



Introducing GATECloud.net

Valentin Tablan, Ian Roberts
University of Sheffield



Text Solutions on the Cloud

What?

We provide end-to-end text and web processing solutions from [the GATE family](#) running on **cloud computing** infrastructures.

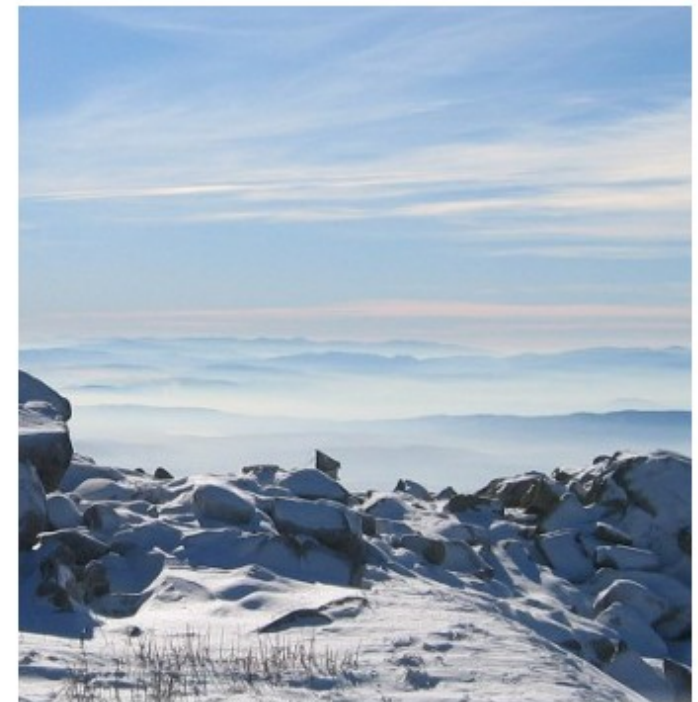
Why?

- Solve any sort of text processing problem: web, text or opinion mining; indexing and search (fulltext, boolean, conceptual, structural); information extraction; semantic annotation; sentiment analysis; ontology population; etc. ([some examples](#)).
- Run large-scale jobs without investing in server hardware or other fixed costs.
- Exploit a **15-year R&D programme**, the expertise of the [GATE community](#) and a **defined and repeatable process**.

How?

- Click through to [the shop](#), browse products and add them to your shopping basket.
- [Create an account](#) and then buy credit vouchers from [the University of Sheffield](#) (use the same email address in each case).
- Put the vouchers in [your account](#), and [go to checkout](#). We'll email you the login or job creation details for your cloud servers.
- Monitor and control your cloud machines on [your dashboard](#).

For more detail see the [help page](#).



Special offers

- **no VAT**, Q1/Q2 2011 (we pay the VAT!)

Discounts

- volume, research or non-profit use
- to apply create an account and [send us](#) your details

<https://gatecloud.net>



- A new service from the GATE team
- Usual cloud benefits:
 - Pay-as-you-go, no upfront costs
 - No sysadmin work
 - Web-based management tools
- Not-so-usual cloud benefits
 - Based on open-source software
 - Bring your own pipeline
 - No API required (all interaction is browser based)



Current Offering

- On demand servers
 - Teamware
 - Mimir
 - Always latest version, maintained by us
 - Backups
- On demand document processing
 - Massive parallelism
 - On-line job definition tool
 - Many output formats, including Mimir



GATEcloud Products

<u>Annotation Job - Custom</u> <i>Execute your own pipeline on the GATE cloud.</i>	£0.00 (plus £0.99 per hour)
<u>ANNIE (Named entity annotation service)</u> <i>Upload your own documents and run our pre-packaged named entity annotator (ANNIE) on GATE Cloud</i>	£0.00 (plus £1.49 per hour)
<u>ANNIE plus Measurements and Numbers (annotation service)</u> <i>Upload your documents and run our pre-packaged named entity annotator with the Numbers and Measurements add-ons</i>	£0.00 (plus £1.75 per hour)
<u>GATE Teamware 1.4 Server (Enterprise)</u> <i>Web-based collaborative corpus and document annotation tool</i>	£349.99 £99.99 (plus £1.49 per hour)
<u>GATE Mimir 3.2.0 Server</u> <i>Multiparadigm indexing server</i>	£99.99 £0.00 (plus £0.99 per hour)

Special offers

- **no VAT**, Q1/Q2 2011 (we pay the VAT!)

