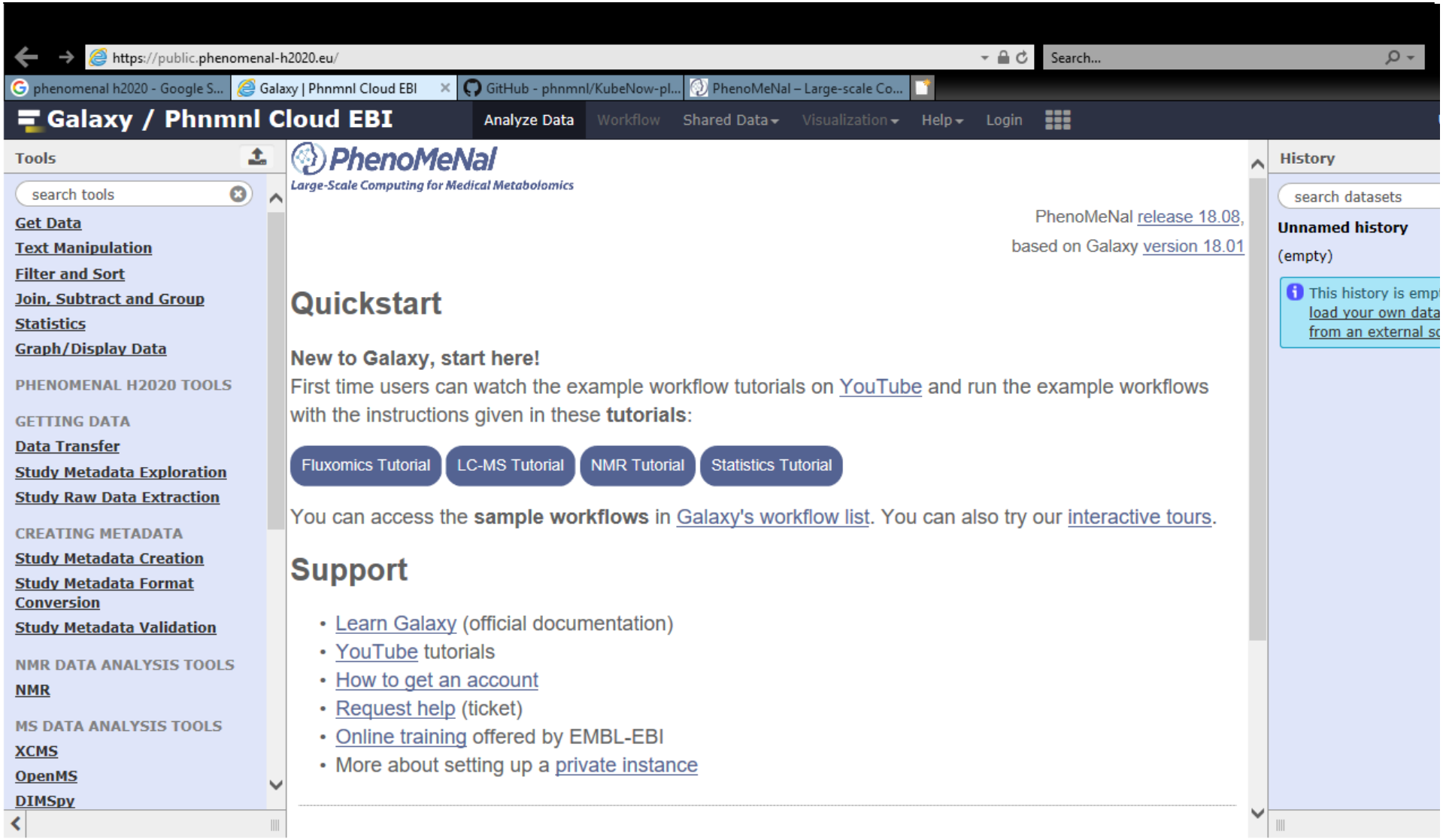
Scalability of Metabolomics Tools in the Cloud

