CBRAIN

An International Computing Platform for Neuroimaging Dr. Alan C. Evans' Laboratory

> BIC Lecture, Feb 2012 by Marc-Etienne Rousseau







Summary

- What is CBRAIN?

- What is it for?
- Who uses it and how much?
- A Few Special Projects
- What can it do for you?
- Can I get access? Can I get involved?



Funded by CANARIE CANADA'S ADVANCED RESEARCH AND INNOVATION NETWORK

http://www.canarie.ca



Canadian Brain Imaging Research Network Global Brain Imaging Research Network





















What is CBRAIN?



What is Neuro-Imaging?

St McGill



🐯 McGill

Neuro-Imaging Applications

Population Studies:

- Alzheimer's Disease
- Multiple Sclerosis
- Autism
- Schizophrenia
- Normal brain development



Alzheimer loss of cortical thickness



Multiple Sclerosis lesions



Normal Brain Development in Children

A Processing Pipeline: CIVET



"Native" T1-weighted MRI



Registered MRI



Classification



Subject brain with its cortical surface



Template brain, in stereotaxic space



Deformation field



Atlas defined on template brain, in stereotaxic space



Automatic segmentation

- Hundreds of MRI per study
- GBs of data
- Thousands of CPU hours

Removing Obstacles to Compute



Challenges: Users

Most Neuroscientists are not IT experts

- Don't want to deal with networks, HPC processing, UNIX, etc.
- Just want to use specific tools with their data and get the results.
- They want to share, but they want to control their data.
- Research requirements are unclear and unstable.





Challenges: Heterogeneous HPC





CBRAIN: An Integrative Platform





File & Project Management (inputs & results)

mcgill.ca https://portal.cbrain.mcgill.ca/	userfiles	습 호 C 🖨 🏢 🕻	} ▼ Google	۹ 🔒 💽 🦗
Home My Account User Manual Feedback Credits	Rev: 3.1.0-91		Logg	ed in as Marc-Étienne Rousseau Sign out
StormPortal - Files				
Files Tasks Projects Data Providers Serve	rs Messages Tools	Coloriad Files	Filters	View Ontions
Upload File + Launch Task + Viewers	File Management Update Attrit	outes - Selected Files	• Filters •	view Options

Active Project: Twins

Filename V	• File Type	• Owner	Creation Date	Size	• Tags Pro		
AA 2071574 mprage20050926-noflip.mnc </td <td>Minc File</td> <td>mero</td> <td>2010-12-20</td> <td>8.2 Mb</td> <td>P</td> <td>Launch lask Viewers</td> <td>-</td>	Minc File	mero	2010-12-20	8.2 Mb	P	Launch lask Viewers	-
4 2071574-GPC-119276-1	Civet Output	mero	2011-10-05	163.8 Mb (228 files)	н .	Select an Operation	
4 AA2071574-Colosse-19695-1	Civet Output	skarama	2010-12-20	163.1 Mb (228 files)	1	Select an Operation	
L AA2071574-Colosse-20096-1 🗸	Civet Output	skarama	2010-12-21	163.2 Mb (228 files)	1	Asipro -> MINC	
AA 2077226 mprage-20051204-noflip.mnc 🗸 🗸 Formats <u>JIV</u>	Minc File	mero	2010-12-20	8.5 Mb	R <mark>na</mark>	DICOM -> MINC DICOM -> HRRT	2
L 2077226-GPC-119276-1	Civet Output	mero	2011-10-05	175.4 Mb (228 files)	00	MINC -> Analyze	rst
L AA2077226-Colosse-19702-1	Civet Output	skarama	2010-12-20	174.8 Mb (228 files)	<u>,00</u>	MINC -> NIFTI	rst
L AA2077226-Colosse-20103-1	Civet Output	skarama	2010-12-21	174.8 Mb (228 files)	1 <u>01</u>	0 NIFTI -> MINC	rst
BL 2039584 mprage-20040911-noflip.mnc 🗸 🗸 Formats <u>JIV</u>	Minc File	mero	2010-12-20	8.0 Mb	_P 01	Launch Bedpostx Launch Civet QC on CivetStudy	rst
L 2039584-GPC-119276-1	Civet Output	mero	2011-10-05	164.0 Mb (228 files)	,01	Launch Civet	1 St
4 ABL2039584-Colosse-19707-1	Civet Output	skarama	2010-12-20	163.4 Mb (228 files)	, <u>01</u>	4 Launch CivetCombiner	rst
4. ABL2039584-Colosse-20108-1	Civet Output	skarama	2010-12-21	163.4 Mb (228 files)	I S'	Launch Cw5filter Run Cluster Diagnostics	
						Launch Mincaverage Launch Mincmath	Ă