Discounts

- volume, research or non-profit use
- to apply create an account and [send us](#) your details





End-to-End Solution

1. Bring your own documents
2. Bring your own pipeline (or use one of ours)
3. Automatically annotate the documents on a massively parallel infrastructure
 - Talk to us if you have special requirements
4. Use Teamware for pipeline fine-tuning and quality assurance (see Track 1, Module 4, on Thursday)
5. Use Mimir to power multi-paradigm search (see demo, plenary session on Thursday!)

Servers (Machines)



Servers Actions

□ Reserve

- A new data volume is created and reserved for you.
- A new machine appears on your dashboard.
- You pay the one-off price.

□ Start Instance

- The reserved server is started and you can use like any other server.
- You pay the hourly price for as long as the sever is running.

Servers Actions (continued)



□ Stop Instance

- The started server is stopped.
- You stop paying hourly charges.

□ Create Backup

- A copy is made of your data volume. You can only do that while the server is stopped.
- You can restore this backup back onto your data volume whenever you want.

□ Destroy Reservation

- Your data volume and all the backups are destroyed.

Making a Reservation:

1) Find Required Server



The screenshot shows a web browser window with the URL <https://gatecloud.net/shopfront/displayItem/9>. The page header includes the GATE CLOUD logo with a "BETA" tag, navigation links for Home, Shop, and Dashboard, and a shopping basket containing 1 item. The main content area displays the product "GATE Teamware 1.4 Server (Enterprise)" with a description: "Web-based collaborative corpus and document annotation tool". The price is shown as £349.99 (crossed out) and £99.99 (current price), plus £1.49 per hour. A quantity selector is set to 1. Below the product name is a sub-header: "GATE Teamware (Enterprise): Web-based Collaborative Corpus Annotation and Curation for Large Teams". A list of features includes: 20 GB data partition, Five backups for the data partition, and 64 bit installation. An embedded image shows a preview of the GATE Teamware web interface with a "Welcome" message and a list of links: Open Annotation Editor, Edit Profile, Projects, My Tasks [1], and Group Tasks [0]. At the bottom of the product page is an "Add to basket" button. The footer contains copyright information for The University of Sheffield (1995-2011) and a license notice.

2) Add to Basket



Shopping basket

https://gatecloud.net/shopfront/seeBasket

BETA
GATE CLOUD

[Valentin Tablan's account](#) | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#) [Basket \(1 Items\)](#) [Checkout](#)

Shopping basket

Item	Price	Quantity	Total
GATE Teamware 1.4 Server (Enterprise)	£99.99, plus £1.49 per hour.	<input type="text" value="1"/>	£99.99, plus £1.49 per hour.
			Update Quantities Checkout
Grand Total:			£99.99, plus £1.49 per hour.

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
 All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
 Cloud computing server time is payable in advance.

3) Checkout



Your order contents

https://gatecloud.net/shopfront/checkout?execution=e1s1

BETA
GATE CLOUD

Valentin Tablan's account | Log out | Help

Home Shop Dashboard Basket (1 Items) Checkout

Your order contents

Item	Price	Quantity	Total
GATE Teamware 1.4 Server (Enterprise)	£99.99 (plus £1.49 per hour)	1	£99.99 (plus £1.49 per hour)
Total to pay now:			£99.99

You are about to be charged **£99.99**.
 Your current account balance is **£820.83**.
 After completing this purchase, your remaining balance will be **£720.84**.
 If you have ordered any cloud machine instances or annotation jobs, you will also incur the following additional charges:

- **for each running hour, for each running machine**, you will be charged the hourly rate listed above

Please check the details above. If everything is in order, please press:

You can also and return to your basket.

