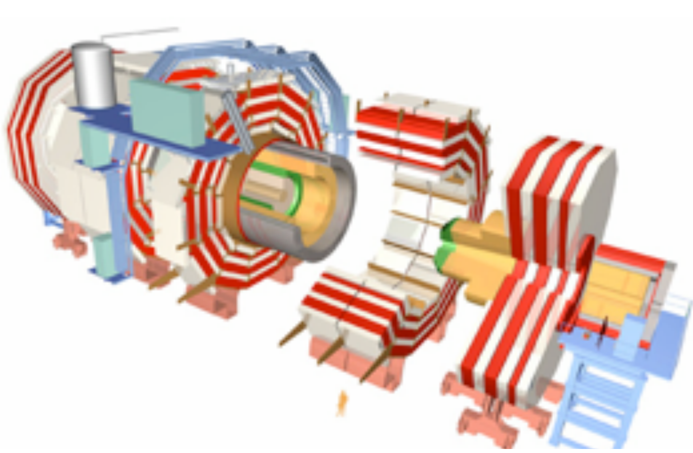
Landslides, Couches and Particle Physicists

Michael Wallace, Research Assistant
University of Bristol



What?

- A collaboration between Particle Physics Group and Geographical Sciences at the University of Bristol
- Improving scalability of landslide modelling software
- Solving associated data management problems

Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

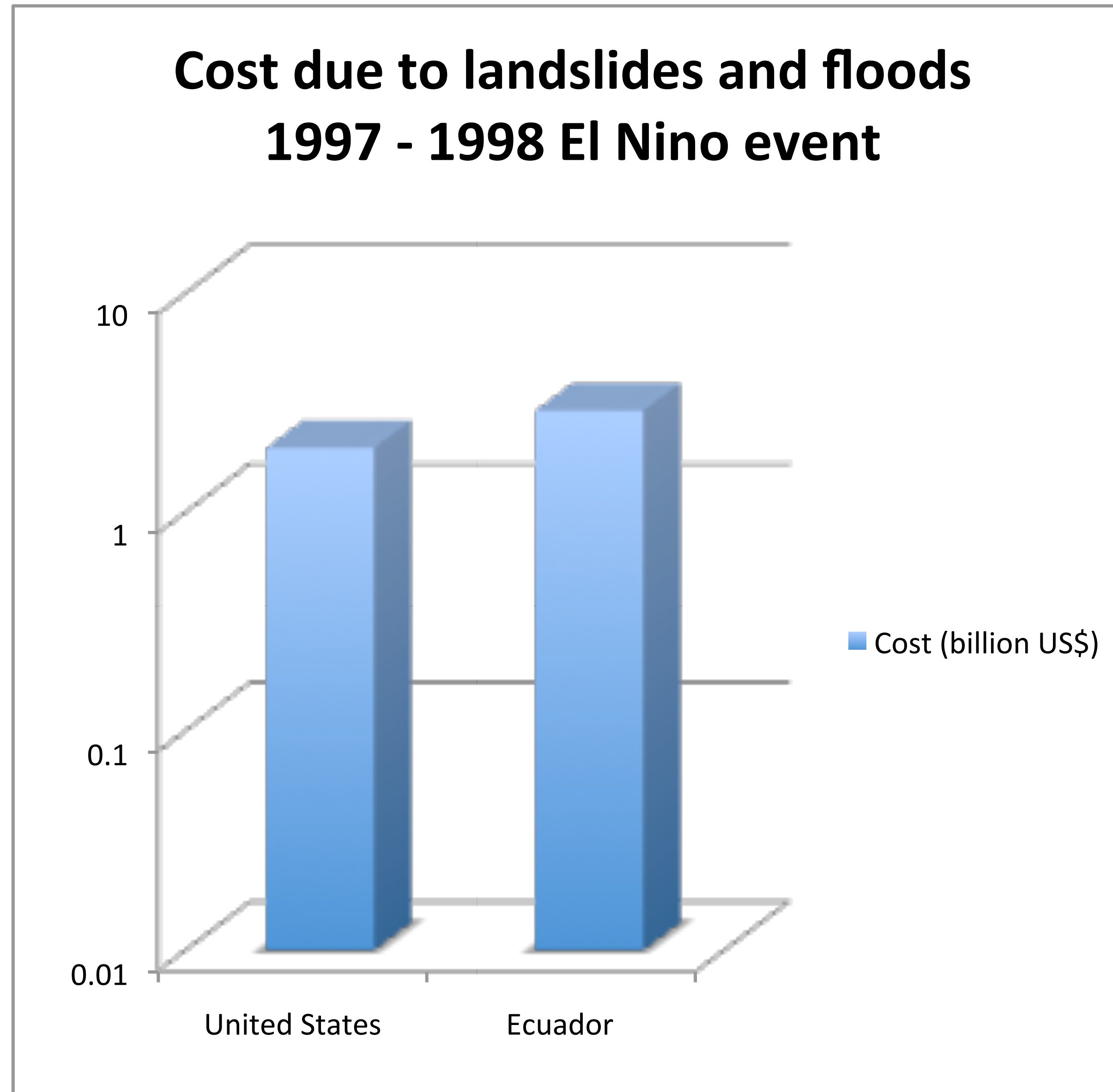
Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

Landslides

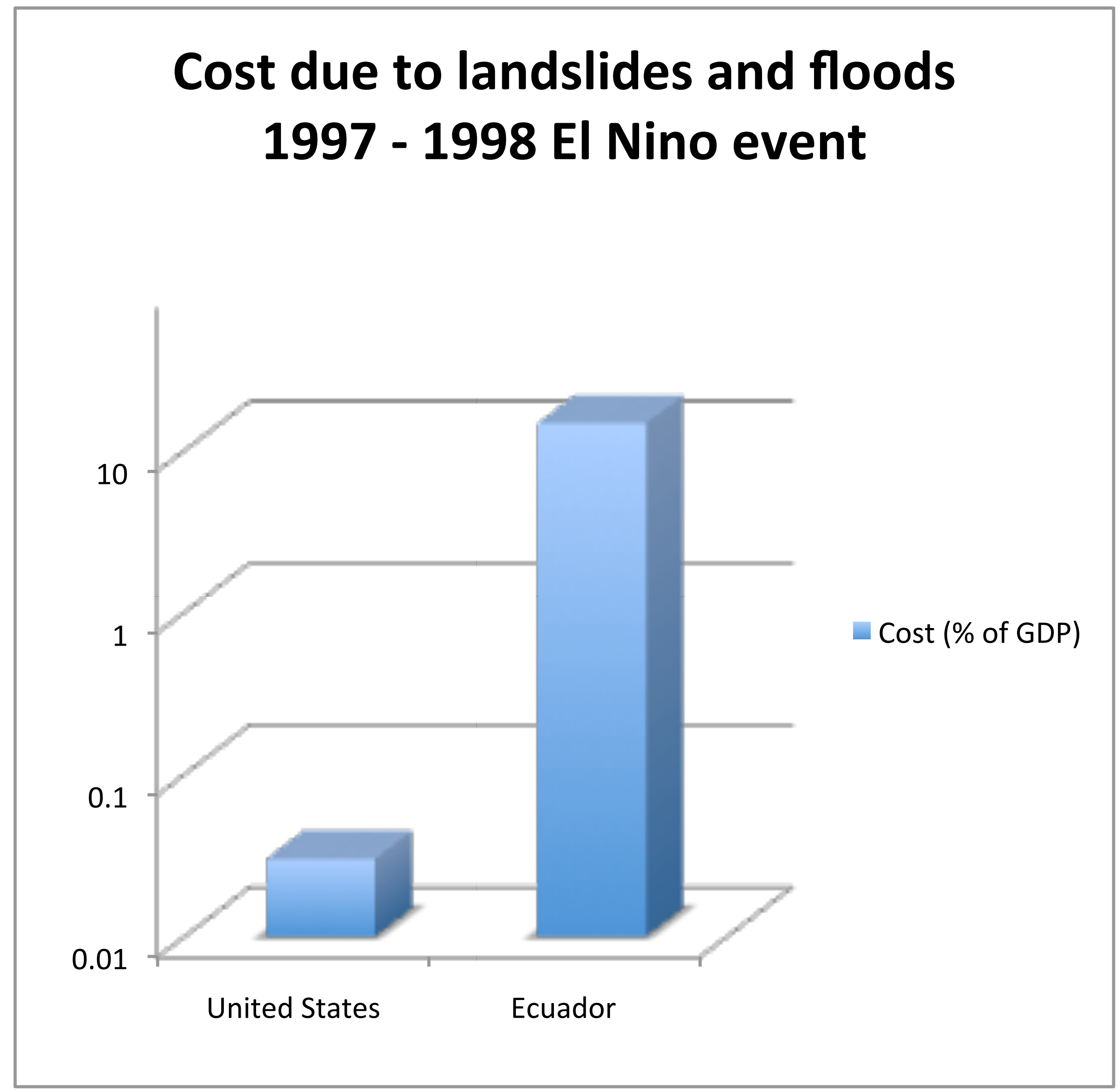
Economic impact

- Developing countries more vulnerable
- Less resources for disaster recovery
- “Accumulated risk” due to poorly planned development schemes



Economic impact

- Developing countries more vulnerable
- Less resources for disaster recovery
- "Accumulated risk" due to poorly planned development schemes



Indirect costs are orders of magnitude higher than direct costs

Disruption to critical infrastructure whilst clear up and repair is ongoing

Indirect costs are orders of magnitude higher than direct costs

Clear up and repair

Reducing clear up and repair times can massively reduce overall costs

Reducing clear up and repair times can massively reduce overall costs

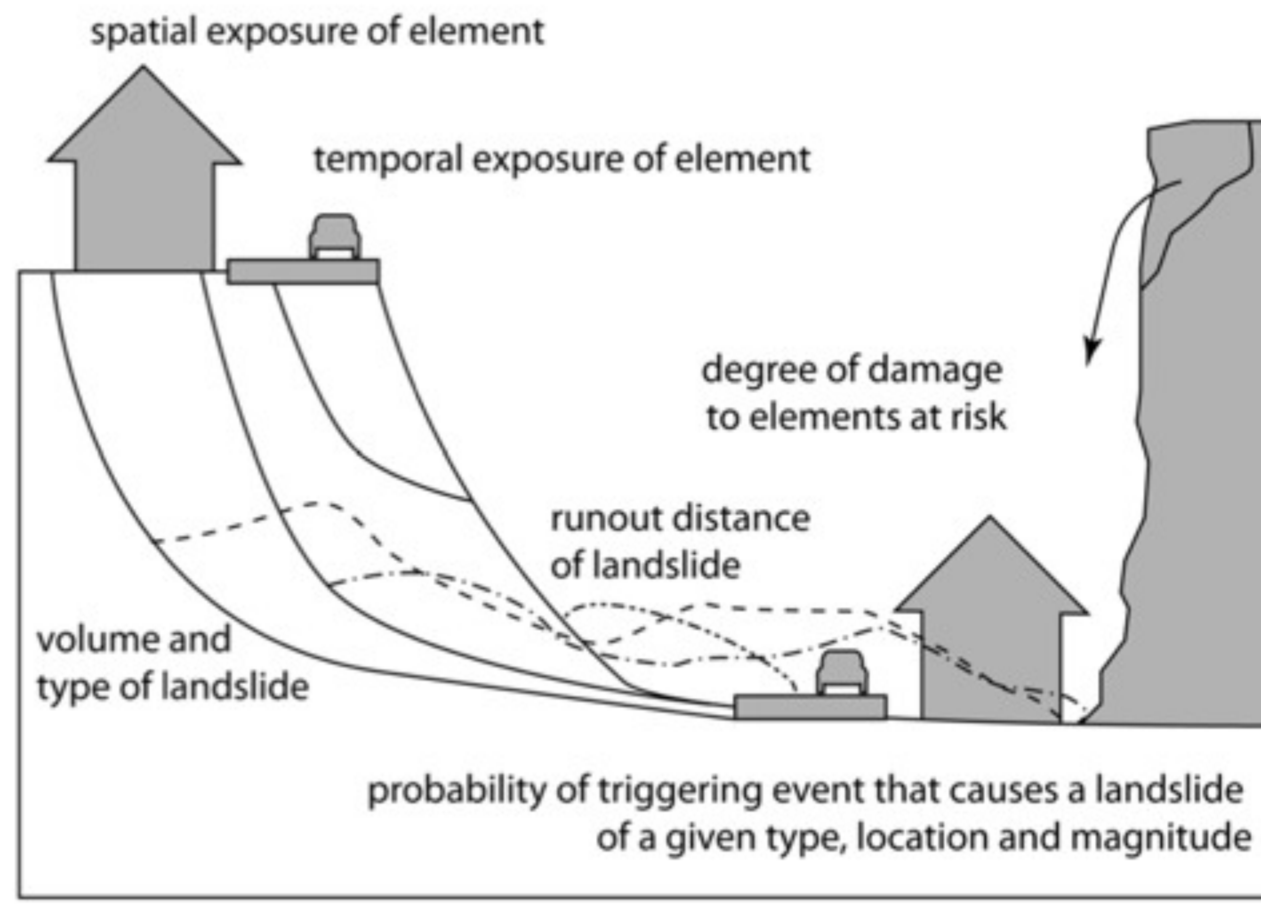
This is one of the aims of disaster response planning

Disaster planning

- Risk is a function of:
  - hazard
  - exposure
  - vulnerability

Disaster planning

- Landslide risk of a slope for a given storm
- Size and "runout" of a landslide
- Economic impact

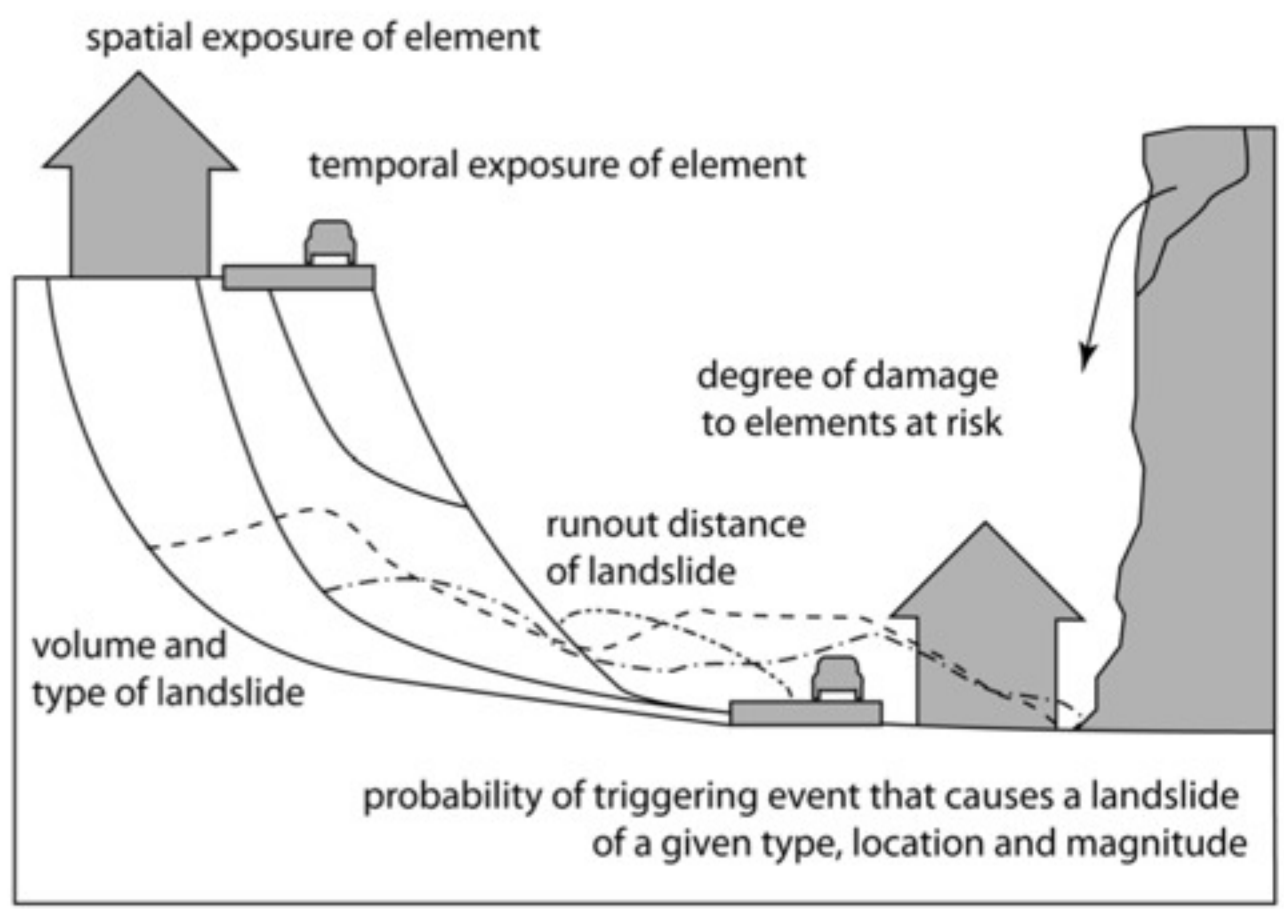


courtesy of Dr Liz Holcombe

Disaster planning

- Landslide risk of a slope for a given storm
- Size and "runout" of a landslide
- Economic impact

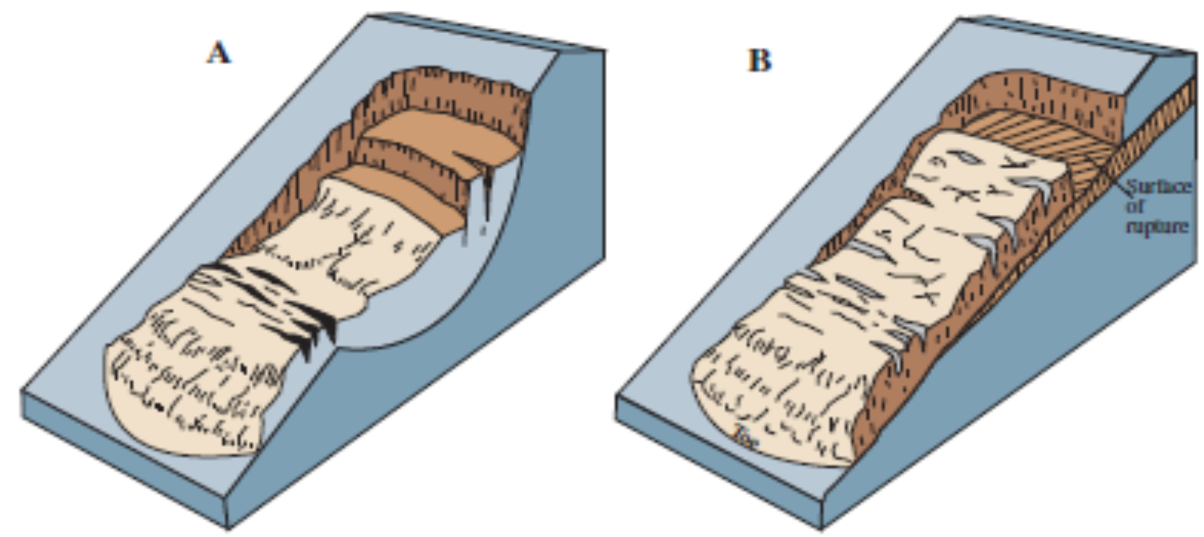
These can all be predicted using software



courtesy of Dr Liz Holcombe

Landslide modelling

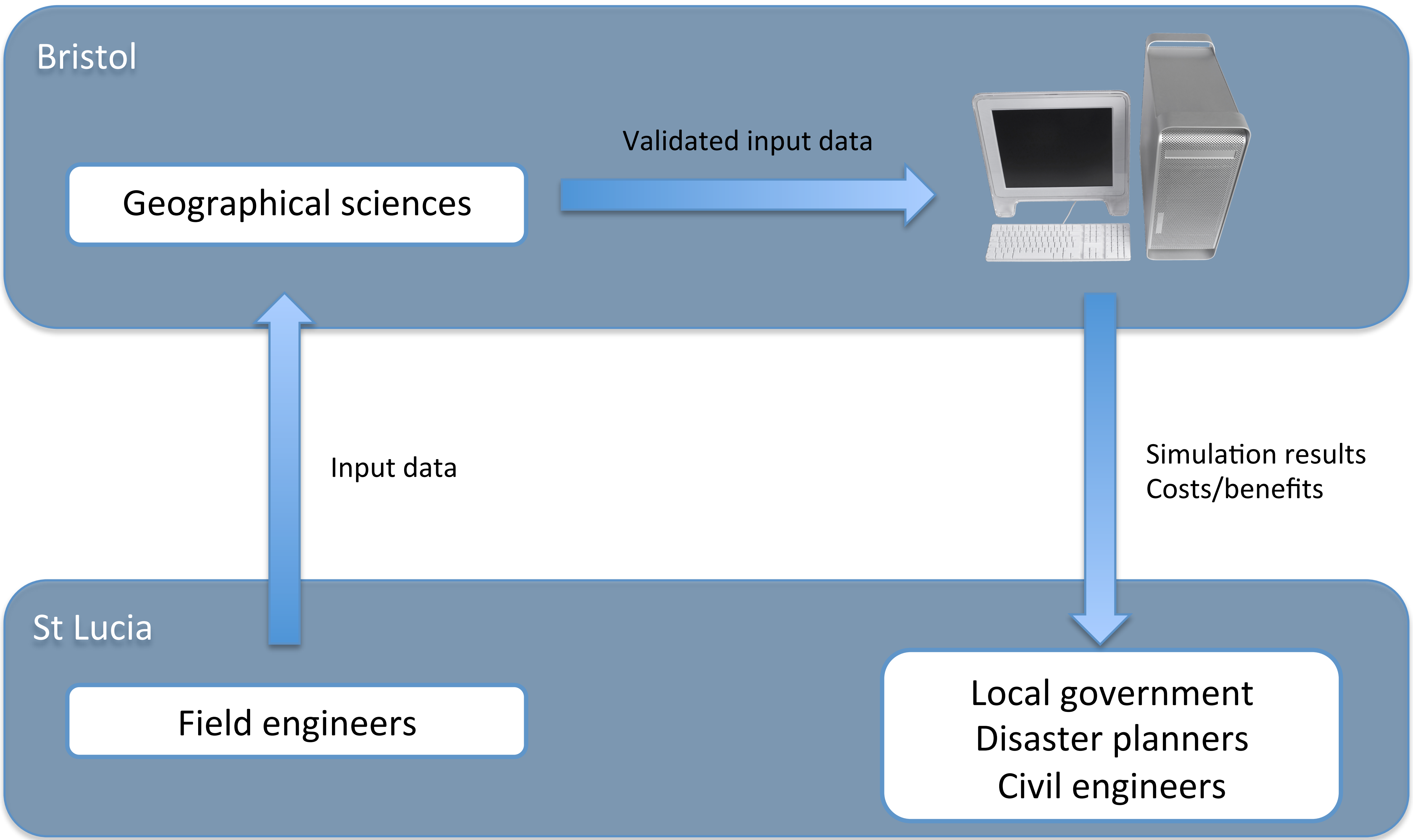
CHASM
Landslide hazard



Rotational landslide Translational landslide

QUESTA
Exposure, vulnerability and risk

Simulation workflow



Complexity

- Simulation scenarios
  - One or more parameters are variable
  - Run simulation for each set of parameters
- Example:
  - Cut slope angles from 50° to 70° with 1° increments

Complexity

- Stochastic simulations
  - One or more parameters are stochastically generated with a given mean and standard deviation
  - Run simulation for each set of parameters
- Example:
  - Soil saturation

Complexity

Slopes	Cut slope angles	Stochastic parameters	Variations for each stochastic parameter	Output files	Runtime
1	1	0	0	1	0.25 (cpu hours)

Complexity

Slopes	Cut slope angles	Stochastic parameters	Variations for each stochastic parameter	Output files	Runtime
1	25	0	0	25	6.25 (cpu hours)

Complexity

Slopes	Cut slope angles	Stochastic parameters	Variations for each stochastic parameter	Output files	Runtime
1	25	5	10	1250	312.5 (cpu hours)

Complexity

Slopes	Cut slope angles	Stochastic parameters	Variations for each stochastic parameter	Output files	Runtime
100	25	5	10	125000	31250 (cpu hours)

Complexity

Slopes	Cut slope angles	Stochastic parameters	Variations for each stochastic parameter	Output files	Runtime
100	25	5	10	125000	4 (**years**)

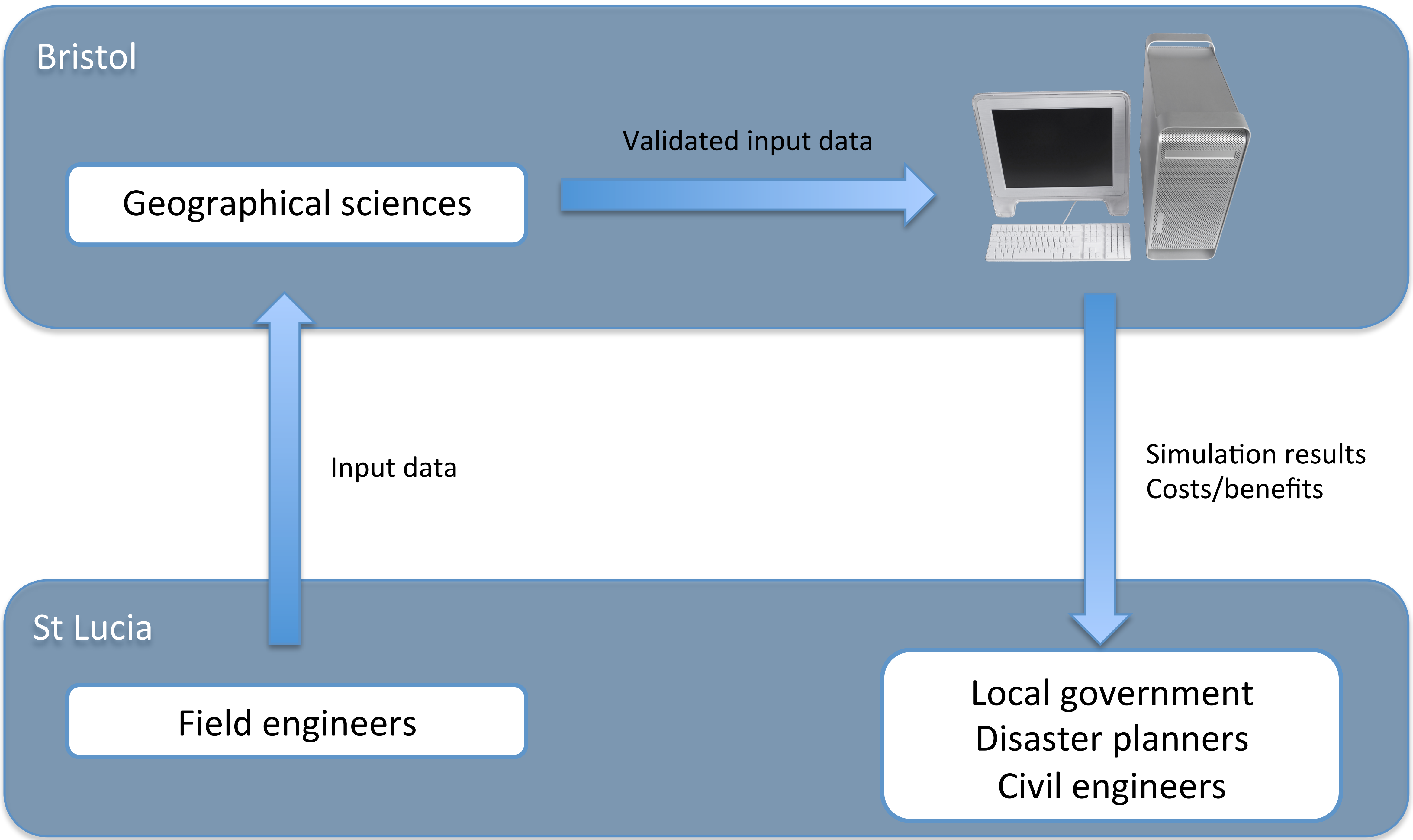
Complexity

- Runtime can be managed by parallel execution of simulations
- Each result set a few 10's GB in size
- There will be many result sets
- Data volume needs to be managed

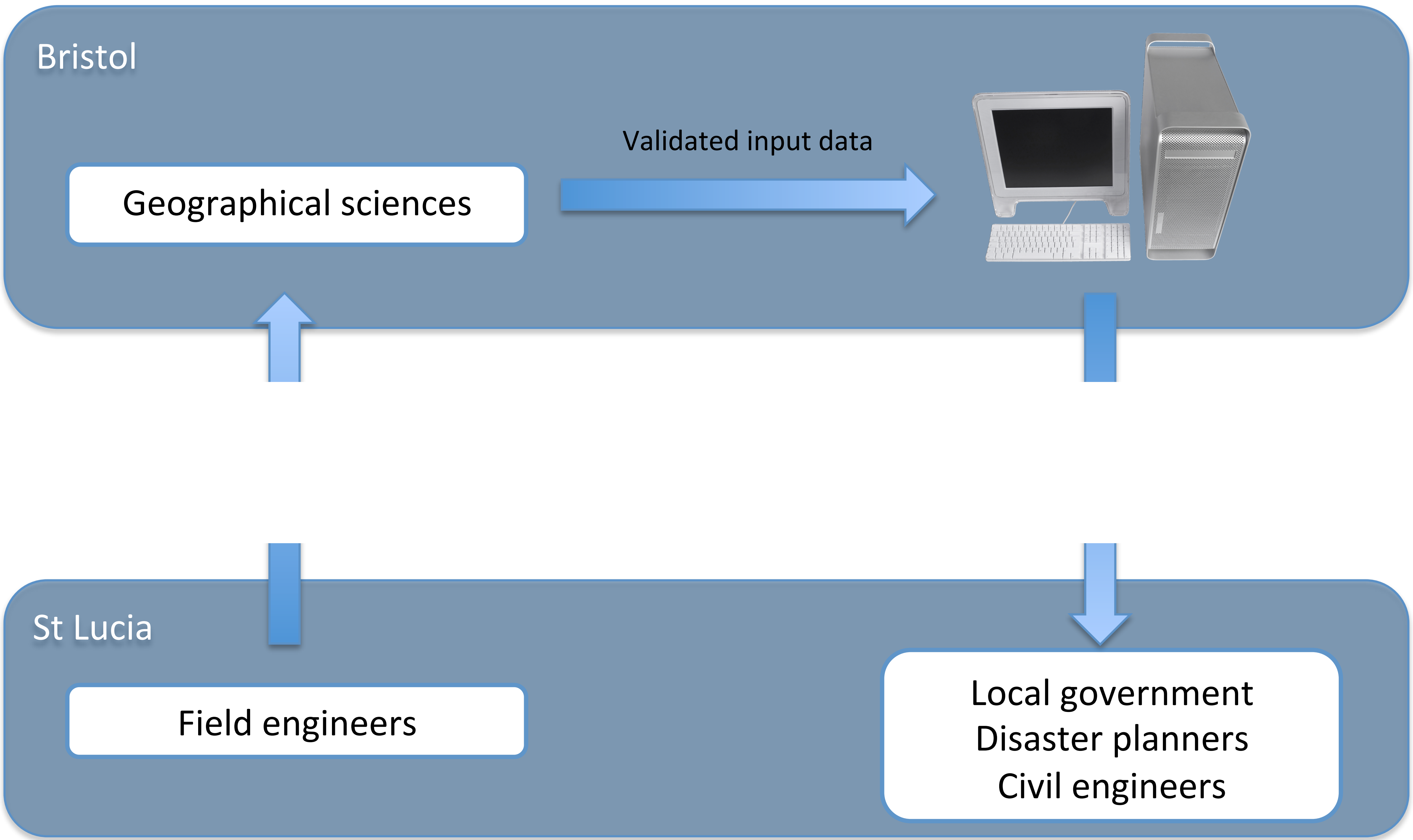
Data management

- 10s of TB generated per year for a typical usage scenario
- No facilities for managing this data

And another problem...



