

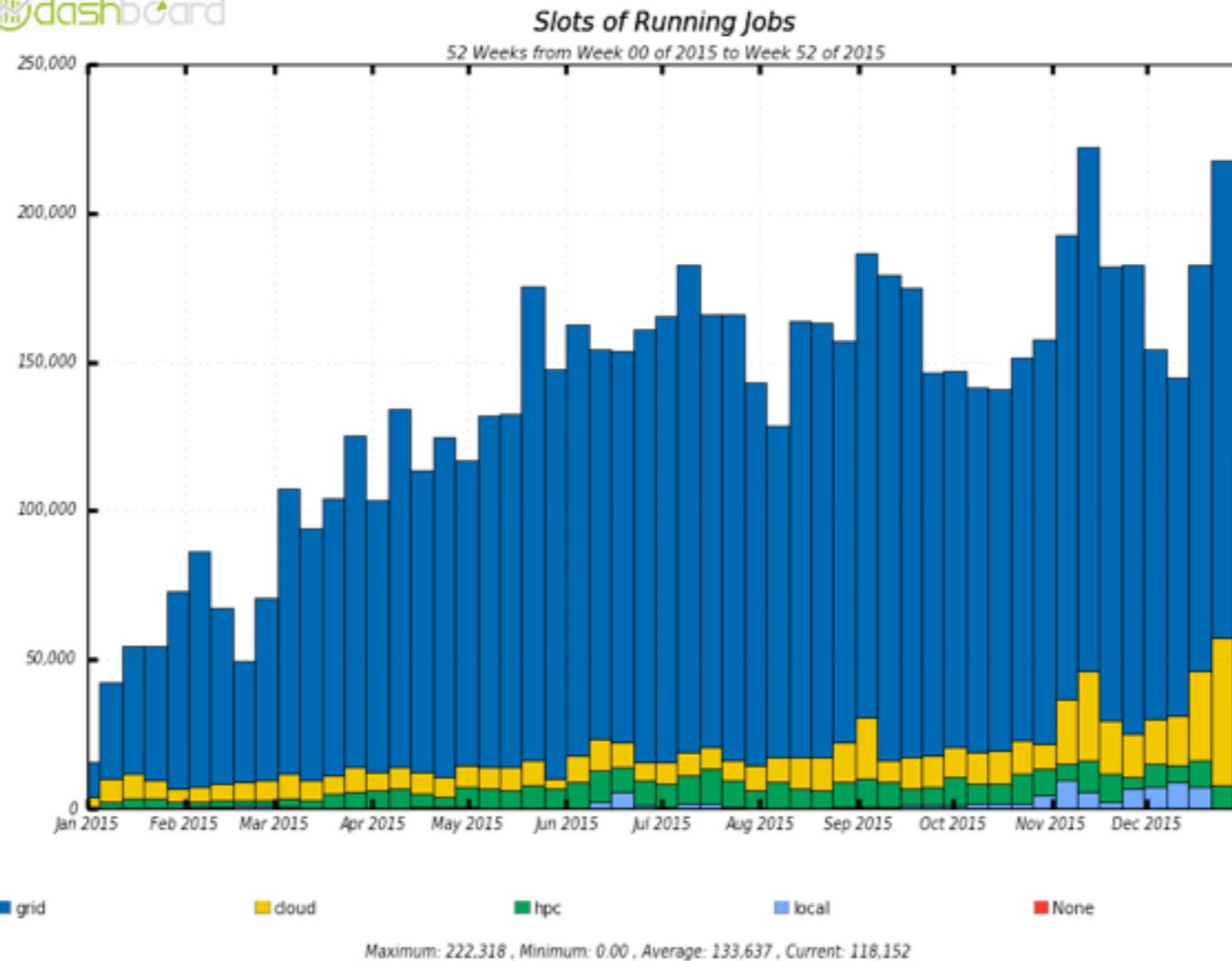
# GridPP36: New directions from ATLAS

Peter Love  
Lancaster University  
13 April 2016 - Pitlochry

# Introduction

- What does ATLAS need from GridPP?
- Compute
- Storage
- Network
- ATLAS remains agile and is capable of using a diverse range of resources, giving options to sites

# Non-Grid growth



## 4.5 Cloud computing Operations

x	R. Taylor	Jan-Dec	12	35	1.27	0.35
M	L. Field		0	0		0.00
x	M. Paterson	Jan-Dec	12	10		0.10
x	A. DaSilva	Jan-Dec	12	2		0.02
x	P. Love	Jan-Dec	12	15		0.15
x	R. Sobie	Jan-Dec	12	10		0.10
x	C. Cordeiro	Jan-Dec	12	5		0.05
x	E. Tal Hod	Jan-Dec	12	50		0.50 Boinc

