



University of Hertfordshire Research Data Management

Adventures in storage: towards the ideal Hybrid Cloud

Dr W J Worthington, University of Hertfordshire

JISC Managing Research Data Programme Workshop

25 March 2013, Aston Business School

Storage adventures

Idealised hybrid storage

Now and next

Enablers and Opportunity

Storage adventures

duplication (the blight of the backup)

context

Prof Principal Investigator regularly dumps his C:working directories to a network store. He also keeps copies of full machine backups from now defunct hardware. He never throws anything away.

requirement

Backup as a service, or active file management of our local storage.

solution

??? - currently uses departmental Network Attached Storage :(

Storage adventures

research group file sharing (enter the unsupported third party)

context

Research Group A needs to share files with co-workers in academia and industry. Some of the group are peripatetic and want to use 'bring-your-own-device' for access to their personal working files.

requirement

Best of breed storage application 'within our walls'.

solution

??? - currently uses Dropbox :(

Storage adventures

dataset repository (a mini deluge)

context Current University of Hertfordshire Research Archive UHRA,
10,000 records ~ 20GB. ONE oral history dataset = 200GB
(more impending).

requirement low cost archive storage, Repository as a Service

solution ??? - currently DSpace with high performance SAN :(

Storage adventures

work group store (and the one that got away)

context

Research Group B works with longitudinal medical data replicated locally from an external original source. The data is processed in large chunks (~0.25TB). They need storage to keep the corpus and a backup of their derived data.

requirement

2-3 TB robust, low performance storage for data that is mostly 'at rest'

solution

RackSpace cloud files, £860 /TB/yr + ~£50/yr bandwidth ;)

Storage adventures

ZendTo large file delivery system (sits well in the cloud)

context Researchers need to send large files that would burst the mail system.

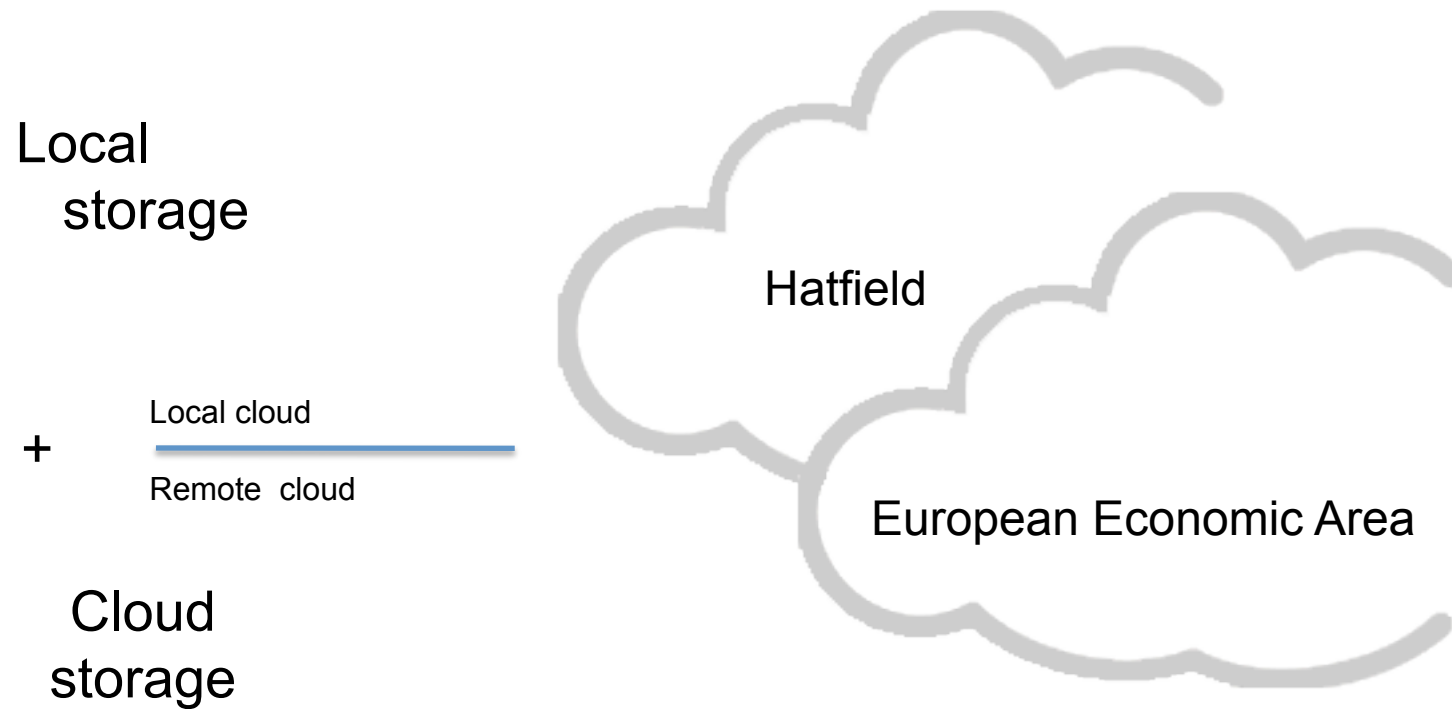
requirement Our own 'mailbigfile' / 'yousendit' file delivery system