And another problem...



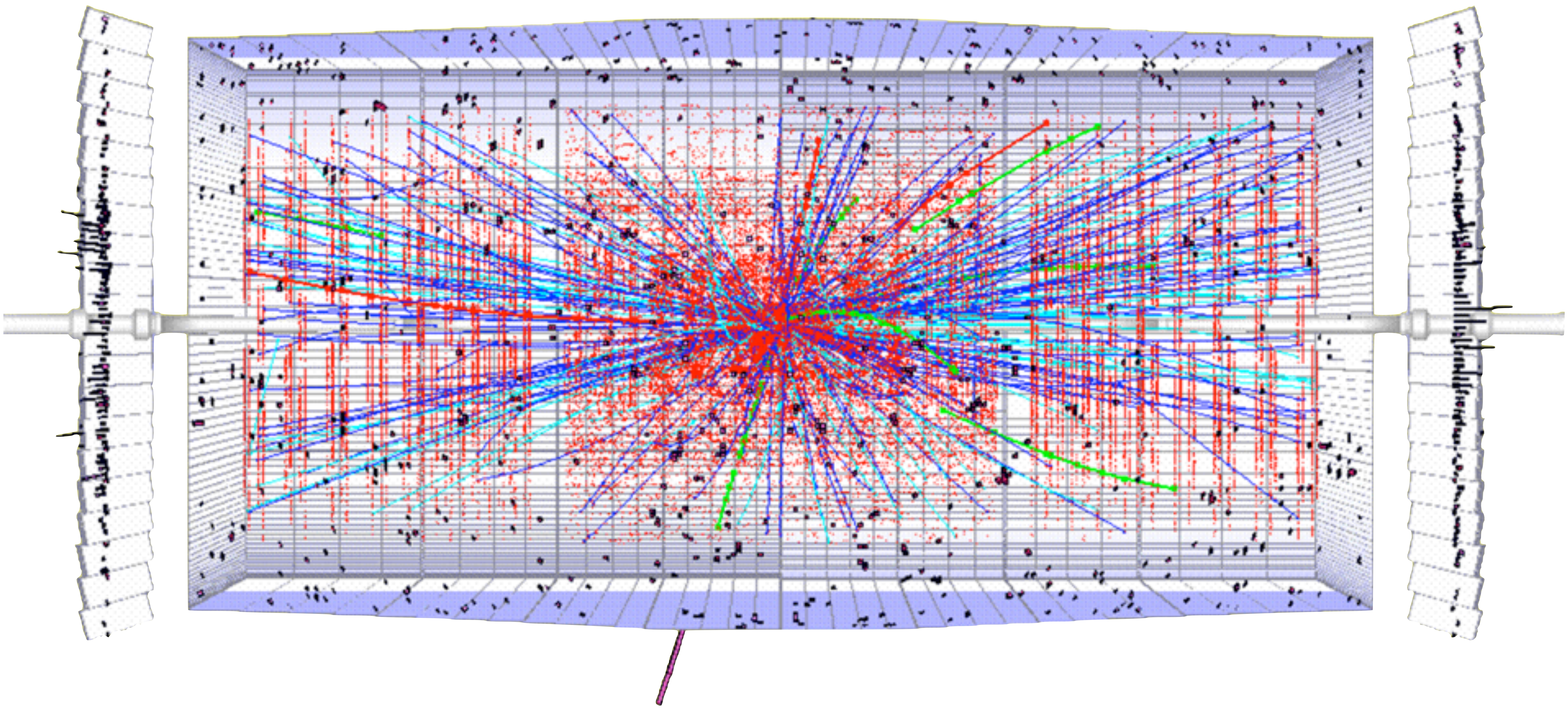
Requirements Summary

- Enable running 1000s of simulations in parallel
- Long term storage of simulation results
- Query-able simulation results
- Provide offline access to subsets of data

Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

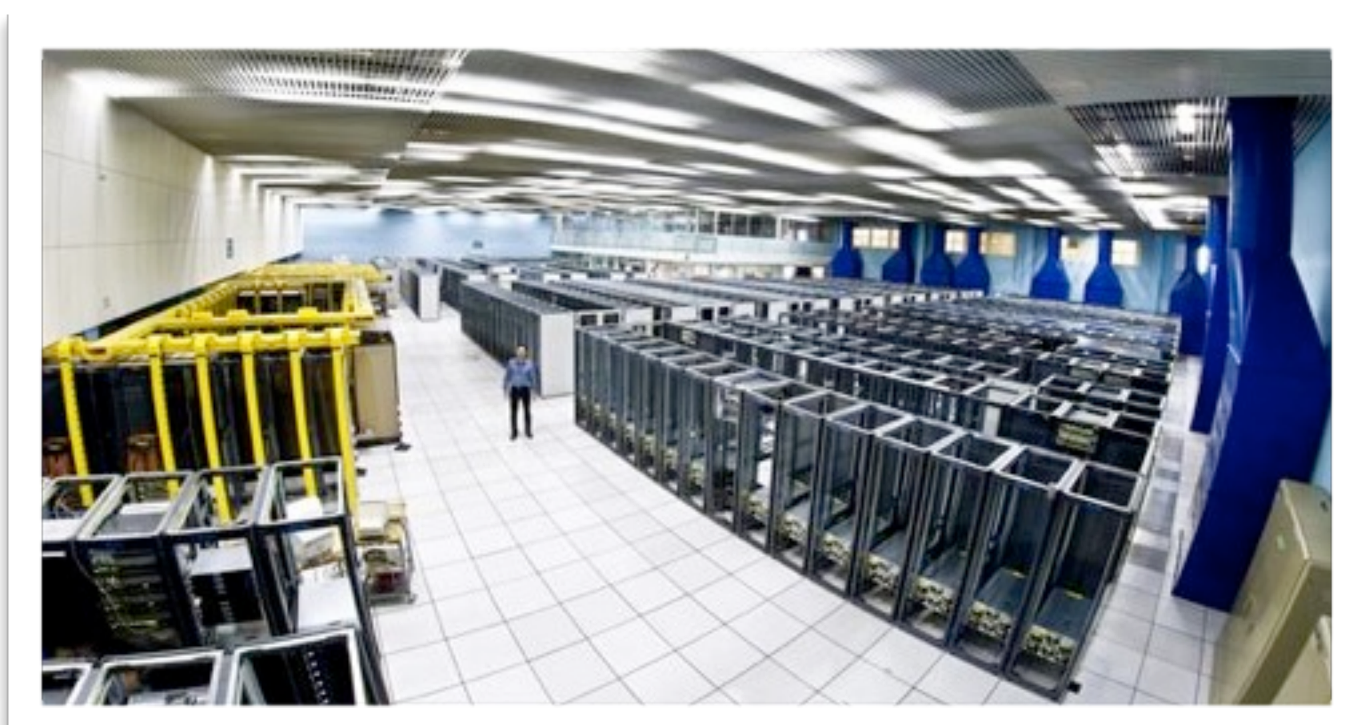
Particle physics



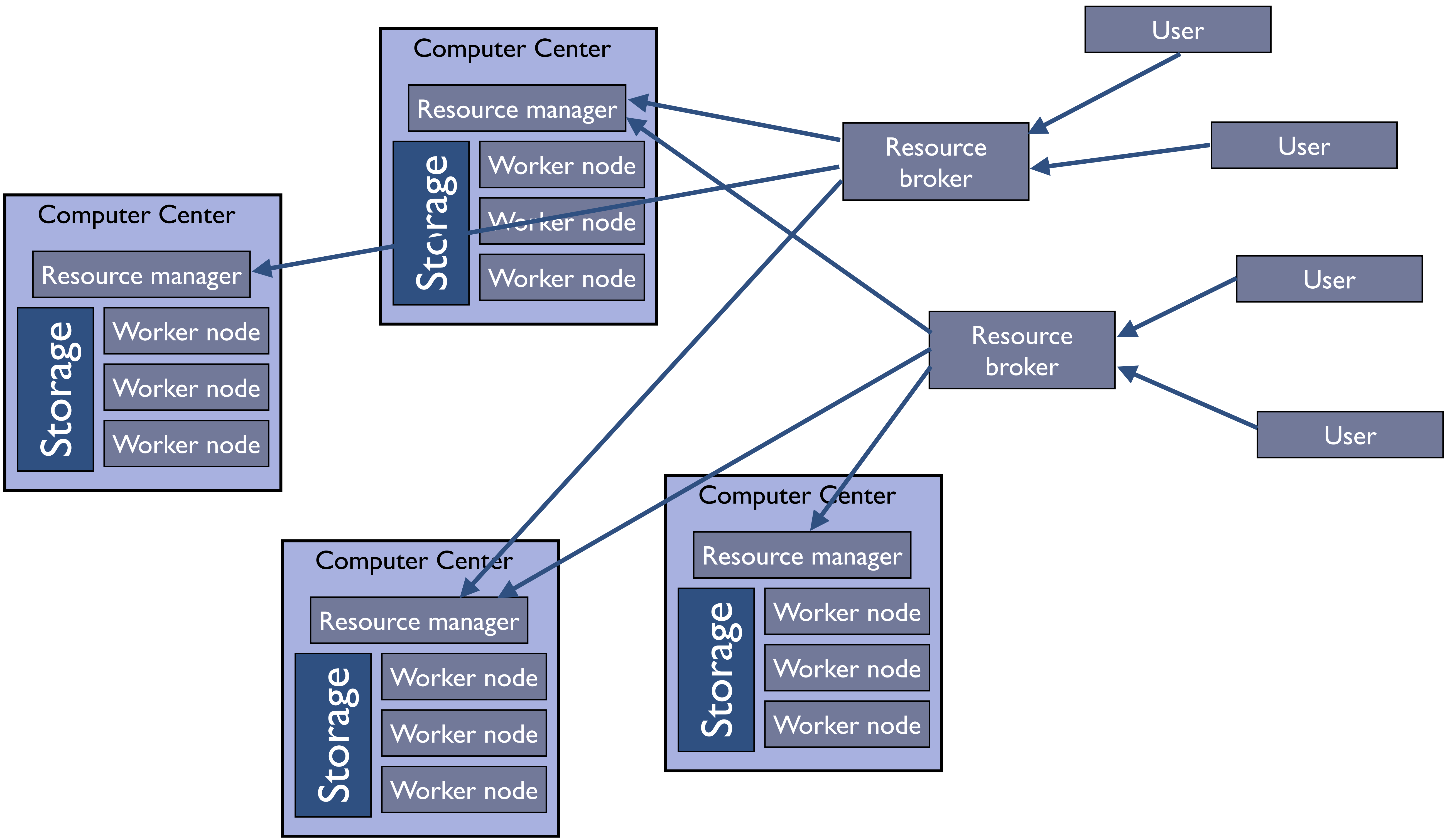
courtesy of Dr James Jackson

The WLCG

- Allows sharing of storage and compute power
- Physicists use the grid to perform complex analysis on petabyte-scale data



Grid computing



Grid vs Batch

- Inter-organisational administration
- Heterogeneous resources
- Hostile environment
- Multiple data centres, globally distributed

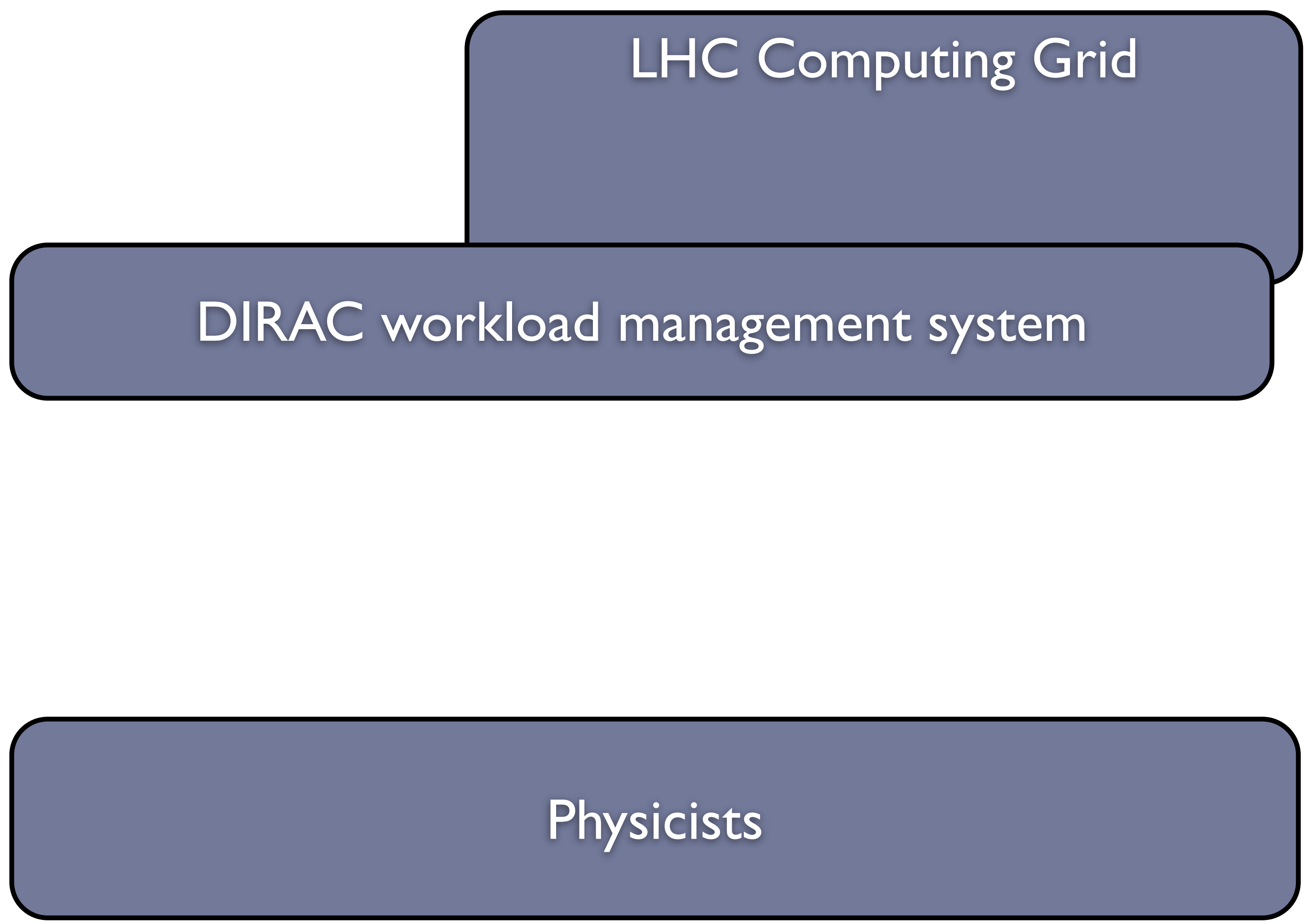
Grid vs Cloud (IaaS)

- Walled garden
- Many software stacks for specific use cases
- Grid approach requires hardware to be purchases up-front
- Requires expertise to use
- Cloud/utility computing better suited for commercial applications

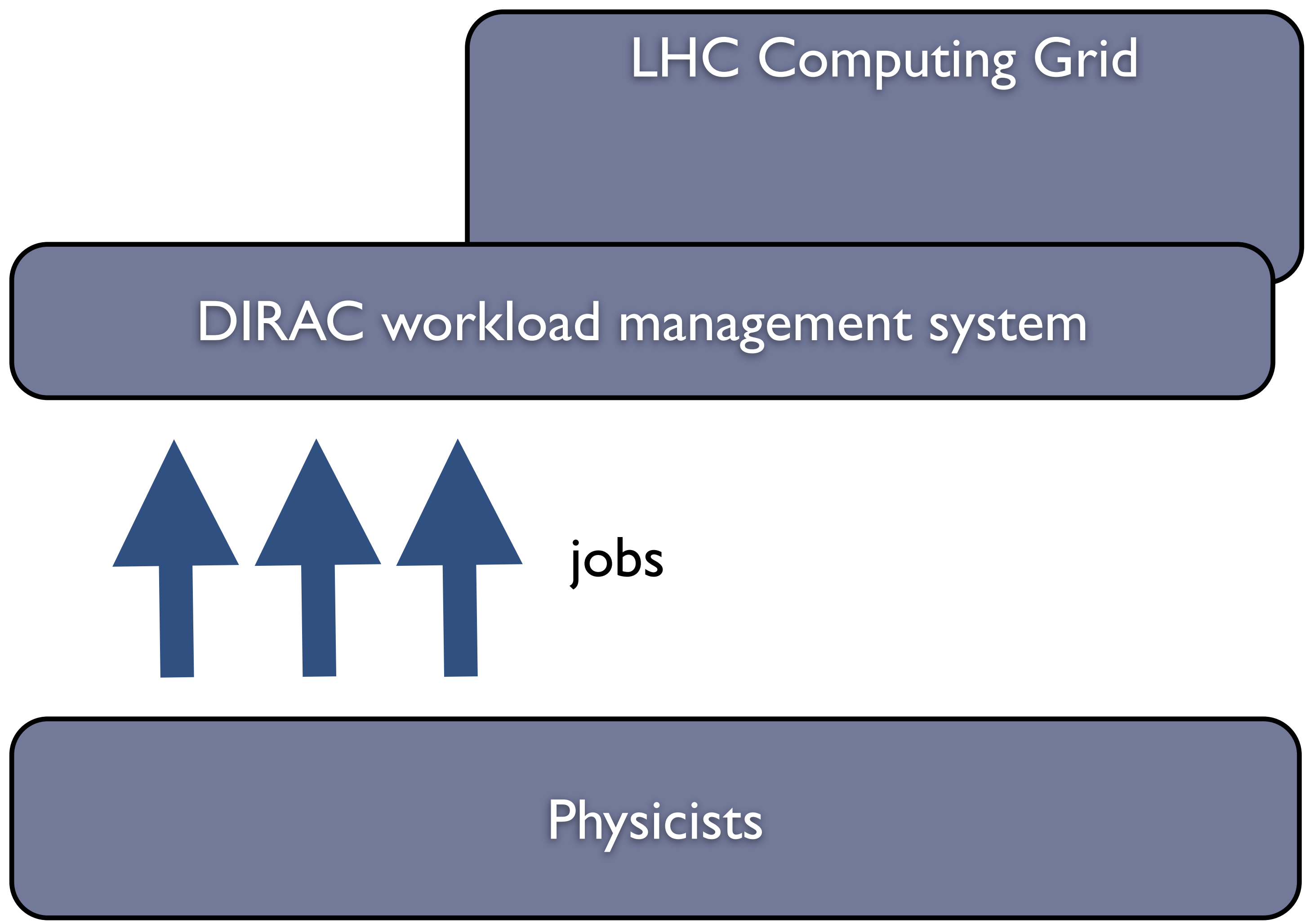
DIRAC

- Distributed Infrastructure with Remote Agent Control
- Manages user-submitted jobs on the grid
- Plugin architecture supports numerous compute backends

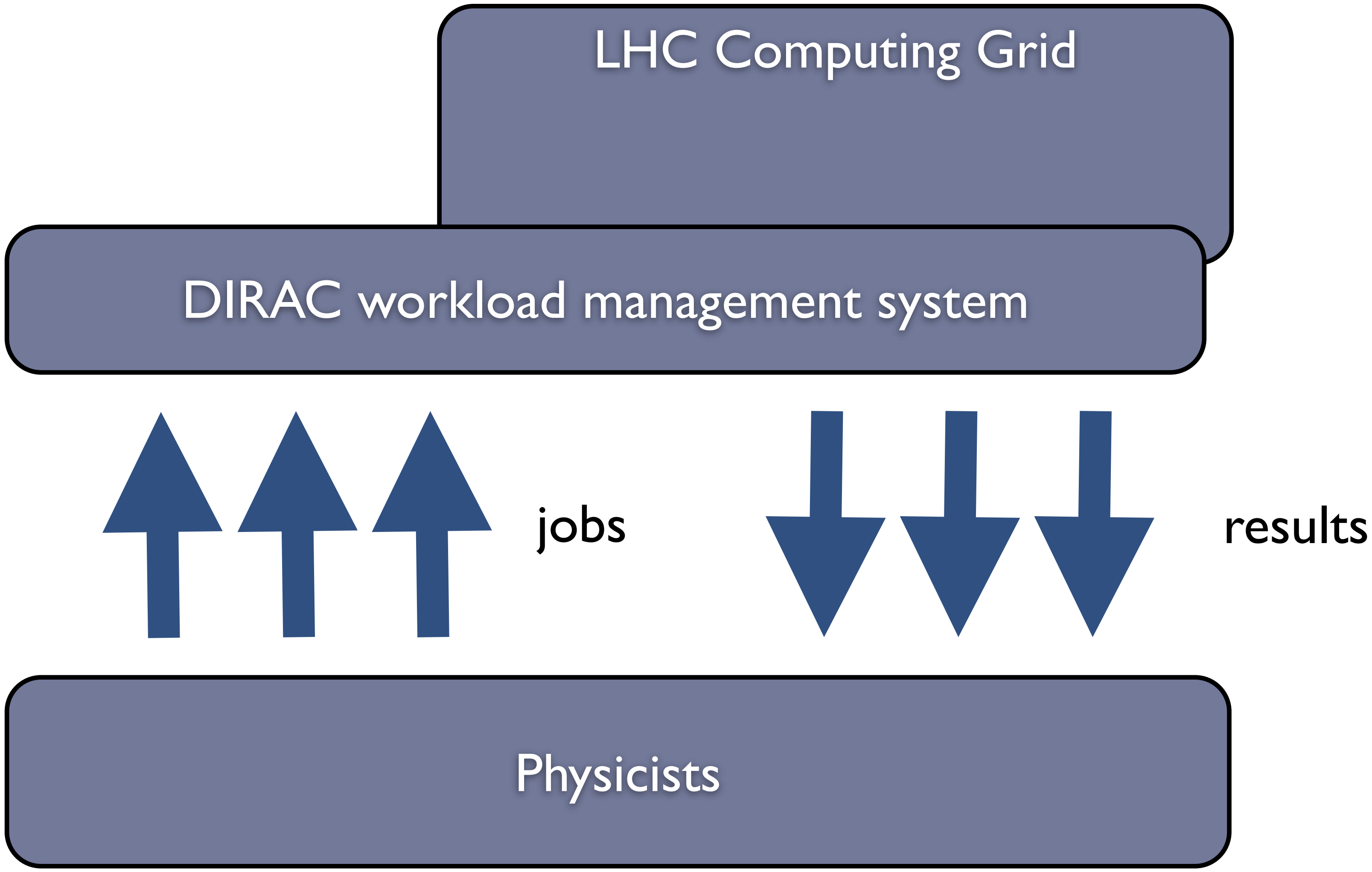
Particle physics workflow



Particle physics workflow



Particle physics workflow



Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

Data management

- CouchDB chosen for data storage
  - Satisfies key requirements
- Due to availability and scalability concerns need to consider clustered solution; BigCouch

What is CouchDB?

- Top-level Apache project
- Document-oriented database
- Data stored as JSON documents, indexed by UID
- Views allow indexing by any document field
- Loves to replicate

Why CouchDB?

- Multi-master replication out of the box
- Geo-indexing (via GeoCouch)
- Runs on phones, laptops, servers & clusters
- Provides application logic and can serve client-side code (CouchApp)
- Browser + CouchDB = Web application

Why not CouchDB?

- Greater storage requirements than a typical RDBMS for “most” data
- Uncertainty over write speeds - need to test how many writes per second we can support

Why not CouchDB?