FTE for 2015

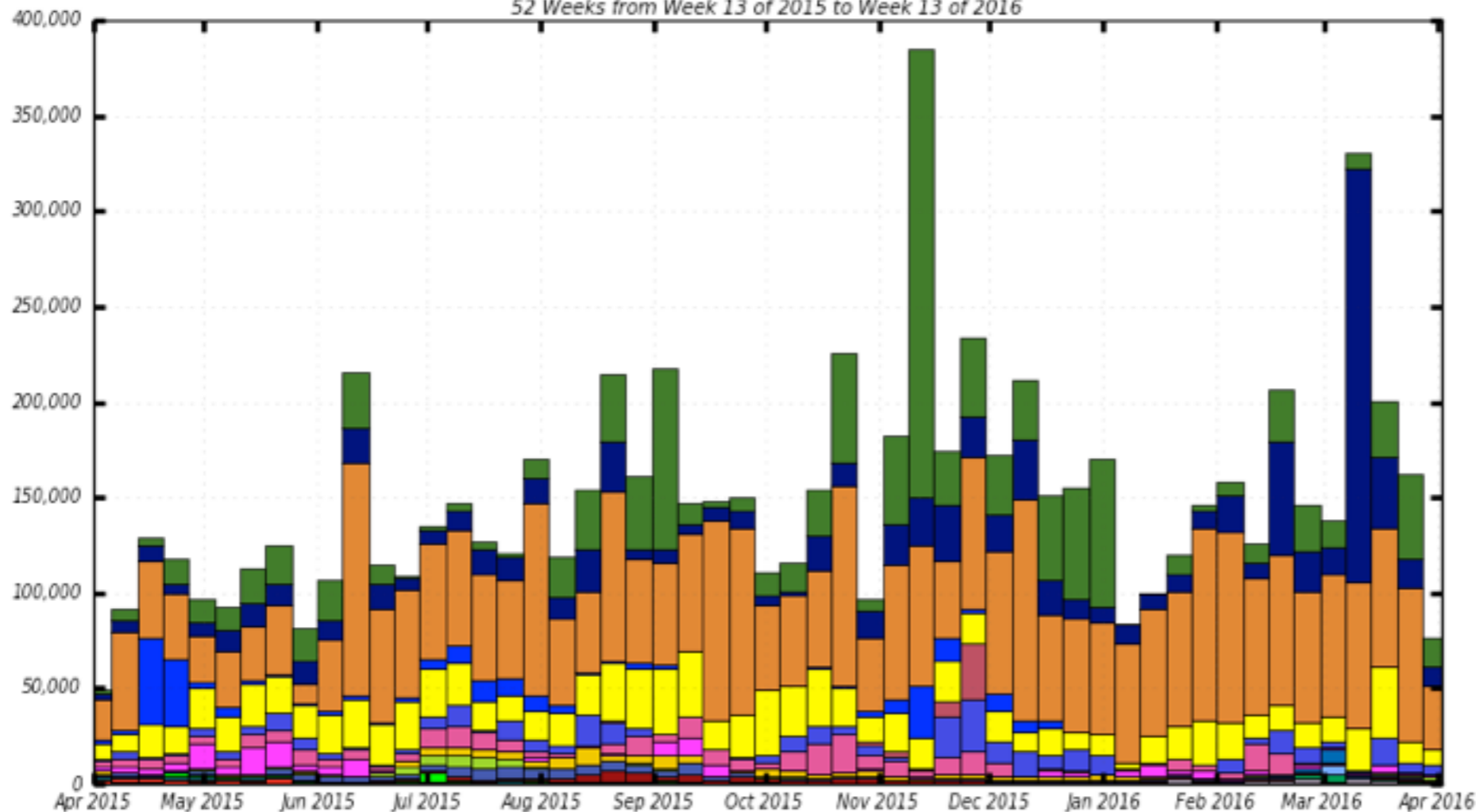
1.27

# A rainbow of VM activity



Completed jobs

52 Weeks from Week 13 of 2015 to Week 13 of 2016



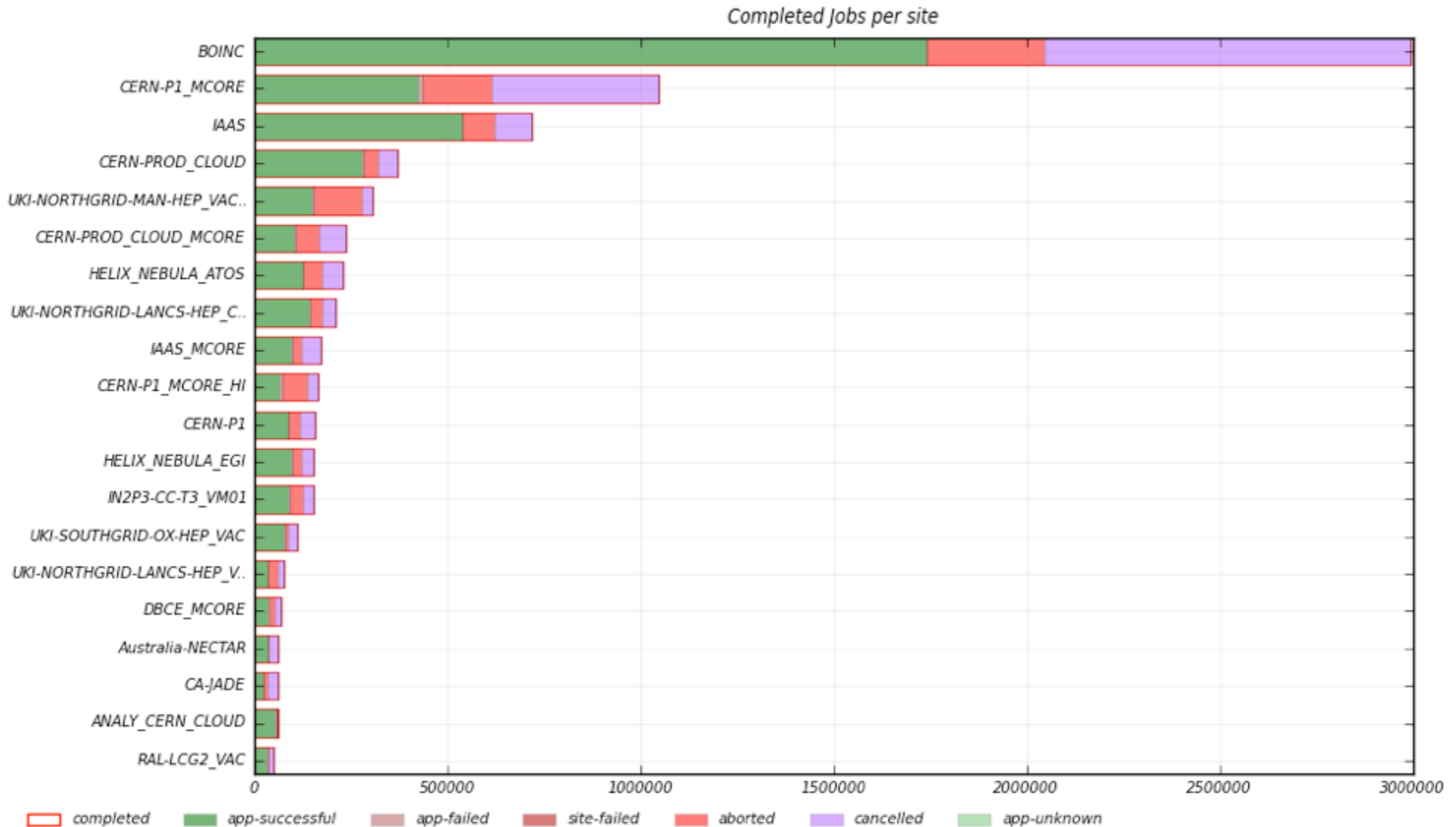
Mix of stable, ad-hoc, and experimental

- BOINC
- HLT
- Commercial
- Agile
- VAC
- Openstack