solution ZendTo over RackSpace cloud server, elastic storage, integrated with UH LDAP and mail

😊 <http://zend.to>

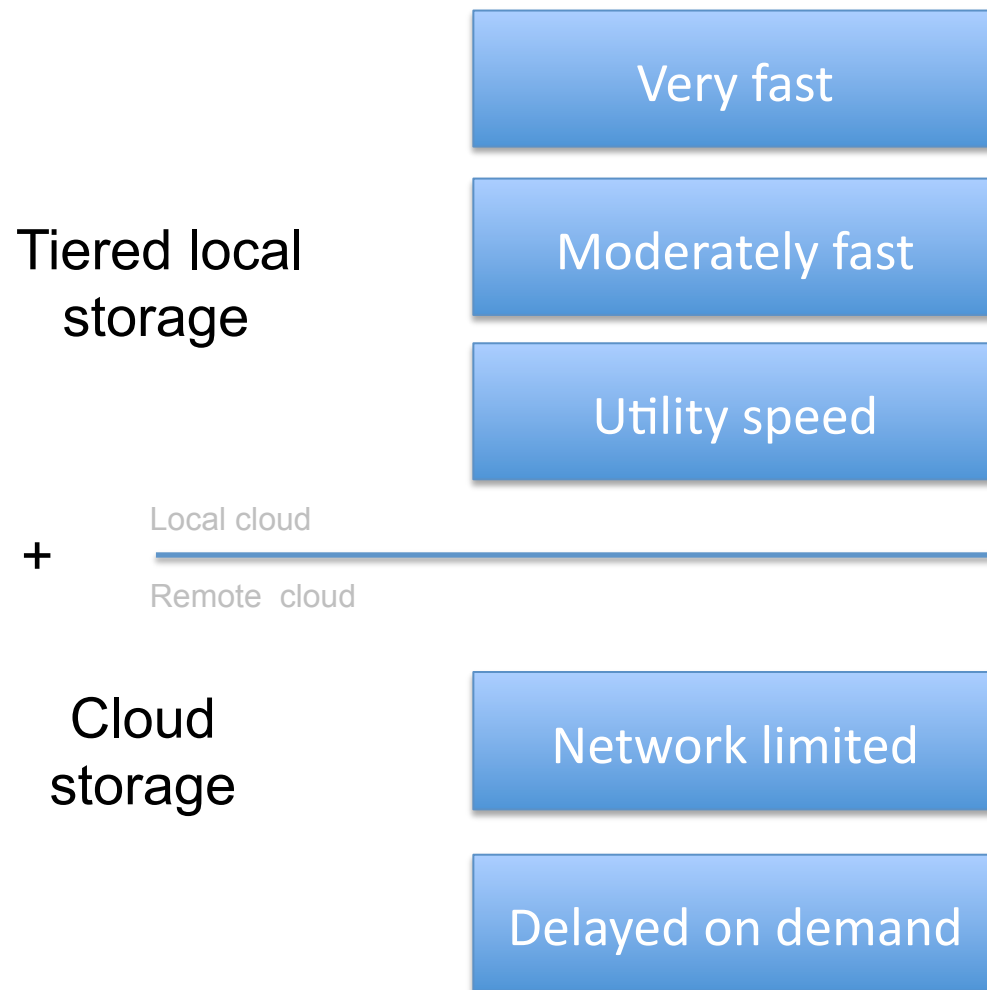
Idealised hybrid storage infrastructure

Ian storage + wan storage



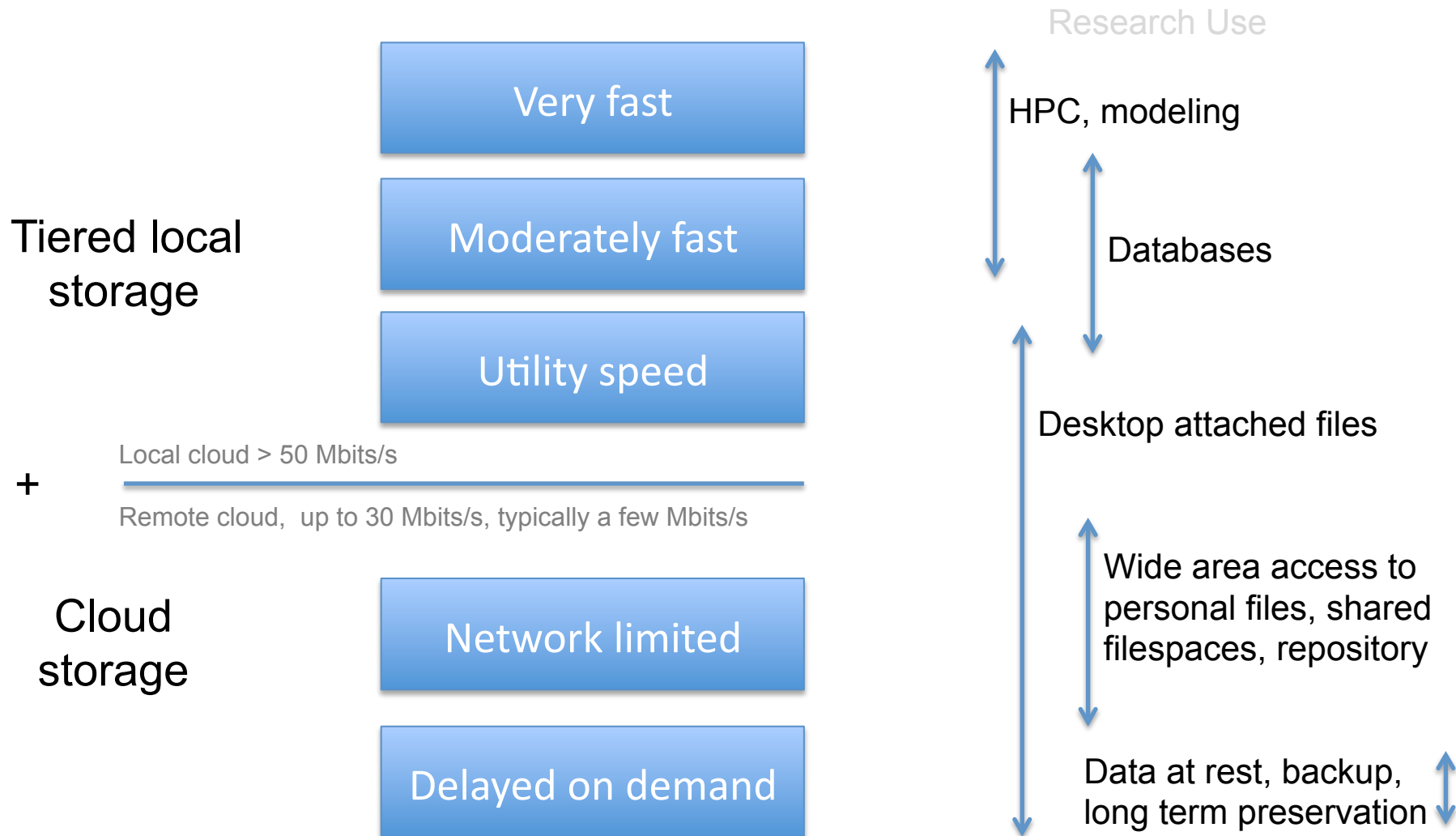
Idealised hybrid storage infrastructure

tiered performance

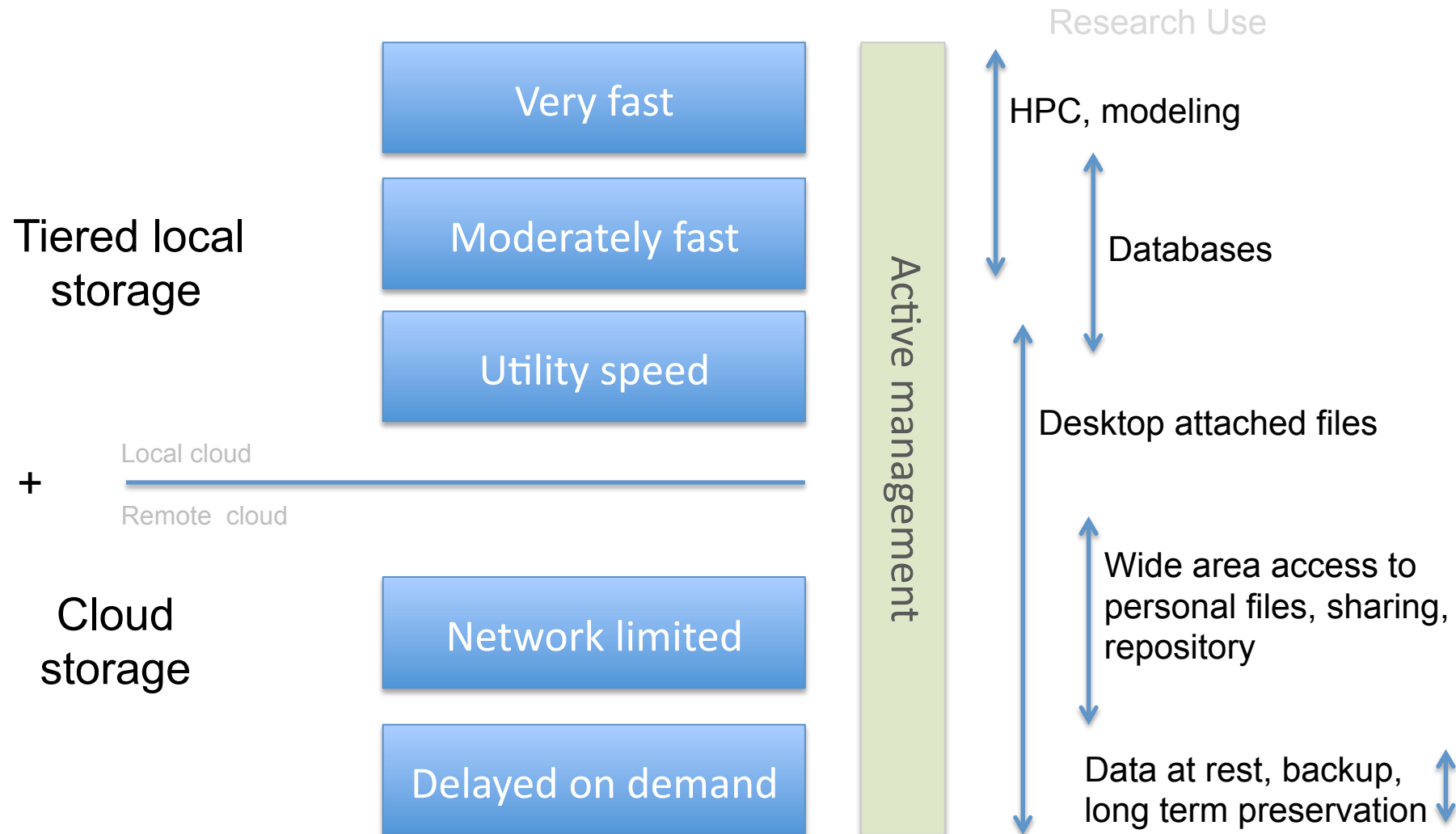


Idealised hybrid storage infrastructure

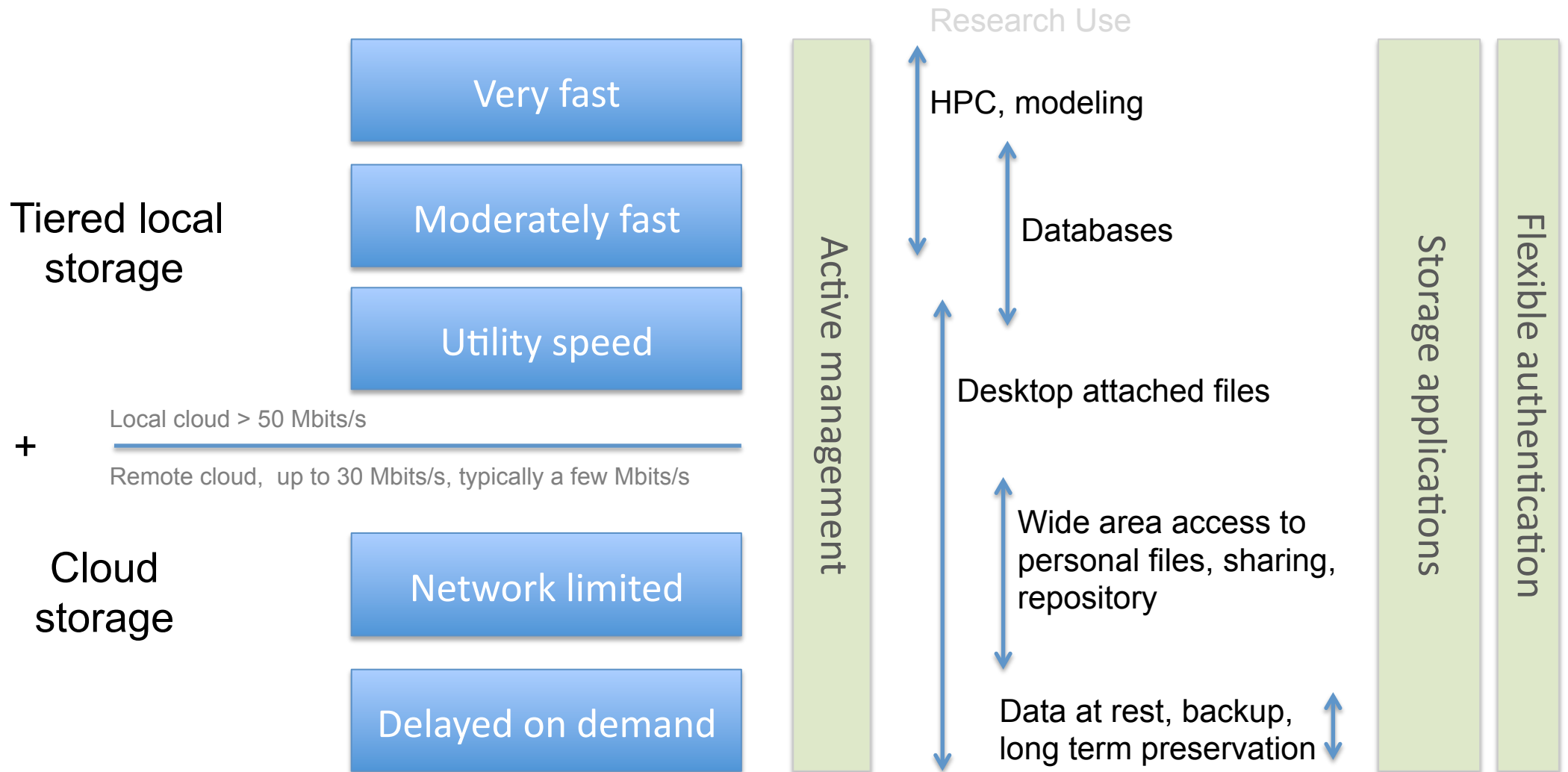
use case determines performance



