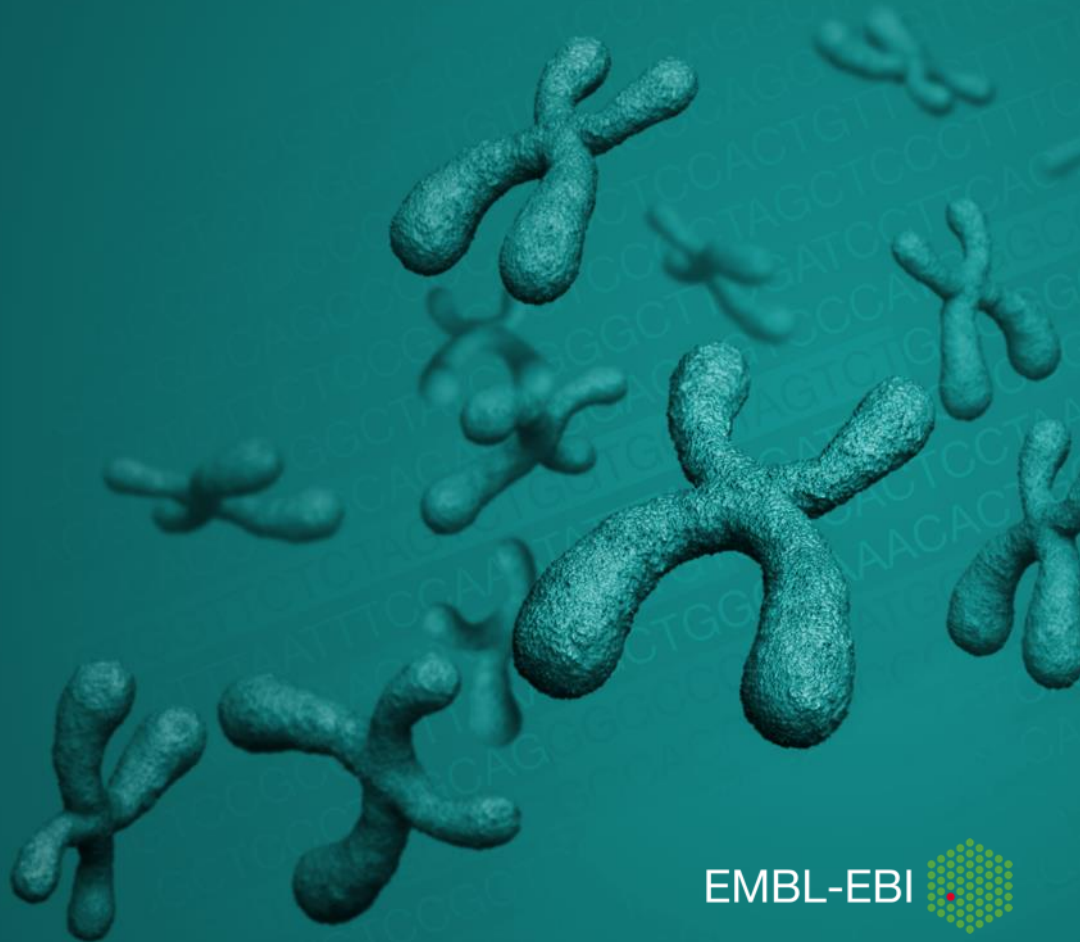


European Bioinformatics Institute: ICT Challenges

Steven Newhouse, Head of Technical Services



European Bioinformatics Institute

- Outstation of the European Molecular Biology Laboratory
- International organisation created by treaty (cf CERN, ESA)
- 20 year history of service provision and scientific excellence
- EMBL-EBI has 500+ Staff & €50 Million Budget
- Provide services to a wide range of “as-possible” usage model
- Thin-client model
 - Web browser & web services
 - Equivalent to SaaS



