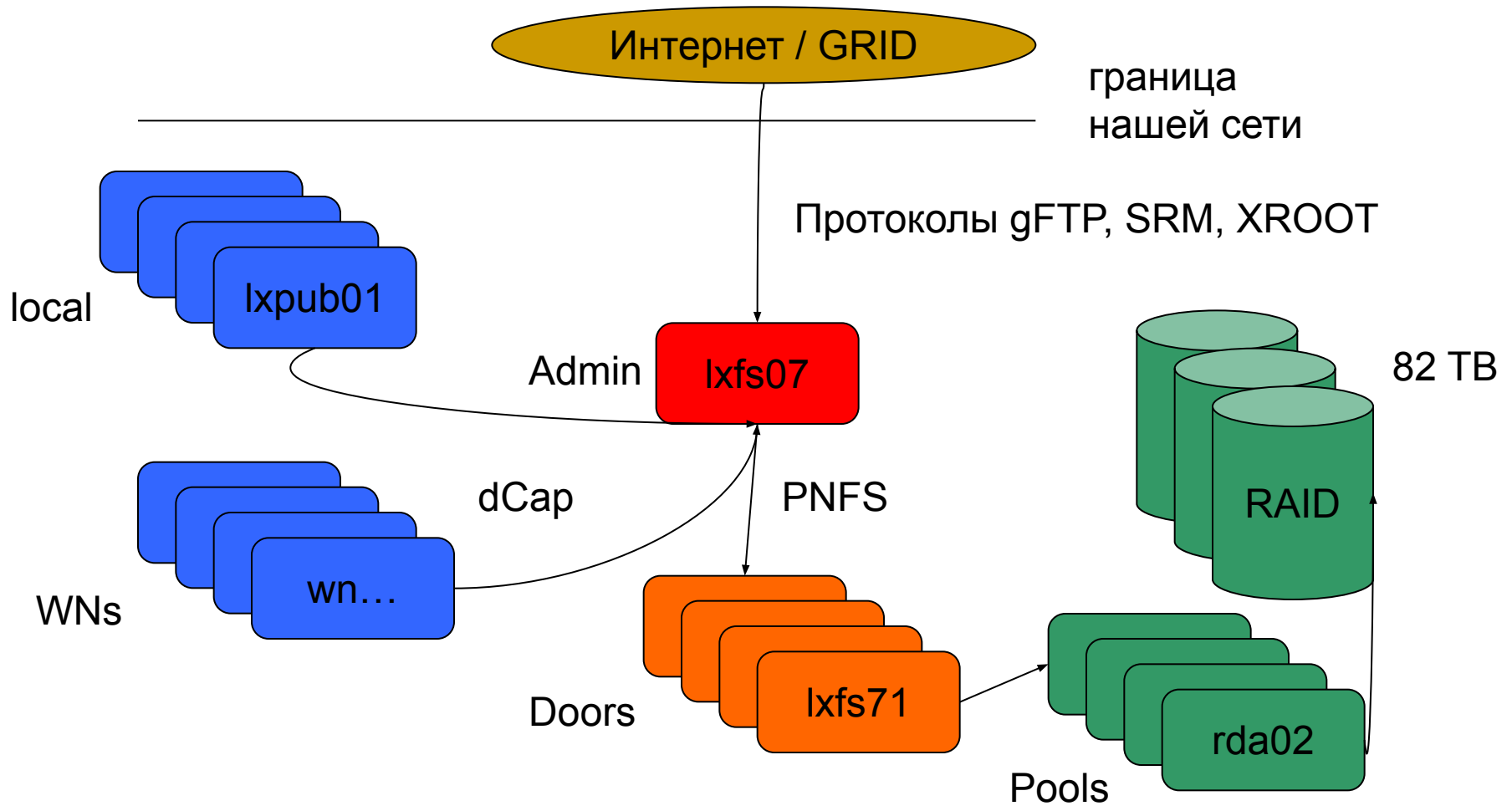
dCache: статистика T1 FNAL. Кстати e+13 это 10 ТВ.



dCache: как это у нас



dCache в ЦИВК ЛИТ тех. структура



dCache протоколы работающие в ОИЯИ.

- У нас доступны следующие:
- dCap, аналог команды ср. Доступен с локальной сети ОИЯИ, не требует GRID сертификата;
- gFTP, аналог ftp, требует сертификат;
- в ОИЯИ на основе dCache построен SE, для доступа к которому используется SRM;
- дверь для всех одна lxfs07.