Maximum: 384,906 , Minimum: 0.00 , Average: 147,911 , Current: 76,213

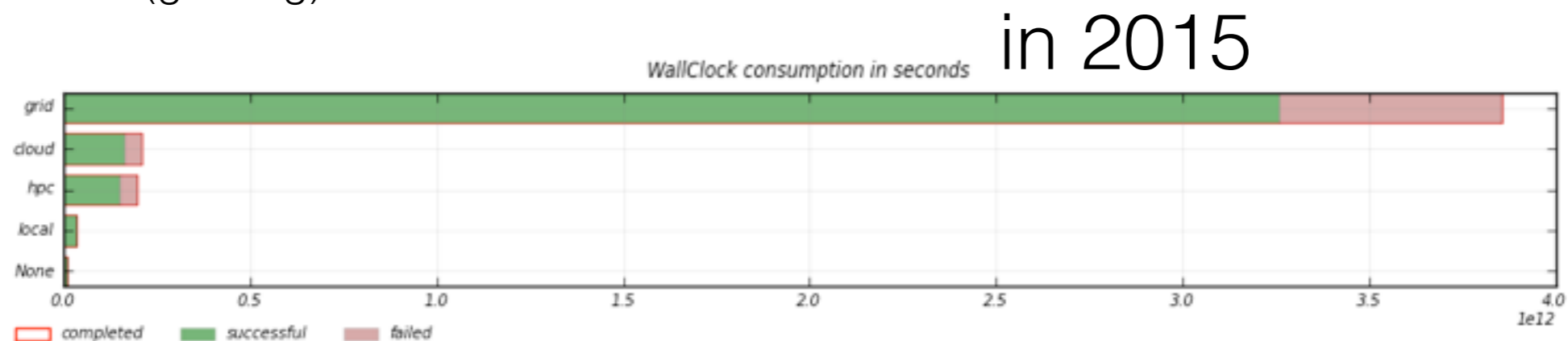
# Cloud jobs in 2015



# CPU capacities

- AWS - 40,000 (transient)
- BOINC - 10,000 (growing)
- HLT - 10,000 (stable)
- IAAS - 1500 (growing +6500 this year)
- CERN AI - 1000 (stable)
- IN2P3-CC - 700 (stable)
- VAC - 500 (growing)