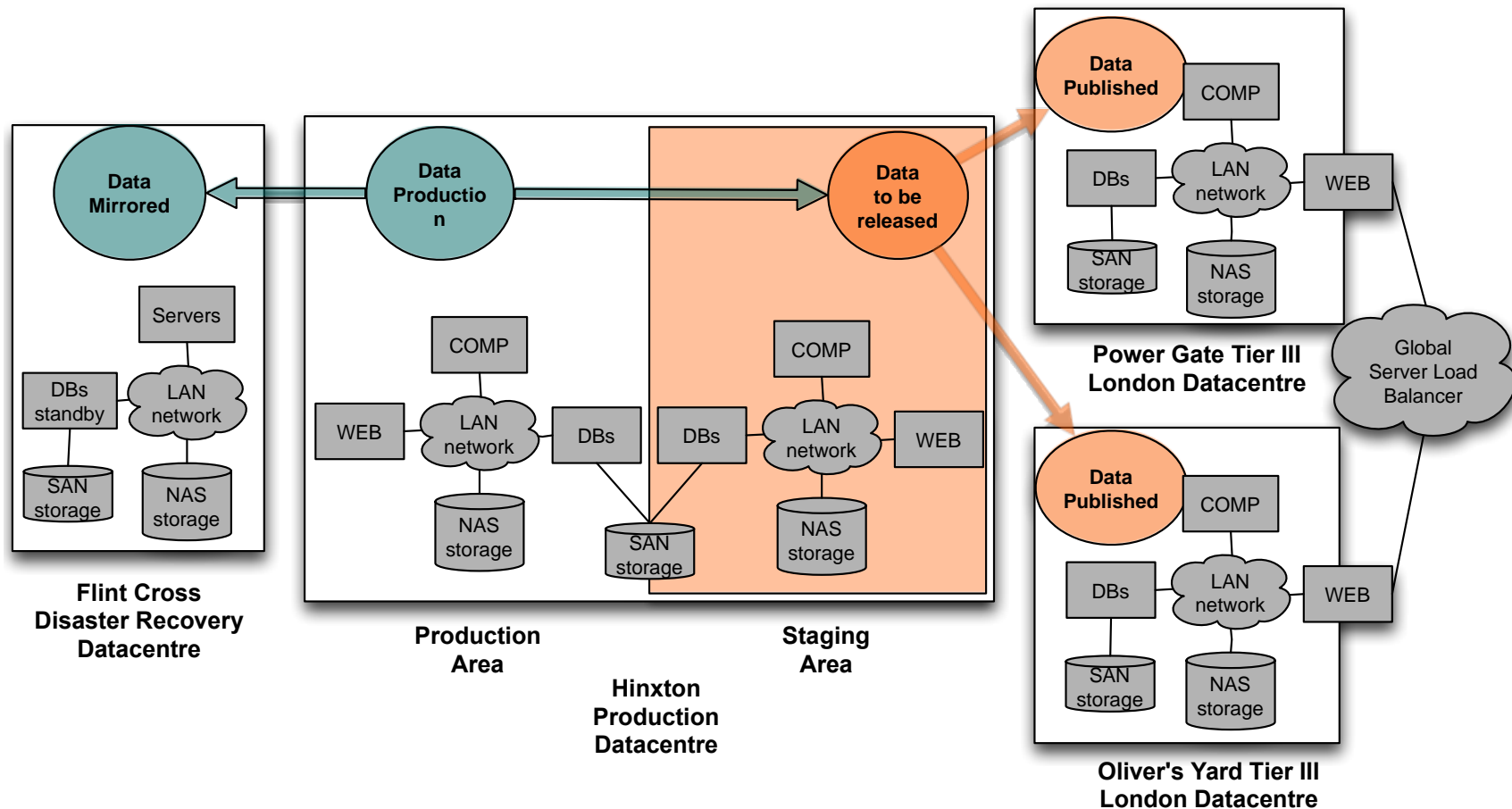
The Challenge Facing Bioinformatics

- Volume and variety of genomic data expanding
 - Data at EBI doubling every year - replication is challenging
 - >12,000 CPUs & 30PB (but need more!)
- Complex analysis
 - Access to both public and managed access data sets
 - Bespoke workflows and tools across a variety of domains
 - Issues with disk to memory bandwidth
- EMBL-EBI Provides
 - Public & managed access data sets
 - Web and programmatic access to services (3M unique users)

