Cube Cloud as a Platform: Big Data, Apps, and Turnkey Solutions for Transportation Modeling

Colby Brown, AICP PTP
Regional Director, USA and Canada
Overview: Cloud Computing

- Computing resources that are delivered as a service over a network (the Internet)
- Properties of “true” cloud computing:
  1. Pay for infrastructure when needed
  2. Lower total cost of IT
  3. Don’t guess capacity (scalable)
  4. Increase innovation
  5. End undifferentiated heavy lifting
  6. Go global / go anywhere
- Cloud computing is everywhere in 2013!
  - Amazon, Google, Microsoft, Apple, Facebook, Netflix, Dropbox, IBM

<table>
<thead>
<tr>
<th>Cloud Clients</th>
<th>SaaS</th>
<th>PaaS</th>
<th>IaaS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web browser, mobile app, thin client, terminal emulator, ...</td>
<td>CRM, Email, virtual desktop, communication, games, ...</td>
<td>Execution runtime, database, web server, development tools, ...</td>
<td>Virtual machines, servers, storage, load balancers, network, ...</td>
</tr>
</tbody>
</table>
July 2012: Release of Cube Cloud

TRAVEL MODEL
Amazon’s EC2 Cloud Computing Environment

- Develop the Model with **Cube** in the Desktop Environment
- Publish the Model from **Cube** to the Cube Cloud
- Create, run and analyze scenarios from anywhere
Ability to ‘publish’ your model to Cube Cloud
## Benefits of Cube Cloud – Reduced Run Times

Run on 1 to 1024 processors using Cluster

### 4-step model

<table>
<thead>
<tr>
<th>Cores</th>
<th>Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20:41</td>
</tr>
<tr>
<td>4</td>
<td>11:05</td>
</tr>
<tr>
<td>8</td>
<td>7:33</td>
</tr>
<tr>
<td>16</td>
<td>5:48</td>
</tr>
<tr>
<td>32</td>
<td>4:42</td>
</tr>
<tr>
<td>64</td>
<td>4:02</td>
</tr>
</tbody>
</table>

~>80% faster

### Highway Assignment

<table>
<thead>
<tr>
<th>Cores</th>
<th>Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1:59</td>
</tr>
<tr>
<td>4</td>
<td>0:46</td>
</tr>
<tr>
<td>8</td>
<td>0:29</td>
</tr>
<tr>
<td>16</td>
<td>0:18</td>
</tr>
<tr>
<td>32</td>
<td>0:15</td>
</tr>
<tr>
<td>64</td>
<td>0:13</td>
</tr>
<tr>
<td>128</td>
<td>0:08</td>
</tr>
</tbody>
</table>

~>93% faster

### ABM Model

<table>
<thead>
<tr>
<th>Cores</th>
<th>Run Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>175:13</td>
</tr>
<tr>
<td>16</td>
<td>139:03</td>
</tr>
<tr>
<td>32</td>
<td>53:40</td>
</tr>
<tr>
<td>64</td>
<td>25:57</td>
</tr>
<tr>
<td>128</td>
<td>12:41</td>
</tr>
<tr>
<td>256</td>
<td>9:58</td>
</tr>
<tr>
<td>512</td>
<td>7:17</td>
</tr>
</tbody>
</table>

~>95% faster
Benefits of Cube Cloud - Scalability

- Start 1, 10 or 100 scenarios simultaneously—they all start immediately
- Improve validation and forecasts by doing more tests
Benefits of Cube Cloud: Sharing

- You own the model
- You invite others to use it
- No more physical copy of scripts and models
  - Eliminate onerous, mistake prone process
  - Eliminate problems with version control
  - Protect model integrity by not sharing scripts
  - Protect intellectual property by not showing scripts
- Users run the model through simple web interface
- True solution for sharing and maintaining model(s) with multiple users and for delivering a turn-key solution
- **Sharing = Value Creation**

