LiveRAC - Live Reorderable Accordion Drawing

Peter McLachlan
Tamara Munzner

Stephen North
Eleftherios Koutsofios

CHI 2008, Florence, Italy

Domain and methodology problems

- Domain problem: managing large numbers of networked computer systems
  - can information visualization help solve this problem?
- Methodology problem: applying user centered design principles in a production setting
  - "know thy user", "involve users early"
    - not always possible!
      - in corporate environments, access to users needs management approval

Contributions

- Case study of phased approach
  - halfway between full-scale user-centered design and “working in a vacuum”
- Visualization system: LiveRAC
  - novel, scalable time-series visualization
- Informal qualitative study of LiveRAC in a production setting
  - encouraging feedback on our design approach and visualization design principles

Design target

- Assist network operations centre staff with their information needs
  - managed hosting service: shepherd outsourced IT infrastructure

Network devices & monitored data

- Time series data about system status collected from network devices
- Sample data:
    - [10 AUG 2007 9:52:37, CPU, 95%]

Network devices & monitored data

- Time series data about system status collected from network devices
- Sample data:
    - [10 AUG 2007 9:52:37, CPU, 95%]

Network devices & monitored data

- Time series data about system status collected from network devices
- Sample data:
  - [10 AUG 2007 9:52:37, CPU, 95%]

Network devices & monitored data

- Time series data about system status collected from network devices
- Sample data:
  - [10 AUG 2007 9:52:37, CPU, 95%]

System management professional job description & activities

- Plan, monitor and act
- Activities:
  - interpreting network environment status
  - capacity planning
  - event investigation
  - coordinate between customers, engineering & operations

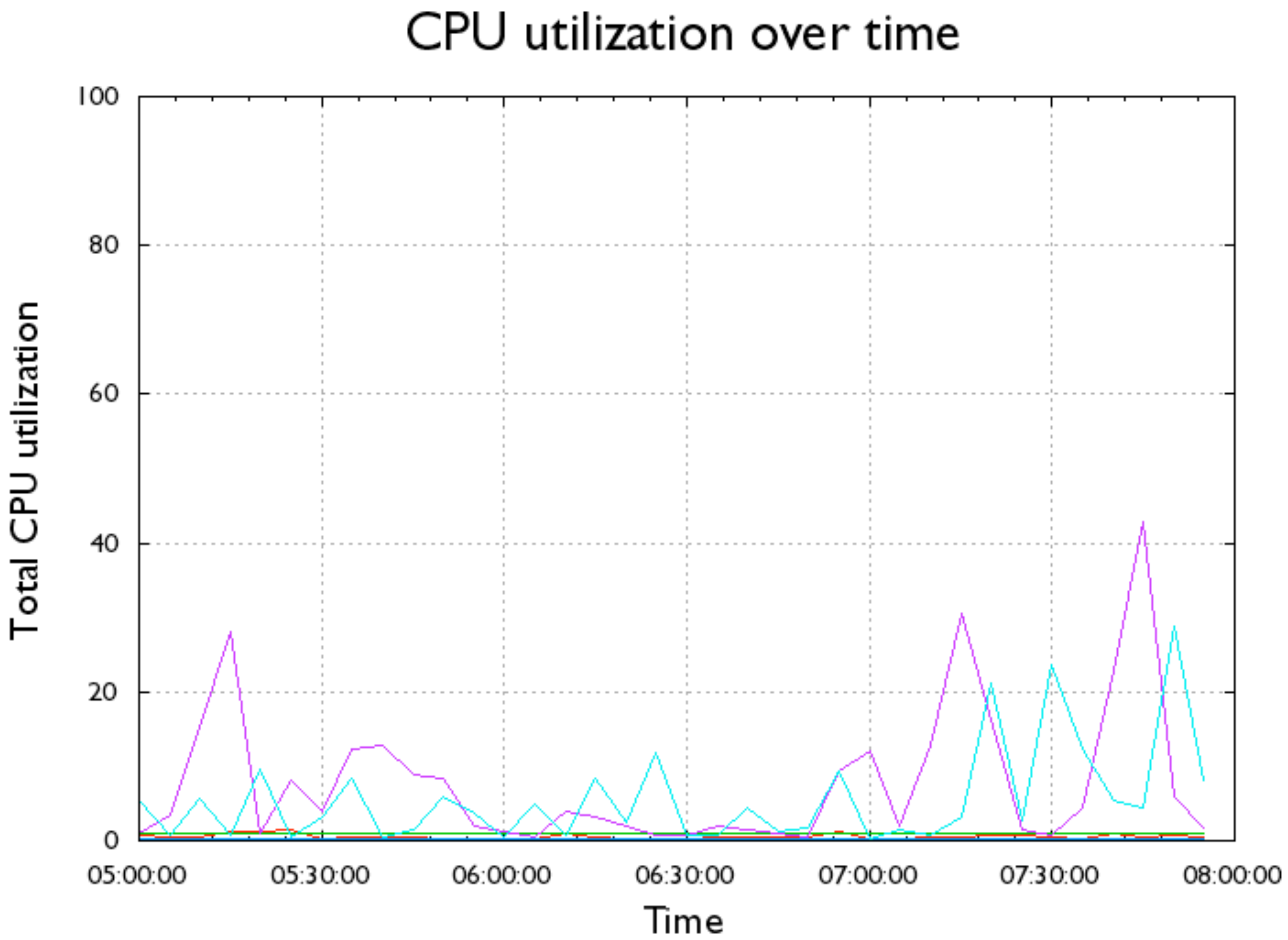
Limitations of current tools

- Most significant limitation: lack of mid-level overviews
  - high level overviews are useful, but have limitations
  - what do these numbers mean? which systems are up, which are down? how important are they? which customers are affected?

Percentage change over past 24 hours		
Categories	**Outages**	**Availability**
Network Interfaces	0 of 18	99.621%
Web Servers	0 of 27	99.526%
Email Servers	0 of 10	96.802%
DNS and DHCP Servers	0 of 6	99.954%
Database Servers	0 of 1	100.000%
JMX Servers	0 of 0	100.000%
Other Servers	1 of 18	92.659%

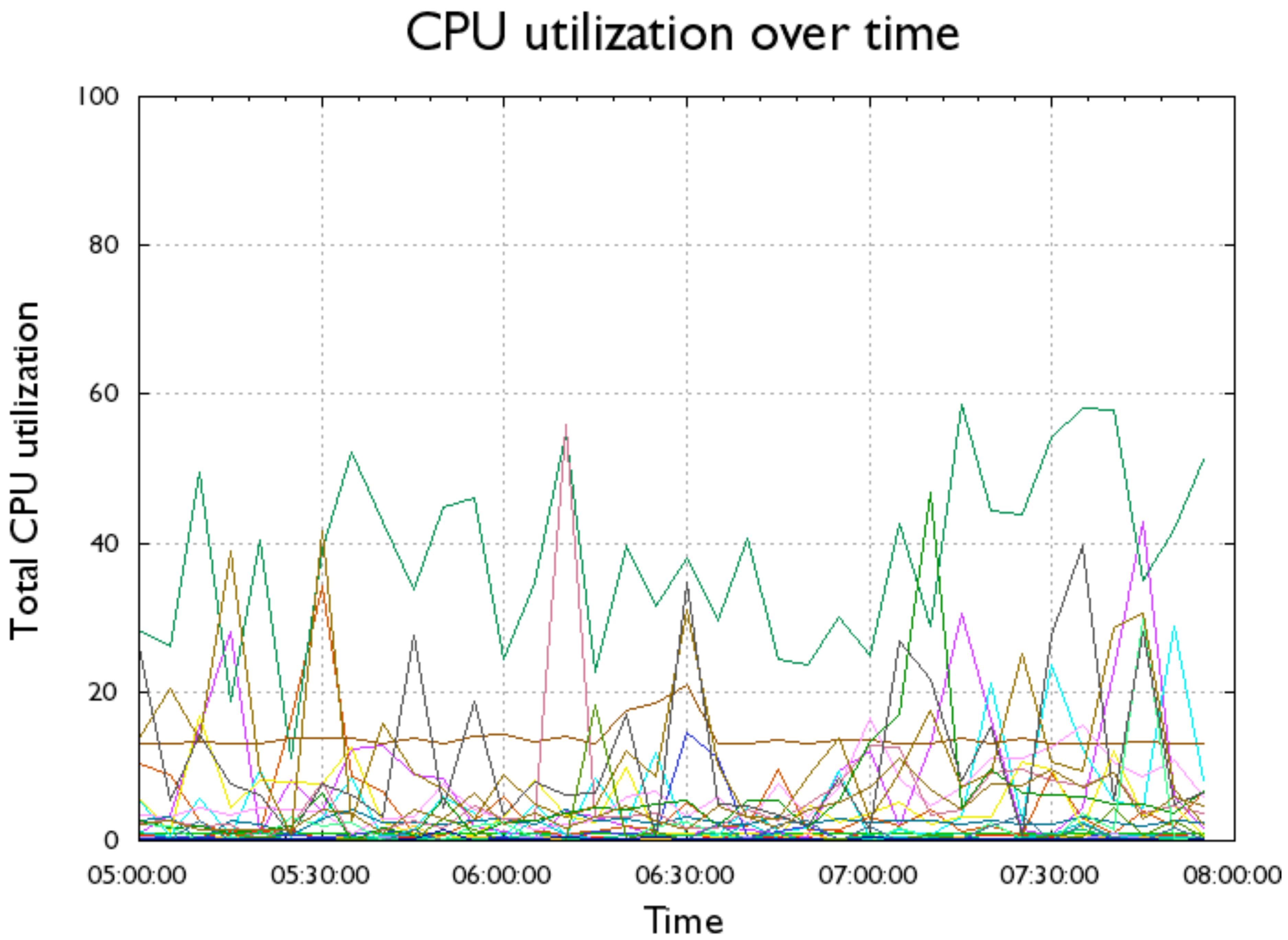
OpenNMS enterprise network management system dashboard

Monitoring time-series: is there really a problem?



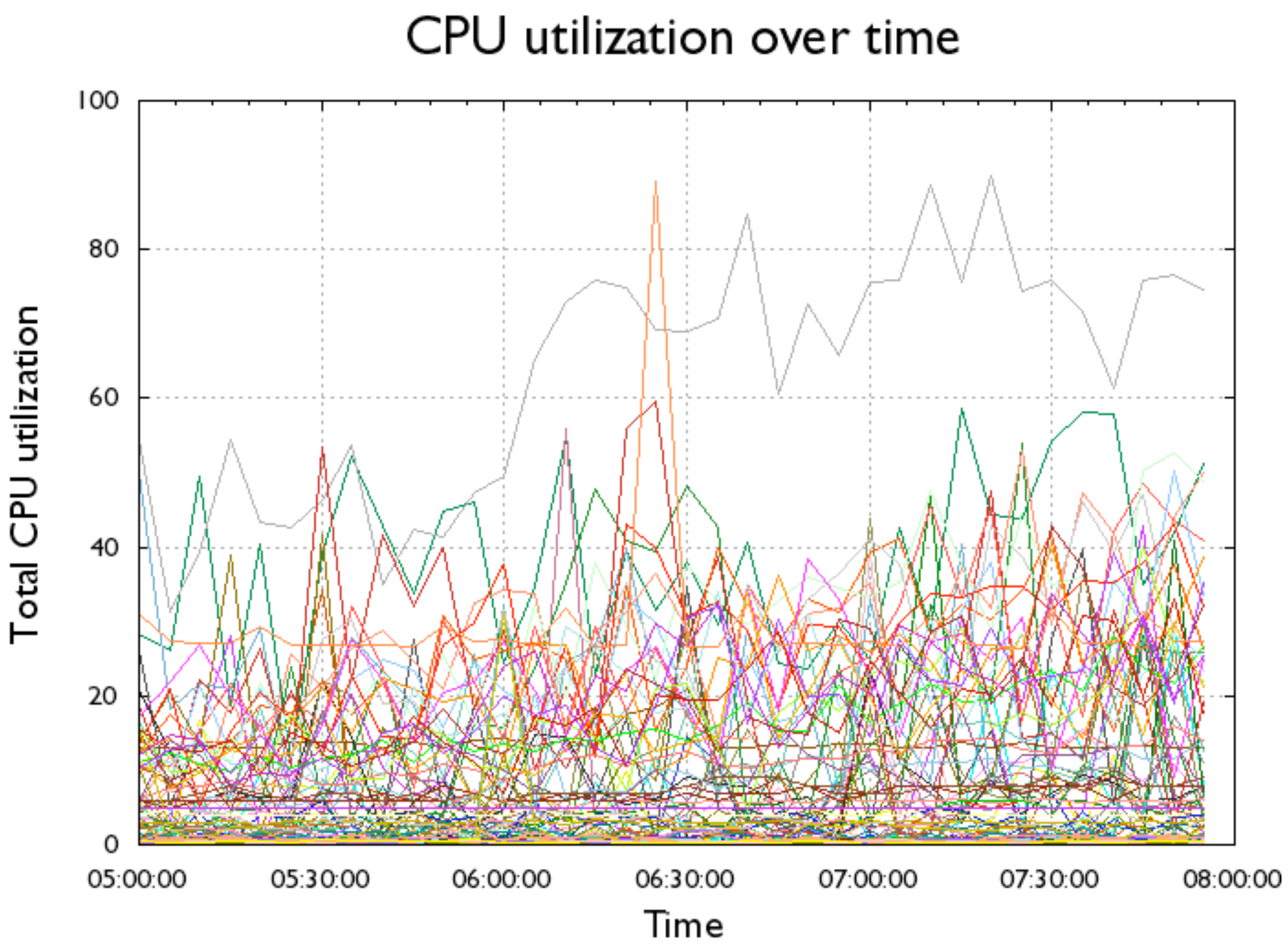
good . . .

Monitoring time-series: is there really a problem?