÷

Launch Mincpik

Launch Tools

Task Pa	rameters (Help):					Tool o	configuration					
Pipeline O Template: Model: ie	1.00 + bm152nl +											
CIVET opt Interpolat N3 distand Degrees o Do NOT bu Thickness Resample	tions tion trilinear te: 0 f freedom for linear registration: uild surfaces: method: tlink surfaces: surfaces:	9 🛟 size: 20										
Combine s	surfaces:	(Comput	e Job s	tatus and Manag	gement
Process VI Blurring k Run symm Keep cere	BM files: ernel size in mm for volume: netry tools: bellum in VBM maps:	Ho S	me My Account (tormPorta iles Tasks Pro	User Manu <mark> - Ta</mark> s ojects E	ual Feedback Credit SKS Data Providers Ser	s Rev: 3.1. vers Mes	0-91 sages Tools	10			Logged in as Marc-Étienne Rou	isseau Sign out
-MINC file	input list		For Queued T View Options	'asks 🔻	For Running Tasks	For	Failed Tasks For	Completed Tasks -	Terminating An	d Cleaning	Up - Archiving -	Filters
☑ Launch CIVET?	T1 name	A	ctive Project: Tw	vins View							(1401 tasks)	
Ø	AA_2071574_mprage20050926- noflip.mnc	pret	Civet (Go to	Type Batch)	Description N3_150mm_Spline N3=150	• <u>Owner</u> skarama skarama	Execution Server 200 jobs launched at 201 Colosse	Current Status Corrent Status Completed	Run Number W 200 x Failed 1	'orkdir Size ⊽	Time Submitted	Last Upd
			Civet (Go to Minc2jiv Minc2jiv (Go	Batch) Batch) Batch) Batch) to Batch)	N3=150 N3=150 N3=200 N3=200 N3=200	skarama skarama skarama skarama skarama skarama	200 jobs launched at 201 200 jobs launched at 201 <u>Colosse</u> 200 jobs launched at 201 199 jobs launched at 201 <u>MindStorm</u> 200 jobs launched at 201	10-12-21 11:35:41 EST 10-12-21 10:29:13 EST Completed 10-12-21 09:43:52 EST 10-12-20 16:12:07 EST Completed 10-12-20 15:54:03 EST	200 x Completed 200 x Completed 1 200 x Completed 199 x Completed 1 200 x Completed	⊽ ⊽	2010-12-21 09:45:53 EST 2010-12-20 15:55:53 EST	2012-01-12 13 2012-01-11 22
			Minc2jiv (Go	to Batch)		skarama	199 jobs launched at 201 Workdir archiving state	10-12-20 15:53:20 EST us symbols:	199 x Completed luster ⊽: As File			

Results, QC, Provenance and Visualisation



CBRAIN Transparent Compute Access



Transparent Data Grid



Interactive Scientific Visualisation



Tools & Compute

Community Portal





Generic Research Platform





CBRAIN International Platform



Current Worldwide Distribution CBRAIN Users



CBRAIN Platform Statistics

Membership



CBRAIN Platform Statistics



CBRAIN Platform Statistics

Workload

(Official Compute Canada Resources Only)



Core*hours

HPC Integration 2012 (11 compute installations, 80,000+ core)





Infrastructures Continental Access for Communities





OUTGRID – GBRAIN (EGI Tech Forums 2011, Lyon: Best live demonstration)



