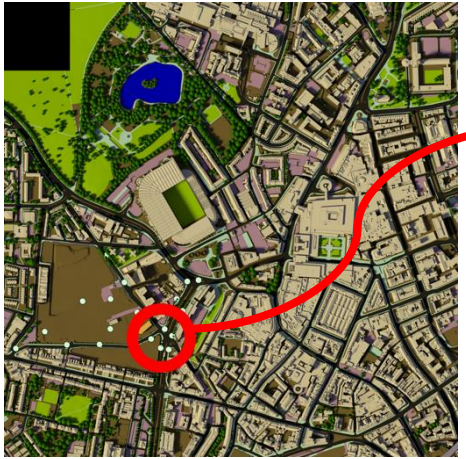
Visualizing Urban IoT Using Cloud Supercomputing

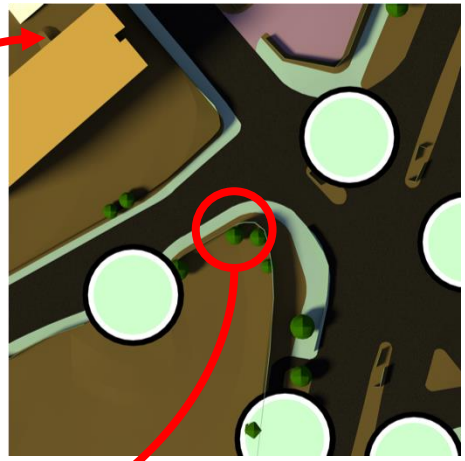
Nicolas Holliman, Manu Antony,
Stephen Dowsland & Mark Turner

1.

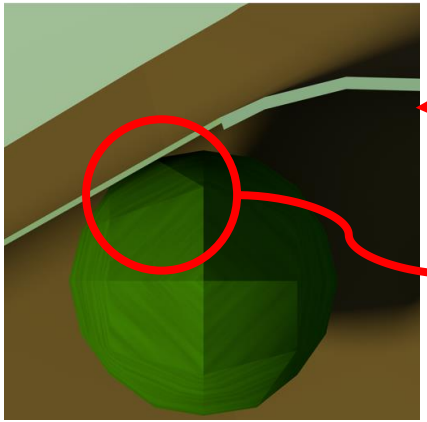
Introducing: The terapixel image


Virtual Tour - 1048576x1048576

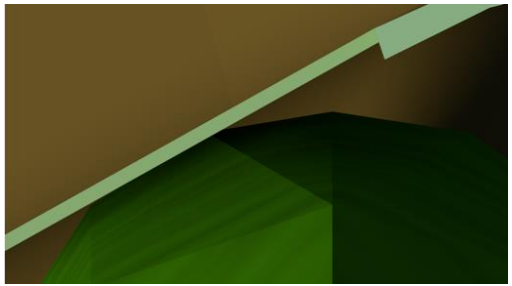




16x zoom



256x zoom



512x zoom

How far can we zoom in?