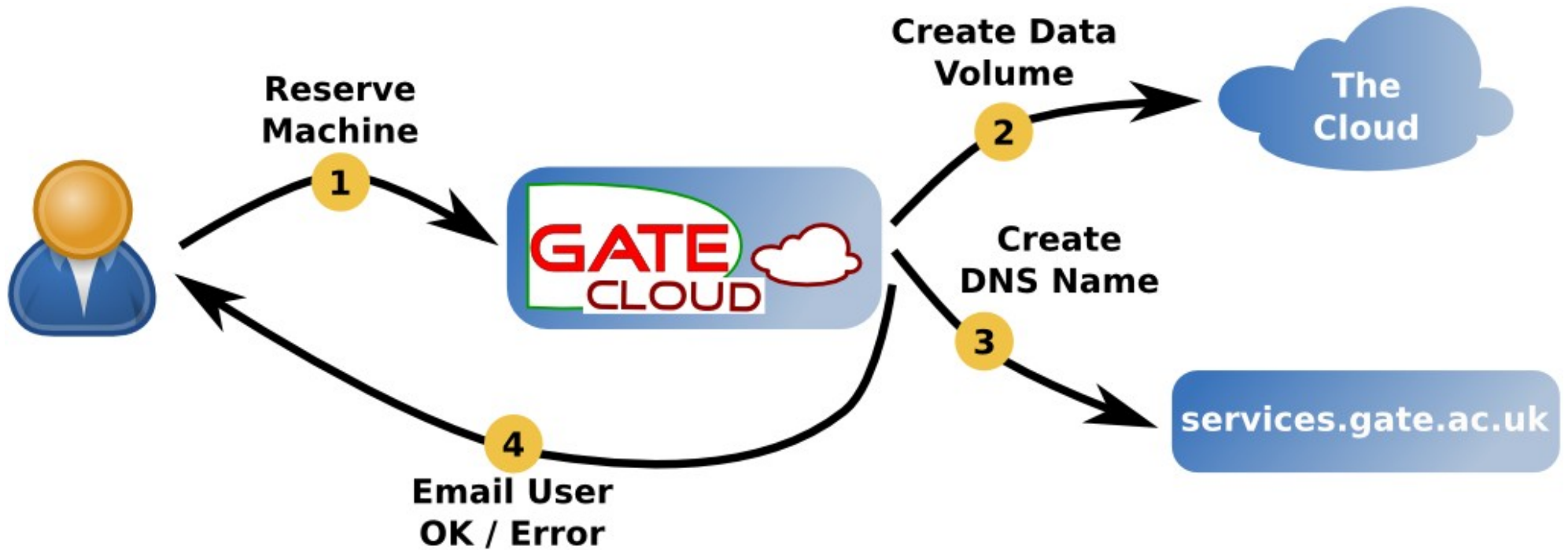
[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
 All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
 Cloud computing server time is payable in advance.



Reserving a Server



Your Dashboard



Browser window: Your Dashboard | <https://gatecloud.net/yourAccount/dashboard?execution=e1s1>

BETA
GATE CLOUD

[Valentin Tablan's account](#) | [Log out](#) | [Help](#)

[Home](#) | [Shop](#) | [Dashboard](#) | [Your account](#)

Valentin Tablan: Your Dashboard

Cloud Machines

Recently purchased machines may take a few minutes to appear in this list.

[Filter view...](#)

Reservation ID	Machine Type	State	Order Number	Order Date
R-000065	TW 1.4 Enterprise	inactive	O-000099	12 May 2011

Annotation Jobs

[Filter view...](#)

Name	Created At	State

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
 All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
 Cloud computing server time is payable in advance.

Your Machine Reservation



Machine Reservation ... x

https://gatecloud.net/yourAccount/machineReservationDetails/65

Home Shop Dashboard Your account


Machine Reservation R-000065

ID	R-000065	Destroy Reservation
Machine type	GATE Teamware 1.4 Server (Enterprise)	
Hourly price	£1.49	
State	inactive	Start Instance
Instance ready	no	

Backups

Slot 1	<empty>	Create new backup
Slot 2	<empty>	Create new backup
Slot 3	<empty>	Create new backup
Slot 4	<empty>	Create new backup
Slot 5	<empty>	Create new backup

Reservation Details:

URL1:	https://twe-havaaw1sjs.services.gate.ac.uk/teamware/executive/
Admin account name:	admin
Admin account password2:	

⁽¹⁾ This link will only work when your reservation is active and the associated instance is ready.

⁽²⁾ This password is the value that was generated and used when the reservation was originally made. If you have changed the password for the admin account in the mean time, then please use the new value that you set yourself. We do not have access to the new password, so we cannot display it here. If you have lost the password, please [contact us](#) and we can reset it for you.

Starting a Server:

1) Press Start



Machine Reservation ... x

https://gatecloud.net/yourAccount/machineReservationDetails/65?Start+Instance=Start+Instance

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#) [Your account](#)

Machine Reservation R-000065

i We are working on starting your instance. This will take a few minutes, please be patient!
You should receive an email message when the new instance is ready for use. You can also check back on this page later.