Happy people share
### Managing Access through Admin Control Panel

<table>
<thead>
<tr>
<th>Username</th>
<th>Email</th>
<th>Pay for User Runs</th>
<th>Is Model Admin?</th>
</tr>
</thead>
<tbody>
<tr>
<td>citiboss</td>
<td><a href="mailto:mclarke@citilabs.com">mclarke@citilabs.com</a></td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>CitilabsAdmin</td>
<td><a href="mailto:amohideen@citilabs.com">amohideen@citilabs.com</a></td>
<td>☐</td>
<td>☑</td>
</tr>
<tr>
<td>colbybrown</td>
<td><a href="mailto:cbrown@citilabs.com">cbrown@citilabs.com</a></td>
<td>☑</td>
<td>☑</td>
</tr>
</tbody>
</table>
Run Scenarios with a Simple Web Interface

### Scenario Manager

**Model:** Atlanta  
**Scenario:** Base  
**Base Highway Network:** ARC0910V1.net  
**Data for “major” park-and-ride lots:** PNPOCDEOS1.DAT  
**Data for “minor” park-and-ride lots:** PNPOCDEOS2.DAT  
**Rail data including distance and speed between stations:** TRAINLNO5.TXT  
**Transit fare for each mode:** fmttrans.txt  
**Transit line data for rail lines:** TR0URGAR1.TXT  
**Transit line data for non-rail lines:** traul005.txt  
**Toll rates by time-of-day for each toll link:** TOUL095.DAT  
**Transit station nodes used for reporting stations.dbf**  
**Household SE data stratified by Income:** hwhldg.txt  
**Population and Employment SE Data:** popa029.txt  
**Extra Zone Data File:** extraZoneData029.pr  
**Pedestrian Environment Data:** PedestrianEnvironment029.txt  
**Synthesized HW Population File:** ForecastHHFile.csv  
**Synthesized Person Population File:** ForecastPersonFile.csv  
**Delta Trip Matrix for AM:** deltaAM.txt  
**Delta Trip Matrix for MD:** deltaMD.txt  
**Delta Trip Matrix for NT:** deltaNT.txt

**Help - Scenario Manager**

The Scenario Manager is where you select and schedule the scenarios you will be running. Scenarios can be managed with the provided tools which include options for accessing run results.

**Scheduling a Scenario**
1. Select your Scenario in the left sidebar  
2. Click Schedule Scenario

**Updating a Scenario**
1. Select Edit Scenario  
2. Modify values by updating text boxes or by browsing to files and click update

**Creating a Child Scenario**
1. Select Create Child Scenario  
2. Enter a Scenario Name and Description  
3. Click “Create” to create the scenario  
* Note: edit scenarios inherit key values from the parent scenario

**View Run Status and Results**
1. Select Status  
2. The scheduled and completed runs for this scenario will be listed  
3. Click view under files or reports to access the associated content  
4. Click cancel to terminate a run

**Deleting a Scenario**
1. To delete the currently selected scenario, select and confirm after clicking Delete Scenario

**Scenario Status**
1.  - no data available  
2.  - scheduled to run  
3.  - completed without warnings (return code: 0)  
4.  - completed with warnings (return code: 1)
Make it Easy to Map Results
Make it Easy to Get Charts and Tables
Improve Accessibility and Security

- Access the model through a common web browser
- Work from anywhere at anytime
- Work effectively as teams
- Protect the data and models with safe, off-site storage with redundant data backup
It’s Cheaper Than the Desktop

- Eliminate expensive hardware costs
- Spend your time on forecasting and not IT
- Cube software is included on the cloud
- Reduce costs to share and maintain the model through web-based control panels
- Reduce training costs—simple interface
- Reduce time to in analysis and communication through web maps, charts and tables
Cube Cloud Includes Sugar for ArcGIS

Cube 6

Sugar Menu
Sugar Toolbars
Feature Explorer
Data Manager
Junction Editor
Completing the workflow with Amazon and ESRI

ArcGIS Server

Sugar

SDE

Edit data

Developers

Publish models

Cube 6

Web UI

Appliers

Run models

Amazon Web Services

Cube Cloud Services

Base
“Currently, it appears unrealistic to expect all potential users of the model to invest in the computing power required to run the ABM in a short amount of time. Thus, cloud computing… offers the alternative of accessing the ARC ABM on an as-needed and pay-as-you go basis, to any eventual ABM user… the cloud-based ABM allows for virtually unlimited model runs within reasonable timelines, and relieves the burden placed upon ARC’s server infrastructure…”
Case Study: San Joaquin Valley MIP

- Eight MPOs in California’s Central Valley cooperatively working together to improve their models and meet the challenges of new SB 375 / AB 32 regulations
- A communal Cube Cloud model hosting plan, managed by Fresno COG, providing reserve capacity for all the MPOs
- Before: have to corral as many physical machines as possible to meet deadlines
- After: set up runs on Cloud, click and go
- Client testimonial: would use Cube Cloud again for time-critical model run situations
At Cube Cloud store: nationwide block-level socio-economic data for the USA (free!)

