

Provenance for Interactive Visualizations

Fotis Psallidas
fotis@cs.columbia.edu



Eugene Wu
ewu@cs.columbia.edu

Introduction

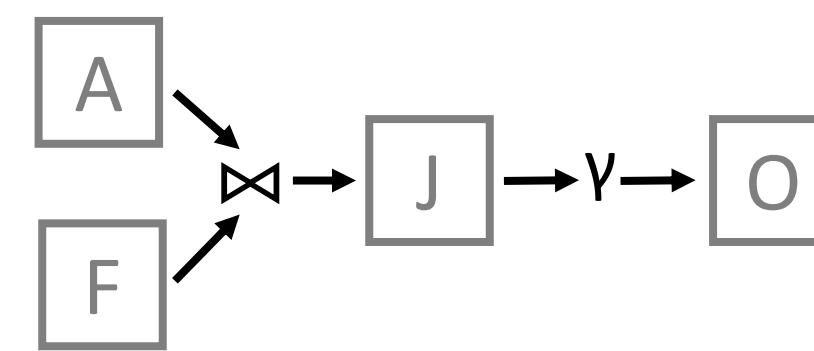
Recent research has shown that provenance-enabled data systems can be fast enough for interactive responsiveness

- Deep connections between fine-grained provenance and interaction
- Can express core visualization interactions
- Extends to general interactive applications

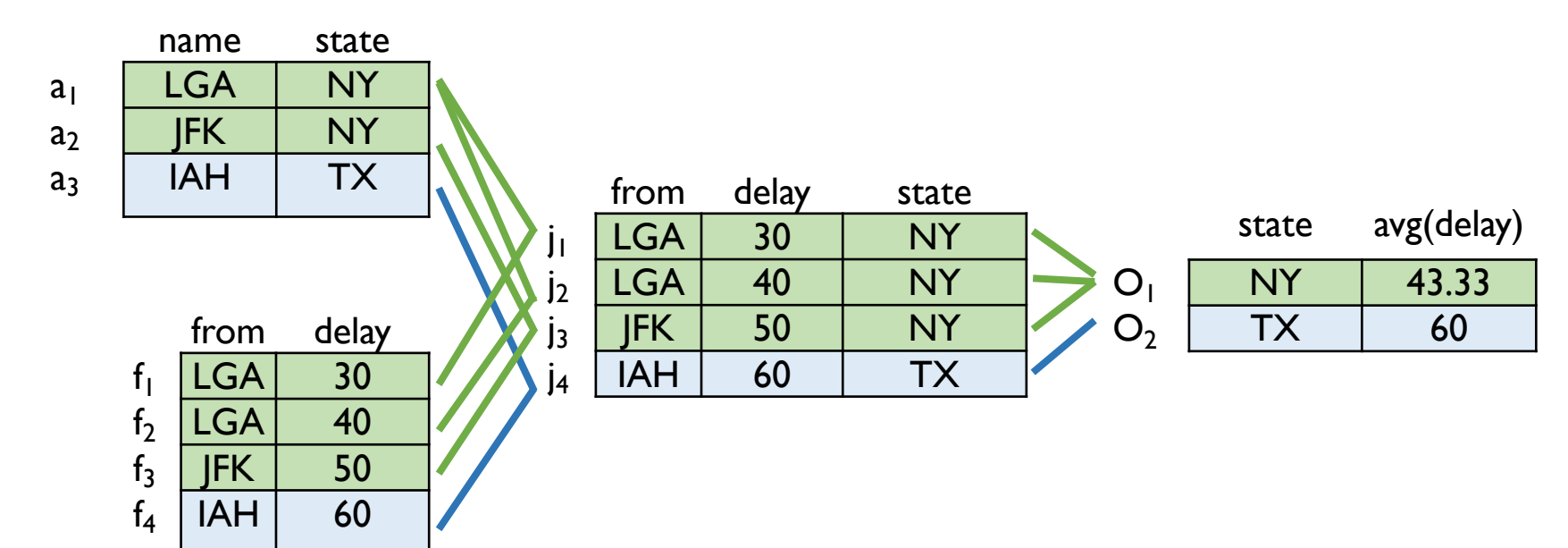
Expressing interactions using provenance enables holistic optimization

Lineage Primer

$\gamma_{state, avg(delay)}(Airports \bowtie Flights)$
name = from



Lineage Graph



Provenance For Interactive Visualizations

Interactive Selections

- Item selection
- **Group selection**
- Range selection
- Generalized selections