- Greater storage requirements than a typical RDBMS for “most” data

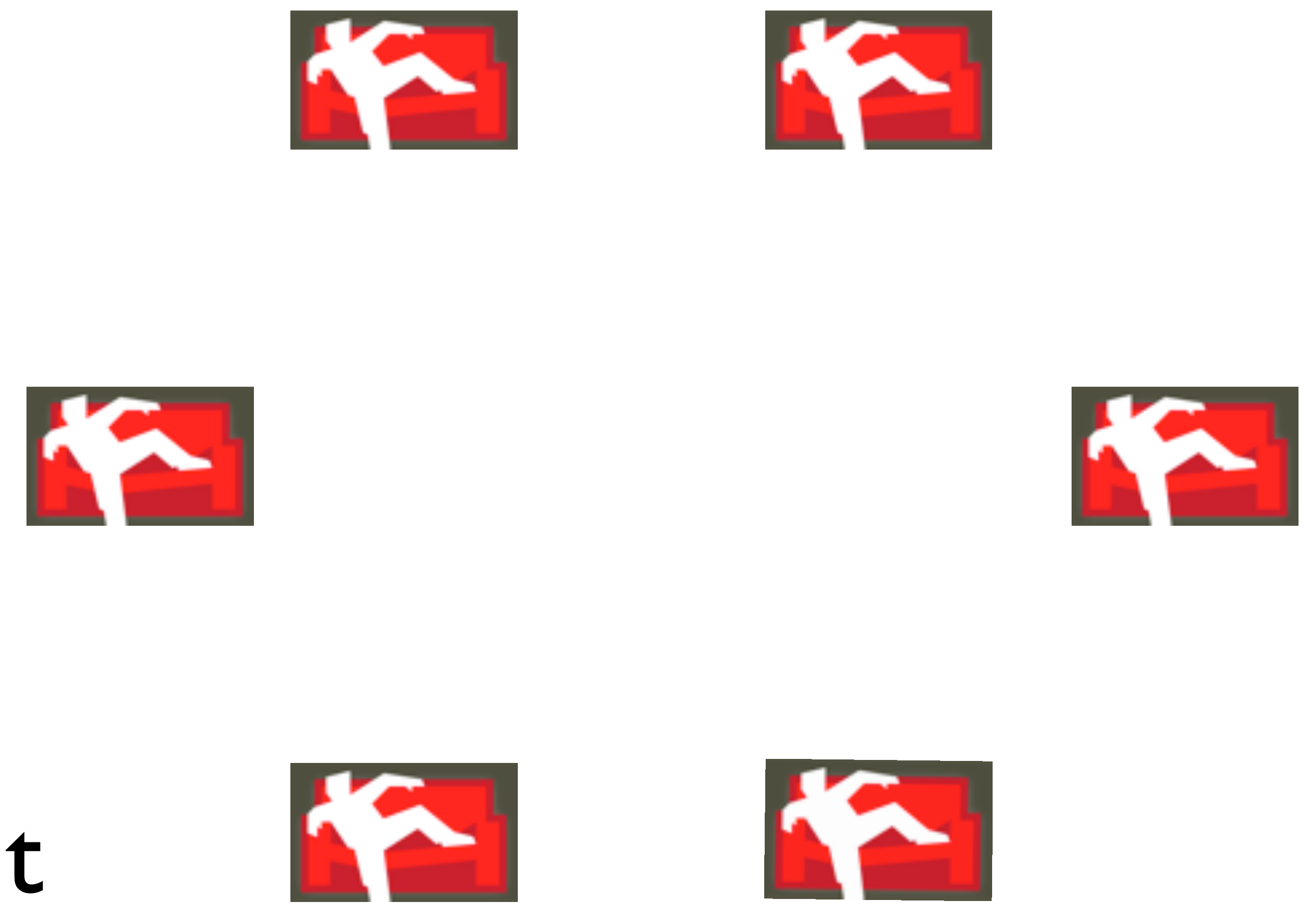
Snappy compression coming in 1.2

- Uncertainty over write speeds - need to test how many writes per second we can support

BigCouch

- Dynamo-influenced clustered CouchDB
- Flexible partitioning
- Read/write quorums per query
- Can have partition tolerance with some level of availability and read-your-writes consistency

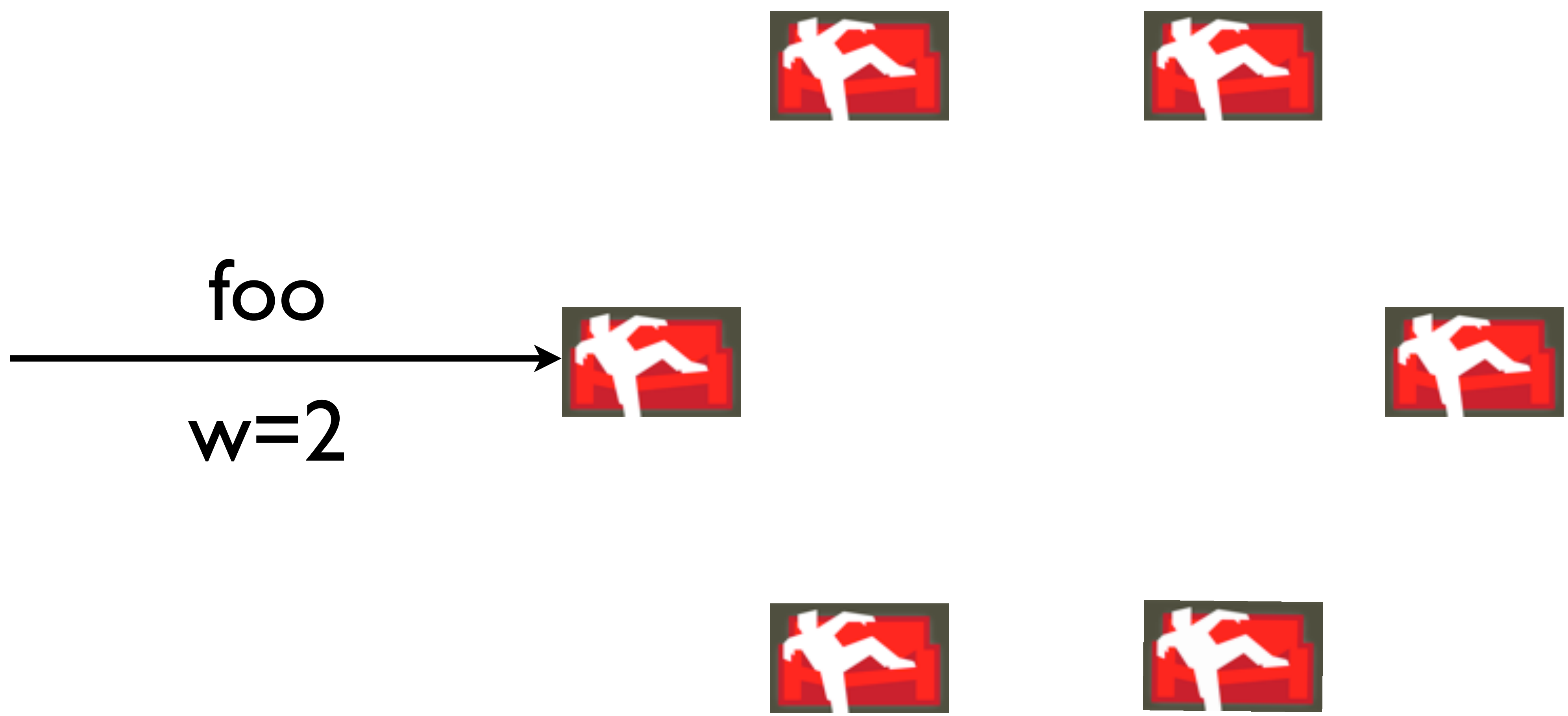
BigCouch



N = replication constant
Q = number of partitions
W = number of successful writes before response
R = number of successful reads before response