BrainBrowser – Web3D Planes, brains and automobiles

MACCAC Maps





- Drs. Zilles & Amunts -**Post-Mortem High-Resolution Histology**



Julich, Germany. Dr. Karl Ziles

- **Rip & Tear correction**
- ~5000 CPU hours

Non-linear slice alignment, 50 iterations ~370,000 CPU hours

Conversion: 1TB 3D data structure HPC ~160 CPU hours

- Full brain view
- ~10 ∝μ ρεσ



Goal: Collaborative, Real-Time 3D Visualization

- 1TB of data per brain
- 7408 Slices
- 11500x8800 pixels
- Animal studies will generate a few dozens per year

Big Brains – Atelier3D Histology – Neuroimaging – HPC – Modeling/Visualisation





Large Animal Datasets (Dr. Bedell)

CBRAIN BrainBrowser









Atelier 3D

۰

What can CBRAIN do for you?



Typical Usage

- Access to pre-installed tools, converters, pipelines.
- Access to large amounts of compute power (project speedup).
- Long distance collaborations / data sharing.
- Web based visualisations.
- Do things you could not easily do before!



Illustrative Performance Comparison

NIH-Pediatric-Obj1: up to 3 visits per subject

866 CIVET pipeline runs to generate cortical thickness maps Input: 866 x 3 x 5Mb = 15Gb Output: 866 x 250 Mb = 211Gb

Cluster	Total CPU-hrs	Maximum F	Performance	Typical Performance		
		# cores	Execution time	# cores	Execution time	
mammouth-ms2 (RQCHP -Sherbrooke)	866 x 4 = 3464	~500	3hr	176	17hr	
CLUMEQ-Krylov (McGill)	866 x 6 = 5196	~90	2.5d	24	9d	
BIC (MNI)	866 x 8 = 6928	~100	3d	40	7d	

In general, studies which use to takes 1 week to 1 month now take 1 day.

🐯 McGill

Available Tools

Based on Community requests! Come talk to us, we will see what we can do.

CIVET

NIAK (Dr. Bellec)

FSL (some tools)

Minc tools (some tools)

FreeSurfer recon-all

SPM Batch

Format conversion (minc, nifty, dicom, analyze...)

Cw5

Can the tools be complex?

Dynamic Flowchart for fMRI pre-processing (NIAK) – 2 subjects



A Few CBRAIN Related Publications



NeuroImage

Volume 55, Issue 2, 15 March 2011, Pages 462-467



Patterns of cortical thickness and surface area in early

Parkinson's disease

Thomas Jubaulta, b, Jean-Francois Gau C. Evansd, Oury Monchia, b, a. W

^a Unité de Neuroimagerie Fonctionelle, Institut U ^b Département de Radiologie, Université de Mor ^c Département de Psychologie, Université du Qu d McConnell Brain Imaging Centre, McGill Unive ^e Montreal Neurological Institute, McGill University ^f Movement Disorders Unit, McGill University He

Received 20 August 2010. Revised 30 November

http://dx.doi.org/10.1016/j.neuroimage.2010.12.0 Permissions & Reprints

Integration of a neuroimaging processing pipeline

into a pan-canadian computing grid

- S. Lavoie-Courchesne^{1,2,3}, P. Rioux³, F. Cl T. Sherif³, M.-E. Rousseau³, S. Das³, R. A C. Craddock^{4,8}, D. Margulies^{5,8}, C. Chu^{6,8}
- A. C. Evans³, P. Bellec^{1,2,8}
- ¹Centre de recherche de l'institut de gériatrie de Montréa ²Département d'informatique et de recherche opérationne Montréal, CA
- ³McConnell Brain Imaging Centre, Montreal Neurologica Montréal, CA
- ⁴Virginia Tech Carilion Research Institute, Roanoke, VA. ⁵Max Planck Institute for Human Cognitive and Brain S ⁶Section of Functional Imaging Methods, Laboratory of 1 Bethesda, USA
- ⁷Mind, Brain Imaging and Neuroethics Unit, Institute of Ottawa Health Care Group, University of Ottawa, Ottaw ⁸The Neurobureau Research Institute, Chicago, USA