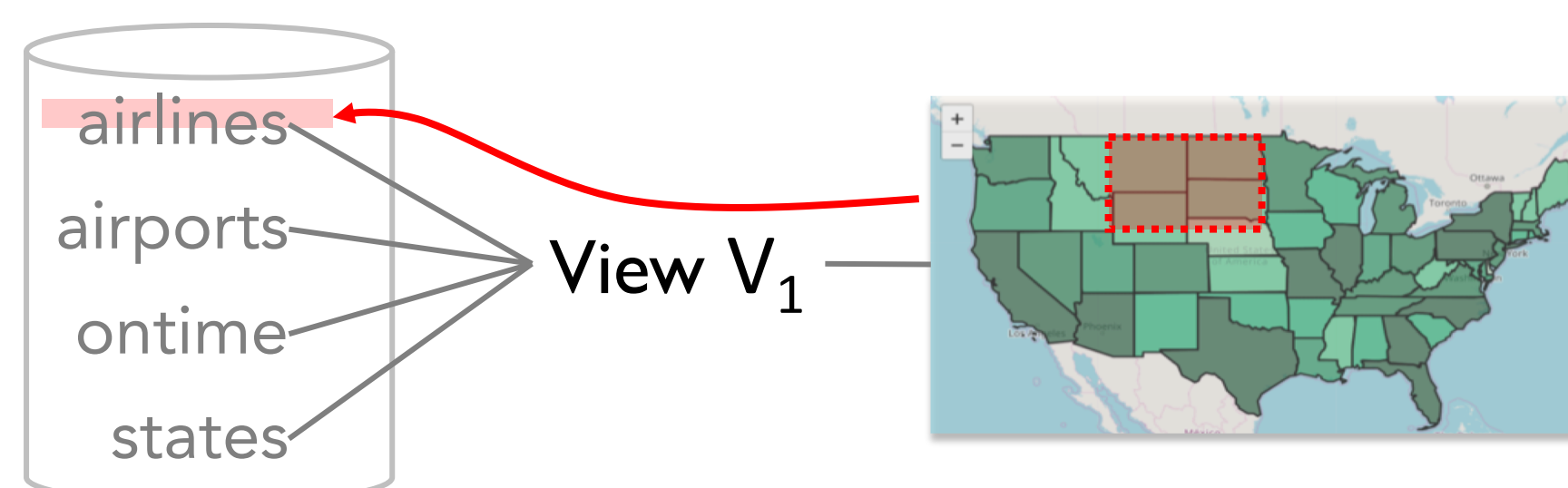
Logic over Selections

- **Tooltips**
- Details-On-Demand
- Semantic Zooming