Whole image:
1.28 km x 1.28 km
1048576 x 1048576 pixels

Zoomable 512x to full HD:
2.5 m x 1.25 m
2048x1024 pixels

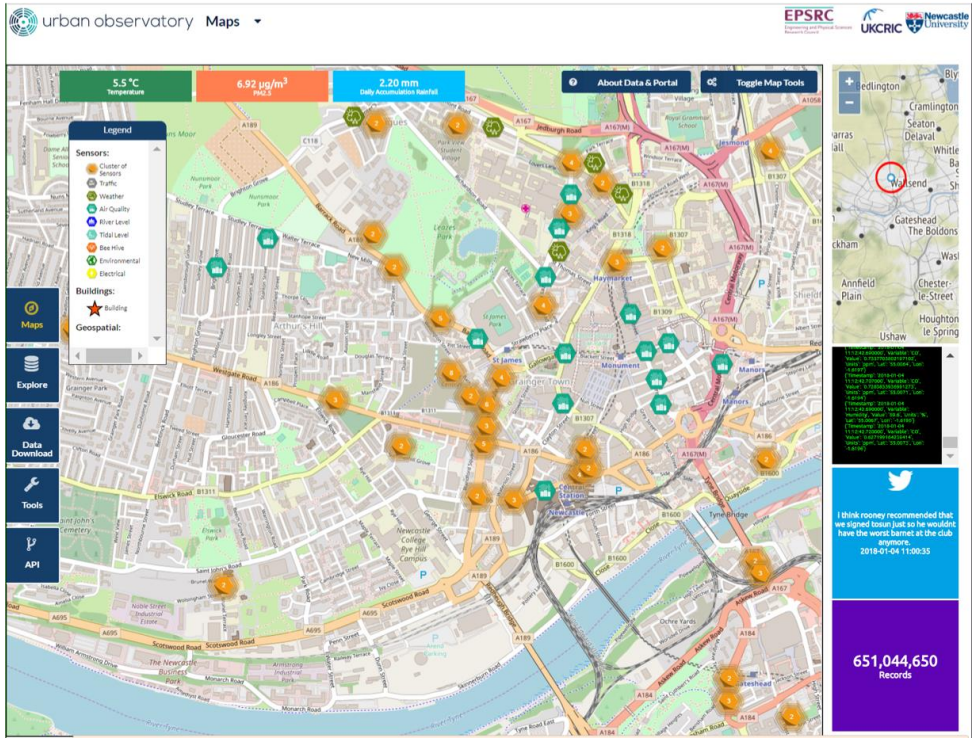
1x1 pixel == 1.22x1.22 mm

Terapixel is
1,099,511,627,776 pixels
(530,243 full HD TV images)

Allow visualization across many scales

Newcastle 3D city model courtesy of Northumbria University, VRV Studio

IoT Data source: The Newcastle Urban Observatory



http://uoweb1.ncl.ac.uk

Stores approx. 2.4Gbytes every day from IoT sensors across the city and the Tyne and Wear region.

Data is published within a few minutes.

More than 650 million readings collected over the last 3 years.

(Probably) largest open data set from a Smart City in the world.

2.

Why render using the cloud?

Benefits

- Performance
- Cost
- Thin Clients

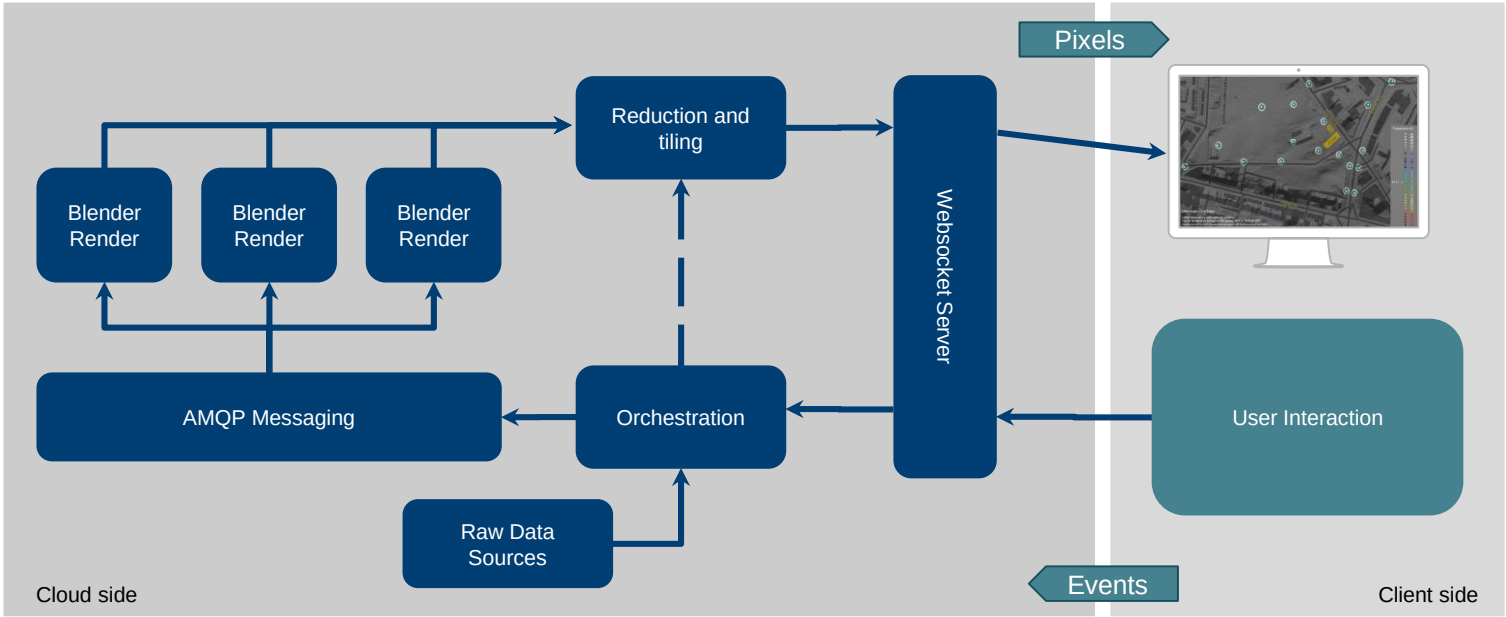
Drawbacks

- Engineering overheads
- Less predictable cost
- Vendor lock-in

3.