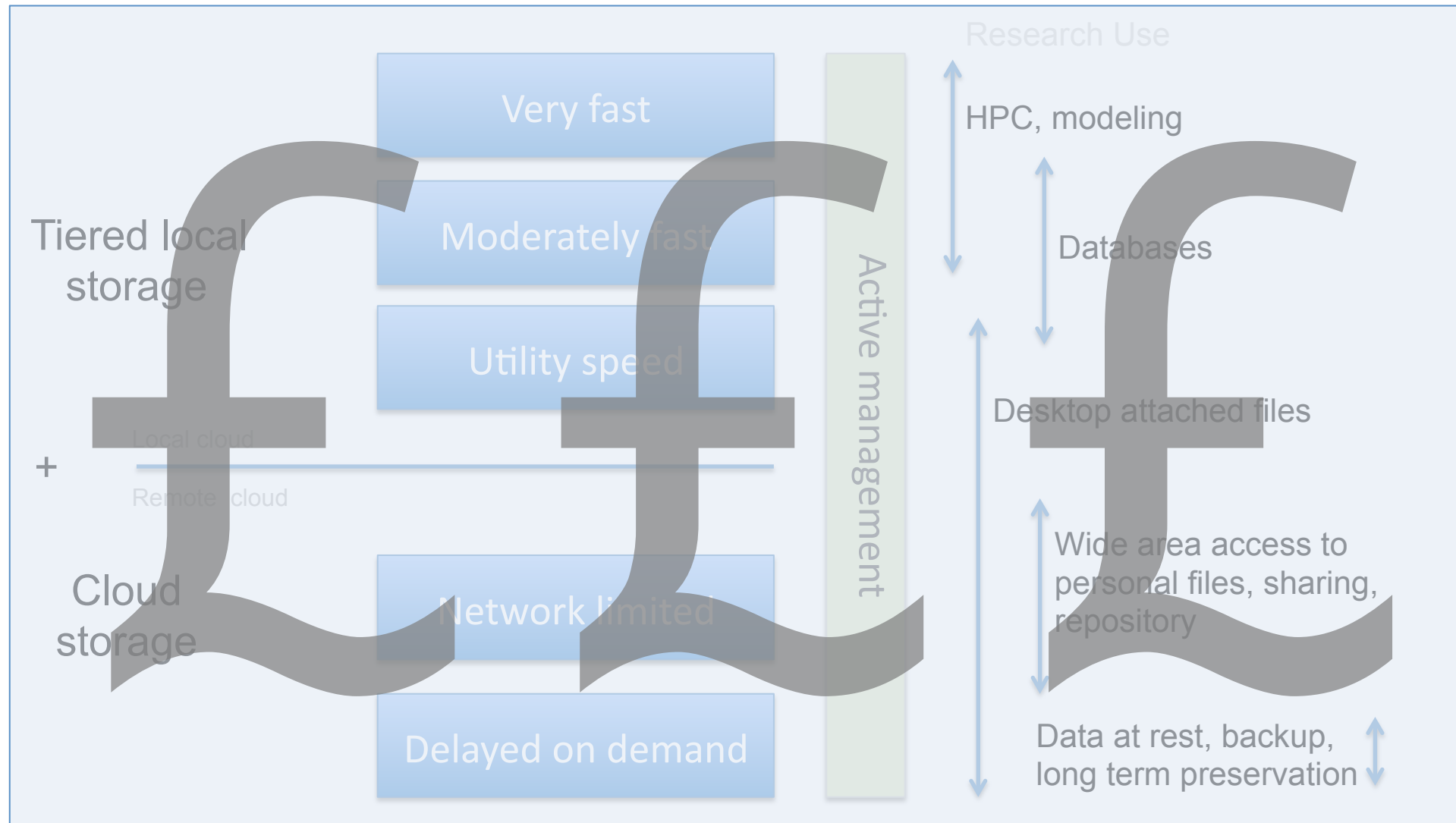
Idealised hybrid storage infrastructure active file management (location and blocks)



Idealised hybrid storage infrastructure storage applications and flexible access



Idealised hybrid storage infrastructure can be purchased, but at what cost?



Idealised hybrid storage infrastructure

incremental change necessitated by need for r o i, and limited new resource

£££

£££

£££

£££

£££

where are we now central storage offer

- ~ 300TB tier 1 SAN allocated on an 'as required' basis