ID	R-000065
Machine type	GATE Teamware 1.4 Server (Enterprise)
Hourly price	£1.49
State	pending
Instance ready	no

Backups

Slot 1	<empty>	Create new backup
Slot 2	<empty>	Create new backup
Slot 3	<empty>	Create new backup
Slot 4	<empty>	Create new backup
Slot 5	<empty>	Create new backup

Reservation Details:

URL1:	https://twe-havaaw1sjs.services.gate.ac.uk/teamware/executive/
Admin account name:	admin
Admin account password2:	eg6tr1t935

2) Wait for the Email (or refresh the page)



Machine Reservation ...

https://gatecloud.net/yourAccount/machineReservationDetails/65?Start+Instance=Start+Instance

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#) [Your account](#)

Machine Reservation R-000065

ID	R-000065
Machine type	GATE Teamware 1.4 Server (Enterprise)
Hourly price	£1.49
State	active <input type="button" value="Stop Instance"/>
Instance ready	yes

Backups

Slot 1	<empty>	<input type="button" value="Create new backup"/>
Slot 2	<empty>	<input type="button" value="Create new backup"/>
Slot 3	<empty>	<input type="button" value="Create new backup"/>
Slot 4	<empty>	<input type="button" value="Create new backup"/>
Slot 5	<empty>	<input type="button" value="Create new backup"/>

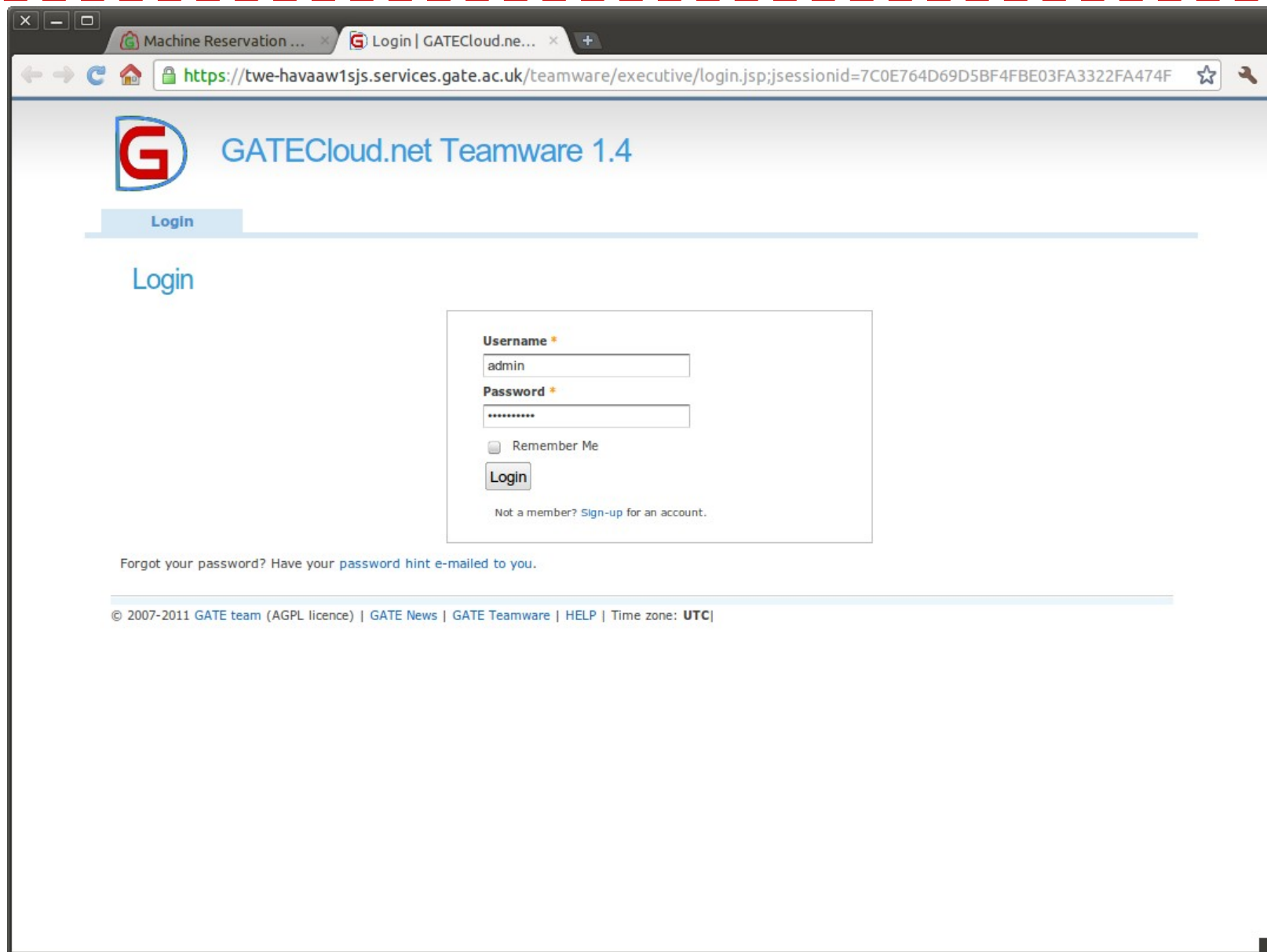
Reservation Details:

URL1:	https://twe-havaaw1sjs.services.gate.ac.uk/teamware/executive/
Admin account name:	admin
Admin account password2:	<input type="password"/>

(¹) This link will only work when your reservation is active and the associated instance is ready.