Grid ~ 160k  
HPC ~ 20k  
Cloud ~ 20k

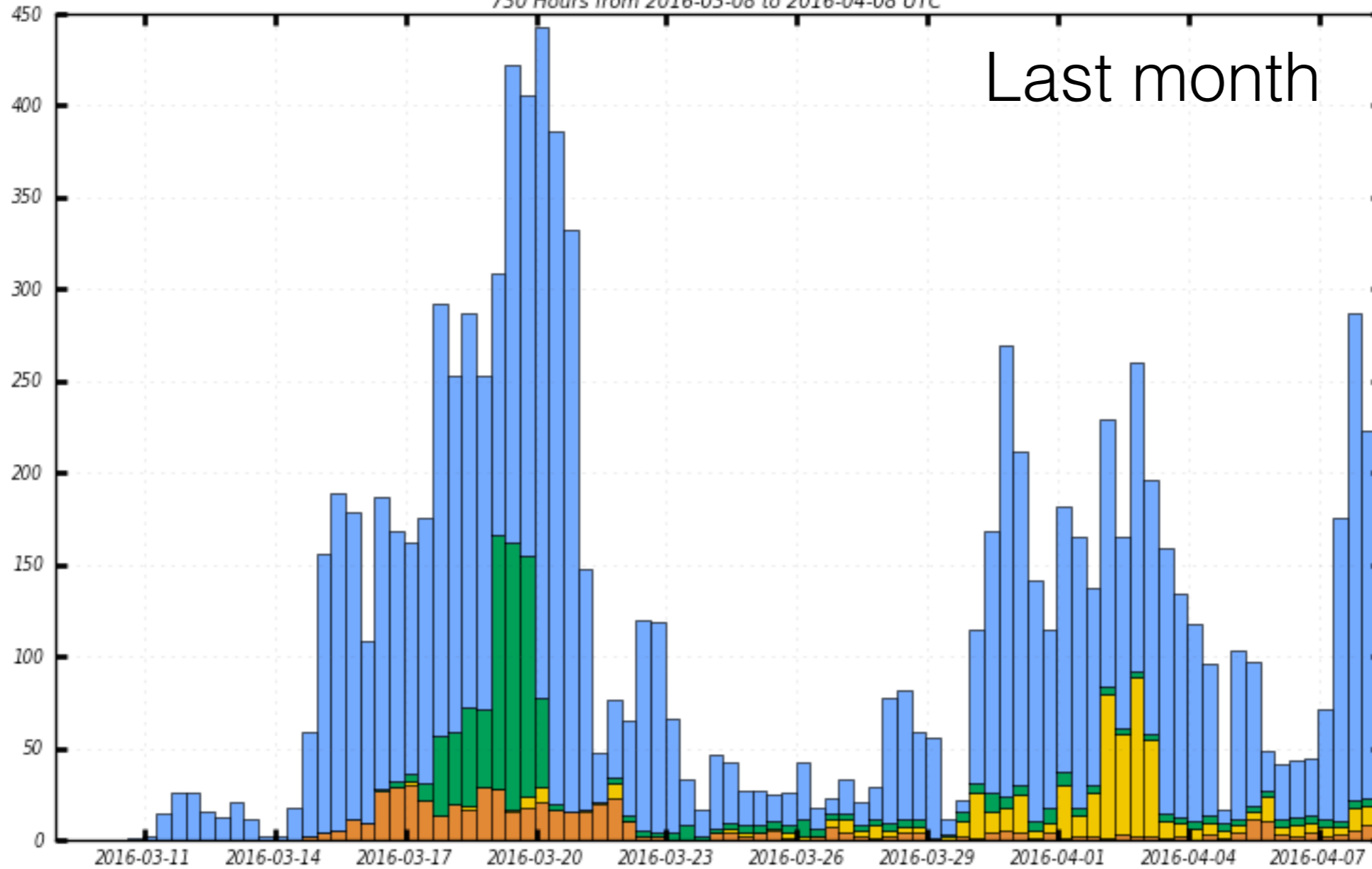


# ATLAS Vac sites



Slots of Running Jobs

730 Hours from 2016-03-08 to 2016-04-08 UTC

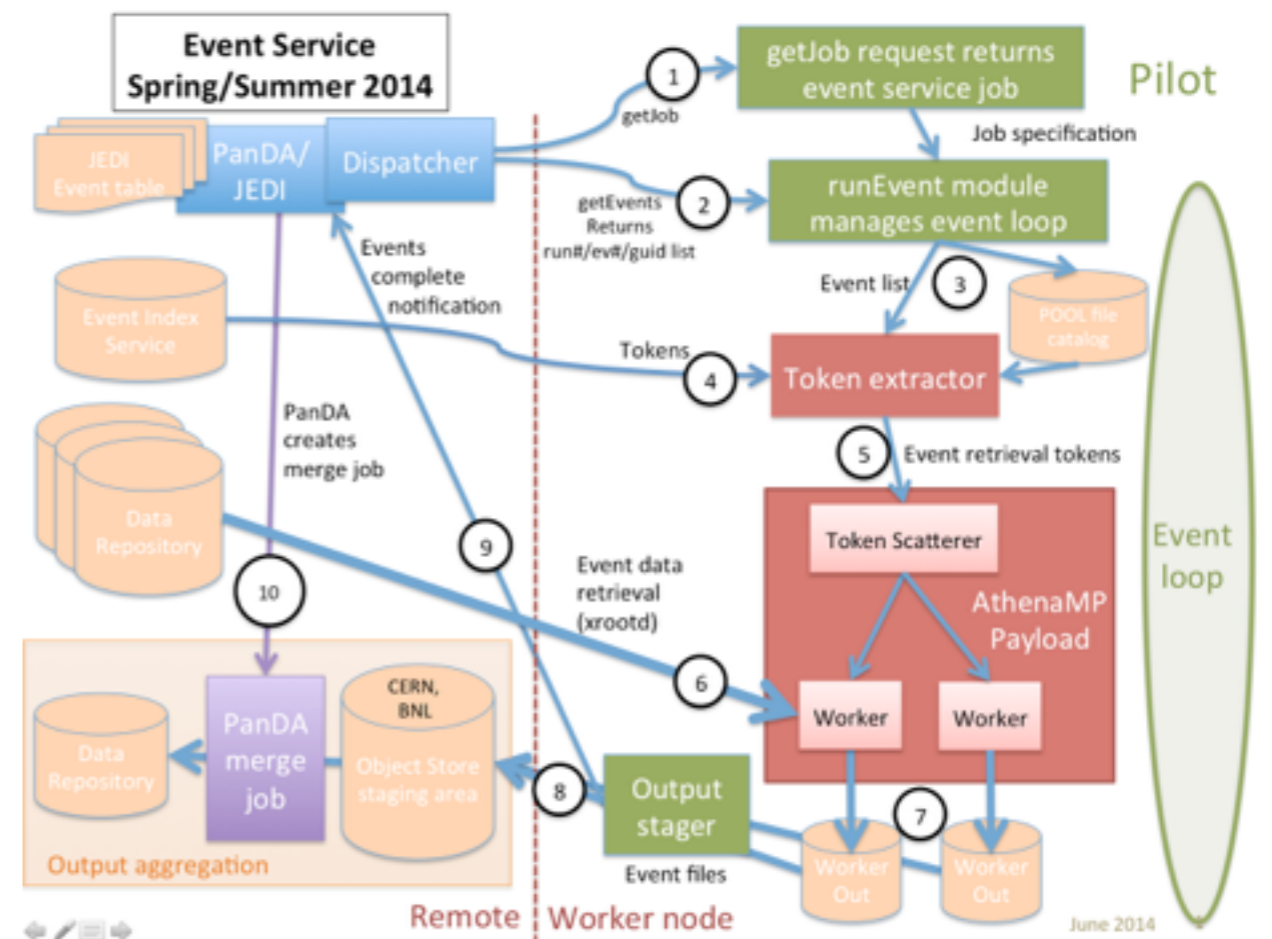


UKI-NORTHGRID-MAN-HEP\_VAC    UKI-NORTHGRID-LIV-HEP\_VAC    UKI-LT2-UCL-HEP\_VAC    UKI-SOUTHGRID-OX-HEP\_VAC  
UKI-NORTHGRID-LANCS-HEP\_VAC    RAL-LCG2\_VAC

Maximum: 442.88 , Minimum: 0.00 , Average: 112.06 , Current: 25.50

# ATLAS Event Service

- Pilot runs either normal job or AES job