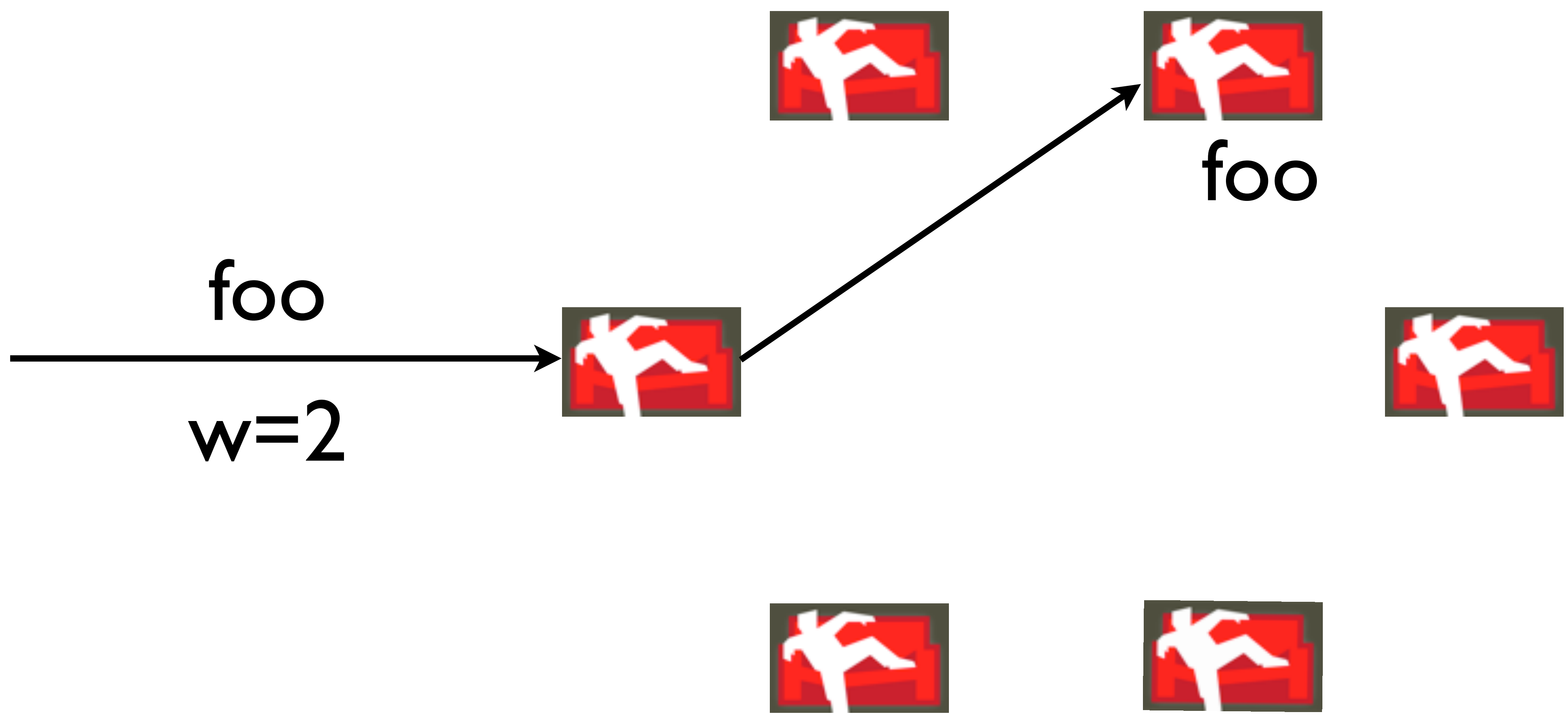
BigCouch

N=3
Q=2



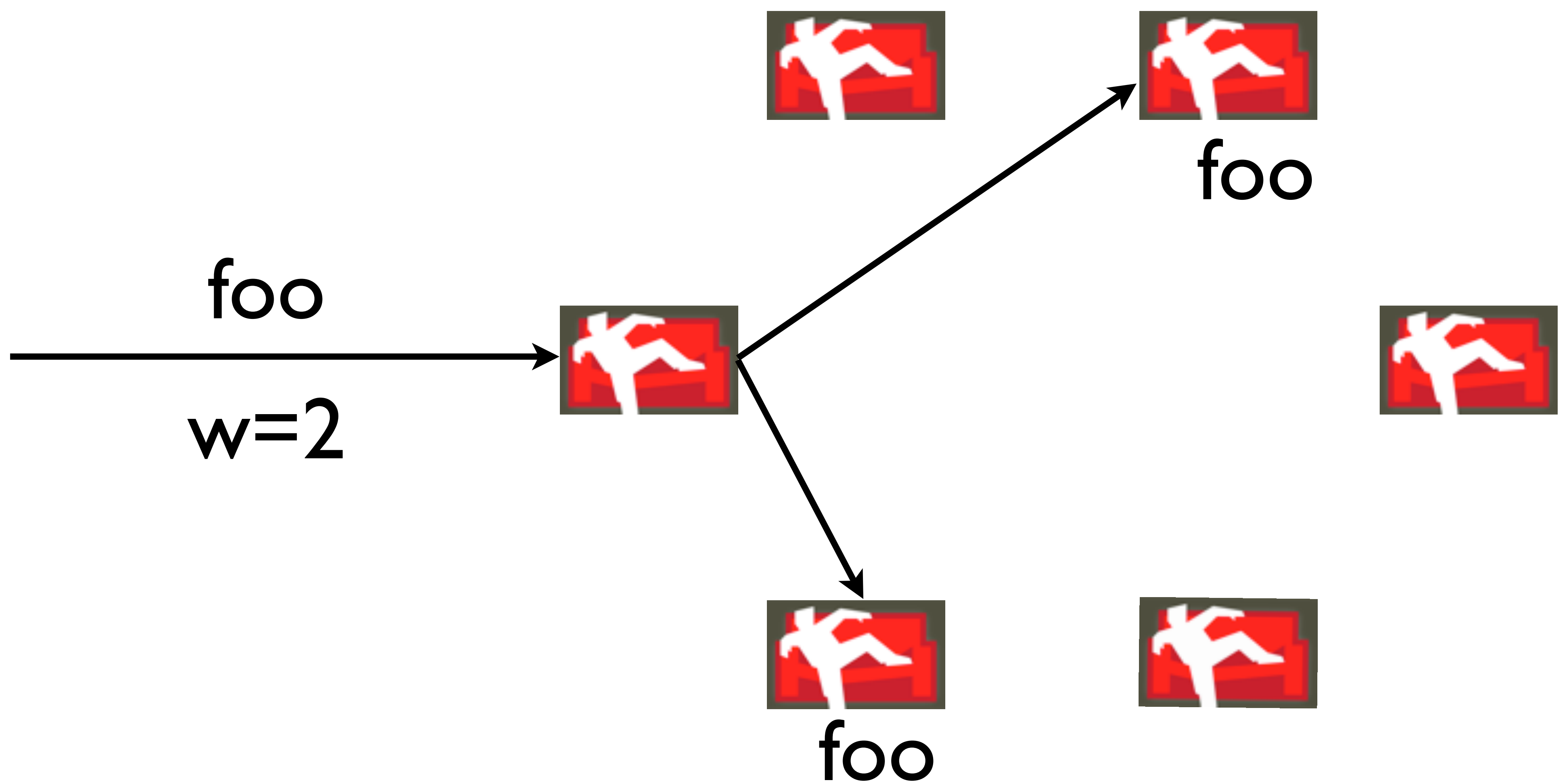
BigCouch

N=3
Q=2



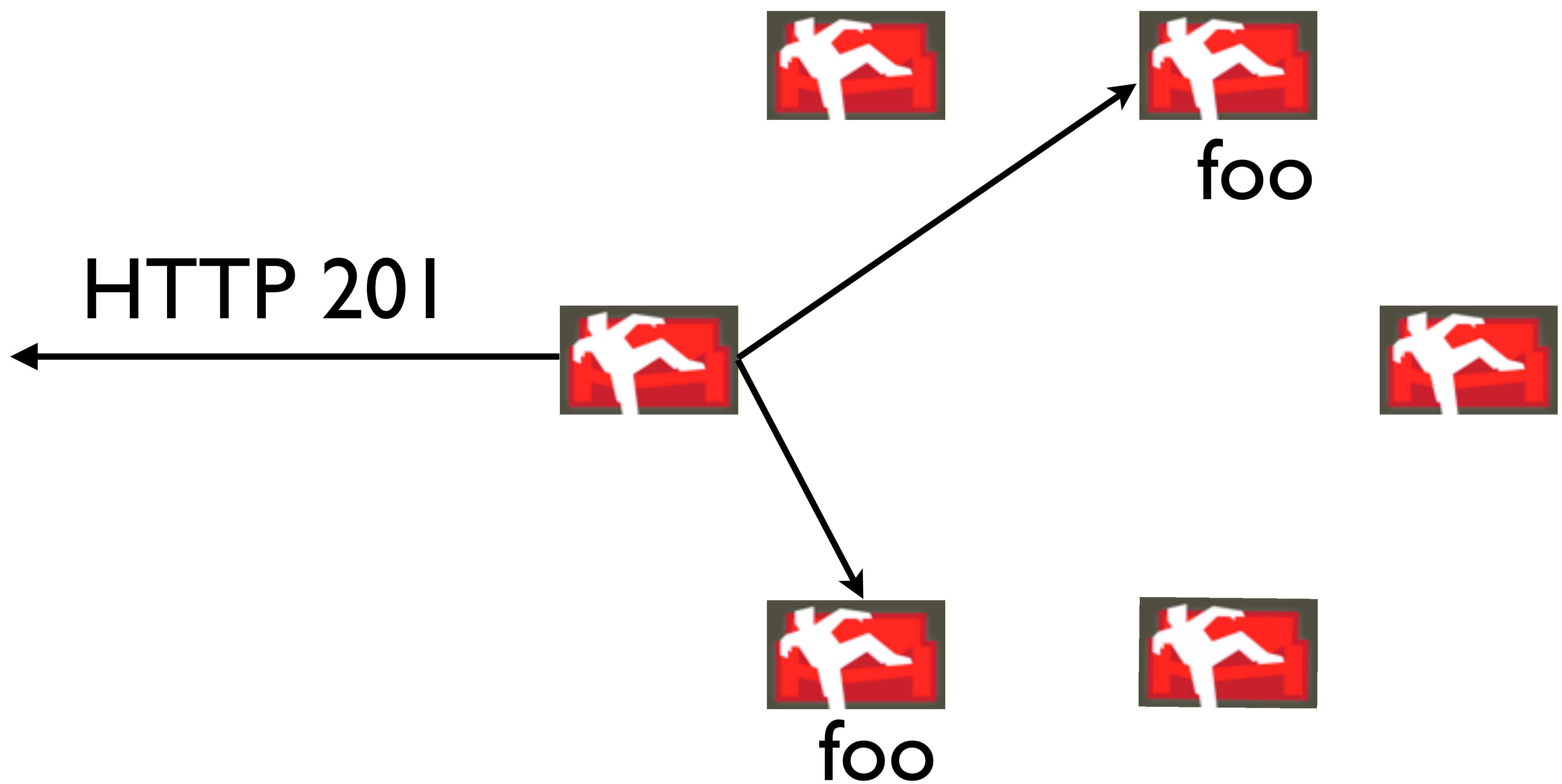
BigCouch

N=3
Q=2



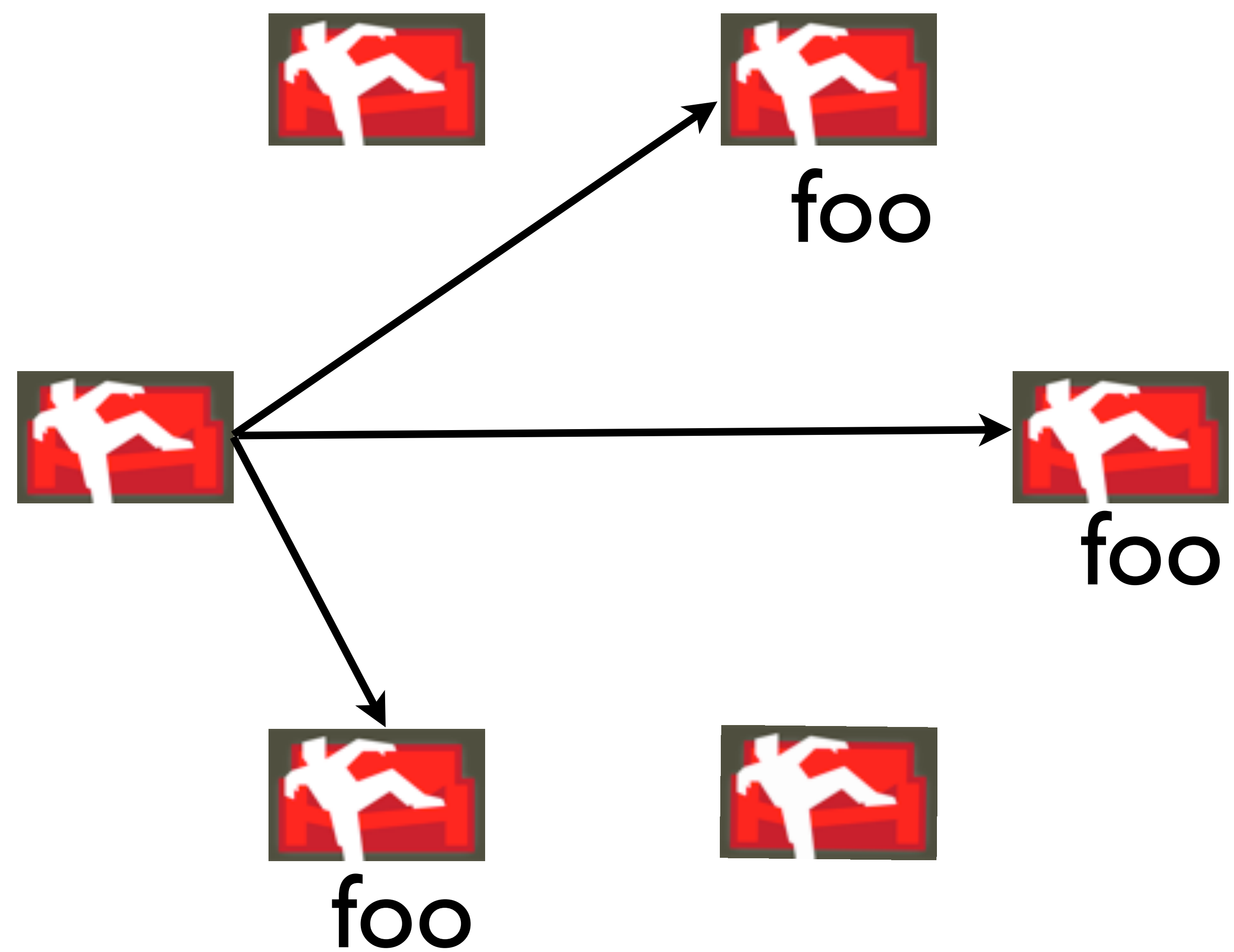
BigCouch

N=3
Q=2



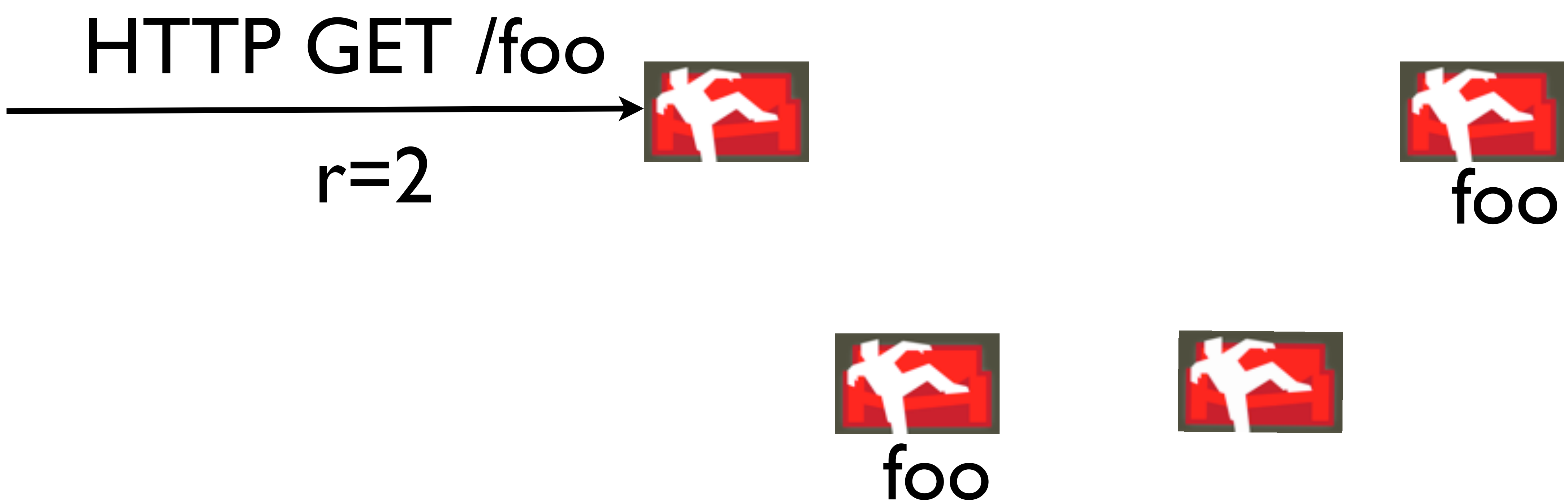
BigCouch

N=3
Q=2

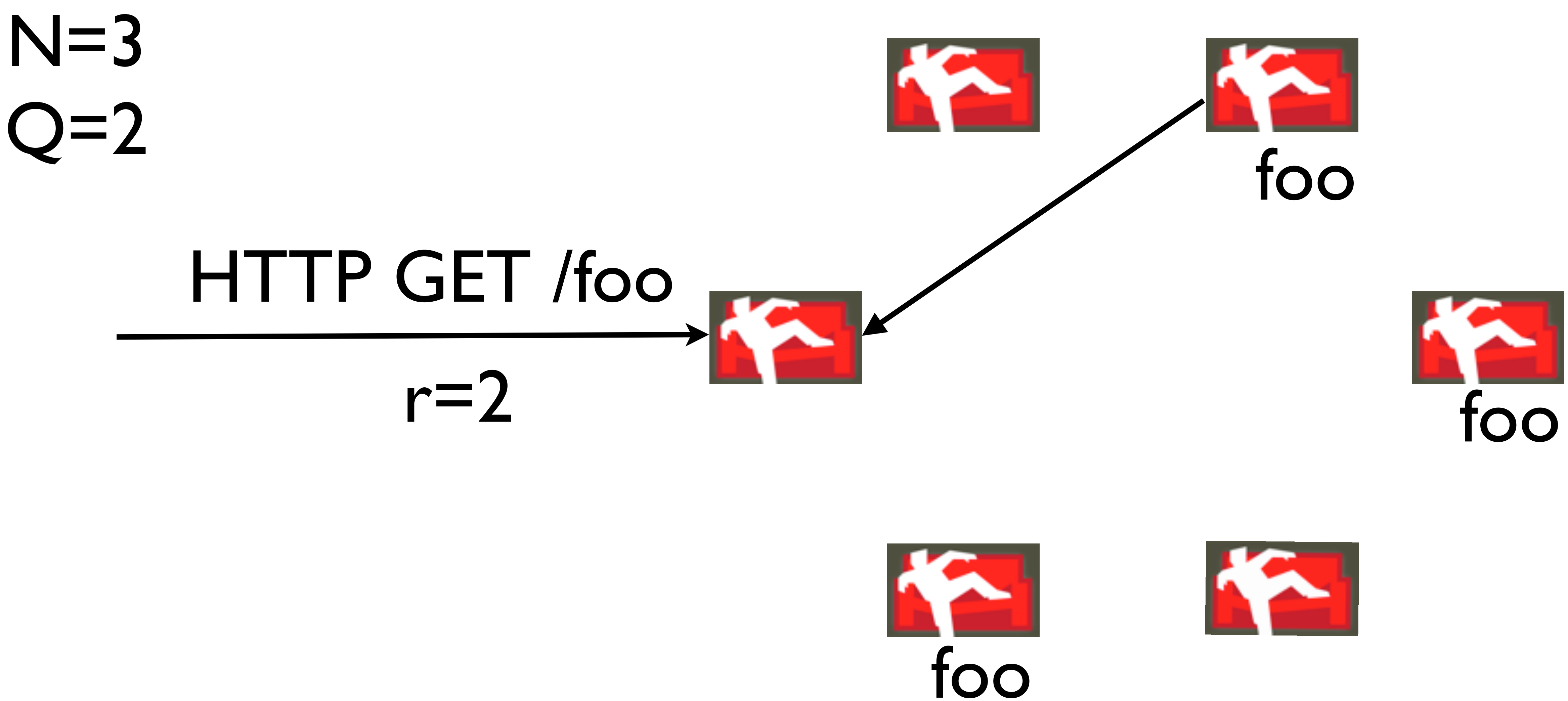


BigCouch

N=3
Q=2

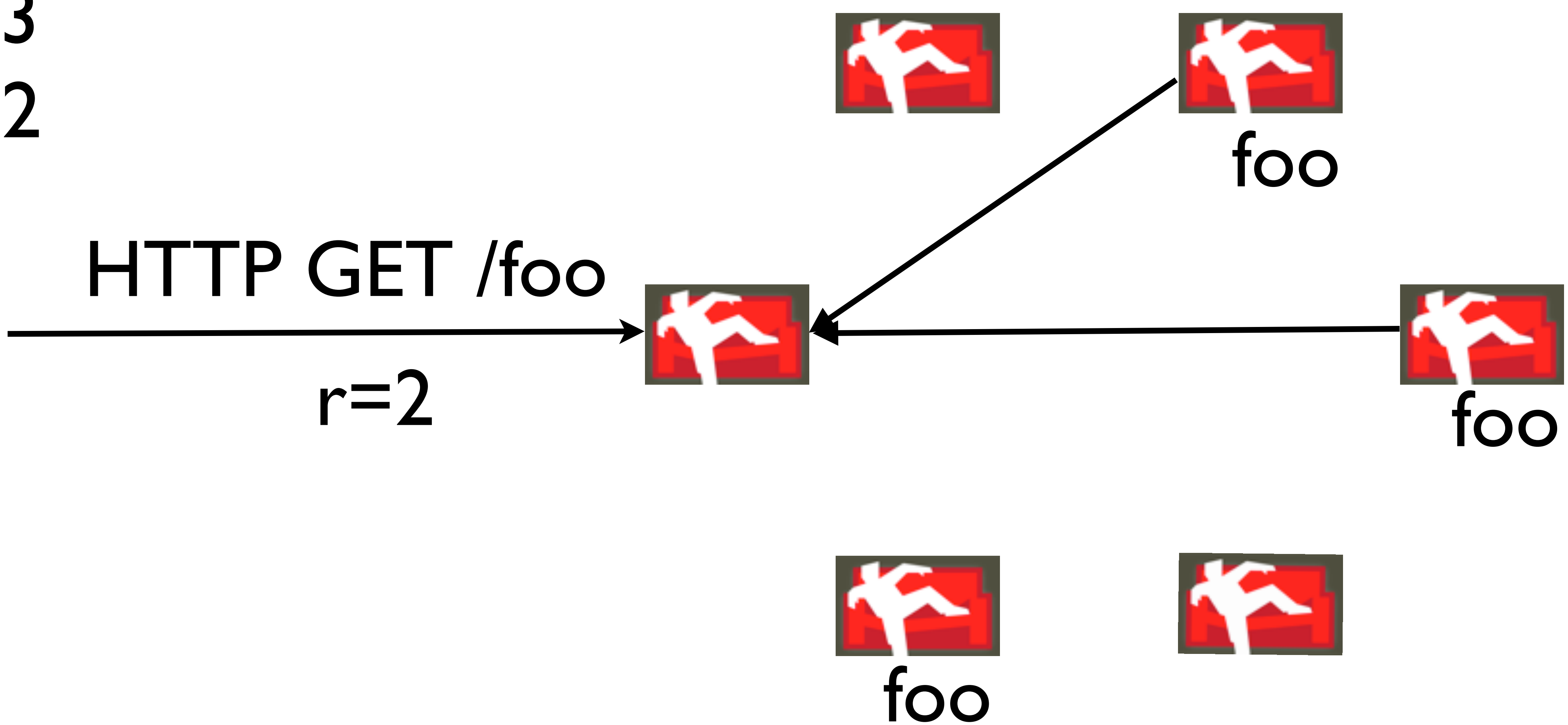


BigCouch

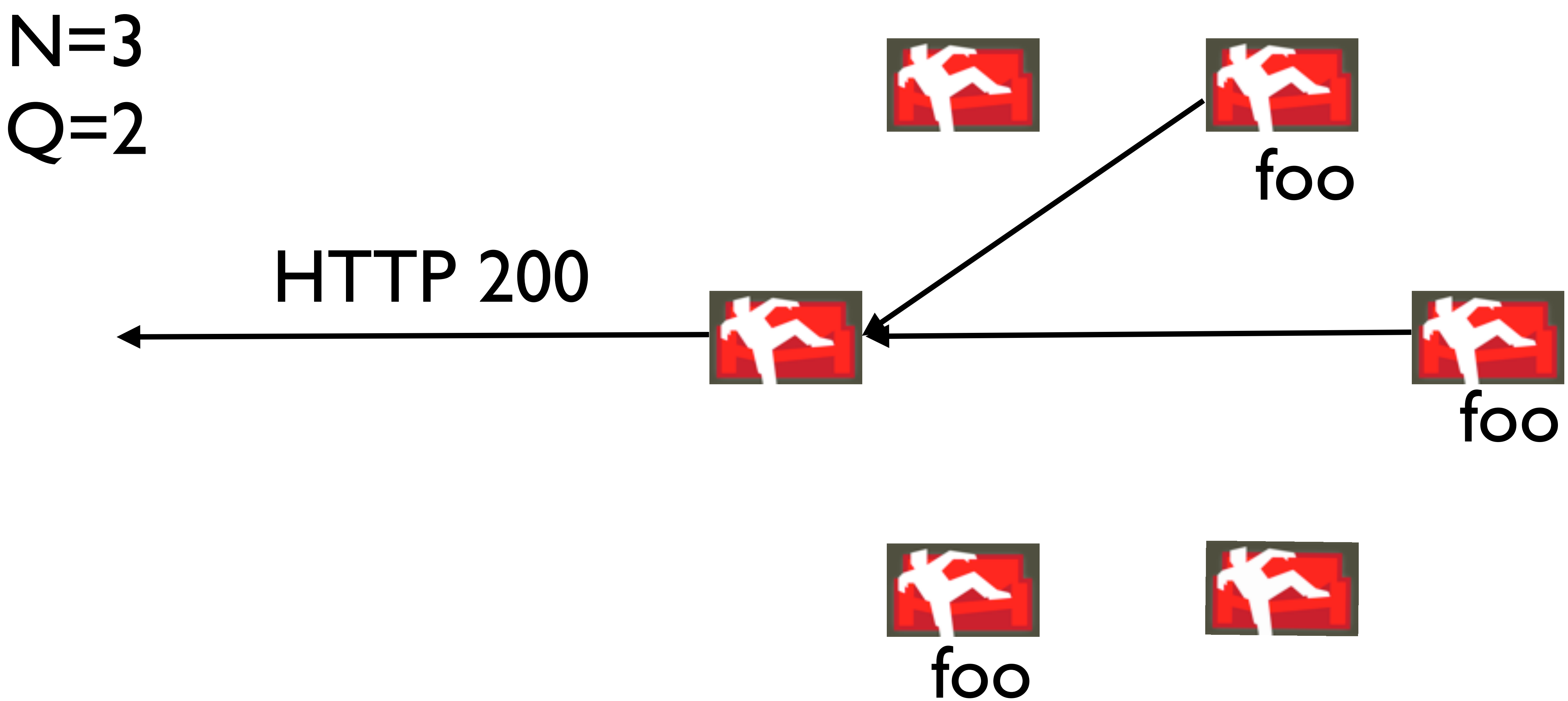


BigCouch

N=3
Q=2

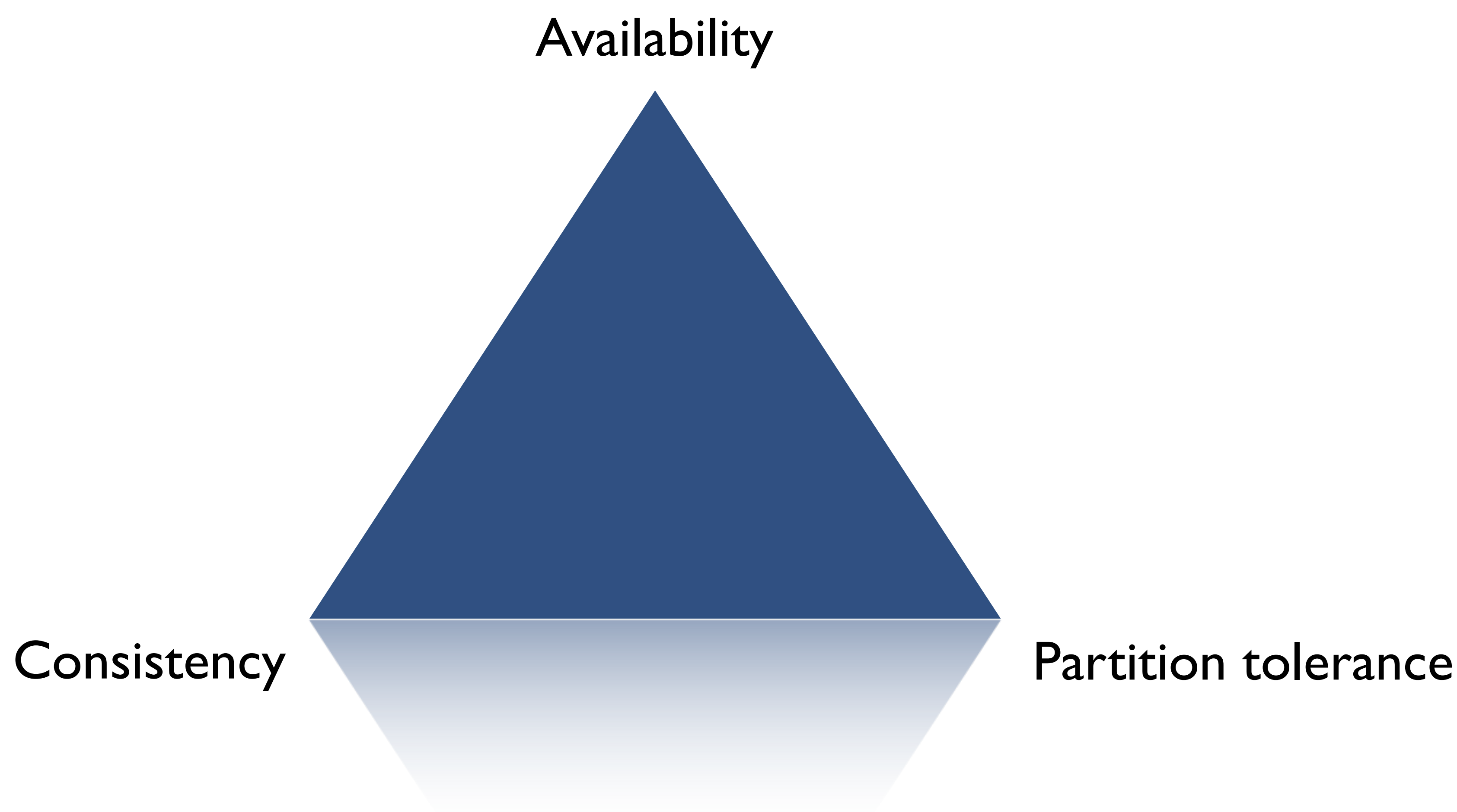


BigCouch

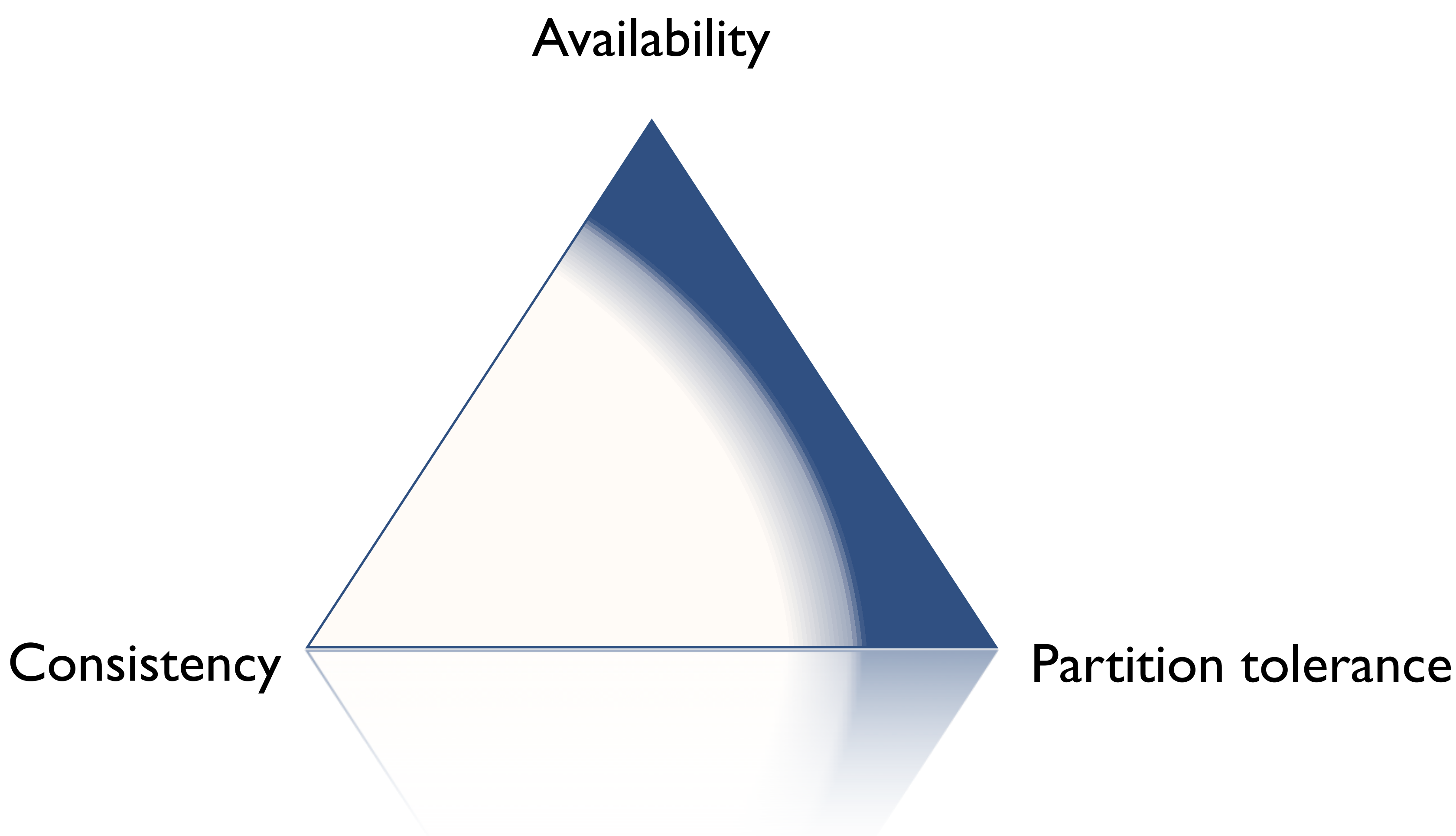


CAP

“choose two”



CAP
“choose two”

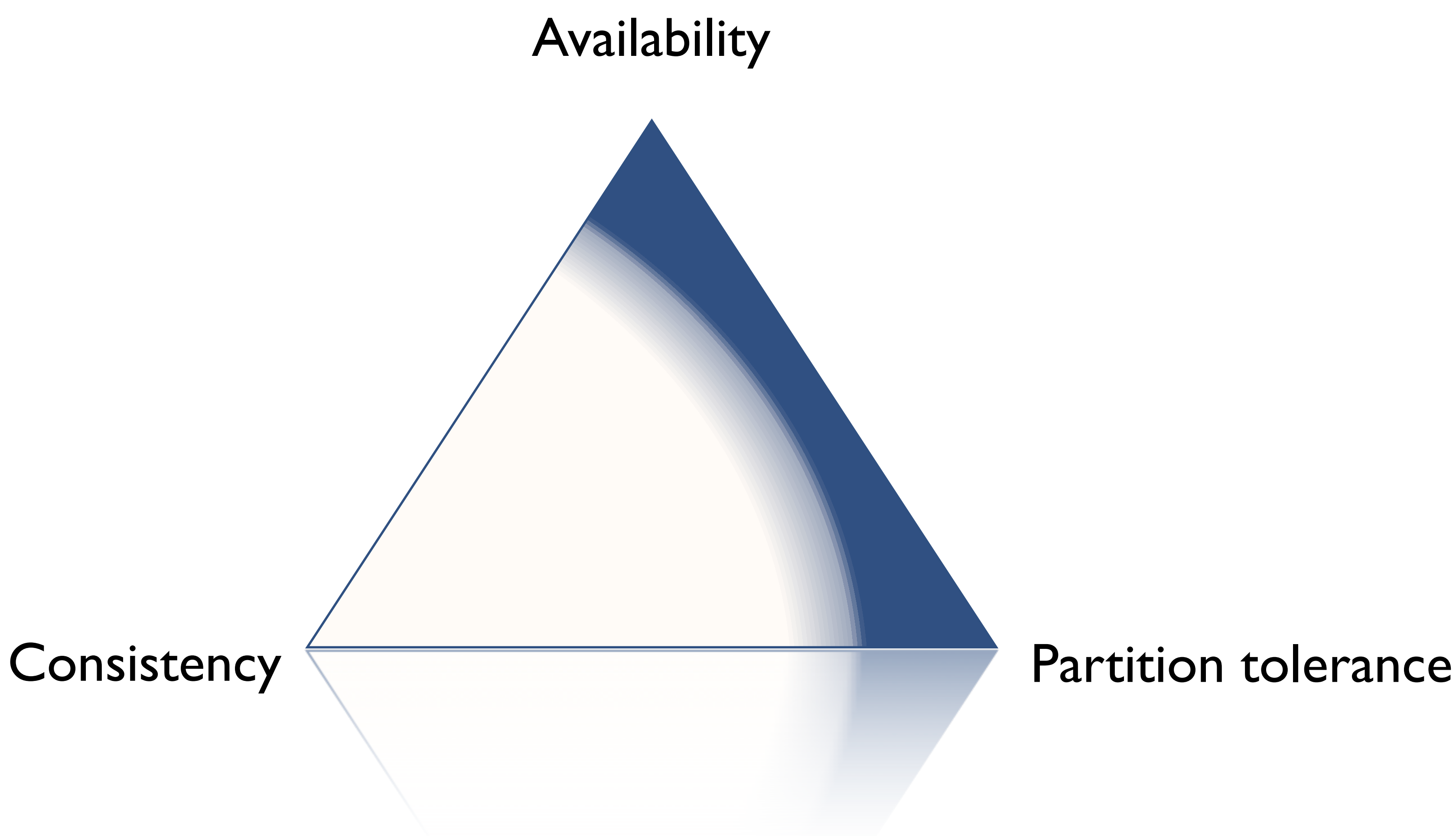


However...

BigCouch and CAP

BigCouch lets us trade between availability and consistency
at different points in our application

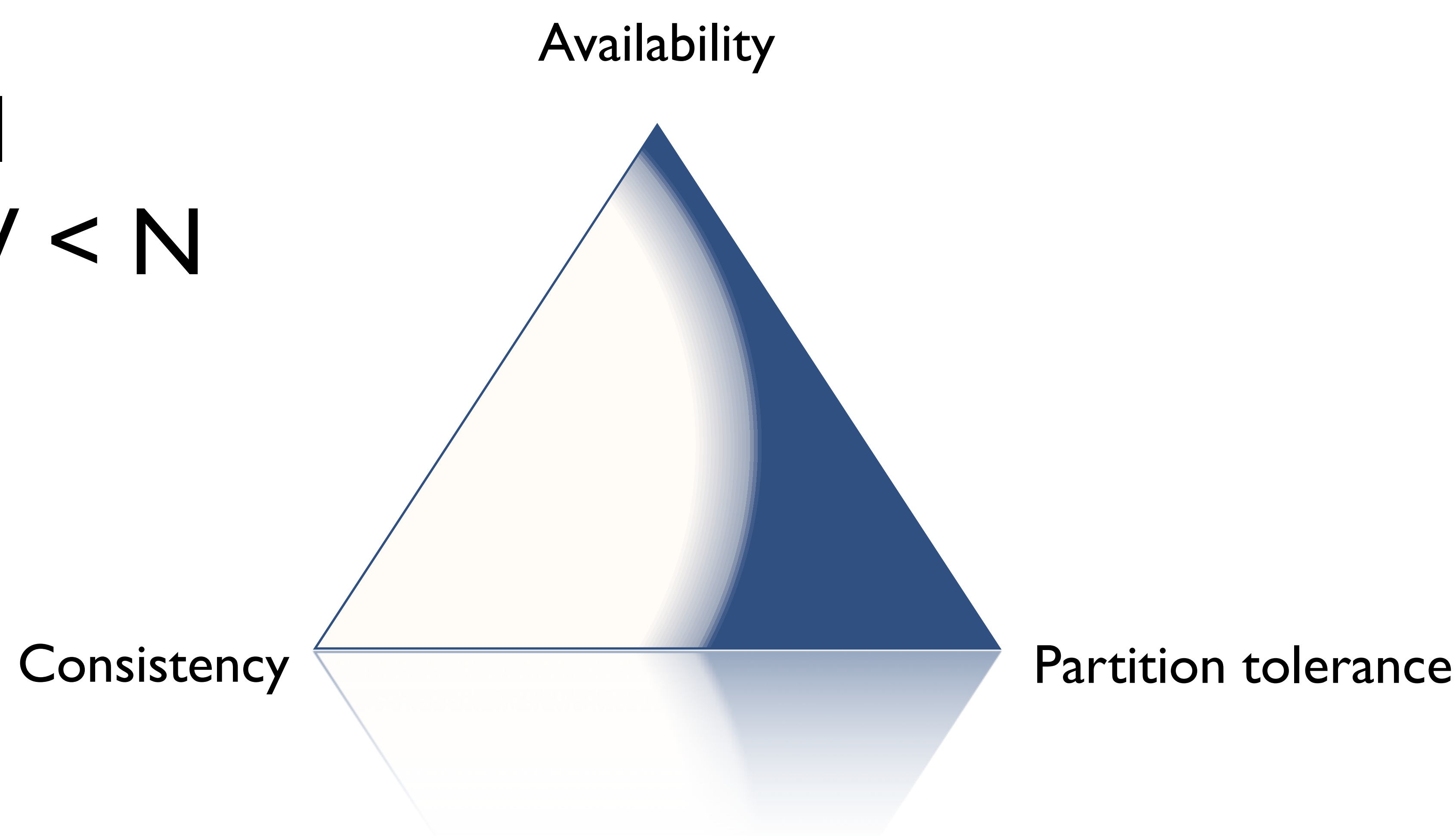
W=1
R=1



BigCouch and CAP

BigCouch lets us trade between availability and consistency
at different points in our application

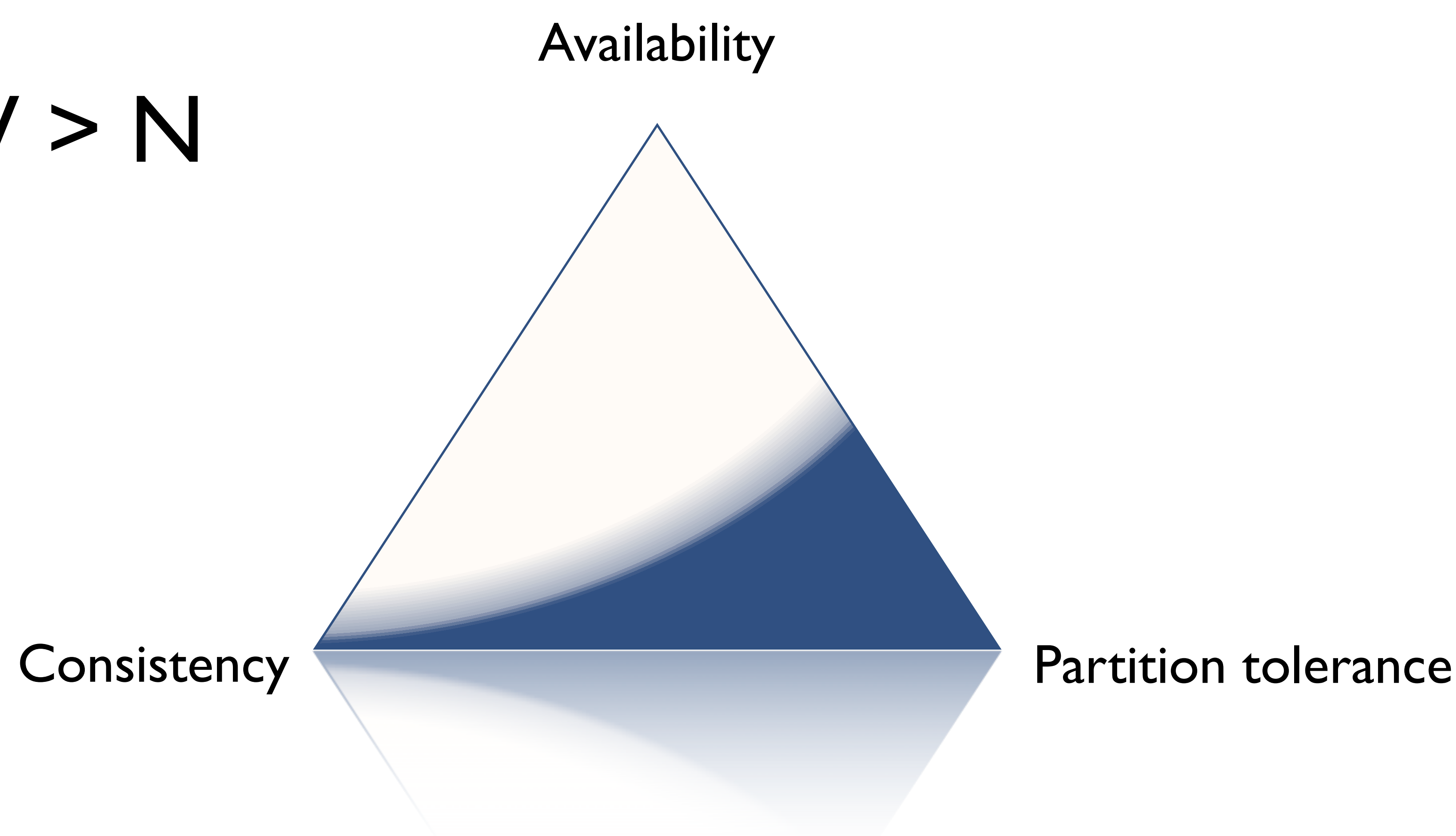
$R > 1$
$W > 1$
$R + W < N$



BigCouch and CAP

BigCouch lets us trade between availability and consistency
at different points in our application

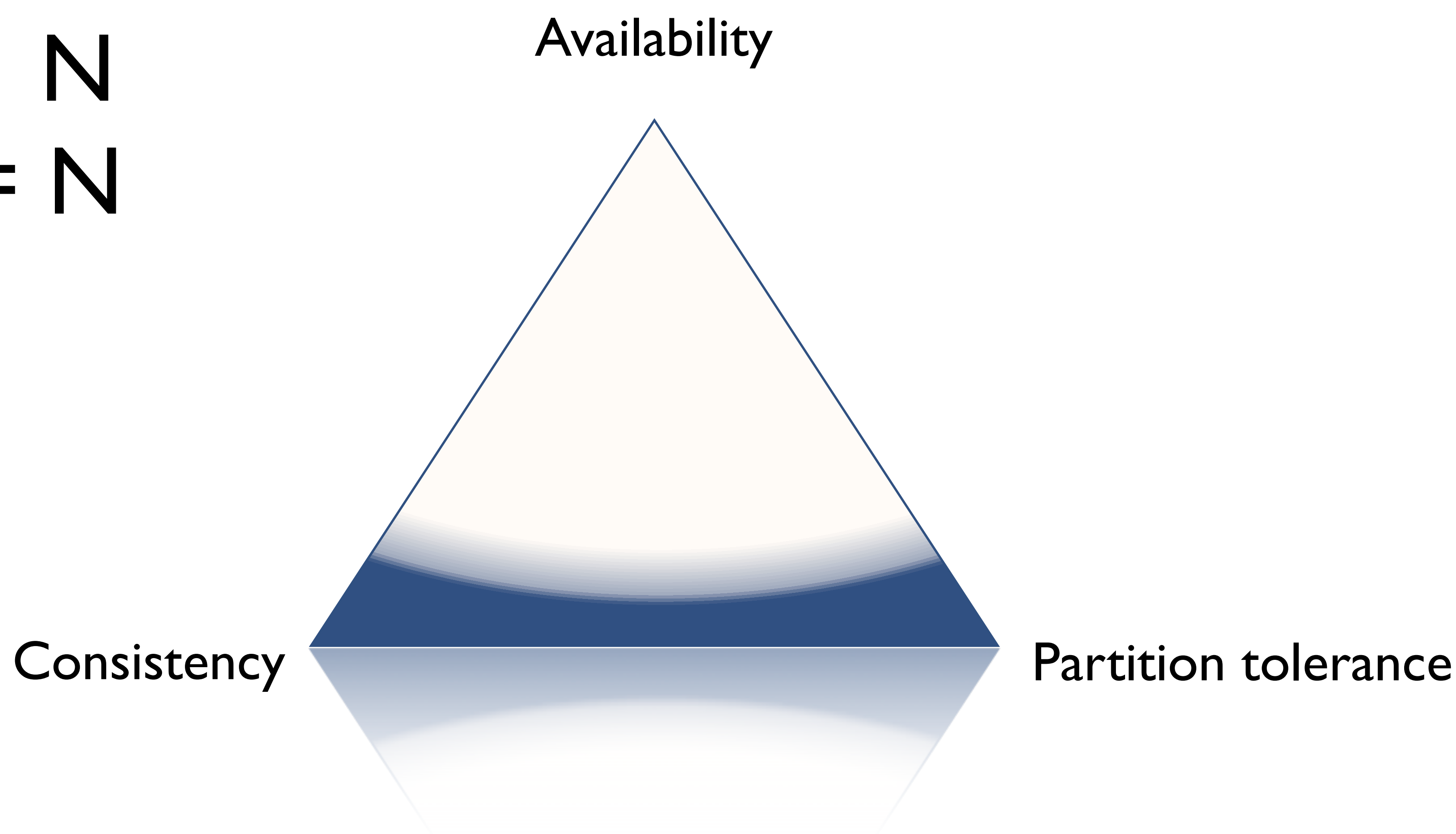
$R + W > N$



BigCouch and CAP

BigCouch lets us trade between availability and consistency at different points in our application

R = N
W = N



Things to be aware of...