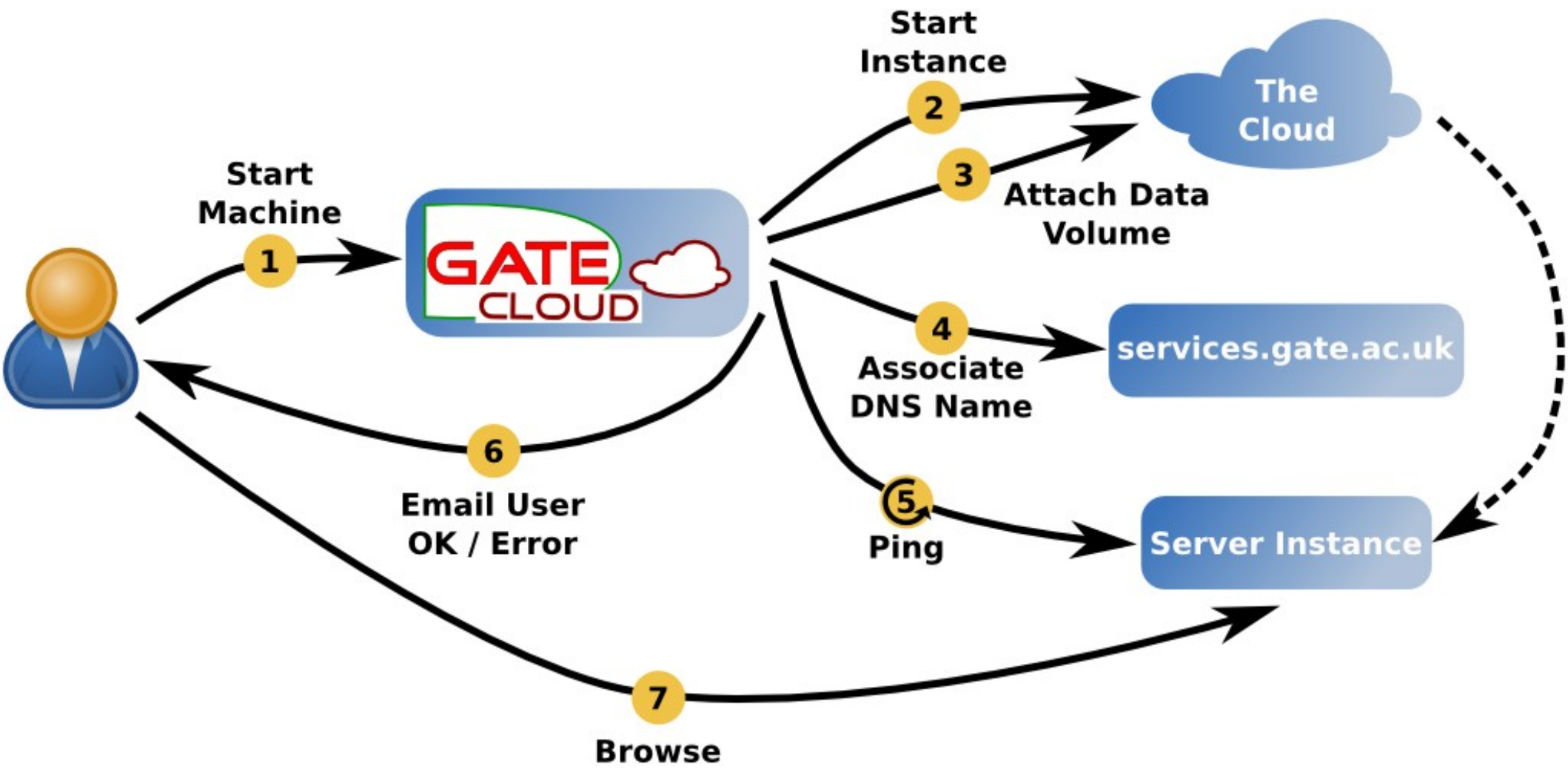
(²) This password is the value that was generated and used when the reservation was originally made. If you have changed the password for the admin account in the