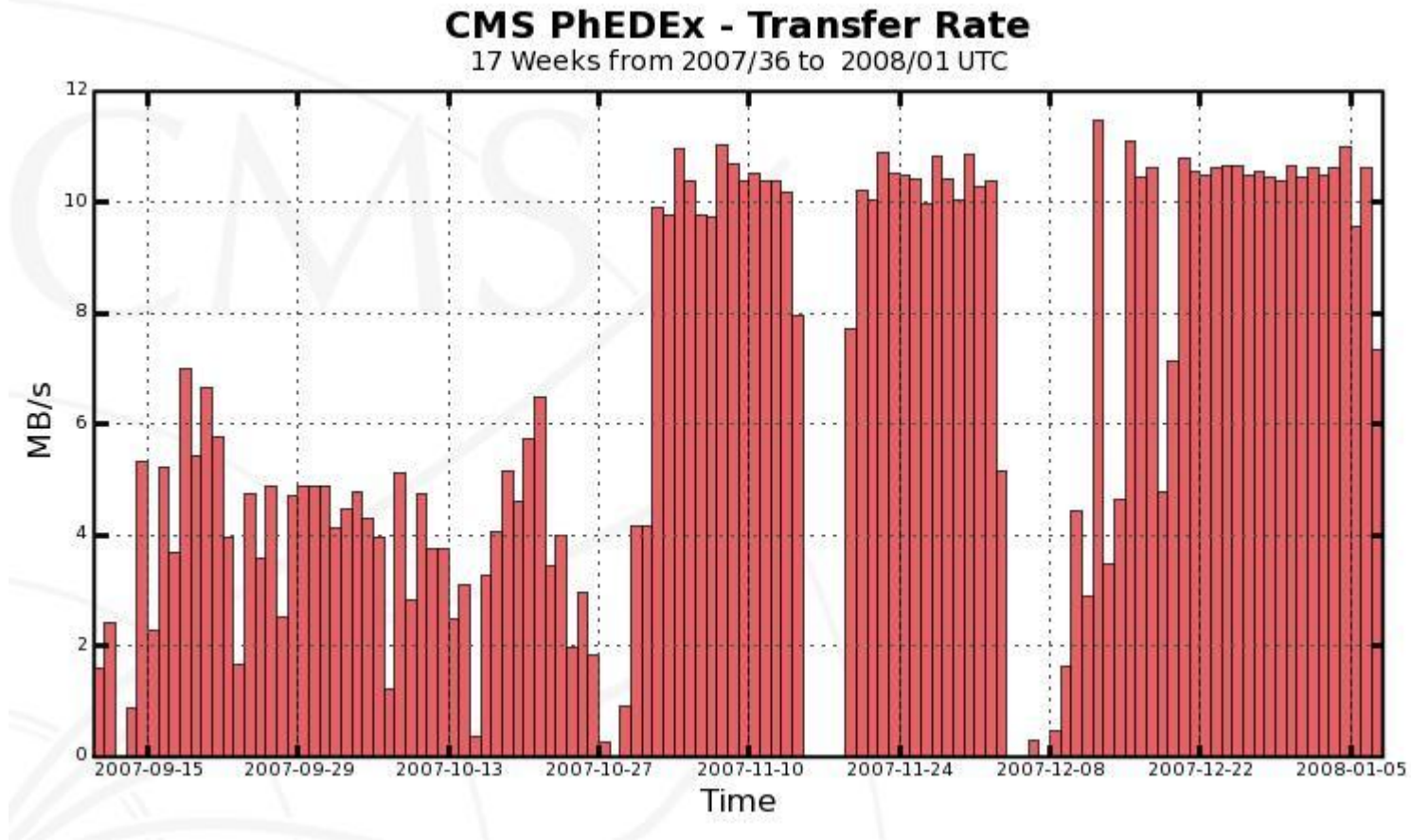
dCache: доступ к файлу из сети 159.93 (ОИЯИ)