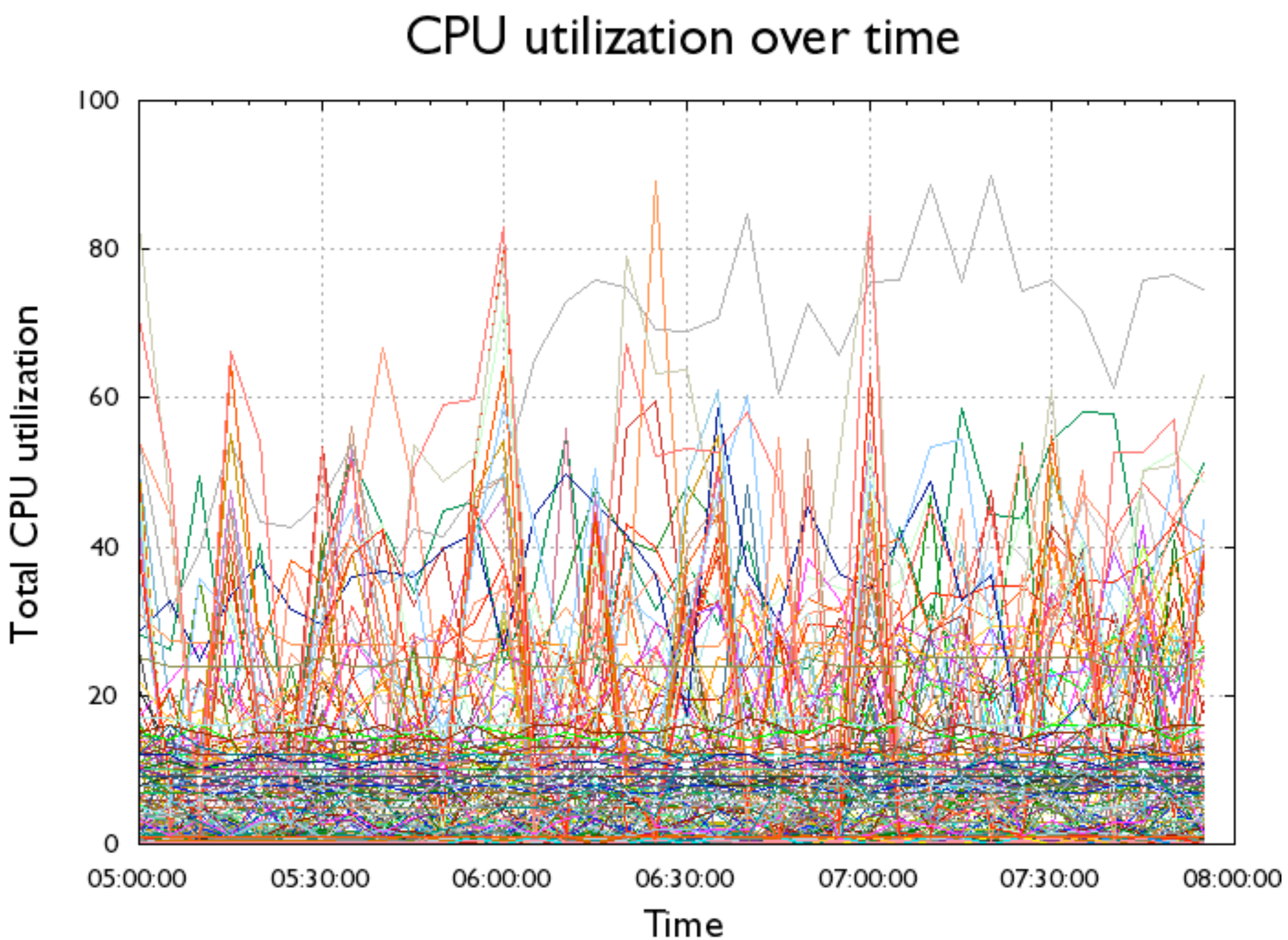
ok . . .

Monitoring time-series: is there really a problem?



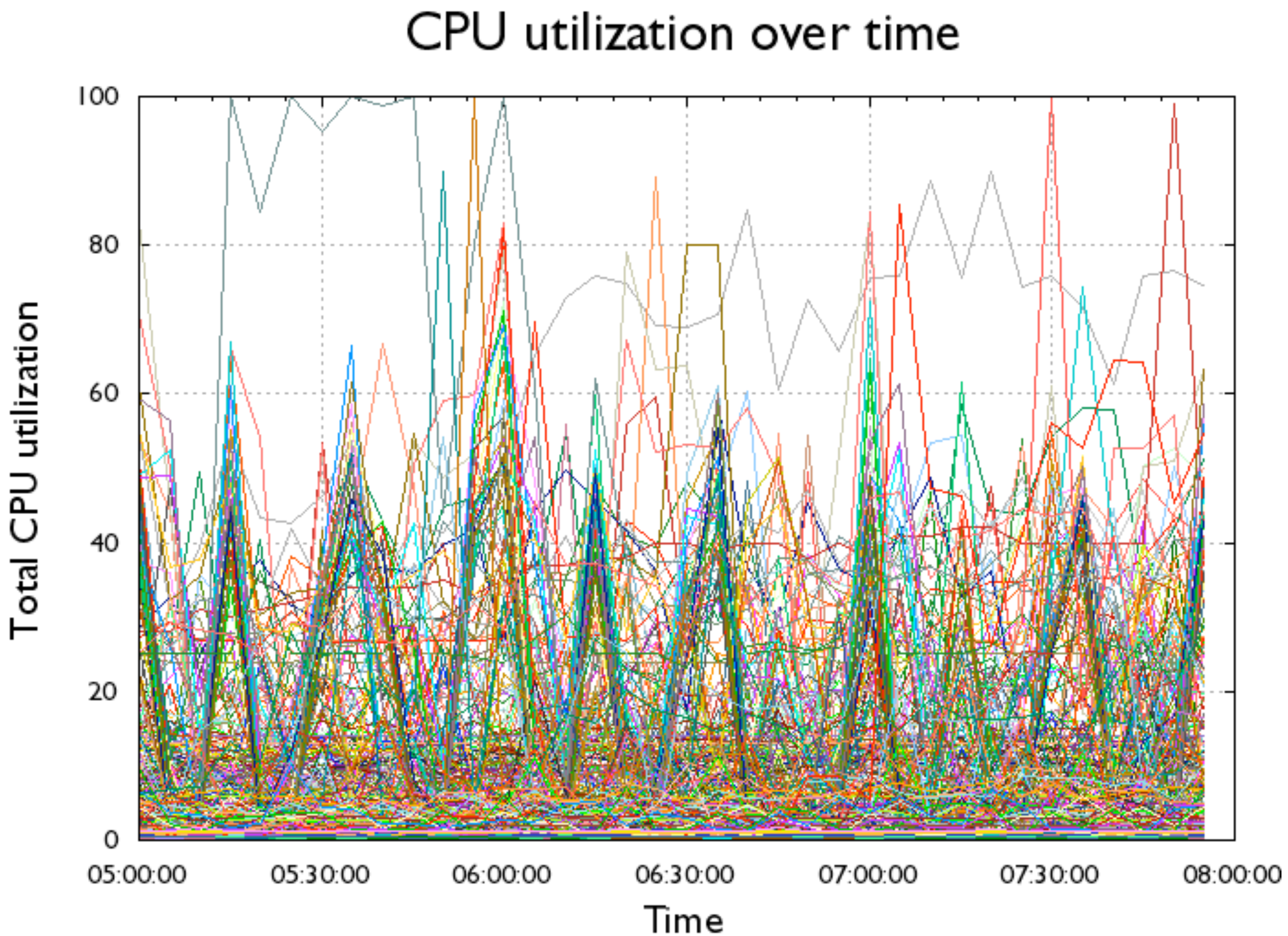
ug . . .

Monitoring time-series: is there really a problem?



ahhh . . .

Monitoring time-series: is there really a problem?



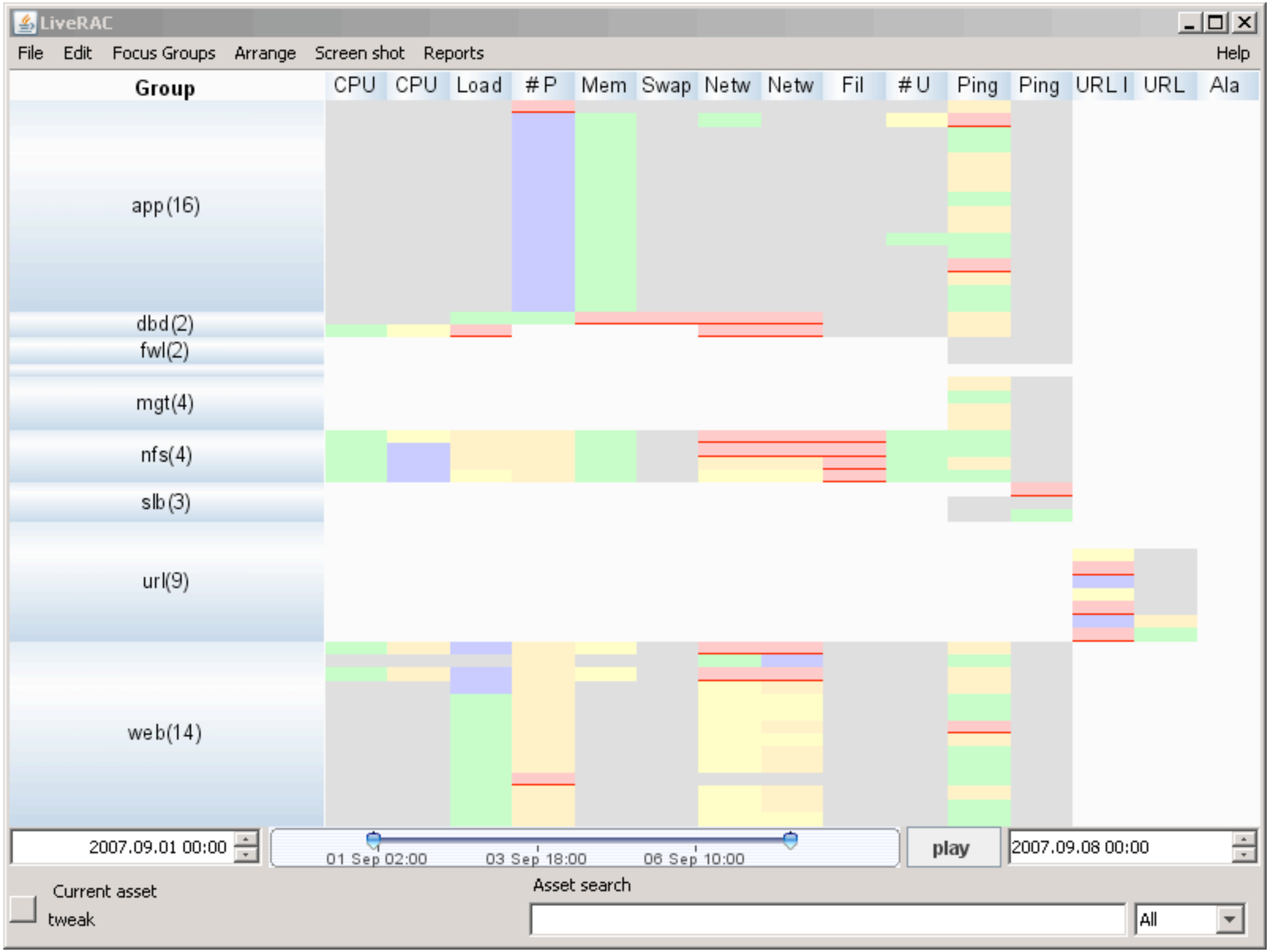
aieeee! . . .

System management professionals are hard to access

- Management needs a very persuasive business case to commit system critical personnel to evaluating prototype tools
  - paper prototypes are often not enough
- Phased design offers a chance to build credibility and gain increasing access to users
  - can make design decisions without full initial access to the desired target population