3) There is no “3”: Your Server is Ready





Starting a Server



Annotation Jobs

Annotation Jobs



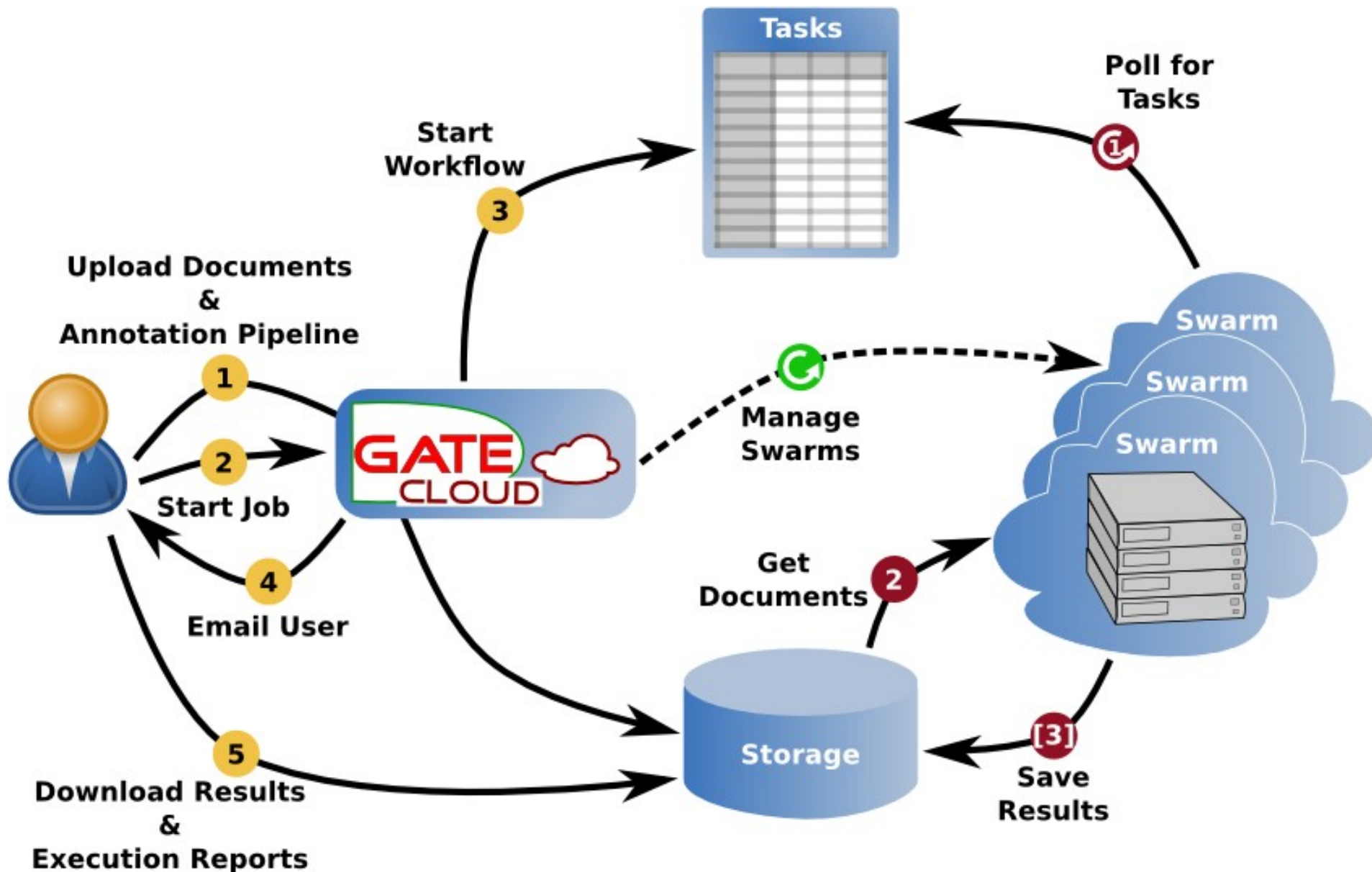
- Parallel and distributed annotation of documents using GATE applications.
- Upload your own document archives (.zip, .tar.gz, .tar.bz2, .arc).
- Upload your own GATE pipeline,
- ...or use a pre-packaged one.
- Download the outputs as .zip files,
- ...or send the annotated documents directly to Mimir for indexing.