- Eventual consistency is ***eventually*** consistent
- _bulk_docs does not support quorum values
- _changes feed has no global ordering
- All views have an implicit r=1

GeoCouch

- Geo-indexing
- Bounding box search
- “Show all slopes within this area that are at risk of failing”

BigCouch+GeoCouch?

Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

Requirements Summary

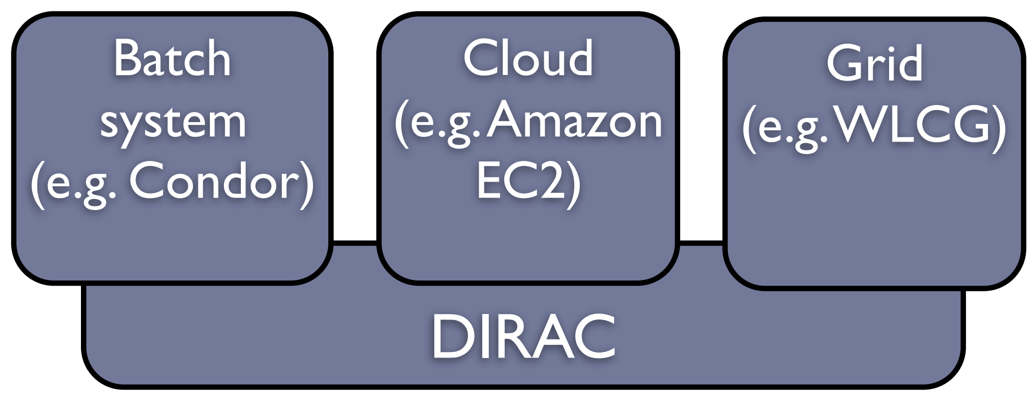
- Enable running 1000s of simulations in parallel
- Long term store of simulation results
- Be able to query simulation results
- Provide offline access to subsets of data

13 12 11 10 9 8 7 6 5 4 3 2 1
Foxconn
25 24 23 22 21 20 19 18 17 16 15 14

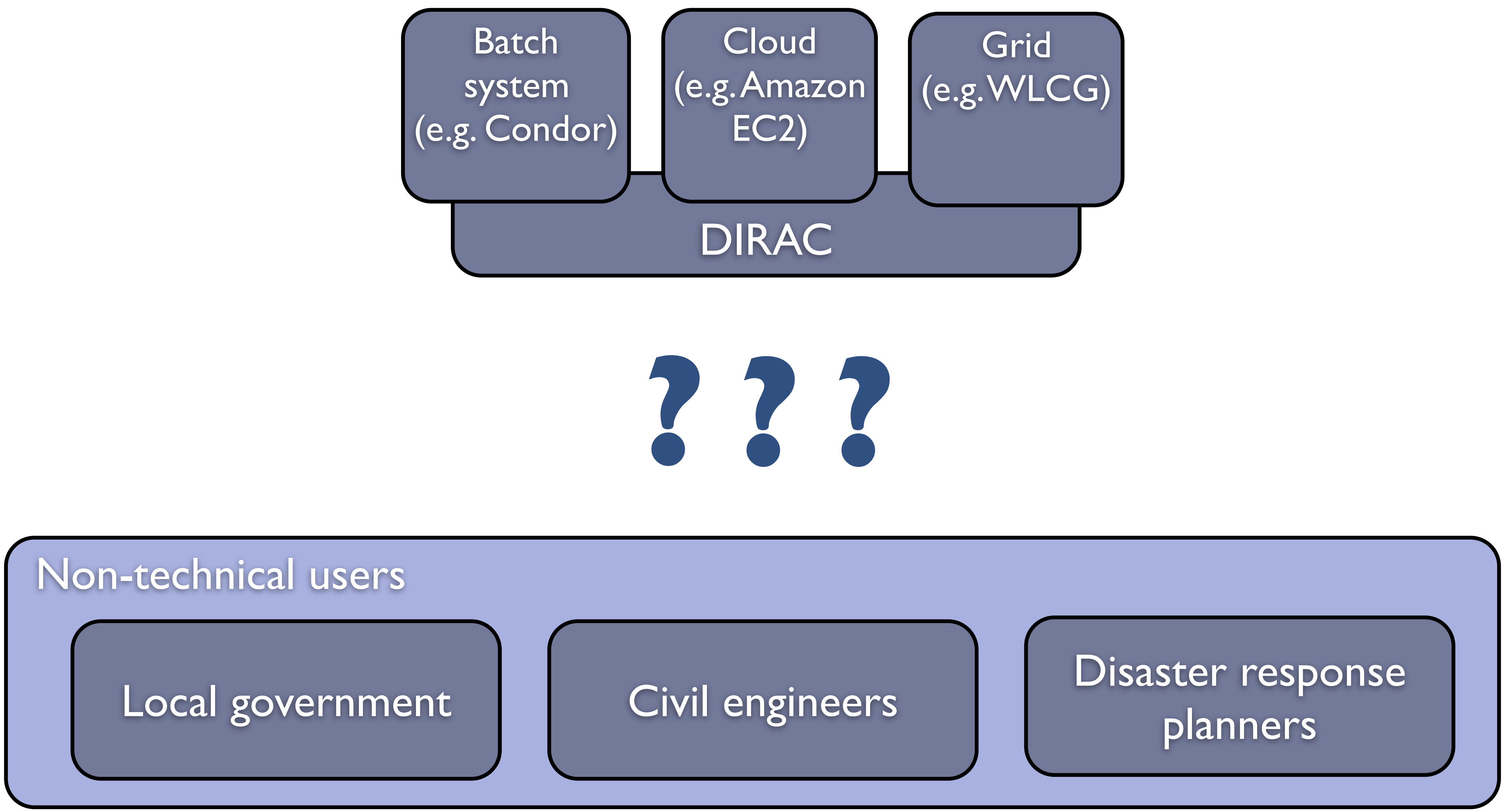
Grid-ified Landslides

- Potential users include the world bank, governments, universities and commercial organisations
- Batch/Grid/Cloud computing all useful in the different contexts of these users

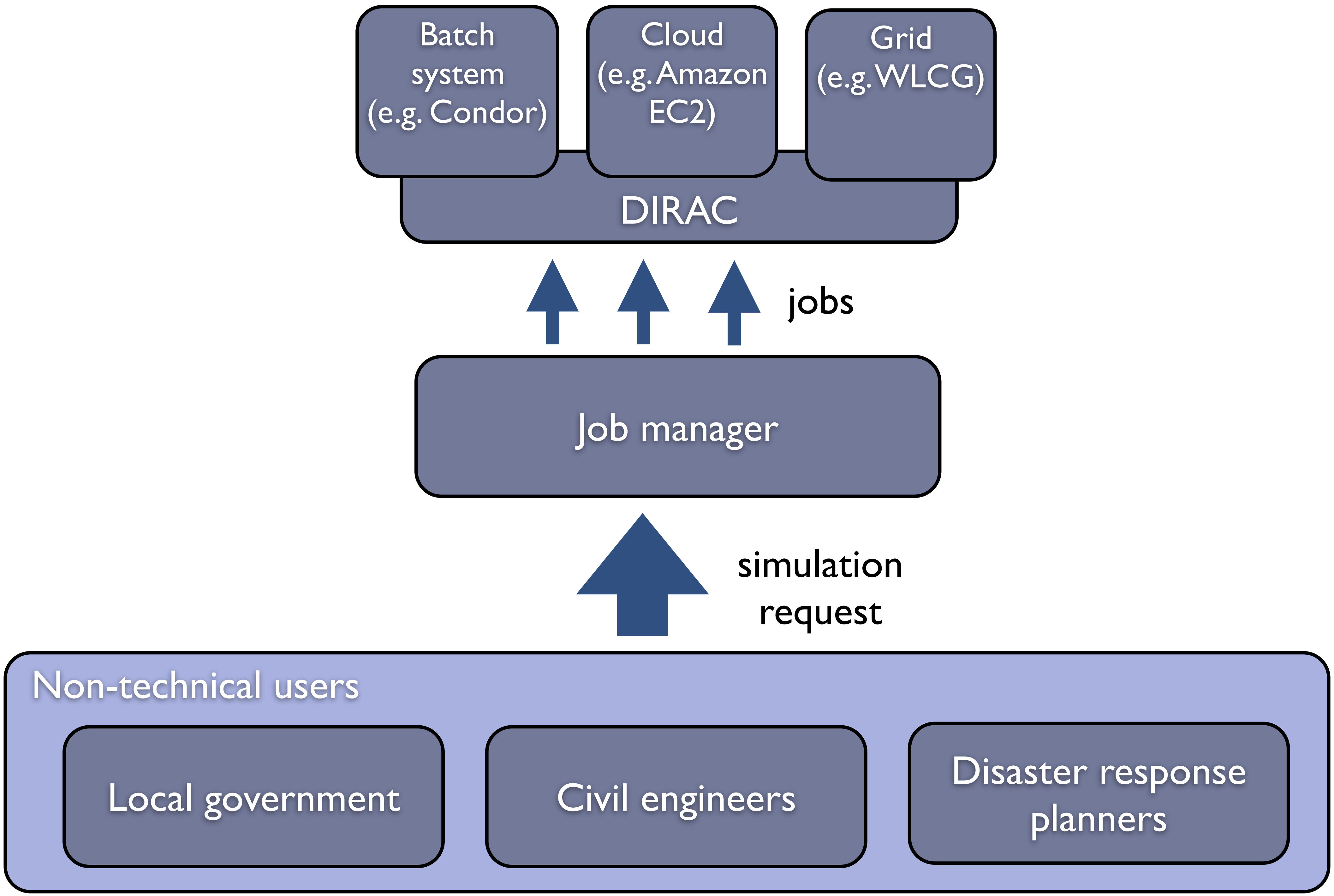
Integration



Integration



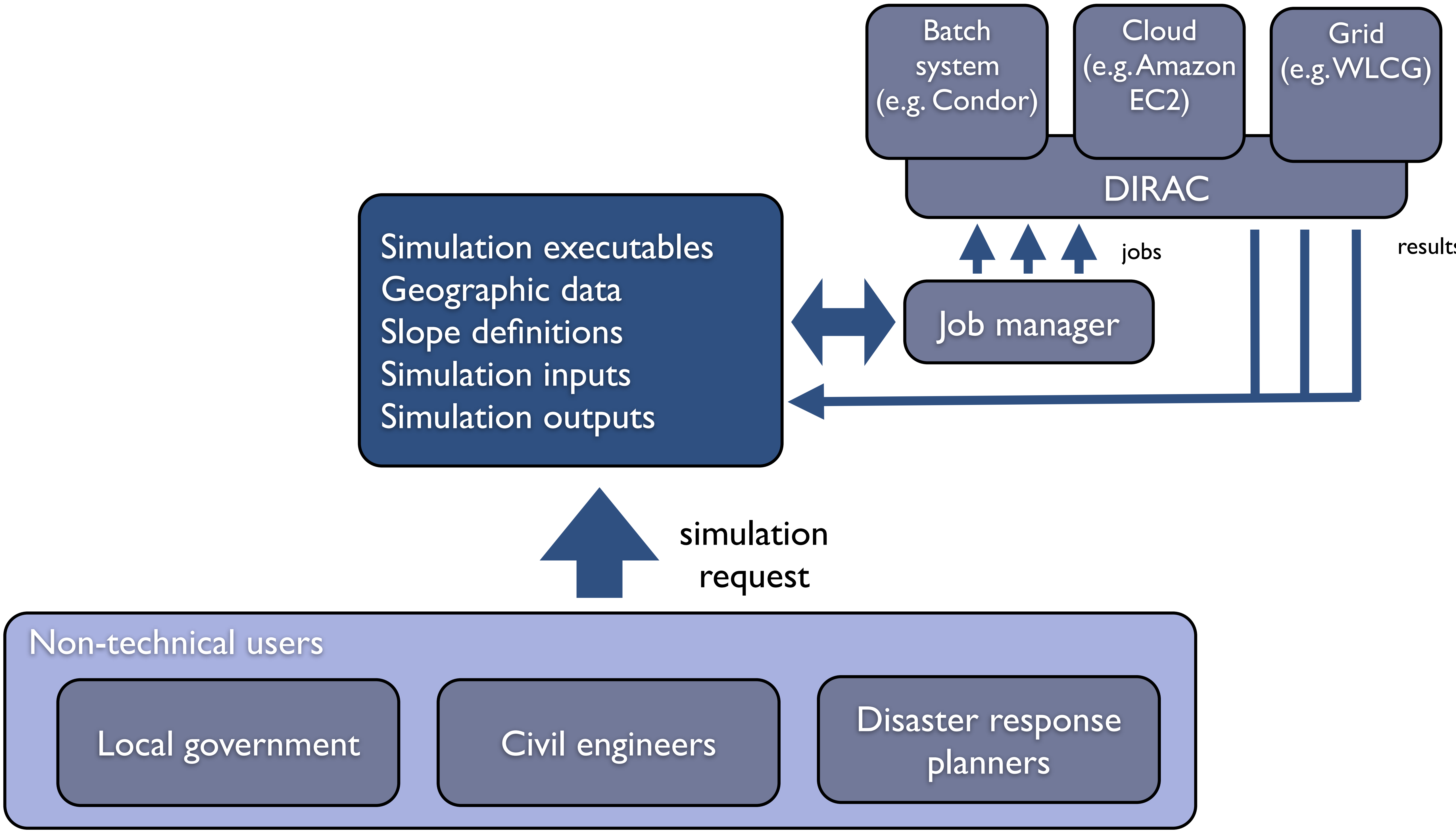
Integration



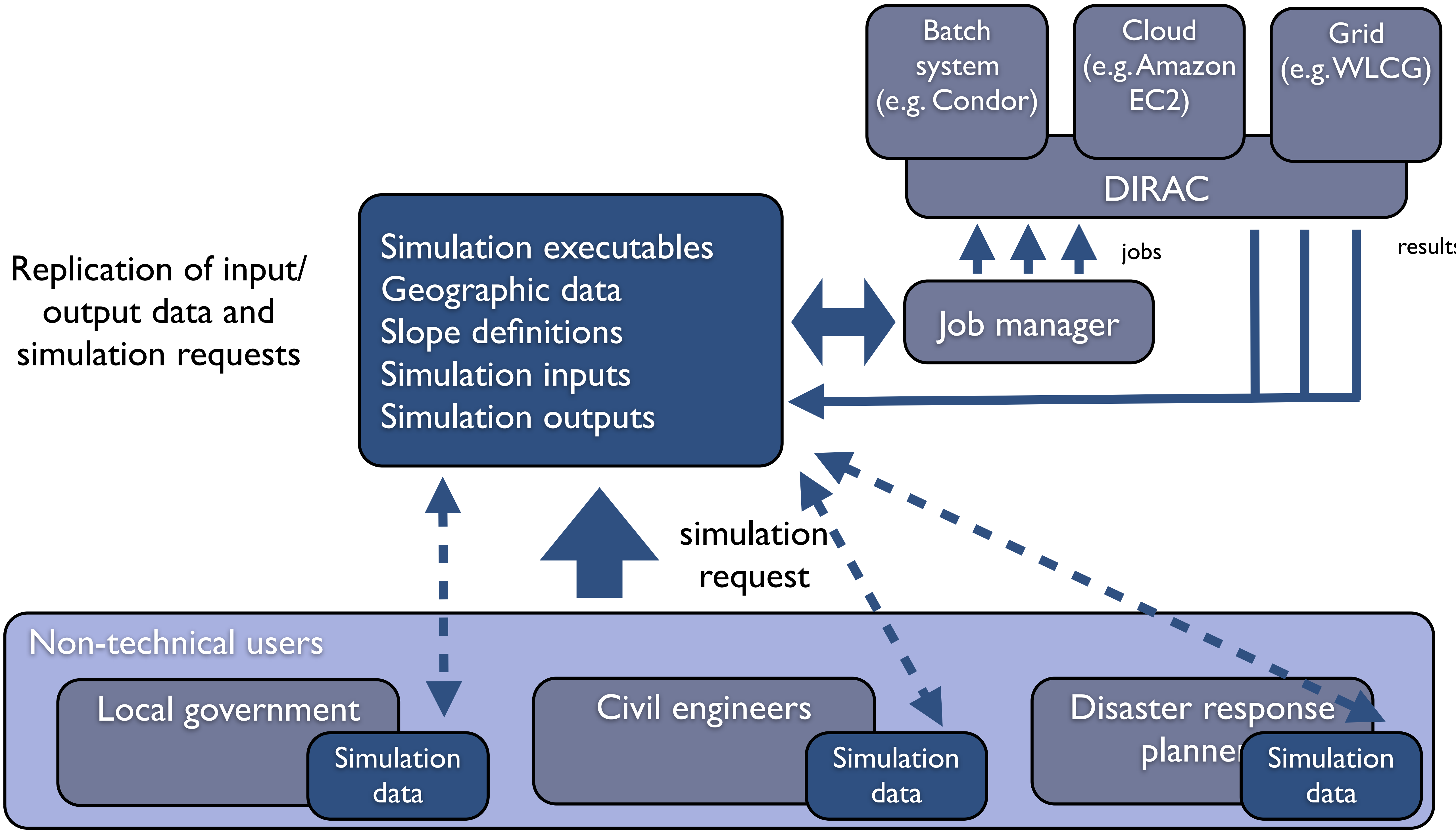
Requirements Summary

- Enable running 1000s of simulations in parallel
- Long term store of simulation results
- Be able to query simulation results
- Provide offline access to subsets of data

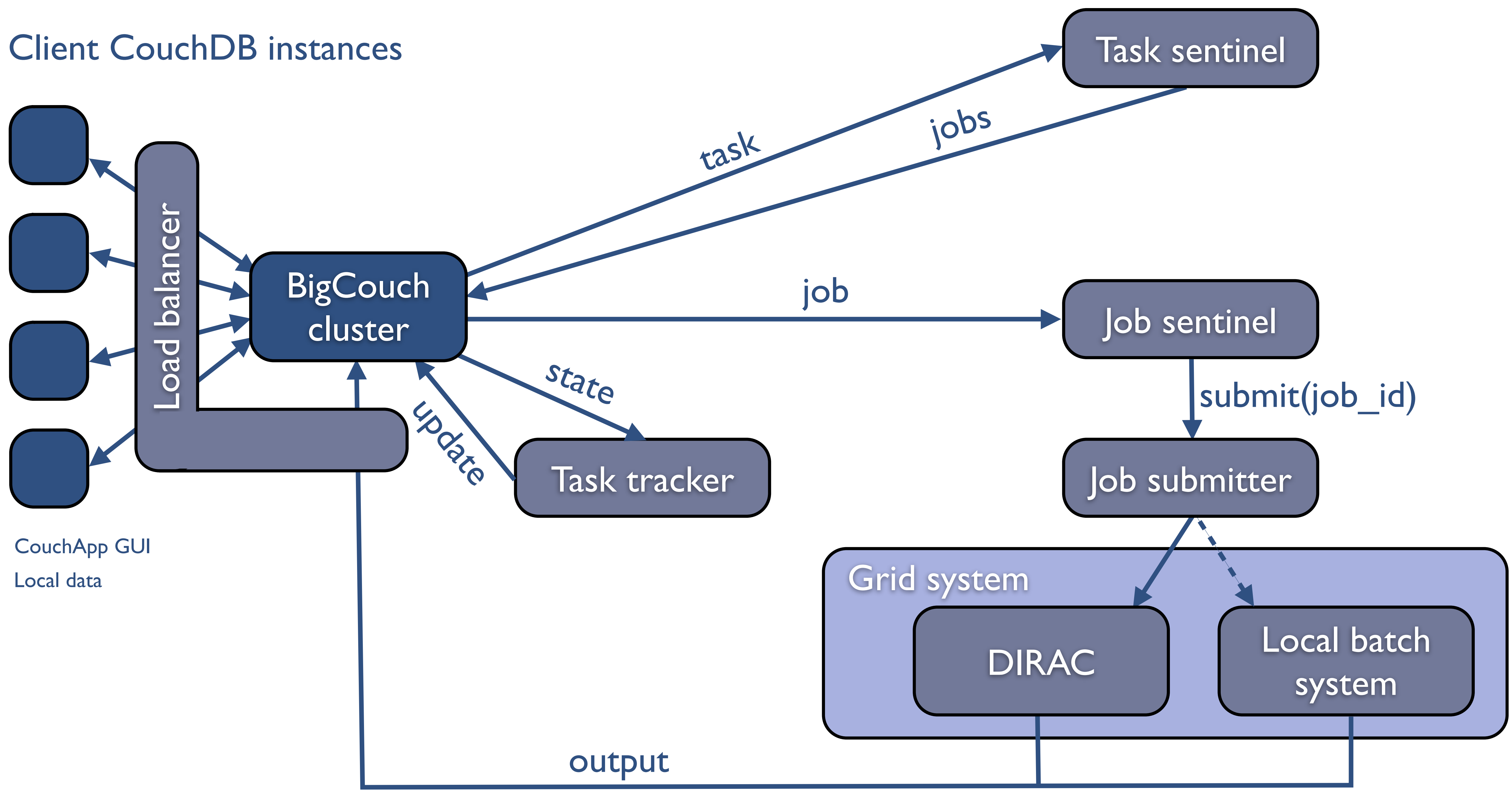
Integration



Integration



Architecture



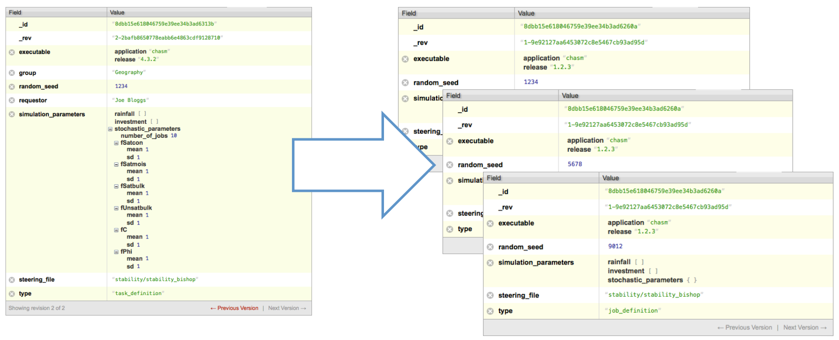
Data Model

Task request

- Specifies software release and input files
- Specifies variable parameters:
  - Stochastic parameters
  - Investment scenarios
  - Rainfall scenarios

Job definition

- Specifies software release and input files
- Specific parameters generated by task sentinel
- For stochastic jobs:
  - Seed for generation of stochastic parameters

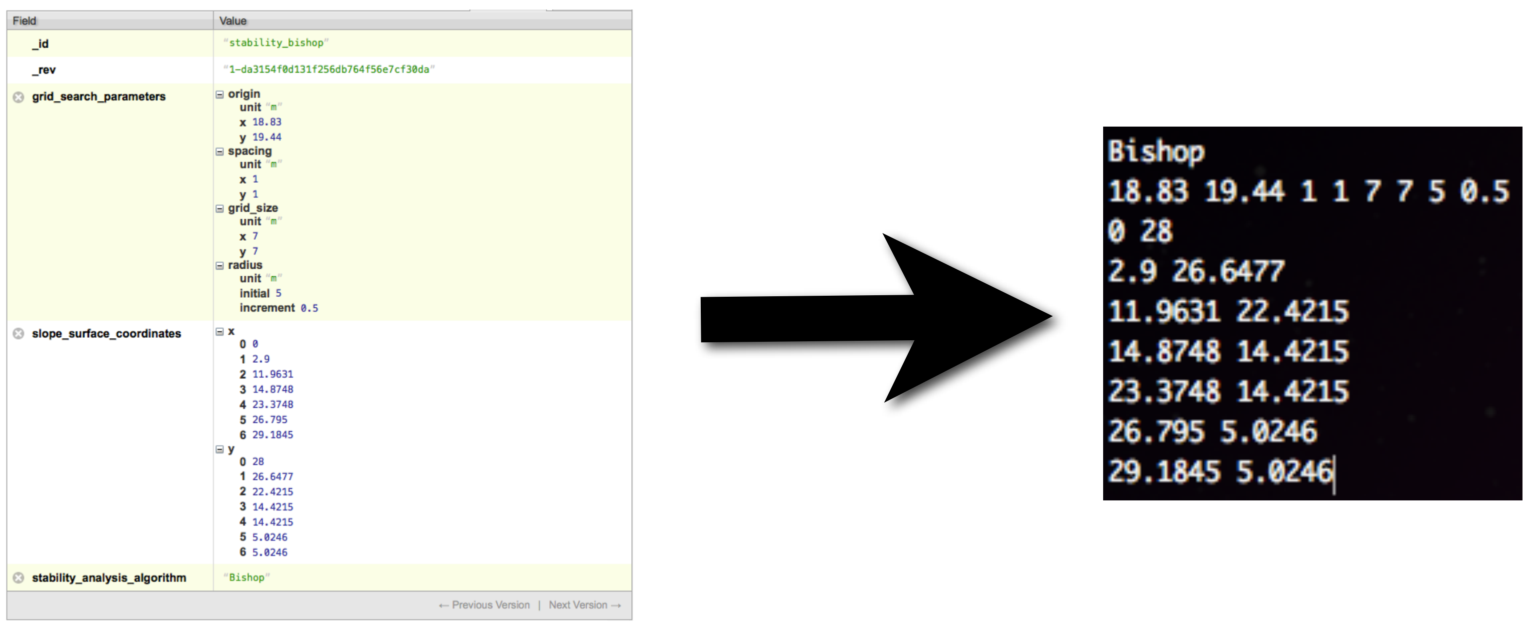


What is a job?