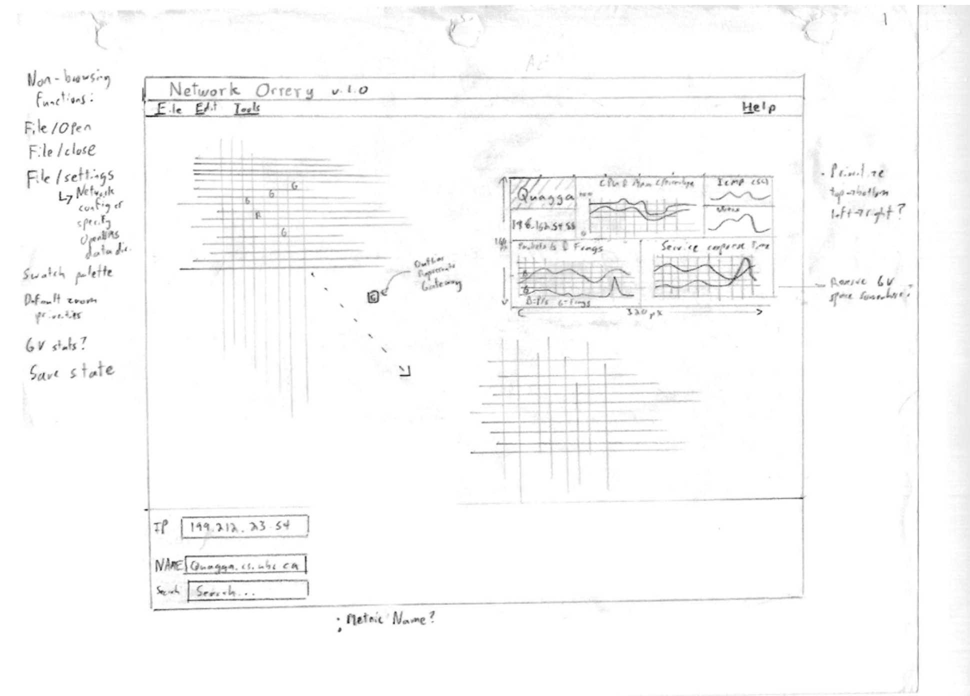
LiveRAC visualization system

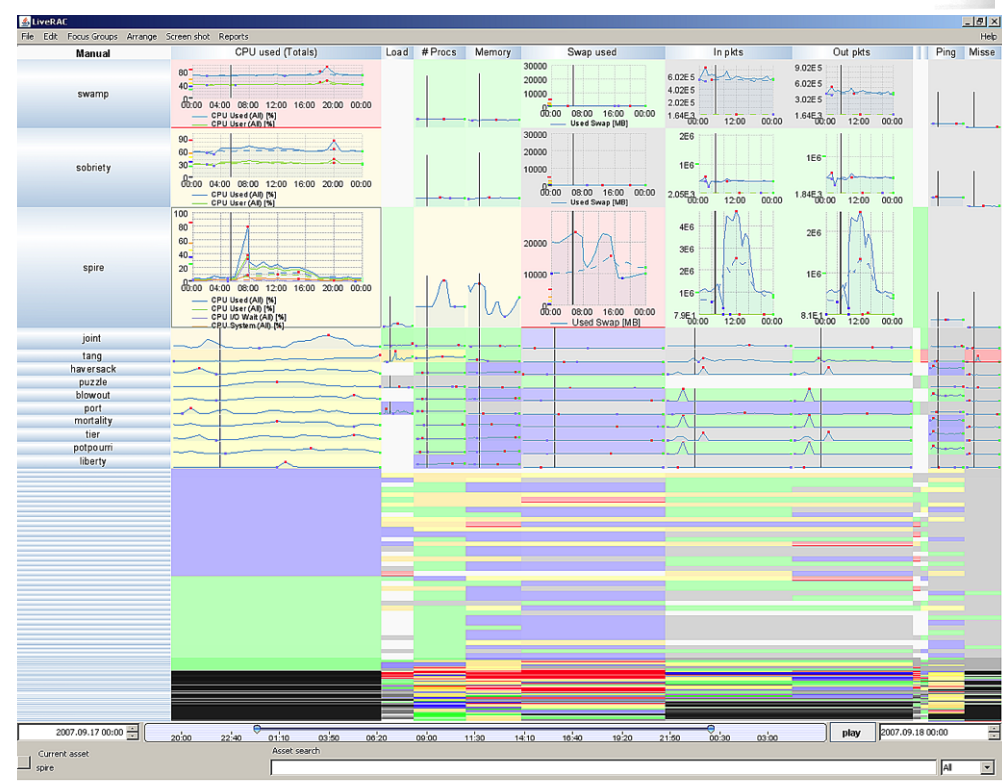
LiveRAC demo



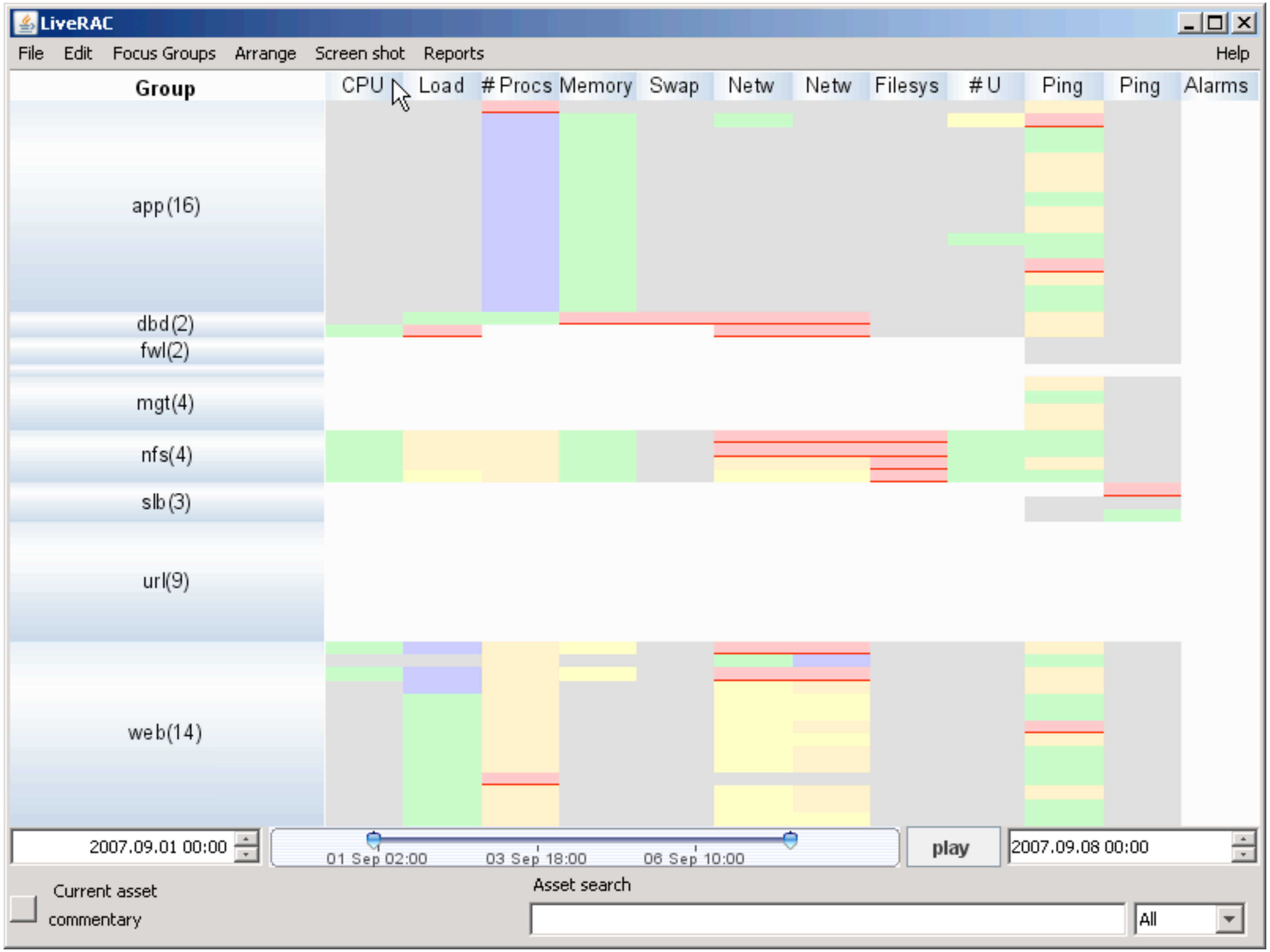
Design principles

- We present a series of motivating principles behind our design
- We'll cover only a few key principles in this talk, the rest is in the paper

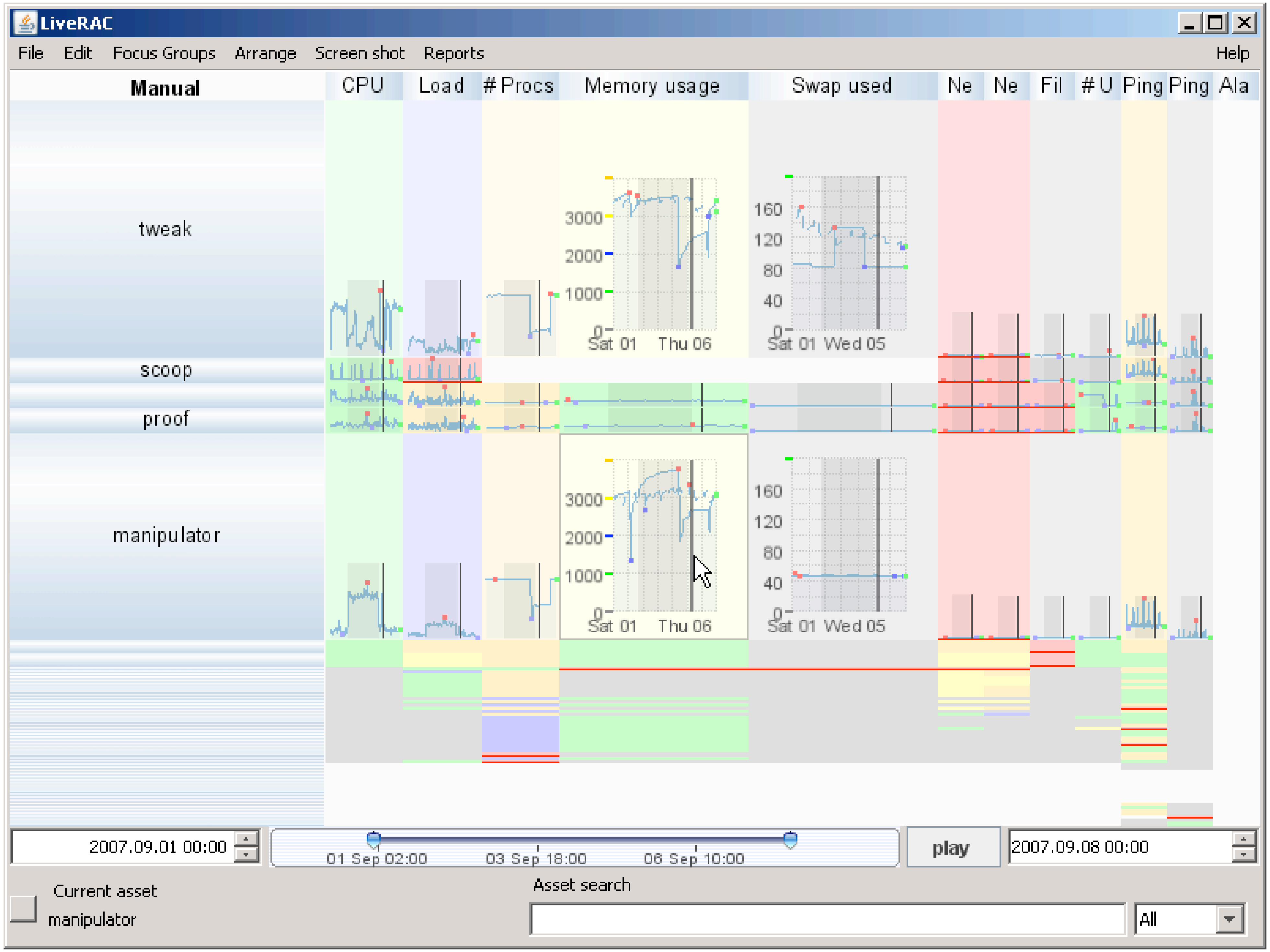




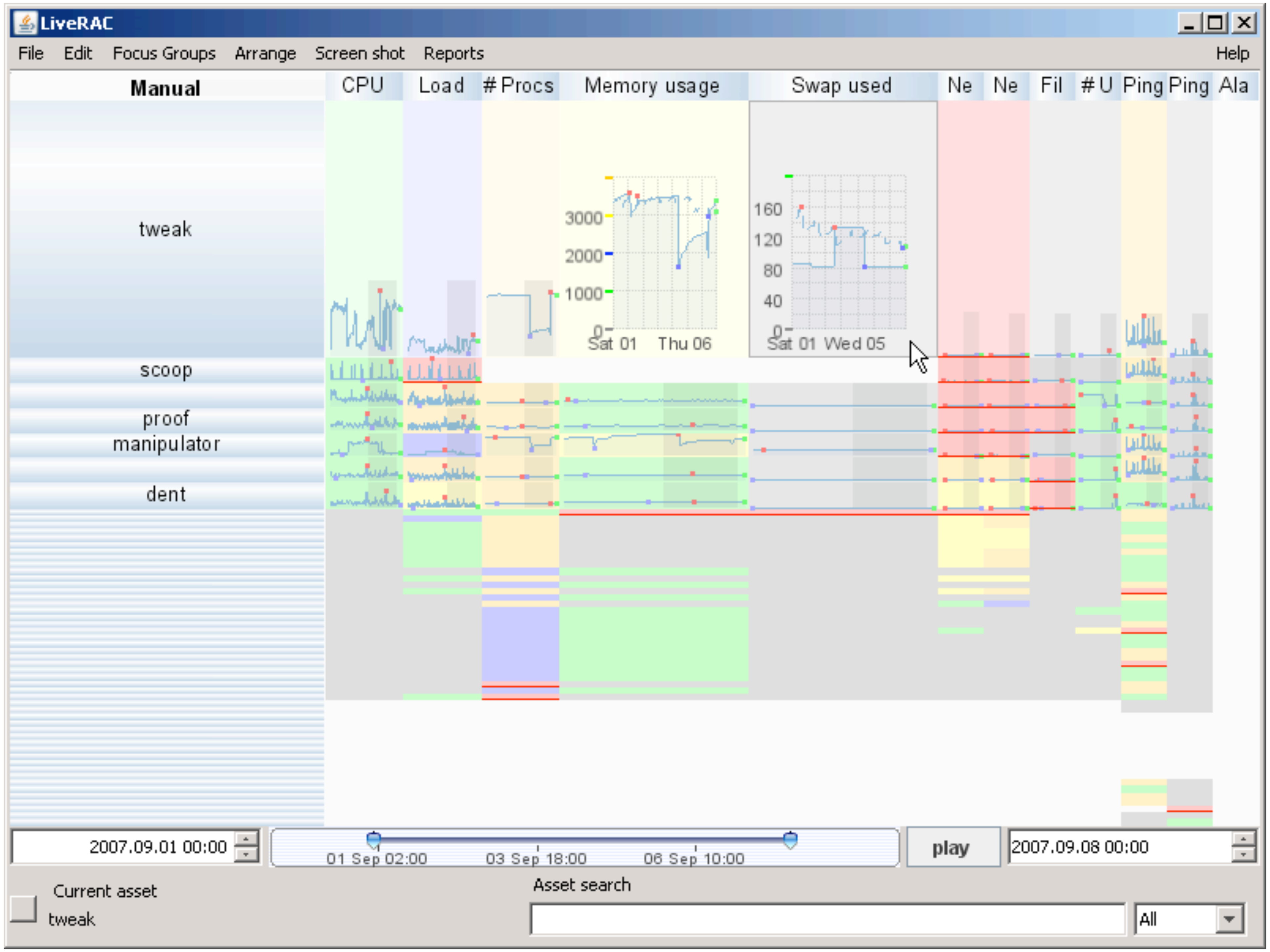
Design principle: spatial position is the strongest perceptual cue



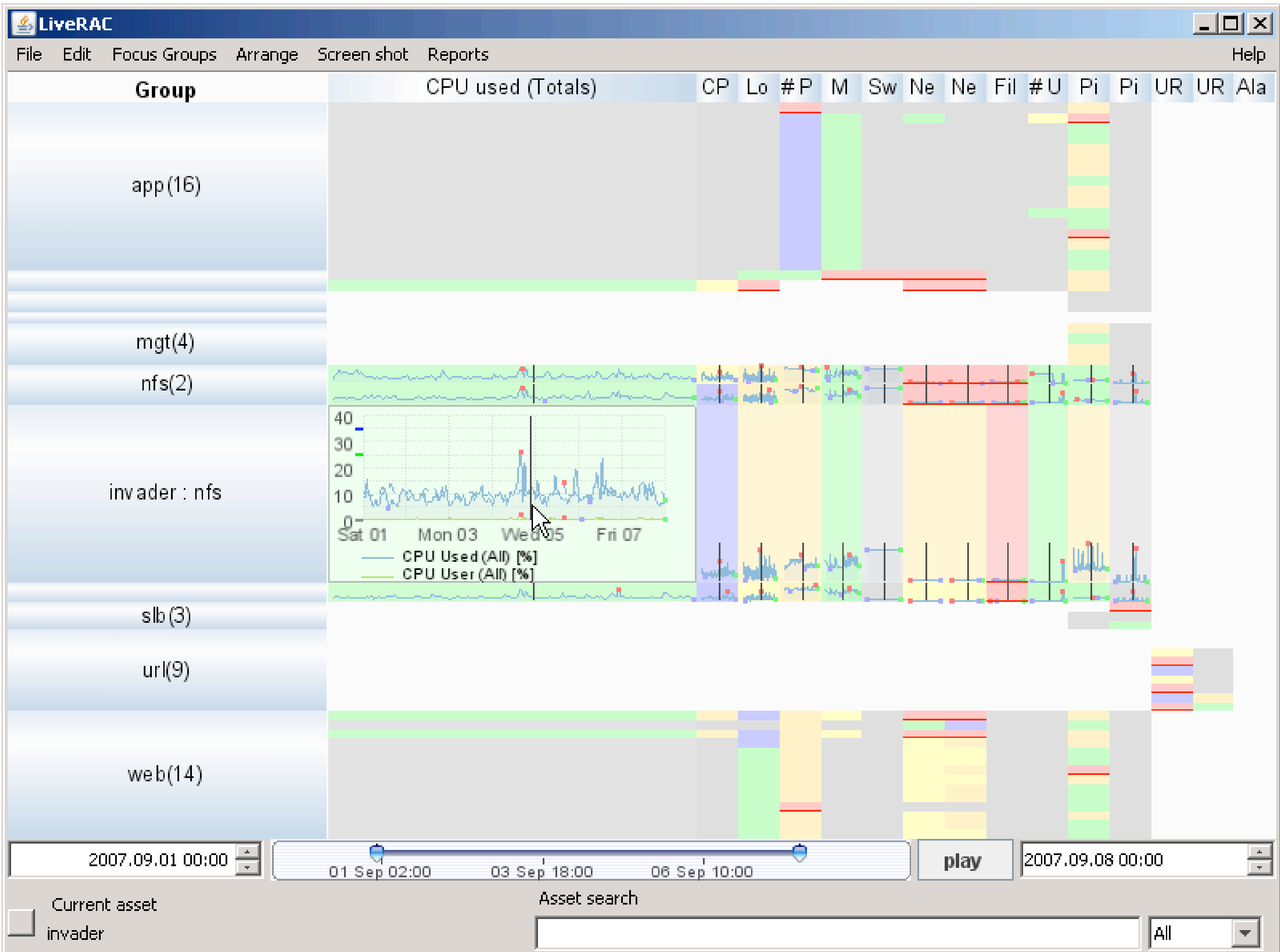
Design principle: side-by-side comparison of small multiple views is easier than remembering previously seen views