Annotation Jobs: Execution Environment



- Ubuntu 10.10 Linux
- 64 bit
- SUN (Oracle) Java 1.6.0_21
- 2 GB RAM / thread (on average)
- GATE Embedded version 6.1
- GATE Cloud Paralleliser version 2.0
- Secure sandbox

Running an Annotation Job



Pick, Checkout, Buy



ANNIE plus Measur... x

https://gatecloud.net/shopfront/displayItem/8

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#) [Basket \(empty\)](#) [Checkout](#)

ANNIE plus Measurements and Numbers (annotation service)

Upload your documents and run our pre-packaged named entity annotator with the Numbers and Measurements add-ons £0.00
(plus £1.75 per hour)

Run ANNIE named entity annotation with the Numbers and Measurements add-on, on the GATE Cloud

Annotate your documents on our servers using **ANNIE** plus the **Numbers** and **Measurements** add-ons.

Features:

- **2GB RAM** per execution thread (on average)
- use the mature ANNIE named entity recogniser
 - producing the typical ANNIE annotations, such as *Person, Location, Organization, Date, Money, Address*, etc.
- also includes the **Numbers** add-on
 - produces Number annotations with normalised values
- also includes the **Measurements** add-on
 - produces Measurement annotations for scalar and interval measurements, covering **hundreds of measurement units**. Each measurement is annotated with a **normalised value and unit**.
- bring **your own documents** as .zip or .arc archives
- process any document format that GATE supports, including .txt, .html, .xml, .pdf.
- mix and match output formats, including GATE XML, inline XML, XCES
- send annotated documents directly to your Mimir server for indexing

Usage is charged by the hour.

[Add to basket](#)

Your Dashboard



Browser address bar: <https://gatecloud.net/yourAccount/dashboard?execution=e1s1>

BETA
GATE CLOUD

[Valentin Tablan's account](#) | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#) [Your account](#)

Valentin Tablan: Your Dashboard

Cloud Machines

Recently purchased machines may take a few minutes to appear in this list.

[Filter view...](#)

Reservation ID	Machine Type	State	Order Number	Order Date
R-000065	TW 1.4 Enterprise	inactive	O-000099	12 May 2011

Annotation Jobs

[Filter view...](#)

Name	Created At	State
Annotation Job - ANNIE + Measurements	13 May 2011 11:42:46 BST	RESERVED

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
 All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
 Cloud computing server time is payable in advance.

Job Editor: Details



Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#)

Annotation Job "Annotation Job - ANNIE + Measurements"

Details | **Inputs** | **Outputs** | **Execution Progress** | **Download Results** | **Status**

Use this tab to specify details about your job.
The application for this job is predefined and cannot be changed.

Job Name: **Annotation Job - ANNIE + Measurements**

Application File: **ANNIE-Plus-Measurements-6.1.zip**

Job State: Job reserved but not yet defined.

ID: J-000103

UUID: 10c89416-7de6-4d73-bac5-ebfc0e1fde38

Date Created: 13 May 2011 11:42:46 BST

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
Cloud computing server time is payable in advance.

Job Editor: Inputs



Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA
GATE CLOUD

Valentin Tablan's account | Log out | Help

Home Shop Dashboard

Annotation Job "Annotation Job - ANNIE + Measurements"

Details **Inputs** Outputs Execution Progress Download Results **Status**

Use this tab to specify which documents should be processed.
Select one or more inputs in the list on the left to modify their parameters.

<p>news-corpus.zip (Type: ZIP)</p> <p>Input Filters: .xml</p> <p>Input mime type: automatic</p> <p>Input encoding: UTF-8</p>	<p>Options for the selected inputs:</p> <p>Type: ZIP</p> <p>File Suffixes: <input type="text" value=".xml"/></p> <p>Mime Type Override: <input type="text"/></p> <p>Encoding: <input type="text" value="UTF-8"/></p> <p><input type="button" value="Cancel Changes"/> <input type="button" value="Apply Changes"/></p>	<p><input checked="" type="checkbox"/> An application file has been provided.</p> <p><input checked="" type="checkbox"/> Some input documents have been provided.</p> <p><input type="checkbox"/> You need to define some outputs.</p>
--	--	--

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011.
GATE is free open source software.
All our products and services are supplied **with no warranty**.

This website is licenced under the Creative Commons **BY-NC-SA licence**.
Cloud computing server time is payable in advance.



Job Editor: Outputs

Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA GATE CLOUD [Valentin Tablan's account](#) | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#)

Annotation Job "Annotation Job - ANNIE + Measurements"

Details | **Inputs** | **Outputs** | **Execution Progress** | **Download Results** | **Status**

On this tab you can specify what should happen to the output of your annotation job.
 Select one of the outputs on the left to change its parameters.