How to render in the cloud

UICE System Diagram



Live feed from a subset of the UO IoT sensors, the image generated shows mean hourly temperature readings.

Added the 3D temperature glyphs to an Environment Agency topographic model of the city

Hosted on Azure with a private cloud hosted Blender Render nodes.

Azure

openstack.

blender™

A Scalable Platform for Visualization Using the Cloud, S.Dowsland, M.Turner, N.Holliman in CGAT, April 2017, Singapore.

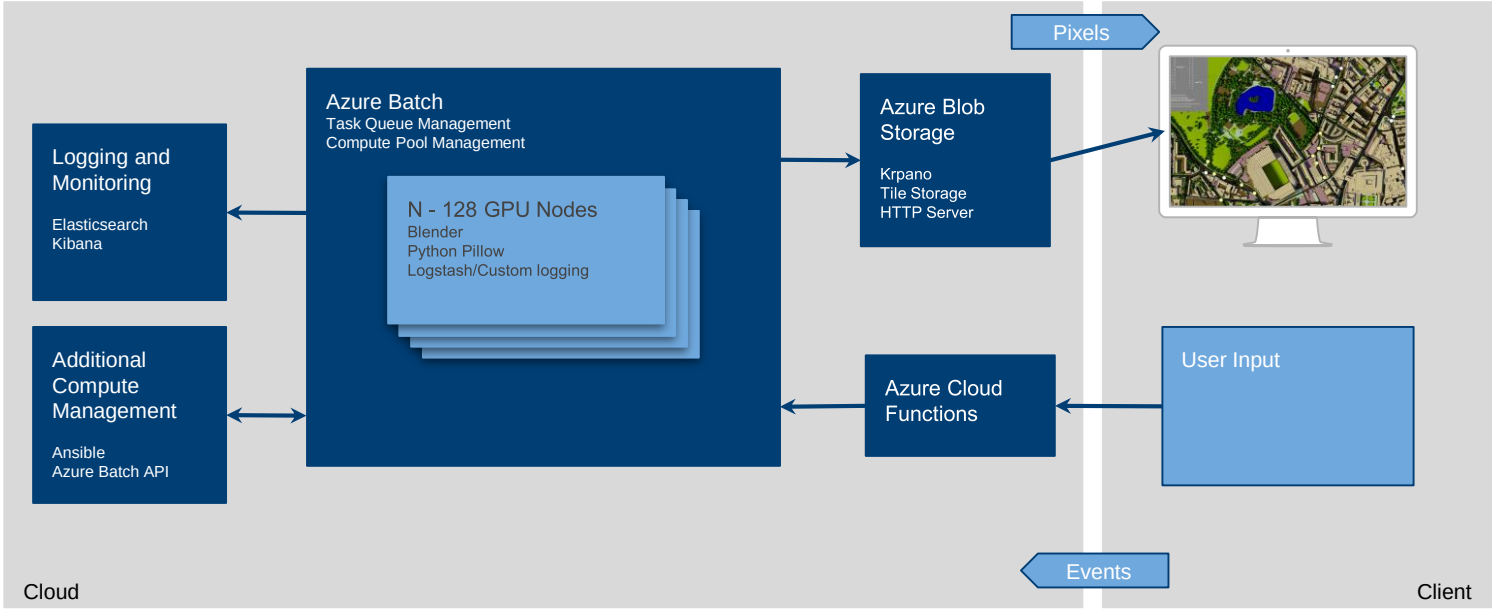
Make use of Azure services

Functions

Batch

Storage

TeraScope System Diagram



Peak 372 Tflop cloud supercomputer: 128 NC6 (6 core + 1 Tesla K80 GPU) nodes plus 1 Tbyte of Azure blob storage

What needs to be computed?