- Input staged locally or read directly
- Event objects and logs sent to objectstore whilst job running
- Pilot continues to process additional event range
- Once all events done a merge job runs locally or on different resource





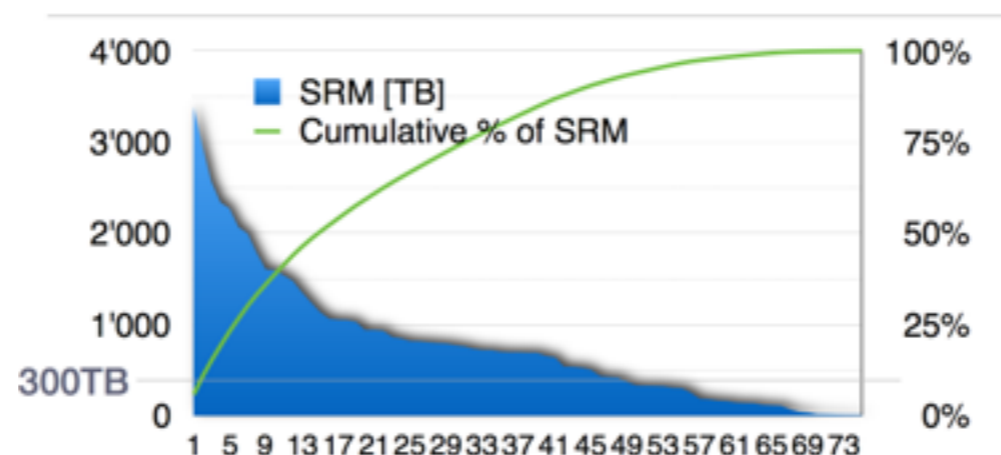
# Event Service implications

- Enables sites to provide pre-emptable slots
- Useful for opportunistic resources and backfilling existing resources
- Requires access to Object Store
- Configured in UK, running at RAL and MAN, watch this space