Design principle: multiple views are most effective when coordinated through explicit linking

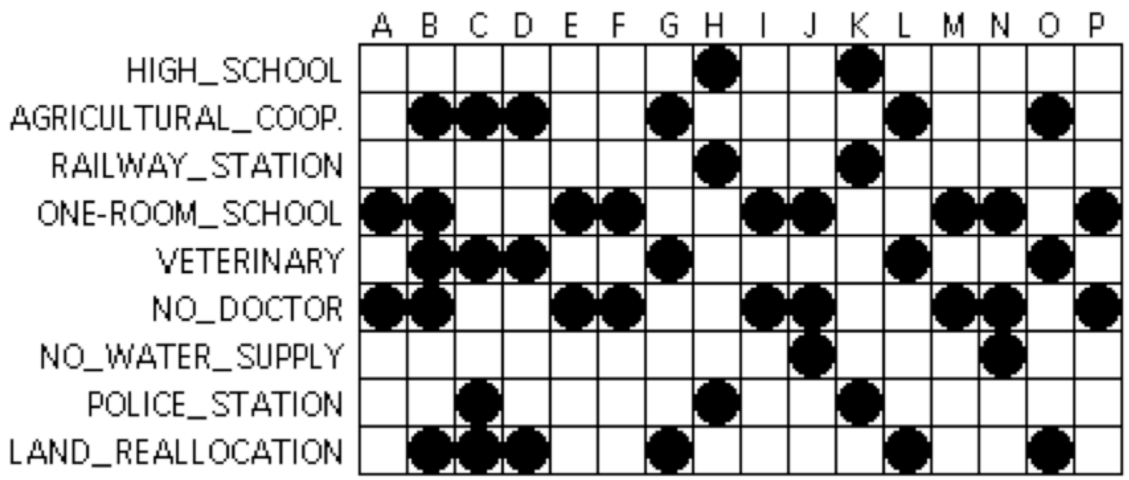


Assertion: showing several levels of detail simultaneously provides useful high information density in context



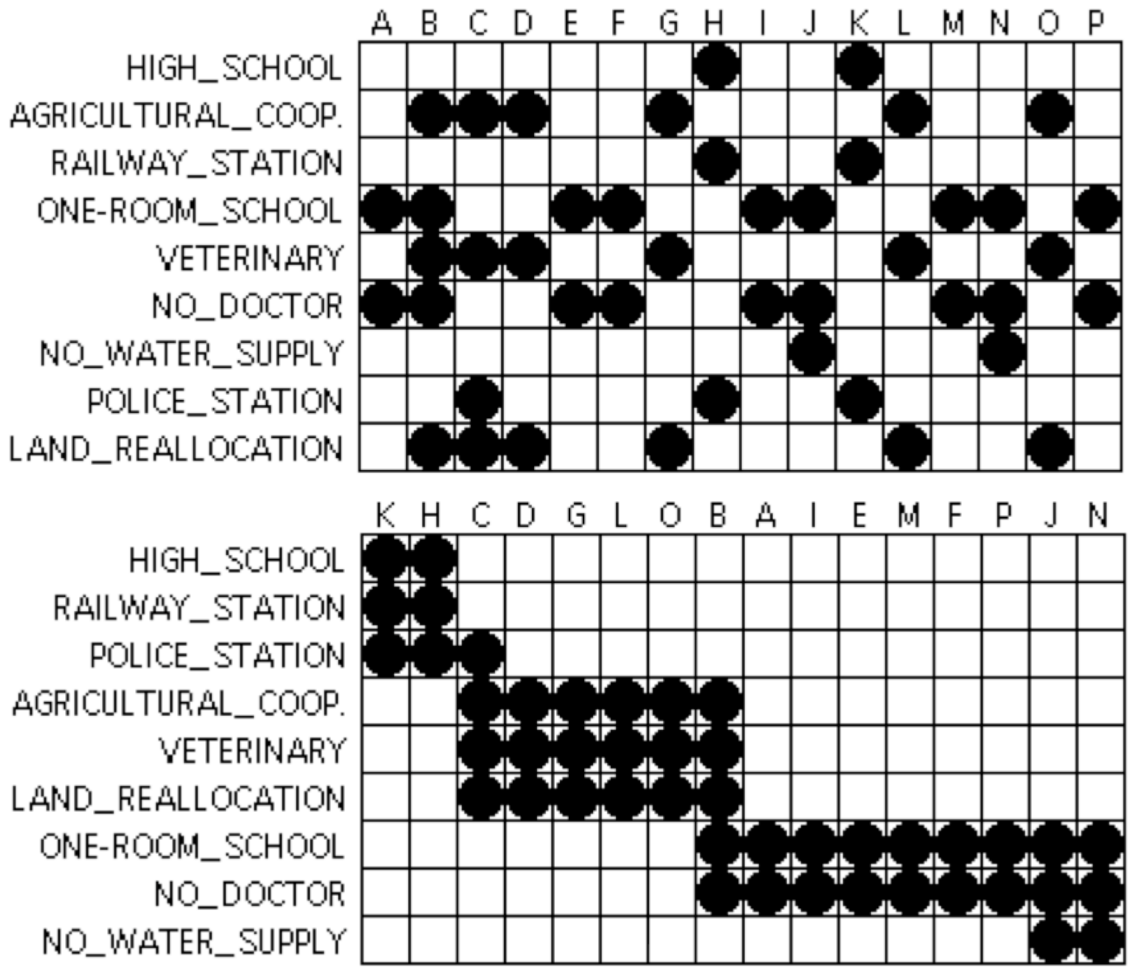
Related work

Related work: Matrix visualizations



E.Mäkinen, H. Siirtola. Reordering the Reorderable Matrix as an Algorithmic Problem. Theory and Application of Diagrams. 2000, 453–467.

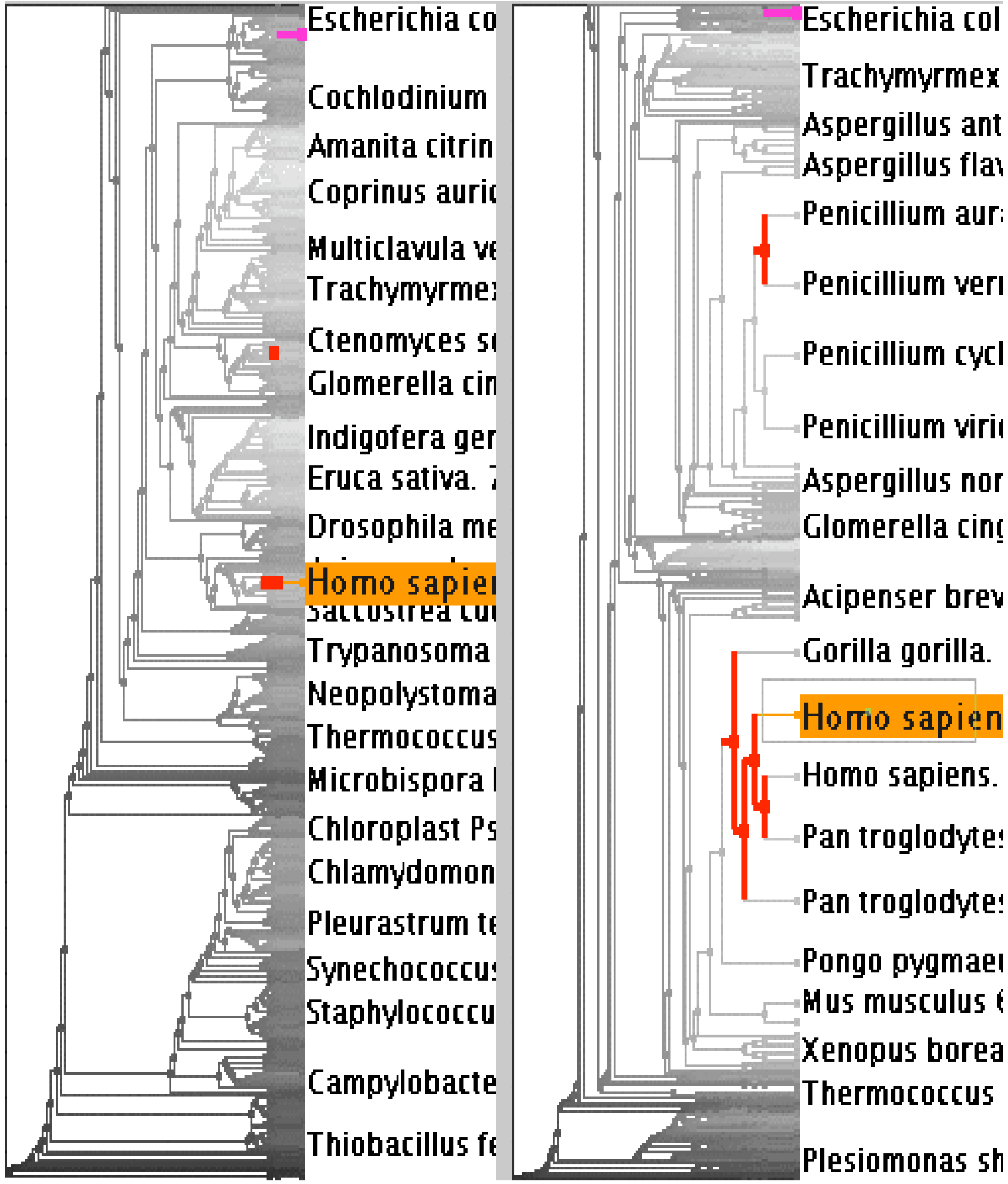
Matrix visualizations



E.Mäkinen, H. Siirtola. Reordering the Reorderable Matrix as an Algorithmic Problem. Theory and Application of Diagrams. 2000, 453–467.

Stretch and squish navigation

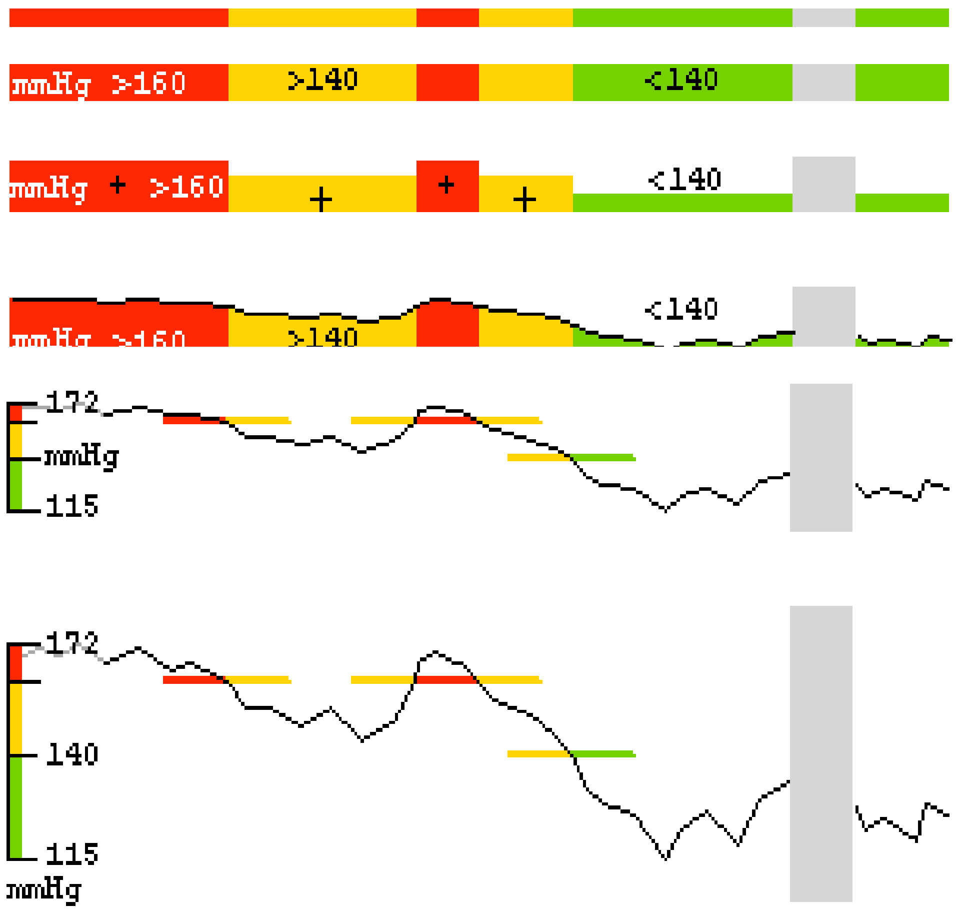
- Information visualization technique [Sarkar 1993]
  - enlarge some areas while retaining surrounding context
- Guaranteed visibility
  - important landmarks remain visible [Munzner 2003]



Munzner, Guimbretiere, Tasiran, Zhou, and Zhang. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. SIGGRAPH 2003, 453-462.