- We submit a job wrapper
  - ~400 lines of python
  - Downloads job definition
  - Downloads executable and inputs
  - Runs executable
  - Parses and uploads output files

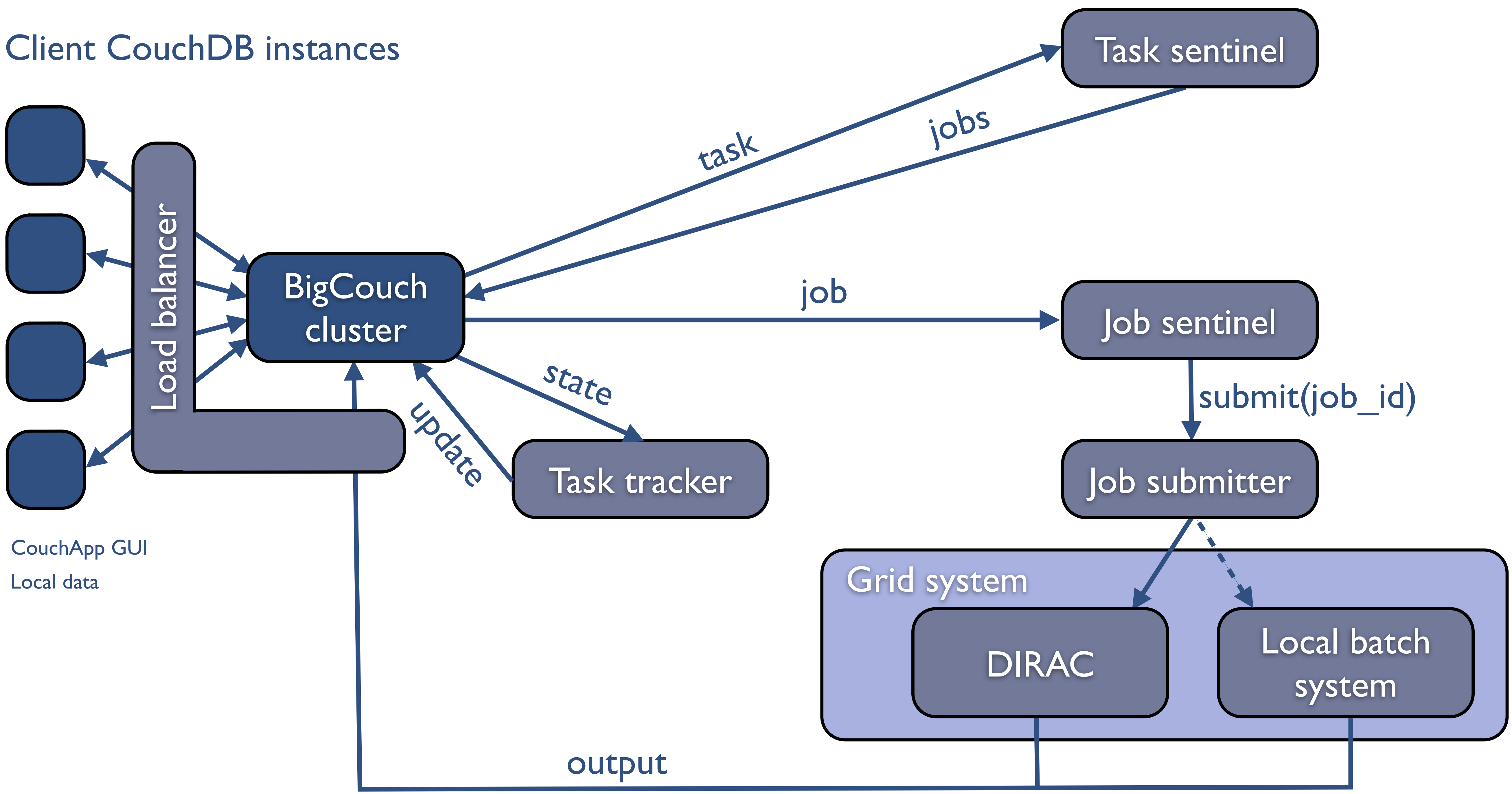
Shows and views

- Show functions generate input files from JSON

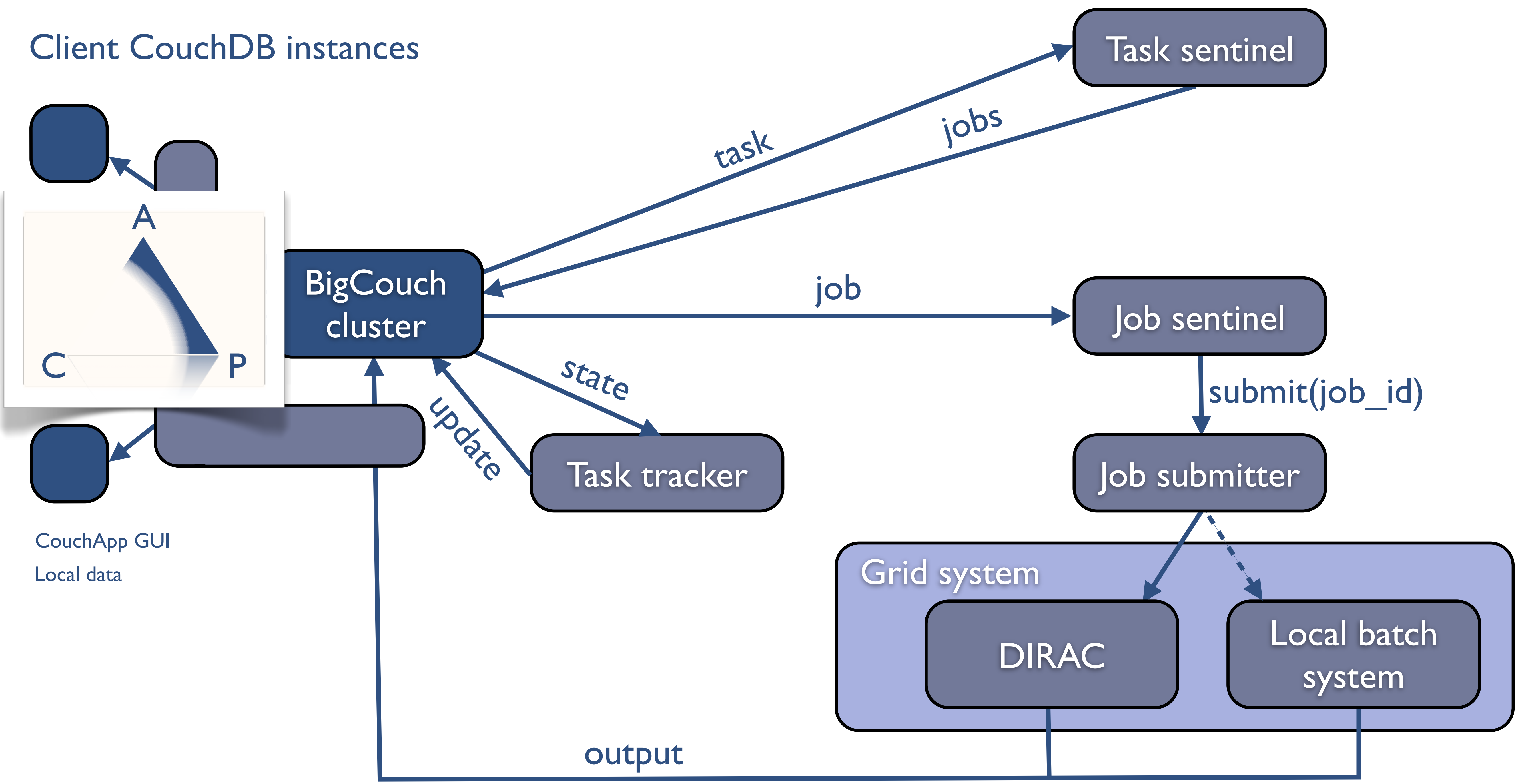


- Views index data by any field:
  - E.g: All results where FOS < 1.2

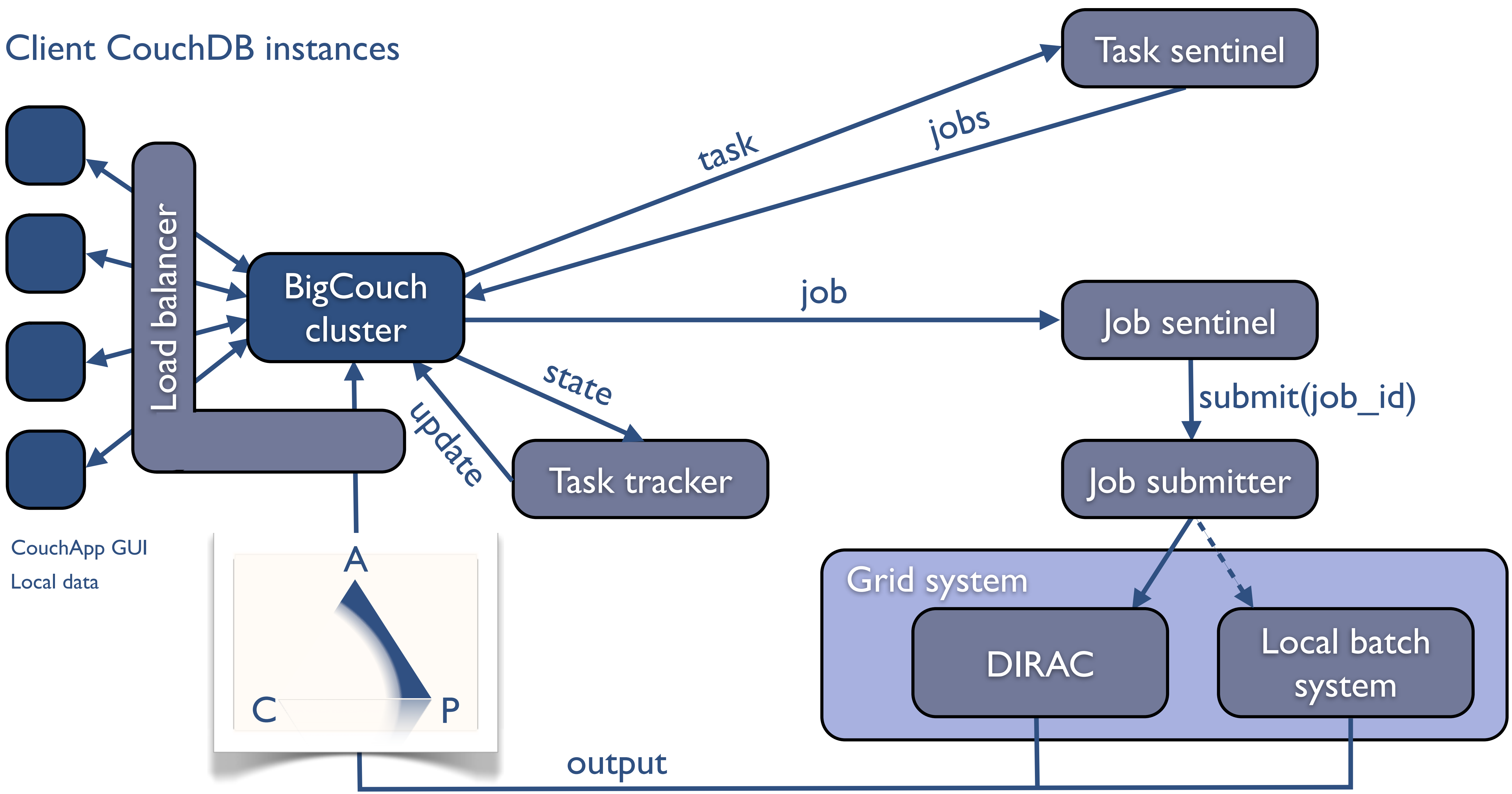
Architecture



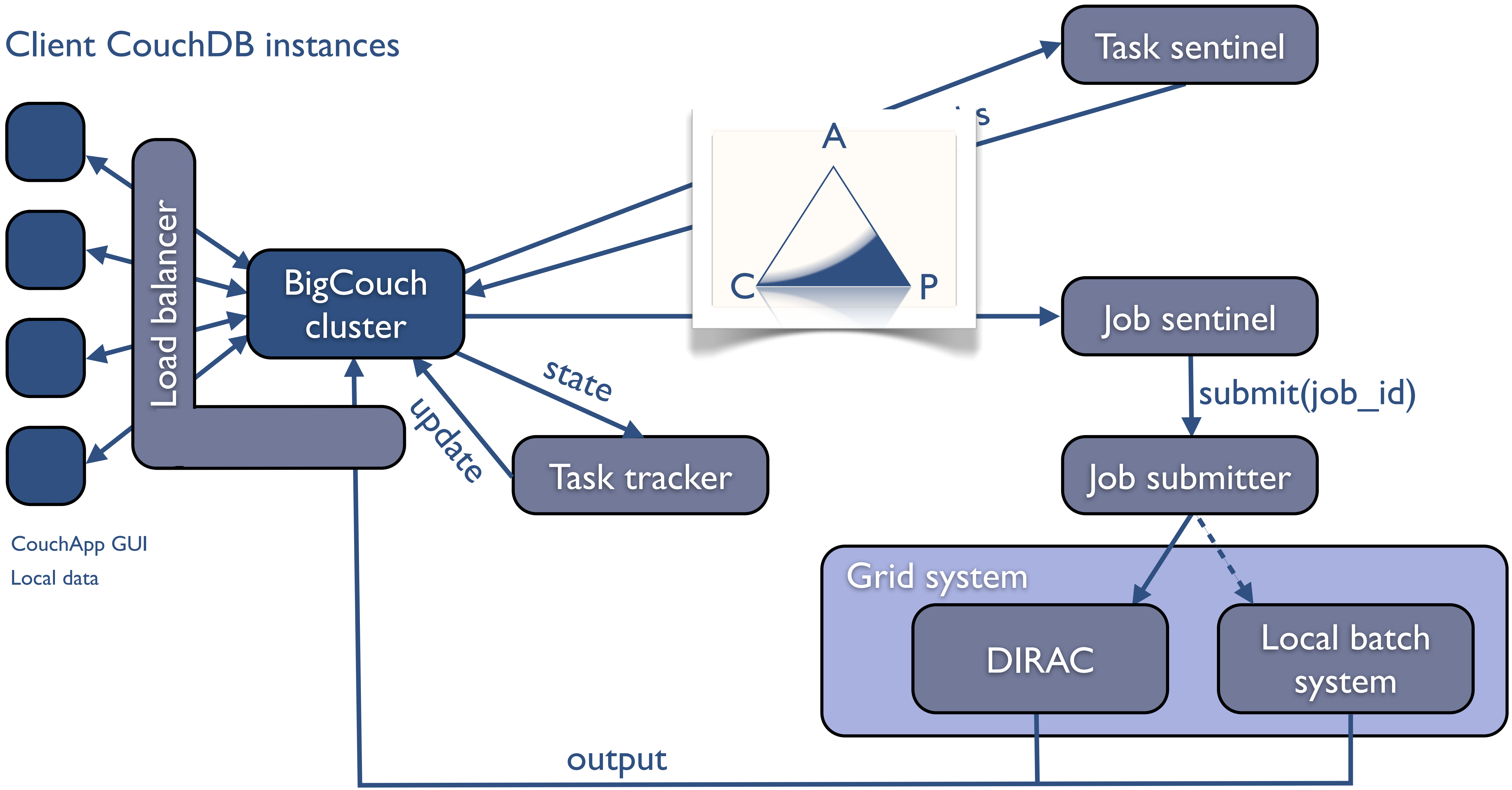
Architecture



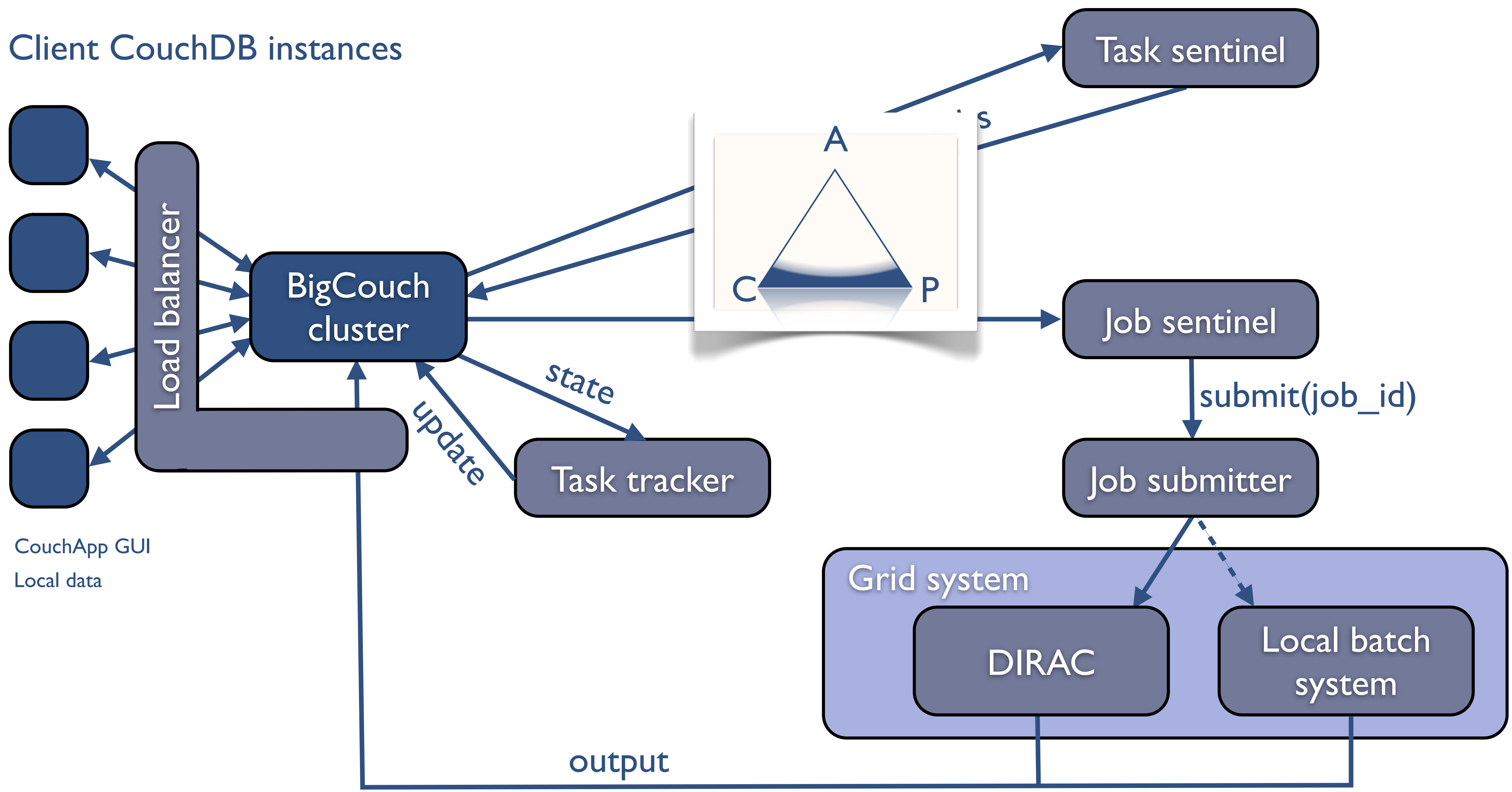
Architecture



Architecture



Architecture



Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

- Producing good generic benchmarks is hard
- These are not generic benchmarks
- Draw conclusions at your own risk

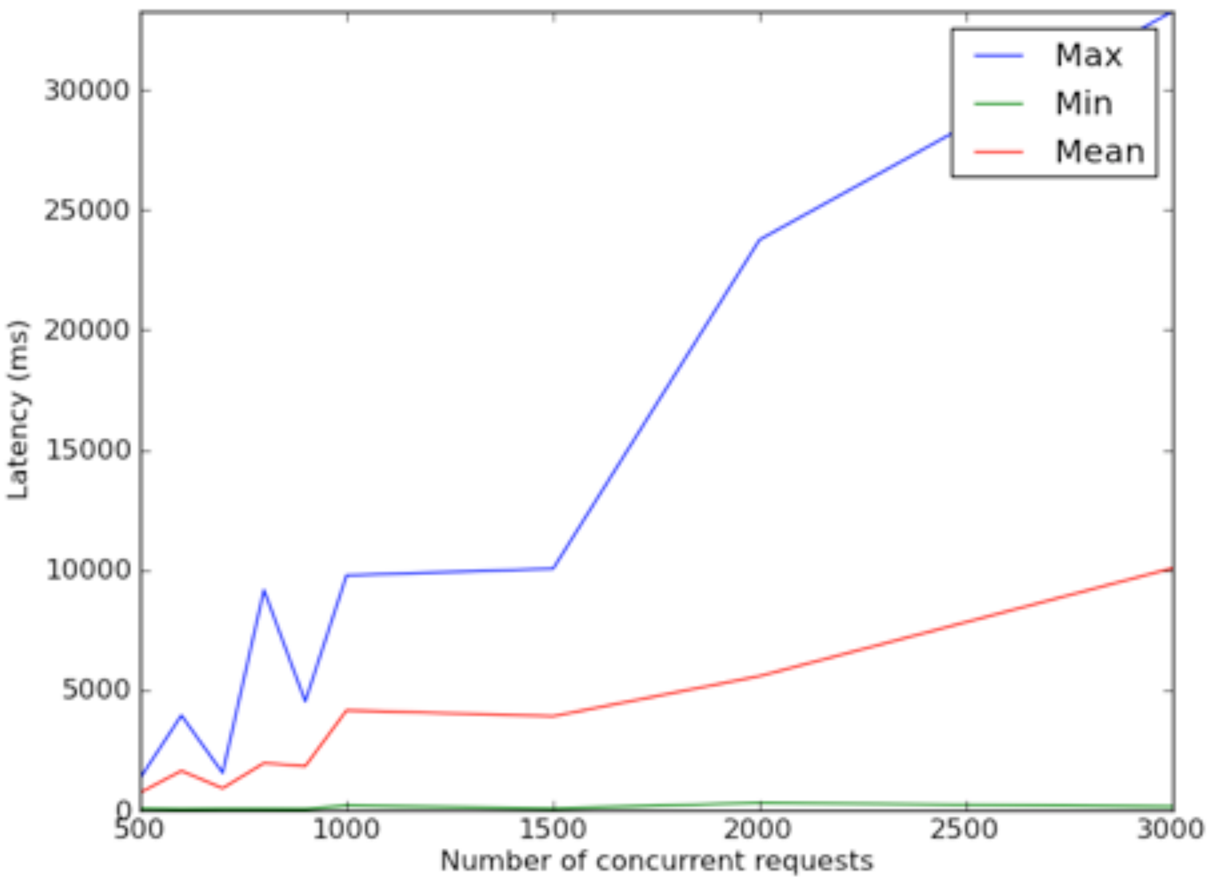
http://www.flickr.com/photos/pictureperfectpose/76138988/sizes/l

Performance testing

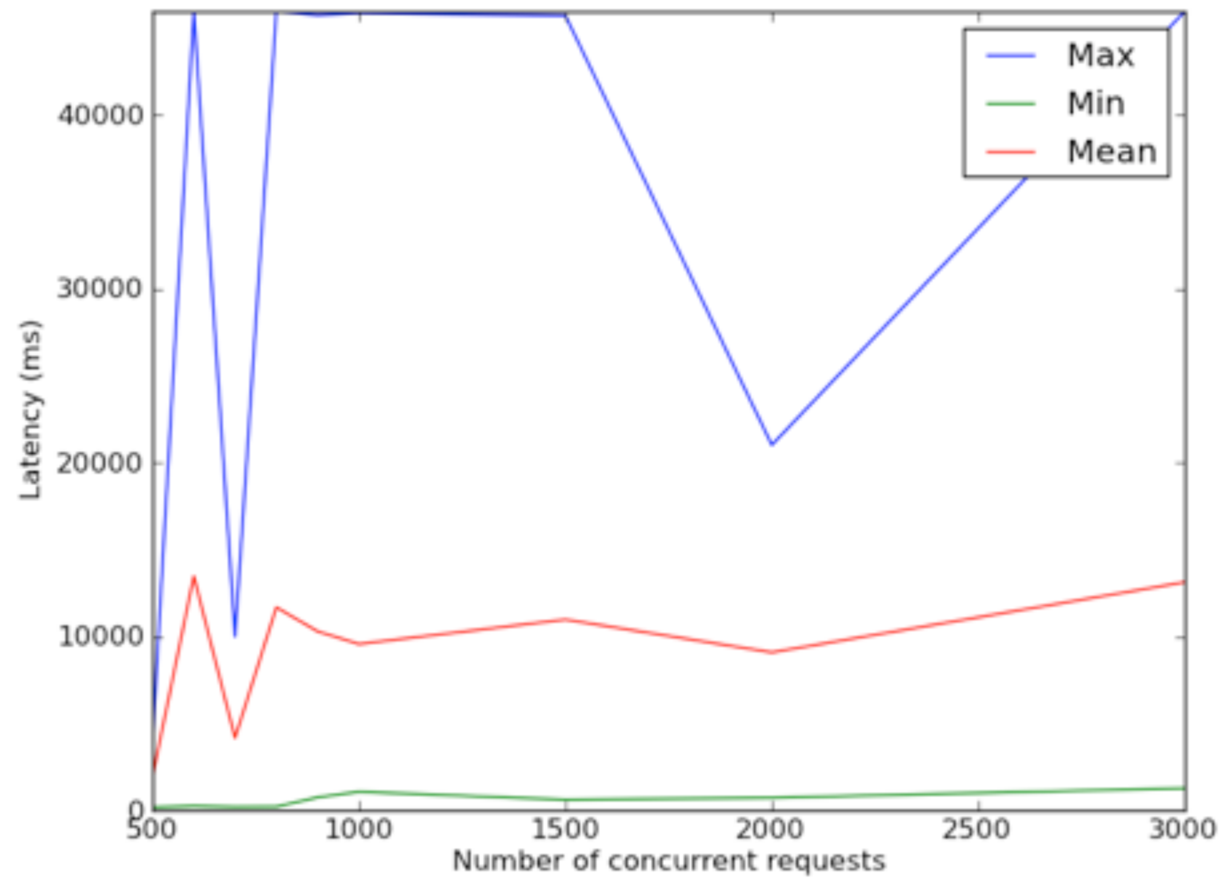
- Preliminary benchmarking with BigCouch
- Crunch points:
  - Job input download (concurrent reads)
  - Job output upload (concurrent writes)
- Test system:
  - Scientific Linux 5
  - 2 x quad-core AMD Opteron 2.4GHz
  - 16GB RAM, Gb ethernet
- Load:
  - 5 clients, each generating up to 1000 requests per second
  - Test document 104KB in size