E-mail: {bellecp@iro.umontreal.ca, lavoiecs@iro.um

Abstract. The ethos of the neuroimaging field is quickl resources, including both imaging databases and processi represents a large volume of datasets and as neuroimaging heterogeneous, computationally intensive tools, such oper challenges. This motivates the design of novel dedica describe here an interface between PSOM, a code-oriented CBRAIN, a web-oriented platform for grid computing. a PSOM-compliant pipeline for preprocessing of structur imaging into CBRAIN. We also showed the capacity of thi database by processing close to 1000 subjects released as This pilot experiment demonstrates the feasibility of our computing solution for high-throughput processing in the

1. Introduction

With most modern scientific projects relying heavily on compute research software prototypes is of central importance at the very of published results by independent research groups []. In dissemination raises a number of domain-specific computationa

(i) Complex multi-stage processing. Neuroimaging data number of preprocessing steps before a scientifically mean These steps can include such operations as non-linear real

[PCS 2011 – HPCS 2011

FOCUS ON NEUROINFORMATICS

Virtual imaging laboratories for marker discovery in neurodegenerative diseases

Giovanni B. Frisoni, Alberto Redolfi, David Manset, Marc-Étienne Rousseau, Arthur Toga and Alan C. Evans

Psychiatry Abstract | The unprecedented growth, availability and accessibility of imaging data from people with neurodegenerative conditions has led to the development of computational infrastructures, which offer scientists access to large image databases and e-Science services such as sophisticated image analys algorithm pipelines and powerful computational resources, as well as three-dimensional visualization ar statistical tools. Scientific e-infrastructures have been and are being developed in Europe and North Am that offer a suite of services for computational neuroscientists. The convergence of these initiatives represents a worldwide infrastructure that will constitute a global virtual imaging laboratory. This will pro computational neuroscientists with a virtual space that is accessible through an ordinary web browser, u image data sets and related clinical variables, algorithm pipelines, computational resources, and statist image data sets and related united relation of the set of the set of their physical location. S Volume 70, Number 3 an experimental environment will be instrumental to the success of ambitious scientific initiatives with I

societal impact, such as the prevention of Alzheimer disease. In this article, we provide an overview of t August 1, 2011 currently available e-infrastructures and consider how computational neuroscience in neurodegenerat disease might evolve in the future.

Frisoni, G. B. et al. Nat. Rev. Neurol. 7, 429-438 (2011); published online 5 July 2011; doi:10.1038/nmeurol.2011.99

Introduction

Research in neurodegenerative diseases is undergoing a radical transformation brought about by extraordinary owth in the volume, availability and accessibility of clinical and research imaging data, both in the form of public releases and within virtual research organizations. Traditional neuroimaging research typically involved small to mid-sized locally collected data sets ranging from dozens to hundreds of scans. Only a few imaging laboratories have the technical expertise and computa ional resources required to merge multiple large data sets and explore scientific questions relating to larger populations. Not only do neuroscientists face a steep learning curve to grasp their own particular computing ecosystem, in terms of operating ystem environment, basic scripting, programming, remot data transfers and remote computing, but also, because of divergence in the basic information technology (IT) setup, the principles of one ecosystem often do not adapt wel to other laboratories. The commonplace replication and idiosyncrasies of toolsets and infrastructures among many sites greatly increases the complexity and overheads for neuroimaging projects, leading to issues such as the need to locally support IT-related technical staff, and difficulties in coordinating multisite studies.

Open access to large data sets, pioneered in genetics and physical sciences, has been implemented successfully by ious initiatives in the neuroimaging field, such as the

Competing interests The authors declare no competing interests

NATURE REVIEWS NEUROLOGY

Alzheimer's Disease Neuroimaging Initiative (A and the NIH Pediatric Database (NIHPD).² Sino all researchers who subscribe to these databases have able to obtain full access to images and clinical dat people with varying degrees of cognitive deteriorati were originally collected to identify biomarkers of initiation and progression.34 Currently, a number to very large data sets can be found in the public d and freely downloaded, such as the 1000 Fun Connectomes Project,5 the Human Imaging Da (HID),6 the Open Access Series of Imaging S (OASIS),⁷ the Bipolar Disorder Neuroimaging Da (BiND),⁸ Multisite Imaging Research In the Ana Depression (MIRIAD),⁴ and Efficient Longitudinal of Depression in the Elderly (ELUDE).⁴⁰