Semantic zooming

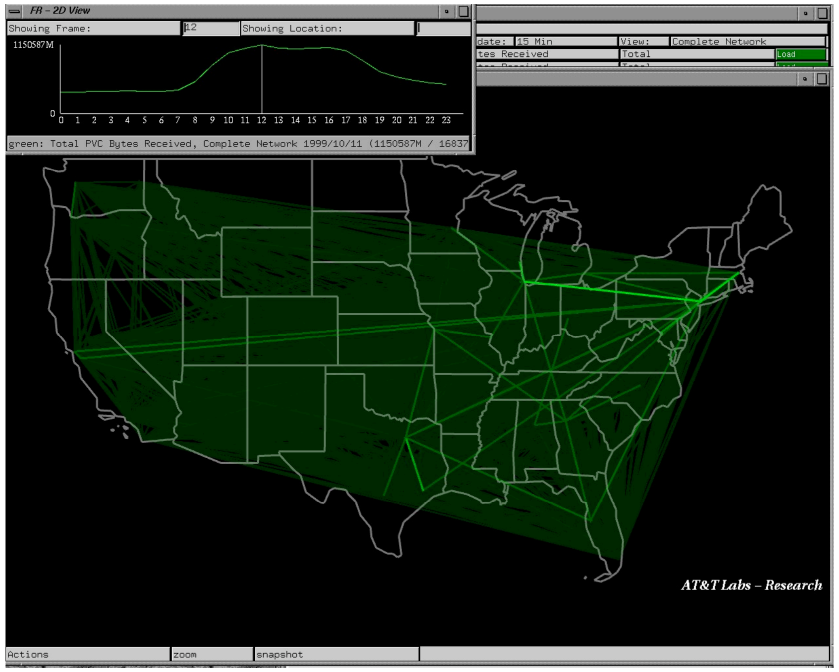
- Semantic zooming represents data differently at different zoom levels [Perlin 1993]
- Optimize representation for available space

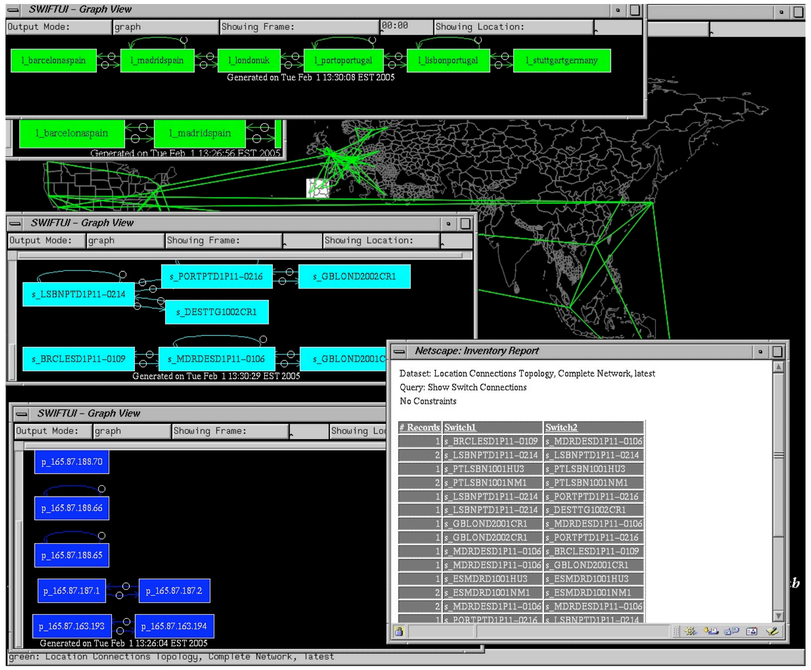


R. Bade, S. Schlechtweg, S. Miksch. Connecting Time-Oriented Data and Information to a Coherent Interactive Visualization. CHI 2004, pp 105-112.

SWIFT Network Management System

- SWIFT: data storage, aggregation and visualization tools [Koutsofios 1999]
- Existing views:
  - geographic
  - node-link
  - line charts
  - text tables
- Limitations:
  - difficult to compare between large numbers of time-series objects
- LiveRAC: re-orderable matrix visualization for SWIFT database





Koutsofios, North, Truscott, and Keim.
Visualizing large-scale telecommunication networks and services.
IEEE Visualization 1999, 457-461

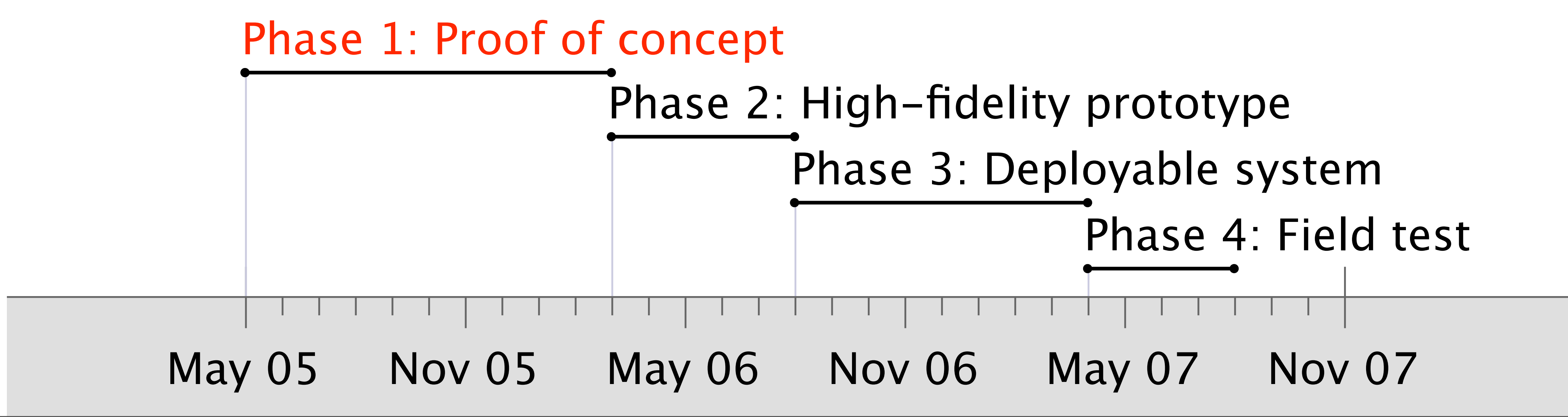
Phased design methodology

Phased design methodology

- Recruiting users with operational duties was challenging
  - modified conventional protocol increasing our participant pool in phases as the project evolved
    - generated interest by means of interim results
  - additional requirements were gathered at each phase

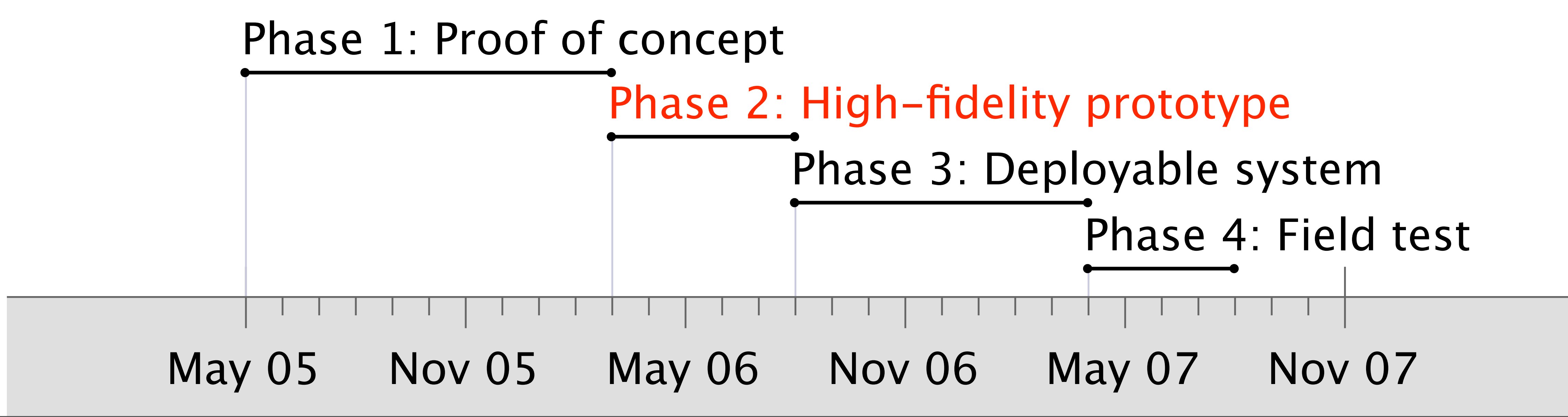
Phased design approach

- Phase 1: Proof of concept
  - Participants
    - CTO of a small company with senior system management experience
    - ourselves: installed OpenNMS enterprise network management system to improve our understanding of current "state of the art"
    - tools engineer
  - Prototypes: paper, interactive proof-of-concept
  - Data: simulated



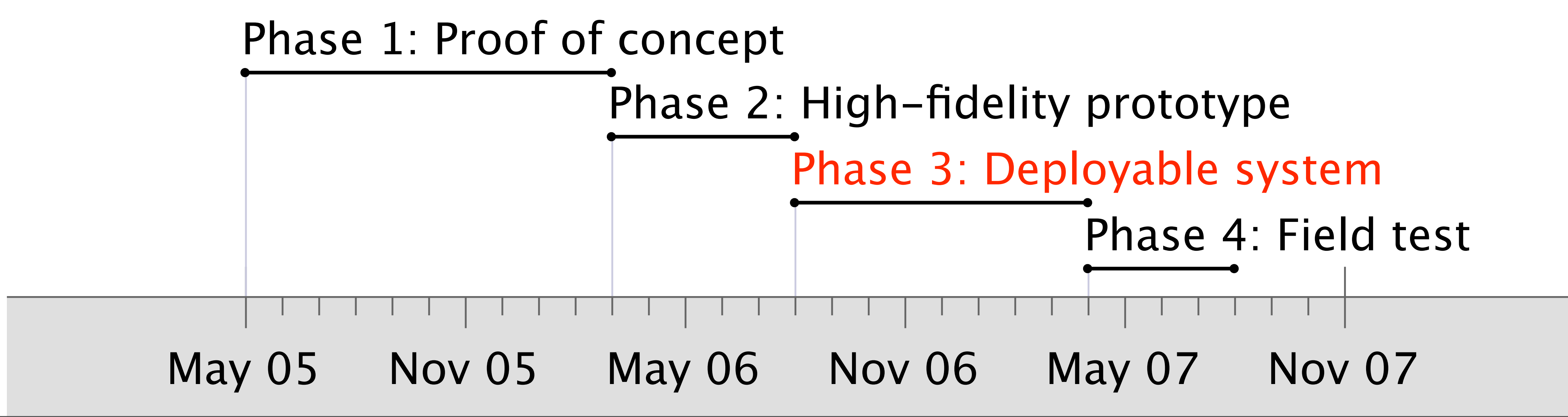
Phased design approach

- Phase 2: High-fidelity prototype
  - Participants
    - manager
    - technical director
    - tools engineer
  - Prototypes: hi-fidelity
  - Data: real data



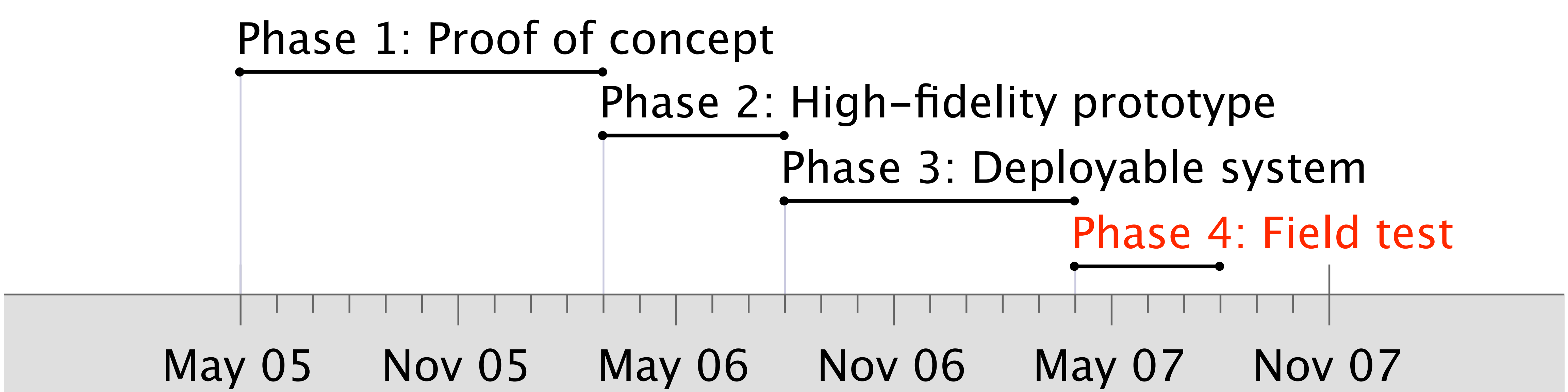
Phased design approach

- Phase 3: Deployable system
  - Participants
    - manager
    - 4 technical directors
    - tools engineer
  - Prototypes: deployment-ready hi-fidelity
  - Data: real data



Phased design approach

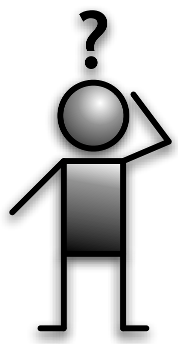
- Phase 4: Field test
  - Participants
    - manager
    - 4 technical directors
    - tools engineer
    - 7 network engineers
  - Prototypes: deployed in production
  - Data: again, real data



Phased design successes

- We basically got it right!
- Number of changes we needed to make at the end of the process were relatively minor
- Close relationship of early stakeholders with target users let us make reasonably accurate design choices
- Basic interaction strategy and visualization techniques appear valid

Phased design challenges



- Biggest misstep: wasted engineering effort
  - initial participants emphasized alarms
  - target population cared about alarms only at a very high level
    - engineers often disabled alarm viewing functionality

Informal longitudinal evaluation

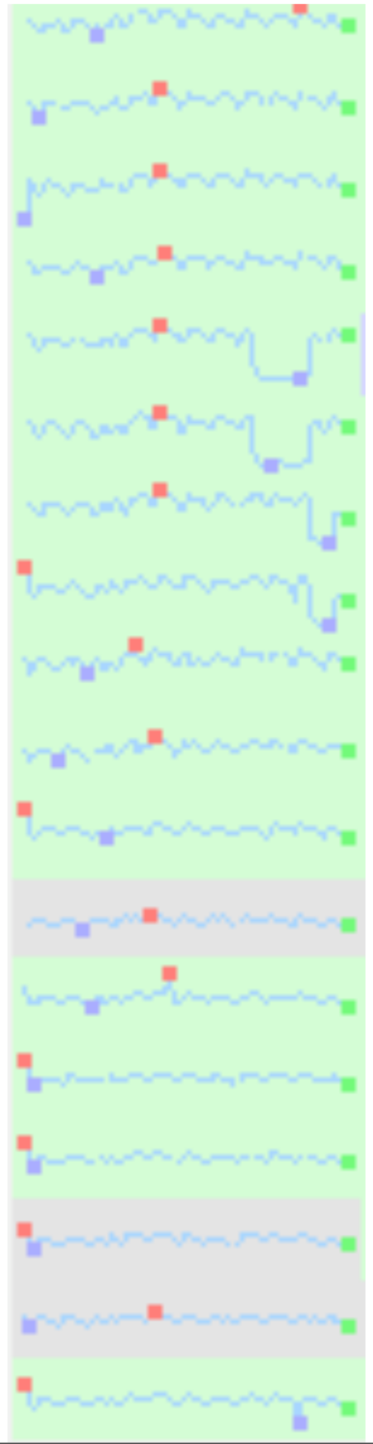
Informal longitudinal evaluation

- Objective: did our visualization system work?
  - better understand value of visualization techniques used in our design