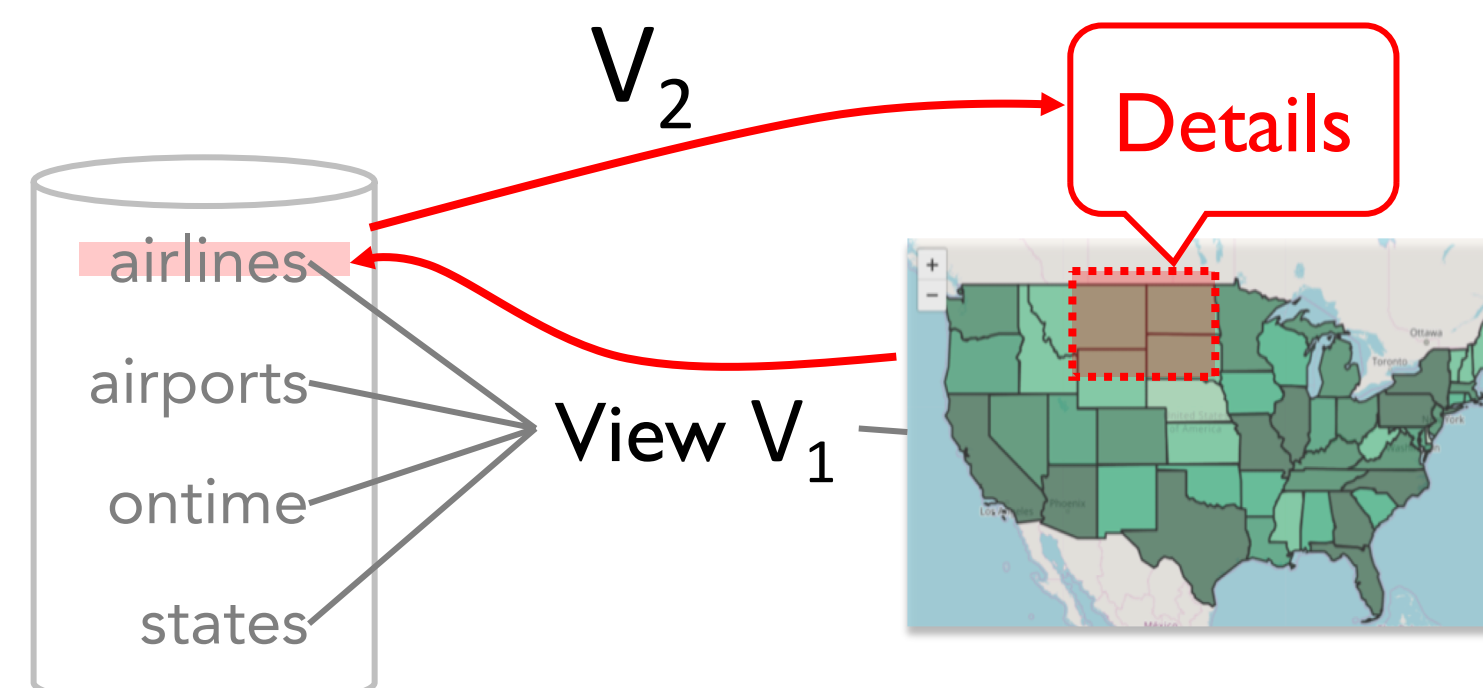
Multi-View Linking

- **Linked Brushing**
- Crossfilter

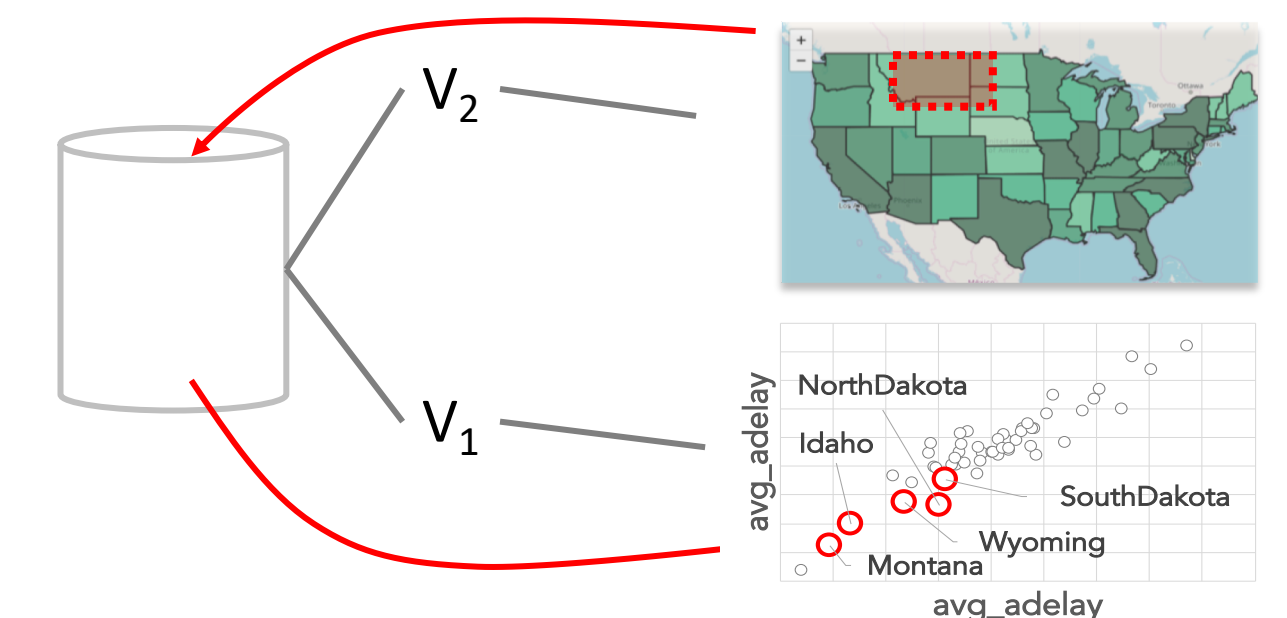
backward_trace()