- Всё растёт из корня /pnfs/jinr.ru
- Синтаксис простой, но утомительный
- Можно монтировать на локальный компьютер
mount -o lxfs07.jinr.ru:/pnfs /pnfs/jinr.ru
- После монтирования работают команды LINUX ls, mv, rm....
- Если директория /pnfs/jinr.ru смонтирована, то можно указывать файл как в LINUX (аналог AFS)
dccc \$HOME/nobel_price.row /pnfs/jinr.ru/user/i/i_am
- Если директория не монтируется – то URL нотация
dccc dcap://lxfs07.jinr.ru:22125/pnfs/jinr.ru/data/cms/file \
~/destinationfile

dCache для членов виртуальных организаций

- Для всех протоколов, кроме dCap
/pnfs/jinr.ru/data/<виртуальная организация>/.....
- использовать команды gFTP:
globus-url-copy file:/tmp/test \
gsiftp://lxf07.jinr.ru:2811/pnfs/jinr.ru/data/dteam/testfile2
- или srmcp:
srm://lxf07.jinr.ru:8443/pnfs/jinr.ru/data/dteam/testfile2 \
file:///tmp/backtest
- Осторожно использовать локальные команды Linux, поскольку это может привести к противоречиям в LFC

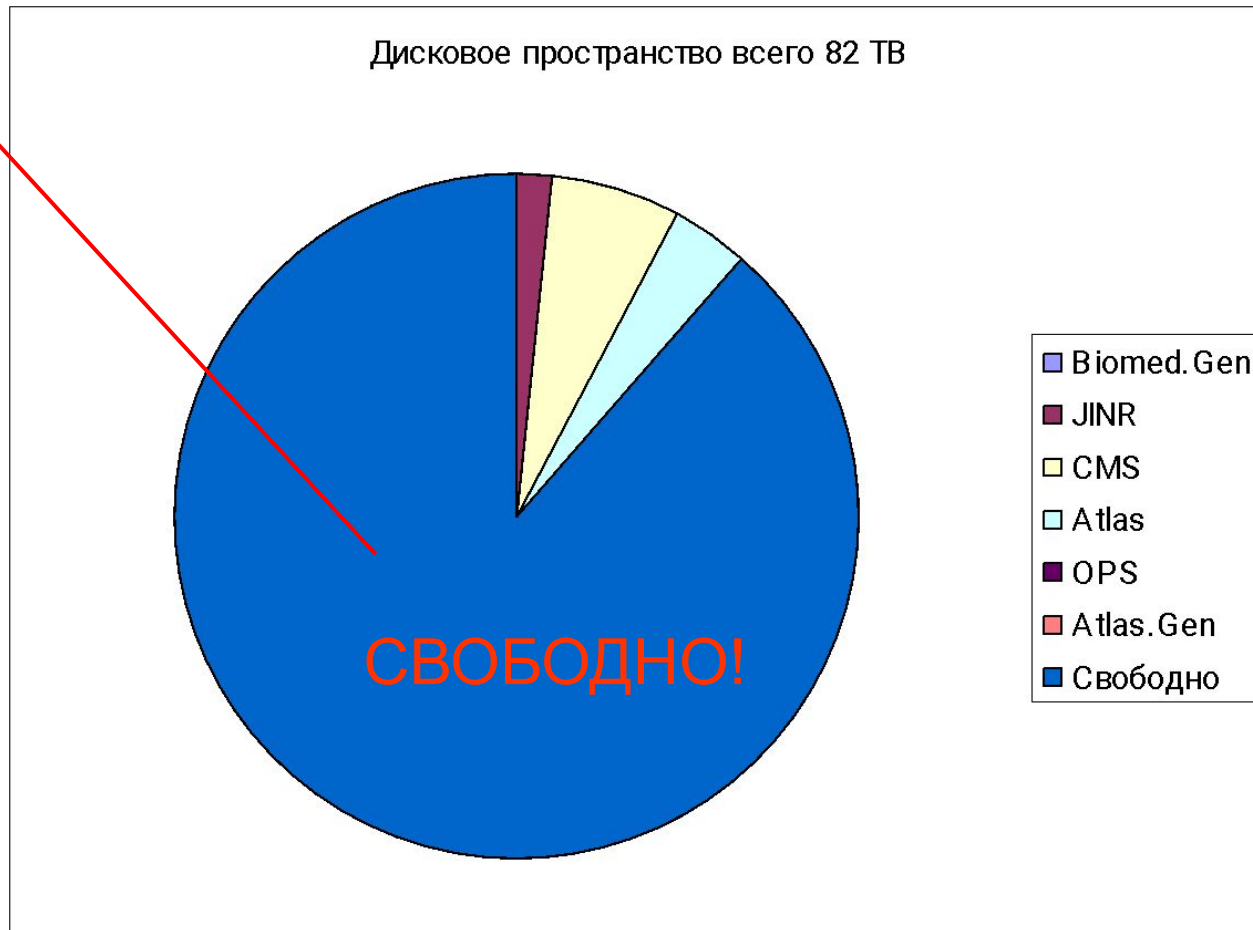
dCache о скоростях: к чему надо быть готовым, когда качаешь файлы.