Impact on EMBL-EBI's Infrastructure

- Grow the capacity of the current data centres
 - Commodity infrastructure – blades and NAS (50 racks)
 - RDBMS and SAN for high throughput transaction processing
 - Tape backup is no longer feasible
- Provide a resilient topology by geographical separation
 - Against local & regional disaster in the UK
 - Against national disaster through international collaboration
- Enable new easier science through the cloud
 - Provide access to the increasingly hard to replicate data sets

Overview EMBL-EBI IT infrastructure



Data centre virtualised throughout with VMware

Technology Areas

- Storage (deployed and/or assessed)
 - Panasas, IBM (GSS, TMS), EMC (Isilon, VNX, ScaleIO), Infinidat, Avere, DDN (WOS), Violin, Cleversafe, Tegile, NexSan, HP (3par), NetApp, HDS
- Wide Area Networking
 - Dedicated light paths & commodity internet over UK NREN
- Databases
 - Oracle, MySQL, MongoDB, Delphix, Vertica, Clustrix
- Computing
 - HP, VMware (cloud & data centre), LSF (large cluster), OpenStack, OpenVZ
- Data Centres
 - Telecity (tender open for renewal)

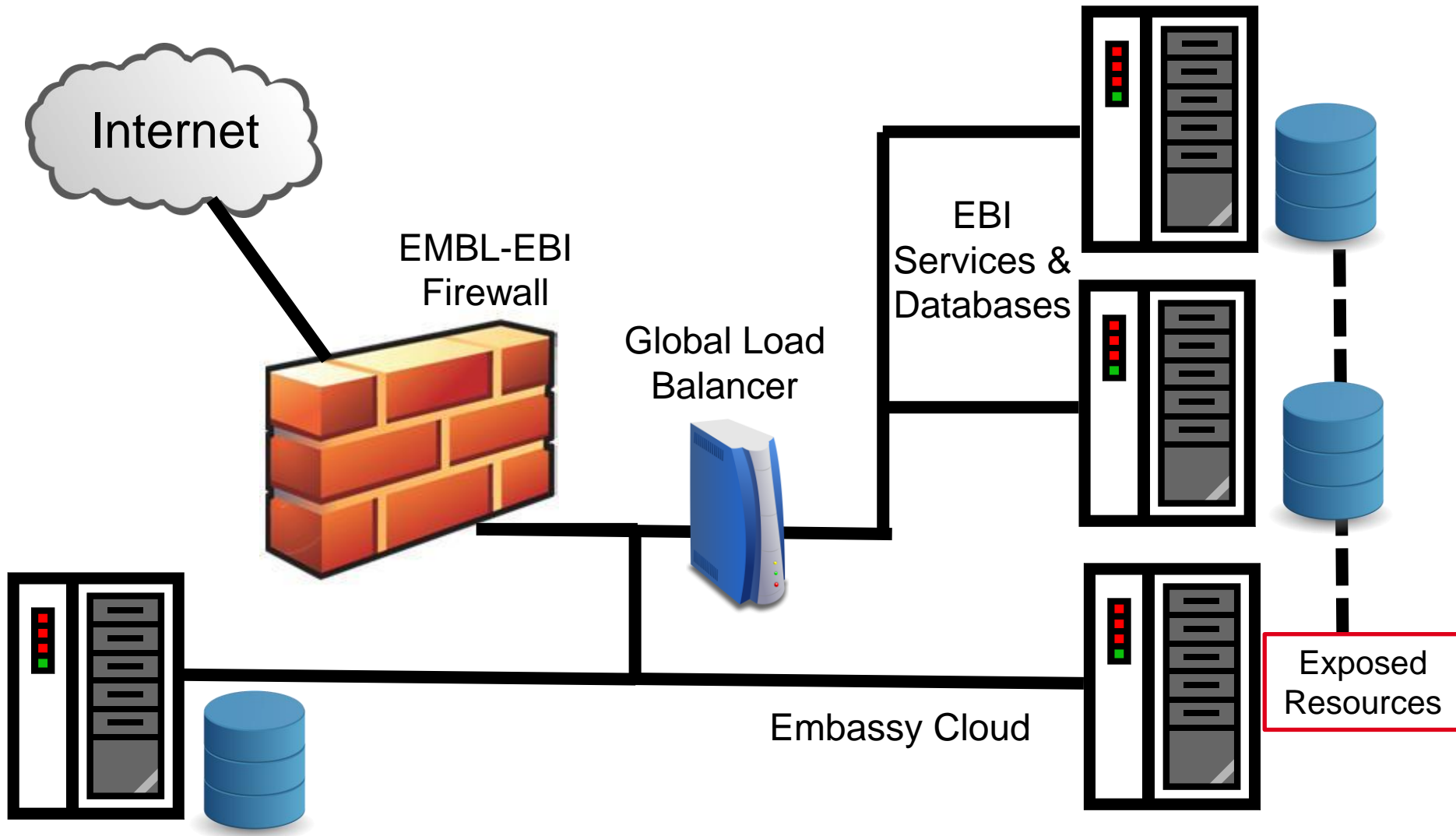
EMBL-EBI Embassy Cloud

- Pilot service hosted at EMBL-EBI data centres
 - Logically isolated outside EBI's LANs
 - Secure flexible infrastructure for both tenant and host
 - File based access to EBIs' data sets
 - Currently, only the 1000 Genomes dataset exposed
- Expect both academic and commercial users
 - Wishing to move their compute and data to EBI's 'big-data'
- Resources exposed using VMware's vCloud Director