Dr Jianliang Gao
Department of Computing
Imperial College London

@ RSLondonSouthEast 2019

Introduction - PhenoMeNal

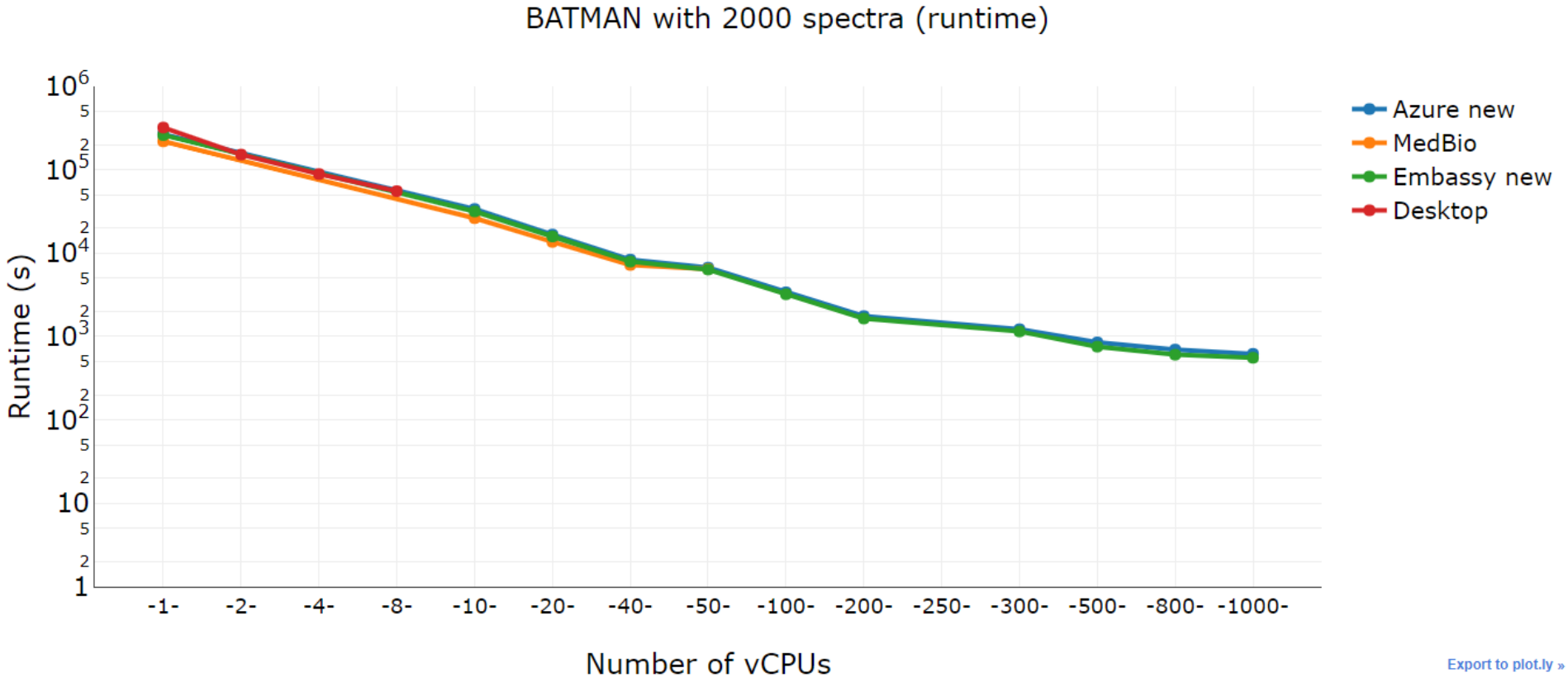


Methods

- Deployment
  - 3 platforms VS high end desktop: Medium scale cluster (50 CPUs/1TB RAM) and 2 large scale clusters (1000 CPUs/3TB RAM) – Microsoft Azure and EBI Embassy cloud
  - KubeNow https://github.com/phnmnl/KubeNow-plugin
    - » Include: Galaxy Web UI, Jupyter Notebook, Luigi GUI
    - » Support: KVM, OpenStack, AWS, Azure, Google Cloud
  - Tools: BATMAN and PAPY
- Metrics for performance measurement
  - Runtime and Efficiency of Strong Scalability https://github.com/csmsoftware/phnmnl-scalability