- Data collection
  - most interviews / screen captures conducted remotely over the phone and desktop sharing software
  - participants were physically distributed over three locations on the east coast of the USA
- Training
  - difficult due to disparate locations
  - flash based training video very helpful

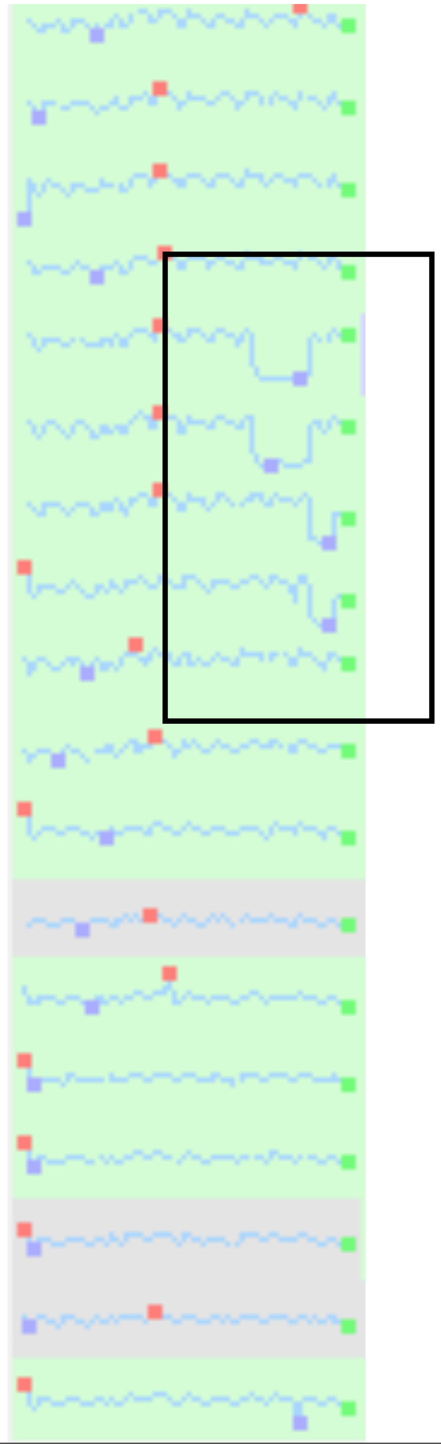
Case study: web load balancing irregularity

CPU

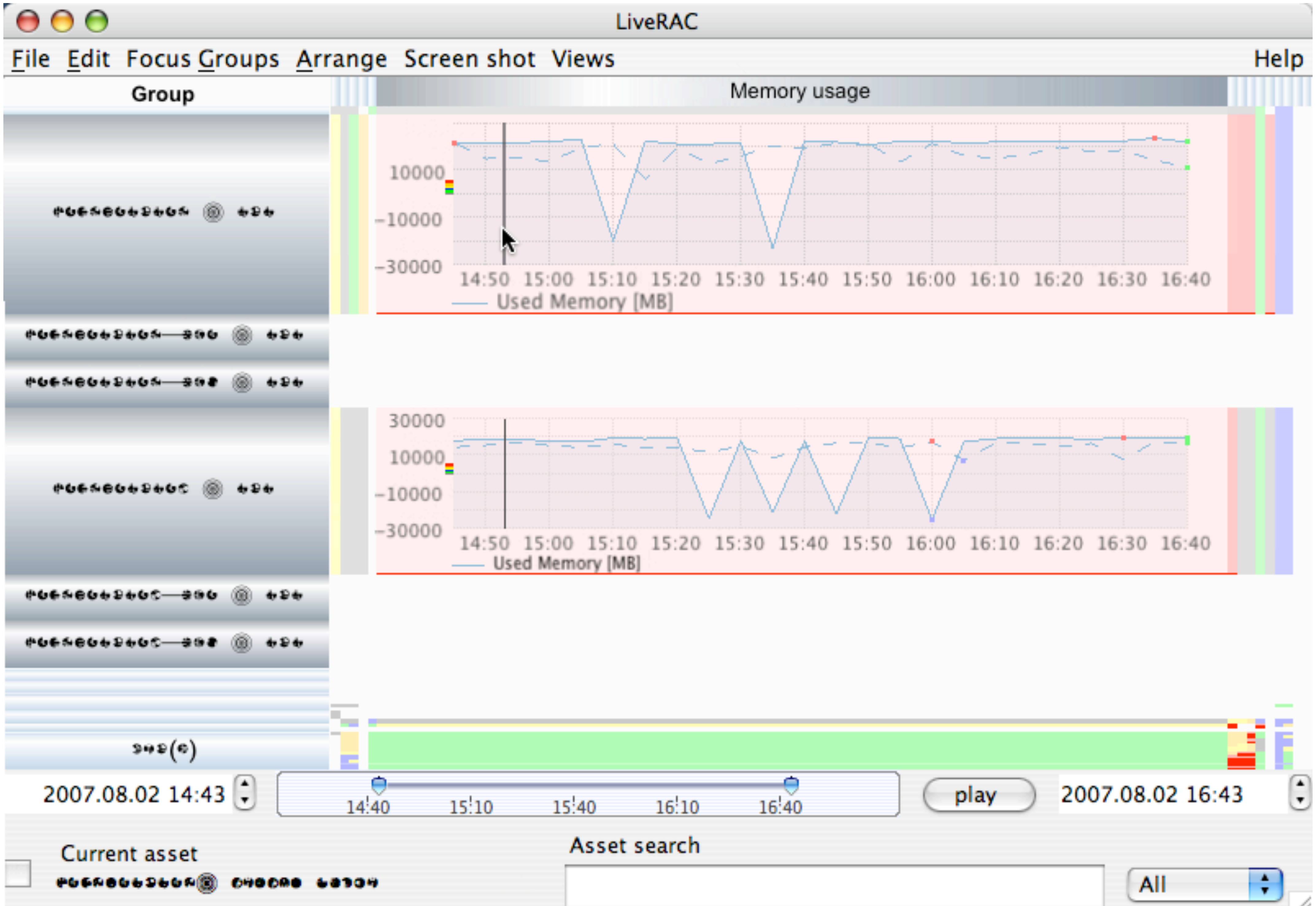


Case study: web load balancing irregularity

CPU



Case study: oops, negative memory??



Implications for design

- Iterative design can be successful even when full participation of the target user population is not available at project conception!

- The infovis design principles stood up well in the wild
    - longer term, more formal study is needed (but we think it needs to be qualitative!)

Conclusion

Future work

- Visual & interface improvements
  - apply some data charting principles
  - animated transitions on data load
- Longer term evaluation
  - months, years?
- Apply system to different data sets & domains
  - already in progress...

Conclusion

- We reviewed a phased design approach for a visualization system
- We deployed LiveRAC in a workplace environment using real data
  - interest in making the system a part of the standard engineering tool suite
- We conducted an informal longitudinal study with encouraging feedback on our visualization techniques
  - also validates design approach

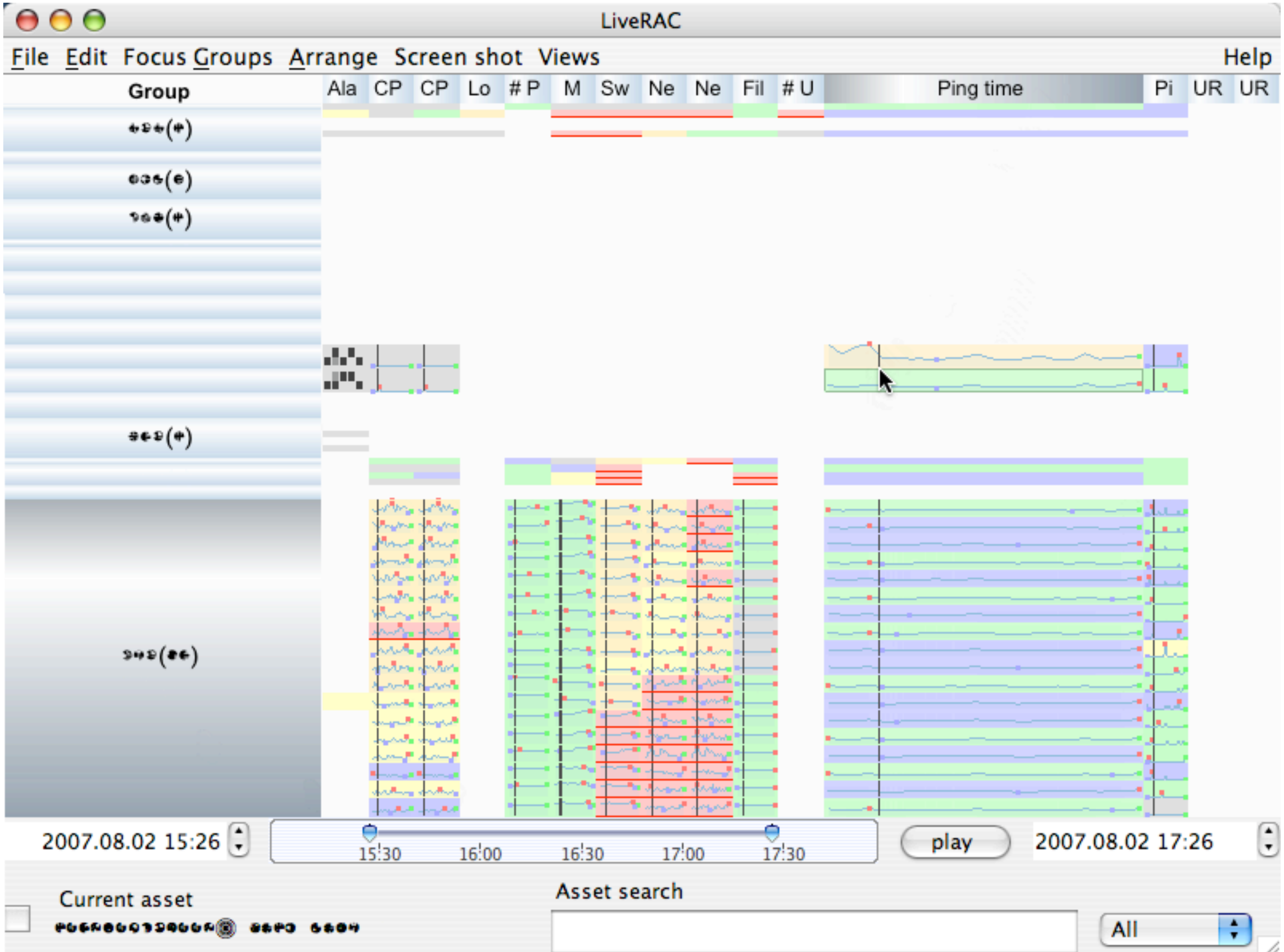
Extras

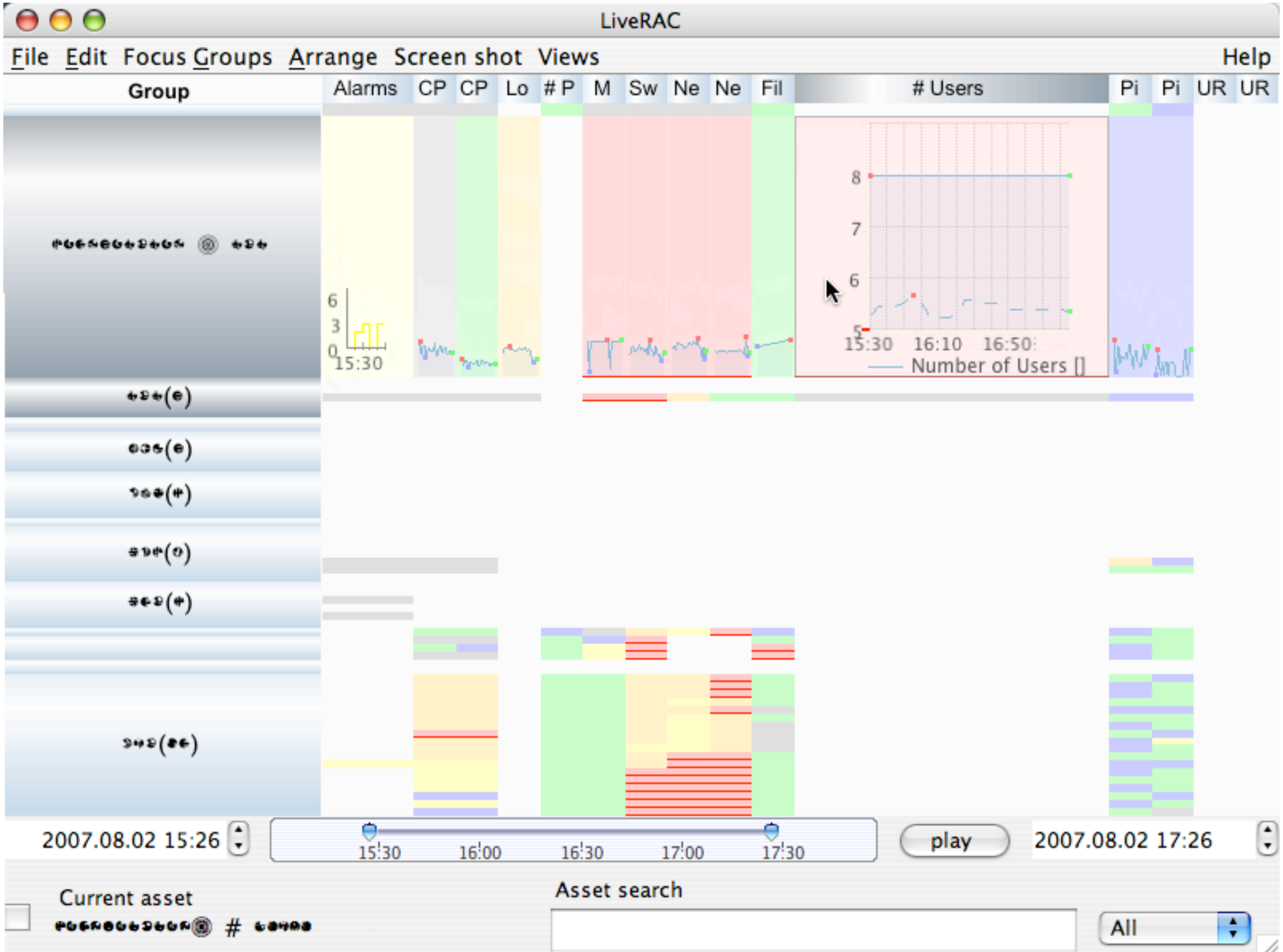
Participants

- 14 participants total in 3 groups
  - external group:
    - one participant, CTO of a small company with senior system management experience
  - internal group:
    - senior technical personnel in our target organization
      - tools engineer (te)
      - executive director (ed)
      - four senior technical directors (sd)
  - life cycle engineering group (primary target population):
    - seven life cycle engineers (lce)