Pyramid level	Image side length (pixels)	Total number of pixels at this level	Number of 512x512 tiles	Tile side length in the real world (mm)
12	**1048576**	**1099511627776**	**4194304**	**625**
11	524288	274877906944	1048576	1250
10	262144	68719476736	262144	2500
9	131072	17179869184	65536	5000
8	**65536**	**4294967296**	**16384**	**10000**
7	32768	1073741824	4096	20000
6	16384	268435456	1024	40000
5	8192	67108864	256	80000
4	**4096**	**16777216**	**64**	**160000**
3	2048	4194304	16	320000
2	1024	1048576	4	640000
1	512	262144	1	1280000

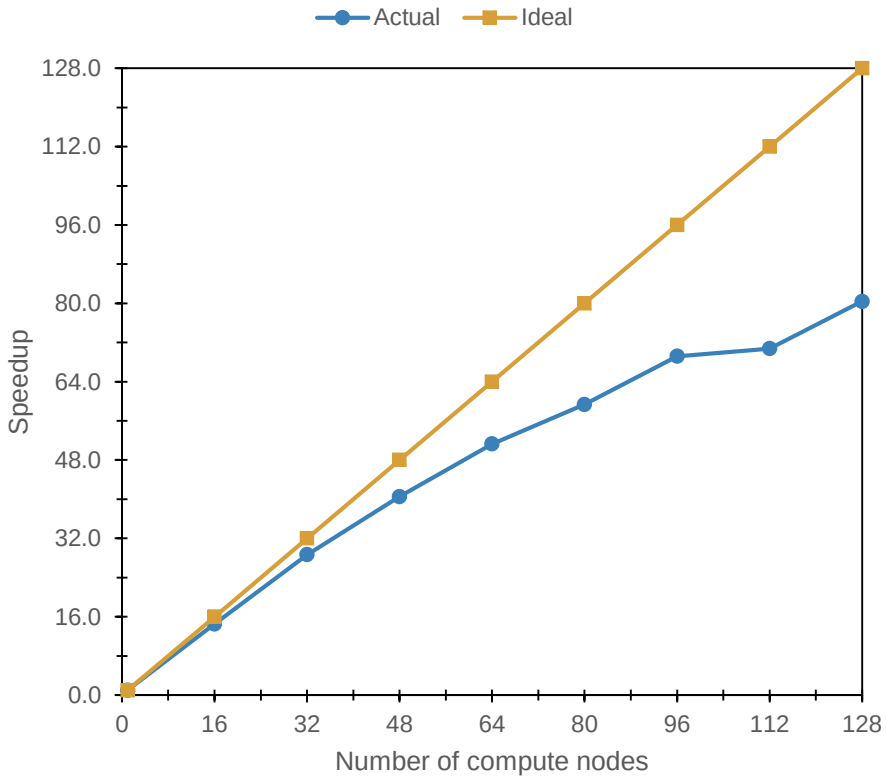
Render 65,793 4096*4096 images, i.e. all pixels in levels 12, 8 and 4.

Then build a hierarchical image data set of 5,592,405 512*512 pixel tiles.

4.

Results

Nodes v Speedup - Gigapixel

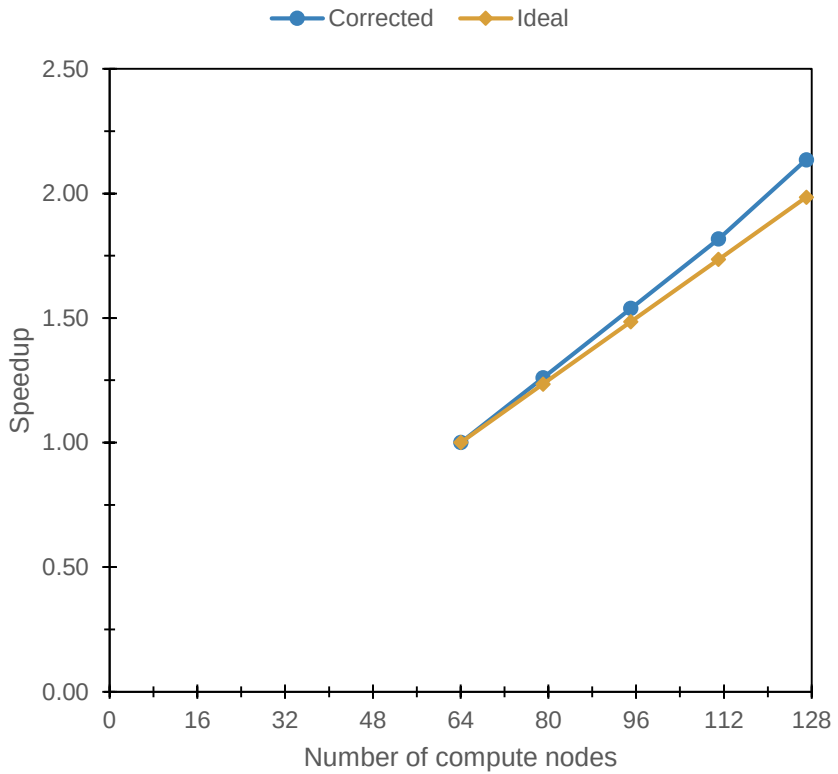


Scaled compute from 1 to 128 NC6 in East US region, in steps of 16 nodes.