Also available: O-D trip matrix tables from cell phone signal data gathered by AirSage

Now available: NAVTEQ™ nationwide roadway network and traffic database
Free apps provide access to commonly used analytic tools. Comes standard with Cube Cloud!

Paid apps allow users to “sell” their models on the Cloud…
# Cube Cloud Store: Resources

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Category</th>
<th>Type</th>
<th>Region</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Household Travel Survey (NHTS)</td>
<td>Comprehensive data on travel and transportation patterns in the USA</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Public Use Microdata Sample (PUMS)</td>
<td>Public Use Microdata Sample files from the American Community Survey - US Census</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>American Community Survey (ACS)</td>
<td>Ongoing survey that provides data every year from the US Census</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Census Transportation Planning Package (CTPP)</td>
<td>Census data specific to transportation analysis</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>CTPP Data on the web</td>
<td>CTPP Data available on the web</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Longitudinal Employer-Household Dynamics (LEHD)</td>
<td>Census employment data</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Quarterly Census of Employment and Wages</td>
<td>Bureau of Labor Statistics – Quarterly Census of Employment and Wages</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Census Population forecast estimates</td>
<td>Census Population forecast estimates</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Travel Estimation Techniques for Urban Planning - nCHRP 305</td>
<td>Travel estimation techniques and parameters from TRB</td>
<td>Report</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Online Travel Survey Manual</td>
<td>Guide to travel survey design</td>
<td>Report</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Characteristics of Urban Transportation Demand (TCRP Report 73)</td>
<td>Guide to urban travel demand characteristics</td>
<td>Report</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Batch geo-coding API (Choose Yahoo or Google)</td>
<td>Useful tool to geocode multiple addresses</td>
<td>Tool</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>ArcGIS Base Maps for download</td>
<td>Base maps for ArcGIS</td>
<td>Tool</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>2011 TIGER/Line Shape files</td>
<td>Place to download TIGER files</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Florida Geographic Data Library (FGDL)</td>
<td>Downloadable geographic data for Florida</td>
<td>Data</td>
<td>Web Link</td>
<td>United States</td>
<td>link</td>
</tr>
<tr>
<td>Library of German Transport data</td>
<td>Downloadable transportation data for Germany</td>
<td>Data</td>
<td>Web Link</td>
<td>Germany</td>
<td>link</td>
</tr>
<tr>
<td>German Road Statistics</td>
<td>Downloadable statistics for German roadways</td>
<td>Data</td>
<td>Web Link</td>
<td>Germany</td>
<td>link</td>
</tr>
<tr>
<td>Bavarian Road Statistics</td>
<td>Downloadable statistics for Bavaria</td>
<td>Data</td>
<td>Web Link</td>
<td>Germany</td>
<td>link</td>
</tr>
<tr>
<td>Austrian Road Statistics</td>
<td>Downloadable statistics for Austria</td>
<td>Data</td>
<td>Web Link</td>
<td>Austria</td>
<td>link</td>
</tr>
<tr>
<td>Swiss Road Statistics</td>
<td>Downloadable statistics for Switzerland</td>
<td>Data</td>
<td>Web Link</td>
<td>Switzerland</td>
<td>link</td>
</tr>
<tr>
<td>Census data for England</td>
<td>Downloadable census data for England</td>
<td>Data</td>
<td>Web Link</td>
<td>United Kingdom</td>
<td>link</td>
</tr>
</tbody>
</table>
Turnkey Modeling Solutions

- Quick-response tools for traffic forecasting
- Standard models developed by Citilabs
- Minimal data requirements and default parameters
- Models are calibrated and published to Cube Cloud for scenario analysis by users
- May be useful in places that do not have an MPO model
Turnkey Model Data Collection