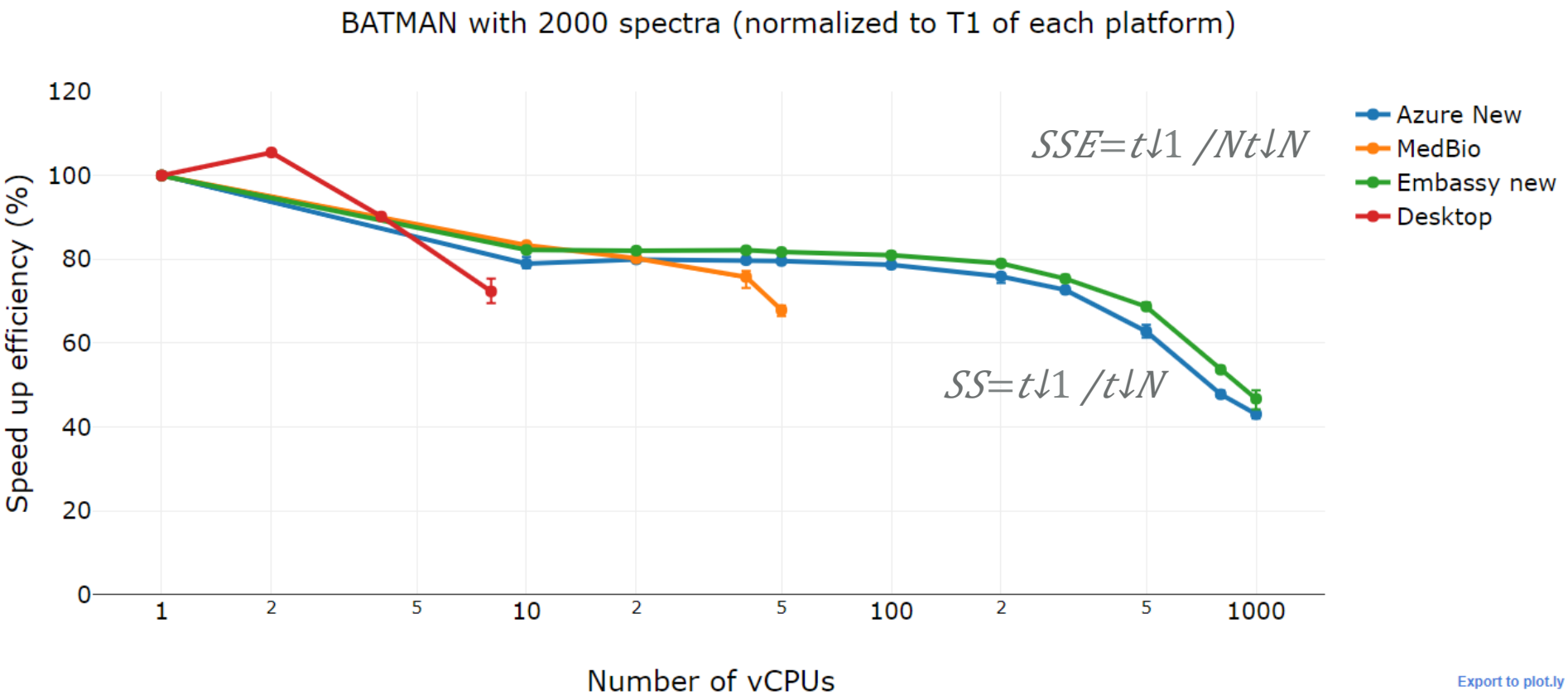
Results

- Runtime



Results (cont'd)

- Strong scaling efficiency (*SSE*)



Acknowledgement

- PhenoMeNal team at Imperial College London
- EBML-EBI, Uppsala Univ. (Sweden), IPB-Halle (Germany) etc…
- BioMedIA group and dHCP team at Imperial College London
- Microsoft Azure Research

Thank you!!!

- Questions?