GATE_XML: \${filename}-gate.xml Annotation Set Annotation Type [Default] Person [Default] Location [Default] Organization [Default] Measurement	MIMIR: http://ip-10-230-31-156.eu-west-1.amazonaws.com Annotation Set Annotation Type [Default] Person [Default] Location [Default] Organization [Default] Measurement [Default] Money [Default] Date [Default] Ratio [Default] Sentence	XCES: \${filename}-measurements.xml Annotation Set Annotation Type [Default] Measurement
---	--	---

Details for the selected output:

Type: MIMIR

Index URL: http://ip-10-230-31-156.eu-

Username:

Password:

Output filters (which annotations should be saved):

Annotation Set:	Annotation Type:	
[Default]	Person	Delete
[Default]	Location	Delete
[Default]	Organization	Delete
[Default]	Measurement	Delete
[Default]	Money	Delete
[Default]	Date	Delete

Status

- ✓ An application file has been provided.
- ✓ Some input documents have been provided.
- ✓ Some outputs have been defined.

[Add New](#) [Delete Selected](#)

[Start Job](#) [Stop Job](#) [Delete Job](#)

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011. This website is licenced under the

Job Editor: Progress



Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA
GATE CLOUD

Valentin Tablan's account | Log out | Help

Home Shop Dashboard

Annotation Job "Annotation Job - ANNIE + Measurements"

Details Inputs Outputs **Execution Progress** Download Results **Status**

On this tab you can review the execution progress of your job.
No need to refresh: the execution log is **refreshed automatically**.

Job progress 17%

Execution Log

Time	Message
2011 May 13 17:48:38 UTC+1	Task 'split' completed with an exit status of 'success', execution time: 1 minute 3 seconds.
2011 May 13 17:43:07 UTC+1	Job started
2011 May 13 11:42:46 UTC+1	Job purchased.

Start Job Stop Job Delete Job

About | Contact

© The University of Sheffield, 1995-2011

This website is licenced under the

Job Editor: Done!



Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#)

Annotation Job "Annotation Job - ANNIE + Measurements"

Details **Inputs** **Outputs** **Execution Progress** **Download Results** **Status**

On this tab you can review the execution progress of your job.
No need to refresh: the execution log is **refreshed automatically**.

Job progress 100%

Time	Message
2011 May 13 18:14:43 UTC+1	Task 'join' completed with an exit status of 'success', execution time: 2 seconds.
2011 May 13 18:14:43 UTC+1	Job completed.
2011 May 13 18:13:42 UTC+1	Task 'gcp' completed with an exit status of 'success', successful documents: 326, failed documents: 0, execution time: 1 minute 31 seconds.
2011 May 13 18:12:42 UTC+1	Task 'gcp' completed with an exit status of 'success', successful documents: 3998, failed documents: 0, execution time: 11 minutes 19 seconds.
2011 May 13 18:01:41 UTC+1	Task 'gcp' completed with an exit status of 'success', successful documents: 3930, failed documents: 0, execution time: 11 minutes 57 seconds.
2011 May 13 17:48:38 UTC+1	Task 'split' completed with an exit status of 'success', execution time: 1 minute 3 seconds.
2011 May 13 17:42:07 UTC+1	Job started

Start Job Stop Job Delete Job

Status
✔ The job has finished running.

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011

This website is licenced under the

Job Editor: Results



Annotation Job: "Ann..."

https://gatecloud.net/jobManagement/jobDetails/103

BETA
GATE CLOUD

Valentin Tablan's account | [Log out](#) | [Help](#)

[Home](#) [Shop](#) [Dashboard](#)

Annotation Job "Annotation Job - ANNIE + Measurements"

Details **Inputs** **Outputs** **Execution Progress** **Download Results** **Status**

On this tab you can download the results of your job, after the execution completes.
This annotation job has finished running; download your results using the links below.

[report.html](#)
[bbcnews-20110419.tar.gz-part001-out.zip](#)
[bbcnews-20110419.tar.gz-part002-out.zip](#)
[bbcnews-20110419.tar.gz-part003-out.zip](#)

The job has finished running.

[About](#) | [Contact](#)

© The University of Sheffield, 1995-2011. This website is licenced under the

Thank You!



More info at:

- <https://gatecloud.net>
- <http://gate.ac.uk>
- <http://facebook.com/gate.ac.uk>
- <http://twitter.com/GateAcUk>
- Talk to us during coffee breaks ;)