# Storage

- Nucleus model to help improve task throughput
- Utilize network to better use large T2 sites and move further away from the tiered model
- Satellite sites move towards cached storage and invest more into CPU
- Sites with <400TB advised to only procure CPU, <1PB to be used in more of a cache-like way

Available storage at Tier 2 sites



# Objectstores

- Used for storing log files and Event Service output
- All via S3 interface
- ATLAS currently using Objectstores at BNL, CERN, RAL, LANCs

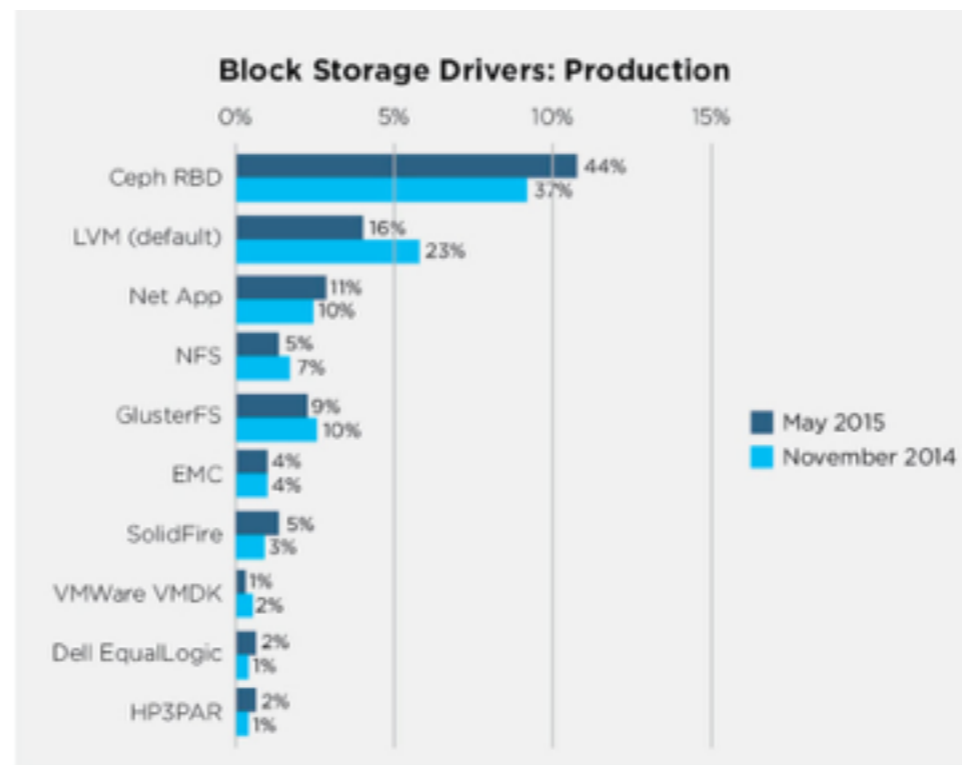


# FTS testing with OS

- RAL guys testing FTS with objectstores
- Transferring from Grid storage into OS

# Objectstores and cloud

- Openstack dominating private cloud deployment
- Ceph becoming preferred block storage
- Expect to see more Objectstore deployment coming with Openstack deployments
- <http://superuser.openstack.org/articles/openstack-users-share-how-their-deployments-stack-up>







# Network







- Defer to Duncan's talk
- ATLAS network use (WAN) increasing over Run II
- Tier2 sites at 10Gbit/s and larger T2 20Gbit/s
- Getting smarter with network usage, in a dynamic way using various analytics techniques
- Important to review usage given difficult to predict and long lead times for deployment

# Amazon Web Services






## Compute

-  **EC2** ★  
Virtual Servers in the Cloud
-  **EC2 Container Service** ★  
Run and Manage Docker Containers
-  **Elastic Beanstalk**  
Run and Manage Web Apps
-  **Lambda**  
Run Code in Response to Events




## Storage & Content Delivery

-  **S3** ★  
Scalable Storage in the Cloud
-  **CloudFront**  
Global Content Delivery Network
-  **Elastic File System** **PREVIEW**  
Fully Managed File System for EC2
-  **Glacier**  
Archive Storage in the Cloud
-  **Import/Export Snowball**  
Large Scale Data Transport
-  **Storage Gateway**  
Hybrid Storage Integration

## Database

-  **RDS**  
Managed Relational Database Service
-  **DynamoDB** ★  
Managed NoSQL Database
-  **ElastiCache**  
In-Memory Cache
-  **Redshift**  
Fast, Simple, Cost-Effective Data Warehousing
-  **DMS**  
Managed Database Migration Service