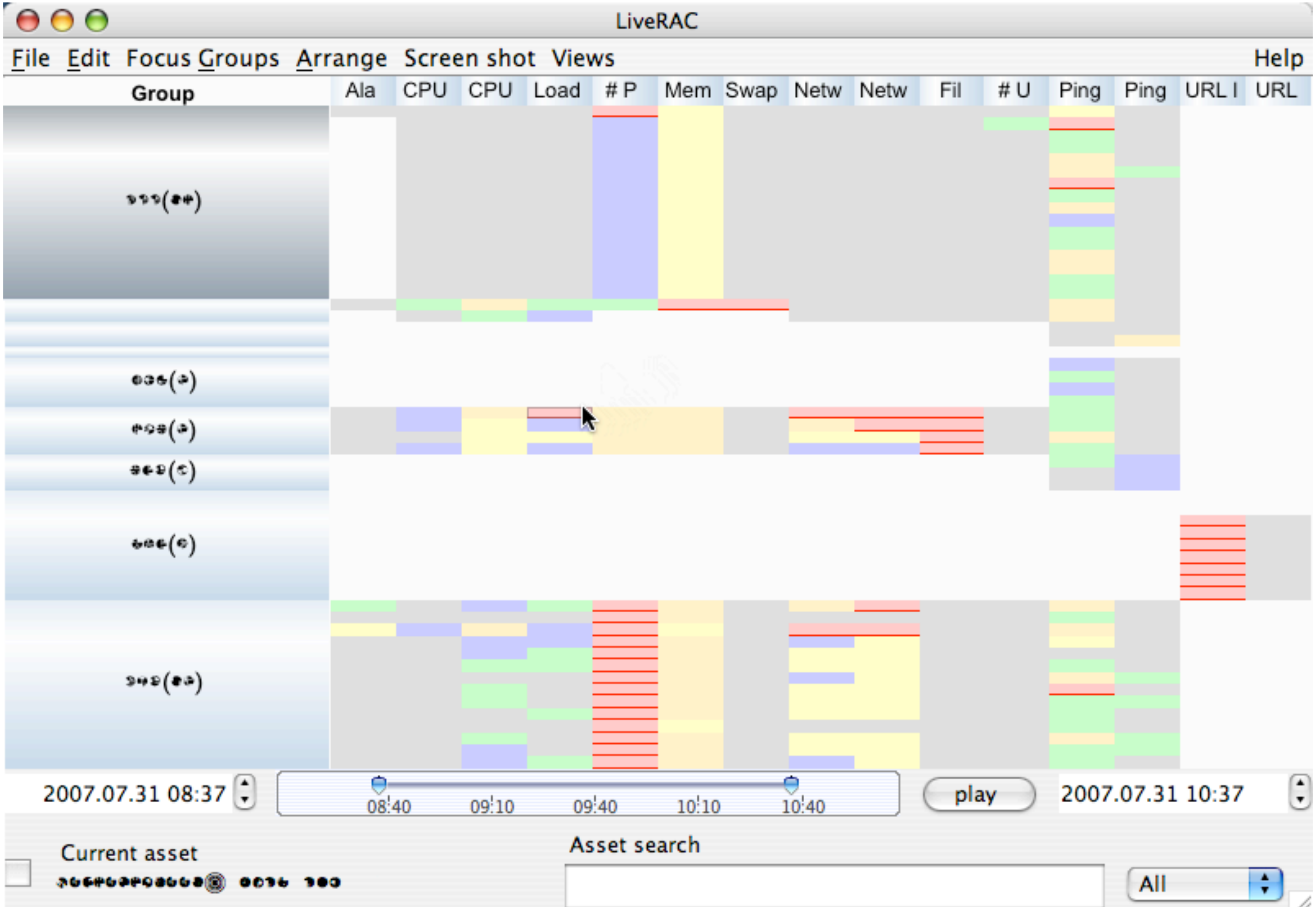
Load sort


LiveRAC
File Edit Focus Groups Arrange Screen shot Reports
Help
Manual
Alarms
CPU used
Load avg
Procs
Memory
Swap used
Network
Network o
Filesys
Users
Ping time
Ping loss
2007.06.22 06:00
05:10 05:50 06:30 07:10 07:50 08:30 09:10 09:50 10:30 11:10 11:50 12:30 13:10
play
2007.06.22 12:00
Current asset
i35034nfs002: Load avg
Asset search
All


LiveRAC
File Edit Focus Groups Arrange Screen shot Views Help
Group
Ala CP CP Lo # P M Sw Ne Ne Fil # U Ping time Pi UR UR
2007.08.02 15:26
15:30 16:00 16:30 17:00 17:30
play
2007.08.02 17:26
Current asset
Asset search
All


LiveRAC
File Edit Focus Groups Arrange Screen shot Views Help
Group
Alarms CP CP Lo # P M Sw Ne Ne Fil # Users Pi Pi UR UR
8
7
6
5
15:30 16:10 16:50
Number of Users []
2007.08.02 15:26
15:30 16:00 16:30 17:00 17:30
play
2007.08.02 17:26
Current asset
Asset search
All

Where is the high load?



Study details

- Surveys: pre-survey and post-survey administered to all 7 LCE participants and 4 technical directors
- 44 connections
- Actual usage difficult to estimate, in the 10's of hours range

Logged events

Focus group modify	3
Change update frequency	2
Strip load	51
Grow search result	19
Arrange strips	2
Focus group resize	14
Group load	51
Double click zoom	229
Single line move	144
RECT_CREATE	594
Query	744
Resize group	103
First time survey	10
Reset display	70
Source load	51
Search sources	257
ReshapeRect_OP	656
Connection	44
Scroll wheel zoom	2578
Survey	68
Strip focus group load	51
Time range modified	330
Cell detail request	31
Source focus group load	51
External URL request	6

Information visualization

- Human visual channel is highest-bandwidth perceptual system [Norretranders, 1999]
- Information visualization: field of study whose object is to aid cognition through the graphic representation of abstract data
- displays relevant information graphically to assist in memory tasks
- supports data exploration through direct interaction
- assists in pattern finding through the display of overview and detail, search, and user-directed reordering

Pre-experiment survey

1) Gender: **Male / Female** (please circle)

2) Age: _____

3) When you use the VizGems visualization system, are you looking at alarms, stats, or both? (circle all that apply.)

Alarms **Stats**

4) On average, how much time per day do you spend interacting with the VizGems visualization system? (please circle one)

___ Less than 30 minutes
___ 1 hour
___ 1-2 hours
___ 3-4 hours
___ 4-5 hours
___ 5-6 hours
___ 7-8 hours
___ 9-10 hours or more

5) VizGems is necessary for my daily activities: (please circle one)

1 2 3 4

1=Very infrequently
2=Somewhat infrequently
3=Often
4=Constantly

6) VizGems makes my job easier: YES / NO (please circle one)

7) I use VizGems for the following activities: (check all that apply)

___ Capacity planning
___ Alarm investigation
___ Curiosity
___ Comparing assets
___ Inventory management
___ Situational overview
___ Statistical investigation
___ Ticket management
Other (please describe below)
____
____

8) In a few sentences, summarize the tasks in which you most typically interact with VizGems. (continue on the back of the page if necessary)

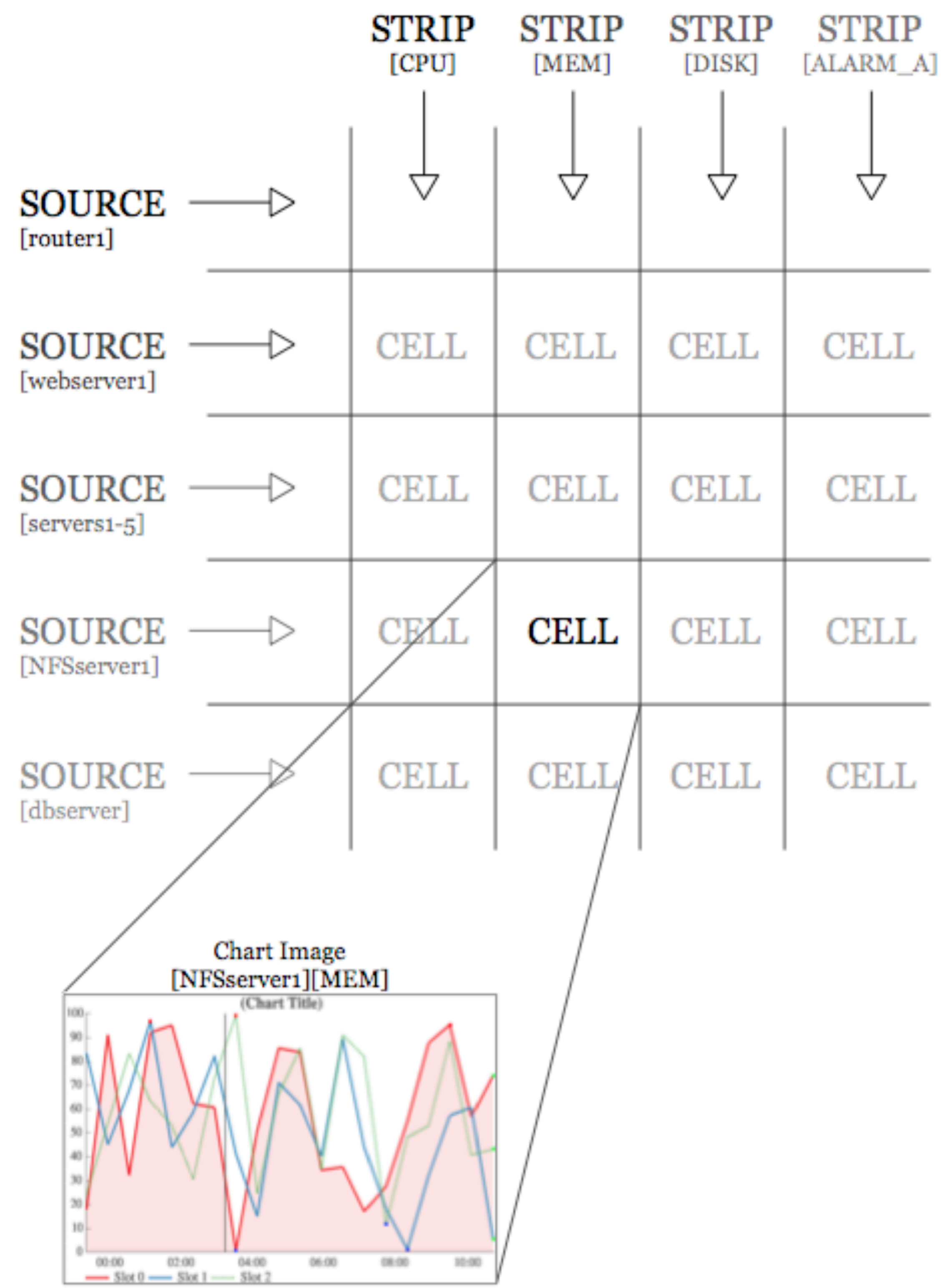
____
____
____

Exit survey questions

- "Why did you use LiveRAC this session? (check all that apply)"
  - "Incident investigation"
  - "Capacity planning"
  - "Situational overview"
  - ”Curiosity"
  - "Comparing asset status"
- Did LiveRAC help you?
  - yes, no
- Writing a few sentences concerning your task and experience with LiveRAC will help us improve it!

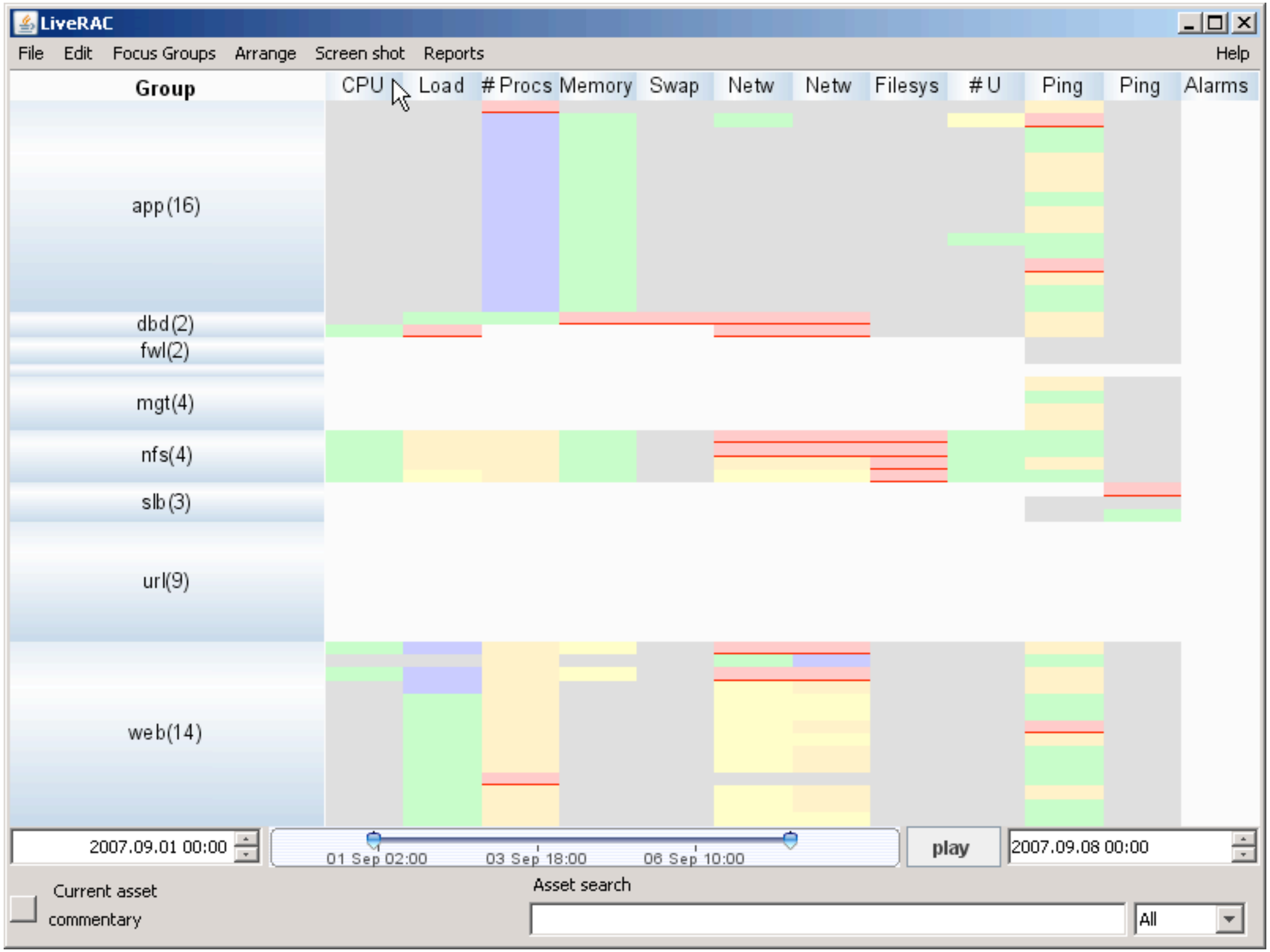
LiveRAC visual encoding

- High data density grid view
    - rapidly explore time-series data
    - context available at all times
- Re-orderable matrix supports side by side data comparison