- Used 2010 Census Demographic Data and NAVTEQ Network Data
- Used NCHRP 7-16 Parameters for Trip Generation Rates
- TAZs are Census Block Groups (65)
- Centroids created by Cube
- Counts at gateways obtained from CT DOT and recent traffic studies
- Transit Network and stop nodes – manual or GTFS
- Results achieved with minimal customization
- Quick response time – process of weeks, rather than months
**Turnkey Model Required User Inputs**

“Standardized” Available:
- Network Data
- Demographic Data
- TAZ Boundaries
- Trip Generation Rates

“Region Specific”
- Traffic Counts at Gateways
- Other Traffic Counts
- Potential XX Routes
- Mode Share Targets
- Transit Info
- Special Generators

**Cube Turnkey Three-Four Step Model - Basic Keys (Page 2)**

<table>
<thead>
<tr>
<th>Keys With Default Values (review of these is recommended)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hour corresponding to AM peak</td>
</tr>
<tr>
<td>Hour corresponding to PM peak (15=3PM, 16=4PM, etc...)</td>
</tr>
<tr>
<td>Minimum Daily Traffic Volume for a road to be included in validation</td>
</tr>
<tr>
<td>Average Auto Occupancy Used For Productions in an Auto-only Model</td>
</tr>
<tr>
<td>Speed of Walking (MPH)</td>
</tr>
<tr>
<td>Maximum Walking Distance (Miles)</td>
</tr>
<tr>
<td>Driving Cost (cents per gallon)</td>
</tr>
<tr>
<td>Additional Distance Traveled on Corpod Trics (miles)</td>
</tr>
<tr>
<td>Average Occupancy of Carpool</td>
</tr>
<tr>
<td>Additional Driving Time Required for Carpooling (minutes)</td>
</tr>
<tr>
<td>Factor to apply to transit travel time</td>
</tr>
</tbody>
</table>

**NOTE:** Many other inputs (e.g. mode choice and assignment model constants) can be controlled or left at “default” values
## Daily Link Validation Results

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mountain View, CA</th>
<th>West Hartford, CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Modeled Volume / Count Ratio</td>
<td>1.17</td>
<td>1.18</td>
</tr>
<tr>
<td>% RMSE</td>
<td>27%</td>
<td>52%</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>0.98</td>
<td>0.77</td>
</tr>
<tr>
<td>Pct. of Links Within “allowable” deviation</td>
<td>62%</td>
<td>51%</td>
</tr>
</tbody>
</table>
Turnkey Models on Cube Cloud Services

- No desktop software
- Manage scenarios
- Display maps, tables
- Model-sharing
Growth Factor Model Process

Starting Trip Matrix → Change in Traffic → Traffic Counts → Change in Network → Change in Land Use

Other Results: Speeds, Travel Times, Levels of Service, Queues, Simulation, etc…
PM Peak Hour Link Validation in West Hartford With Growth Factor Tool

<table>
<thead>
<tr>
<th>Measure</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Modeled Volume / Count Ratio</td>
<td>1.21</td>
<td>0.97</td>
</tr>
<tr>
<td>% RMSE</td>
<td>84%</td>
<td>23%</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>0.71</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Compared *modeled* PM Peak Hour traffic to *counts* on 72 “Study Links” from the 2009 Blue Back Square traffic study

“Before” = output of Turnkey Three Step Trip Based Model
“After” = with Cube Analyst applied in Growth Factor Tool
Results Network

- Shows difference between “What If” and “Base” scenario traffic volumes
- Shows Turn movements
- Will compute standard LOS results, export to Synchro / SimTraffic
Sample Scenario – West Hartford Bishop’s Corner
Added 75 ksf retail + 100 ksf office
Why Use Growth Factor Model?

- Versatile – can do TIAs or small city corridor / general plan studies
- Eliminates need for ‘guessing’ trip distribution; it’s based on an actual trip matrix
- Background traffic is specified as zone-to-zone, also based on same actual trip matrix
- Data entry process is easy, less error-prone
- Allows visual, map-based network editing
- ITE Trip Generation Component, includes internal capture from NCHRP 8-51
Growth Factor and Simulation Tool

- Level of Service Reports
- Dynamic Traffic Assignment & Mesoscopic Simulation
- HCM Intersection Modeling

Additional Data Requirements

- Roadway Geometry
- Signal Timings
- Queueing Observation Data
- Level of Service Definitions / reporting requirements
Thank you!