 - SSL Connections to the web management interface
 - Provide isolated IaaS clouds to multiple tenant organisations

Why 'Embassy' Cloud?

- An embassy is sovereign territory in a host country
 - Host Country: EMBL-EBI Data Centre
 - Sovereign Territory: Host Country not allowed to enter
- Virtualisation provides the protection for 'tenant' and 'host'
 - Host puts boundaries in place to protect it from the tenant
 - Tenant has freedom and control within those boundaries
- Added value from EMBL-EBI over other clouds:
 - Machines and data hosted in known jurisdiction
 - File access to hosted data sets (public & managed access)
 - Direct network access to public EMBL-EBI services

Embassy Cloud

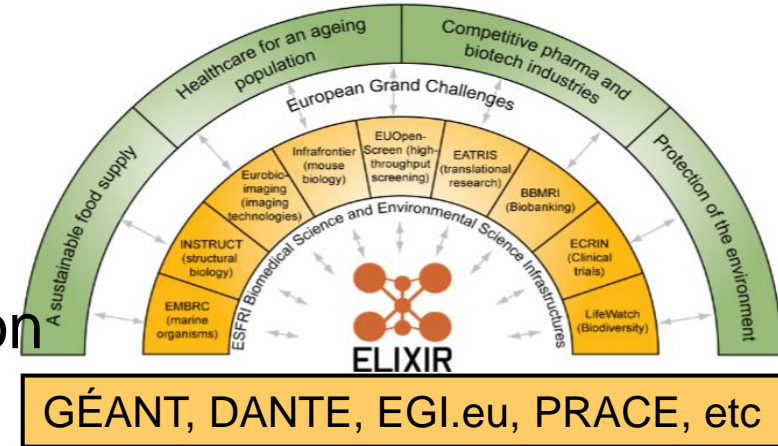


Other Cloud Activity at EMBL-EBI

- Use Amazon to provide geographical distribution
 - Direct link to globally replicate databases
- HelixNebula
 - Integration of commercial cloud providers with big research
- Benefit of additional security assurances
 - For use by pharmaceutical companies
 - For on-demand personalised medicine
- Explore using IaaS to supplement/replace data centres
 - Put DC on cloud, scale out services (service + database), etc.

The Future

- Exploitation by ELIXR
 - An e-Infrastructure for Life Science
- Technology and Science Integration
 - Physical Infrastructure
 - Software Infrastructure
- New Technology Areas
 - Use of commercial IaaS and public sector resources
 - Use of OpenStack



Any questions?

- Contact Points
 - steven.newhouse@ebi.ac.uk
- Acknowledgements
 - EMBL-EBI Systems Team