The gap between the pace of data generation a capability to extract clinically or scientifically re information is rapidly widening. Sophisticated algo are available, and more are being developed, that all extraction of biologically relevant markers from and clinical data requiring heavy computation instance, the extraction of the three-dimensional of thickness map, a marker of neurodegeneration, high-resolution structural MRI scan can take be

30 min and 22 h per scan on a single-core comput extraction of functional connectivity networks ci Genotype, Circuits, and Cognition in Autism and 20-120 min. At present, relatively few imaging labor worldwide have the expertise and resources reAttention-Deficit/Hyperactivity Disorder for such sophisticated high-throughput comput



Biological

© 2011 Macmilian Publishers Limited. All rights reserved

Nature reviews. Neurology, 2011 Jul 07; 7(8):429-38

ISSN 0006-3223 www.sobp.org/journal Official Journal of the Society of Biological Psychiatry

Selected Conferences and Press

SIGGRAPH 2011, Vancouver, Canada, Bird of a Feather: WebGL session, Invited Speaker, BrainBrowser.

EGI User Forum 2011, Vilnius, Lithuania, Invited Speaker, GBRAIN Project (a CANARIE NEP Project).

MICCAI 2011 Workshop High Performance and Distributed Computing for Medical Imaging, Toronto, Canada, Invited Speaker, The CBRAIN Neuroimaging Platform.

Canadian Research Data Summit 2011, Invited Speaker. Dialogue on developing a common vision. Ottawa, Canada, The CBRAIN Neuroimaging Platform.

HPCS 2011, Montreal, Canada, Invited Speaker, CBRAIN: Canadian Neuroinformatics Platform.

Best Poster at HPCS 2011, Montreal, Canada. Sébastien Lavoie-Courchesne, P. Rioux, T. Sherif. S. Das, N. Kassis, J. Doyon, R. Adalat, M.E. Rousseau, A.C. Evans, P. Bellec, Integration of a neuroimaging pipeline prototype into a pan-canadian computing grid.

Best Live Demonstration at EGI Technical Forums 2011, Lyon, France.

http://gridtalk-project.blogspot.com/2011/09/win-win-win.html and article http://www.isgtw.org/visualization/hat-trick-alzheimer %E2%80%99s-grand-challenge

Rapport Annuel RISQ 2009-2010: Le Project CBRAIN

Colloque RISQ 2009: Highest Satisfaction level (from public vote)

CBRAIN technologies produced a cover page for Ducharme & al in **Biological Psychiatry** (see Publication: Biol Psychiatry. 2011 Aug 1;70(3):283-90. Epub 2011 Apr 30.)

Gigabits on the Brain. Ottawa Citizen, Business Frontpage. <u>http://www.ottawacitizen.com/business/Gigabits+brain/3691533/story.html</u>

CANARIE opens the "ultra-fast lane". Telemanagement.

http://www.tele-management.ca/content/23481-canarie opens the %E2%80%9Cultra fast lane%E2%80%9D

OGF-28 Sessions 2010, Munich GER, Invited Speaker. CBRAIN/GBRAIN Projects.

Selected 2011 Publications

Frisoni G. B., Redolfi A., Manset D., Rousseau M. E., Toga A., Evans A. C. Virtual imaging laboratories for marker discovery in neurodegenerative diseases. Nature reviews. Neurology, 2011 Jul 07; 7(8):429-38

Ducharme S, Hudziak JJ, Botteron KN, Ganjavi H, Lepage C, Collins DL, Albaugh MD, Evans AC, Karama S; Brain Development Cooperative Group. **Right anterior cingulate cortical thickness and bilateral striatal volume correlate with child behavior checklist aggressive behavior scores in healthy children.** Biol Psychiatry. 2011 Aug 1;70(3):283-90. Epub 2011 Apr 30.