Vanilla CouchDB

Job input download

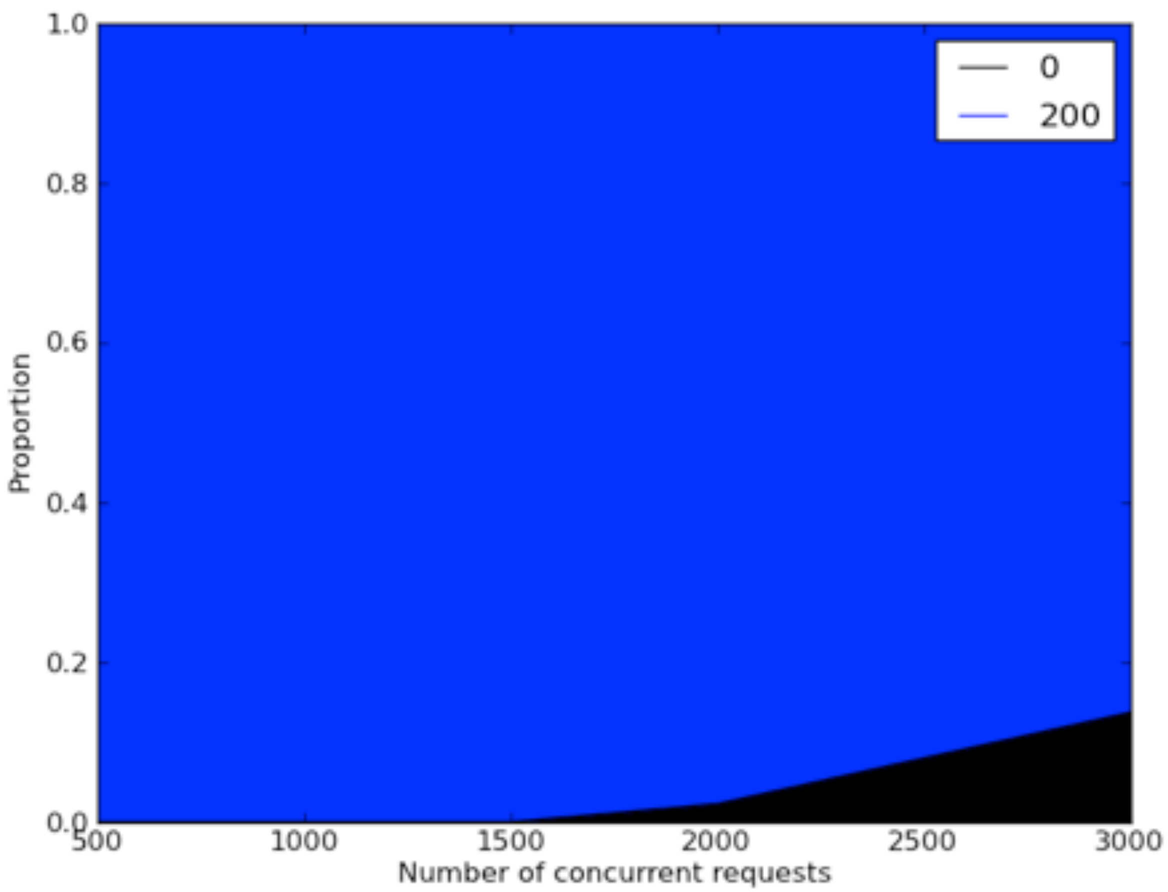


Job output upload

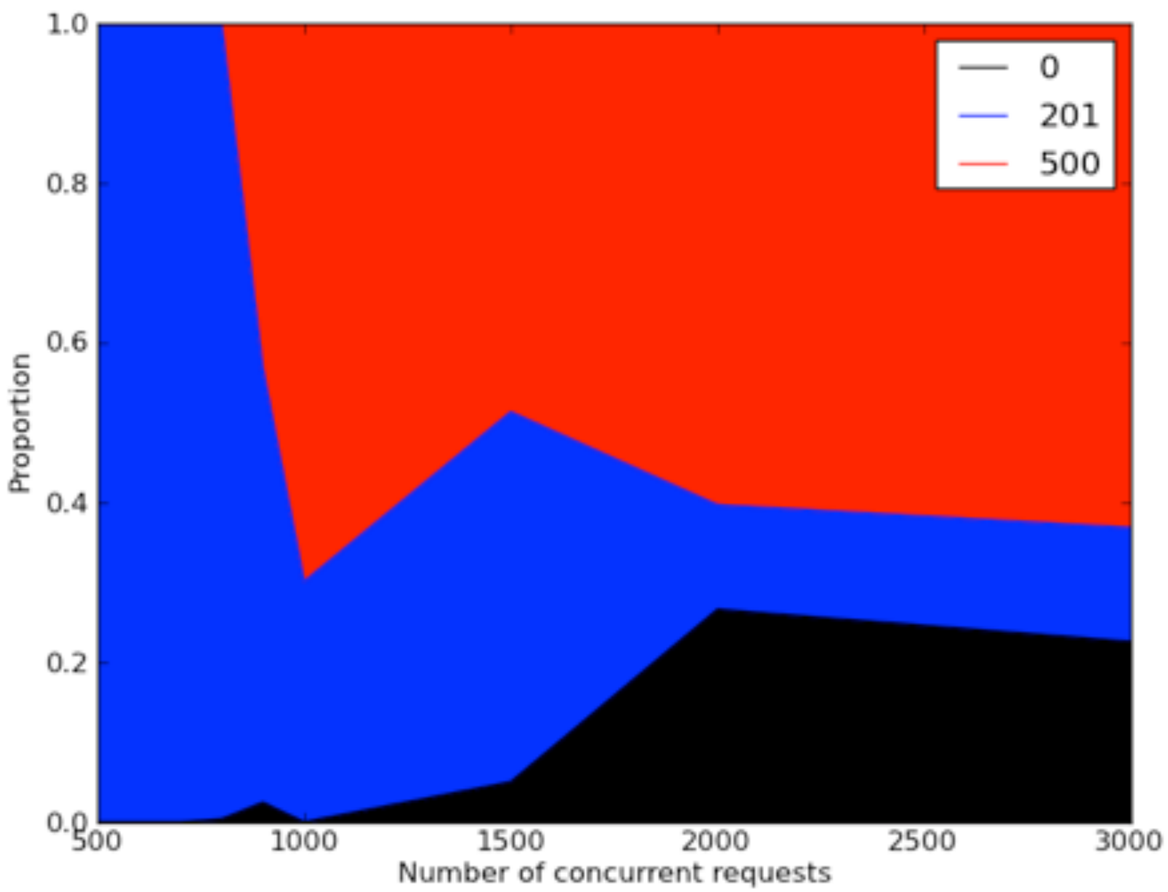


Vanilla CouchDB

Job input download

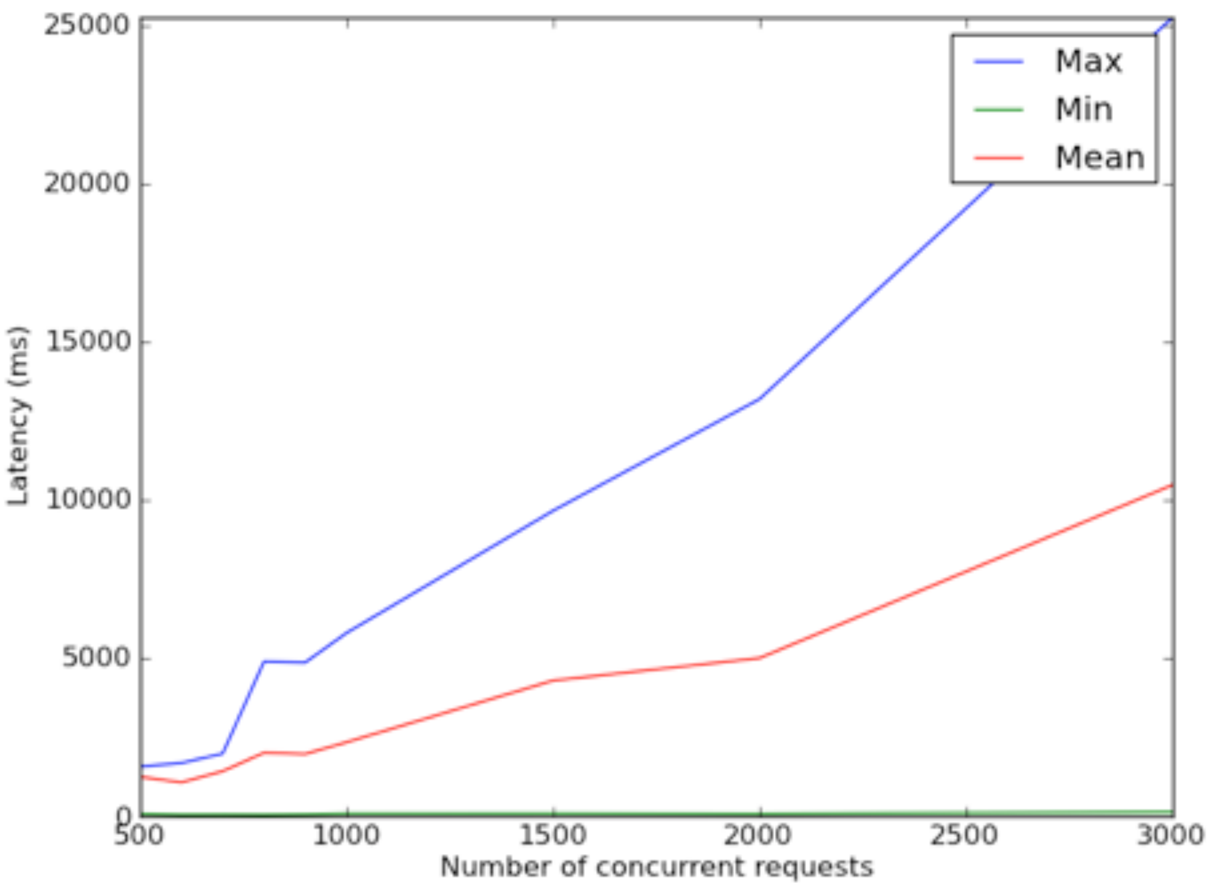


Job output upload

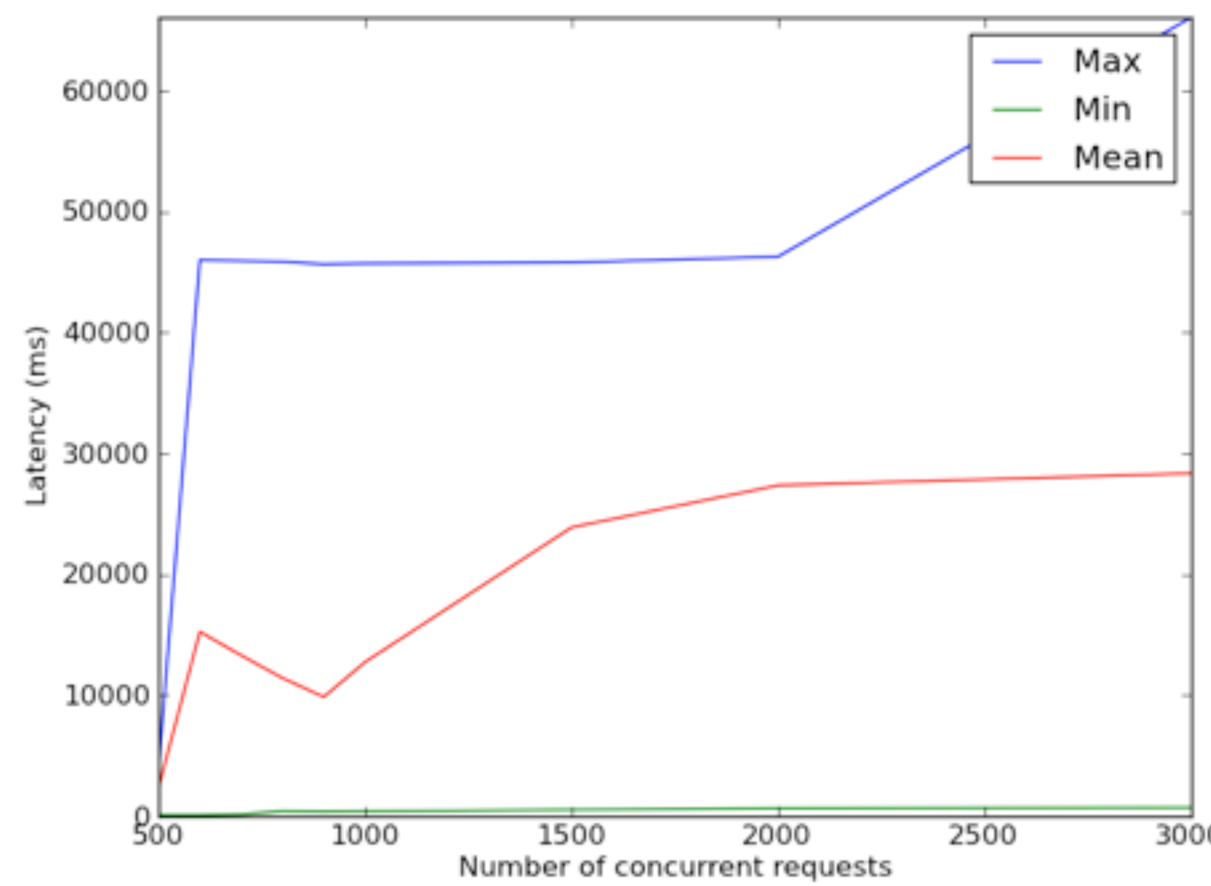


BigCouch 1 node

Job input download

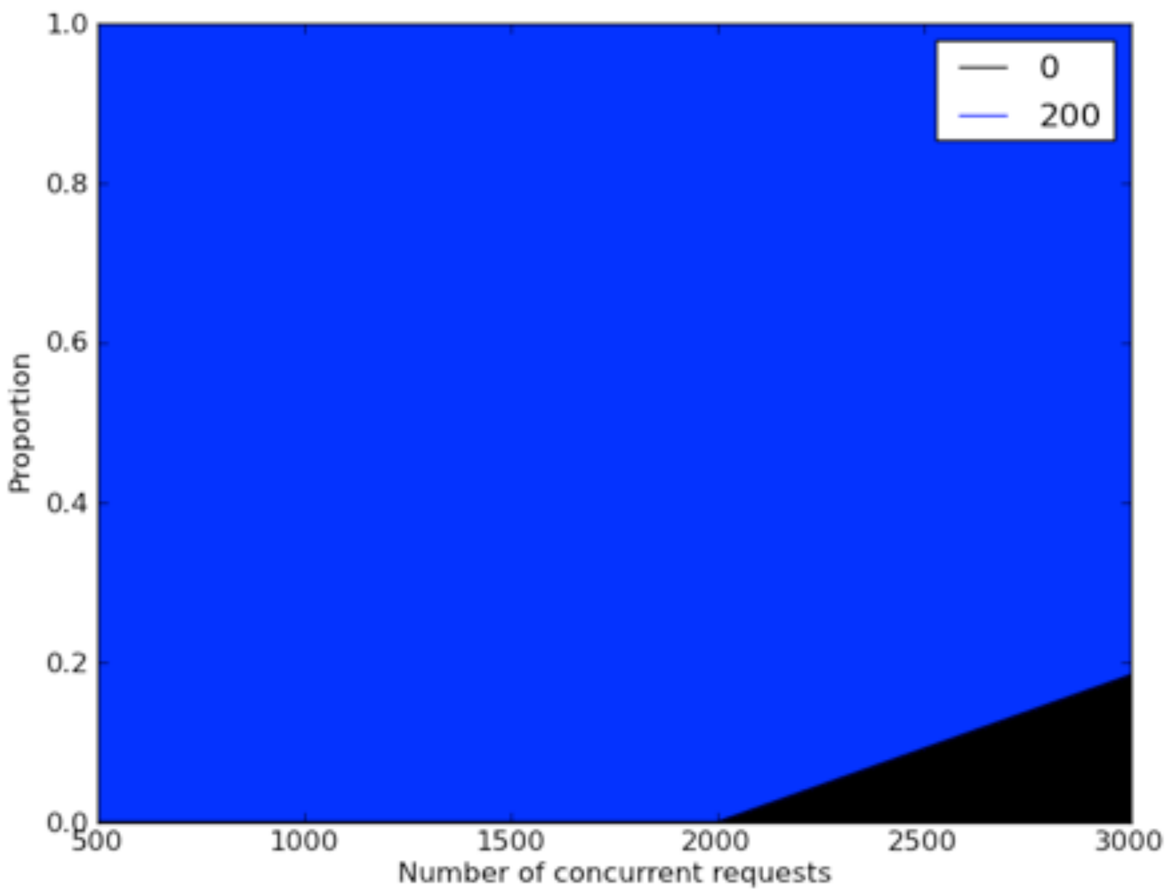


Job output upload

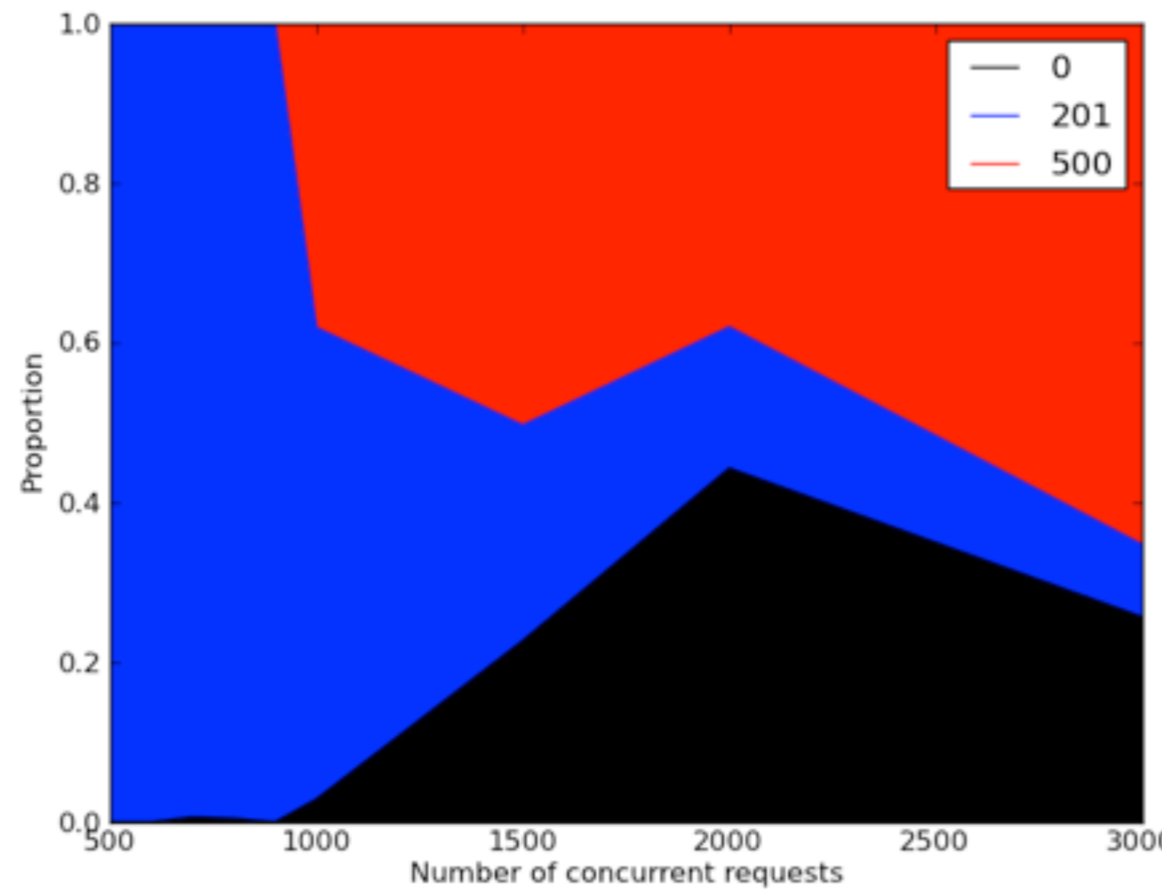


BigCouch 1 node

Job input download

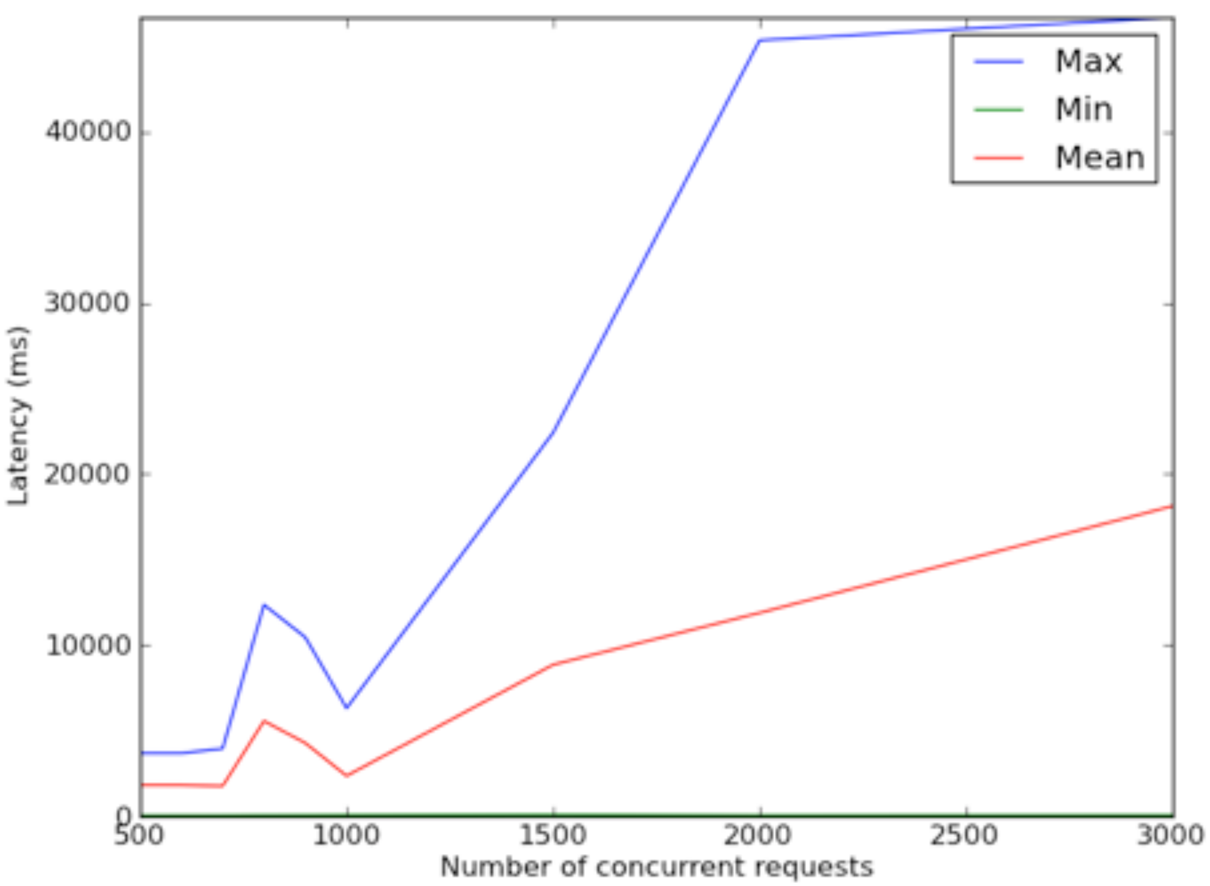


Job output upload

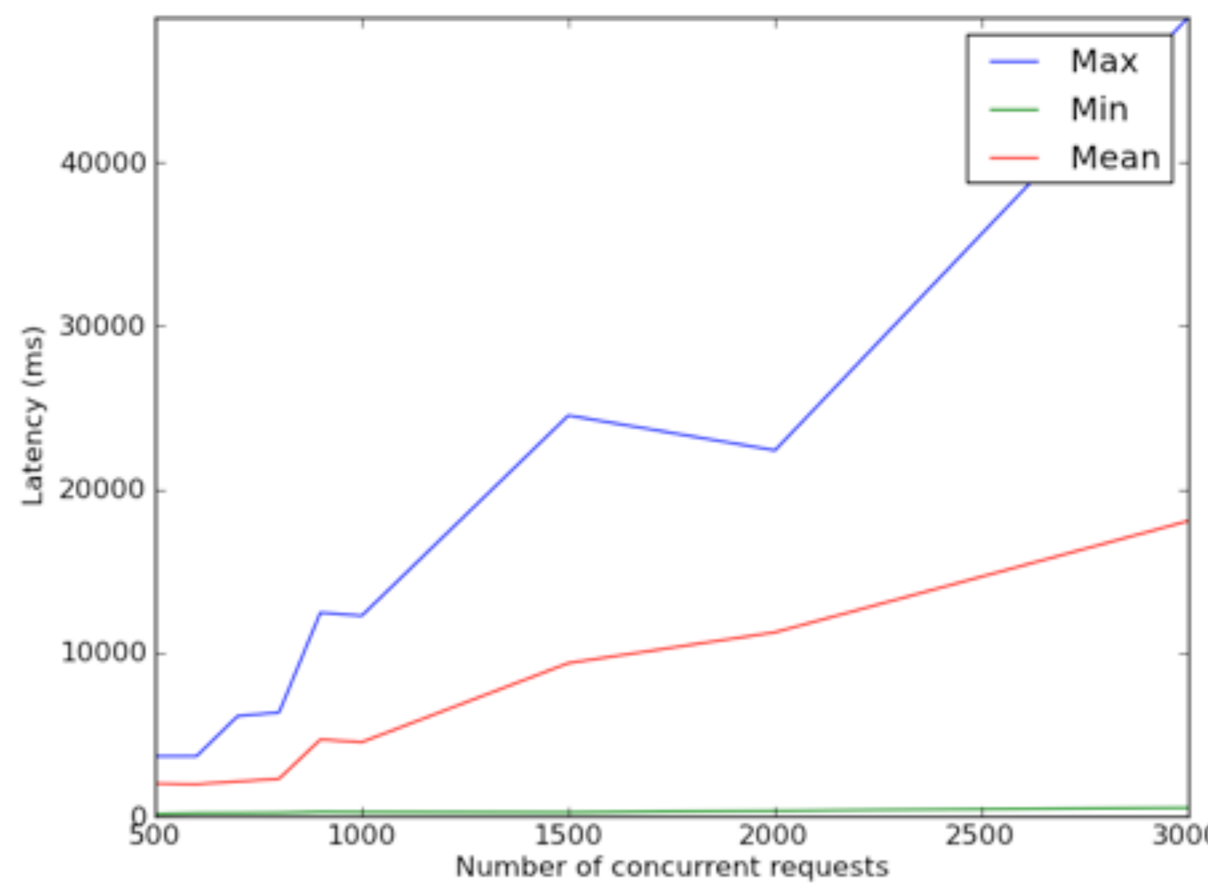


BigCouch 3 node

Job input download

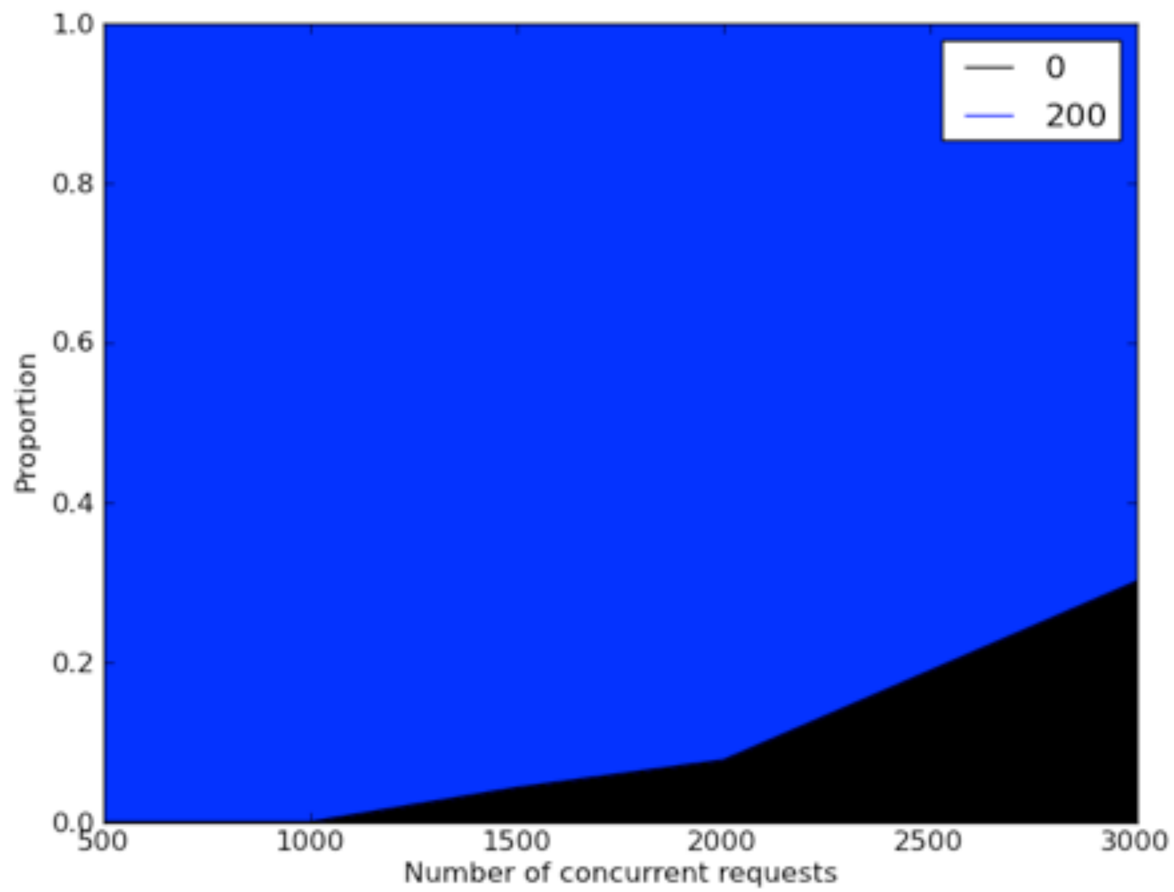


Job output upload

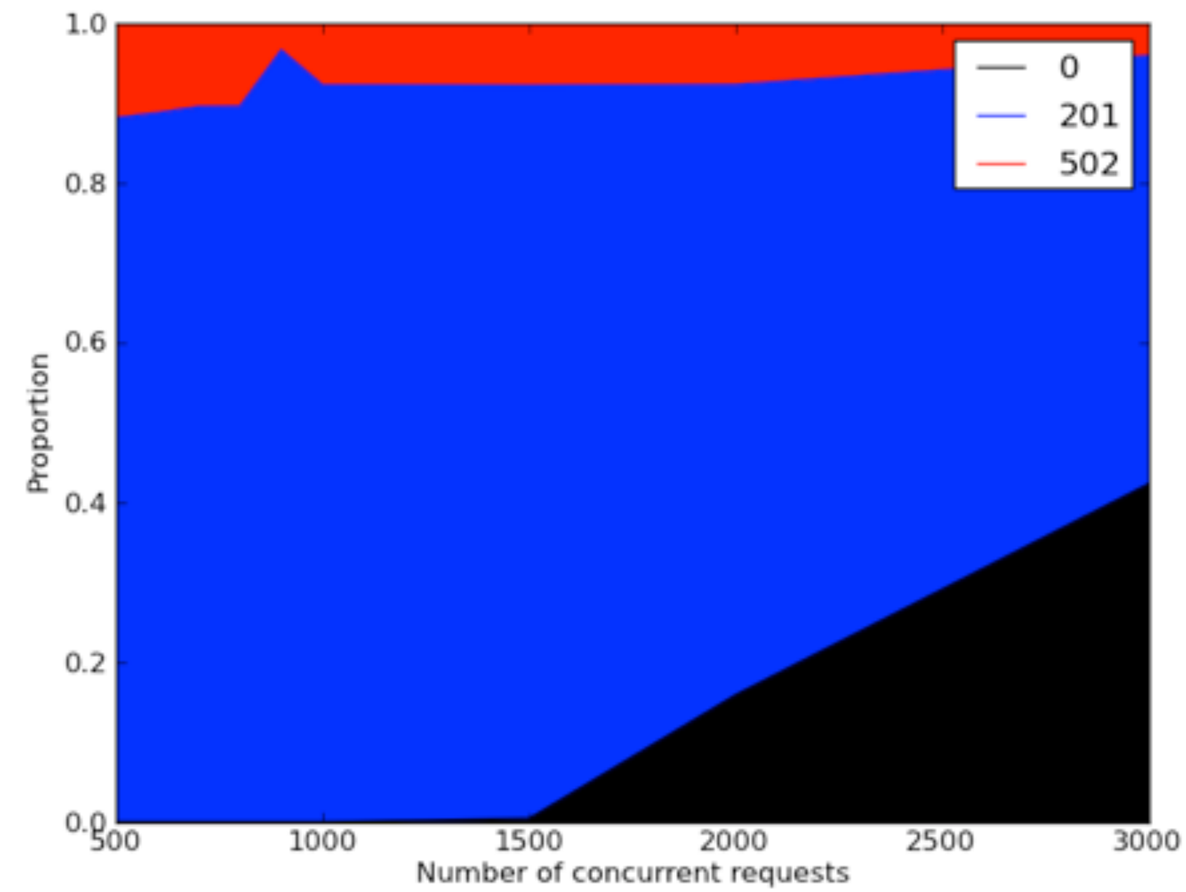


BigCouch 3 node

Job input download

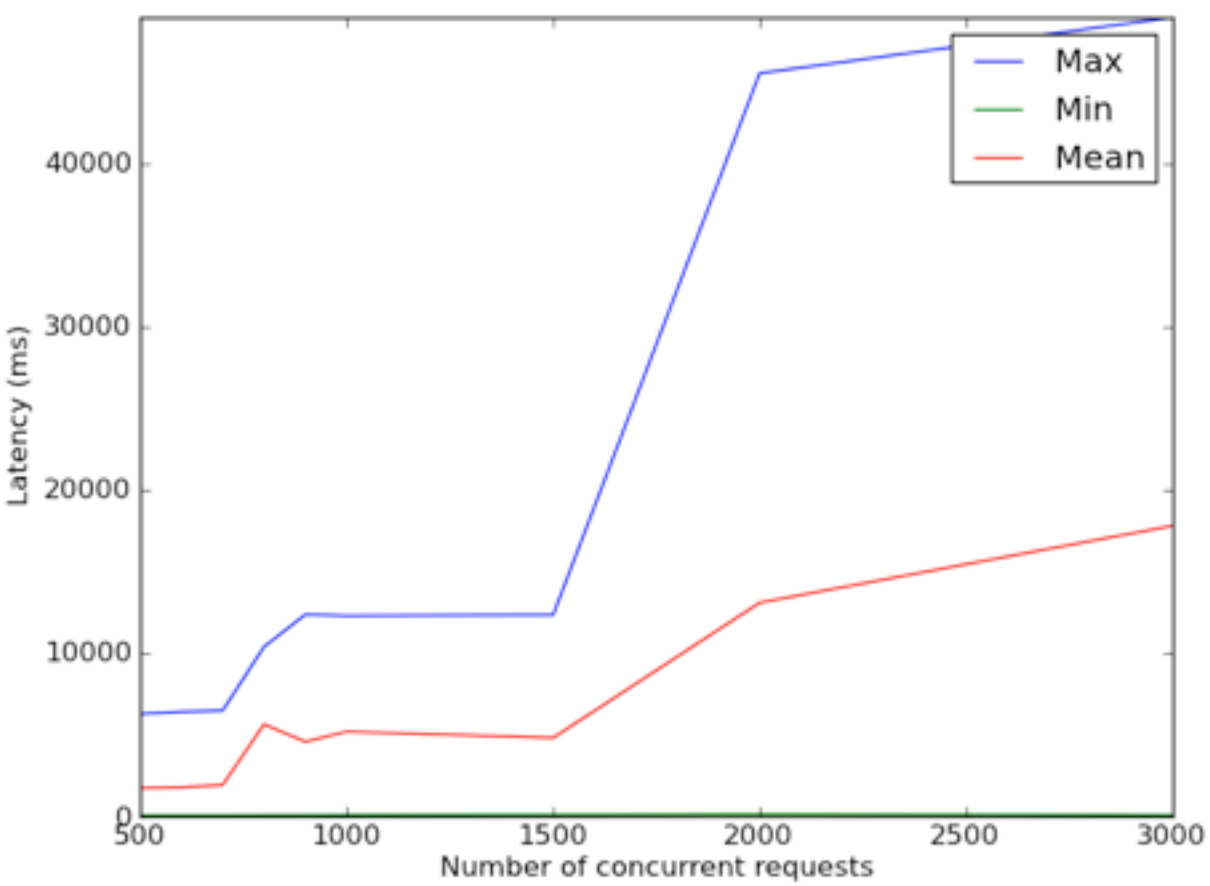


Job output upload

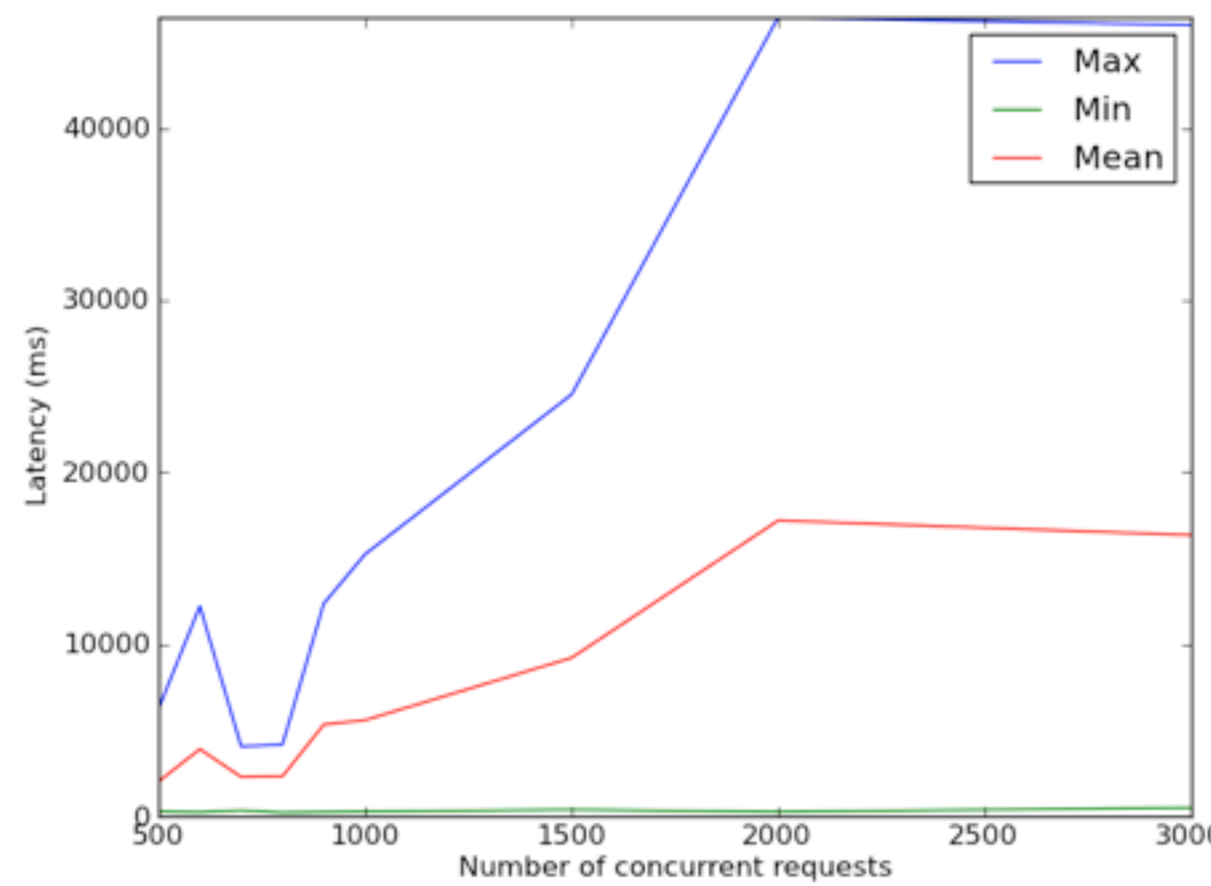


BigCouch 3 node

Job input download



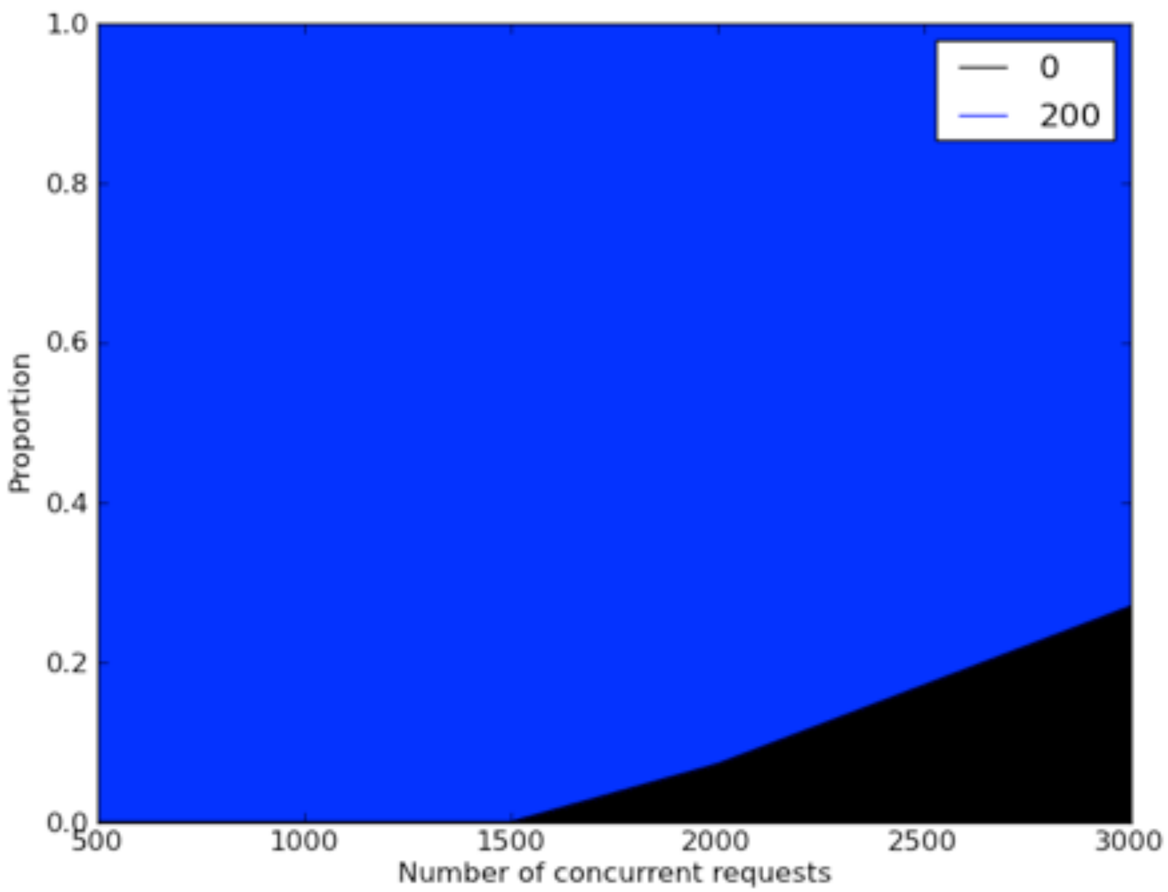
Job output upload



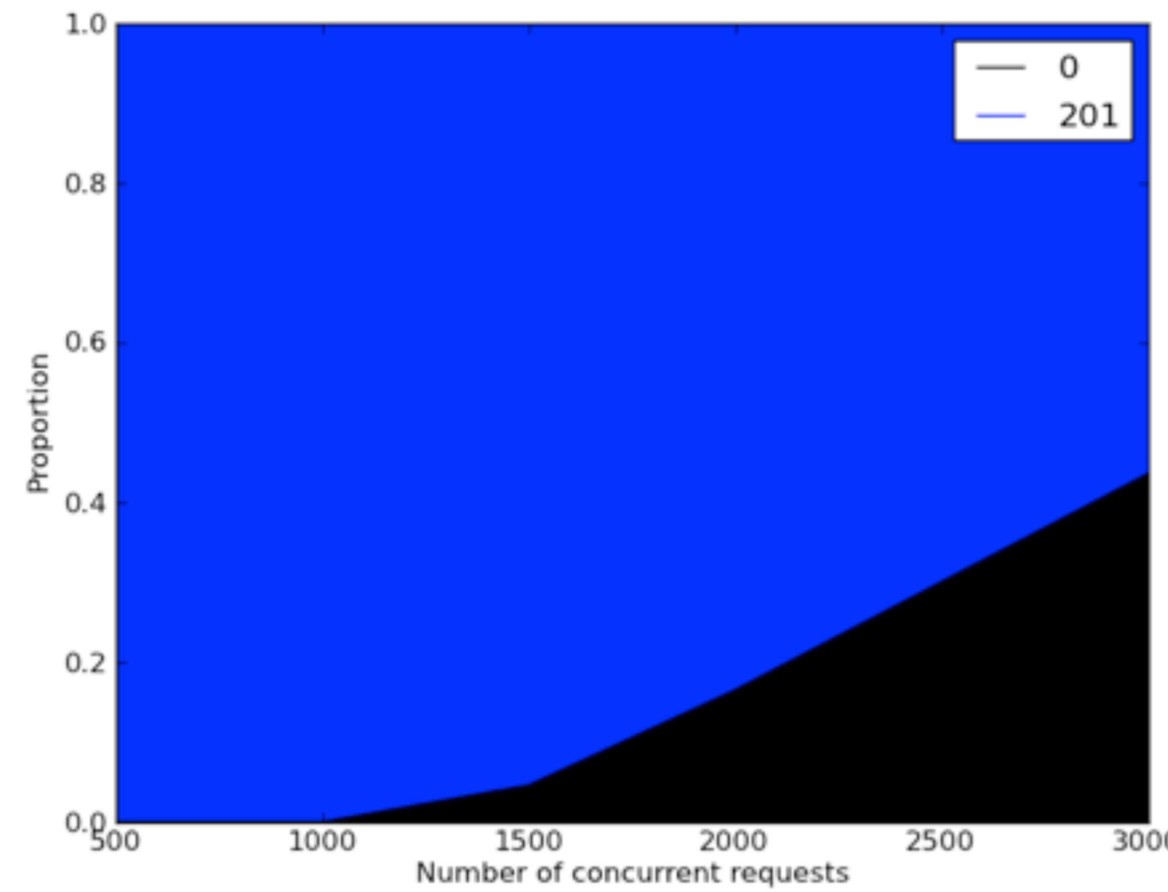
No keepalive connections from the load balancer

BigCouch 3 node

Job input download



Job output upload



No keepalive connections from the load balancer

Can we do better?

- Tested 3 node, 5 node and 6 node configurations and manipulated R and W values

Read tests by configuration

Write tests by configuration

Observations

- Benchmarking really is hard
- 1 node can handle up to 1000 concurrent writes
- CouchDB and BigCouch survive > 1000 concurrent writes (but they are not all successful)
- Observed errors:
  - 409 - conflicts (with sequential uuid generation)
  - 502 - proxy upstream error (with keepalive)
  - 500 - server error
  - 0 - no valid HTTP status code

Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

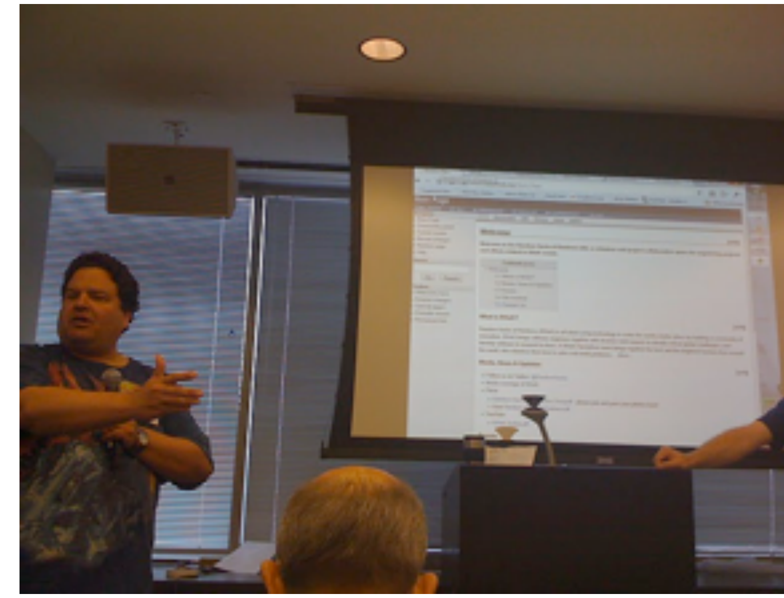
Related work

- Management of Slope Stability in Communities (MoSSaiC)
- Random Hacks of Kindness (RHOK)
  - Open source data entry and visualisation tools for MoSSaiC

MoSSaiC

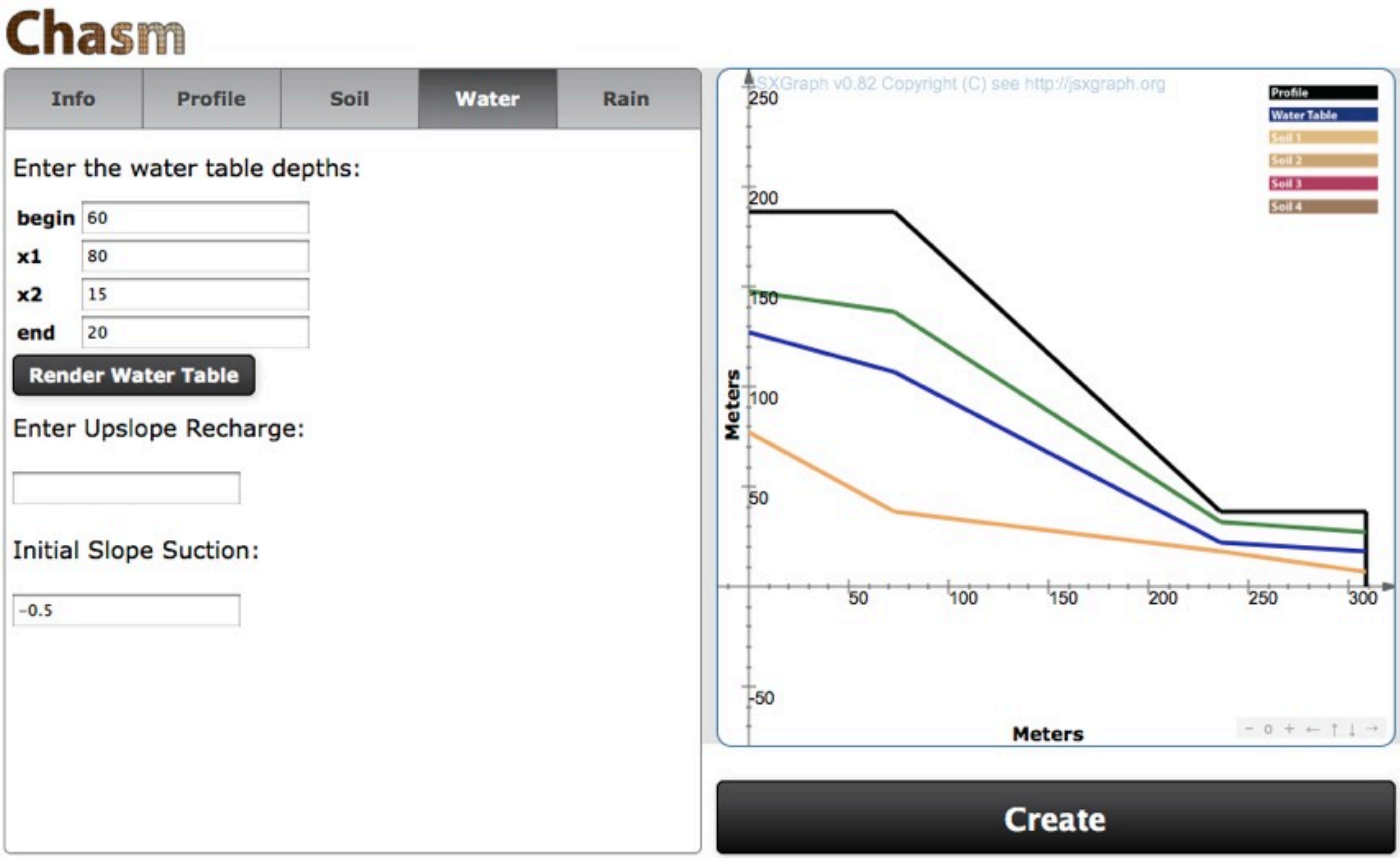
- Community based landslide risk reduction