- used for everything e.g. by enterprise systems and staff store
- inappropriate use and unsatisfied customers
- Users don't like staff store because it is thought to be: not enough, too difficult to use away from the desk, too slow, inflexible, difficult to share
- Providers don't like giving it to researchers because there is limited capacity (compounded by inappropriate use) and they need to protect core MIS functionality

where we are going

new elements added to our hybrid cloud

- tier 2 storage next year
- + active file management
- Backup as a Service
- Cloud files?
- Storage Applications?

enablers: back to the ideal world
integrate cloud components with local infrastructure

best of breed usability

flexible authentication
and access control for
users and groups

elastic storage applications

administration tools: usage, duplication
indicators, inactivity indicators

apps: backup as a service,
repository as a service

‘Capex’ funding models, cash ceilings

satisfies local
governance

opportunity for vendors

focus on storage as service

not about cost (provided it is cheaper)

the offer needs to overcome barriers:

- integration
- governance
- Researcher's predisposition toward independence

JISCMRD activity at University of Hertfordshire

Contact

<http://research-data-toolkit.herts.ac.uk/>

Project tags: #rdtk_herts, #rdmtpa_herts, #jiscmrdr

Dr Bill Worthington, RDM Projects Manager

w.j.worthington@herts.ac.uk

Dr David Ford, Chief Technology Officer

d.m.ford@herts.ac.uk