## Networking

-  **VPC**  
Isolated Cloud Resources
-  **Direct Connect**  
Dedicated Network Connection to AWS
-  **Route 53**  
Scalable DNS and Domain Name Registration

## Developer Tools

-  **CodeCommit**  
Store Code in Private Git Repositories
-  **CodeDeploy**  
Automate Code Deployments
-  **CodePipeline**  
Release Software using Continuous Delivery






## Management Tools

-  **CloudWatch**  
Monitor Resources and Applications
-  **CloudFormation**  
Create and Manage Resources with Templates
-  **CloudTrail**  
Track User Activity and API Usage
-  **Config**  
Track Resource Inventory and Changes
-  **OpsWorks**  
Automate Operations with Chef
-  **Service Catalog**  
Create and Use Standardized Products
-  **Trusted Advisor**  
Optimize Performance and Security

## Security & Identity

-  **Identity & Access Management**  
Manage User Access and Encryption Keys
-  **Directory Service**  
Host and Manage Active Directory
-  **Inspector** **PREVIEW**  
Analyze Application Security
-  **WAF**  
Filter Malicious Web Traffic
-  **Certificate Manager**  
Provision, Manage, and Deploy SSL/TLS Certificates

## Analytics

-  **EMR**  
Managed Hadoop Framework
-  **Data Pipeline**  
Orchestration for Data-Driven Workflows
-  **Elasticsearch Service**  
Run and Scale Elasticsearch Clusters
-  **Kinesis**  
Work with Real-Time Streaming Data
-  **Machine Learning**  
Build Smart Applications Quickly and Easily






## Internet of Things

-  **AWS IoT**  
Connect Devices to the Cloud








## Game Development

-  **GameLift**  
Deploy and Scale Session-based Multiplayer Games




## Mobile Services

-  **Mobile Hub**  
Build, Test, and Monitor Mobile Apps
-  **Cognito**  
User Identity and App Data Synchronization
-  **Device Farm**  
Test Android, FireOS, and iOS Apps on Real Devices in the Cloud
-  **Mobile Analytics**  
Collect, View and Export App Analytics
-  **SNS**  
Push Notification Service