- Low-cost solutions, e.g: Improved drainage
- Provision of field survey equipment
- Already makes heavy use of CHASM

RHOK

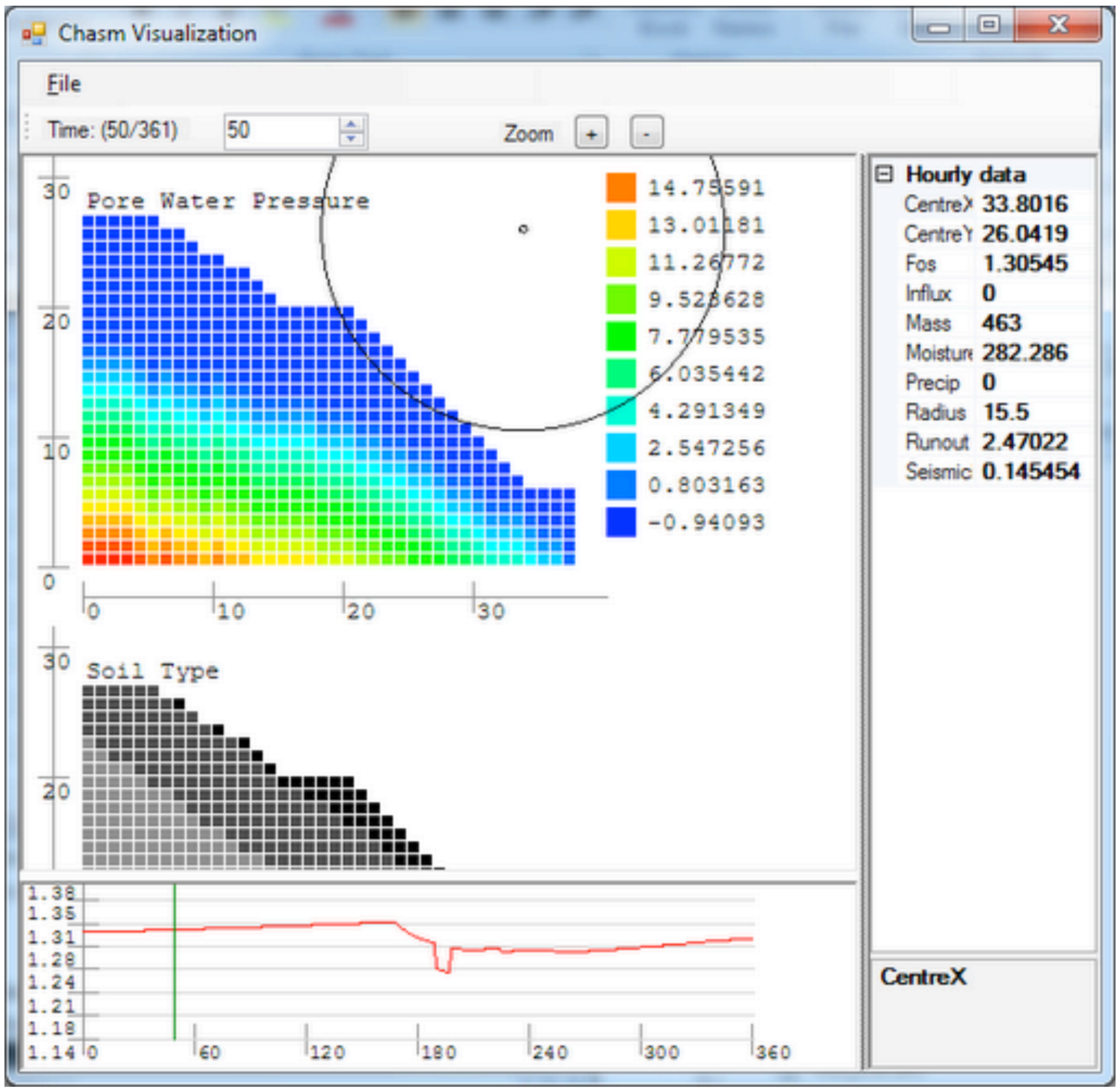


- Winning project from RHOK #1
- Standalone visualisation and data entry tools for CHASM
- http://code.google.com/p/chasm-rhok/
- Working towards interoperability with new platform

RHOK


Chasm
Info
Profile
Soil
Water
Rain
Enter the water table depths:
begin 60
x1 80
x2 15
end 20
Render Water Table
Enter Upslope Recharge:
Initial Slope Suction:
-0.5
JSXGraph v0.82 Copyright (C) see http://jsxgraph.org
Profile
Water Table
Soil 1
Soil 2
Soil 3
Soil 4
250
200
150
100
50
-50
50
100
150
200
250
300
Meters
Meters
Create

RHOK



Chasm Visualization
File
Time: (50/361)
50
Zoom
Pore Water Pressure
14.75591
13.01181
11.26772
9.523628
7.779535
6.035442
4.291349
2.547256
0.803163
-0.94093
Soil Type
Hourly data
CentreX 33.8016
CentreY 26.0419
Fos 1.30545
Influx 0
Mass 463
Moisture 282.286
Precip 0
Radius 15.5
Runout 2.47022
Seismic 0.145454
CentreX

Outline

- Landslides, landslide modelling and limitations
- The particle physicist connection
- CouchDB and BigCouch
- System architecture
- Performance tests
- Related work
- Portability

Portability

- Framework is executable-agnostic

Portability

- Framework is executable-agnostic
- Except for where it isn't:
  - Show functions for input files (can use any resource accessible over HTTP)
  - Data entry parts of the GUI
  - Results queries and visualisations

Portability

- Is there an interest in wider use of CouchDB for workload management?

Summary

- Parallel execution of landslide modelling software is useful
- The framework can be used to scale other software
- Performance testing is important - don't just add nodes and expect improvements
- Aim to get involved in CouchDB community once initial prototype running

Acknowledgements

- Dr Simon Metson (Particle Physics)
- Dr Liz Holcombe (Geographical Sciences)
- Jason Li and Bryan Michmerhuizen (RHOK)

Thank You!

Questions?