dCache: кто и сколько к нам обращается

■	unknown	78658
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=mlassnig/CN=663551/CN=Mario Lassnig	52409
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=jnovak/CN=582979/CN=Judit Novak	46366
■	/C=RU/O=RDIG/OU=users/OU=sinp.msu.ru/CN=Olga Kodolova	44466
■	/C=IT/O=INFN/OU=Personal Certificate/L=Milano/CN=Davide Rebatto/E=David.Rebatto@mi.infn.it	38877
■	/C=IT/O=INFN/OU=Personal Certificate/L=Milano/CN=Guido Negri 2	30505
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=samoper/CN=582979/CN=Judit Novak	27224
■	/C=CH/O=CERN/OU=GRID/CN=Piotr Nyczyk 6217 - Atlas	25125
■	/C=CH/O=CERN/OU=GRID/CN=Judit Novak 0973 - ops	21061
■	/DC=org/DC=doegrids/OU=People/CN=Sanjay Padhi 496075	20727
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=teodoro/CN=664016/CN=Douglas Teodoro	19347
■	/C=IT/O=INFN/OU=Personal Certificate/L=Roma 1/CN=Alessandro Di Girolamo	16305
■	/C=CH/O=CERN/OU=GRID/CN=Piotr Nyczyk 6217	14243
■	/C=PL/O=GRID/O=Cyfronet/CN=Lukasz Flis - SAM	11205
■	/C=IT/O=INFN/OU=Personal Certificate/L=Napoli/CN=Michela Biglietti	10352
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=pnyczyk/CN=626004/CN=Piotr Nyczyk	9998
■	/O=GRID-FR/C=FR/O=CNRS/OU=IBCP/CN=Jaipheessa Ibecepe	9335
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=pgoglov/CN=662561/CN=Pavel Goglov	8683
■	/C=IT/O=INFN/OU=Personal Certificate/L=Milano/CN=Silvia Resconi/E=Silvia.Resconi@mi.infn.it	7494
■	/C=TW/O=AS/OU=PHYS/CN=suijian zhou/E=suijian.zhou@cern.ch	7076
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=sciaba/CN=430796/CN=Andrea Sciaba	6569
■	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=asciaba/CN=430796/CN=Andrea Sciaba	5539
■	/C=CA/O=Grid/OU=westgrid.ca/CN=Rodney Walker	5483
■	/C=CH/O=CERN/OU=GRID/CN=Judit Novak 0973	4443
■	/C=CH/O=CERN/OU=GRID/CN=Andrea Sciaba 8968	3250
■	/C=RU/O=RDIG/OU=users/OU=jinr.ru/CN=Mikalai Kutouski	3099
■	/C=DE/O=GermanGrid/OU=MPPMU/CN=Oliver Kortner	2786
■	/O=GRID-FR/C=FR/O=CNRS/OU=LAPP/CN=Stephane Jezequel	1914
■	/DC=org/DC=doegrids/OU=People/CN=Alessandro DI Mattia 321034	1237
■	/C=IL/O=IUCC/OU=Technion/CN=shikma bressler	1143
■	/C=CH/O=CERN/OU=GRID/CN=Miguel De Oliveira Branco 2423	1132

dCache: на этом месте МОГ БЫ БЫТЬ Ваш файл



dCache что надо знать для начала

- Это единственный способ получить на дисках ЦИВК место объёмом несколько ТВ.
- Краткое руководство есть на www.jinr.ru.
- Выбирайте – или работать в GRID или локально. Система ориентирована на GRID.
- 82 ТВ сейчас без квот.
- По спецификации T2 хранение ROW на T1.
- По плану 2008 г до 400 ТВ (нереально) но будет наращиваться быстрее, чем пространство AFS.
- Квоты по VO к весне будут, а то и раньше.