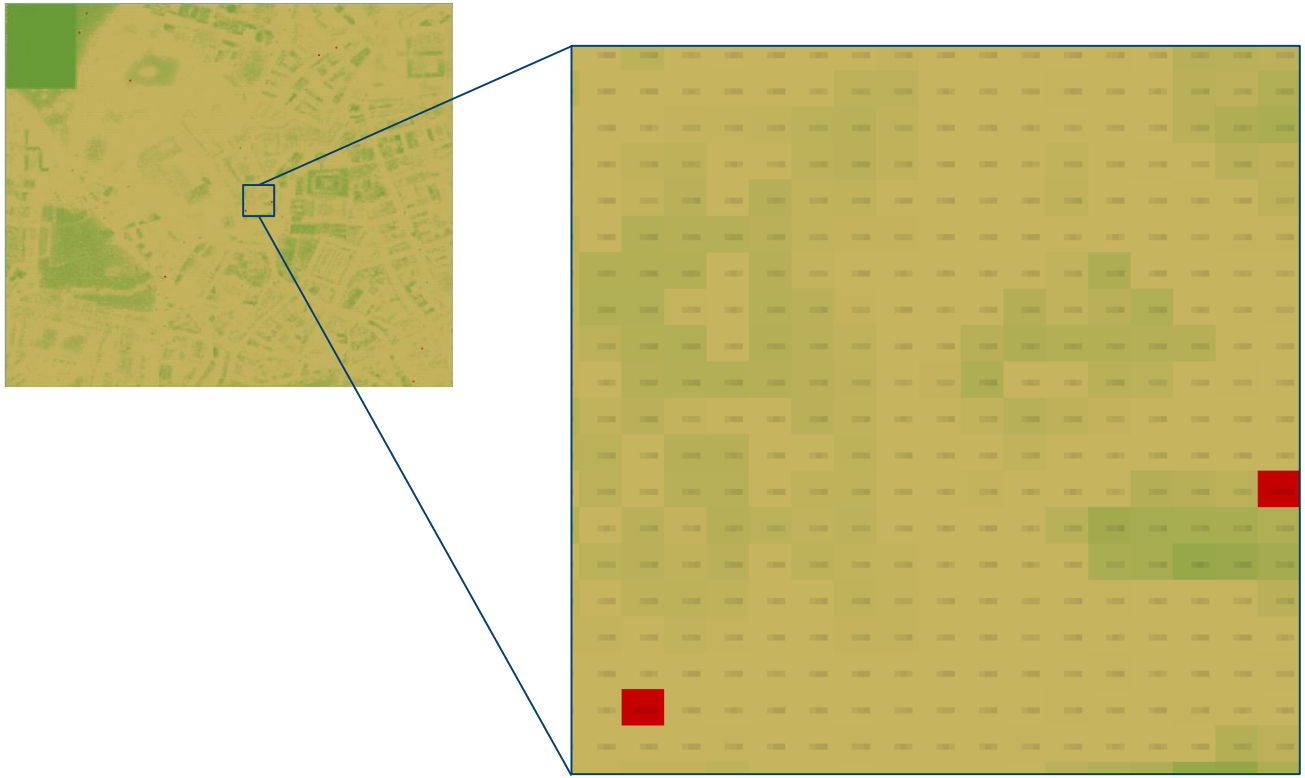
Rendering time down from 570 to 7 min

At the upper end we would expect to see limited scaling efficiency because of the small number of tasks per node, and we do find this above 96 nodes.

Nodes v Speedup - Terapixel



Render Times Heatmap



What does it cost?

Nodes	Duration (s)	Price
128	74348	£1,064.27
112	87249	£1,092.82
96	102885	£1,104.57
80	125442	£1,122.29
64	159925	£1,144.64

- 128 machines purchase cost would be £625K
- Use of low priority nodes would reduce the cost per run by up to two thirds

5.

Where next?

Next steps

- Larger geographical area
- Interaction
- Exploration in 3D
- Scale up to 1024 nodes (Eqiv. £6m+ Machine)

Terascope

- Rendering at cloud scale
- Cost effective
- Engineering is a barrier to widespread use
- Terascope Viewer - https://bit.ly/2DKI2JH

Newcastle University

National Innovation Centre for Data

Innovating through data

Microsoft Azure

Newcastle City Council

Northumbria University NEWCASTLE