Lavoie-Courchesne, S. Rioux P., Chouinard-Decorte P., Sherif T., Rousseau M.-E., Das S., Adalat R., Doyon J., Craddock C., Margulies D., Chu C., Lyttelton C., Evans A.C., Bellec P. Integration of a neuroimaging processing pipeline into a pancanadian computing grid. Journal of Physics: Conference Series HPCS 2011, accepted.

Gong G, He Y, Chen ZJ, Evans AC. Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. Neuroimage. 2011 Aug 22. [Epub ahead of print]

Gong G, He Y, Evans AC. Brain connectivity: gender makes a difference. Neuroscientist. 2011 Oct;17(5):575-91. Epub 2011 Apr 28.

Ganjavi H, Lewis JD, Bellec P, MacDonald PA, Waber DP, Evans AC, Karama S; Brain Development Cooperative Group. **Negative associations between corpus callosum midsagittal area and IQ in a representative sample of healthy children and adolescents.** PLoS One. 2011;6(5):e19698. Epub 2011 May 19.

Operation and Support

- Compute Canada Support (3.7M hours for 2012)
- Creation of a Center for Neuroinformatics
- CFI Application
 - 5 PB Datacenter, part of Compute Canada environment
 - Obtained Server Room support for 5 years from McGill
 - Obtained Compute Canada and CLUMEQ-ETS support
- Partner Projects
 - Small contracts to support specific CBRAIN projects (Julich, outGRID...)
- CANARIE & Compute Canada next round (in 2012
 - National Research Platform Initiative.

Can I get access? Can I get involved?



cbrain.mcgill.ca



Contact Our Team

CBRAIN contacts:

Technical Manager:

Marc Rousseau – <u>marc.rousseau@mcgill.ca</u>

Project Manager:

Reza Adalat – reza.adalat@mcgill.ca

CBRAIN Architecture



CBRAIN Project Team & Partners

Alan Evans				
Reza Adalat	Participating Research Centers			
Pierre Rioux	Montreal Neurological Institute, McGill University (Lead) Principal Investigator: Alan Evans			
Tarek Sherif	Program Manager: Reza Adalat System Architect: Marc Rousseau Developers: Pierre Rioux, Tarek Sherif, Angela			
Nicolas Kassis	McCloskey, Samir Das, David Brownlee McGill Office of Technology Transfer (OTT): Francoys Labonte			
Natacha Beck	Canada National Research Council: Louis Borgeat Consultants: Rosanne Aleong, Claude Lepage, Pierre Bellec, Andrew Janki, Robert Vincent			
Tien Duc Nguyen	Rotman Research Institute, University of Toronto Principal Investigators: Stephen Strother and Randy MacIntosh Developers: Anda Pacurar, Anita Oder, Jacques			
Claude Lepage	Robarts Research Institute, University of Western			
Louis Borgeat (NRC)	Ontario Principal Investigators: Ravi Menon and Mel Goodale			
Lindsay Lewis	Developers: Martyn Klassen, Ronghai Tu			
Hartmut Molhberg, Timo Dickscheid (Drs. Zilles – Amunts)	Montréal Principal Investigators: Julien Doyon and Rick Hog Developer: Mathieu Desrosiers			
	Division of Neurology, University of British Colombia Principal Investigators: Jon Stoessl and Max			
All ACE Lab members	Cynader Developers: Ryan Thomson, Nasim Vafai			
Mathieu Desrosiers (UdeM)	CBRAIN acknowledges the contributions of IBM Canad Jonathan Harlap and Sebastian Muehlboeck.			
Pierre Bellec (UdeM)				

Sébastien Courchenes-Lavoie (UdeM)

canarie

Contact Our Team

CBRAIN contacts:

Technical Manager:

Marc Rousseau – <u>marc.rousseau@mcgill.ca</u>

Project Manager:

Reza Adalat – reza.adalat@mcgill.ca

CBRAIN Technical Overview





HPC Scheduling (SCIR: Simple Cluster Interface in Ruby)



Flexible Resource Access



Flexible Data Access



Julich Forschungszentrum, Germany

Large dataset assembly pipeline (current and future Big Brains) Remote visualization of Big Brains (A3D) MRI study



Operations in 2012