Principles

Design principle: overview first, zoom, filter, details on demand



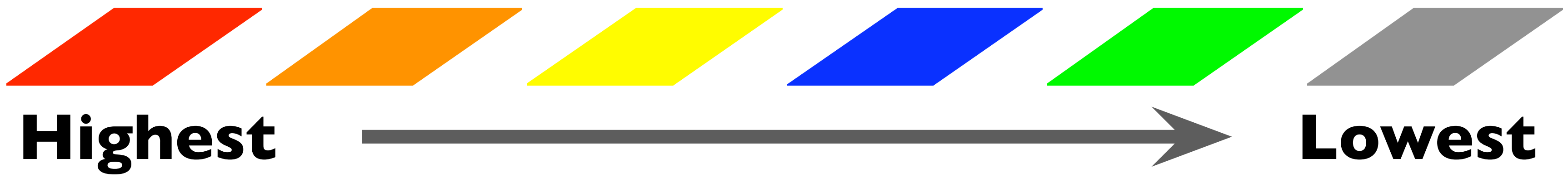
Design principle: abrupt visual change should be avoided

- Principle arises from perceptual theories of object constancy [Robertson, 1989] and change blindness [Rensink, 2000]
- Transitions are animated during navigation
- During time window changes, we continue to show the old representation but mark it as stale with an unobtrusive yellow dot

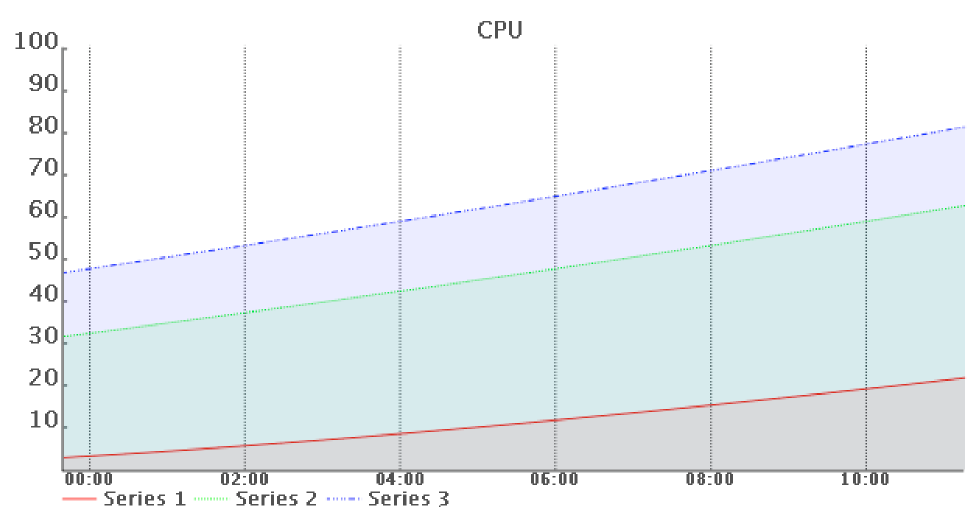
Design principle: user actions should receive immediate visual feedback

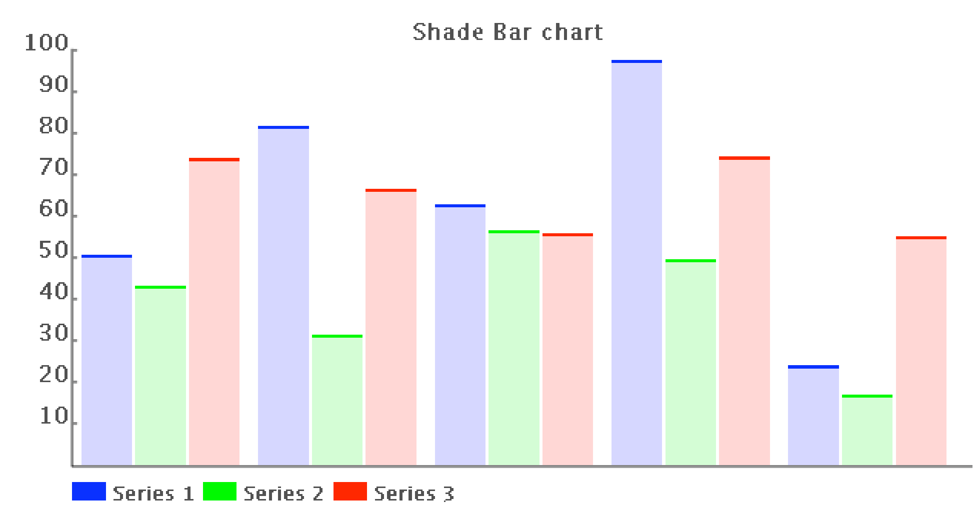
- LiveRAC provides guaranteed minimum frame rates
  - if the frame takes too long to draw during navigation, drawing is not completed and rendering on the next frame begins
  - critical marks are rendered first
- Server updates take place using separate threads preserving interactivity

Design principle: familiar visual representations should be preserved when appropriate



- Although not maximally discriminable, we preserved the color scheme used by LCEs in previous network management applications





- Users already familiar with line and bar chart graphics

Monitored data

- Most data collected from monitored network devices is time-series data
    - any type of computer or appliance: servers, routers …
    - time stamp and value
- Two types of time-series objects collected:
    - performance metrics
        - [10 AUG 2007 9:52:37, CPU, 95%]
    - alarm data
        - [16 AUG 2007 12:13:25, "HIGH TEMP", CRITICAL]
- Key difference for visualization:
    - performance metrics quantitative
    - alarms categorical

System management activities

- Interpreting network environment status
  - understand state of individual devices and impact on end-to-end services
  - report generation
    - create sharable documents for customers or internal reporting
- Capacity planning
  - forecasting future system and infrastructure needs
    - requires an accurate understanding of network environment status
- Event investigation
  - Specific events might be service outages or network security breaches
- Coordinate between customers, engineering & operations
  - "work as a conduit" for service delivery

System management professional job description

- System management professional primary goal: meet Service Level Agreement (SLA) for customer
- SLA's describe:
  - what services will be provided
  - how they will be delivered
  - how service delivery will be measured
  - consequences if agreement is not met
- A combination of business, analytical & system management skills are required

Challenge: managing large numbers of systems

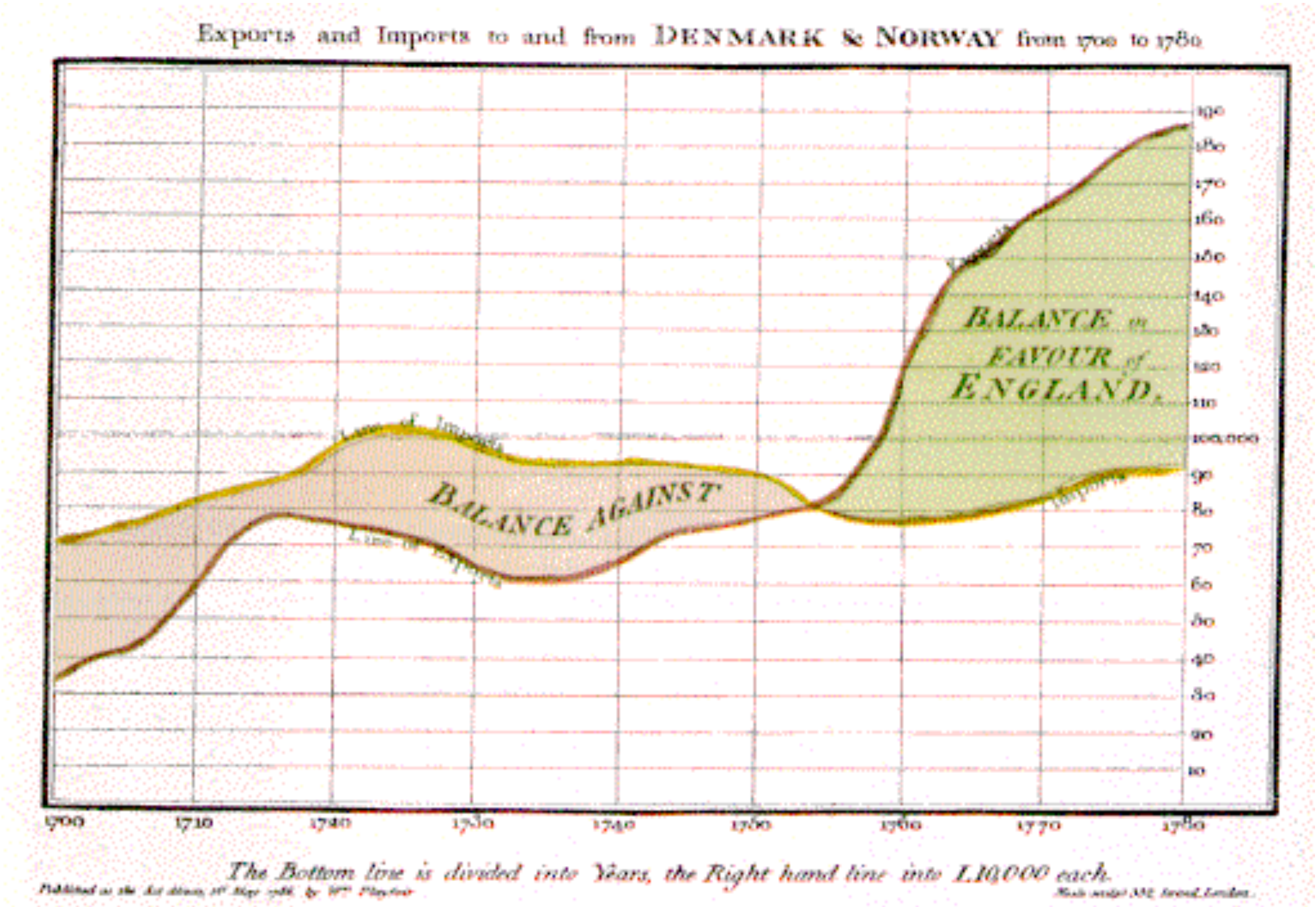
- Monitoring tasks are very difficult ("vigilance decrement") [James 1890] [Davies 1982]
- Made worse by:
  - multiple disparate data sources
    - different platforms
    - different protocols
    - different services
  - lack of context in standard tools
- Shortage of integrated visualization solutions
- LiveRAC: visualization system for browsing & correlating time-series data with high information density

Related work

Longitudinal visualization studies

- Actively developing area of information visualization
- Numerous papers exhorting visualization researchers to perform longitudinal research studies [Schneiderman 2006] [Plaisant 2004]
- Work by Gonzalez and Kobsa, 2003, studied the benefits of deploying an information visualization system in a corporate environment
  - emphasized importance of tool chain integration

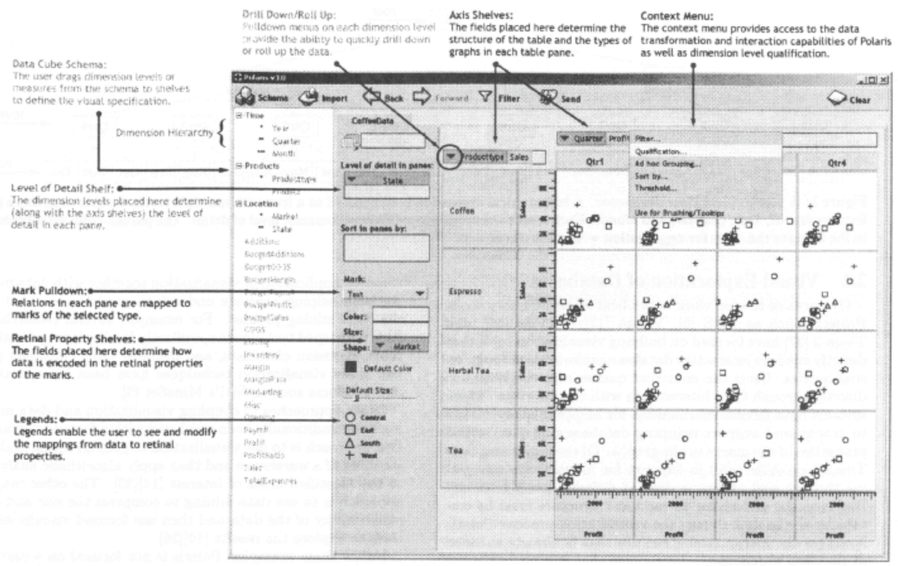
Time-series visualizations



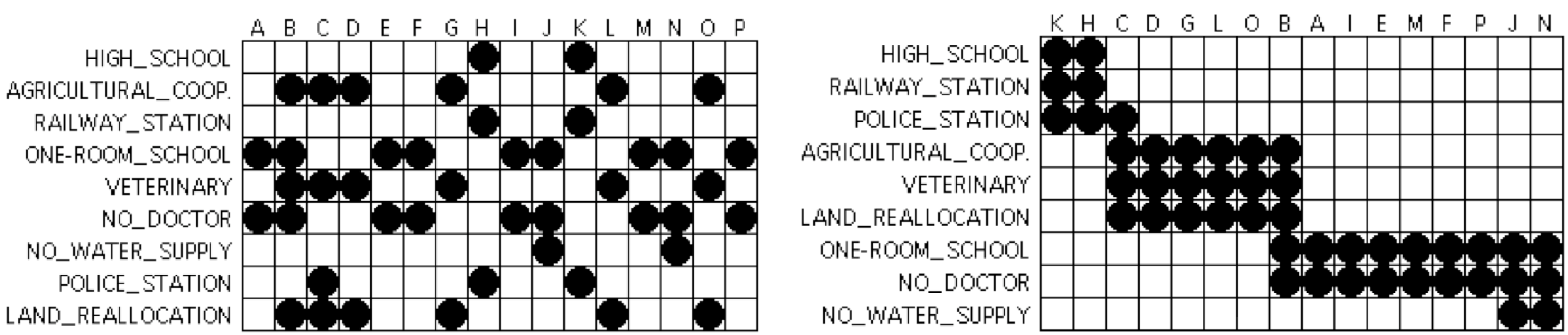
[Playfair 1786]

- They've been around a long time...
  - common examples include information graphics created by William Playfair (1759 - 1823)
- Extensive study of time-series conducted by Aigner et al. in *Visualizing Time-Oriented Data – A Systematic View*
- TimeSearcher [Hochheiser 2001] is an example of an information visualization tool specifically designed for time-series data

Matrix layouts



C. Stolte, D. Tang, and P. Hanrahan. Query, analysis, and visualization of hierarchically structured data using Polaris. Proc. KDD 02,. 2002, 112-122.



E.Mäkinen, H. Siirtola. Reordering the Reorderable Matrix as an Algorithmic Problem. Theory and Application of Diagrams. 2000, 453–467.