SQL(backward_trace())



selective_refresh(backward_trace())



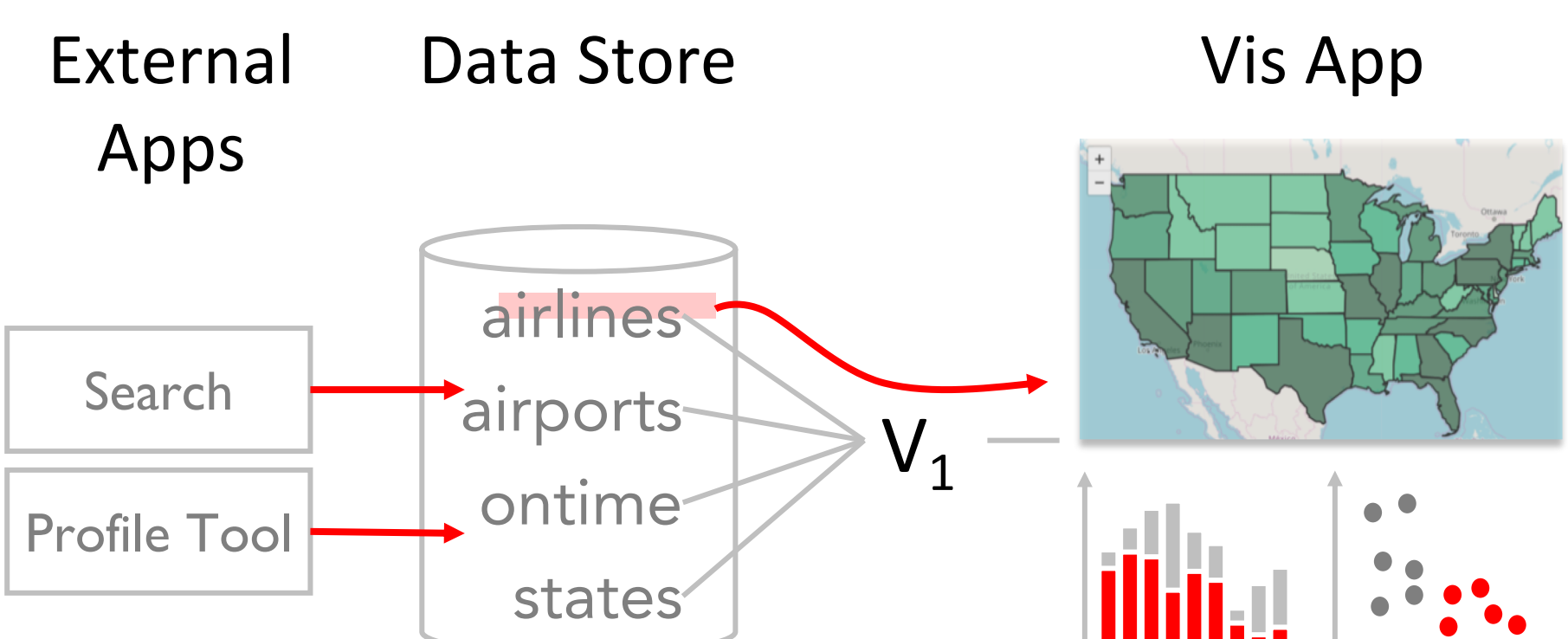
Provenance For Interactive Applications

Provenance management systems can provide

- core functionality across application domains and
- novel functionality to extend the space of interactive applications

Interactive Visualizations	Replication and Reproducibility	Multi-Application Linking	Interactive Data Profiling	What-if Provisioning
Query Explanations	Why-not Analytics	Iterative Analytics	Viz Workflow Debugging	Interactive Data Cleaning
Interaction Debugging	ML Interpretability	Visualization Deconstruction and Restyling	Interaction By Example	Application Design Search
Collaborative Communication	Action Recovery	Sense-Making	Meta-Analysis	Interactive Query Specification
Data Integration	Debugging	Network Diagnostics	Resource Scheduling	Auditing

Example: Multi-Application Linking



Want to Know More?

- [VLDB18] Smoke: Fine-Grained Lineage At Interactive Speed
- [SIGMOD18] A Deep breath of Data-Intensive Lineage Applications
- [HILDA18] Provenance for Interactive Visualizations**
- [CIDR17] Combining Design and Performance in a DVMS