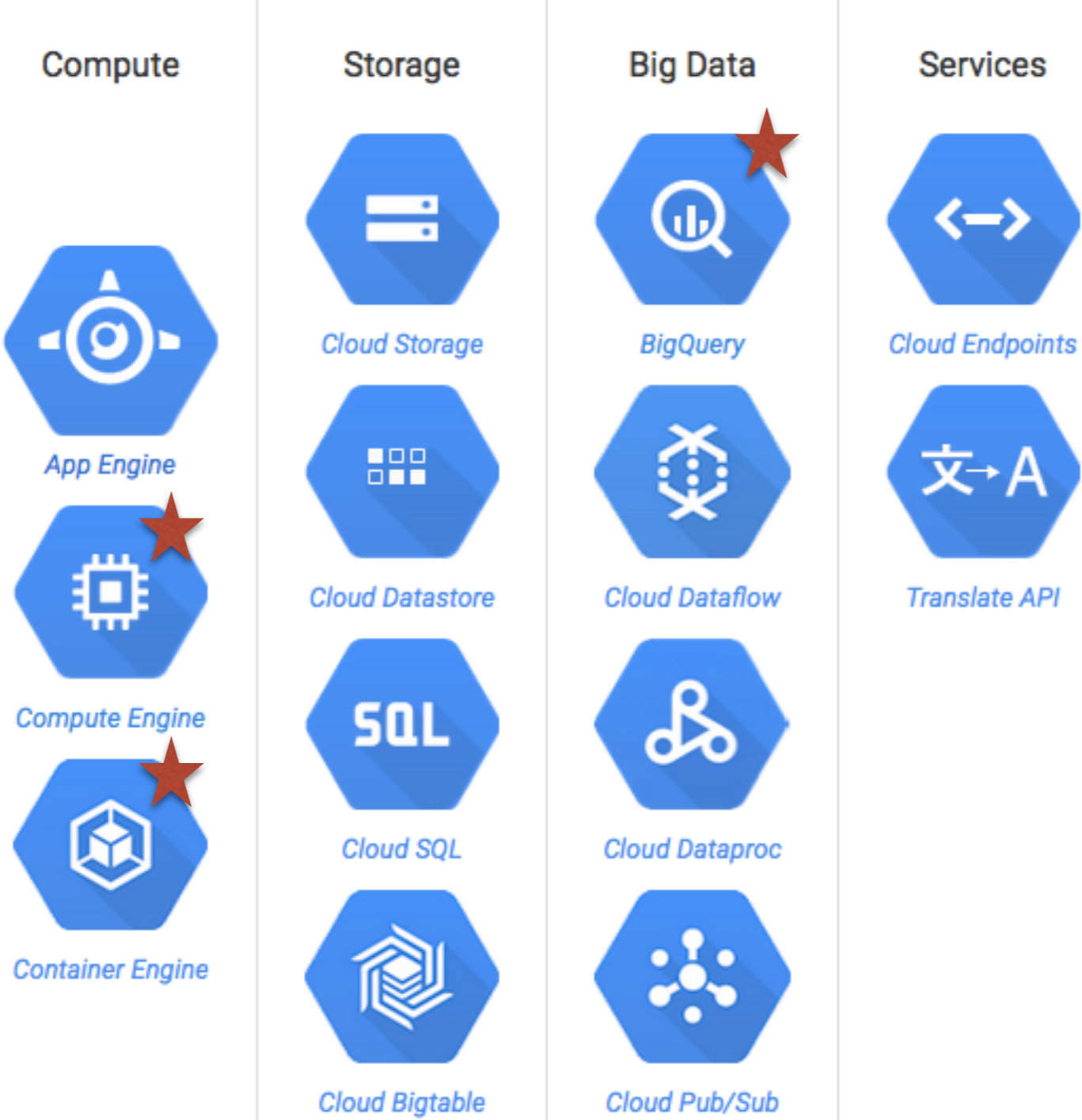
## Application Services

-  **API Gateway**  
Build, Deploy and Manage APIs
-  **AppStream**  
Low Latency Application Streaming
-  **CloudSearch**  
Managed Search Service
-  **Elastic Transcoder**  
Easy-to-Use Scalable Media Transcoding
-  **SES**  
Email Sending and Receiving Service
-  **SQS**  
Message Queue Service
-  **SWF**  
Workflow Service for Coordinating Application Components

## Enterprise Applications

-  **WorkSpaces**  
Desktops in the Cloud
-  **WorkDocs**  
Secure Enterprise Storage and Sharing Service
-  **WorkMail**  
Secure Email and Calendaring Service

# Google Cloud Platform

































Google engineering for the rest of us  
This stuff is very attractive



# Azure Cloud Platform

Popular products

<p>Virtual Machines </p>  <p>Provision Windows and Linux Virtual Machines in minutes.</p>	<p>App Service</p>  <p>Create web and mobile apps for any platform and any device.</p>	<p>SQL Database</p>  <p>Managed relational SQL Database-as-a-service.</p>	<p>Storage </p>  <p>Durable, highly available and massively scalable cloud storage.</p>
<p>Cloud Services</p>  <p>Create highly available, infinitely scalable cloud applications and APIs.</p>	<p>DocumentDB</p>  <p>Managed NoSQL document database-as-a-service.</p>	<p>Azure Active Directory</p>  <p>Synchronize on-premises directories and enable single sign-on.</p>	<p>Backup</p>  <p>Simple and reliable server backup to the cloud.</p>
<p>HDInsight</p>  <p>Provision-managed Hadoop clusters.</p>	<p>RemoteApp</p>  <p>Deploy Windows client apps in the cloud, run on any device.</p>	<p>Batch</p>  <p>Run large-scale parallel and batch compute jobs.</p>	<p>StorSimple</p>  <p>Hybrid cloud storage for enterprises, reduces costs and improves data security.</p>
<p>Visual Studio Team Services</p>  <p>Services for teams to share code, track work and ship software.</p>	<p>API Management</p>  <p>Publish APIs to developers, partners and employees securely and at scale.</p>	<p>Azure IoT Hub</p>  <p>Connect, monitor and control millions of IoT assets.</p>	<p>CDN</p>  <p>Deliver content to end users through a robust network of global data centers.</p>
<p>ExpressRoute</p>  <p>Dedicated private-network fibre connections to Azure.</p>	<p>Site Recovery</p>  <p>Orchestrate protection and recovery of private clouds.</p>	<p>Azure DNS</p>  <p>Host your DNS domain in Azure.</p>	<p>Machine Learning</p>  <p>Powerful cloud-based predictive analytics.</p>
<p>Service Fabric</p>  <p>Build and operate always-on, scalable, distributed applications.</p>	<p>Multi-Factor Authentication</p>  <p>Safeguard access to your data and apps with an extra level of authentication.</p>	<p>Visual Studio Application Insights</p>  <p>Diagnose problems and track usage to improve web and mobile apps.</p>	<p>SQL Data Warehouse</p>  <p> elastic data warehouse-as-a-service with enterprise-class features.</p>
<p>Virtual Network</p>  <p>Provision private networks, optionally connect to on-premises data centers.</p>	<p>Media Services</p>  <p>Encode, store and stream video and audio at scale.</p>	<p>Stream Analytics</p>  <p>Real-time stream processing.</p>	<p>Azure Active Directory Domain Services</p>  <p>Your domain controller as a service.</p>
<p>Event Hubs</p>	<p>Data Factory</p>	<p>Key Vault</p>	<p>Speech Service</p>

# Conclusions

- ATLAS has embraced technologies beyond Grid
- Actively delivering resources at the 10% level
- Little risk for sites to deploy new technologies (helps with new VOs)
- ATLAS continues to be agile and is keen to use all resources

